A NEW LAGRANGIAN BASED SHORT TERM PREDICTION METHODOLOGY FOR HF RADAR CURRENTS

Lohitzune Solabarrieta[1,2], Ismael Hernandez-Carrasco[3], Anna Rubio[2], Alejandro Orfila[3], Michael Campbell[1], Ganix Esnaola[4,5], Julien Mader[2], Burton H. Jones[1].

(1) KAUST, Red Sea Research Center, Integrated Ocean Processes, Saudi Arabia.
(2) AZTI Marine Research, Basque Research and Technology Alliance (BRTA), Pasaia, Spain
(3) Instituto Mediterráneo de Estudios Avanzados. IMEDEA (CSIC-UIB), 07190 Esporles, Spain.
(4) Nuclear Engineering and Fluid Mechanics Dept., Gipuzkoako Ingeniaritza Eskola, Europa Plaza 1, 20018-Donostia, Spain.
(5) Joint Research Unit BEGIK, Instituto Español de Oceanografía (IEO)-Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Plentziako Itsas Estazioa, Areatza Pasealekua, 48620-Plentzia, Spain

Corresponding author's email: lsolabarrieta@azti.es

ABSTRACT

The use of High Frequency Radar (HFR) data is increasing worldwide for different applications in the field of operational oceanography and data assimilation, as it provides real-time coastal surface currents at high temporal and spatial resolution. In this work, a Lagrangian based empirical real-time, Short-Term Prediction (L-STP) system is presented in order to provide short term forecasts of up to 48 hours of ocean currents from HFR data. The method is based on the finding of historical analogues of Lagrangian trajectories obtained from HFR surface currents. Then, assuming that the present state will follow the same temporal evolution as the historical analogue did, we can obtain a short-term prediction of the surface currents. The method is applied to two HFR systems covering two areas with different dynamical characteristics: the southeast Bay of Biscay and the central Red Sea. The L-STP improves on previous prediction systems implemented for the SE Bay of Biscay and provides good results for the Red Sea study area. A comparison of the L-STP methodology with predictions based on persistence and reference fields has been performed in order to quantify the error introduced by this approach. Furthermore, a temporal sensitivity analysis has been addressed to determine the limit of applicability of the methodology regarding the temporal horizon of Lagrangian prediction. A real-time skill-score has been developed using the results of this analysis which allows to identify periods when the short-term prediction performance is more likely to be low and persistence can be used as a better predictor for the future currents.

1. *INTRODUCTION*

The coastal zone is under increasing human pressure. On the one hand, during recent decades coastal seas have been experiencing intensified activity for recreation, transport, fisheries and marine-related energy production. Simultaneously, continued growth of the global coastal population largely contributes to increase the problem of the wastewater discharge which, in many cases, results in serious damage to coastal marine ecosystems. A better understanding of the dynamical processes responsible for the surface oceanic transport, is a prerequisite for the efficient management of the coastal ocean. These processes are responsible of the transport and fate of pollutants, nutrients, jellyfish, harmful algal blooms, plastics, etc, and improving the capacity of monitoring and forecasting the coastal area, is necessary to identify regions of accumulation or dispersion of these harmful materials. This requirement is driving the set-up of a growing number of multi-platform operational observatories designed for the continuous monitoring of the coastal ocean (e.g., US IOOS, EU EOOS, SOCIB, Australian IMOS, etc.). In the need of providing a long-term framework for the development and improvement of the European Marine coastal observations, the JERICO Research infrastructure has been putting efforts (through JERICO, JERICO-NEXT and JERICO-S3 projects) to develop methods and tools for the production of high-quality marine data, and the sharing of expertise and infrastructures between the exiting observatories in Europe. Moreover, due to the need of forecasting applications for response to emergency situations such oil spills, or search and rescue operations, many of the existing operational observatories are linked with operational ocean forecasting models with or without data assimilation (e.g. MARACOOS, NOAA Global Real-Time Ocean Forecast System, COPERNICUS Marine Environment Monitoring System). Typically, constituted with different in-situ point-wise observational platforms (such as moored buoys, tidal gauges, wave buoys, etc.) a significant number of these observatories now employ land-based High Frequency Radars (HFR), that provide real-time coastal currents with unprecedented coverage and resolution (e.g. Paduan and Rosenfeld, 1996; Kohut and Glenn, 2003; Abascal et al., 2009; Solabarrieta et al., 2014, Rubio et al. 2017; Paduan and Washburn, 2013). Each HFR coastal site measures radial surface currents moving away or approaching its antenna, based in the shift of the first peak (Bragg peak) of the Doppler spectra (Crombie 1955, Barrick et al 1977). Combining the overlapping radial vectors from at least 2 antennas provides surface true vector currents (Barrick et al., 1977, Lipa and Barrick, 1983). Several studies have compared in-situ current measurements with HFR observations (e.g., Schott et al. 1985; Hammond et al. 1987; Paduan and Rosenfeld 1996, Emery et al. 2004; Paduan et al., 2006; Ohlmann et al. 2007; Liu et al., 2014; Solabarrieta et al, 2014, Bellomo et al., 2015; Lana et al., 2016; Hernandez-Carrasco et al., 2018b) and have repeatedly demonstrated the validity of this technology. Presently, more than 250 HFR antennas are installed being active worldwide (Roarty et al., 2019; http://global-hfradar.org/).

The range and the spatial resolution of the HFR current systems depend on their working frequency and the conductivity of the water over which the system is measuring. Ranges vary from 15 to 220 km range and spatial resolution from 250 m to 12 km. Typically, a 12 MHz radar has a range ~70 km with a spatial resolution of 2-5 km. HFR systems usually average current measurements for one hour, although some average currents for shorter periods, such as 30 minutes. Due to their high spatio-temporal resolution, HFR data are commonly used in real time for search and rescue (Ullman et al., 2006) or oil spill prediction/mitigation emergency response (Abascal et al., 2017).

The performance of HFR for measuring near-real time surface currents has resulted in the development of assimilation strategies that incorporate the HFR measured surface currents into ocean coastal models (Breivik and Saetra, 2001, Oke et al 2002, Paduan and Shulman 2004, Stanev et al., 2011, Barth et al., 2011) some of which have been tested for short periods of time (Chao et al., 2009). However, assimilation of HFR data into models is still a computationally expensive and complex issue, not to mention operational applications of such a procedure. Because of these constrains, the availability of real-time high-resolution HFR current fields has led to alternative solutions in order to obtain short term prediction (STP) of surface coastal currents, through the direct use of HFR historical and nowcast observations using different approaches (e.g. Zelenke 2005, Frolov et al. 2011, Barrick et al., 2012, Orfila et al. 2015, Solabarrieta et al. 2016, Vilibić et al, 2016, Ren et al., 2019, see Table 1).

The above-mentioned studies develop and implement different STP approaches (harmonic analysis of the last hours, genetic algorithms, numerical models, …) which often require additional data, or long training periods of data without gaps which can jeopardize the general utility of these methods in real time (Hardware failures due to power issues, communications or environmental conditions often result in spatio-temporal gaps within HFR datasets. Spatial gaps can be filled on a real-time basis but the filling of long temporal gaps is not straightforward). Several gap-filling methodologies have been developed for HFR data sets: Open Modal Analysis, (OMA) (Kaplan and Lekien, 2007), Data Interpolating EOFs (DINEOF) (Hernandez-Carrasco et al., 2018), and Self-Organizing Maps (SOM) (Hernandez-Carrasco et al., 2018). The OMA method has been used for spatial gap filling in this paper mainly because it's well functioning has been demonstrated (Kaplan and Lekien, 2007, Hernández-Carrasco et al., 2018) and it is easily appliable in real time, with available codes that will also be applied for trajectories' generation, later in this paper (HFR_progs MATLAB package: https://github.com/rowg/hfrprogs).

A widely used method in time series prediction, especially in early weather forecasting, is the method of analogues. It is based on the assumption that if the behavior of a system at a given time is similar to some other situation in the historical record, then the evolution in the future of state will be similar to the evolution observed in the same historical record. Simply stated, two analogue fields are two distinct fields that are close enough considering some metric, to be considered as equivalent. The finding of the best (nearest) analogue of a specific time does not require a historically continuous dataset, as long as it contains subsets of observations that extend longer than the testing period. These analogue events occur naturally in the environment and this methodology has been applied and tested in atmospheric forecasts (Lorenz, 1969, Jianping et al,1993, Prince and Goswami 2007, Shao and Li 2013).

Given the motivation described above, and developed partially in the framework of JERICO-NEXT project, we present a Lagrangian-based Short-Term Prediction (L-STP from now on) methodology using existing HFR datasets, to be applied to current real-time observations. The uniqueness of this approach is two-fold: first the historical Eulerian velocity fields are used to construct a catalogue of Lagrangian trajectories and second, using the trajectories obtained from present observations, analogues in the past dataset are searched in order to obtain the best predictive match. The method is based on Lagrangian computations since they have proven to be robust in identifying dynamical flow structures and they are direct measurements of transport of substances at sea.

Then, it is worth highlighting that this is the first time that the analogues technique is applied to the HFR-derived ocean surface currents to obtain short-term forecast. The L-STP is intended to be implemented operationally requiring low computational cost (seconds to few minutes for each forecast, depending on the size of the historical dataset) and it is easy to implement using existing HFR data processing tools.

