SCIENCE AND TECHNOLOGY ORGANIZATION
CENTRE FOR MARITIME RESEARCH AND EXPERIMENTATION

**Reprint Series** **CMRE-PR-2019-133**

# A relocatable EnKF ocean data assimilation tool for heterogeneous observational networks

Silvia Falchetti, Alberto Alvarez, Reiner Onken

June 2019

Originally published in:

OCEANS 2015, 18-21 May 2015, Genoa, Italy,
doi: 10.1109/OCEANS-Genova.2015.7271359

# About CMRE

The Centre for Maritime Research and Experimentation (CMRE) is a world-class NATO scientific research and experimentation facility located in La Spezia, Italy.

The CMRE was established by the North Atlantic Council on 1 July 2012 as part of the NATO Science & Technology Organization. The CMRE and its predecessors have served NATO for over 50 years as the SACLANT Anti-Submarine Warfare Centre, SACLANT Undersea Research Centre, NATO Undersea Research Centre (NURC) and now as part of the Science & Technology Organization.

CMRE conducts state-of-the-art scientific research and experimentation ranging from concept development to prototype demonstration in an operational environment and has produced leaders in ocean science, modelling and simulation, acoustics and other disciplines, as well as producing critical results and understanding that have been built into the operational concepts of NATO and the nations.

CMRE conducts hands-on scientific and engineering research for the direct benefit of its NATO Customers. It operates two research vessels that enable science and technology solutions to be explored and exploited at sea. The largest of these vessels, the NRV Alliance, is a global class vessel that is acoustically extremely quiet.

CMRE is a leading example of enabling nations to work more effectively and efficiently together by prioritizing national needs, focusing on research and technology challenges, both in and out of the maritime environment, through the collective Power of its world-class scientists, engineers, and specialized laboratories in collaboration with the many partners in and out of the scientific domain.

# A Relocatable EnKF Ocean Data Assimilation tool for heterogeneous observational networks

Silvia Falchetti, Alberto Alvarez and Reiner Onken
NATO Science & Technology Organization (STO)
Centre for Maritime Research and Experimentation (CMRE)
La Spezia, Italy
silvia.falchetti@cmre.nato.int

*Abstract*—**This study investigates the performance of a multivariate Ensemble Kalman Filter coupled with a relocatable limited-area configuration of the Regional Ocean Modeling System to predict ocean states by assimilating a heterogeneous data set involving underwater gliders and ship observations. In particular, two different ensemble initialization techniques are exploited and evaluated with the dataset collected during the REP13-MED experiment conducted by CMRE on 5-20 August 2013 in the Ligurian Sea. Results show that the forecast skill is significantly improved when the free ensemble is initialized from a long term climatology of the Mediterranean Forecast System. In particular the results obtained reveal significant increased skills in salinity forecasting in comparison with the previous ensemble initialization technique [6].**

***Keywords—data assimilation; ensemble Kalman filter; regional ocean predictions; underwater gliders; ocean observing networks***

## I. Introduction

Observational oceanography is transitioning from ship-based to heterogeneous ocean observing networks involving static nodes (moored profilers, bottom mounted systems), nodes with uncontrolled motion (drifter buoys and profiling floats) and nodes with controlled motion (ships, autonomous underwater vehicles-AUVs, gliders and autonomous surface vehicles-ASVs). This in situ observing system is complemented by remote sensing platforms/sensors. The information received from the in situ network and remote sensors are then assimilated into numerical ocean models to produce an analysis and forecast of the marine environment. Present ocean observatories such as the US Integrated Ocean Observing System (IOOS) and the Australian Integrated Marine Observing System (IMOS) among others, are heading towards this integrated concept.

Data assimilation, which aims to estimate the state and uncertainty of the ocean as accurately as possible by combining all available information (including model forecasts and observations, and their respective uncertainties) constitutes an essential component of these integrated systems. In the data assimilation community, the ensemble Kalman filter (EnKF) [1], is drawing increasing attention due to its ease of implementation and its ability to forecast ocean states and their corresponding uncertainty. Ensemble filters have now been implemented in several real applications with state-of the art ocean models [2], [3] and atmosphere models [4], [5] at the global scale. Less attention has been devoted to the synoptic scale and mesoscale in EnKF ocean limited area models (LAMs) [6], which is the focus of the current study.

