Ocean Modelling 29 (2009) 269–276

Contents lists available at ScienceDirect

Ocean Modelling

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

Lagrangian validation of numerical drifter trajectories using drifting buoys: Application to the Agulhas system

E. van Sebille[a,*], P.J. van Leeuwen[a,1], A. Biastoch[b], C.N. Barron[c], W.P.M. de Ruijter[a]

[a] Institute for Marine and Atmospheric Research Utrecht, Utrecht University, The Netherlands
[b] Leibniz Institute of Marine Sciences, Kiel, Germany
[c] Naval Research Laboratory, Stennis Space Center, Mississippi, USA

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form 19 May 2009
Accepted 20 May 2009
Available online 28 May 2009

Keywords:
Numerical Lagrangian drifters
Model validation
Kolmogorov–Smirnov statistics
Drifting buoys

ABSTRACT

The skill of numerical Lagrangian drifter trajectories in three numerical models is assessed by comparing these numerically obtained paths to the trajectories of drifting buoys in the real ocean. The skill assessment is performed using the two-sample Kolmogorov–Smirnov statistical test. To demonstrate the assessment procedure, it is applied to three different models of the Agulhas region. The test can either be performed using crossing positions of one-dimensional sections in order to test model performance in specific locations, or using the total two-dimensional data set of trajectories. The test yields four quantities: a binary decision of model skill, a confidence level which can be used as a measure of goodness-of-fit of the model, a test statistic which can be used to determine the sensitivity of the confidence level, and cumulative distribution functions that aid in the qualitative analysis. The ordering of models by their confidence levels is the same as the ordering based on the qualitative analysis, which suggests that the method is suited for model validation. Only one of the three models, a 1/10° two-way nested regional ocean model, might have skill in the Agulhas region. The other two models, a 1/2° global model and a 1/8° assimilative model, might have skill only on some sections in the region.

1. Introduction

Assessing the skill of ocean models is an important step before the data produced by such a model can be analyzed and interpreted. Special projects have been set up to facilitate the comparison of different ocean models within a fixed framework (e.g. the Coordinated Ocean-ice Reference Experiments (CORE), Griffies et al. (2009)). One of the problems of such skill assessment is that the observations to which the model should be verified are scarce in space and time. The skill assessment is therefore, often limited to a subset of the state vector.

Historically, verification is predominantly qualitative, where one or more specific model variables are compared to observations of these variables. The advantage of this qualitative method is that it introduces the expertise of the modeler in selecting fields and regions that are more important than others. However, the qualitative method also introduces subjectiveness into the skill assessment procedure.

There are objective methods to assess the model skill. Hetland (2006) introduced a way to calculate the improvement of a model with respect to some climatology. Using statistics on the complete model domain, however, has the disadvantage that dynamically relevant regions (such as the western boundary currents) are treated similar to dynamically less important regions. This is a relevant problem especially when the subsequent data analysis is done using numerical Lagrangian floats, tracers that are advected with the flow. These floats often cluster in some regions of the model domain and only the model skill in these regions is relevant for the aptitude of the float data. Ideally, these regions should, therefore, have more weight in the skill assessment. A way to accomplish this focus on dynamically relevant regions is to base the skill assessment on the float trajectories themselves.

The assumption behind trajectory verification is that only skillful models produce trajectories with similar properties as drifting buoys. Therefore, a high skill in float trajectories implies that the underlying model is highly skilled. Here, we present a quantitative method to assess the skill of a set of numerical drifters. Using real-world drifting buoy trajectories, the chance can be calculated that the drifting buoys and the numerical drifters are drawn from the same distribution.

For assimilative models, where it is the objective for the model to represent the ocean state as accurately as possible, Barron et al. (2007) have developed a technique to compare drifting buoy

* Corresponding author. Address: Department of Physics and Astronomy, Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Princetonplein 5, Utrecht, The Netherlands. Tel.: +31 302532909.
E-mail address: E.vanSebille@uu.nl (E. van Sebille).
[1] Present address: Department of Meteorology, University of Reading, United Kingdom.

1463-5003/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2009.05.005

trajectories with the trajectories of numerical drifters. The authors seed numerical drifters at the locations where drifting buoys are observed and then calculate the deviation of model and in-situ paths as a function of time. However, many models are non-assimilative and for these models one-to-one comparison of buoys and numerical drifters is futile as the forcing is different between the model and drifting buoy trajectories. And even if the forcing is similar, nonlinearity leads to de-coupling (or rather de-timing) between the circulation and the forcing, and therefore, an increased error between observed and modeled trajectories. Verification should be done in a statistical sense, where the distribution functions of the two kinds of drifters are compared rigorously.

Lagrangian data is often used in examinations of relative and absolute dispersion. Such estimates of dispersion would be useful in quantifying important aspects of Agulhas circulation. For example, Drijfhout et al. (2003) identify dispersion through Rossby-wave radiation as a key factor in the decay of Agulhas rings. Lacorata et al. (2001) used Lyapunov exponents to characterize the drifter paths and assess the dispersion of drifting buoys. Manning and Churchill (2006) track the spread within drifter clusters in an alternate approach to estimating dispersion.

Drifter observations used within the present Agulhas study, however, are not well distributed for these type of methods, which analyze group characteristics of among multiple pairs or clusters of simultaneously trajectories with initially small separation. Numerical simulations of drifter trajectories can be designed to support dispersion studies, but the validity of such studies requires that the simulated trajectories are representative of the true local circulation. The focus of the present study is to present a technique to assess whether the advection patterns in the model drifters agree with patterns in the real ocean. Model results that are shown to be sufficiently representative of observed characteristics could then be more credible in a subsequent study focused on dispersion characteristics.

