@AGUPUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER

10.1002/2017GL076075

Key Points:

- The crossover time is defined as the time it takes for a model to equal persistence performance
- The concept is demonstrated by assessing a high-resolution global ocean model forecast and analysis
- The average crossover time calculated using the separation distance for the model analysis is 6 days

Correspondence to:

L. Phillipson, I.phillipson14@imperial.ac.uk

Citation:

Phillipson, L. M., & Toumi, R. (2018). The crossover time as an evaluation of ocean models against persistence. *Geophysical Research Letters*, 45. https://doi.org/10.1002/2017GL076075

Received 18 OCT 2017 Accepted 7 DEC 2017 Accepted article online 11 DEC 2017

©2017. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The Crossover Time as an Evaluation of Ocean Models Against Persistence

L. M. Phillipson¹ and R. Toumi¹

¹Space and Atmospheric Physics Group, Department of Physics, Imperial Collage London, London, UK

Abstract A new ocean evaluation metric, the crossover time, is defined as the time it takes for a numerical model to equal the performance of persistence. As an example, the average crossover time calculated using the Lagrangian separation distance (the distance between simulated trajectories and observed drifters) for the global MERCATOR ocean model analysis is found to be about 6 days. Conversely, the model forecast has an average crossover time longer than 6 days, suggesting limited skill in Lagrangian predictability by the current generation of global ocean models. The crossover time of the velocity error is less than 3 days, which is similar to the average decorrelation time of the observed drifters. The crossover time is a useful measure to quantify future ocean model improvements.

1. Introduction

Improvements in ocean current modeling are important for numerous applications from tracking marine pollution Marta-Almeida et al. (2013) to understanding the global circulation Melsom et al. (2009). Therefore, the benchmarking of ocean current forecasts is important in quantifying improvements.

Persistence has been extensively used in meteorology (Van den Dool, 2007) but less in oceanography (Farrara et al., 2013; Oey et al., 2005; Rowe et al., 2016; Shriver et al., 2007; Solabarrietaa et al., 2016; Vandenbulcke et al., 2009) as a performance benchmark. For quantifying improvements in ocean current forecasts, particularly the Lagrangian predictability, persistence has been typically taken as either a constant last known position (Barron et al., 2007; Kuang et al., 2012; Schmidt & Gangopadhyay, 2013; Ullman et al., 2006) or constant last known velocity (Paldor et al., 2004; Rixen & Ferreira-Coelho, 2007) of a surface-drifting buoy held constant. Surface-drifting buoys provide unique observations of the ocean currents capturing various scales of motion (Lumpkin et al., 2017). The observed trajectory paths can be compared with simulated trajectories from the model providing a stringent test of the accumulation of errors and the chaotic nature of Lagrangian flow (Özgökmen et al., 2000).

There are a number of knowledge gaps we aim to fill. All these studies where done for specific regions, and here we present a first global analysis of ocean model performance against persistence. So far, the evaluation of persistence has been given as an error or skill score for different times. This lacks an easy physical interpretation. Here we propose a new diagnostic in quantifying improvements in ocean modeling, denoted as the crossover time. This represents the time scale after which the model performs better than persistence for any chosen metric.

The paper is organized as follows; the crossover time, persistence, and evaluated ocean model are presented in section 2, the crossover time analysis in section 3, and discussion in section 4 followed by the conclusion in section 5.

2. Methodology

The crossover time provides a way of understanding the relative predictability of a model with respect to the benchmark forecast of persistence. For this study, persistence is formulated as keeping the last known velocity, V_0 . This constant velocity is then persisted throughout the prediction, representing perhaps the simplest forecast that can be made from the drifters themselves. This is a natural first simple forecast for a numerical model to outperform.

AGU Geophysical Research Letters



Figure 1. A schematic of the crossover time framework in terms of the Lagrangian separation distance, $T_c(d)$. The drifter is in red with time markers in days, the model in blue, and persistence in black. The black cross denotes the approximate location of $T_c(d)$ for this example.

We then define the crossover time as the time, T_c at which the model is just about to outperform persistence, that is, when the model and persistence performance are equal:

$$T_c \mid M(T_c)_p - M(T_c)_m = 0$$
 (1)

where *M* is a measure of goodness as a function of time, which can be flexibly defined and based on any chosen error or skill metric. The subscripts *p* and *m* denote that *M* has been estimated using the persistence and model results, respectively, and T_c is the crossover time given that the persistence (M_p) and model (M_m) performance are now equal, that is, the moment the model is about to "crossover," becoming the best predictor.