2. DATA AND METHODS

2.1 Data

HFR data from two distinct oceanographic regions have been used for the evaluation, validation and testing of the developed methodology in this paper (Figure 1): Left: The Bay of Biscay (hereinafter BoB HFR) and Right: The central Red Sea region (hereinafter Red Sea HFR). These two study regions are used to evaluate the skill of the method with different dynamical conditions, and with a sufficient set of observations to provide a database suited to the efficient research of appropriate analogues. The BoB HFR system, located in the southeastern corner of the Bay of Biscay, in the Basque Country, is composed of two CODAR Seasonde sites, working since 2009 which transmit at 4.5MHz frequency covering up to 200km range and providing hourly surface velocity field at 5 km of spatial resolution. The dataset used in this study spans the period from 2012 to 2015. The Red Sea HFR system is located on the central western coast of Saudi Arabia and is also composed of two CODAR Seasonde sites, operational since June 2017, transmitting at 16.12MHz frequency, covering up to 120 km range and providing the hourly surface velocity field at 3 km spatial resolution. The dataset from June 2017 to October 2018 has been used in this study.

The BoB HFR has been chosen as the pilot system for testing the developed methodology because of our previous knowledge regarding the circulation and dynamical processes in the study area (Rubio et al 2013, Solabarrieta et al 2014, Solabarrieta et al., 2015, Rubio et al., 2018, Hernandez-Carrasco et al. 2018). The resulting methodology is then applied to the operational Red Sea HFR dataset, as a study case. Coastal dynamics in the BoB show a clear seasonality where cyclonic and anticyclonic eddies dominate in winter and summer, respectively in responding to local winds and the mean coastal current (Iberian Poleward Current) (Esnaola et al., 2013, Solabarrieta et al., 2014). The circulation in the central Red Sea also demonstrates a clear seasonality (Sofianos and Johns, 2003; Yao et al., 2014a, 2014b; Zarokanellos et al., 2016, 2017) linked to the seasonal winds of the area (Abualnaja et al., 2014; Langodan et al., 2017). The region is dominated by eddy activity, with both cyclonic and anticyclonic eddies dominating the region (Zhan et al., 2014; Zarokanellos et al. 2016). Due to the only recently available dataset (since mid-June 2017 to present) the detailed small-scale surface circulation processes of this area is under characterization at the moment.

The primary difference between the two HFR systems is the operating frequency (5MHz for the BoB system and 16 MHz for the Red Sea system) resulting in a larger spatial coverage for the BoB HFR than for the Red Sea HFR (200km range vs. 120km, respectively), but with higher spatial resolution for the latter (3km and 5 km, respectively). This difference in the spatial resolution should result in better capturing the small-scale dynamical features in the Red Sea that could influence the selection of an analogue.

The data from both systems have been processed similarly. The spectra of the received backscattered signal are converted into radial velocities using the MUltiple SIgnal Classification (MUSIC) algorithm (Schmidt 1986). HFR Progs MATLAB package is then used to combine radial currents and generate gap-filled total 2D currents, using the Open Modal Analysis (OMA) methodology of Kaplan and Lekien (2007).

2.2 Lagrangian analogues

The proposed methodology is based on the analogue finding approach, using a historical catalogue of maps of Lagrangian trajectories and finding the most similar one (detailed later in this section) to that of the last 48 hours (target field). Then, the next 48-hour time evolution of the closest (chosen) analogue provides the forecast for the target period. In other words, if we find a state in the historical database that is close enough to the target field (given a metric), the forecast for the current observations will evolve in the same way as did for the chosen analogue. Analogue finding has been applied in several geophysical variables in different regions (Zorita and von Storch, 1999; Fernandez-Ferrero et al., 2009, 2010; Ibarra-Berastegi et al., 2011; Martin et al., 2014; Seubert et al., 2014; Ibarra-Berastegi et al., 2015).

The analogue finding was first applied to eulerian surface velocity fields of the BoB HFR System (not shown), but the results did not improve the previously published STP results for the study area. The methodology was tested subsequently using a four-year dataset (2012-2015) of trajectory maps computed for the SE BoB, where the trajectory maps from the three first years was used as the search catalogue for analogues (2012-2014) (hereinafter "Lagrangian catalogue"), and the remaining year (2015) was used as a test case (hereinafter "test period"). Then the method was applied to the Red Sea dataset, for the period of July 2017-October 2018. As the period was short (1 year and 4 months), we have used the whole period to build the Lagrangian catalogue and act as a test period at the same time. In this case, for the analogues search the 5-days period around the date of the target field was removed from the catalogue at each iteration, to avoid temporal overlapping with the target field.

To build the Lagrangian catalogue we first generated hourly fields of 25 virtual particle trajectories on a regular grid, blued dots of Figure 2), which were advected by the OMA HFR surface currents (without considering diffusion) during 48 hours. To this end we used the Lagrangian module included in the HFR_Progs MATLAB package, following the same procedure for the test period. Then, for each hour of the test period, the method searched the most similar Lagrangian patterns in the Lagrangian catalogue dataset. To increase the efficiency of the processes, the search was done in two steps. First, we looked for potential analogues with a similar main drift. To do that we computed and compared the position of the centroid of the 25 trajectories of each analogue to that of the target field, and discarded the analogues whose centroid was at a distance > δ_cg. The value of the δ_cg needs to be small enough to minimize computational time but sufficiently large to as to not lose potential analogues. We explored different values of this distance threshold and we found that δ_gc=10km produces a good compromise between computational cost and number of potential analogues in both study areas. Then, in a second step, we computed the Lagrangian errors (Ɛ) between the trajectories of the target field and the potential analogues, defined as:

$$\mathcal{E} = \Sigma \left((\delta_{\_6h})^2 + (\delta_{\_12h})^2 + (\delta_{\_24h})^2 + (\delta_{\_36h})^2 + (\delta_{\_48h})^2\right) \qquad \text{Eq. (1)}$$

Where δ_t is the mean separation distance [km] at time t between the trajectories belonging to the target field and each of the potential analogues (being t=6, 12, 24, 36 and 48 hours inside the trajectories lifetimes).

Finally, the potential analogue with the lowest Ɛ was selected as the best analogue ($\mathcal{E}_{ANL}$= min (Ɛ)) and the velocity fields during the next 48 h from that analogue provides STP currents for the target period (hereinafter “L-STP fields”). *Figure* 3 shows an example of the values of Ɛ, through the potential analogues for a specific case.

Figure 2 provides an example of the selected analogue (Figure 2b) and corresponding L-STP fields (Figure 2d) for a given target field (Figure 2a) and the 'truth' trajectories for the following 48 hours from the date of the target field (Figure 2c). The associated temporal series of errors for the target field and the potential analogues are shown in *Figure* 3, where the value of ε_{ANL} is marked using a red dot. (corresponding to the error between the trajectories of the L-STP field in Figure 2d and the truth trajectories for the forecast period in –Figure 2c).

To assess the performance of the methodology, we computed forecasted trajectories based on persistence of currents (hereinafter 'persistence fields'). To obtain simulated trajectories using persistence currents, the particles were advected during 48 hours using a constant velocity field (target field) during the 48 hours of simulation:

v(x,y,tf+ti)= v(x,y,ts),

where tf= study time and ti=[tf: tf+48h].

The mean drift of the truth forecasted trajectories is also computed for each simulation period (the means drift is considered as the average of the distances moved by each particle during 48 hours).

The Lagrangian errors between the truth trajectories and the L-STP and between the truth trajectories and the persistence field were also computed as follows:

$$\varepsilon_{STP} = \Sigma\,((\delta\_{6h})^2 + (\delta\_{12h})^2 + (\delta\_{24h})^2 + (\delta\_{36h})^2 + (\delta\_{48h})^2) \qquad \text{Eq. (2)}$$

where δ_t is the mean separation distance between truth field's and the L-STP field trajectories for t= t : t+48 (following 48 hours from the study time)

$$\varepsilon_{PRS} = \Sigma\,((\delta\_{6h})^2 + (\delta\_{12h})^2 + (\delta\_{24h})^2 + (\delta\_{36h})^2 + (\delta\_{48h})^2) \qquad \text{Eq. (3)}$$

where δ_t is the mean separation distance between truth field's and Persistent field trajectories for t= t : t+48 (following 48 hours from the study time)

All the process for the selection and validation of the analogue with the different variables has been summarized in *Figure* 4.

The time series and spatial distribution of the ε_{STP} and ε_{PRS} errors have been analyzed for both study areas. Finally, ε_{STP} and ε_{PRS} time series have also been calculated and compared to the time series of the ε_{ANL}, in order to evaluate if the ε_{ANL} can be used as an indicator of the expected skill of the L-STP with respect to the persistence.

3. RESULTS

The performance assessment results for the BoB HFR system are described in section 3.1 and the temporal and spatial forecast for both study areas are shown in section 3.2.

3.1 Assessment of the L-STP skills

Figure 5 shows the ε_{ANL} through year 2015 for the BOB study area, together with the ε_{STP} and ε_{PRS}. The mean value of the ε_{PRS} is 73% higher than the ε_{STP}. Black dots over the timeline in Figure 5 show the times when ε_{STP} is higher than the ε_{PRS}, which occurs 12% of the time. Focusing on the times when the ε_{PRS} is lower than the ε_{STP} (black dots of the timeline in Figure 5), it can be seen that they mostly occur during winter months. Previous works in this area have shown that there are high persistent eastward currents that can least for several weeks during winter months (Solabarrieta et al., 2014), which can explain the better performance of the persistence fields in this period.

