# SANGOMA: Stochastic Assimilation for the Next Generation Ocean Model Applications EU FP7 SPACE-2011-1 project 283580

Deliverable 5.5:
Regional model V1 analysis
Due date: 01/11/2014
Delivery date: 13/07/2015
Delivery type: Report , confidential

Luc Vandenbulcke        Jean-Marie Beckers
Alexander Barth
University of Liège, BELGIUM

Peter Jan Van Leeuwen        Sanita Vetra-Carvalho
University of Reading, UK

Lars Nerger
Alfred-Wegener-Institut, GERMANY

Arnold Heemink        Nils van Velzen
Martin Verlaan
Delft University of Technology, NETHERLANDS

Pierre Brasseur        Jean-Michel Brankart
CNRS-LEGI, FRANCE

Pierre de Mey
CNRS-LEGOS, FRANCE

Laurent Bertino
NERSC, NORWAY

# Chapter 1

# Introduction

The impact of high-frequency radar surface current observations in data assimilative regional ocean models is analyzed in the particular case of an ensemble of models of the Ligurian Sea. The method for generating the ensemble is exposed, as well as the observation operator linking the data and model vectors. The impact of the velocity observations is analyzed in space as well as in time. Their impact on other (non-observed) variables is assessed as well.

A number of recent studies, following the initial paper by Barrick et al (1977, 1978), already presented assimilation of radar currents in models. Let's mention:

- Lewis et al. [1998] use nudging by adding a pseudo-layer above the surface and impose a shear stress on top of the one generated by the wind. At the time, they noted that the radar data accuracy was rather poor.

- Breivik and Saetra [2001]: During the EuroROSE project, they performed operational data assimilation in a nested model along the Norwegian coast. Data is discarded when the velocity difference between model and observations is more than 0.5m/s, or the direction difference is more than 45Â°. The analysis is smoothed with a second-order Shapiro filter. Some scatter plots of the kinetic energy in analysis versus observations (1 plot per lead time) are presented, including the linear regression line.

- Oke et al. [2002] assimilates high-frequency radar observations off the Oregon coast. The correction is applied in multiple steps

- Kurapov et al. [2003] uses radar data to improve a model (including tides) off Oregon as well

- Wilkin et al. [2005] assimilates radar data in a ROMS model off the New Jersey coast

- Kaplan and Lekien [2007] produces smooth two-dimensional fields from radar-based observations, though noting that assimilation of radial currents avoids the additional step and errors of creating the vector currents

- Chao et al. [2009] Run a ROMS model nesteed in the Monterey bay. Available observations include 4 radars with a coverage of about 200km and

with resolution of 1 to 3km, but this data is used for validation only. The hourly radar data is filtered with a 33-hour filter to remove tidal currents (diurnal and semi-diurnal time-scales). Regarding predictability, they note that glider velocity rms errors double in 48h, and conclude that it is necessary to assimilate current data.

- Shulman and Paduan [2009]: In a Monterey bay experiment,
  1/ assimilation of low-pass filtered (33 hours) radar data into a non-tidal model, improves comparison with moored current observations
  2/ assimilation of unfiltered radar data into a model with tides yields the same level of improvement
  Thus, unfiltered radar data can be assimilatde as long as the model represents the observed time-scales too.

- Hoteit et al (2009) use a 4DVar filter to improve initial condition, OBC and atmospheric forcings

- Barth et al. [2010] use an ensemble Kalman filter to assimilate HFradardata and correct tides. The ensemble members are perturbed using the WCE algorithm.

- Zhang et al. [2010] use a 4DVar filter to improve a ROMS model of the New York Bight Different datasets are assimilated. It is noteworthy that HFradar data degrades sub-surface temperature.

- Barth et al. [2011] implement an EnKF to correct the wind forcing field; multiple time instants are grouped in the statevector which is subsequently called the estimation vector.

- Yu et al. [2012] use a 4DVar filter to assimilate radar data, in order to improve the geometry of the upwelling SST front, and SSH. The wind stress is not corrected, but the state vector comprising (U,V,SSH,T,S). The error covariance is function only of the model-observations misfit. The model error covariance is diagonal and dependent on distance.

