Response to Anonymous Referee #2

Below, referee comments (starting with 'Comment'), and our specific responses (starting with 'Response' and/or [envisioned] 'Change') are provided in black and blue fonts, respectively.

Comment: This study introduces a new biogeochemical model, consisting of modified versions of previously published models. Coupled with a hydrodynamic model with an improved mixing scheme, the system is validated in the German Bight region, and used to assess the impact of meteorology and river forcing on a specific flood event in 2013. The conclusion is that an interplay of the two resulted in anomalous conditions, as previously noted in observations.
The paper acts as both presentation and validation of a new modelling system, and an investigation into a specific event. While it could potentially work as separate papers, the paper is well enough written and laid out, and the validation both sufficiently comprehensive and targeted, that it works well and is an enjoyable and interesting read. I recommend publication in Biogeosciences subject to a few minor comments detailed below.
Response: We thank the referee for the careful read of the manuscript, positive assessment of our work and constructive suggestions.

The biogeochemical model appears to be a work in progress towards a different mixotroph-based model, rather than a model likely to be widely used in its present form, if I've got the correct impression? This is fine, given its structure seems sensible and plenty of validation is presented, but it would be worth adding some discussion about what sets it apart from other similar models, particularly ECOHAM, and what future developments are intended.
Response: The Referee's impression is correct, that the model presented in this study is intended to be developed further. Nevertheless, we also believe that at its present state, it can already serve the purposes of this study. We agree that further clarification and additional discussion of the model structure and future directions is needed.
Change: we will clarify the similarities and differences of our model with similar models, particularly ECOHAM in the revised manuscript, and discuss the potential directions for further model development.

Given that it's central to the study, in Section 2.1 and/or 2.2 it would be worth explicitly detailing which variables are used in the atmospheric and riverine forcing, and how they're applied to the model (e.g. bulk formulae? are rivers applied just at the surface or over the full depth?).
Change: further details on the application of meteorological and riverine forcing in the model will be provided.

"using the 'spatial.cKDTree' package from the Scipy library of Python 3.5." – add the Scipy version number for completeness.
Change: the Scipy version that will be used for the revised manuscript will be provided.

Figure 3 shows the Ems, and that this doesn't have anomalous discharge in 2013. This isn't mentioned or discussed in Section 3.1, and should be. Also, the "dashed blue lines" appear solid.
Change: we will mention which rivers were affected and which were not in the discussion, and correct the caption of Fig. 3.

“Simulated temperature and salinities . . . (Fig. 5) . . . exhibit no signs of systematic deviations or biases.” The calculated B* values are near-zero, but by eye it looks like there’s a cold bias, particularly at colder temperatures, and that salinity is usually too high. Is this just a trick of the eye, or are the simulated and observed distributions different? Please also state what B*, rho and n are in the caption of Fig. 5, as per Fig. 6.
Response: A careful assessment of the figure reveals that a slight cold bias at the lower range is indeed present, which seems to be canceled out by the slight warm bias at the higher range. But these deviations are mostly within a 1K range, therefore presumably do not have a significant effect. At an intermediate range, salinity is somewhat (in the order of 2 g/kg) overestimated, indicating insufficient spread of coastal waters with low salinity. This may either be due to (still) underestimated horizontal mixing, or inaccuracies in the advection patterns. Either way, the potentially underestimated salinity during the studied event may lead to an underestimation of the importance of riverine discharges on the stratification dynamics in the transition zone characterized by intermediate (29-32 g/kg) salinities.
Change: a more nuanced description of the model performance, and implications thereof will be provided in the discussion. Definition of B*, rho and n will be included in caption of Fig.5.

“(Fig.8) . . . The ability of the model to capture the sharp increase in DIN during June/July 2013 at the Helgoland station suggests that the spreading of the plume of the Elbe-Weser rivers following the flood event was realistically reproduced.” The model completely misses the peak earlier in the year, and also in early 2014. Can you be confident therefore that this result was obtained for the right reasons?
Response: The sentence was indeed misleading, as the word ‘suggest’ emphasizes the uncertainties of mechanisms causing the summer peak. The reasons for not reproducing the peaks in DIN during winters are not clear, but the reason for the mid-July peak as captured by the model is very likely to be the flood. Occurrence of such a high summer DIN peak at this station is not common under typical hydrological settings. In this particular case, we can tell with certainty that the reason for the model to produce such a high summer peak is the flood: the Figure R2-1 below shows how the flood water characterized by low-salinity and high DIN move within the 45 days after the flood event. These findings are consistent with the *in-situ* data shown by Voynova et al. (2017, Fig. 12), building confidence to believe that the unique DIN peak measured at Helgoland in July 2013 was caused by the flood in reality as well.
Change: the figure shown below might be too specific for the manuscript, but this finding will be more clearly described in the text.

