ESSDD

Interactive
comment

Interactive comment on "A global monthly climatology of total alkalinity: a neural network approach" *by* Daniel Broullón et al.

Anonymous Referee #2

Received and published: 18 December 2018

Recommendation: Reconsider after major revisions.

Summary: It is clear that hard and good work was done in getting this paper this far and the authors should be congratulated on their progress toward what, to me, looks like 2 papers. However, more must be done before this will be concise, clear, complete, and novel enough to warrant being broadly read by the oceanographic community.

This manuscript attempts several things:

First, it justifies the need for at TA climatology. Next it produces a new neural network for calculating TA from other seawater measurements. It then assesses the neural network and discusses likely sources of error for the method, paying special attention to riverine influences in the Arctic. Finally, it presents the climatology and compares it

[Printer-friendly version](#)

[Discussion paper](#)

to other TA climatologies in the literature.

Unfortunately, the paper uses a lot of text to only do half of the job with each of these objectives. I was left confused what use the authors had in mind for the climatology (I don't dispute that uses exist, but rather suggest that the uses were not clearly communicated). The methods used to create the neural network are similarly incomplete, where significant text is devoted to their description but not enough text is devoted to the explanation for it to make sense to people who don't already understand the material. The neural network is created, but there is only an effort to test the optimal number of neurons, and insufficient efforts are made to optimize other aspects of the NN, such as the combination of predictors used to calculate TA. The neural network assessment is incomplete (see below), and insufficient effort is also made toward comparing the new neural network to options in the literature, e.g. the Sauzude et al. CANYON reference, the Carter LIAR et al. reference (up-dated here: https://aslopubs.onlinelibrary.wiley.com/doi/full/10.1002/lom3.10232), or the recent Bittig et al. CONTENT methods (https://meetingorganizer.copernicus.org/EGU2018/EGU2018-2774.pdf). I attempted to do some of these comparisons on my end, but couldn't get the code to work. Finally, the presentation of the climatology itself is rushed and contains too many vague and general statements.

Going forward, consider splitting the paper into two, both halves of which will require more work before being ready for publication. Alternately, shift focus towards or away from algorithm development. If towards, then do a complete job of optimizing param-eters and testing the NN against alternatives. If away, then simply omit the results from this new algorithm and use existing ones or present the climatology alongside estimates from alternatives.

If split. . .

For the first paper, a more complete case must be made as to why the new methods are

better or better specifically for generating a climatology than existing methods. Most of this case can be made by showing the new method has decreased (or comparable but independent) errors to alternatives, and this can be shown by improving the validation text with a number of new quantitative comparisons employing the various methods. Randomly selecting testing/validation data is not useful since there are large systematic TA errors along hydrographic sections. This means you will always underestimate error along a section if you train your routine with data from the same section as the test data. For this reason, CANYON authors reserved entire regions for their test data and LIAR authors omitted entire sections at a time during testing. This may or may not end up making the computed RMSE worse, but it is an important step regardless so readers know what to expect from the algorithm when they deploy it in areas where there weren't nearby dense measurements collected at the time of the estimate. It wasn't clear what omitting data with a >3 umol/kg bias for the 2nd NN training was intended to show. Uncertainty is necessary for these estimates.

For the second paper, creating a climatology from an algorithm is fast. The second paper will be complete when the authors have answered the following questions: 0. Why is a TA climatology needed? 1. What does climatological TA distribution look like? 2. What processes make it look that way? 3. What does TA variability look like? 4. What processes make it vary like that? 5. How large are the uncertainties in the climatological values, and how does this uncertainty vary regionally and with depth? And finally, 6. How do we know the answers to these questions? The current paper begins to answer all of these questions, but ultimately falls back on too many qualitative and vague statements. It therefore ends up neither concise nor complete.

Specific comments:

L40. "The capacity of the ocean..." this sentence doesn't make sense. Are you suggesting the atmospheric pCO2 would today be 520 ppm without the ocean CO2 storage? This estimate is incorrect if so.

L51: This definition sounds closer to the Revelle factor definition.

L53: "Processes that change salinity..." it would be better to name those processes since one can imagine processes that change salinity without changing TA.

L61: Hydrothermal TA inputs should perhaps also be mentioned.

L66: "Therefore, the knowledge of AT variability over the global oceans at monthly timescales is very useful to increase the understanding of the ocean carbon cycle and to make assessments and projections related to ocean acidification with greater rigor." Build on this. What applications specifically do you have in mind for this climatology? Why use the climatology instead of an algorithm?