The correlation between ε_{ANL} and ε_{STP} is 0.46 while correlation between ε_{ANL} and ε_{PRS} is 0.05, for the whole test year (2015) (Figure 5).

The hourly values of ε_{STP} and ε_{PRS} have been plotted against their corresponding hourly ε_{ANL} values for the test year, ordered from minimum to maximum along the x-axis in Figure 6. We observe that, when ε_{ANL} is low (less than 853 km^2 for this data set), ε_{STP} is smaller than $\varepsilon_{PRS.}$ However, as ε_{ANL} increases, ε_{STP} and ε_{PRS} converge until an inflection point beyond which ε_{STP} is slightly greater than $\varepsilon_{PRS.}$ For the SE BoB experiment, the inflection point occurs at ε_{ANL} =853 km^2 and 88% of cumulative

ε_{ANL}. Results from the Red Sea HFR system indicates a similar pattern (not shown), when the inflection point occurs at ε_{ANL} = 821 km^2 and at 86.4% of cumulative ε_{ANL}.

Further analysis to elucidate the time periods that largely contribute to the errors, compared to persistence are presented hereinafter. ε_{ANL} has been plotted together with the mean separation distances of the trajectories using STP and persistent currents (hereinafter δ_{STP} δ_{PRS}), after 6, 12, 24, 36 and 48 hours for each target field (Figure 7). δ_{STP} is always higher than the δ_{PRS} for the 6 hours' simulation. But the values of δ_{STP} show better results for simulations at 12, 24, 36 and 48 hours. The values of the correlation coefficient (R^2) between the ε_{ANL} and δ_{STP} and between ε_{ANL} and δ_{PRS} after 6, 12, 24, 36 and 48 hours are summarized in Table 2. Values of R^2 for ε_{ANL} and δ_{PRS} are small (almost no correlation), varying between 0.01 and 0.11, while correlations between ε_{ANL} and δ_{STP} are higher, varying between 0.19 and 0.56, and showing higher correlation (>than 0.39) after 12 hours of simulations. The behavior of the Red Sea HFR system figures (not shown) is similar to the BoB HFR system.

3.2 L-STP performances in the selected study areas

Mean separation distances between truth and forecasted trajectories after different periods of integration times have been computed for both systems, for the best analogues, i.e., before the inflection point of $\varepsilon_{STP>}\varepsilon_{PRS}$ (Figure 6), in order to evaluate the temporal forecast capabilities of the methodology. Only analogues with $\varepsilon_{ANL} <$ 853km^2 (BoB system) have been used to generate this analysis, as those are the periods when the methodology produces good results. Separation distances computed for the whole test year 2015, are shown in Figure 8, for the BoB HFR observations.

The separation distances between the measured trajectories and predicted persistent and STP trajectories, have similar values during the first 6 hours (4km) of the forecast period, with slightly better results for persistent trajectories. But after 6 hours, the separation distance for the forecast based on persistent currents increases faster than using L-STP. At 24 hours, the separation distance is 11 km for persistence forecasts and 8km for L-STP forecasts. The values are 12 and 18km, respectively, after 48 hours of simulation. The mean drift values of the truth trajectories show that the mean drift is similar to the L-STP separation distances, during the 48 hours.

Temporal mean separation distances between truth and forecasted trajectories for the Central Red Sea HFR System, computed for the whole test time are shown in Figure 9. Only the best analogues with ε_{ANL} less than inflection point, i.e., $\varepsilon_{ANL} < 821\text{km}^2$, have been used to generate this analysis. The separation distances for the STP forecasts are higher than those forecasts with persistent currents during the first 15 hours. After 15 hours, quality of forecasts reversed where STP produced better results than persistence.

Spatial distribution of the difference between $\delta_{\_PRS}$ and $\delta_{\_STP}$ at 6, 12, 24 and 48 hours, for the BoB and the Red Sea study areas, are shown in Figure 10 and Figure 11.

For the BoB HFR system, the differences are not appreciated during the first 6 hours. But after 12 hours of simulation, the advantage of the L-STP is clear in most of the study area, especially outside the continental shelf slope where persistent currents dominate the circulation. The separation values between $\delta_{\_PRS}$ and $\delta_{\_STP}$ increase up to 10km after 48hours of simulation.

For the Red Sea, the significant differences between STP and Persistence start after 24 hours of simulation, and continue until 48 hours.

4. *DISCUSSION*

In this work, a new methodology to forecast HFR currents has been described and the skill of the proposed STP methodology is analyzed. Different analyses are performed in order to check the spatial and temporal capabilities of the proposed methodology.

The methodology is based on the search of analogues in a trajectory (Lagrangian) space using a previously generated trajectory field catalogue. The values of the $\delta_{\_STP}$, compared to previous works in the BoB area showed that the L-STP produces accurate predictions, which demonstrates the ability of the Lagrangian approach to capture key dynamical features needed to accurately predict the proper dynamical conditions.

Significant correlation values between ε_{ANL} and $\delta_{\_STP}$, suggest that the ε_{ANL} can be considered as a real-time skill-score metric for the L-STP. Both BoB and the central Red Sea show a similar behavior; although the ε_{ANL} values are different, the accumulative % of the transition point is similar in both cases.

Figure 7 shows that after 12 hours of simulation, the L-STP provides a better prediction than the persistence field for more than 80% of the cases (reaching more than 90% of the cases for 36 and 48 hours of simulation). The minimum ε_{ANL} value for the $\delta_{\_STP}$ and $\delta_{\_PRS}$ cross point is 714km^2. Figure 6, for the total ε_{ANL} shows the same behavior being 853km^2 the transition analogue error value between STP and Persistence.

For the BoB HFR System, temporal δ_{STP} shows values of 3.5km, 5.5km and 8km, after 6, 12, and 24 hours respectively. The δ_{STP} values are similar to the δ_{PRS} values during the first 6 hours of simulation but δ_{STP} are lower after that, with 3km and 5.5km of difference between them, after 24 and 48 hours of simulation, respectively (*Figure* 8). As stated in previous work, that the circulation over the BoB area is dominated by a stable, persistent current field during winter (Solabarrieta et al., 2014) which is reflected by these results where persistence has good or even slightly better forecasting skill during the first 6 forecast hours than the proposed methodology.

The δ_{STP} values for the BoB HFR system are similar to the ones obtained by Solabarrieta et al., 2016, for the whole year but δ_{STP} are better for summer months, for the same study area. They used the linear autoregressive model, described in Frolov et al., 2012, to forecast HFR current fields and the errors using that approach were 2.9 and 7.9km after 6 and 24 hours. Although the results obtained in this work improve only during certain periods the forecast presented in Solabarrieta et al. 2016, the presented methodology has three advantages over the previous method: it is easily run in real time; it does not require a continuous training period; and it is able to discriminate the times when the usage of the persistence is applicable. On the negative side, it requires the generation of a catalogue of past trajectories as the search space for analogues, but once it is ready, it is easily increasable in real time, without extra pre-analysis; just adding new trajectory fields to the previous catalogue.

The values of the δ_{STP} for the Red Sea HFR system follow a similar pattern to the BoB results, with higher separation distances. This may be related to the limited time span of the available dataset, as a better closest analogue may be found in a longer dataset.

The spatial comparison of the δ_{STP} and δ_{PRS} for the BoB HFR system (Figure 10), shows that the L-STP has better skills for the entire study area after 12 hours of simulations. The skills of the L-STP with respect to the persistence increases with time, showing up to10km of improvement relative to persistence at 48 hours in some parts of the study area. For the spatial distribution, after 12 hours, the smallest differences between δ_{STP} and δ_{PRS} occurred over the slope. This is explained by existence of persistent seasonal Iberian Poleward Current that flows along the continental slope toward the east along the Spanish coast and northward along the French coast (Solabarrieta et al. 2014). In other words: although the L-STP can be performant in periods of persistent currents , the persistence field can show a better forecast for a short temporal scale (48h). L-STP will improve those forecasts, as soon as spatiotemporal variability increases.

The results for the Red Sea HFR system are similar but the benefit of the L-STP methodology appears only after 12 hours of simulation. Spatially, the improvement is again lower where persistent currents occur, as it is the case of the Eastern Boundary Current that flows northward following the eastern Red Sea Coastline in the study area (Bower and Farrah, 2015; Sofianos and Johns, 2003; Zarokanellos et al., 2017). The dominance of the persistent currents is evident in the lower values of the difference between the STP forecasts and the Persistence forecasts as shown in Figure 11 and in comparison with Figure 10.

We have compared the capabilities of the L-STP forecast against the forecast based on the persistency of currents. The L-STP method requires long training periods but performs better during non-persistent periods. Previous efforts to forecast surface currents from HFR data have shown similar results compared with the methodology presented in this paper. However, the advantage of the L-STP method is that it can be used in near real time, with short and non-continuous datasets of around 2-3 years, provided that a Lagrangian catalog representative for the study area can be built.

The HFR Progs MATLAB package (https://cencalarchive.org/~cocmpmb/COCMPwiki) has been used to generate total currents from radial files to fill the spatial gaps of the surface current field using the OMA method, and to generate Lagrangian trajectories. This methodology could be easily included in this package so the final users could get forecast currents, in the same time that they generate total currents.

