Response to reviewers – second review

Comments by reviewer Anna Denvil-Sommer

The manuscript is refined compared to the original version and all of my comments were properly addressed. This work is a useful contribution to the field because this is the first time I have seen the regional importance of predictors addressed in a rigorous way.
The text could still use some minor improvements, mainly in the introduction. I have provided some detailed suggestions below. Once the text has been modified, then I think the manuscript is ready for publication
Suggestions:

- Line 36: consider removing “have been thought to”.

Response*: The words was now removed.*

- Line 38-44: Please rewrite this sentence for clarity. Maybe something like this: “However, the air-sea CO_2 flux averaged between 2001-2015 varies from -1.55 and -1.74 PgC/yr, depending on the pCO_2 product. These differences largely stem from differences in pCO_2 estimates across the products.”

Response*: Thank you for the suggestion. The sentence was modified as “However, the global ocean sea-air CO_2 flux averaged between 2001-2015 varies from -1.55 to -1.74 PgC yr^{-1} with the maximum difference in individual years nearly 0.6 PgC yr^{-1}, depending on the surface ocean partial pressure of CO_2 (pCO_2) product. These differences largely stem from differences in pCO_2 estimates across the products.”*

- Line 45-47: Consider changing sentence to: “Surface water pCO_2 greater than the overlying air indicates CO_2 is released from the ocean to the air. Conversely, absorption of CO_2 by oceans happened when the pCO_2 of the surface water is lower than the overlying air”

Response*: Thank you for the suggestion. The sentence was changed as the suggestion.*

- Line 48: need a comma after “sink”.

Response*: The comma was added.*

- Line 48: sentence starting with “Sparse and uneven…” could be a new paragraph.

Response*: The text starting with this sentence was divided into a new paragraph.*

- Line 56: Consider changing “Recent researches on” to “Advances in”

Response*: The words was replaced.*

Line 62: consider removing “methods such as” and changing “and other” to “with”
Response*: The sentence was modified as “In addition, finding better predictors or combining SOM with other neural networks were also attempt to further decrease the pCO_2 predicting error”.*

- Line 68 “ were considered as” to “are considered”

Line 74: “appeared” to “appears”
Response*: The typos were corrected.*

- Line 77: Consider writing sentence as “In addition, sampling information, such as latitude and longitude, has been used as a predictor.” Also, consider referencing Gregor et al. 2019

Response*: The sentence was modified as “In addition, sampling information, such as latitude and longitude and sampling time, has been used as a predictor.” The reference was added.*

- Lines 111-137: I would consider removing this text and moving Table S1 to the main text. Tables are much easier to read.

Response*: Thank you for the suggestion. the Table S1 was moved to the main text.*

- Line 145: Landschützer et al. (2020)

Response*: The reference was corrected.*

- Line 148: consider rephrasing “Provinces with connected pixels less than 10 and provinces with SOCAT observation less than 1000…” to “Provinces with less than 10 pixels and less than 1000 SOCAT observations…”

Response*: The sentence was rephrased as the suggestion.*

- Line 153: “The” needs to be lower case

Line 156: “200 m”
Line 167: comma after “summation”.
Line 170: “a tan-sigmoid”
Response*: The typos were corrected.*

- Line 173: Maybe remove “pure” from “pure linear function”, I am not sure what that means.

Response*: The word was removed.*

- Line 177: Please state how sensitive the model is to the choice of random number. That random number defines where the NN starts “searching” in

errors space. I understand the random number was fixed for reproducibility purposes. However, my understanding of best practice is to either run the model many times with different random numbers and take the average or "tune" the random number to find the one that gives best results. Please mention if this was explored or if the random number was simply chosen randomly.

Response*: Thank you for the suggestion. The fixed random number was currently chosen randomly. When using different random number streams, several predictors at the end of the output list of the stepwise FFNN algorithm differed, but the leading predictors were consistent, and the different predictors were also related. The fixed random number makes all networks using different predictors start training from the same point at the error space when comparing the performance of each predictor. In the future version, the method of taking the average will be used after we completely excluded the influence of initial state of FFNN on the results, and make sure the results would not change no matter which random number was used. Also, the algorithm may take much more time when taking the average to compare the perform of each predictor.*

- Line 224: "after that,"
- Line 226: "Then,"

Response*: The comma was added.*

- Line 231-235: Please consider rephrasing, this sentence is hard to understand.

Response*: The sentence was modified as "The part 1, including Loop 1, Selection step and Determine step 1 in Fig. 2, was repeated until no indicator was left in the Indicators pool or no decrease of E_0 can be found no matter which two indicators were added in the next two steps."*

- Line 236: "At this time, part 1 … finished and …"

Response*: The place of comma was corrected.*

- Line 259: In what increments where neurons increased by?

Response*: The neurons was set as [5,10,15,20,25,30,35,40,45,50,60,70,80,90,100,150,200,250,300]. The sentence was modified as "The number of neurons increased from 5 to 300 (the increment was five during 5-50 and ten during 50-100 and fifty during 100-300) and the corresponding MAE values of each size were recorded, and then the number of neurons with the lowest MAE was applied."*

- Line 290: It is also worth noting pCO_2 at BATS and HOT are estimated from TA and DIC (I am not sure if ESTOC is estimated or directly measured). I think it's important for the reader to know these estimates are derived and not directly measured like SOCAT observations.