The better the model performs, the shorter time it takes to outperform persistence and thus the shorter is T_c . Conversely, the better persistence performs (linked to the memory of the local circulation), the more challenging it is for the model and thus the longer is T_c .

An example of a reasonable choice of *M* is the commonly used Lagrangian separation distance (Barron et al., 2007; Berta et al., 2015; Castellari et al., 2001; Liu et al., 2014; Phillipson & Toumi, 2017; Rixen & Ferreira-Coelho, 2007; Muscarella et al., 2015), that is, the separation distance between simulated trajectories (computed via an advection scheme) and observed drifters (km):

$$M(t) = d(t) = \sqrt{[O_x(t) - S_x(t)]^2 + [O_y(t) - S_y(t)]^2}$$
(2)

where the subscript x represents the longitudinal position and y the latitudinal position at time t. S is the simulated trajectory for either the model (M_m) or persistence (M_p) , and O is the observations.

To demonstrate the flexibility of the concept, the T_c is also calculated for the root-mean-square error (RMSE) of the model (persistence) velocities and inferred velocities of the observed drifters (m/s):

$$M(t) = \text{RMSE}_{uv}(t) = \sqrt{\frac{\sum_{i=1}^{n} (S_i(t) - O_i(t))^2}{n}}$$
(3)

where S changes from equation (2) and instead denotes either the model-simulated velocity fields (U and V) interpolated spatially to the location of the drifters and temporally to the 6-hourly output frequency (M_m) or the persistent velocity (M_p) for every *i* drifter, at time *t*.

 $T_c(d)$ therefore refers to the crossover time of the Lagrangian separation distance *d* (equation (2)), and $T_c(\text{RMSE})$ refers to the crossover time of the velocity RMSE (equation (3)).

For illustration, Figure 1 is a schematic of the crossover time calculated using the separation distance $T_c(d)$ between a model, persistence, and drifter trajectory. Here $T_c(d)$ occurs where the separation distance for the model is equal to the separation distance for persistence. $T_c(d)$ is about 3 days here. For a perfect model, the drifter would represent the exact path. And therefore, always outperform persistence with $T_c(d)$ equal to 0.

To quantify and test the concept of the T_c , the global 1/12° resolution MERCATOR ocean model forecast and analysis (http://www.mercator-ocean.fr) are assessed for July 2015 and January 2016 using Surface Velocity Program (SVP) drifters maintained by the Global Drifter Program's (GDP) Drifter Data Assembly Centre (Lumpkin & Pazos, 2007). The model forecast is run out 7 days, so we compare four model forecast and analysis cycles 7 days apart in each month. This provides over 8,000 independent trajectories globally, allowing for robust statistics of the metric. The model analysis contains the assimilation of satellite sea surface temperature (SST) and sea surface height (SSH).

To compute the separation distance (equation (2)), the trajectories for both the persistence and MERCATOR models are advected using a fourth-order Runge-Kutta advection scheme (Abdel Karim, 1966). For simplicity, only the advection (i.e., the forcing of the ocean currents alone) acts on the simulated trajectories.

3. Results

3.1. Global Summary

The persistent trajectories outperformed both the model analysis and forecast trajectories, for the majority of the prediction of separation distance. The instantaneous persistence (last known velocity) performed significantly better of up to 6 days compared to the model analysis and greater than 6 days compared to the model forecast. Calculating an average persistence through averaging prior observed drifter velocities degrades the persistence performance (not shown). The last known velocity is typically the optimal choice.

The improvement of the persistence trajectories over the model analysis trajectories initially increases with time (Figure 2a), peaking at approximately 3 days with a separation distance as much as 13.5 km (Figure 2a). The separation distance increases with time, and so does the relative difference between the model analysis and persistence. However, the separation distance of persistence increases at a slower rate, so persistence is a more effective predictor on short time scales. Beyond 3 days, the improvement of the model analysis increases with time and is the best predictor on average after about day 6 (Figure 2a). This $T_c(d)$ of 6 days represents the limits of persistence as a useful predictor compared to the model analysis. The model forecast fails to ever reach a crossover time for the entire span of the available prediction. During the study period the average decorrelation times of the drifter velocities was found to be 1.2 days (not shown) in approximate agreement with Poulain et al. (1996), LaCasce (2008), Döös et al. (2013), and Lumpkin and Johnson (2013). This time scale is very similar to the *e*-folding time scale of persistence (black line in Figure 2a).