- Gopalakrishnan and Blumberg [2012] nudge HFradar in a 3D estuarine/coastal model, and show improvement of the model with respect to ADCP data. The paper contains a nice short history of radar assimilation.

- Supulveda et al. [2013] perform a twin-experiment using a 3DVar filter to assimilate data off northern Chile.

- Paduan and Washburn [2013] present a review article of radar current data utilization.

- Kurapov [2014] Assimilates sea surface height, temperature and radar data in the presence of a river plume

- Mermain et al. [2014] assimilate HFradar around Toulon to correct wind and open sea boundary conditions; rms errors on radial velocity and surface currents are about 0.2m/s, and improved very slightly; when errors are larger (around 0.3m/s), the improvement is larger ( 0.1m/s). Assimilation of radar data brings no significant changes on T-S profiles.

- Sperrevik et al. [2015] use ROMS-4DVar in an experiment off Norway, assimilating HF-radar and CTD data. The free simulation has relatively bad velocities (compared to drifters). Assimilating radar velocities improves the current field. Assimilating also CTD profiles does not improve surface velocity, but improves the density field.

In general, most authors recommend a 24h window to assimilate radar data. Recent studies usually assimilate directly the radial currents (with respect to the radar position), or if they assimilate orthogonal currents (i.e. interpolated on the model grid), they recommend to switch to the original radial currents in the future.

In this report, we assimilate HF radar data in a ROMS model of the Ligurian Sea.

# Chapter 2

# ROMS model in the Ligurian Sea

The hydrodynamic model is the ROMS model with 1/60 degree horizontal resolution and 32 vertical *s* layers. Open boundary conditions are obtained from the Mediterranean Forecasting System (MFS). Atmospheric forcing fields are from the COSMO forecasts. The model is run during the Recognized Environmental Picture campaign that the took place during summer 2010 (REP'10). The Ligurian sea is conditioned by the Liguro-Provencal Current, which is created by the joined Eastern and Western Corsican Currents. The region is also the siege of large mesoscale activity; as well as inertial oscillations with a period of approximately 17 hours.

An ensemble of models is generated by perturbing the open sea boundary conditions and the wind field, and by adding a stochastic term to the momentum equations represented by the last term in the right-hand side of:

$$\frac{d\mathbf{u}}{dt} + \Omega \wedge \mathbf{u} = -\frac{1}{\rho_0}\nabla_h p + \frac{1}{\rho_0}\nabla \cdot \mathbf{F^u} + \nabla_h \wedge \epsilon \mathbf{e}_z \tag{2.1}$$

This does not create horizontal divergences or convergences, and can create (absent or misplaced) mesoscale features in the flow.

The ensemble is then spun up for 1 week with all the perturbations, in order for the members to create different mesoscale circulation features. Hence during the subsequent actual experiment, all members have different initial conditions and boundary conditions. The model error covariance is subsequently estimated by the ensemble covariance matrix. For example, after the 1-week spin-up, the ensemble surface velocity spread is about 10 cm/s. The spatial correlation is about 50km (for temperature) and 10 km (for velocity). The obtained ensemble should represent the variability at all spatial and temporal scales represented in the simulation.