5. *CONCLUSION*

A methodology to short term forecast of the surface currents in real-time has been proposed. This methodology provides accurate forecast of sea surface currents and its capability has been tested in terms of spatial and temporal distributions. The good functioning and confidence of this methodology has been demonstrated in the previous sections and also, its capability to be applied in real time. The methodology has been successfully applied to two distinct coastal regions to evaluate its capabilities in different hydrodynamic regimes, although further analysis using data from more areas is required to generalize the methodology.

Relationships between ε_{ANL} and ε_{STP}/ ε_{PRS} suggest that the ε_{ANL} can be considered as a reliable indicator of the method's performance. Taking in consideration all the analyses done in this work, we propose to use STP currents for trajectory or velocity field predictions from 12 hours forward, if the ε_{ANL} value is lower than 80% of the cumulative ε_{ANL} . If ε_{ANL} is higher, or the forecast is just for the next 6 hours, the use of the persistence field is suggested. We also suggest that the ε_{ANL} value and forecast transition time need to be carefully evaluated for each study region. This, of course, infers that a minimum data set is required before the L-STP method can be applied.

Further analysis of analogue finding approaches is required to improve the observed results, especially during periods when currents are persistent. The use of longer dataset as a training period may improve this aspect. Then, the next step would be to test the methodology for additional periods and other regions, to analyze the possibility to find analogues for different sub-regions and to evaluate its functionality in an operational mode.

The methods to find the minimum training period for each system should be analyzed deeper in future works. The minimum training period will be directly related to the variability of the local dynamics and those should be considered during the analysis.

DATA AVAILABILITY

The Red Sea HF Radar data can be requested through:

- https://lthdatalib.kaust.edu.sa

Historical and NRT Bay of Biscay HF Radar data can be requested through:

- Euskoos portal: https://www.euskoos.eus/en/data/basque-ocean-meteorological-network/high-frequency-coastal-radars/
- Emodnet Physics - http://www.emodnetphysics.eu/Map/platinfo/piradar.aspx?platformid=10273
- CMEMS Instac - http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=INSITU_GLO_UV_NRT_OBSERVATIONS_013_048

AUTHORS CONTRIBUTION

- Lohitzune Solabarrieta: She has worked on the set up of the methodology, data analysis, manuscript writing and final submission.
- Ismael Hernandez-Carrasco: He has worked on the set up of the methodology and the manuscript writing.
- Anna Rubio: She has worked on the set up of the methodology, data analysis, and manuscript writing.
- Alejandro Orfila: He has worked on the configuration of the methodology, data analysis and the manuscript writing.
- Michael Campbell: He has worked on the configuration of the methodology, especially in the pre-configuration that led us to rule out other data prediction methodologies. He has also contributed on the manuscript writing.
- Ganix Esnaola: He has worked on the configuration of the methodology, especially in the pre-configuration that moved us to the usage of analogues on this paper. He has also contributed on the manuscript writing.
- Julien Mader: He has contributed on the writing of the manuscript.
- Burton H. Jones: He has contributed on the writing of the manuscript

COMPETING INTERESTS

The authors declare that we have no conflict of interest

ACKNOWLEDGEMENTS

This work was funded by a Saudi Aramco-KAUST Center for Marine Environmental Observation (SAKMEO) Postdoc fellowship to Lohitzune Solabarrieta, and from the Integrated Ocean Processes (IOP) Group in KAUST. We acknowledge the support of the LIFE-LEMA project (LIFE15 ENV/ES/000252), the European Union's Horizon 2020 research and innovation program under grant agreement No. 654410 & 871153 (JERICO-NEXT and JERICO-S3 Projects), the Directorate of Emergency Attention and Meteorology of the Basque Government, the MINECO/FEDER Projects MUSA and MOCCA (CTM2015-66225-C2-2-P; 256RTI2018-093941-B-C31). and the Department of Environment, Regional Planning, Agriculture and Fisheries of the Basque Government (Marco Program). This work was partially performed while A. Orfila was a visiting scientist at the Earth, Environmental and Planetary Sciences Department at Brown University through a Ministerio de Ciencia, Innovación y Universidades fellowship (PRX18/00218). Ismael Hernandez-Carrasco acknowledges the Vicenç Mut contract funded by the Balearic Island Govern and the European Social Fund (ESF) Operational Programme.

The HF radar-processing toolbox HFR_Progs use to produce OMA was provided by D. Kaplan and M. Cook, Naval Postgraduate School, Monterey, CA, USA.

REFERENCES

Abascal, A. J., Castanedo, S., Medina, R., Losada, I. J., Álvarez-Fanjul, E.: Application of HF radar currents to oil spill modelling. Mar. Pollut. Bull. 58 (2), 238–248, 2009

Abascal A. J., Sanchez, J., Chiri, H., Ferrer, M. I., Cárdenas, M., Gallego, A., Castanedo, S., Medina, R., Alonso-Martirena, A., Berx, B., Turrell, W. R., Hughes, S. L.: Operational oil spill trajectory modelling using HF radar currents: A northwest European continental shelf case study. Marine Pollution Bulletin, Volume 119, Issue 1, Pages 336-350, ISSN 0025-326X, https://doi.org/10.1016/j.marpolbul.2017.04.010, 2017.

Abualnaja, Y., Papadopoulos, V. P., Josey, S. A., Hoteit, I., Kontoyiannis, H., and Raitsos, D. E.: Impacts of climate modes on air–sea heat exchange in the Red Sea, J. Clim., 28, 2665–2681, doi:10.1175/JCLI-D-14-00379.1, 2015.

Barrick, D. E.: Extraction of wave parameters from measured HF radar sea-echo Doppler spectra. Radio Sci., 12, 415–424, doi:10.1029/RS012i003p00415, 1977.

Barrick D.E., Fernandez, V., Ferrer, M.I., Whelan, C., and Breivik, Ø.: “A short-term predictive system for surface currents from a rapidly deployed coastal HF-Radar network,” Ocean Dyn., vol. 62, no. 5, pp. 725–740, 2012.

Barth, A., Alvera-Azcárate, A., Beckers, JM., Staneva J., Stanev E.V., and Schulz-Stellenfleth J.: Correcting surface winds by assimilating high-frequency radar surface currents in the German Bight. Ocean Dynamics, 2011, vol 61: 599. https://doi.org/10.1007/s10236-010-0369-0, 2011.

Bellomo, L., Griffa, A., Cosoli, S., Falco, P., Gerin, R., Iermano, I., Kalampokis, A., Kokkini, Z., Lana, A., Magaldi, M.G., Mamoutos, I., Mantovani, C., Marmain, J., Potiris, E., Sayol, J.M., Barbin, Y., Berta, M., Borghini, M., Bussani, A., Corgnati, L., Dagneaux, Q., Gaggelli, J., Guterman, P., Mallarino, D., Mazzoldi, A., Molcard, A., Orfila, A., Poulain, P. M., Quentin, C., Tintoré, J., Uttieri, M., Vetrano, A, Zambianchi, E. and Zervakis, V.: Toward an integrated HF radar network in the Mediterranean Sea to improve search and rescue and oil spill response: the TOSCA project experience. Toward an integrated HF radar network in the Mediterranean Sea to improve search and rescue and oil spill response: the TOSCA project experience, Journal of Operational Oceanography, 8:2, 95-107, DOI: 10.1080/1755876X.2015.1087184, 2015.

Bower, A. S., and Farrar, J. T.: Air–sea interaction and horizontal circulation in the Red Sea. In N. M. A. Rasul & I. C. F. Stewart, (Eds.), The Red Sea, Springer Earth System Sciences (pp. 329–342). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-662-45201-1_19, 2015.

Breivik, Ø, and Saetra, Ø.: Real time assimilation of HF radar currents into a coastal ocean model. Journal of Marine Systems, Volume 28, Issues 3–4, April 2001, Pages 161-182. https://doi.org/10.1016/S0924-7963(01)00002-1, 2001.

Chao, Y., Li Z., Farrara, K., McWilliams, J.C., Bellingham, J., Capet, X., Chavez, F., Choi, J., Davis, R., Doyle, J., Fratantoni, D. M., Li P., Marchesiello, P., Moline, M.A., Paduan, J., Ramp, S.: Development, implementation and evaluation of a data-assimilative ocean forecasting system off the central California coast. Deep Sea Research, Vol. 56, Issues 3-5, pp 100-126. https://doi.org/10.1016/j.dsr2.2008.08.011, 2009.

Crombie, D. D.: Dopler Spectrum of Sea Echo at 13.56-Mc/s', Nature 175, 681-682, 1955.

Emery, B. M., Washburn L., and Harlan, J. A.: Evaluating radial current measurements from CODAR high-frequency radars with moored current meters. J. Atmos. Oceanic Tech- nol., 21, 1259–1271, doi:10.1175/1520-0426(2004)021,1259: ERCMFC.2.0.CO;2, 2004.

Esnaola, G., Sáenz, J., Zorita, E., Fontán, A., Valencia, V., and Lazure, P.: Daily scale wintertime sea surface temperature and IPC-Navidad variability in the southern Bay of Biscay from 1981 to 2010, Ocean Sci., 9, 655–679, https://doi.org/10.5194/os-9-655-2013, 2013.

Fernández-Ferrero, A., Sáenz, J., Ibarra-Berastegi, G., Fernández, J.: Evaluation of statistical downscaling in short range precipitation forecast. Atmos. Res. 94, 448–461, 2009.

Fernández-Ferrero, A., Sáenz, J., Ibarra-Berastegi, G.: Comparison of the performance of different Analog-Based Bayesian probabilistic precipitation forecasts over Bilbao, Spain. Mon. Weather Rev. 38, 3107–3119, 2010.

