We thank the reviewers very much for their constructive comments. Major revisions were done as required, and a detailed response to the reviewer comments, that carefully addresses, point-by-point, the issues raised in the comments, is provided. We hope that you will find the changes satisfactory and that this revised manuscript will be now considered for recommendation in PCI Ecotoxicology & Environmental Chemistry. Please note that an application of this methodology has been accepted for publication (https://doi.org/10.1016/j.cscee.2021.100172) and is now referred in the paper. We are at your disposal if you need any further information. Thank you very much in advance for your attention.

Best regards,

Rémi Servien also on behalf of co-authors.

Reviewed by Sylvain Bart :

Servien et al presents a new method based on machine learning to predict ecotoxicological metrics for chemicals for which we don't have these metrics. The approach is promising and complementary to the linear QSAR method which cannot deal with nonlinearity.

The graphical abstract is very informative and the introduction provides all the necessary information to understand the topic and the scientific gap addressed. All the methods and procedures are deeply described which is very appreciated for reader whom machine learning is not the primary expertise, like me.

In conclusion, the manuscript is well written, I don't see any major issue in the manuscript, and I would recommended it for publication in a peer reviewed journal.

We thank the reviewer for this recommendation.

minor comment:

-I would suggest to carefully check all figure captions to ensure all necessary informations are given for the figures to be read by themselves. E.g. : Figure 4, Provide full name somewhere for RF, PLS etc.. ?

These figure captions have been checked and the information added.

Reviewed by Patrice Couture :

I would not provide an in-depth review of this manuscript, due to my very limited expertise in the area of the paper (I am an ecotoxicologist). This paper needs to be properly reviewed by experts in modeling. I only identified a few points that would need to be addressed to improve the clarity and the relevance of ecotoxicological terms like LC50 (see file attached).

I consider that the topic addressed in this paper is interesting and the approach proposed is promising. Overall, this work has the potential to provide very useful tools for environmental and human risk assessment of new chemicals that will reduce costs, time and use of live organisms.

We thank the reviewer for these remarks and following corrections have been made according to the points raised in the attached file.

Page 2 Line 11 : There seems to be a statement missing before this sentence. Otherwise replace the word "Then" by the appropriate term

« Then » was replaced by the more appropriate « Finally ».

Page 3 Line 9 : This reads like the USEPA list is an example of EU regulation...

We replace « As an example » with « Furthermore » to avoid misunderstanding.

Page 3 Line 11 : « (eco)-toxicological » what does this spelling mean? If you mean ecotoxicological, then write it correctly in full. You may instead mean ecological data and toxicological data? Then on line 8 you spell it differently again. This is important to clarify, as you argue that we need a certain type of data to make informed decisions.

This spelling was used for toxicological and ecotoxicological data. As we agree this could be confusing we replace it on lines 8 and 11 with toxicological and ecotoxicological.

Page 3 Line 16 : LCA : Define at first use

We correct this.

Page 4 Line 24 : This is not the definition of EC50. Mortality refers to LC50.

The definition of EC50 was corrected. It is now the effective concentration required to have a 50% effect.

Page 5 Line 7 : the number of existing and ever-increasing numbers : Rephrase

We rephrase this sentence : *it is still extremely difficult to test existing chemicals due to their large and ever-increasing number.*

Page 6 Line 21 : Use past tense to refer to your study. Change throughout.

These modifications were maded throughout the mannuscript.

<u>Reviewed by Dominique Lamonica</u> :

Here is my review of the paper "Machine learning models based on molecular descriptors to predict human and environmental toxicological factors in continental freshwater". I would like to highlight that I am specialised in ecological modelling (development and analysis of IBM, state space models, ODE based models) and Bayesian statistics. My fields of research are movement ecology, forest ecology and ecotoxicology. I am not a specialist of: chemistry, toxicology, risk assessment, machine learning.

This study aims aims at predicting characterization factors of chemicals based on molecular descriptors using statistical models. Those predictions would complement experimental approaches in order to decrease experimental costs. The intersection of USEtox database comprising CFs and TyPol database comprising molecular descriptors made available the variables and the predictors for 274 chemicals. The authors tested three different models predicting CFs from molecular descriptors, one linear model (PLS) and two machine learning based models (RF and SVM). They also tested whether a clustering step of the chemicals before applying the different statistical models results in better predictive performance. Therefore six models were tested in total. The choice of the "best model" relied on the absolute error between the prediction and the true value. Then missing CFs (for some chemicals of the database those were

not available) were predicted with the "best model". Also, the five variables (ie molecular descriptors) that contributed the most to the predictions were identified. Overall, the clustering step improved the prediction performance for one variable and absolute errors were smaller with the machine learning models than the linear model for both. The "best models" lead to acceptable predictions. Overall the paper is well written and easy to understand. It is actually a useful study for the ecotoxicology community, since it shows that CFs can be predicted from the molecular descriptors stored in TyPol database with an acceptable error, using a rather easy method (clustering then machine learning regression models, or only machine learning regression models, depending on the predicted variable).