The recent review [7] for LAM application of the EnKF in numerical weather prediction indicates the ensemble initialization as still a difficult and open topic among the other issues requiring particular treatment. The ideal method suggested would be to use a consistent global ensemble forecast system to directly provide the initial and boundary perturbations. If such ensemble is not available, the most common alternative is to randomly sample the climatological uncertainties of the initial states [8] or to derive random perturbations from the background error statistic of an existing 3D/4DVar system [9].

This study aims to compare two different ensemble initialization techniques for a LAM version of the EnKF applied to the Regional Ocean Modelling System (ROMS, [10]) nested in the Mediterranean Forecast System (MFS, [11]). The region under study is the Ligurian Sea, where the field experiment MED-REP13 was conducted by CMRE to investigate the operational feasibility and benefits, if any, of different concept of operations (CONOPS) to characterize the environment in denied areas using observational networks constituted by different platforms. The paper is organized as follows: Section 2 summarizes the experimental design including the description of the sea trial experiment and the main Ligurian Sea oceanographic circulation. Section 3 is dedicated to the modelling and data assimilation system. Section 4 reports the performance obtained by the different initialization techniques. This performance is measured in terms of the accuracy obtained in the forecast along certain validation sections inside the area of experimentation. The results are finally discussed in Section 5.

## II. Mediterranean Rapid environmental picture (MED-REP 13)

The MED-REP13 field experiment was conducted by the NR/V Alliance, on 5-20 August 2013 in a marine area offshore La Spezia (Italy) in the Ligurian Sea (Western Mediterranean Sea). The sampled region was nearly rectangular with approximately 90 km x 70 km in the along-shore and cross-shore directions, respectively. Depths range from about 50 m to almost 1800 m in this area. The following section summarizes only the aspects of the design of the field experiment relevant for the objective of this study.

### *A. Ocean observing network*

The experiment consisted in comparing the performance of different configurations of heterogeneous ocean observing networks. Hereinafter, only the ship-glider ocean observing network is presented which was implemented during the first phase of the experiment 07-13 August. In this networked configuration, the glider platform sampled the interior of the area under consideration while external boundary information was provided by towing a ScanFish by NR/V Alliance (Fig. 1).

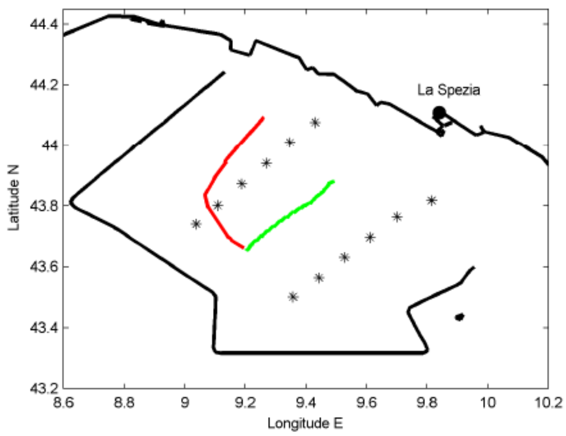


Fig 1. Trajectory of glider "Jade" during August 5–7 August (green), 7–9 August (red). The grey line represents the track of the Scanfish while the CTD locations are represented by asterisks.

The glider was driven by an adaptive sampling procedure guided by the predicted ocean uncertainties [12]. Assimilation cycles of two days were considered to characterize the area. Each assimilation cycle started at 8:00. The sampling of the boundary condition with the ScanFish was initiated at the southern boundary at 19:00 of the same day and finished at 19:00 of the following day. At 7:00, uncertainties from the different forecasts were sent to NR/V Alliance where the information was used to compute optimum sampling trajectories for the glider for the next 48 hours cycle. This procedure was repeated for three cycles. Only the first two assimilation cycles are investigated in the following. This is because the impact of initial conditions are more remarkable during these cycles.

For the purpose of collecting a validation dataset for performance assessment, oceanographic conditions were sampled 8-14 August inside the observational domain on night time and included conductivity, temperature and depth (CTD) casts. CTD samples were distributed along two transects perpendicular to the coast at half distance between the southwest-northeast axis of symmetry and the north and southern boundaries, respectively (Fig. 1).