During the REP'10 experiment, HF-radar currents, satellite SST images, and glider data were obtained. Two WERA radars are operated by the NATO Undersea Research Center (NURC, now CMRE) in the Insula Palmaria and San Rossore location; they provide radial current fields with an azimuthal resolution of 6 degrees. These are smoothed in the azimuthal direction as a function of distance to the radar to account for loss of precision at larger distances. We average the observed currents over 1 hour. Figure 2.1 shows the model domain,
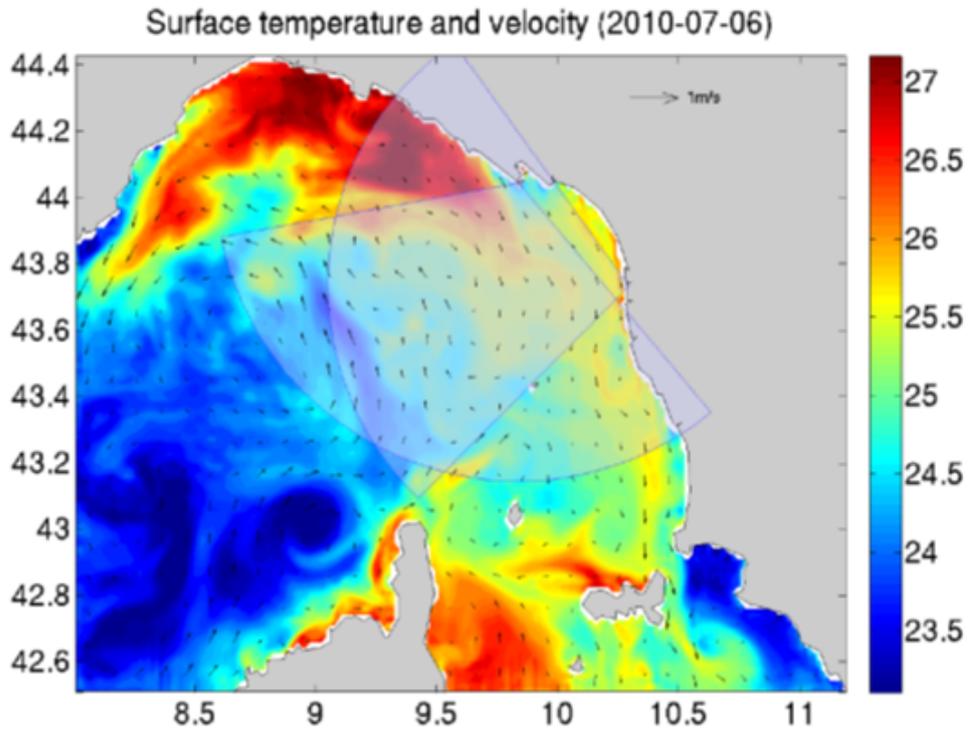
Figure 2.1: Model domain, and temperature and velocity forecast for 06/07/2010. The coverage of radar observations is shown as well.

the simulated temperature and velocity on 06/07/2010 and the area covered by the 2 radars.

# Chapter 3

# Assimilation of HF radar currents

Data assimilation is performed by the Ocean Assimilation Kit (OAK) software implementing the Ensemble Kalman Filter (EnKF). The state vector may contain different variables at different time steps, in which case it is called the estimation vector. In our setup, during a 2-day window, each member of the ensemble saves 48 hourly-averaged outputs which are all assembled in the estimation vector. The filter is then closely related to the Asynchronous Ensemble Kalman Filter (AEnKF).

The estimation vector may also contain uncertain forcing fields such as the wind field, the initial condition, the open boundary condition, and the stochastic error term (at one or more instants).

The model currents are transformed into radial currents by the observation operator, according to

$$u_{HF} = \frac{k_b}{1 - exp(-k_b h)} \int_{-h}^{0} \mathbf{u}(z) \cdot (e)_r \exp k_b z \, dz \qquad (3.1)$$

where $k_b = \frac{2\pi}{\lambda_b}$, and $\mathbf{e}_r$ is the unit vector poiting in the opposite direction to the location of the radar. Positive values hence represent currents away from the system. The operator essentially represents an average over the upper meters.

The points in the dense field of radar velocity observations are not uncorrelated. However, in the current implementation of the EnKF, the observation error covariance matrix $\mathbf{R}$ is diagonal. Hence, we strongly increase the "representativity" error component in (the diagonal of) $\mathbf{R}$:

$$\mathbf{R} = \mathbf{R_{instrum}} + \mathbf{R_{repr}} \qquad (3.2)$$

where $\mathbf{R_{instrum}}$ is the instrumental (measurement) error and $\mathbf{R_{repr}}$ needs to be determined.