Here is my general comments on the paper sections, a more detailed list of modifications I suggest follows. The introduction is clear and exposes well the context, motivation and interest of the study. The method section is clear enough, except for a few paragraphs. The result section could benefit from changes in the structure and in the choice of figures. Indeed I think that the main results are not clearly displayed, and are therefore rather difficult to get at first. The discussion section is clear, however, it seems that a result, namely the identification of the five most explicative variables in the models, is not discussed. I get that it can be uneasy to do so - this is well highlighted in the discussion - but I think it might be useful for the readers to have more insight in the authors' opinion on this specific result (note that I am not a specialist in the field of chemistry or toxicology).

We thank the reviewer for her careful reading and we improved the manuscript thanks to her following remarks.

Title
I wonder if it is really necessary to specify "in continental freshwater".

We strongly believe that it is necessary to add this to the title. Indeed, the whole methodology (predicting CF based on molecular descriptors) could also be applied to any other compartment but it could be burdened by a lack of data (for the soil for example). The developed models are shown to be valid only for the continental freshwater compartment and, so, we think it should be mentioned in the title.

Materials and methods
p10 l.24: "Split each cluster", I guess you considered that, in the case of the "global" models, there is one cluster including all the chemicals ? Maybe you could specify, for instance by moving there the phrase "(the whole dataset [...] a cluster-then-predict model)" which is currently p11 l.25-26.

We move the phrase as requested by the reviewer.

p10 l.24: Is there a reason/reference for choosing those percentages of training and test dataset ?

This choice depends on the dataset. Around 80/20 is a common choice (Pareto principle). As some of our clusters have a small number of data, we decided to increase the number of data in the training set. Some quick tests have been performed and it seems that 80/20 or 85/15 did not affect the quality of our models.

p11 l.17: It seems to me that this paragraph, which ends p12 l.7, is not part of the "comparison procedure" section (2.5). I suggest to start a new section 2.6 Predictions, for instance.

We agree and add this new subsection 2.6.

P11 l.24: I do not get how many repetitions you use to compute the 95% prediction interval, by "leave-one-out bootstrap" do you mean that you compute the prediction n times (each time without one of the chemicals), n being the number of chemicals for which there is a CF value in the cluster ?

Yes, that is what we did. To be clearer we add the following details in brackets: « *A prediction is carried for each leave-one-out model (i.e. n-1 models if n is the number of compounds of the, eventually global, cluster)* … ».

Results
p12 l.14-19: I suggest to move this paragraph to the Materials and Methods section, after the two first subsections describing the databases.

We understand the suggestion of the reviewer, as a first draft of the manuscript was organized as proposed. Nevertheless, we chose to put this paragraph in the Results section because this paragraph (and the figure) are rather results (even if they are only descriptive ones) than M&M. So we prefer to keep this paragraph and the dedicated Figure in the Results section.

I also suggest to move Figures 1 and 2 to Supplementary material and add the lines 7 to 9 p14 as a part of the legend, or a comment.

We agree that the same information is present in these two Figures and that they do not need to be both in the main text. But we believe that this is important to have one of these figures on the main text to deliver the following message: the compounds in common in USEtox and TyPol are more dangerous compounds (with high CF) and they cover the whole order of magnitude of the CFs of the USEtox® database. So we chose to move the CF_{ET} Figure in the supplementary material but to keep the other (and the comments) in the main text.

I suggest to start p15 l.1 as a first subsection of the Results (a title could be "clustering"), and Figure S2 could be moved to the main text, as the entire paragraph develops on clustering.

As suggested by the reviewer, we chose to add a new subsection (« Clustering of the compounds ») to help the reading of the Results section. We think that the addition of Figure S2 in the main text would not bring any information not already contained in the text so we chose to keep it in the supplementary material.

Sections 3.2 and 3.3: I found that the chosen structure does not highlight the results enough. I suggest to reorganise those in two sections focusing first on the model comparison and second on the performance and predictions of the "best model". Also, having figure S5 and the equivalent figure for CFHT in the main text would help visualise and support one of the paper statements, namely the "best model" shows good performances.