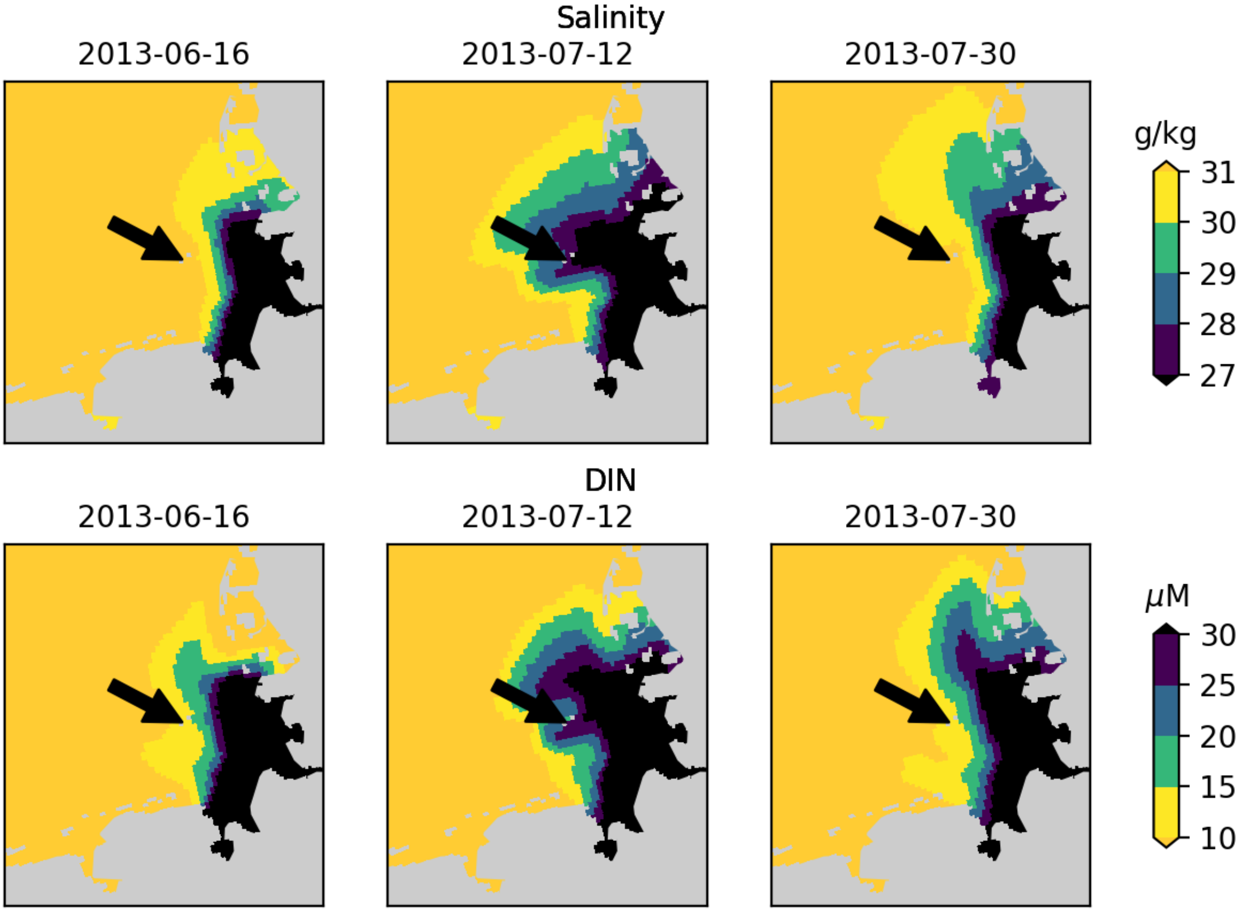

Figure R2-1: Salinity and DIN concentrations following the flood event. Arrow shows the location of monitoring station at Helgolands.

In Fig. 7, plotting the average in white is confusing – I initially thought there were separate yellow/blue lines either side of it, and the white was blank space. Plotting it in dark yellow/blue might be better. Also, make clear in the caption that the line indicates the average and the shading the standard deviation (I assume this is the case?).
Change: dark yellow/blue lines will be used to show the averages, and it will be clarified that the dark lines indicate averages and shades indicate standard deviations (indeed).

In Fig. 14 it would be best to avoid plotting red and green together, as this renders it inaccessible to those who are red-green colour blind. (Disclaimer: I'm not colour blind myself, so can't say for sure.)
Change: a colorblind-friendly palette will be used in Fig.14.

References

Voynova, Y. G., Brix, H., Petersen, W., Weigelt-Krenz, S., and Scharfe, M. (2017) Extreme Flood Impact on Estuarine and Coastal Biogeochemistry: the 2013 Elbe Flood, Biogeosciences, 14, 541–557.