### *B. Ligurian Sea main oceanographic caracterization*

Oceanographically, the region under study is marked by the presence of mesoscale and submesoscale variability associated with the Western Corsica Current (WCC) and the Northern Current (NC). The NC is formed to the north of Corsica, where the WCC encounters the Eastern Corsica Current (ECC). The result of the large scale circulation pattern, composed of the northwards flowing WCC and the NC, is a deep reaching cyclonic gyre [13]. A significant dynamical variability is superimposed to the large scale circulation patterns described above due to the meandering nature of the NC [14], as well as the presence of intense eddy activity north of Corsica [15].

The high mesoscale and submesoscale variability, together with the presence of different water masses from Atlantic and Mediterranean origin and the presence of coastal fronts associated with the local river discharge, makes the prediction of regional salinity field particularly challenging. A recent study [6] provided evidence of poor EnKF performance in ocean salinity prediction in the shelf area south of La Spezia. Salinity is predicted with an error of 45% (estimated as the ratio of the averaged root-mean-square difference (RMSD) to the averaged climatology variability of the area) against a temperature prediction error of 12%. Such a behavior might be driven by an underestimated ensemble spread in the ocean salinity which could prevent the system to benefit from the data assimilation analysis stage.

## III. MODELLING AND DATA ASSIMILATION SYSTEM

The ocean data assimilation and prediction system documented in [6] has been used. In the following, only the main details of the model implementation are given.

The circulation of the Ligurian Sea was simulated by means of a regional configuration of ROMS. ROMS is a primitive-equation, finite-difference, hydrostatic and free-surface model using generalized terrain-following vertical s-coordinates. The model domain covers the entire Ligurian Sea, with two open boundaries on the western and southern sides located at 8ºE and 42.5ºN, respectively. The horizontal resolution of the model grid is around 1.8 km. The vertical grid uses 32 vertical levels, which are non-linearly stretched to allow a finer resolution of the surface boundary layer. The 7-km resolution atmospheric model COSMO-ME of the Italian Air Force National Meteorological Centre [16] provides the atmospheric data necessary to compute the surface fluxes of momentum, heat and freshwater.

### *A. Ensemble Kalman Filter*

An asynchronous multivariate EnKF formulation was used to assimilate salinity and temperature observations at a time different from the time of the analysis. This advanced sequential data assimilation scheme uses an ensemble of perturbed model simulations (here 100 members) to approximate the model error covariance and their spatio-temporal variability. The algorithm advances in two steps: an analysis step and a forecast step.

Given the vector of glider observations $y_i$ collected between $t_i$−48h and $t_i$ and an ensemble of model forecasts simulating this period, the present implementation of the method produces an analysis ensemble of model states at the time $t_i$. The augmented state vector $x_i$ contains the model state variables at time $t_i$ and the model values at the position and time of the observations collected during the past 48 hours. The following equation is used to update the ensemble mean state vector:

$$\overline{x}_i^a = \overline{x}_i^f + K_i\left(y_i - H_i\overline{x}_i^f\right) \quad (1)$$

where overbars denote the ensemble mean and the superscripts $f$ and $a$ represent the forecast and analysis states, respectively. $H_i$ is the observation operator projecting the model state onto the observation space. $K_i$ is the kalman gain, which takes into account the observation error $R_i$ and the background error covariance $P_i^f$ in determining the extent to which observations are weighed relative to the background:

$$K_i = P_i^f H_i^T\left(H_i P_i^f H_i^T + R_i\right)^{-1} \quad (2)$$

In this expression, the superscript $T$ denotes a matrix transpose.

Given the analysis ensemble, the forecast step simply involves forecasting each member forward to the time when the next observations are available (here 48 hours).

### B. *Ensemble Initialization*

The ensemble initialization used in [6] consisted of a 100-member ensemble, including perturbations of the initial conditions, winds and lateral boundary conditions. The ensemble initialized date was at least ten days before the first assimilation cycle using model states selected among the ROMS free run from plus-minus five days the ensemble initialization date. The wind perturbations were generated by computing the bivariate empirical Ortogonal Functions (EOFs) time series computed from the COSMO-ME model zonal and meridional wind fields. The boundary conditions were perturbed using the bivariate EOFs of temperature and normal velocity time series provided by the MFS model.