L88: Why? Why is it necessary to have a seasonal climatology of subsurface TA? How deep does seasonality affect TA, and how do you know this? See: line 190.

L103-L122: This is in an unhappy medium of detail... too little detail to make any sense, and too much for the reader to quickly or confidently skip this. Either reference a paper that explains the method or fully explain it. My preference would be to fully explain the method, but move the text to a supplement so it doesn't interrupt the paper with too much detail... more detail in the text or a supplement would be necessary for the first paper if you split the paper into 2.

L140: Did you include calculated TA? TA where GLODAPv2 did not QC the data?

L159: "We kept..." I don't know what this sentence means.

L162: What is the difference between testing and validation data sets? It's possible I missed the explanatory text, but consider trying to make that a bit clearer.

L168: Explain your rationale here. It is unclear why this test would find places where the network is unable to obtain accurate values.

What does it mean to for an individual data point to have a RMSE of >3?

By my best guess, this is saying that the version of the NN that includes the data with a >3 absolute offset does better at fitting the data with a "less than 3" absolute offset than the version that only includes the "less than 3" data? The RMSE of the data with a "less than 3" offset, by definition, must be less than 3, and yet it climbs to 5.1 when you omit this data... so why bother with this analysis? Why not simply say "if we omit data with large errors our RMSE becomes small." Which do you recommend users adopt? Where and why?

L171: What does it mean to "illuminate the complexity" of a neural network? Be more specific.

L208: Random division of the datasets is inadequate for your test. See main points.

L210: How does "The samples with residuals beyond ±3RMSE are 1% of the global dataset..." align with "99% of the GLODAPv2 dataset used was modelled by the network with a root-mean-squared error (RMSE) of 5.1 μmol kg-1." I'm guessing this is referring to the 2nd NN, but I was confused for a long time before this statement started to make sense.

L228: This sentence doesn't make sense to me... I suspect "disengage" is an incorrect word, but I'm not sure.

L235: It's not always riverine TA that is the problem... often it is rivers with little or no TA that dilute seawater TA in a way that is distinct from the mixing patterns in the open ocean.

L235: this paragraph presents a weak argument against removing the region... fortunately, the argument for omitting that region was not made... omit this text, but instead estimate uncertainty regionally more quantitatively.

L251: "The zones defined.." this sentence is vague.

L252: considerably

L261: what is meant by this sentence? Clarify.

l269: what is meant by this sentence? Clarify.

L273: S and T collectively provide information regarding interior ocean density and mixing patterns, which are important for predicting TA... it is not clear what you are suggesting about the link between T and CaCO3.

L278: "The bias is relatively low in the three time-series with the highest number of data. The AT computed by the NN at KNOT and K2 is higher than the measured one. Summed to the previous test, this independent test 280 with a seasonal time resolution shows the good generalization of the NN." The first two sentences here are difficult to understand and the last one does not make sense to me.

L286: This logic does not follow.

L302: This sentence does not seem to fit with the rest of the paragraph.

L310-L313: I don't understand these sentences.

L310: intra-annual would be clearer as "seasonal"

L340-342: I don't understand the logic.

L354: Isn't that figure 10?

L363: Why a change of grid for a time series?

L380: what is meant by a "continuity" in the differences?

L393: I couldn't get the code to work... some info... âĂć Tried with the directory with NN as the active directory o As well as just with the directory with NN on the Matlab path âĂć Tried entering within a script and on the command line âĂć Tried with Matlab r2014b with the NN toolbox (also 2018b, but without the NN toolbox) âĂć With inputs as single or double precision numbers âĂć Entered: o AT_values=Neural_network_object(data_inputs); o AT_values=NN(data_inputs); o

AT_values=NN_w3RMSE(data_inputs); Invariably got the response: "Undefined function 'NNw3RMSE'/'NN'/'Neural_network_object' for input arguments of type 'single'/'double'."

I recommend making the instructions a bit more clear. I also recommend adding an example calculation so users can be sure they are getting the expected answers. You'll know you are done when the coauthors can use the function reliably without additional guidelines.

Figure 7 and elsewhere: Uppercase theta should be reserved for conservative temperature... use lowercase

Figure 8 and elsewhere: The font is too small.

Figure 10, right panel: How is the RMSE smaller than the bias? Clarify what is being shown here if there is a good reason.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-111, 2018.