Frolov, S., Paduan J., Cook M., and Bellingham J.: Improved statistical prediction of surface currents based on historic HF- radar observations. Ocean Dyn., 62, 1111–1122, doi:10.1007/ s10236-012-0553-5, 2012.

Hammond, T.M., Pattiaratchi ,C.B., Osborne, M.J., Nash, L.A., Collins, M.B.: Ocean surface current radar (OSCR) vector measurements on the inner continental shelf. Continental Shelf Research. Volume 7, Issue 4, Pages 411-431. https://doi.org/10.1016/0278-4343(87)90108-7, 1987.

Hernández-Carrasco, I., Solabarrieta, L., Rubio, A., Esnaola, G., Reyes, E., and Orfila, A.: Impact of HF radar current gap-filling methodologies on the Lagrangian assessment of coastal dynamics, Ocean Sci., 14, 827-847, https://doi.org/10.5194/os-14-827-2018, 2018.

Hernández-Carrasco, I., Orfila, A., Rossi, V., and Garçon, V.: Effect of small-scale transport processes on phytoplankton distribution in coastal seas, Scientific Reports, 8:8613, https://doi.org/10.1038/s41598-018-26857-9, 2018b.

Ibarra-Berastegi, G., Saenz, J., Ezcurra, A., Ezcurra, A., Elias, A., Diaz Argandona, J., Errasti, I.: Downscaling of surface moisture flux and precipitation in the Ebro Valley (Spain) using analogues and analogues followed by random forests and multiple linear regression. Hydrol. Earth Syst. Sci. 15 (6), 1895–1907, 2011.

Jianping H., Yuhong Y., Shaowy W and Jifen C.: An analogue-dynamical long-range numerical weather prediction system incorporating historical evolution. Q. J. R. Meteorol. Soc., 119, pp.547-565, 1993.

Kaplan, D. M. and Lekien, F.: Spatial interpolation and filtering of surface current data based on open-boundary modal analysis, Journal of Geophysical Research: Oceans, 112, https://doi.org/10.1029/2006JC003984, c12007, 2007.

Kohut, J.T., Glenn, S.M.: Improving HF radar surface current measurements with measured antenna beam patterns. J. Atmos. Oceanic Technol. 20 (9), 1303–1316, 2003.

Lana, A., Marmain, J., Fernández, V., Tintoré, J., Orfila A.: Wind influence on surface current variability in the Ibiza Channel from HF Radar. Ocean Dynamics, 66: 483. https://doi.org/10.1007/s10236-016-0929-z, 2016

Langodan S., Cavaleri L., Vishwanadhapalli Y., Pomaro A., Bertotti L., Hoteit I.:. Climatology of the Red Sea - Part 1: the wind. Int. J. Climatol. 37: 4509–4517. DOI: 10.1002/joc.5103, 2017.

Liu Y., Weisberg, R.H., and Merz, C.R.: Assessment of CODAR SeaSonde and WERA HF Radars in Mapping Surface Currents on the West Florida Shelf. Journal of atmospheric and oceanic technology, Vol 31., pp 1363:1382, 2014.

Lorenz, E. N.: Atmospheric Predictability as Revealed by Naturally Ocurring Analogues. Journal of Atmospheric Sciences, Volume 29, pp 636-646, 1969.

Martin, M.L., Valero, F., Pascual, A., Sanz, J., and Frias, L.: Analysis of wind power productions by means of an analog model. Atmos. Res. 143, 238–249, 2014.

Oke, P. R., Allen, J. S., Miller, R. N., Egbert, G. D., and Kosro P. M.: Assimilation of surface velocity data into a primitive equation coastal ocean model, J. Geophys. Res., 107, 3122, doi:10.1029/2000JC000511, 2002.

Orfila A., Molcard, A., Sayol, J.M., Marmain, J., Bellomo, L., Quentin, C., and Barbin, Y.: Empirical Forecasting of HF-Radar Velocity Using Genetic Algorithms IEEE Transactions on Geoscience and Remote Sensing, Vol. 53, No. 5, 2015.

Ohlmann, C., White, P., Washburn, L., Emery, B., Terrill, E., Otero, M.: Interpretation of coastal HF radar–derived surface currents with high-resolution drifter data. J. Atmos. Oceanic Technol. 24 (4), 666–680, 2007.

Paduan, J.D., and Rosenfeld, L.K.: Remotely sensed surface currents in Monterey Bay from shore-based HF radar (coastal ocean dynamics application radar. J. Geophys. Res. 101 (C9), 20669–20686, 1996.

Paduan, J.D., and Shulman, I.: HF radar data assimilation in the Monterey Bay area. J. Geophys Res. 109:C07S09, 2004.

Paduan, J.D., Kim, K.C., Cook, M. S., and Chavez, F.P.: Calibration and Validation of Direction-Finding High-Frequency Radar Ocean Surface Current Observations. IEEE Journal of oceanic engineering, Vol. 31, No. 4, 2006.

Paduan, J.D., and Washburn, L.: High-Frequency Radar Observations of Ocean Surface Currents. Annual Rev. Marine. Sci. 2013.5:115-136, 2013.

Prince, K., X. and Goswami, B., N.: An Analog Method for Real-Time Forecasting of Summer Monsoon Subseasonal Variability. Monthly weather review, Vol 135, pp: 4149-4160. https://doi.org/10.1175/2007MWR1854.1, 2007.

Ren L., Miaro, J., Li Y., Luo, X., Li J. and Hartnett, M.: Estimation of Coastal Currents Using a Soft Computing Method: A Case Study in Galway Bay, Ireland. Mar. Sci. Eng., 7(5), 157; https://doi.org/10.3390/jmse7050157, 2019

Roarty, H., Cook, T., Hazard, L., George, D., Harlan, J., Cosoli, S., Wyatt, L., Alvarez Fanjul, E., Terrill, E., Otero, M., Largier, J., Glenn, S., Ebuchi, N., Whitehouse, B., Bartlett, K., Mader, J., Rubio, A., Corgnati, L., Mantovani, C., Griffa, A., Reyes, E., Lorente, P., Flores-Vidal, X., Saavedra-Matta, K.J., Rogowski, P., Prukpitikul, S., Lee, S.H., Lai, J.W., Guerin, C.A., Sanchez, J., Hansen, B. and Grilli, S.: The Global High Frequency Radar Network. Front. Mar. Sci. 6:164. doi: 10.3389/fmars.2019.00164, 2019 (in press).

Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag., 34, 276–280, doi:10.1109/TAP.1986.1143830, 1986.

Schott F., Frisch, A.S., Leaman, K., Samuels, G., Popa Fotino, I.: High-Frequency Doppler Radar Measurementsof the Florida Current in Summer 1983. J. Geo. Research, Vol 90, No C5, pp 9006:9016, 1985.

Seubert, S., Fernández-Montes, S., Philipp, A., Hertig, E., Jacobeit, J., Vogt, G., Paxian, A., Paeth, H.: Mediterranean climate extremes in synoptic downscaling assessments. Theor. Appl. Climatol. 117 (1–2), 257–275, 2014.

Shao, Q. and Li, M.: An improved statistical analogue downscaling procedure for seasonal precipitation forecast . Stoch Environ Res Risk Assess 27, pp.: 819-830. https://doi.org/10.1007/s00477-012-0610-0, 2013.

Sofianos, S. S., and Johns, W. E.: An oceanic general circulation model (OGCM) investigation of the Red Sea circulation: 2. Three- dimensional circulation in the Red Sea. Journal of Geophysical Research, 108(C3), 3066. https://doi.org/10.1029/2001jc001185, 2003.

Solabarrieta, L., Rubio, A., Castanedo, S., Medina, R., Charria, G., Hernández, C.: Surface water circulation patterns in the southeastern Bay of Biscay: new evidences from HF radar data. Cont Shelf Res 74:60–76 doi:10.1016/j.csr.2013.11.022, 2014.

Solabarrieta, L., Rubio, A., Cárdenas, M., Castanedo, S., Esnaola, G., Méndez, F.J., Medina, R., and Ferrer, L.: Probabilistic relationships be- tween wind and surface water circulation patterns in the SE Bay of Biscay. Ocean Dyn., 65, 1289–1303, doi:10.1007/s10236-015-0871-5, 2015.

Solabarrieta, L., Frolov, S., Cook, M., Paduan, J., Rubio, A., González, M., Mader, J., and Charria, G.: Skill Assessment of HF Radar–Derived Products for Lagrangian Simulations in the Bay of Biscay. J. Atmos. Oceanic Technol., 33, 2585–2597, doi: 10.1175/JTECH-D-16-0045.1, 2016.

Stanev, E.V., Schulz-Stellenfleth, J., Staneva, J., Grayek, S., Seemann, J. and Petersen, W.: Coastal observing and forecasting system for the German Bight – estimas of hydrophysical states. Ocean Sci., 7, 569–583, 2011 doi:10.5194/os-7-569-2011, 2011.

Ullman, D.S., O'Donnell, J., Kohut, J., Fake, T., Allen, A.: Trajectory prediction using HF radar surface currents: Monte Carlo simulations of prediction uncertainties. J. Geophys. Res. 111 (C12005), 1–14, 2006.

