General Comments:

The manuscript introduces a complete and novel carbonate chemistry observation-based dataset, that will be useful to detect changes in the carbonate system. The pCO_2 dataset compares well to previous similar datasets, but the extrapolation method used here incorporates new and improved features, and the alkalinity observation-based dataset is novel. The pCO_2 and TA can be used together to calculate other important parameters such as DIC and pH.
The GRaCER method is similar to previously used techniques, but introduces an ensemble of cluster-regressions that solve the boundary problem between different regions. The authors do an excellent job at calculating possible dataset errors and biases and do an extensive analysis of it. The method and the resulting dataset are an important contribution to the scientific community. However, I believe there are some improvements that need to be made to the manuscript before being accepted for publication. The authors need to address some possible issues with the clustering method, and whether the dataset is suitable for analysis on time-scales longer than the seasonal cycle. Below I detail these points and some technical corrections.

Specific Comments:

Methods section:

I believe the novelty of the GRaCER method is that produces an ensemble of clusters. The ensemble is produced because the clustering process randomly assigns the first cluster center in the predictor-variable space. Thus, each member of the ensemble has a different center, and therefore the ensemble mean does not have discrete boundaries. Based on this, it would be helpful to clarify a couple of details:

- Lines 172-179: How does the result vary, if you use monthly data instead of climatologies for the clustering process?

- Lines 245-248: “It may seem tautological to use other machine learning estimates, but these data are just used to create regional clusters, i.e., they are not used in the regression step. ” How do the results vary if you do not use the previous machine learning estimates for clustering? Previous methods (Landschutzer, Rodenbeck etc), only use the observations from SOCAT for clustering. In this manuscript these datasets are not used in the regression step, but I believe that the results are affected by which data is used for the clustering.

- For the non-expert reader it would be helpful to know how is the "bias" measured. This word is repeated through the manuscript but is not clear how it is calculated. I suggest a short explanation in the appendix.

Discussion section:

The title of the paper indicates that this dataset can be used for studies from seasonal to decadal periods, but the authors do not analyze or compare the trend and the inter-annual and decadal characteristics of this dataset with respect to SOCAT or other observation-based datasets.
A section should be added in the discussion to address whether the dataset is suitable for studies on time-scales longer than seasonal, to answer "How well represented is the inter annual and decadal variability?". Similar observation-based datasets may not be suitable for interannual variability; for example the Rodenbeck dataset indicates that "*interannual variations* may miss important features" (see http://www.bgc-jena.mpg.de/CarboScope/?ID=oc "period of validity" and "usage notes"). Also, see for example Gallego et al., 2019 in which is discussed the lower interannual variability of the pCO_2 dataset of Landschutzer et al. 2018, when compared to the CMIP5 models. In a pre-print Gloege et al., 2020 show that the neural-network estimates are not as skillful for inter annual variability as for seasonal cycle estimations (see https://www.essoar.org/doi/10.1002/essoar.10502036.1).
If the authors consider that their dataset is good to estimate internal and decadal variability, then they should add some analysis and comparison with other existent pCO_2 observation-based datasets.

Moreover, some methodologies used in the manuscript should be discussed in the context of inter annual to decadal variability, such as:

- The MLD sub-annual and inter annual variability is removed.

- To calculate DIC and pH they use climatologies of silicic acts and phosphate instead of monthly data.

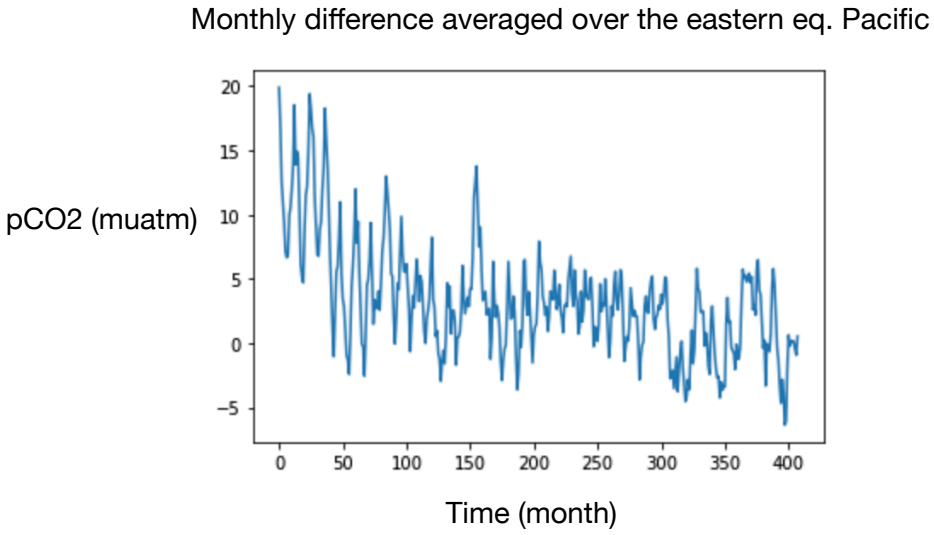
- A pCO_2 RMSE of 12 muatm is larger than inter annual variability for many locations.