In the present tests, the estimation vector $\mathbf{x}$ contains the 48 hourly-averaged radial velocities, and the temperature field at the end of the 2-day window. For memory-saving reasons, the estimation vector does not contain all the other model variables. However, the data assimilation procedure saves the local coefficients $\alpha$ of the linear combination of ensemble members (forecasts), point-wise
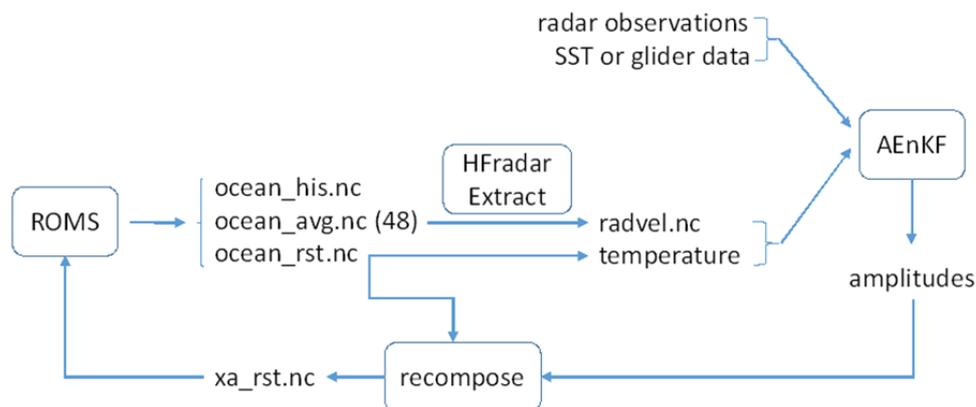
Figure 3.1: Succession of model integration, observation operator, data assimilation scheme, and recomposition.

yielding the analysis and the analyzed members.

$$\mathbf{x_a} = \mathbf{x_f} + \sum \alpha_i \mathbf{x_i} \qquad (3.3)$$

Then, the same linear combination can be applied to the other model variables or forcing fields such as the wind forcing. The procedure is illustrated in Figure 3.1.

The rms error between the free run and the radar observations is shown by the black curve in Fig. 3.2. Errors are between 10 and 15 cm.s$^{-1}$ except around the end of June 2010, when they are over 20 cm.s$^{-1}$. The data assimilative run uses 48 hourly-averaged velocity fields to correct the instanteneous 'restart" velocity (which itself is not observed). The corresponding correction is of the same order of magnitude as the ensemble spread. Root mean square errors are represented by the dotted blue curve. Errors are reduced by the data assimilation procedures, but the model tends to increase the errors again. However, they usually keep smaller than the free run. During the period when the free model errors were larger, data assimilation managed to significantly reduce them.
In order to diminish the tendency that the model has to generate errors, another experiment with strongly increased values in the observation error variance matrix was carried out; results are shown in the pink curve. Furthermore, other experiments were carried out, with even larger observation errors, with more localized corrections, with different time-window lengths (24 instead of 48 hours), with assimilation of only radar data closer than 50km to the radar, or with correction only of the model velocity (without updating the model temperature, salinity and surface elevation). None of these experiments could decrease the rms errors further (not shown).
Another experiment consisted of also adding the wind forcing in the estimation vector, and obtaining a "corrected" wind field. In practice, the corrected wind field is obtained afterwards, during the recomposition step. The model is then run one more time using the corrected wind. The obtained rms errors are represented
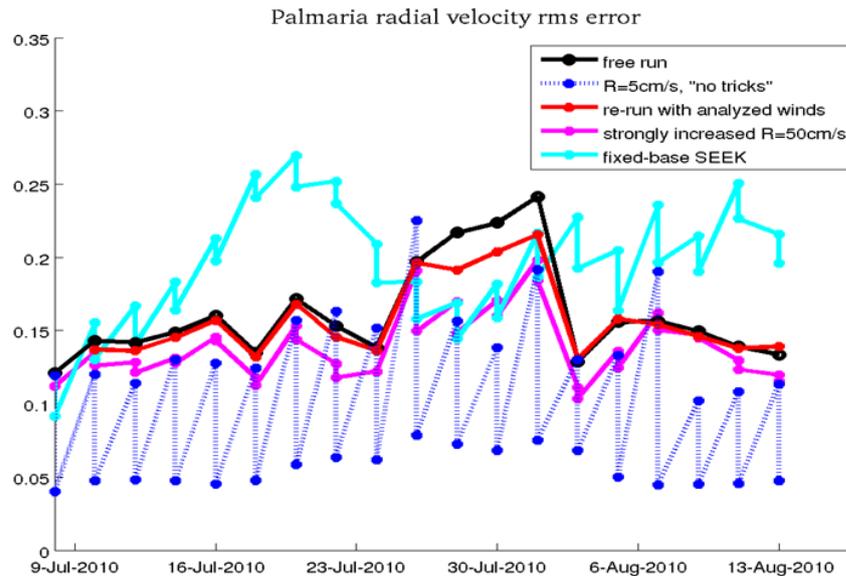
Figure 3.2: Root mean square difference between model and observation for different cases.

