

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-207-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.



review of

Anonymous Referee #1

Referee comment on "Daily hypoxia forecasting and uncertainty assessment via Bayesian mechanistic model for the Northern Gulf of Mexico" by Alexey Katin et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-207-RC1>, 2021

general comments

One important piece of information that is not mentioned explicitly enough in the introduction or the abstract is that the main part of the manuscript is about using a statistical model to generate suitable parameters for an existing mechanistic model (referred to DMO20). Reading these sections for the second time, it becomes a bit more clear but it would be beneficial to the reader to describe this more clearly early in the manuscript.

The authors are careful in creating a forecasting scenario in which no summer data is included. However, the entire set of historical data is used to produce the linear regressions, which may include future data w.r.t. year for which the forecast is produced. If I understand the approach correctly, even in the Case 4 setup, the forecast may include information from summer data of the current year, if the current year is "relevant" (as defined in the manuscript). As a result, does the forecasting system produce significantly better results for the "relevant" years compared to other years? In addition to the 4 cases currently included in the manuscript, I would suggest to add a Case 5 that excludes all data from the future (forecasts for the first few years with little data could be skipped) or, alternatively, excludes the data from the current forecast year, even if it is relevant.

Looking at the results in Fig. 2 and 3, it appears as if the forecast-hindcast as well as the forecast-observation comparisons show a pattern in August and September: The forecast appears to overestimate BWDO and HA for low values and underestimate it for high values and this pattern appears to increase in time. The authors already introduce a linear regression for the purpose of bias reduction but apply it only to the June forecast. A similar linear correction could be applied to correct this pattern which appears to increase with lead time. Yet, I am a bit hesitant to recommend such a correction because it, just like the bias reduction, adds a non-mechanistic element to the model.

How difficult would it be to extend the approach presented in this study and estimate August and September values from all data available until then? I am not suggesting that this needs to be done in this manuscript, yet creating successive 2 month forecasts appear a suitable course of action for producing more accurate estimates. This could be mentioned in the discussion.

Overall, while the manuscript is well written it sometimes overestimates the study-specific knowledge of the reader. Including more information explicitly would benefit readers. In some instances, I had to read ahead to answer questions which could have been addressed right away. I have listed some of those instances below.

specific comments

I 8: "Several models" Here it would be helpful to specify what type of model is meant, e.g. "dynamical", "statistical" etc.

I 52: The same group has previously considered different sources of uncertainty in the 3d model, finding that variations in wind forcing had the largest impact on hypoxia estimates (J. P. Mattern, K. Fennel, and M. Dowd (2013), Sensitivity and Uncertainty Analysis of Model Hypoxia Estimates for the Texas-Louisiana Shelf, Journal of Geophysical Research, doi: 10.1002/jgrc.20130).

I 77: After the first read, I am assuming that the DMO20 model has four compartments, an eastern and a western one, each divided into two layers. This could be made a bit more explicit in the text.

I 91: Is there any indication about the cause of this bias? It is nice to have an underlying parsimonious mechanistic model, yet the bias correction introduces a non-mechanistic element. By the way, it would be helpful to mention again that June is the start of the prediction interval and that the bias disappears over the course of several weeks, so that it can be neglected in the following month.

I 110: Is this done for one or multiple years? All years with data? This information is probably given later but it would be useful to mention it here already or even earlier.

I 118: Do the ten most relevant years represent the full time period accurately?

I 144: I may be good to give some examples of the model parameters that contribute to the uncertainty here, so that the reader does not need to consult the DMO20 paper to get this information.

I 150: What if one, multiple, or maybe all relevant years are in the future w.r.t. to the estimated year?

I 156: "Sixteen multiple linear regressions": I assume, the 16 refers to 4 (months) * 2 (rivers) * 2 (discharge, nitrogen loading) but this could be made a bit clearer, or a reference to Table 1 could be added here already. In my opinion, it would be good to clearly state again that there are 4 regressions for each month.

I 185: The "hypoxia model" is DMO20, correct? I would suggest to include this here again.

I 192: Is there a distinction between "forecasted" in this line and "pseudo-forecast" a few lines above? It would be good to stay consistent with the use of "pseudo". Maybe even drop the "pseudo-" prefix after describing that this is the way the word forecast is used in the context of the manuscript.

Fig. 2: Am I correct in assuming that there are 32*30 red dots in the top panels, one for every day in June in the 32 years with data? But if the monitoring cruises are typically in late July, why are there so many red dots in the bottom panels?

Fig. 4: I can only distinguish between 3 shades of gray here, yet the caption suggests

there should be 4. Are the uncertainties plotted cumulatively or is the effect of parameter uncertainty generally smaller than that of the riverine and meteorological inputs?

I 238: "Note that the relative magnitudes of the variance components are somewhat different from the relative magnitudes of the 95% IQR components ...": If the goal here is to say that the relative magnitudes differ because the variance has squared units I think it would be easier for the reader to state this directly, rather than drawing a line from IQR to standard deviation.