Journal of Marine Systems 180 (2018) 76–89

Contents lists available at ScienceDirect

Journal of Marine Systems

journal homepage: www.elsevier.com/locate/jmarsys

The impact of covariance localization on the performance of an ocean EnKF system assimilating glider data in the Ligurian Sea

Silvia Falchetti[a], Alberto Alvarez[b,*]

[a] NATO Science & Technology Organization (STO), Centre for Maritime Research and Experimentation (CMRE), Viale San Bartolomeo 400, 19126 La Spezia, Italy
[b] Dept. of Marine Ecology (MARE), Instituto Mediterraneo de Estudios Avanzados-IMEDEA (CSIC-UIB), Miquel Marques 21, 07190 Esporlas, Spain

ARTICLE INFO

Keywords:
Data assimilation
Ensemble Kalman filter
Predictability
Regional ocean predictions
Underwater gliders
Ocean observing networks

ABSTRACT

Data assimilation through an ensemble Kalman filter (EnKF) is not exempt from deficiencies, including the generation of long-range unphysical correlations that degrade its performance. The covariance localization technique has been proposed and used in previous research to mitigate this effect. However, an evaluation of its performance is usually hindered by the sparseness and unsustained collection of independent observations.

This article assesses the performance of an ocean prediction system composed of a multivariate EnKF coupled with a regional configuration of the Regional Ocean Model System (ROMS) with a covariance localization solution and data assimilation from an ocean glider that operated over a limited region of the Ligurian Sea. Simultaneous with the operation of the forecast system, a high-quality data set was repeatedly collected with a CTD sensor, i.e., every day during the period from 5 to 20 August 2013 (approximately 4 to 5 times the synoptic time scale of the area), located on board the *NR/V Alliance* for model validation. Comparisons between the validation data set and the forecasts provide evidence that the performance of the prediction system with covariance localization is superior to that observed using only EnKF assimilation without localization or using a free run ensemble. Furthermore, it is shown that covariance localization also increases the robustness of the model to the location of the assimilated data. Our analysis reveals that improvements are detected with regard to not only preventing the occurrence of spurious correlations but also preserving the spatial coherence in the updated covariance matrix. Covariance localization has been shown to be relevant in operational frameworks where short-term forecasts (on the order of days) are required.

1. Introduction

An accurate estimation of the ocean state is required for many interdisciplinary applications, including acoustic, biological, physical and optical sciences and technologies (Lermusiaux et al., 2006). The dynamics of oceanic processes are nonlinear and highly variable, and they involve interactions across several temporal and spatial scales. These characteristics, which form the basis of the non-deterministic nature of ocean predictions, make such estimations unique and require sophisticated numerical ocean models (Brasseur, 2006). These systems are affected by an intrinsic predictability limit (Lorenz, 1969) with an inherent scale connected to the nonlinearity of the dynamical equations and to the errors in the initial conditions (Robinson and Sellschopp, 2002), the atmospheric forcing and the boundary conditions.

Data assimilation techniques aim to estimate the state of the ocean and the associated uncertainties as accurately as possible by integrating observations into ocean model states while using consistency constraints that obey the dynamical principles governing the observed system. Variational and sequential methods are among the most widely used data assimilation schemes for regional ocean systems, and there have been exciting recent advances in ensemble and four-dimensional variational approaches (Edwards et al., 2015). In particular, the ensemble Kalman filter (EnKF) (Evensen, 1994) has drawn increasing attention due to its ease of implementation and its ability to forecast ocean states and their corresponding uncertainties. Unlike the traditional Kalman filter, the EnKF can operate using nonlinear models by forecasting an ensemble of states to compute an ensemble mean and covariance, from which a single Kalman gain is derived. An analysis of each member of the ensemble allows for a mean analysis to be derived. Interested readers are referred to Evensen (2003) for further details on the EnKF approach.

Although EnKF data assimilation systems have been used in several real-world applications with state-of-the-art atmospheric models (Whitaker et al., 2008; Houtekamer et al., 2005) and ocean models at both global (Keppenne et al., 2005; Zhang et al., 2007) and regional scales (Mourre and Chiggiato, 2014), the implementation of such a

* Corresponding author.
E-mail addresses: alberto.alvarez@imedea.uib-csic.es, alberto.alvarez@imedea.uib.es (A. Alvarez).

https://doi.org/10.1016/j.jmarsys.2017.12.008
Received 18 April 2017; Received in revised form 19 December 2017; Accepted 31 December 2017
Available online 04 January 2018

system is not free of difficulties. The limited statistical representativeness of the ensemble is probably the most fundamental issue. This issue originates from the small size of the ensemble when compared with the number of states accessible to the dynamics of the ocean. Inbreeding (Houtekamer and Mitchell, 2001) and the development of long-range spurious correlations in the ensemble covariance (Anderson, 2001) may be caused by a statistical misrepresentation of the ensemble. The former refers to an underestimation of the analysis error covariance after each assimilation cycle, while the latter refers to the unphysical correlations between distant locations generated by the forecast covariance. Covariance inflation (Anderson and Anderson, 1999) and localization (Houtekamer and Mitchell, 2001) have been proposed to mitigate the effects of inbreeding and spurious correlations, respectively.

Data assimilated into ocean prediction models are collected using ocean observing systems. In conjunction with autonomous profiling floats, glider technology is being used to transform ocean observing technologies from individual platform-based designs to networks of sensor nodes. This observational approach has been integrated into many current ocean observatories, such as the US Integrated Ocean Observing System (IOOS) and the Australian Integrated Marine Observing System (IMOS). Gliders make use of buoyancy changes and utilize their low-drag hydrodynamic shapes to perform zig-zag motions between the surface and the bottom of the ocean, inducing a net horizontal displacement. Their nominal speed is approximately 0.5 m s^{-1} with spatial cycle periods that depend on the programmed pitch and immersion depths. Thanks to a buoyancy-based propulsion mechanism, the endurance of a glider can reach up to several months. The maneuverability of a glider, although limited to some degree by the strength of the velocity field, is another advantage offered by this technology.