We agree with the reviewer: for a methodological paper, the proposed sections are probably more appropriate than the previous ones (focus on CF_{ET} with models and predictions and then on CF_{HT}). So we reorganize the paper thanks to comments with Section 3.3 « Performances of the machine learning methods » and Section 3.4 « Best model predictions ».
We also add the equivalent figure of S5 for the CF_{ET} in the supplementary material.
To illustrate the performances of the best model we chose to keep these figures in the supplementary material but to highlight the performances of the best models (with bold) in

Table S3 and S6 and to add comments in the text that seems more interesting to give a quick overview of the results than Figures S5 and S6.

Similarly, since you have assumed (and it is supported by references) that CF can vary by 2-3 orders of log-magnitude (p11 l.14 and p22 l.14) it would be interesting, again for better visualisation of the results, to highlight that value in Figures 4 and 5.

To help visualization we add a horizontal line at 1 log (that gives an interval around the real value of 2 logs) and the explanation in the caption.

The result stated in the discussion p22 l.17 only appears there, it should be moved to the results section.

This result was moved to the beginning of Section 3.4.1 for CF_{ET} and 3.4.2 for CF_{HT}.

In general, it would be useful to write down in the text some quantiles, not only the medians, of the distributions of absolute errors. I would also find interesting to display a table (which could be in the supplementary material) that sums up Figures 4 and 5, with median and quantiles of the absolute error for the 6 models for each CFs and each cluster. I do not find Figure S6 very useful.

We add in the Supplementary Material Tables S3 and S6 to summarize the values contained in the boxplots of Figures 4 and 5, for each method and each cluster. We also add some median values at the beginning of Sections 3.4.1 and 3.4.2. We try to avoid adding other quantile values in the text in a sack of compactness. We agree that Figure S6 (and the corresponding Figure S4 for CF_{ET}) was not very useful, so, to prevent uninteresting information in Supplementary Material, we decided to remove it.

Discussion and conclusion
p23 l.2 "the usual ones": I guess you mean the approaches without the clustering step, I would rather write it like that than "usual".

This modification was made.

p23 l.2 "local": I do not get what you mean by "local" in that context, could you specify ?

Local is used here to refer to the cluster-then-predict approaches. These models are local as they are specific to a small « area » (as clusters could be seen as areas in Figure S2) of the dataset, they are adapted to a small local neighborhood of the global space of the whole dataset. We agree that this needs to be specified so we add « *local (i.e. cluster-then-predict)* ». We think it worth mentioning this other wording for this kind of approach because this is often used in some communities (as in the title of the paper of Metz et al., 2020).

p24 l.10 "a new modelling method": I would not call the method you describe in the paper "new".

The adjective « *new* » was not here because the modeling method was thought to be new but because the context of the application of these well-known methods was. We agree with the reviewer that this was confusing and we remove this word.

More generally, I found that the molecular descriptors that were identified as the "most important" are not discussed, although those are highlighted in the results section (p19

l.10-23 and p21 l.17 to p22 l.7) and in two tables in the main text. Similarly, the clustering result is not discussed either. For those two results, I suggest that you try to deepen the interpretation, or you shorten the corresponding paragraphs (and move the tables to Supplementary material) in the results section.

We agree with the remarks of the referee. As the deep study of all the best models is not the core of the paper and would require a lot of development, we thought that this table could give a quick overview of the important molecular descriptors and highlight that they are different between the clusters/models. Now, we think that a deep analysis of all the models (as also requested by the following reviewer) would be of course very interesting but could blur the main messages of the paper and, thus, is considered out of the scope of this paper. So, the two tables were moved in Supplemental Material and we let two sentences for CF_{ET} (« *To compare the different models in each cluster and give an idea of what are the important molecular descriptors we provide the five most important molecular descriptors for each cluster in the Table S4. We could see in this Table that the important molecular descriptors strongly differ from one cluster to another.* ») and one for CF_{HT} in this subject (« *The most important descriptors of these two models are gathered in the Table S7 and, as for CF_{ET}, are strongly different between the different best models.* »).

Figures
The colour palette of the boxplot figures is not colour blind accessible, that would be good to change for another colour palette (like the viridis one for instance).

Thanks to the reviewer's remark, we change the color palette of the boxplot figures to different shades of blue to be color blind accessible.

Figures 4 and 5: it would be useful to have the number of replicates (I guess it is the 200 repetitions of the algorithm) in the legend.

We add the number of repetitions in the caption.

Table List of the molecular descriptors: "Number of hydrogen atoms" is mentioned twice (1st and 2nd rows).