It is worth noting that while we concentrate on the average of the trajectory error sample, the spread is often larger than the value of the median. This large spread is expected for chaotic trajectories. Nevertheless, the average is useful to demonstrate the concept of the crossover time.

For the velocity error and T_c (RMSE) (Figures 2b and 2c), we find a significantly shorter crossover time of about 1–3 days, close to the average decorrelation time of the drifter velocities (1.2 days). Beyond this time, both the model analysis and forecast velocities were significantly more skillful than the persistent velocity by 0.05 m/s on average (25–30%). The velocity RMSE of the model forecast and analysis are very large (0.15–0.2 m/s) and do not increase as the prediction horizon increases. This behavior suggests that the error may represent the unresolved subgrid-scale processes of the model, that is, the model noise.

The $T_c(d)$ of 6 days for the model analysis trajectories of 6 days corresponded to a drifter displacement (a measure of how far on average the drifters have displaced from their initial location) of approximately 90 km (Figure 2a). A second time of interest is when the model analysis begins to start at least to improve on persistence; this occurs near day 3 when the typical drifter displacement was about 50 km or nearly 5 times the global model resolution. However, for T_c (RMSE) (Figures 2b and 2c), at the shorter crossover time (1–3 days) the drifter displacement corresponded to about 20–30 km — approximately 2–3 times the model resolution.

3.2. Spatial Distribution

The spatial variability of the crossover times is next explored. Although a whole year would be preferable, 2 months is a computational compromise to capture adequate variability and demonstrate the concept further.

On day 3 the model analysis trajectories (Figure 3a) generally have more skill in the open ocean and less skill in the more energetic regions such as the boundary currents. Similarly, the persistence trajectories (Figure 3c) also have larger relative separation distances in energetic regions. However, in a direct comparison (comparing Figures 3a and 3c), persistence shows more skill over the model analysis independent of location to first order.

On day 6 the model analysis trajectories (Figure 3b) perform significantly better in comparison to the persistence trajectories (Figure 3d) than at day 3 with a similar proportion of lower separation distance regions. At this later stage in the prediction, the inhomogeneity of the differences becomes more noticeable. Regions of higher kinetic energy (Ishikawa et al., 1997) were better captured by the model analysis especially the energetic region of Atlantic Circumpolar Current (ACC). However, it should be noted that even at this later prediction time some regions are still strongly favored by persistence such as the Peru Basin, South Equatorial current, the subpolar North Pacific, and central subtropical North Atlantic.

Figure 3e shows the spatial distribution of $T_c(d)$. The majority of the ocean model has a $T_c(d)$ of greater than 3 days. However, large local variability is evident; such is the case for the Pacific Ocean, for example.



Figure 2. (a) Global median separation distance (km) between observed and modeled drifter positions, comparing the simulated trajectories of the MERCATOR model forecast (blue dashed), and analysis (blue solid) and persistence (black solid) as a function of forecast time (days). The average observed drifter displacement (km) is overlain in red. The stem is the difference between the MERCATOR model forecast and persistence (crosses) and MERCATOR analysis and persistence (circles). The dashed lines in light gray mark the $T_c(d)$ and the corresponding drifter displacement (km) for the MERCATOR analysis. (b and c) U and V RMSE (m/s) between observed drifter velocities and modeled/persisted velocity for the model forecast (blue dashed)/analysis (blue solid) and persistence (black solid) as a function of forecast time (days). The average observed drifter displacement (km) on the right-hand y axis in red. T_c (RMSE) and the drifter displacement (km, right-hand y axis) for the MERCATOR analysis (dashed gray).

10.1002/2017GL076075

AGU Geophysical Research Letters



Figure 3. The separation distance (km) for the (a,b) model analysis trajectories and (c,d) persistence trajectories, after 3 days (Figures 3a and 3c) and 6 days (Figures 3b and 3d) and (e) the associated crossover time, $T_c(d)$. The data are averaged into 20° bins over January and July.