Response: *Thank you for the suggestion. The sentence was modified as "The pCO_2 at HOT and BAT were estimated from TA and DIC, and pCO_2 at ESTOC were directly measured. These observations were not included in the SOCAT dataset."*

- Line 505: Wouldn't modifying the structure of the FFNN be considered making a "better network"? By better network do you mean more sophisticated architectures? I would also add that different learning algorithms could be considered.

Response: *Thank you for the suggestion. The sentence was modified as "A possible way to improve the performance of the stepwise FFNN algorithm is to modify the structure of FFNN or to use networks with more sophisticated architecture and to use different learning algorithms."*

- Table 4: Please state what the bold values indicate. Also a comma is needed before respectively: "…Landschützer et al., 2014 and Denvil-Sommer et al., 2019, respectively" This typo is repeated throughout the article

Response: *The lowest MAE and RMSE between different validation groups was shown in bold. The description was added and the typos was corrected.*

Figure 3: Make it clear these provinces are from a SOM. This is just a suggestion, but it would be great to get some of the information in Table 4 into this map. Maybe putting the leading predictor in parentheses in the colorbar? This might look too messy, it's just a suggestion.

Response: *The title of Figure 3 was modified as "The map of biogeochemical provinces based on SOM". It looks messy after adding the leading predictor, so Figure 3 was not changed.*

- Figure 6: Comma before "respectively"

Figure 7: Comma needed: "…Landschützer et al., 2014 and Denvil-Sommer et al., 2019, respectively

Response: *The typos were corrected.*

Figure 8: please put units on the colorbar.

Response: *The units were added.*

Comments by reviewer Lucas Gloege

I would like to thank the authors for the changes they made. I am appreciated that the authors took into account my previous comments. The manuscript is more clear and completed now.

Comments:

- Lines 38-44: New sentence is too long with repetitions, please modify it: “However, due to large uncertainty in estimates of surface ocean partial pressure of CO2 (pCO2), the long-term average global ocean sea-air CO2 flux during 2001-2015 estimated based on sea-air pCO2 difference differ from -1.55 to -1.74 PgC yr $^{-1}$, and the maximum difference between global sea-air CO2 flux in individual years reached nearly 0.6 PgC yr $^{-1}$ (Rödenbeck et al., 2014;Iida et al., 2015; Landschützer et al., 2014; Denvil-Sommer et al., 2019)”. It can be something like: However, due to large uncertainty in estimates of surface ocean partial pressure of CO2 (pCO2), the long-term average global ocean sea-air CO2 flux during 2001-2015 differs from -1.55 to -1.74 PgC yr $^{-1}$ with the maximum difference in individual years nearly 0.6 PgC yr $^{-1}$ (Rödenbeck et al., 2014; Iida et al., 2015; Landschützer et al., 2014; Denvil-Sommer et al., 2019).”

Response*: The sentence was modified as “However, the global ocean sea-air CO_2 flux averaged between 2001-2015 varies from -1.55 to -1.74 PgC yr^{-1} with the maximum difference in individual years nearly 0.6 PgC yr^{-1}, depending on the surface ocean partial pressure of CO_2 (pCO_2) product. These differences largely stem from differences in pCO_2 estimates across the products. (Rödenbeck et al., 2014; Iida et al., 2015; Landschützer et al., 2014; Denvil-Sommer et al., 2019).”*

- Line 123: Need a space between “Interim” and “(Dee et al…”

Response*: This part was changed to a table.*

- Line 128: You use here a word “parameters” to call “predictors” or “indicators”. It appears further in the text as well (for example, in Conclusion, line 498). I think it can be misunderstood as parameters can be used to explain FFNN architecture. Please keep using words “predictors” or “indicators” throughout the text.

Response*: Thank you for the suggestion. The word “parameter” was all changed to “predictor”.*

- Line 138: Please indicate which method you used for downscaling.

Response*: The products were downscaling by taking the average of all values in each 1° × 1° grid. This description was added in the text.*

- Lines 184-187: Please change this sentence, it is difficult to read: “The mean absolute error (MAE) difference that before and after adding or removing one indicator in the input of FFNN calculated using a K-fold cross validation method was used to estimate the performance of each indicator in the FFNN predicating.”

Response*: The sentence was modified as “The mean absolute error (MAE), calculated using a K-fold cross validation method, was used to estimate the*

performance of each predictor in the FFNN predicating."

- Lines 318-320: Please combine these two sentences, it sounds like a repetition: "For example, month was considered as a recommended predictor in most provinces. Especially in the province P4 subpolar Atlantic and P5 north subtropical Atlantic, the parameter month was relatively more recommended."

Response*: The sentence was modified as "For example, the predictor month was considered recommended in most provinces, especially P4 subpolar Atlantic and P5 north subtropical Atlantic."*

- Line 322: Please add "the sine of latitude".

Line 404: Remove point between "Table" and "4".

Response*: Thank you for the suggestion. The typos were corrected.*