by the red curve in Fig. 3.2; they are not smaller than by direct correction of the model fields.

As a further test, we assimilate the 2 radar velocities in a single point, once per 48-hour window. The correction (also called increment) to hourly-averaged velocities, obtained at that particular point, is represented by the blue curve in Fig. 3.3. The coefficients of the analyzis change every 48 hours, which explain the discontinuities. Inside (some of) the 48-hours windows, the corrections present the typical 17-hour period of the inertial oscillations. The forecast, observation and analyzis (for the second 48-hour period) are shown in Fig. 3.4. The single assimilated observation is represented by a red arrow; the other observations are plot for reference only. One can observe that although a single linear combination is built for the 48 hours of the assimilation window, the result of the assimilation is so, that the correction is important mostly during the first 12 hours, i.e. close to the observation. During this period, the velocity is changed from north-eastward to south-eastward.

When using 48 hourly-averaged velocities during the 48-hour window, but still only in 1 single point in space, the corresponding correction is shown by the red curve in Fig. 3.3. One can see that the inertial oscillation correction is much stronger. This shows the beneficial impact of having very frequent observations (i.e. every hour).

During the REP'10 campaign, surface drifters were also launched. However, rms velocity errors between model and drifters are very large (27 cm.s$^{-1}$), and between radar and (projected) drifters as well (25 cm.s$^{-1}$). This may be due to
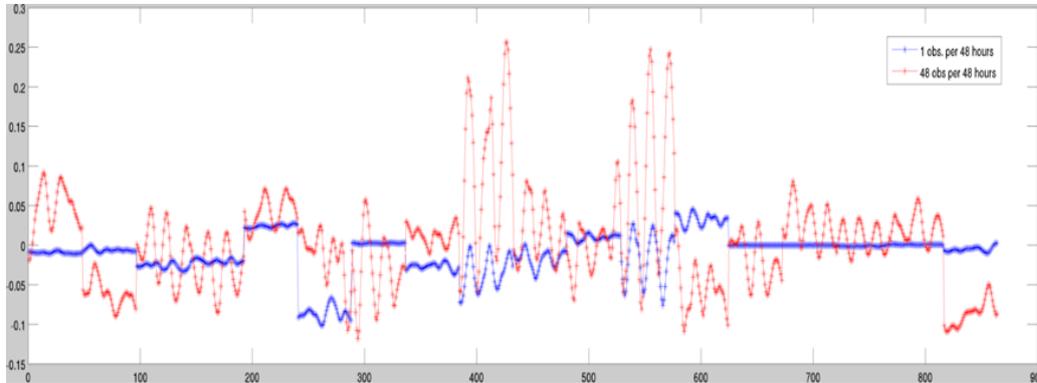
Figure 3.3: Analysis velocity increment when assimilating velocity in a single point. The blue curve corresponds to the case when assimilating one data-point per time-window, the red curve to the case when assimilating 48 data-points.
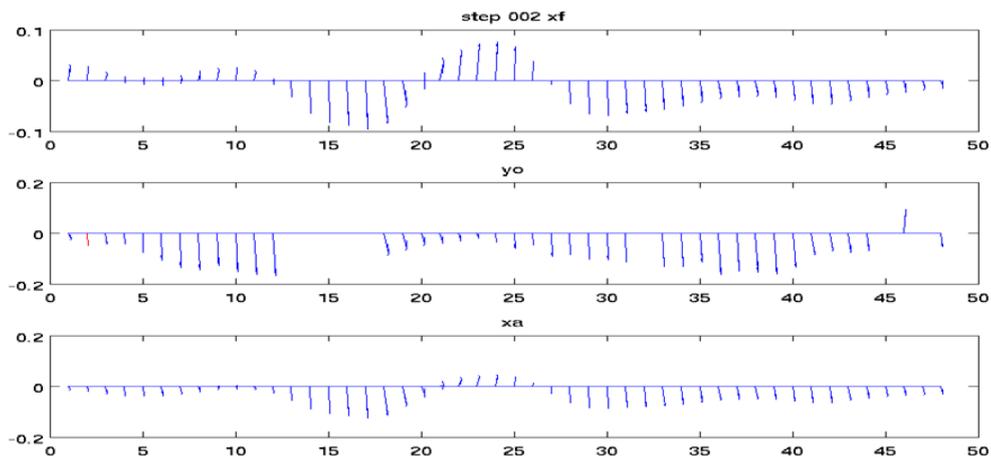