- Lines 412-415: "The highest biases on pCO_2 when comparing with observations are located in the eastern tropical Pacific where inter annual variability is higher". This may suggest that the dataset does not represent well inter annual variability.

- The cluster seems to collect data by climatological month. How does he method change if monthly data is used instead?

Results section:

Figure 8 (d) shows the mean of the monthly differences between the SOMFFN and GRaCER pCO2 datasets. When I plot the time-series of the monthly differences between SOMFFN and GRaCER averaged over the eastern equatorial Pacific, I see a continuous decrease in the difference from 1985-2018. That could mean that for the beginning of the time period there is a larger difference between the two methods than by the end. Could this be a consequence of using the SOMFFN and other products for the clustering process? Since at the beginning of the period there is less observational points compared to the end.



Technical corrections / comments:

Figure 2: Which data is shown in these plots? This is an interesting illustrative figure, but in line 172 of the paper it says "we use monthly climatological data of pCO_2 and TA and related parameters (Figure 2a-c)…" But the caption on Figure 2 does not specify which data (TA or pCO2) is shown, and from (d) to (f) which data are used to do the regression.

Line 176: The cluster center seems to be what determines that clusters and therefore it is possible to create an ensemble of them. However, it is not clear for the non-expert reader what is "the cluster center" or how it is defined. Is it one arbitrary data point?

Line 256: The removing of sub-annual and inter annual MLD variability, seem that it could cause an underestimation of the inter annual variability of pCO_2.

Figure 3: Where it says (a,c) should say (a,d), and where it says (b,d) should say (b,e). Also in the caption it would be helpful for the reader to indicate what are panels (c,f).

Line 387: I suggest to explain further the sentence "Further, the mean seasonal cycle well is relatively well represented at HOT and BATS, being within one standard deviation of the interannual variability (Figure 4b,f). ".

- It seems to me that the mean climatology cannot be one standard deviation from the inter annual variability.

Figure 7 caption: Suggest "further the breakdown of the errors is proportional to the contribution of the sum of the squares (see Eq. (2))"**.**

Figure 9 caption: Perhaps would be helpful to also indicate in the caption that the maps show the 1985-2018 mean, since the word climatology would suggest we are looking at a monthly means.

Line 40 and 50: Show "Omega" instead of Ω.

Line 143: At the end there is an extra ")"

Caption Figure 1, 7th line: Should say "our" instead of or?

Lines 231-232: References need to be inside brackets.

Line 253: It says "the the"

Line 343: instead of "$1x1^o$ by month" should say "$1^o x 1^o$ monthly grid"

Lines 346-347: I suggest a rewording of ".And with an order of magnitude fewer observations of TA, an even larger number of cells are populated by a single observation. "

Caption Figure 4: Change for "The top row (a-d) shows data for the Bermuda Ocean Time Series (BATS), the middle row (e-h) for the Hawaii Ocean Time-series (HOT), and the bottom row (i, l) shows the Irminger station."

Line 387: Remove "well" from "the mean seasonal cycle well is relatively well".

Caption Table 4: It would be good to clarify which product is LDEO (Takahashi).

Line 456: Correct the sentence "However, the estimates are not as coherent for pH where there the bottom-up error is 23% smaller top-down error than the in the open ocean. "

Line 466: Suggest changing "in the spatial distribution of the distribution, " to "in the spatial pattern of the distribution"

Caption Figure 10: Hovmoeller plots (*a-d*) should say (a-e).

Line 493: Change "update" for "uptake"

Line 498: instead of 10,c,d,g,h should be 10,c,d,h,i

Line 500: Should be CO_2 instead of CO2 and 10c,g should be 10,g,h

Line 507: Suggest to change to: "This would also be true for Omega_calc which only differs from Omega_arag in …"

Line 546: Suggest to change w.r.t. for the meaning (with respect to?) since the abbreviation seems unnecessary.

Line 553: For the read it would help to clarify what it means the representation error of the horizontal representation errors.

Figure 11 caption: Last line should say The map (d).

Line 579: I suggest adding the reference of Gallego et al. 2018 (Drivers of future seasonal cycle changes in oceanic pCO_2) and Fassbender et al., 2018 (Seasonal Asymmetry in the Evolution of Surface Ocean pCO_2 and pH Thermodynamic Drivers and the Influence on Sea-Air CO_2 Flux).

Section A3.4: Seems to be empty.

Line 655: Remove one “every”