The first ensemble initialization technique investigated here, involves only perturbations of the initial conditions to achieve an optimal ensemble spread and reliability. Ensemble members are generated by taken snapshots of state values during July, August and September from a three years run of the MFS model [17]. Temperature, salinity, horizontal velocities and sea surface height were selected as state variables. Physically realistic perturbations of the mean state were built from the singular value decomposition (SVD) of the trajectory matrix (a $m$ x $p$ matrix where $m$ is the number of state variables at the model grid-points and $p$ is the number of snapshots extracted). The singular vectors of the matrix trajectory provide a means for initializing the ensemble, ensuring linear dependence of each ensemble deviation. Specifically, perturbations are generated by the summation of the singular values, each one weighted by a Gaussian random coefficient with zero mean and variance determined by the corresponding singular value. A 100-member ensemble free run was initialized for the glider-ship observing network on 26 July 2013, allowing twelve forecast days each to reach an equilibrium state before the first assimilation cycle.

The second technique proposed is also based on the generation of different initial conditions for the 100-member ensemble without wind and boundary perturbations applied. Ensemble members are generated by directly taken snapshots of state values during July, August and September (approximately every ten days) from a ten years climatology of the MFS. Such an increased long-time data set is thought to provide the necessary ensemble background error statistic and to provide an optimal ensemble spread, avoiding vertical static instabilities that could potentially be generated from the previous approach. Such an ensemble, which is generated from MFS dynamically balanced states, needs a much lower spin up time; here a two days spin up time was adopted. In order to produce an ensemble mean as close as possible to the real ocean state, the ensemble mean was replaced by MFS forecast available at the time of the first assimilation cycle (7 August 2013).

### C. *Observations*

The treatment of observations is done as in [6]. Salinity and temperature observations collected from the glider Jade and the Scanfish are first interpolated onto a fixed model grid stretched in the vertical before being assimilated. For each level, the observed variance in the vertical grid cell is used as an approximation of the vertical representation error. The horizontal representation error variance is assumed to be $(0.25°C)^2$ and $(0.05)^2$ for temperature and salinity, respectively.

## IV. RESULTS

The following results presented are relative to the two assimilation experiments performed. In the first experiment the ensemble is initialized with states built with the EOF of the covariance matrix resulting from a time series of 3 years of MFS results (hereinafter labelled simulation A). The second one uses the ensemble initialized from the ten years subsampled MFS climatology without any perturbations applied (hereinafter labelled simulation B). Two assimilation cycles were performed.

In the first cycle the salinity and temperature observations derived from glider Jade and Scanfish from 5 - 7 August are assimilated starting from the free ensembles A and B respectively. After the analysis, a 48 hours forecast is performed applying wind perturbations to the ensemble members. The algorithm used to generate such perturbations is the same detailed in [6]. These perturbations introduce a forcing error, which is one of the largest sources of error for ocean state estimation. The impact of the perturbed forcing is felt primarily in the upper ocean and is limited by the assimilation interval.

The choice of introducing only the forcing error was made in order to focus mostly on the role of the different ensemble initialization techniques, leaving out other source of model errors such as: (i) chaotic perturbations to the ocean mixing and diffusion [2], (ii) internal perturbations from a pre-computed ensemble of ocean integrations [2], and (iii) lateral boundary perturbations [6].

In the second cycle, assimilation of observations from glider Jade from 7 - 9 August is performed and then 48 hours forecast is produced with wind perturbations applied. The results presented in the following relate to this second forecast cycle. The performance of the system is evaluated in terms of taylor diagrams [18]) which represent the similarity between modelled and observed patterns by means of their correlation (CORR), their centered RSMD and their standard deviation (STD). The two CTD transects depicted in Fig. 1 are used for the validation purpose.

### *A. Validation of salinity prediction along CTD transects*

In order to quantify the agreement between the observed and predicted salinity patterns, Fig. 2 shows the Taylor diagram for the northern CTD transect on 10 August. Comparisons are limited to the first 150m of the water column to match results from [6]. Table 1 reports the corresponding skill performance for salinity and temperature prediction.

