

We thank the referee sincerely for the time and effort expended on reviewing our manuscript. This reply addresses all points raised by the referee considering, however, also comments from the other two reviews.

In the following, the referee's comments are shown in blue.

The trajectories of 6 surface drifters are compared to simulations of circulation by two numerical models in the German Bight. For one model, the numerical simulations of drifter tracks include direct downwind slip or Stokes drift estimated from a wave model. This inclusion appears necessary to compensate insufficient vertical resolution of the model. Substantial model errors, that dominate at low winds, are explained in terms of inaccurate Eulerian currents and lacking representation of the sub-grid scales processes by the models. The limit of trajectory predictability is also addressed. This paper is clear and well written, although some parts can be substantially shortened to increase readability (see below). The scientific topic is interesting and the comparison between drifter observations and simulations is done rigorously. The results show that when using a model with reduced vertical resolution, direct windage or Stokes drift must be added in order to better predict surface drift. However, the explicit inclusion of Stokes drift does not produce an added value compared to a simple parameterization of wind-induced slip.

I recommend publication of this manuscript in Ocean Sciences after minor revision and after the authors have addressed the following specific comments.

Page 4:

Paragraph 2.1.

Please add the sampling frequency of the drifters.

Positions were monitored every 30 minutes (now mentioned in Section 2.1).

Was there a drogue presence sensor?

Drifters had no drogue pressure sensors (now mentioned in Section 2.1).

Are you sure that the drogued-drifters kept their drogue during their entire drift?

Unfortunately, we could not be sure. We address the issue in the last paragraph of the discussion. The manuscript reports indications that in particular drifters 8 and 9 may have experienced technical problems towards the end of their journey (cf. third from the last paragraph of the discussion).

What is " $R = \text{drag area in water} / \text{drag area in air}$ " for the drogued drifters?

Now added in Section 2.1: "*The ratio of drag area in the water to drag area outside the water was 33.2 for the MD03i and 16.9 for the ODi model, respectively.*"

Page 8:

Paragraph 3.1 Line 9. Change "a view day" to "a few days"

Thanks, has been corrected.

Pages 12 to 19.

Paragraph 3.2. The descriptions of the observed and simulated drifts for the different periods is too long and the reader might be bored reading all these details. I suggest to shorten these 4 pages of text by at least 50% to increase readability.

We accomplished this reduction by a removal of too many details. Another option (suggested by referee #3) would have been moving parts of the material to an appendix. However, this would destroy the clear chronological order of the description, which enables a quick scan for specific events while skipping others. It would also imply a shift of corresponding figures. These figures, however, provide relevant information, which referee #2 even suggested to expand. We followed his advice and complemented Figs. 7 and 8 by a third figure (Fig. 9) so that now 12 instead of 8 example days can be shown (we compensated for that by a removal of Fig. 6, which just combined panels from different figures in the appendix). As a result, the reader is now less often forced to switch to supplementary material.

Page 20:

Figure 9. Histograms represent the frequency of occurrence in selected classes of parameters. I would use the word "bar" instead of "histogram" to show distances versus time in Figure 9a and d.

Has been changed as suggested.