Ocean Sci. Discuss.,
doi:10.5194/os-2016-73-AC2, 2017
© Author(s) 2017. CC-BY 3.0 License.

Ocean Science
Discussions

Interactive comment on "Technical note: Evaluation of three machine learning models for surface ocean CO_2 mapping" *by* Jiye Zeng et al.

Jiye Zeng et al.

zeng@nies.go.jp

Received and published: 2 March 2017

We thank referee#2 for the thoughtful comments, especially on the appendix. We have revised the appendix substantially to address the reviser's constructive opinions. Here we documented our responses to the reviewer's comments point-by-point.

Q1. This "technical note" discusses the formation of global maps of surface ocean CO2 from limited measurements using inferred dependence on (latitude, surface temperature SST, salinity, chlorophyll concentration, mixed-layer depth, difference between monthly- and annual-mean SST). The dependence is inferred by three methods: selforganisation map (SOM), feedforward neural network (FNN) and a new method (support vector machine; SVM). The results of these three methods, "trained" on a fraction of the data, are compared with the remaining data. The correlations are not particularly good

for any (best at R2 = 0.715 for SVM) considering there are 6 independent variables aiding the fit. However, the results of all three methods for global air-sea CO2 flux are very close and the CO2 maps are visually similar. This similarity extends to a band of high CO2 concentration in February 2005 extending west from Chile where there are apparently no CO2 measurements. This extrapolation from CO2 observations is presumably via a similar feature in (at least) one of the 6 independent variables. There should be more discussion: (i) of the quality of the fit to observed data, especially in relation to the estimates of air-sea flux and the danger that the methods agree with each other more than with reality; (ii) of the extrapolation feature west of Chile (in particular – perhaps also a careful examination for whether there are others) and whether it can be believed in terms of the values of the independent variables – is this set of six values closely approximated somewhere else where there are CO2 measurements constraining the CO2 estimate?

Reply: First on the quality of fit. As the three model are unbiased, i,e, the mean difference between modeled CO2 and observation is statistically zero, we take the quality here means the correlation and the standard deviation between modeled CO2 and observation. Then quality is not only determined by the capability of the models, but also the variability of the data (we added this information in the revised manuscript). Second on the relation to the estimates of air-sea flux. The models may produce differently higher than observed CO2 in some areas and lower CO2 in others. The flux could be largely affect by this and by different wind. That all three models produced similar global fluxes indicates that the effect is small. Third on the danger that the methods agree with each other more than with reality. They are quite different issues. The answer to the quality of fit indicates the models cannot agree with reality than the variability of the reality. Whereas, the agreement between models is determined mainly by their similarity. For example, SVM is considered to be a one layer FNN in some articles; so FNN agrees better with SVM than with SOM. Fourth on the extrapolation. Yes, the extrapolation approximate unmeasured area with somewhere that has a similar biogeochemical property and CO2 measurement.

Q2. Although the organisation and English are generally good, I think some sections and especially the Appendix are unclear/obscure, mainly due to inconsistent or missing explanations, definitions or notation. Most of the following detailed comments are about this aspect.

Reply: We have revise the appendix substantially to address to issue.

Q3. Page 2, lines 12 and 18. "dSST denotes the difference between the monthly and annual means of SST" implies 12 discrete values of dSST; how does this "improve expressing the seasonal variable continuously"?

Reply: Let's consider three measurements taken on January 1, January 31, and February 1. Using month as the seasonal variable, the variable values of the first two measurements are 1 and the last is 2. However, the seasonally the last two are nearly the same. dSST reflects better the actual change of seawater property caused by season change.

Q4. Page 3, Line 13. I think you mean ". . to the range (0, 1) for the SVM . . ."

Reply: We revised the expression accordingly.

Q5. Page 3,. Line 21 (i.e. line after (5)). Why between 0.1 and 0.9 not between 0 and 1? "better" compared with what? Why should scaling the output help?

Reply: For fCO2 close to 0 and 1, and a small change in fCO2 requires very large adjustment of model parameters, which slows down the convergence of training. We added this in the revised manuscript.

Q6. Page 4, Lines 1-2. "We used Eq. (4) to scale . . SOM". There is no mention of this in Appendix A.1, indeed after (A1) it is stated that the diagonal factors of the scale matrix f are equal to 1.

Reply: In the appendix section for SOM, we added "In our application, the data for each input variable were scaled to be unitless by its mean and standard deviation".

Q7. Page 4, Lines 2-3. "Based on preliminary studies, we applied a factor of 2 to . . SST and CHL . .". What preliminary studies? Is this subjective, i.e. why should SST and CHL be emphasised?

Reply: SOM is indeed subjective. In our knowledge, other applications scaled the data with different subjective factors to change the impact of independent variables on the distance defined by Eq.(A1). In our application, we scaled the data non-subjectively and uses the scale factors to change the impact, which in our opinion is easier to understand. Because CO2 shows a much higher correlation with SST and CHL than with others, we subjectively used a factor of 2 for them. There is no theoretical basis for this choice. We revised the manuscript to address the issue and revised "Based on preliminary studies" to "Based on our preliminary correlation analysis"

Q8. Page 4, Line 7. "prediction for an input" needs explaining. Inputs are supposed to be known, not "predicted".