The two ensembles have very similar performance in predicting the ocean salinity, even if ensemble B is slightly superior to ensemble A in its skill. The upper panel shows the salinity prediction validation against the northern CTD transect. Ensemble B has the smallest RMSD of 0.09, which is significantly lower than the one reported by [6] which was above 0.2. The correlation of model and observation pattern is very high and equal to 0.89. Also the variability simulated with the model is quite close to the one observed; ensemble B has 0.16 STD against the 0.20 STD shown by the data.

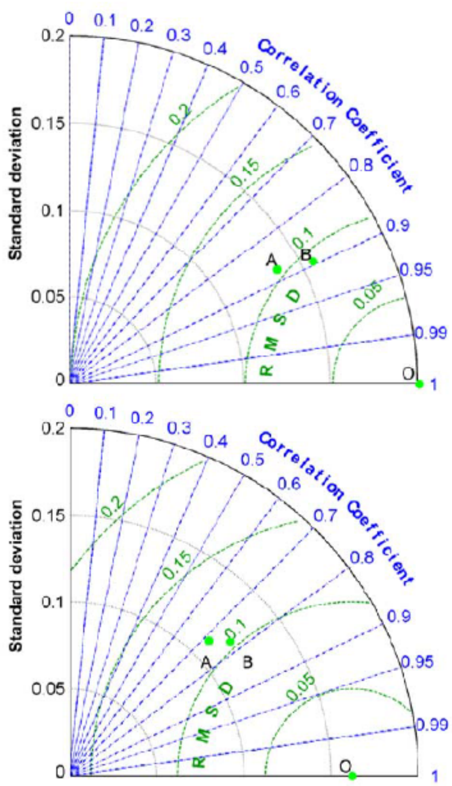


Fig. 2. Taylor diagrams for ocean salinity prediction for 10 August (northern CTD transect) and 11 August (southern CTD transect) on upper and lower panels respectively. Observations correspond to point O, the forecast ensemble mean with data assimilation are represented from point A (initialization from SVD analysis) and B (initialization from MFS climatology) respectively.

The lower panel shows the salinity prediction validation against the southern CTD transect. The performance of both ensembles is slightly worse than on the northern CTD transect. In this section observations show a lower variability (0.12 STD) with respect to the northern transect (0.16 STD). Ensemble B, as was the case for CTD north transect, shows higher skills in predicting CORR, RMSD and STD. Ensemble B RMSD is still very satisfactory, 0.10 with a correlation value of 0.77.

### *B. Validation of temperature prediction validation along CTD transects*

Ensemble B has higher skills than ensemble A in reproducing the temperature pattern for both northern and southern sections. The two ensembles, especially in the south transect, show a different behavior; ensemble B has 0.92 °C RMSD against 1.07 °C RMSD of ensemble A.

TABLE I. METRICS FOR MODEL-OBSERVATIONS STATISTICS

| | *Sim* | *S North* | *S South* | *T North* | *T South* |
|---|---|---|---|---|---|
| RMSD | A | 0.10 | 0.11 | 0.47 | 1.07 |
| | B | 0.09 | 0.10 | 0.46 | 0.92 |
| STD | A | 0.14 | 0.11 | 3.66 | 3.74 |
| | B | 0.16 | 0.12 | 3.55 | 3.50 |
| | obs | 0.20 | 0.16 | 3.45 | 3.29 |
| CORR | A | 0.88 | 0.72 | 0.99 | 0.96 |
| | B | 0.89 | 0.77 | 0.99 | 0.96 |

The northern section is reproduced better than the southern section; the RMSD value is 0.46 °C with a correlation value of 0.99.

## V. DISCUSSION AND CONCLUSION

In general the forecast performance resulting from the operation of the glider-ship network is superior in the northern validation section than in the southern one. Indeed the two sections are characterized by a completely different oceanographic situation as depicted in Fig. 3.

### *A. Validation at the CTD sections*

Fig. 3 shows the free mean ensemble salinity forecast at 50m for the simulation A and B on 7 August before the first assimilation cycle takes place. Both the two ensemble predict the deep reaching cyclonic gyre, which constitutes the main circulation pattern of the area. The NC associated gyre is characterized by a strong salinity front, which has been sampled by the northern CTD section.