Qualitatively, some regions can be highlighted where the $T_c(d)$ is consistently large or small. The Peru Basin, for instance, has the most uniform cluster of larger $T_c(d)$. Conversely, groups of shorter $T_c(d)$ were located in parts of the ACC between South America and Africa.

For T_c (RMSE) (not shown) the majority of the ocean has a shorter T_c (RMSE) of less than 3 days. Areas of comparatively larger T_c (RMSE) closely resemble those of larger T_c (d) (Figure 3e) such as in the Peru Basin.

4. Discussion

The globally averaged $T_c(d)$ is 6 days for the model analysis and even longer for the model forecast. Therefore, on average the persistence trajectories outperforms both models for the majority of the available days. By defining persistence as the last known velocity persisted, the skill of persistence relies on the memory of the local circulation. Through testing the averaging of the velocities of the drifters back in time (previous 1 to 3 days, for example), we found that taking the last known velocity with no prior averaging was the optimal choice for persistence skill. This averaging could be degrading persistence skill due to the short average decorrelation time (1.2 days) and sampling inconsistencies.

Persistence uses the observations directly with no interpolation. Thus, persistence can initially capture scales of motion not represented by the numerical model horizontal grid-scale resolution of about 8 km. Many studies (Griffa et al., 2004; Huntley et al., 2011; Putman & He, 2013) have noted the effect of the horizontal grid resolution in ocean models on Lagrangian predictability. Recently, Carlson et al. (2016) investigated Lagrangian transport in the Adriatic Sea using Regional Ocean Modeling System (ROMS) with a high horizontal resolution of 2 km. They noted separation distances between simulated trajectories and observed that

drifters were still around $5-6 \text{ km d}^{-1}$ but approximately 60% lower than for the global model evaluated in this study ($12-14 \text{ km d}^{-1}$). This highlights the importance of unresolved subgrid-scale processes on Lagrangian predictability.

Despite the resolved initial conditions, a major limitation of persistence is the simple linear nature of the prediction with no dynamical circulation and thus completely depends on retaining the memory of the local ocean velocity. Persistence typically has very high skill at the start of the forecast, t_0 which naturally degrades as the correlation between t_0 and t_{0+i} reduces. The extent of this correlation can be quantified via the decorrelation time of the drifter velocities (1.2 days on average). This is similar to both the *e*-folding time scale of the Lagrangian persistence trajectories (Figure 2a) and the average crossover time of the velocity error, T_c (RMSE) (1–2 days). This similarity is perhaps unsurprising and highlights the link between the memory of the ocean circulation and the performance of the persistent velocity as compared to the model noise.

 $T_c(d)$ is much longer than both $T_c(\text{RMSE})$ and the decorrelation time at 6 days. This difference could be the result of the accumulation of errors when integrating velocities (Huntley et al., 2011). The improvement of the persistent velocity over the model velocities for the first 1–2 days (approximately the length of decorrelation time) is thus accumulated, translating to much longer time advantages in the Lagrangian framework, and then $T_c(d)$ is on average 6 days. An alternative interpretation is that $T_c(d)$ represents the scale at which the observed drifters are no longer following a straight line. The influence of larger-scale circulation patterns is then much more important which can be captured by the dynamical model.

The displacement scale of the observations at T_c (RMSE) is approximately 20–30 km, close to the model resolution or twice the size of the model grid separation (8 km). The persistent velocity is initially a more accurate approximation of the subgrid-scale motion than the model via directly sampling the observations. However, once the drifters have displaced sufficiently outside the model resolution, this advantage is lost. At this point, the drifter is no longer within the initially sampled area, and the initial velocity is less likely to still hold true. The model has the advantage of a dynamically evolving large-scale circulation (capturing processes that persistence cannot capture) but has limited resolution of the subgrid-scale processes (Huntley et al., 2011).

The core advantage of persistence is the highly resolved sampling of the local motion which is adequately representative of the actual drifter for the first few days. Therefore, with dramatic improvements in ocean model resolution, the model should perform better against persistence. The crossover time is a useful metric to demonstrate model improvements, particularly increasing resolution.