Although drifting buoys have been deployed for over a decade now, and large numbers of buoys have been released, the total number of drifting buoys in a mesoscale region such as the Agulhas region is in the order of $10-10^2$. Numerical floats are seeded in quantities of 10^5-10^7, many orders of magnitude larger. This small number of drifting buoy trajectories limits the ability to use standard statistical tools. A common χ^2-test, for example, requires histograms with at least five members in each bin. This confines the number of bins and consequently reduces the accuracy and strength of the method. A statistical test which is better suited for this problem is the two-sample Kolmogorov–Smirnov test, which does not require binning the data.

The method is applied to a set of experiments in the Agulhas system (De Ruijter et al., 1999; Lutjeharms, 2006), where numerical floats are continuously seeded in the upstream Agulhas Current and then tracked as they move through the Agulhas region. The highly nonlinear behavior of the flow in this region, with its dynamic retroflection and mesoscale eddies, serves as an ideal test case to investigate the strengths and weaknesses of the assessment method presented here.

2. The two-sample Kolmogorov–Smirnov test

To measure the agreement between the distribution functions of the numerical drifter data set and the drifting buoy data set, the two-sample Kolmogorov–Smirnov test (2KS-test) is used (Massey, 1951). The 2KS-test is designed to test the hypothesis that two data sets B (drifting buoys) and L (Lagrangian numerical drifters) are taken from the same underlying distribution. This underlying distribution does not need to be known. The two data sets have to be one-dimensional vectors of independent and identically distributed real numbers and they may have different lengths N_B and N_L, as the 2KS-test is also powerful when $N_B \ll N_L$. The 2KS-test starts out with formulating the null-hypothesis that B and L share an underlying distribution. After that, there are four steps (Fig. 1).

First, cumulative distribution functions $F_B(x)$ and $F_L(x)$ are constructed from the data sets B and L. These functions give the fraction of data below some value of the position x. They are zero below the minimum value in the data set and one above the maximum value. At each member of the (sorted) data set they increase with $1/N$. By construction, $F(x) = 0.5$ denotes the median of the data set.

Second, a test statistic is calculated. For the 2KS-test, this test statistic is the largest distance between $F_B(x)$ and $F_L(x)$:

$$D_n = \sup_x |F_B(x) - F_L(x)| \quad (1)$$

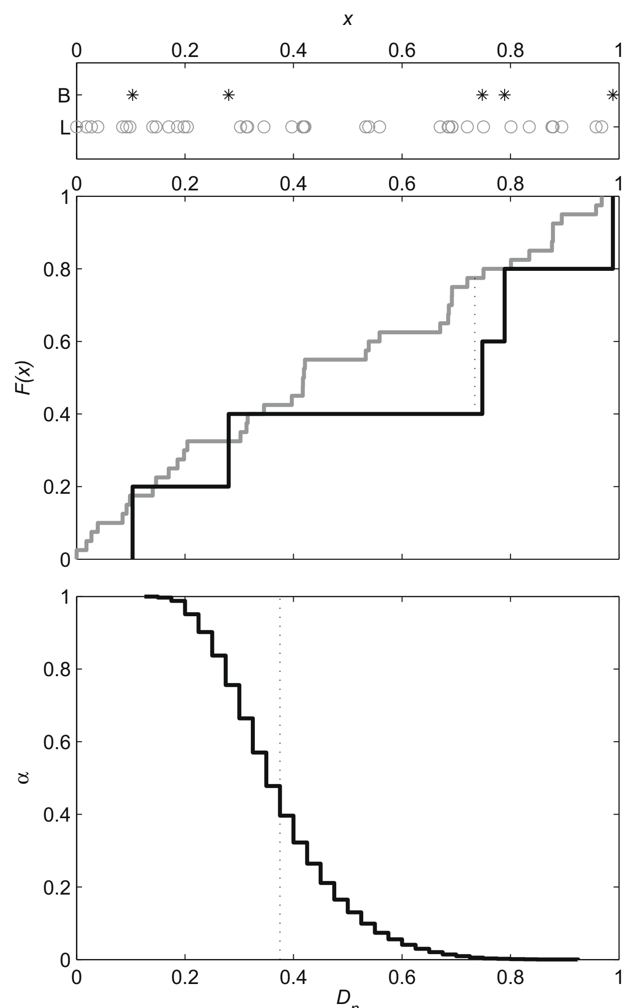

Fig. 1. An illustration of the two-sample Kolmogorov–Smirnov test. The test is performed using two random one-dimensional data sets B (asterisks) and L (circles), with $N_B = 5$ and $N_L = 40$ (upper panel), drawn from a uniform distribution. Cumulative distribution functions, the fraction of data points below some value x, have been computed from these two data sets (middle panel; black line for data set B and gray line for data set L). The test statistic D_n of Eq. (1) is denoted by the dotted line (with a value of 0.38). This test statistic is related to a confidence level α by a Monte Carlo process where D_n is calculated for 10^5 uniformly distributed data sets of similar N_B and N_L (lower panel). In this particular case the confidence level is 0.47, the value for α on the ordinate where the $D_n = 0.38$ line and the cumulative distribution function of all D_ns intersect. Since $\alpha > 0.05$, this leads to the (correct) conclusion that B and L are from the same distribution.