Figure 3.4: Upper panel: model velocity forecast during the 48-hours of the second assimilation window. Middle panel: corresponding observations; the assimilated observation is represented in red. Lower panel: model velocity analysis.

the fact that hourly-averaged model and radar velocities do not correspond to the mean velocity measured by the drifter between 2 transmissions (usually larger than 6 hours, i.e. over one third of the inertial oscillation period). It may also be caused by outliers with huge errors (up to 70 cm.s$^{-1}$); this is further investigated.

# Chapter 4

# Assimilation of radar and temperature observations

Assimilating radar currents leads to a (small) deterioration of the rms error between model and observed (satellite) sea surface temperature (SST). The errors are represented in Fig. 4.1; the blue curve is the free run, whereas the green curve is the rms SST error when assimilating radar observations. This unfortunate result was also obtained in other, recent studies [Zhang et al., 2010, Sperrevik et al., 2015]. When assimilating radar velocities and satellite SST, the error was reduced, as shown by the red curve in Fig. 4.1.

Comparisions of the independent temperature measurements realized by the drifters, with the "free" model run (i.e. assimilation of radar velocity only) and the SST-assimilating model are represented respectively by the blue and green curves in Fig. 4.2. One can observe a reduction of the rms errors in the latter case.
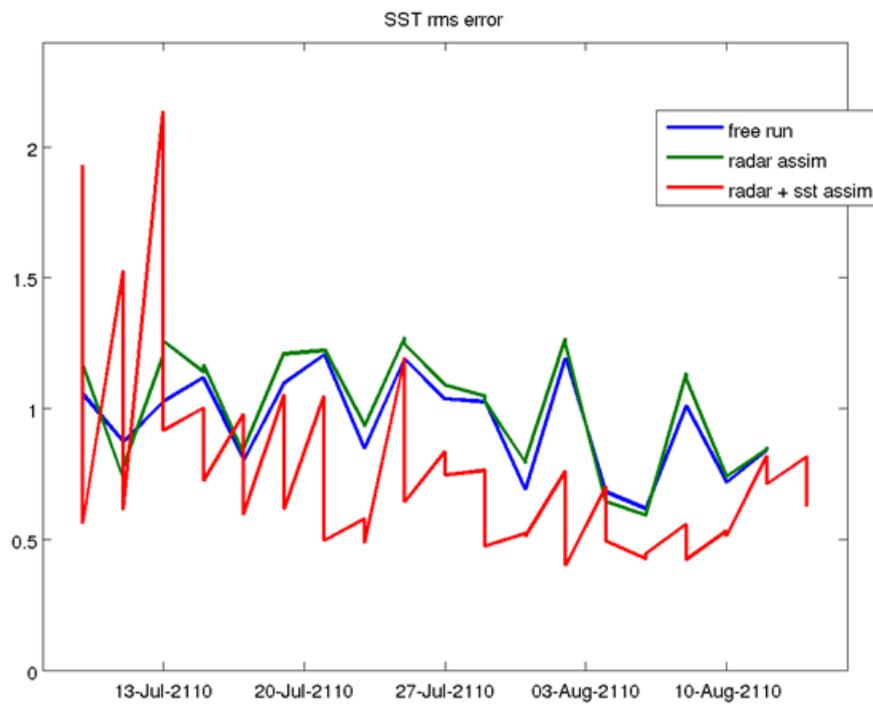
Figure 4.1: Rms SST error of the free run (blue curve), the run assimilating radar velocities (green curve) and the run assimilating both radar and SST observations (red curve) [C].