There are large spatial variations in the separation distance (Figures 3a-3d) and $T_c(d)$ (Figure 3e). A large $T_c(d)$ can be due to the actual local motion (high persistence) and/or model error. The model trajectory error may produce variability in $T_c(d)$ that cannot be explained in terms of local dynamical features. However, some dynamical features can be broadly recognized. The open ocean has a lower crossover time than the energetic boundary current regions. Gille and Kelly (1996) concluded that the ACC scales are determined by local instability mechanisms with short spatial decorrelation scales of about 85 km and longer time scales of 34 days. Shorter $T_c(d)$ present in many parts of the ACC could be linked to these shorter spatial scales. Lumpkin and Johnson (2013) have identified regions of "eddy deserts" with significantly low time mean eddy speeds. Here these regions, that is, the Peru Basin, South Equatorial current, the subpolar North Pacific, and Central subtropical North Atlantic typically have larger $T_c(d)$. The local circulation memory is likely maintained for longer in slowly evolving dynamics, improving the performance of persistence trajectories. Therefore, the ocean model has a tougher persistence benchmark where Lagrangian trajectories are inherently more predictable in the model, that is, in areas of lower square root of the velocity variance (Barron et al., 2007; Özgökmen et al., 2000).

5. Conclusion

A new evaluation metric, the crossover time, is introduced to examine the performance of a global ocean model with respect to persistence. Persistence represents the most straightforward forecast that can be made directly from the observations. The crossover time was estimated using the Lagrangian separation distance and velocity error, for a global ocean model forecast and analysis. For the Lagrangian separation distance, the persistence trajectories outperform the model analysis trajectories for the majority of 7 day prediction, with a corresponding average crossover time of approximately 6 days. For the velocity error, the crossover time is significantly smaller at about 1–3 days, similar to the decorrelation time scales of the observed drifters.

The significant difference between crossover times in the separation distance and velocity error can be understood in terms of the nature of trajectory prediction and their inherent accumulation of velocity errors.

Spatial variations in the crossover time are explained in terms of the model error and some regional dynamical features.

The crossover time can be used to evaluate ocean models giving simple, intuitive, and physical insight. The crossover time concept could also be useful to evaluate the forecast performance of any other ocean variables such as sea surface temperature and the mixed layer depth. Persistence, a simple benchmark forecast, is a good short-term predictor of Lagrangian trajectories but as numerical models improve (increased resolution, improved model physics, etc.) the crossover time is expected to shorten.

