

THE ANALOG DATA ASSIMILATION: APPLICATION TO 20 YEARS OF ALTIMETRIC DATA

Pierre Tandeo¹, Pierre Ailliot², Bertrand Chapron³, Redouane Lguensat¹, Ronan Fablet¹

Abstract—The reconstruction of geophysical dynamics remain key challenges in ocean, atmosphere and climate sciences. Data assimilation methods are the state-of-the-art techniques to reconstruct the space-time dynamics from noisy and partial observations. They typically involve multiple runs of an explicit dynamical model and may have severe operational limitations, including the computational complexity, the lack of model consistency with respect to the observed data as well as modeling uncertainties. Here, we demonstrate how large amount of historical satellite data can open new avenues to address data assimilation issues, and to develop a fully data-driven assimilation. Assuming that a representative catalog of historical state trajectories is available, the key idea is to use the analog method to propose forecasts with no online evaluation of any physical model. The combination of these analog forecasts with observations resorts to classical stochastic filtering methods. For illustration of the proposed analog data assimilation, the brute force use of 20 years of altimetric data is demonstrated to reconstruct mesoscale sea surface dynamics.

I. DATA

The sea surface height and the associated geostrophic surface currents are key ocean variables. Observed from space for 20 years by along track altimeter measurements, these data have led to significant breakthroughs in the understanding of the large-scale (> 200 km) oceanic circulation, leading to unequalled views of the eddy field and its kinetic energy on a global scale, to significantly advance the study of the dynamics of oceanic variability. The wide spacing between the satellite ground tracks is known to limit the cross-track resolution to several hundred km. During this period, up to five altimeters were operating. Classical statistical methods were performed to interpolate these data in a regular spatio-temporal grid (see [1]). Using this historical database of interpolated altimetric observations, we propose to reconstruct the year 2012, where only two altimeters were available. The idea is to use the repeatability of the events in order

to reconstruct the spatiotemporal evolution of mesoscale eddies. As an example, we plot in Figure 1 a sample catalog of analogs (at a given time t) and successors (at time $t + dt$) of interpolated altimetric data in the Agulhas return current.

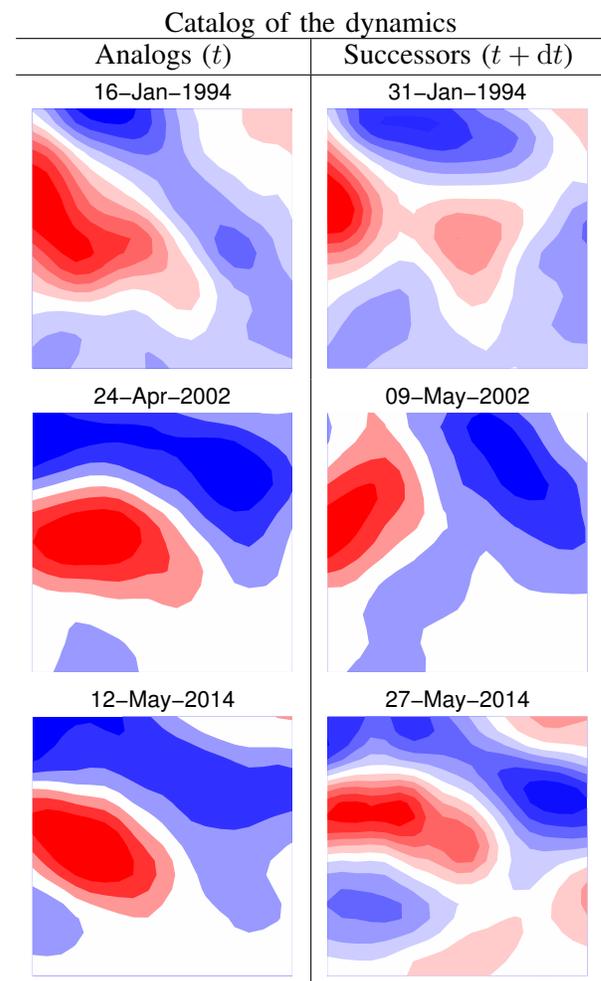


Fig. 1. Examples of interpolated fields of sea level anomalies for different consecutive times ($dt = 15$ days), during the period 1993-2014 on the Agulhas return current.

Quite analog situations are apparently revealed. For instance, the mesoscale eddies, on May 12th 2014, exhibit similar spatial distribution to those observed in the past, April 24th 2002 and/or January 16th 1994.

Corresponding author: pierre.tandeo@telecom-bretagne.eu,
¹Télécom Bretagne, Brest, France ²University of Brest, France
³Ifremer, Brest, France

Accordingly, we can use the following days to infer the spatiotemporal evolution and propose a spread of realistic forecasts. Hereafter, we explore this idea to emulate the dynamical model and we combine it to a classical stochastic filtering technique.