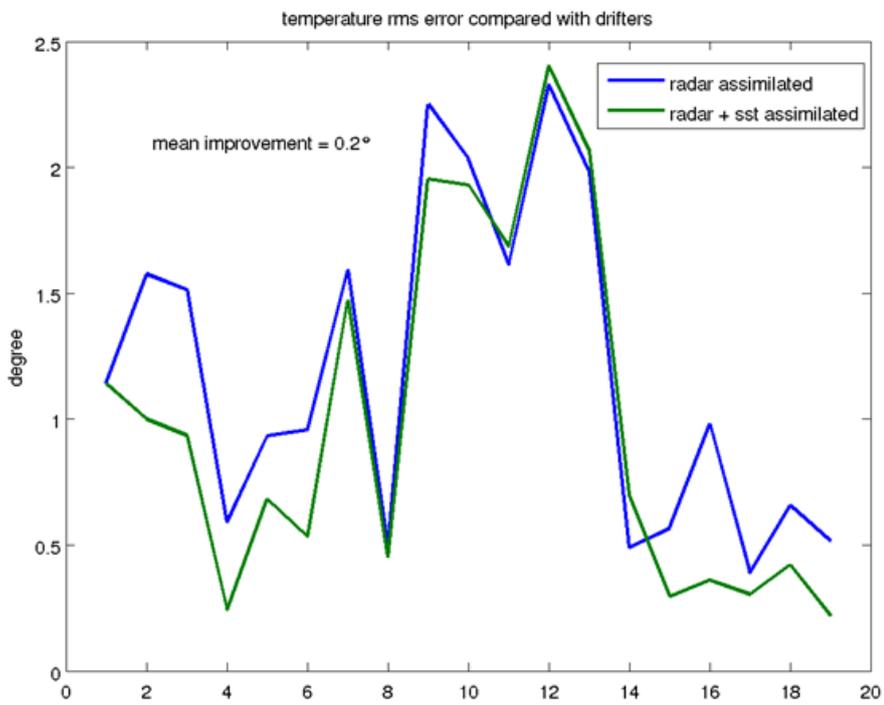Assimilation of radar and temperature observations

Figure 4.2: RMS temperature error between drifters and model forecasts when assimilating only radar velocities (blue curve) or when assimilating SST images as well (green curve).

# Chapter 5

# Conclusion

In this report, we analyzed wether radar observations could improve a regional model of the Ligurian Sea, using an ensemble of ROMS models. The ensemble is generated by perturbing the open sea boundary conditions and the wind forcing, and by adding a supplementary stochastic term to the momentum equations. The estimation vector used by the Ensemble Kalman Filter (EnKF) contained multiple instances of hourly-averaged currents (48 instances in this case), making the EnKF closely related to the Asynchronous Ensemble Kalman Filter (AEnKF) and the Ensemble Kalman Smoother (EnKS).

The results showed that the radar data somewhat allow to somewhat reduce the discrepancy between model velocity and radar observations, particularly when the model itself departs more strongly from observations. Adding the wind forcing in the estimation vector so as to obtain an analyzed forcing field, did not yield better results.

The high temporal frequency of observations (hourly, in this case) allows to correct the inertial oscillations in the model, or at least, modify their phase.

Surface drifter velocities could not be compared to model or radar data, because they represent an average over a much longer period and/or due to the presence of outliers.

Assimilation of radar data does not allow to improve model SST; actually model SST is slightly degraded. However, assimilating both radar velocities and satellite SST however, allows to improve the model SST. This is confirmed by comparisons with drifter temperature measurements.