The mean ensemble salinity derived from the singular value decomposition analysis presents a smoother salinity front with an elongated shape, contrary to the strong defined and circular front simulated by ensemble B. Indeed the mean ensemble B salinity coincides with the MFS daily mean product for 7 August, since the ensemble members are reported back to this mean value before starting the assimilation cycle. The ensemble mean standard deviation pattern for simulation B reflects the main circulation pattern; the highest uncertainty (0.2) is found in the middle of the salinity front and at the confluence of the WCC current and the ECC which originates from the Tyrrhenian Sea. Finally, the 0.18 standard deviation isoline follows approximately the NC gyre.

The uncertainty pattern found for ensemble B looks more realistic than the one characterizing ensemble A. The latter one shows isolated patches of constant uncertainty in the region, contrary to the well-defined pattern of ensemble B. Moreover the uncertainty is about three times lower than B. The highest uncertainties about 0.12 is found close to the western boundary (not shown here) suggesting that the longer spin up period of twelve days, might have caused an increased influence of the uncertainty due to the boundary condition for simulation A.

The southern CTD section crosses a region of Modified Atlantic Water (MAW) characterized by salinities lower than 38. The most offshore CTD stations are in the proximity of the region where the two currents, the ECC and WCC, converge. Specifically, this transect intersects the southern boundary of an anticyclonic eddy at its nearshore segment, and the inflow jet at its off-shore side (this evidence derives from the analysis of in situ ADCP measurements not shown here). Both circulation structures involve the re-circulation and inflow of MAW, respectively.

As shown in Fig. 3, ensemble B is characterized by about 0.1 lower salinity values in comparison to ensemble A. Southern CTD section measured salinities lower than 38 as ensemble B is suggesting for that forecast day.

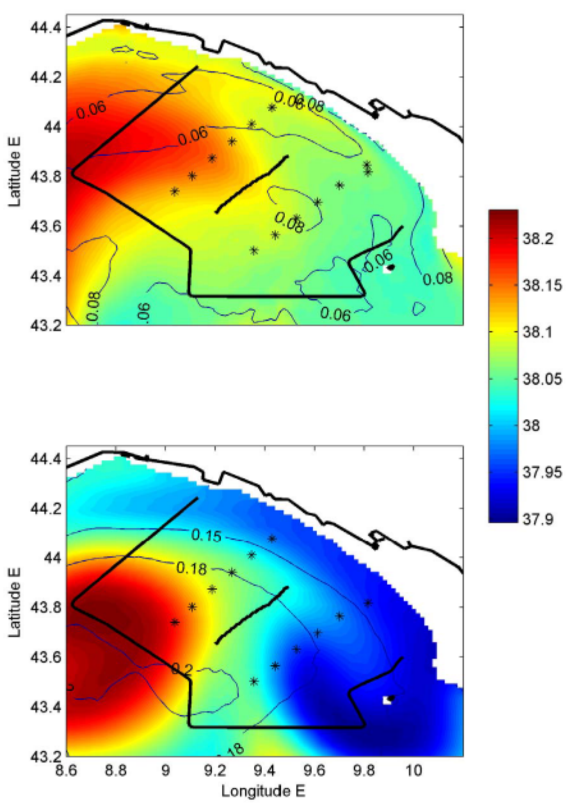


Figure 3. The free ensemble mean salinity forecast at 50 meters on 7 August 2013 (before the data assimilation analysis) for simulation B and A, upper and lower panels respectively. Contour lines represent the ensemble mean standard deviation. The black peripheral line represents the track of the Scanfish, the central black line represents the glider Jade 48 hours trajectory (5 - 7 August), which will be assimilated in the successive analysis step. CTD locations are represented by asterisks.

### B. *Ensemble Spread*

In order to further analyse the difference between the two generated ensembles, Fig. 4 shows the profiles of model background error for temperature and salinity variables at the most offshore station of the southern CTD transect.

The temperature shows a significant standard deviation in the upper 50 meters for both ensemble A and B. Only ensemble B seems to capture the existence of a mixed-layer in the first 10-15 m depth, which is consistent with the CTD temperature vertical sections. The analysis of the in situ observed profiles reveals a thermocline which develops between 15 to 40m. The ensemble B standard deviation has a maximum value of 2.2°C at 23m depth, against the maximum value of 1.8°C at 17m depth for ensemble A. The slightly deeper thermocline predicted by ensemble B is more consistent with the in situ evidence of a deeper thermocline, than the one predicted by simulation A.