II. METHOD

The analog data assimilation was introduced in [2]. It combines a stochastic ensemble Kalman filter/smoothen and machine learning techniques based on the k-nearest neighbors. This is also known as the analog method in meteorology (see [3]). Assuming that a representative catalog of historical state trajectories is available (such as the one presented in Figure 1), at each time step and for all the ensemble members of the Kalman filter, the closest analogs are found. Their corresponding successors are then used to propose realistic forecasts. Then, the analysis step is classical: we compare the available noisy observations to the analog forecasts.

This analog data assimilation is nonparametric, flexible and fast. Indeed, the dynamical model is directly emulated instead of evaluating differential equations. As shown in [2], even in the case of highly nonlinear dynamics, a sufficiently important catalog can lead to the same performance as classical data assimilation techniques using the pure dynamical model.

III. RESULTS

To construct the catalog, we use the 20 years of interpolated altimetric data except year 2012. For the observations, we use the raw along track data in 2012. We then compare the classical interpolation and the analog data assimilation approach in Figure 2 for two dates.

The results indicate an overall good agreement between the two interpolation methods. The 20 years of historical data is thus large enough to encompass most of the prominent dynamics of the system. This is not the case with 10 years (results not shown here). This issue may be related to properly encompass the inter-annual variability within the catalog. The root mean square error with the along track observations is reduced from 0.05 to 0.04 meter compared to the classical interpolation. The analog data assimilation further indicates a better consistency between the derived streamlines and the trajectories of drifting buoys, good Lagrangian tracers of the upper ocean surface dynamics. These encouraging improvements may be explained by the ability of the proposed method to catch nonlinear dynamical aspects induced by moving structures which can not be reproduced using classical static interpolation with large spatial-temporal correlation length.

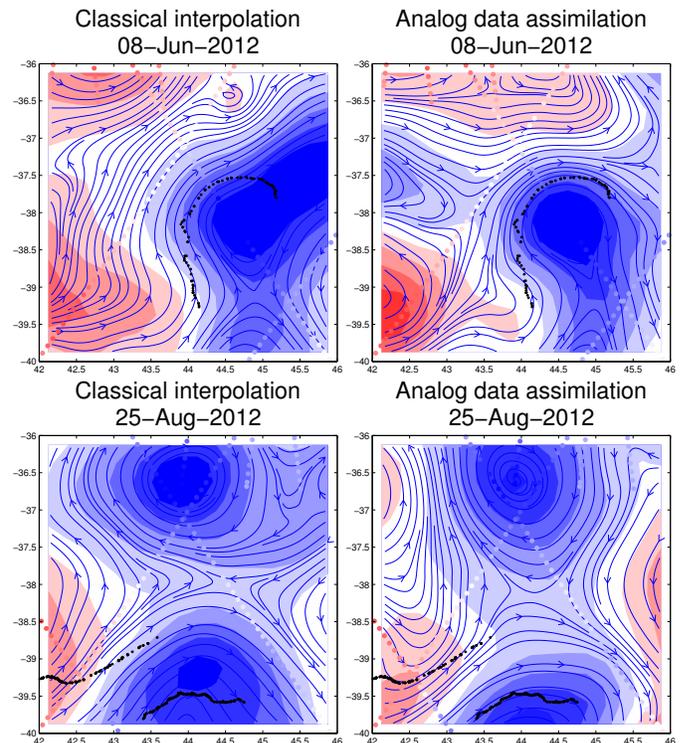


Fig. 2. Interpolation results of along track sea level anomalies using classical spatiotemporal methods and the analog data assimilation. The streamlines represent the corresponding geostrophic currents. Along track observations and drifting buoys trajectories correspond to ± 2 days of data centered on June 8th 2012 and August 25th 2012 in the Agulhas return current.

ACKNOWLEDGMENTS

This work was supported by both EMOCEAN project funded by the "Agence Nationale de la Recherche" and the "inter-Labex SEACS" project. The altimeter products used in this paper were produced by Ssalto/Duacs and distributed by Aviso, with support from Cnes (<http://www.aviso.altimetry.fr/duacs/>). The drifting buoys data were collected and made freely available by the Coriolis project and programs that contribute to it (<http://www.coriolis.eu.org>).

REFERENCES

- [1] P.-Y. Le Traon, D. Antoine, A. Bentamy, H. Bonekamp, L. Breivik, B. Chapron, G. Corlett, G. Dibarboure, P. DiGiacomo, C. Donlon, *et al.*, "Use of satellite observations for operational oceanography: recent achievements and future prospects," *Journal of Operational Oceanography*, vol. 8, no. sup1, pp. s12–s27, 2015.
- [2] P. Tandeo, P. Ailliot, J. Ruiz, A. Hannart, B. Chapron, A. Cuzol, V. Monbet, R. Easton, and R. Fablet, "Combining analog method and ensemble data assimilation: Application to the Lorenz-63 chaotic system," in *Machine Learning and Data Mining Approaches to Climate Science*, Springer, 2015.
- [3] E. N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogues," *Journal of the Atmospheric Sciences*, vol. 26, no. 4, pp. 636–646, 1969.