Reply: We revised "Making prediction for an input" to "Making a CO2 prediction for an input"

Q9. Page 4, Lines 8, 9. "map size". In normal language the map size is the earth's surface area. Do you mean resolution, equivalent to the number of CO2 output locations? Please explain / use correct word.

Reply: We revised "the feature map size" to "the number of neuron cells"

Q10. Page 5, Line 8. "respectively" should be "for SOM"

Reply: We corrected the mistake.

Q11. Page 5, Lines 11, 15. Please explain "normalized"/"normalization"

Reply: We revised the manuscript and explained "normalized"/"normalization".

Q12. Page 6. To have value, this needs to be understood in its own terms; the reader should not have to refer to cited references to understand the words used and the

overall meaning. Too many words are not defined or explained. Also, it is too abstract. This is a manuscript about “output” CO2, depending on “inputs” LAT, SST, SSS, CHL, MLD, dSST. Presumably this applies to A.1, A.2 and A.3 – say so and do not use vague terms like “feature space” – at present the reader has to guess what you mean.

Reply: We have revised the appendix substantially according to address the concern.

Q13. (A.1 . .) Page 6 Line 23. What is “feature space” in oceanographic terms?

Reply: We removed the jargon.

Q14. Lines 23-24. “usually represented by grid points in two dimensional space”. Never mind about “usually”; describe in terms of the problem here.

Reply: We revised the expression to be specific.

Q15. Line 24. “weight vector w”. This name is confusing.

Reply: We changed the symbol and revised the descriptions.

Q16. page 7, lines 7-8 weights (weight factors) h are defined by (A3). “w” is the result of applying the weights “h” to combine values of “v” at various locations [presumably to represent “v” at grid locations rather than original locations, but this is not clear to the reader. If this the case, then “w” is “gridded v” or “interpolated v”]. See also the comment on page 7 line 21. Line 25. Not “a data vector” which might refer to any vector at all, but “an input data vector” (I guess).

Rely: We changed the symbol and revised the descriptions.

Q17. page 7, Line 30. “best matching cell (BMC)” needs explaining.

Reply: We revised the descriptions.

Q18. page 7, Line 30. “minimizing the distance”. What is varied to do this? Reply: We revised the descriptions.

Q19. Page 7 Line 4. “matched”. Either this is the wrong word or it needs explaining.

Reply: We revised the descriptions.

Q20. Line 17. “vector x of input data”. In A.1 the input data were “v”. Use consistent names for variables.

Reply: We revised the descriptions.

Q21. Lines 20-22. You have input data, hidden neurons and output. There should be distinct variable names for each of these, e.g. v, x, y respectively. Here you have y for the hidden neurons and for the output, which is confusing.

Reply: We changed the symbols and revised the descriptions.

Q22. Line 21. “w is the weight vector”. Indeed this seems correct for its use in (A4) but that is very different from its use in (A1). Use different terms for different quantities (c.f. comment on page 6 line 24).

Reply: We changed the symbols and revised the descriptions.

Q23. Line 22. “The training updates the offset and weight parameters”. What are the starting values before updating? Do you mean “weight vector” as in line 21?

Reply: We revised the descriptions. The parameters are initialized randomly between -1 and 1. We added this information in the revised manuscript.

Q24. Line 23. What is “e” or is it defined by (A5)? Please make this clear.

Reply: Yes. It is the “e” defined b (A5).

Q25. Line 24. “modelled . . y” is unclear (especially because you use “y” for hidden neurons and output). Why are two “y” in this line in bold type but not the third or “y” in (A4)?

Reply: We revised the description. Bold font indicate vector or matrix.

Q26. Line 24. “w includes both . .” This seems to be defining a vector with more components; it should have a new name.

Reply: We used a new name.

Q27. Line 28. " is the learning rate". How is its value decided?

Reply: The initial value is about 0.25. It is determined by try-and-error. A small value make training slow. A large value make a training diverge. We added the information in the revised manuscript.

Q28. Line 30. "derivatives of e by w". Do you mean "derivatives of e with respect to w".

Reply: We revised the description.

Q29. Page 8, Lines 6-10. "The SVM . . SVM parameters." Is this relevant?

Reply: We removed this part.

Q30. Line 14. "independent variables", "high dimensional space", "target variable". Please define these in terms of the oceanographic problem in question.

Reply: We removed these jargons.

Q31. Line 16. "minimizes" – what is varied to do this?

Reply: We revised the description.

Q32. Lines 18-19. "subjecting to the constraint". (A11) looks like a definition of "e" and is not a constraint unless "e" is defined in some other way which needs to be stated.

Reply: We revised the description and re-arranged the equation.

Q33. Line 27. Can there be an explicit expression for '? Where has "b" in (A9) gone to? Table 1. SOM column half way down. "closest" not "closet"!

Reply: We corrected the mistake.

Q34. Figure 3 caption. Please explain "normalized to 2005".

Reply: We revised "normalized to 2005" and in the section 4 added that fCO2 means

trend-removed fCO2 unless specified otherwise.

Interactive comment on Ocean Sci. Discuss., doi:10.5194/os-2016-73, 2016.