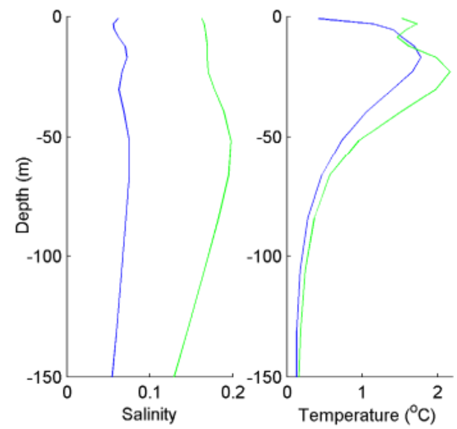


Fig. 4. Profile of salinity and temperature ensemble standard deviation (model background error) at southernmost CTD south station. Blue line and green lines represent ensemble A and B respectively.

The salinity model background error is maximum at 50m depth for both ensembles (0.08 and 0.2 for A and B respectively). Ensemble B has significantly higher standard deviation than A; consequently the analysis will be much more sensitive to the observations along all the profiles with respect to ensemble A.

### C. *Conclusions*

The evaluation of two different ensemble initialization techniques used with a LAM application of a multivariate asynchronous EnKF reveals significant improved skills in the ocean salinity prediction compared to the procedure outlined in [6]. The data assimilation was performed using observations collected by a heterogeneous ocean observing network composed by an underwater glider and a ship, while the validation was evaluated against CTD profiles along two parallel sections approximately perpendicular to the coast.

The ensemble generated from MFS ocean state snapshots taken from an extended ten years time series has superior prediction skills in comparison to the one generated from the EOFs of the covariance computed from a three years time series of MFS results. Two significant differences characterize the two ensemble initialization: (i) before the first assimilation cycle, the first ensemble mean is forced to be coincident with the daily mean ocean state predicted by MFS at the time of the analysis, (ii) the second ensemble requires more days to reach an equilibrium state than the first one due to the vertical static instabilities which derive from the singular value decomposition analysis.

These two main differences translate into two equivalent important aspects: (i) the first ensemble mean is closer to the in situ oceanographic situation especially with regard to the prediction of MAW waters across the south CTD validation station, (ii) the model background error of the first ensemble reflects the main circulation pattern of the region, while the second ensemble is more influenced by the lateral boundary condition given the higher spin-up period adopted.

The first ensemble shows superior skills also in the prediction of ocean temperature, particularly in the south CTD validation station.

## Acknowledgment

We acknowledge the Italian Air Force National Meteorological Center for provision of COSMO-ME data. We thank Dr. Baptiste Mourre for provision of the coupled EnKF-ROMS code used in this study. We are grateful to the CMRE staff for maintaining, piloting the gliders and Scanfish instruments and the Alliance crew for the operations conducted on the NRV Alliance. This work was funded by the Allied Command Transformation (ACT project number SAC000507).

## References

[1] Geir Evensen, "Inverse methods and data assimilation in nonlinear ocean models, Physica D: Nonlinear Phenomena", Volume 77, Issues 1–3, 1 October 1994, Pages 108-129, ISSN 0167-2789

[2] C. L. Keppenne, M. M. Rienecker, N. P. Kurkowski, D. A. Adamec. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. Nonlinear Process in Geophysics, European Geosciences Union (EGU), 2005, 12 (4), pp. 491-503.

[3] S. Zhang , M. J. Harrison, A. Rosati, A. Wittenberg, System Design and Evaluation of Coupled Ensemble Data Assimilation for Global Oceanic Climate Studies, Mon. Wea. Rev., 2007, 135, 3541-3564.

[4] J. S. Whitaker, T. Hamill, X. Wei, Ensemble Data Assimilation with the NCEP Global Forecast System, Mon. Wea. Rev., 2008, 136, 436-482.

[5] Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. Mon. Wea. Rev., 2005, 133, 604–620.

[6] Mourre, B., & Chiggiato, J. (2014). A comparison of the performance of the 3-D super-ensemble and an ensemble Kalman filter for short-range regional ocean prediction. *Tellus A, 66*.

[7] Meng, Z. and F. Zhang, Limited-Area Ensemble-Based Data Assimilation, Mon. Wea. Rev. 2011, 139.