References

- Abdel Karim, A. I. (1966). The stability of the fourth order Runge-Kutta method for the solution of systems of differential equations. *Communications of the ACM*, 9(2), 113–116. https://doi.org/10.1145/365170.365213
- Barron, C. N., Smedstad, L. F., Dastugue, J. M., & Smedstad, O. M. (2007). Evaluation of ocean models using observed and simulated drifter trajectories: Impact of sea surface height on synthetic profiles for data assimilation. *Journal of Geophysical Research*, 112, C07019. https://doi.org/10.1029/2006JC003982
- Berta, M., Griffa, A., Magaldi, M. G., Özgökmen, T. M., Poje, A. C., Haza, A. C., & Josefina Olascoaga, M. (2015). Improved surface velocity and trajectory estimates in the Gulf of Mexico from blended satellite altimetry and drifter data. *Journal of Atmospheric and Oceanic Technology*, 32(10), 1880–1901. https://doi.org/10.1175/JTECH-D-14-00226.1
- Carlson, D. F., Griffa, A., Zambianchi, E., Suaria, G., Corgnati, L., Magaldi, M. G., . . . Borghini, M. (2016). Observed and modeled surface Lagrangian transport between coastal regions in the Adriatic Sea with implications for marine protected areas. *Continental Shelf Research*, *118*, 23–48. https://doi.org/10.1016/j.csr.2016.02.012
- Castellari, S., Griffa, A., Ozgokmen, T., & Poulain, P. (2001). Prediction of particle trajectories in the Adriatic Sea using Lagrangian data assimilation. *Journal of Marine Systems*, 29(1-4), 33-50. https://doi.org/10.1016/S0924-7963(01)00008-2
- Döös, K., Kjellsson, J., & Soomere, T. (2013). Evaluation and tuning of model trajectories and spreading rates in the Baltic Sea using surface drifter observations, *Preventive methods for coastal protection: Towards the use of ocean dynamics for pollution control* (pp. 251–281). Heidelberg: Springer International Publishing. https://doi.org/10.1007/978-3-319-00440-2_8
- Farrara, J. D., Chao, Y., Li, Z., Wang, X., Jin, X., Zhang, H., ... Colas, F. A. (2013). A data-assimilative ocean forecasting system for the Prince William sound and an evaluation of its performance during sound Predictions 2009. Continental Shelf Research, 63, S193–S208. https://doi.org/10.1016/j.csr.2012.11.008
- Gille, S. T., & Kelly, K. A. (1996). Scales of spatial and temporal variability in the Southern Ocean. Journal of Geophysical Research, 101(C4), 8759–8773. https://doi.org/10.1029/96JC00203
- Griffa, A., Piterbarg, L. I., & Özgökmen, T. (2004). Predictability of Lagrangian particle trajectories: Effects of smoothing of the underlying Eulerian flow. *Journal of Marine Research*, 62(1), 1–35. https://doi.org/10.1357/00222400460744609
- Huntley, H. S., Lipphardt, B., & Kirwan, A. (2011). Lagrangian predictability assessed in the East China Sea. Ocean Modelling, 36(1-2), 163-178. https://doi.org/10.1016/j.ocemod.2010.11.001
- lshikawa, Y., Awaji, T., & Akitomo, K. (1997). Global surface circulation and its kinetic energy distribution derived from drifting buoys. *Journal of Oceanography*, 53, 489–516.
- Kuang, L., Blumberg, A. F., & Georgas, N. (2012). Assessing the fidelity of surface currents from a coastal ocean model and HF radar using drifting buoys in the Middle Atlantic Bight. Ocean Dynamics, 62(8), 1229–1243. https://doi.org/10.1007/s10236-012-0556-2 LaCasce, J. H. (2008). Statistics from Lagrangian observations, 77(1), 1–29. https://doi.org/10.1016/j.pocean.2008.02.002
- Liu, Y., Weisberg, R. H., Vignudelli, S., & Mitchum, G. T. (2014). Evaluation of altimetry-derived surface current products using Lagrangian drifter trajectories in the eastern Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 119, 2827–2842. https://doi.org/10.1002/2013JC009710
- Lumpkin, R., & Johnson, G. C. (2013). Global ocean surface velocities from drifters: Mean, variance, El Niño–Southern Oscillation response, and seasonal cycle. Journal of Geophysical Research: Oceans, 118, 2992–3006. https://doi.org/10.1002/jgrc.20210
- Lumpkin, R., & Pazos, M. (2007). Measuring surface currents with Surface Velocity Program drifters: The instruments, its data and some recent results. In A. Griffa, et al. (Eds.), Lagrangian analysis and prediction of coastal and ocean dynamics (pp. 39–67). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511535901.003
- Lumpkin, R., Oz Okmen, T., & Centurioni, L. (2017). Advances in the application of surface drifters. *Annual Review of Marine Science*, *9*, 59–81. https://doi.org/10.1146/annurev-marine-010816-060641
- Marta-Almeida, M., Ruiz-Villarreal, M., Pereira, J., Otero, P., Cirano, M., Zhang, X., & Hetland, R. D. (2013). Efficient tools for marine operational forecast and oil spill tracking. *Marine Pollution Bulletin*, 71(1–2), 139–151. https://doi.org/10.1016/j.marpolbul.2013.03.022
- Melsom, A., Lien, V. S., & Budgell, W. P. (2009). Using the Regional Ocean Modeling System (ROMS) to improve the ocean circulation from a GCM 20th century simulation. *Ocean Dynamics*, 59, 969–981. https://doi.org/10.1007/s10236-009-0222-5
- Muscarella, P., Carrier, M. J., Ngodock, H., Smith, S., Lipphardt, B. L., Kirwan, A. D., & Huntley, H. S. (2015). Do assimilated drifter velocities improve Lagrangian predictability in an operational ocean model? *Monthly Weather Review*, 143(5), 1822–1832. https://doi.org/10.1175/mwr-d-14-00164.1
- Oey, L. Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., & Fan, S. (2005). An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico. *Geophysical Research Letters*, *32*(12), 1–4. https://doi.org/10.1029/2005GL023253
- Özgökmen, T. M., Griffa, A., Mariano, A. J., & Piterbarg, L. I. (2000). On the predictability of Lagrangian trajectories in the ocean. Journal of Atmospheric and Oceanic Technology, 17(3), 366–383. https://doi.org/10.1175/1520-0426(2000)017<0366:OTPOLT>2.0.CO;2
- Paldor, N., Dvorkin, Y., Mariano, A. J., Özgökmen, T., & Ryan, E. H. (2004). A practical, hybrid model for predicting the trajectories of near-surface ocean drifters. *Journal of Atmospheric and Oceanic Technology*, 21(8), 1246–1258. https://doi.org/10.1175/ 1520-0426(2004)021<1246:APHMFP>2.0.CO;2
- Phillipson, L., & Toumi, R. (2017). Impact of data assimilation on ocean current forecasts in the Angola Basin. Ocean Modelling, 114, 45–58. https://doi.org/10.1016/j.ocemod.2017.04.006