Vilibić, I., Šepić, J., Mihanović, H., Kalinić, H., Cosoli, S., Janeković, I., Žagar, N., Jesenko, B., Tudor, M., Dadić, V. dnd Ivanković, D.: Self-organizing maps-based ocean currents forecasting system. Scientific Reports 6, 22924, 2016.

Yao, F., Hoteit, I., Pratt, L. J., Bower, A. S., Zhai, P., Kohl, A., and Gopalakrishnan, G.: Seasonal overturning circulation in the Red Sea: 1. Model validation and summer circulation, J. Geophys. Res. Oceans, 119, doi:10.1002/ 2013JC009004, 2014.

Yao, F., Hoteit, I., Pratt, L. J., Bower, A. S., Kohl, A., Gopalakrishnan, G., and Rivas, D.: Seasonal overturning circulation in the Red Sea: 2. Winter circulation, J. Geophys. Res. Oceans, 119, 2263–2289, doi:10.1002/ 2013JC009331, 2014.

Zarokanellos, N. D., Kürten, B., Churchill, J. H., Roder, C., Voolstra, C. R., Abualnaja, Y., and Jones, B. H.: Physical mechanisms routing nutrients in the central Red Sea. Journal of Geophysical Research: Oceans, 122. https://doi.org/10.1002/2017JC013017, 2017a.

Zarokanellos, N. D., Papadopoulos, V. P., Sofianos, S. S., and Jones, B. H.: Physical and biological characteristics of the winter-summer transition in the Central Red Sea. Journal of Geophysical Research: Oceans, 122, 6355–6370. https://doi.org/10.1002/2017JC012882, 2017b.

Zhan, P., Subramanian, A. C., Yao, F., and Hoteit, I.: Eddies in the Red Sea: A statistical and dynamical study, J. Geophys. Res. Oceans, 119, 3909–3925, doi:10.1002/2013JC009563, 2014.

Zelenke B. C.: An Empirical Statistical Model Relating Winds and Ocean Surface Currents. Master of Science in Oceanography - Thesis, Oregon State University, 2005.

Zorita E, and von Storch H.: The analog method as a simple statistical downscaling technique: comparison with more complicated methods. J Climate 12:2474–2489, 1999.

TABLES

Table 1: Characteristics of the previously developed STP works based on HFR data.

Authors	**Approach**	**Needs continuous training period**	**Comple-mentary data required?**	**Region of application**	**Reliable forecast period**
Zelenke 2005	EOF + bilinear regression model	Yes	Wind	Oregon coast	48 hours
Frolov et al. 2012	EOF + linear auto regression model	Yes	Wind and tides (optional)	Monterey Bay, California	48 hours
Barrick et al., 2012	Constant linear trend model applied to OMA modes	Yes	Wind	Finnmark, Norway	12 hours
Orfila et al. 2015	EOF+Genetic Algorithm	Yes	No	Toulon, France	48 hours
Solabarrieta et al. 2016	Frolov et al., 2012	Yes	No	Bay of Biscay	48 hours
Vilibić et al., 2016	SOM+neural network +winds	Yes	Wind	Northern Adriatic Sea	72 h
Ren et al., 2019	Random Forest (RF) classification algorithm	No	Tide and Wind	Galway Bay, Ireland	59 h
This paper: L-STP	Analogue finding	No	No	Bay of Biscay and the Central Red Sea	48 h

Table 2: Correlation coefficient values between winner ε_{ANL} and $\delta_{\_STP}$ and between ε_{ANL} and δ_PRS, after 6, 12, 24, 36 and 48 hours of simulation.

	6 hours	**12 hours**	**24 hours**	**36 hours**	**48 hours**
$\mathbf{R^2}$ $\boldsymbol{\varepsilon_{ANL} - \delta_{\_STP}}$	0.19	0.37	0.55	0.56	0.54
$\mathbf{R^2}$ $\boldsymbol{\varepsilon_{ANL} - \delta_{\_PRS}}$	0.07	0.11	0.03	0.01	0.04
ε_{ANL} [km2], for the inflection point between $\delta_{\_STP}$ and $\delta_{\_PRS}$	-	714	774	857	1027
% of ε_{ANL} (accumulative) for the previous line	-	81	84	87	95

FIGURES

Figure 1*: (Left) HFR system of the BoB. (Right) HFR system of the central Red Sea. Blue dots represent the data points and the black cross are the HFR antenna positions*

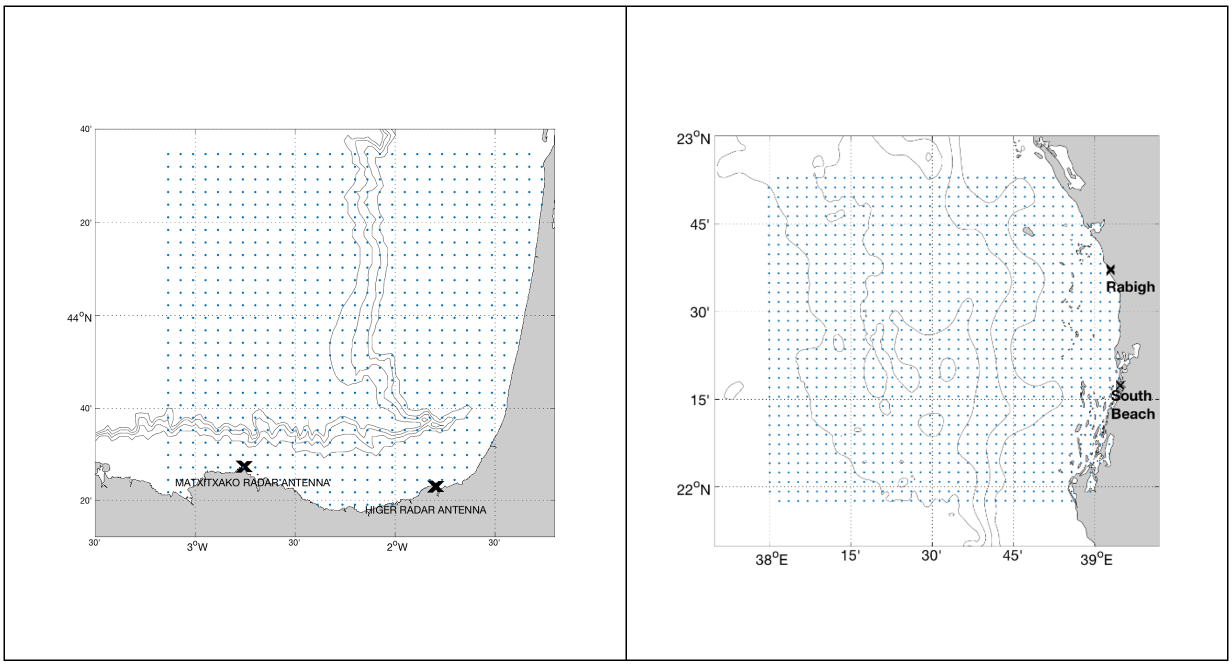


Figure 2: (1) 15-Apr-2015 00:00 example of the developed methodology applied to the BoB HFR system. (a) The past 48 hours of target field of test period (b) The analogue having the lowest error, (c) The truth trajectories for the forecast period (d) the STP trajectories. The initial positions of the particle trajectories are indicated by the blue dots, and the red dots indicate the position after 48 hours.

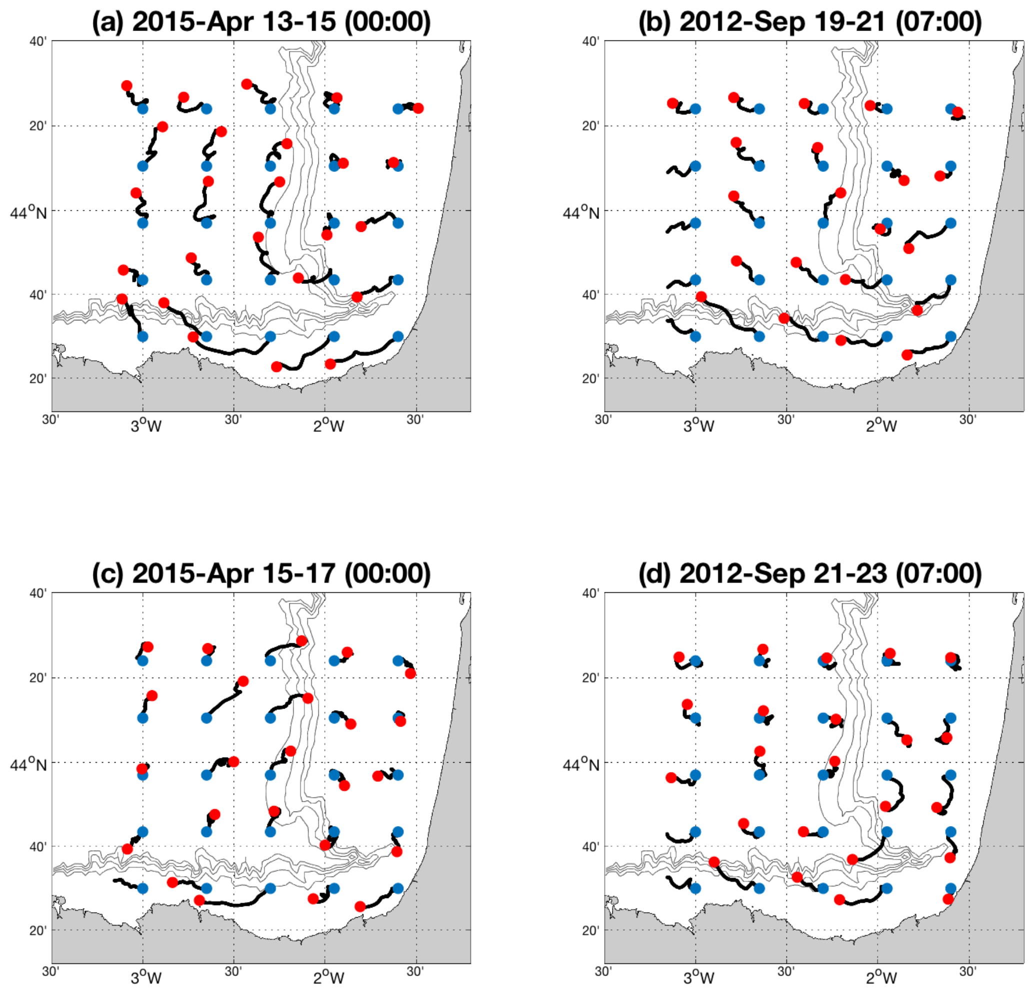


Figure 3: Example for the test period: 15-Apr-2015 00:00; errors for the whole Lagrangian catalogue fields of the BoB HFR System, restricted to the δ_cg = 10 km condition. The red dot indicates the occurrence date and the error of the winner analogue (19-Sep-2012 07:00).