[8] Aksoy, A., F. Zhang, and J. W. Nielsen Gammon (2006), Ensemble based simultaneous state and parameter estimation in a two dimensional sea breeze model, Mon. Weather Rev., 134, 2951–2970

[9] Torn, R. D., and G. J. Hakim, 2008: Ensemble Data Assimilation applied to RAINEX observations of Hurricane Katrina (2005). Mon. Wea. Rev. 137, 2817-2829 DOI: 10.1175/2009MWR2656.1

[10] Haidvogel, D.B., Arango, H., Budgell, W.P., Cornuelle, B.D., Curchitser, E., Di Lorenzo, E., Fennel, K., Geyer, W.R., Hermann, A.J., Lanerolle, L., Levin, J., McWilliams, J.C., Miller, A.J., Moore, A.M., Powell, T.M., Schepetkin, A.F., Sherwood, C.R., Signell, R.P., Warner, J.C., Wilkin, J. Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the Regional Ocean Modeling System,. J. Comput. Phys. 227, 2008, 3595–3624.

[11] N. Pinardi and G. Coppini, 2010. Operational oceanography in the Mediterranean Sea: the second stage of development Ocean Sci., 6, 263-267.

[12] Alvarez, A., Mourre, B., Optimum sampling designs for a glider-mooring observing network, Journal of Atmospheric and Oceanic Technology, 29 (4) , 2012, 601-612.

[13] Millot, C. (1999), Circulation in the Western Mediterranean Sea, J. Mar. Syst., 20(1-4), 423–442, doi:10.1016/S0924-7963(98)00078-5.

[14] Crepon, M., L. Wald, and J. M. Monget (1982), Low-frequency waves in the Ligurian Sea during December 1977, *J. Geophys. Res.*, *87*, 595, doi:10.1029/JC087iC01p00595.

[15] Casella, E., P. Tepsich, X. Coulevard, R. M. A. Caldeira, and K. Schroeder (2014), Ecosystem dynamics in the Liguro-Provençal Basin: the role of eddies in the biological production., Mediterr. Mar. Sci., 15(2), 274–286, doi:10.12681/mms.520

[16] Bonavita, M. and L. Torrisi, Impact of a variational objective analysis scheme on a regional area numerical model: The Italian Air Force weather service experience, Meteor. And Atmos. Phys.., Vol. 88, N.1-2, 2005, 39-52.

[17] Turner M. R. J., Walker J. P., Oke P. R., Ensemble member generation for sequential data assimilation, Remote Sensing of Environment, 112, 2008, 1421-1433.

[18] Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res., 106, 7183-7192, 2001.

# Document Data Sheet

| Security Classification | | Project No. |
|---|---|---|
| | | |
| **Document Serial No.** CMRE-PR-2019-133 | **Date of Issue** June 2019 | **Total Pages** 6 pp. |

**Author(s)**

Silvia Falchetti, Alberto Alvarez, Reiner Onken

**Title**

A relocatable EnKF ocean data assimilation tool for heterogeneous observational networks

**Abstract**

This study investigates the performance of a multivariate Ensemble Kalman Filter coupled with a relocatable limited-area configuration of the Regional Ocean Modeling System to predict ocean states by assimilating a heterogeneous data set involving underwater gliders and ship observations. In particular, two different ensemble initialization techniques are exploited and evaluated with the dataset collected during the REP13-MED experiment conducted by CMRE on 5-20 August 2013 in the Ligurian Sea. Results show that the forecast skill is significantly improved when the free ensemble is initialized from a long term climatology of the Mediterranean Forecast System. In particular the results obtained reveal significant increased skills in salinity forecasting in comparison with the previous ensemble initialization technique [6].

**Keywords**

Data assimilation, ensemble Kalman filter, regional ocean predictions, underwater gliders, ocean observing networks

**Issuing Organization**

NATO Science and Technology Organization
Centre for Maritime Research and Experimentation

Viale San Bartolomeo 400, 19126 La Spezia, Italy

*[From N. America:*
STO CMRE
*Unit 31318, Box 19, APO AE 09613-1318]*

Tel: +39 0187 527 361
Fax:+39 0187 527 700

E-mail: library@cmre.nato.int