Acknowledgments

This study was supported by the Natural Environment Research Council and has been conducted using the E.U. Copernicus Marine Service Information via the MERCATOR ocean model (http://www.mercator-ocean.fr). The drifter data set (used for both the observations and persistence forecasts) was processed by the AOML NOAA Global Drifter Program and is available at http://www.aoml.noaa.gov. The authors would like to thank the Imperial College High Performance Computing Service for providing computational resources. We thank the two anonymous reviewers who helped improve the manuscript.

Poulain, P.-M., Warn-Varnas, A., & Niiler, P. P. (1996). Near-surface circulation of the Nordic seas as measured by Lagrangian drifters. Journal of Geophysical Research, 101(15), 18,237–18,258. https://doi.org/10.1029/96JC00506

Putman, N. F., & He, R. (2013). Tracking the long-distance dispersal of marine organisms: Sensitivity to ocean model resolution. Journal of the Royal Society, Interface, 10(81), 20120979. https://doi.org/10.1098/rsif.2012.0979

Rixen, M., & Ferreira-Coelho, E. (2007). Operational surface drift prediction using linear and non-linear hyper-ensemble statistics on atmospheric and ocean models. *Journal of Marine Systems*, 65(1–4), 105–121. https://doi.org/10.1016/j.jmarsys.2004.12.005

- Rowe, M. D., Anderson, E. J., Wynne, T. T., Stumpf, R. P., Fanslow, D. L., Kijanka, K., ... Davis, T. W. (2016). Vertical distribution of buoyant Microcystis blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *Journal of Geophysical Research: Oceans*, 121, 5296–5314. https://doi.org/10.1002/2016JC011720
- Schmidt, A., & Gangopadhyay, A. (2013). An operational ocean circulation prediction system for the western North Atlantic: Hindcasting during July–September of 2006. Continental Shelf Research, 63, S177–S192. https://doi.org/10.1016/j.csr.2012.08.017
- Shriver, J. F., Hurlburt, H. E., Smedstad, O. M., Wallcraft, A. J., & Rhodes, R. C. (2007). 1/32° real-time global ocean prediction and value-added over 1/16° resolution. *Journal of Marine Systems*, 65(1–4), 3–26. https://doi.org/10.1016/j.jmarsys.2005.11.021
- Solabarrietaa, L., Frolov, S., Cook, M., Paduan, J., Rubio, A., González, M., Mader, J., & Charria, G. (2016). Skill assessment of HF radar-derived products for Lagrangian simulations in the Bay Of Biscay. *Journal of Atmospheric and Oceanic Technology*, 33(12), 2585–2597. https://doi.org/10.1175/JTECH-D-16-0045.1
- Ullman, D. S., O'Donnell, J., Kohut, J., Fake, T., & Allen, A. (2006). Trajectory prediction using HF radar surface currents: Monte Carlo simulations of prediction uncertainties. *Journal of Geophysical Research*, 111, C12005. https://doi.org/10.1029/2006JC003715

Van den Dool, H. M. (2007). Methods in short-term climate prediction. In *Empirical methods in short-term climate prediction* (Chap. 8, pp. 121–157). Oxford: Oxford University Press.

Vandenbulcke, L., Beckers, J. M., Lenartz, F., Barth, A., Poulain, P. M., Aidonidis, M., ... Rixen, M. (2009). Super-ensemble techniques: Application to surface drift prediction. *Progress in Oceanography*, 82(3), 149–167. https://doi.org/10.1016/j.pocean.2009.06.002