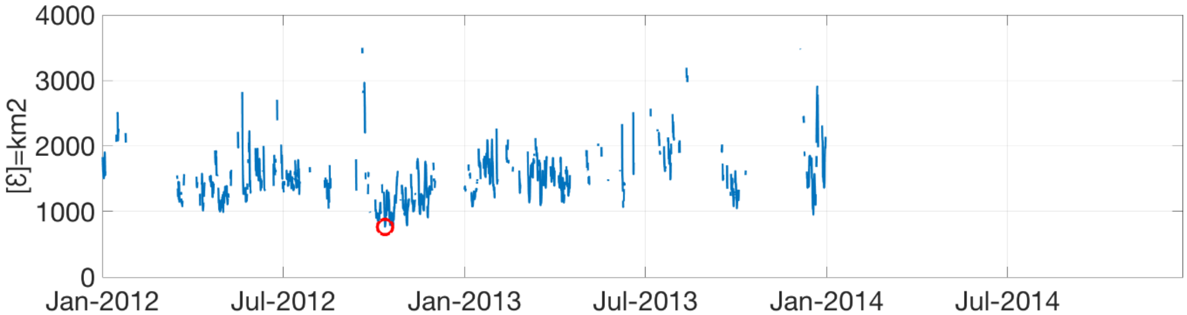
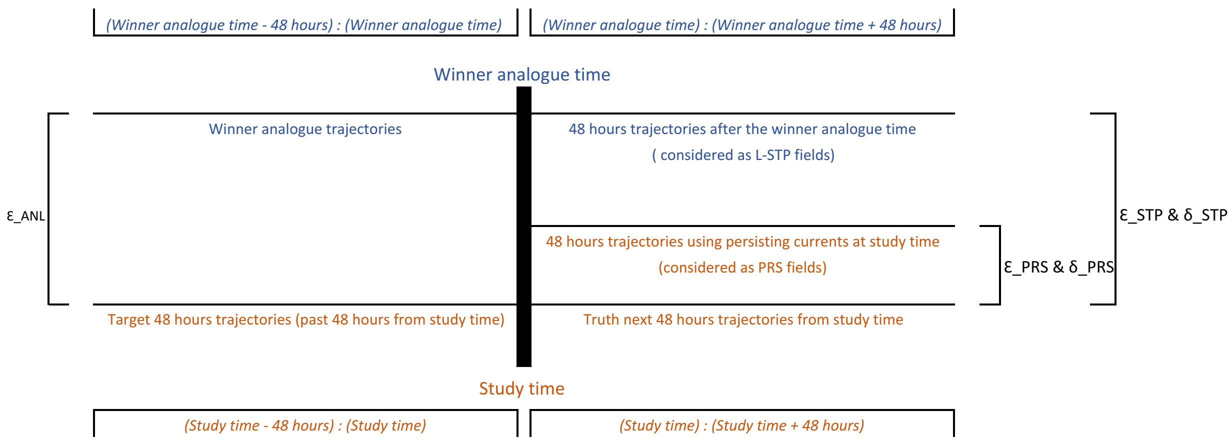


Figure 4**:** *Scheme of the analogue selection and L-STP forecast assessment process.*



Ɛ_ANL is used to select the winner/best analogue (min Ɛ)

Ɛ_STP , Ɛ_PRS , δ_STP and δ_PRS are used to validate the methodology and estimate the final error or separation distances between truth and L-STP/PRS trajectories

Figure 5: errors of the hourly winner analogue for 2015 (ε_{ANL}), together with the ε_{STP} and ε_{PRS}. The black dots over the timeline shows the times when ε_{STP} is higher than ε_{PRS}

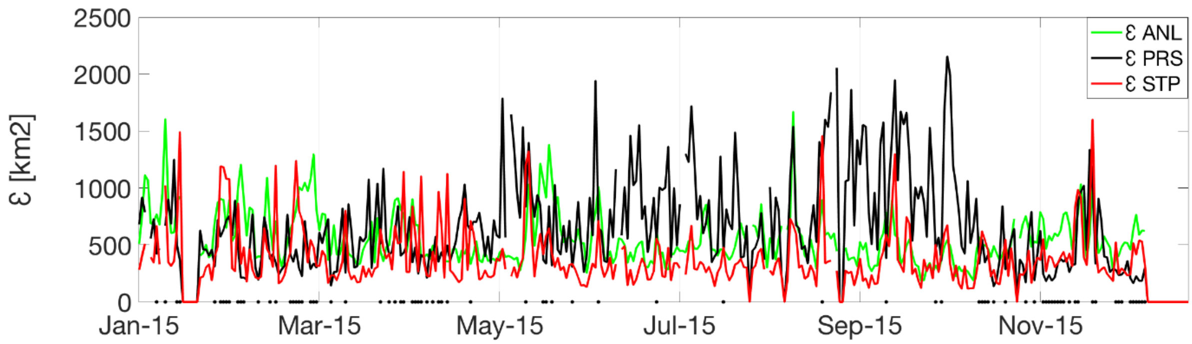


Figure 6: X axis shows the ε_{ANL}, ordered from minimum to maximum, for the winner analogue for the test year 2015. Left Y axis indicates ε_{STP} (red) and ε_{PRS} (blue) for the corresponding ε_{ANL}. Right Y axis indicates the % of the accumulative comparison times as shown by the black solid line. Dashed vertical line indicates the crossing point between ε_{STP} and ε_{PRS} (ε_{ANL}=853Km2).

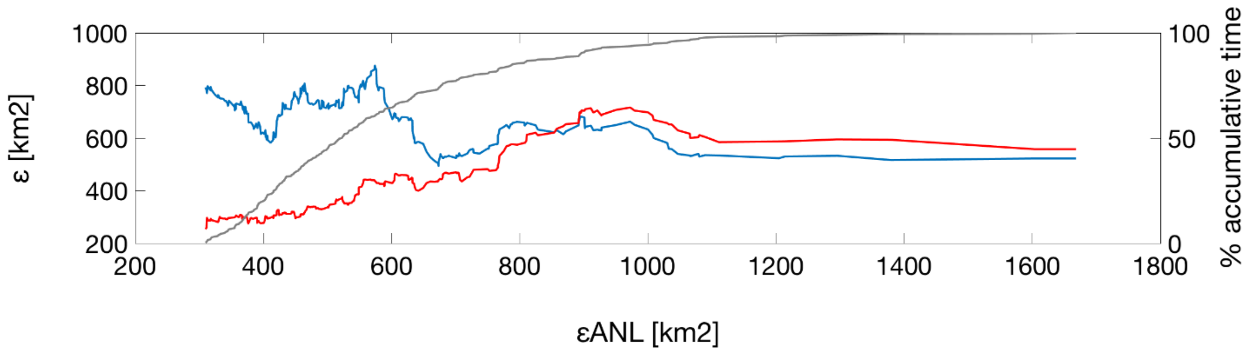


Figure 7: Left Y axis indicates δ_{STP} (red) and δ_{PRS} (blue) for the corresponding ε_{ANL}, after 6, 12, 24, 36 and 48 hours. Right Y axis is the cumulative % of timesteps in the computation of the mean errors, as indicated by the black line in the plots. X axis is the ε_{ANL}, ordered from minimum to maximum, for the winner analogue for the test year 2015.