# Bibliography

A. Barth, A. Alvera-Azcarate, K.-W. Gurgel, J. Staneva, A. Port, J.-M. Beckers, and E.V. Stanev. Ensemble perturbation smoother for optimizing tidal boundary conditions by assimilation of high frequency radar surface currents - application to the German Bight. *Ocean Science*, 6:161–178, 2010.

Alexander Barth, Aida Alvera-Azcarate, Jean-Marie Beckers, Joanna Staneva, Emil V. Stanev, and Johannes Schulz-Stellenfleth. Correcting surface winds by assimilating high-frequency radar surface currents in the German Bight. *Ocean Dynamics*, 61:599–610, 2011.

O. Breivik and O. Saetra. Real time assimilation of HF currents into a coastal ocean model. *Journal of Marine Systems*, 28:161–182, 2001.

Yi Chao, Zhijin Li, John Farrara, James McWilliams, James Bellingham, Xavier Capet, Francisco Chavez, Jei-Kook Choi, Russ Davis, Jim Doyle, David Fratantoni, Peggy Li, Patrick Marchesiello, Mark Moline, Jeff Paduan, and Steve Ramp. Development, implementation and evaluation of a data-assimilative ocean forecasting system off the central California coast. *Deep-Sea Research II*, 56:100–126, 2009.

G. Gopalakrishnan and A. F. Blumberg. Assimilation of HF radar-derived surface currents on tidal timescales. *Journal of Operational Oceanography*, 5:75–87, 2012.

David Kaplan and Francois Lekien. Spatial interpolation and filtering of surface current data based on open-boundary modal analysis. *Journal of Geophysical Research - Oceans*, 112, 2007.

Alexander Kurapov. Improvements in the Oregon coastal ocean forecast system: data assimilation in the presence of the Columbia River plume. *presentation*, 2014.

Alexander Kurapov, Gary Egbert, J. S. Allen, Robert N. Miller, Svetlana Y. Erofeeva, and P. M. Kosro. The M2 internal tide off Oregon: Inferences from data assimilation. *Journal of Physical Oceanography*, 33, 2003.

J. Lewis, I. Shulman, and A. Blumberg. Assimilation of Doppler radar current data into numerical ocean models. *Cont. Shelf Research*, 18:541–559, 1998.

J. Mermain, A. Molcard, P. Forget, A. Barth, and Y. Ourmieres. Assimilation of HF radar surface currents to optimize forcing in the northwestern Mediterranean Sea. *Nonlin. Processes Geophys.*, 21:659–675, 2014.

Peter R. Oke, J. S. Allen, Robert N. Miller, Gary D. Egbert, and P. M. Kosro. Assimilation of surface velocity data into a primitive equation coastal ocean model. *Journal of Geophysical Research*, 107, 2002.

Jeffrey D. Paduan and Libe Washburn. Observations of ocean surface currents. *Annu. Rev. Mar. Sci*, 5, 2013.

Igor Shulman and Jeffrey Paduan. Assimilation of HF radar-derived radials and total currents in the Monterey Bay area. *Deep-Sea Research II*, 56:149–160, 2009.

A. K. Sperrevik, K.H. Christensen, and J. Rohrs. Contraingin energetic slope currents through assimilation of high-frequency radar observations. *Ocean Science*, 11:237–249, 2015.

Hector Supulveda, Patrick Marchesiello, and Zhijin Li. Ocean data assimilation study in northern Chile: use of a 3DVAR method. *Lat. Am. J. Aquat. Res.*, 41: 570–575, 2013.

J.L. Wilkin, H.G. Arango, D.B. Haidvogel, C.S. Lichtenwalner, S.M. Glenn, and K.S. Hedstrøm. A regional ocean modeling system for the Long-term Ecosystem Observatory. *Journal of Geophysical Research*, 110, 2005.

P. Yu, A. L. Kurapov, G. D. Egbert, J. S. Allen, and P. M. Kosro. Variational assimilation of HF radar surface currents in a coastal ocean model off Oregon. *Ocean Modelling*, 49:86–104, 2012.

Weifeng G. Zhang, John L. Wilkin, and Hernan G. Arango. Towards an integrated observation and modeling system in the New York Bight using variational methods. Part I: 4DVAR data assimilation. *Ocean Modelling*, 35:119–133, 2010.