A few studies in the literature have investigated the exploitation of gliders in operational ocean forecasting systems. A remarkable case was provided by the Autonomous Ocean Sampling Networks (AOSN)-II field experiment conducted in Monterey Bay, California, in 2003 (Ramp et al., 2009). The day-to-day physical variability in that ocean region was predicted using different ocean prediction models that assimilated data from heterogeneous observational assets, including a glider fleet (Chao et al., 2008; Lermusiaux, 2007; Shulman et al., 2009). The impact of the assimilation of glider observations improved the short-term (1–1.5 days) subsurface salinity and surface temperature forecasts (Shulman et al., 2009); meanwhile, for more extended forecasts, accurate atmospheric forcing data play a critical role. Further research has also demonstrated an improvement in the forecasting skills of ocean prediction models with the assimilation of glider observations of temperature and salinity fields (e.g., Zhang et al., 2010a, 2010b; Jones et al., 2012; Gangopadhyay et al., 2013; Mourre and Chiggiato, 2014).

The importance of assimilating all available information from gliders, including the vertically averaged velocity in addition to salinity and temperature observations, was emphasized by Dobricic et al. (2010) when they used an operational forecasting model for the Ionian Sea (eastern Mediterranean Sea). Mourre and Alvarez (2012) exploited the autonomy and maneuverability of gliders to explore the benefits of piloting a glider with an adaptive sampling procedure in the western Ligurian Sea under a fully operational framework. This adaptivity required a continuous feedback of information between the glider and an operational forecasting system based on a 3D super-ensemble (3DSE) assimilation technique (Lenartz et al., 2010). The above studies, among others, suggested that glider observations could significantly contribute to improving the performance of present-day operational forecasting systems.

An evaluation of the performance of an operational ocean forecast system is an important aspect of the system's development and exploitation. Determining the extent to which the events predicted by the model will compare to a corresponding set of independently obtained and reliable observations, is the most appropriate method to evaluate the performance of an operational forecasting system (Willmott et al., 1985). While the statistical measures of a model's performances are well defined, the difficulties in assessing those performances are often generated by the sparseness and unsustained collection of independent observations due to economic and/or operational limits generally encountered when observations are needed to support data assimilation and model validation tasks. As a consequence, a coherent view of the spatiotemporal structure of the prediction error is not usually well defined.

This article investigates the impact of the covariance localization on an EnKF used to assimilate glider data by means of a field experiment conducted by the Centre for Maritime Research and Experimentation (CMRE) in an area of the Ligurian Sea (western Mediterranean Sea). The trial's specific objectives include an assessment of the predictive capabilities of a multivariate EnKF, which is augmented by covariance localization and assimilated data from a glider, coupled with a regional configuration of the Regional Ocean Model System (ROMS) (Mourre and Chiggiato, 2014; Falchetti et al., 2015). A high-quality data set was repeatedly collected each day over a period (15 days) that is longer than the synoptic time scale in the region (3–4 days, Alvarez and Mourre, 2012) to allow for unique spatiotemporal tracking of the error forecasts.

The article is organized as follows. The methodology used in the EnKF system and the observational data set employed in this study are described in Section 2. The results obtained from the forecast system and the ocean circulation pattern observed in the region are described in Section 3. Finally, Section 4 provides a discussion and concludes the study.

2. Data and methods

2.1. The Mediterranean Rapid Environmental Picture 2013 (MED-REP13)

During 5–20 August 2013, a field experiment known as the Mediterranean Rapid Environmental Picture 2013 (MED-REP13) was conducted by the CMRE in a nearly rectangular area (90 km by 70 km in the along- and cross-shore directions, respectively) of the north-eastern Ligurian Sea (western Mediterranean Sea, Fig. 1). The main scope of this field experiment was to investigate the operational feasibility and benefits of using a heterogeneous ocean observing network to characterize the marine environment.

The seafloor depths range from approximately 50 m to almost 1800 m in this area. Oceanographically, the region is characterized by a current system called the Northern Current (NC). This current flows northward at speeds of 0.3–0.4 m s^{-1} along the continental slope, which is approximately 20–35 km off the shore of the Italian coast. The current extends down to a depth of 300 m. Below the mixed layer, the current is characterized by low salinities (37.8–38.3) and warm (14–16 °C) water masses corresponding to the salinity and temperature signatures of the Modified Atlantic Waters (MAW), which result from the inflow through the Strait of Gibraltar (Millot, 1999; Schroeder et al., 2008). The regional circulation pattern is subjected to significant dynamical variabilities reflected in the meandering nature of the NC and the presence of intense eddy activities (Marullo et al., 1985).

A Slocum glider (Webb et al., 2001) named Jade was operated in this region for the duration of the field experiment. From 5 to 10 August, the glider transited the northern part of the area before it turned towards the southern portion, which it traversed from 15 to 19 August (Fig. 1). The glider trajectory was the result of a compromise between scientific requirements and operational needs. In total, the glider conducted 285 dives to measure the conductivity, temperature and depth between 20 and 200 m using a pumped Seabird 41 CTD sensor operating at 0.5 Hz. The glider was programmed to surface every 3 h to transmit the collected data. Based on the climatological means and standard deviations of the temperature and salinity in the region, the outliers in the data were identified and removed; before calculating the salinity and potential density, the CTD data were low-pass filtered and corrected for cell thermal mass effects. Subsequently, the density