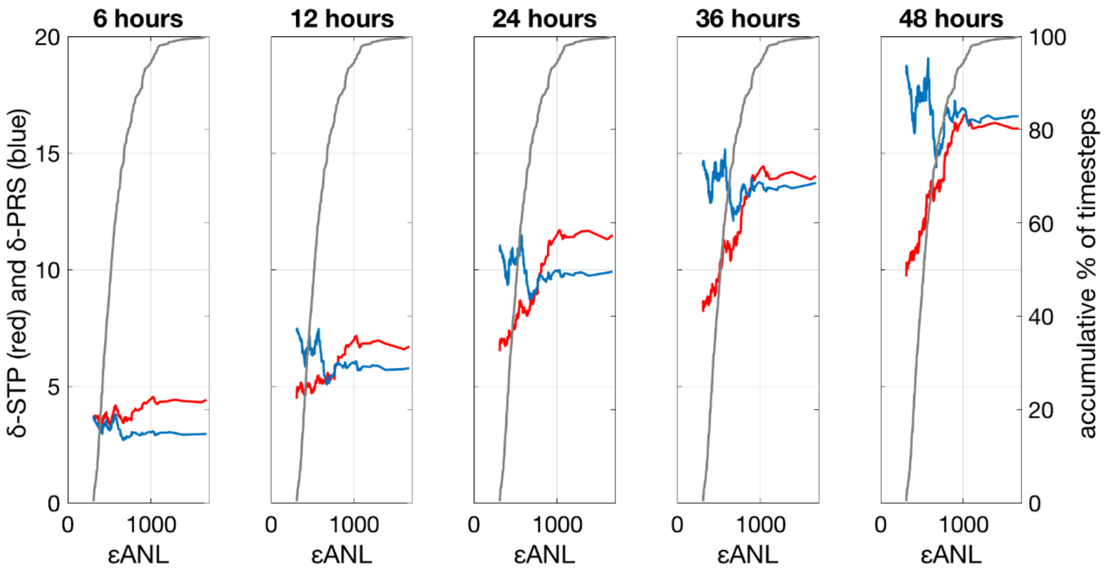


Figure 8: Time evolution of the mean separation δ_{STP} *and* δ_{PRS} *[km] between truth and forecast trajectories using truth and STP/PRS currents and the mean drift, with BoB system data, for 2015.*

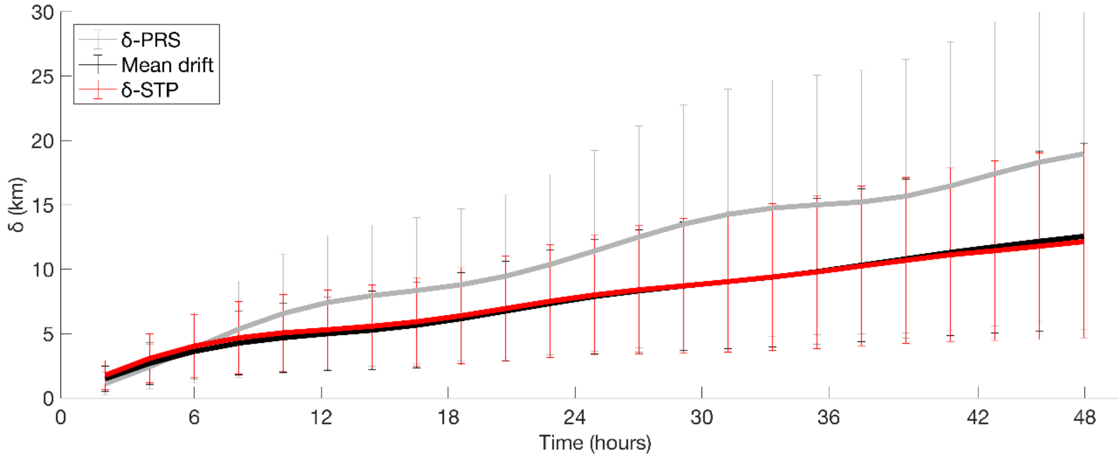


Figure 9: Time evolution of the mean separation distances $\delta_{\_STP}$ and $\delta_{\_PRS}$ [km] between real and forecast trajectories using truth and STP/PRS currents and the mean drift, with the Red Sea HFR system data, for July 2017 to October 2018.

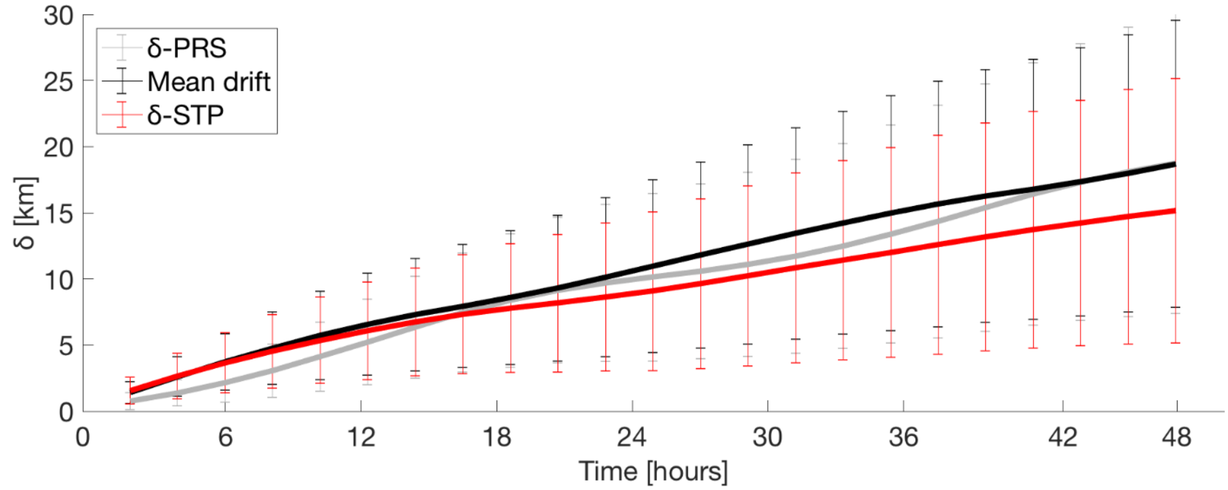


Figure 10***:*** *Spatial distribution of separation distances [km] between trajectories using L-STP and persistent currents at 6, 12, 24 and 48 hours, for the BoB HFR System.*

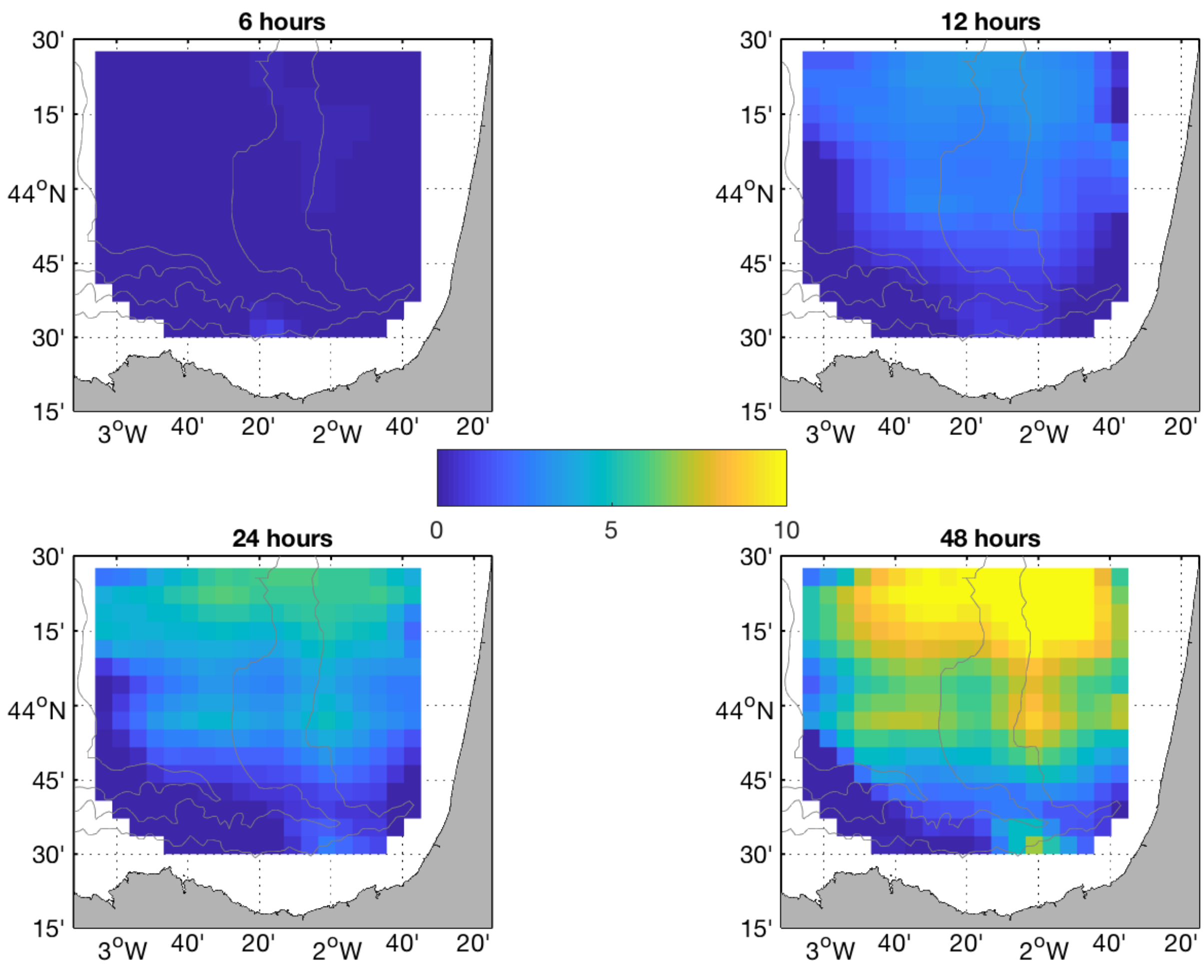


Figure 11*: Spatial distribution of separation distances [km] between trajectories using L-STP and persistent currents at 6, 12, 24 and 48 hours, for the Red Sea HFR system.*