This error was corrected.

There is an issue in numbering, there are two Table 1. It seems there is a mistake in the caption of Table 1 (the second) and 2 "The most important descriptors are in the first line of the table" should rather be "The best model [...]".

The tables and the figures have been checked and corrected. This sentence in the caption was to explain that the most important descriptors of each model were at the top of the table. To avoid misunderstanding we rephrase this sentence : « *Descriptors are listed from top to bottom in decreasing order of importance.* »

R script
It would be better to have the comments in English rather than in French. Also it is not very easy to quickly get how it is structured, so if you could separate the different steps of the analysis and put explicit titles, that would be great.

The comments were translated into English. The titles of each part of the code have been modified to be more explicit. We hope these modifications make the code easy to follow.

<u>Reviewed by anonymous reviewer :</u>

The paper frames itself in a line of research initiated by other researchers and pursued also by the same authors in previous works, i.e. the use of machine learning to predict human and environmental toxicity of chemicals (using the USETox database, but not only). The application described in this paper is just another confirmation of the potential of this kind of approach.

The paper is rather well written, although it appears too concise in the description of the full path of modelling that was followed. In this sense, to facilitate the understanding of the model chain, I suggest inserting a clear flowchart or a figure like Fig. 1 in Hou et al. 2020 (Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. Environment International, 135, 105393) or Fig. 1 in Marvuglia et al. 2013 (Machine learning for toxicity characterization of organic chemical emissions using USEtox database: learning the structure of the input space. Environment International 83: 72-85).

We agree that the paper could benefit from the insertion of a clear flowchart of the modeling process. Thus, the following figure was added to the paper.

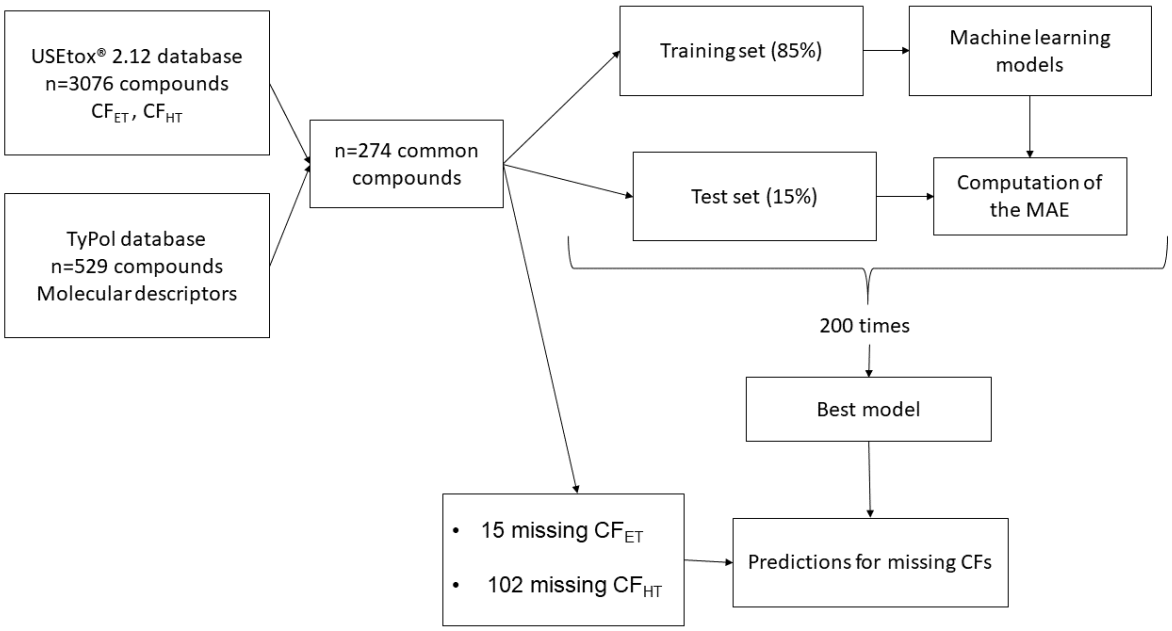

Figure 1 Schematic representation of the modeling procedure adopted in the paper.

Besides these two articles, other exist on similar applications in the literature, that have not been cited in this manuscript. They authors might want to take a look at them to improve their state of the art:

- ***Marvuglia et al. 2014. Variables selection for ecotoxicity and human toxicity characterization using Gamma Test. In: B. Murgante et al. (Eds.): ICCSA 2014, Part III, LNCS 8581, pp. 640–652, 2014. Proceedings of the 14th International Conference on Computational Science and Applications (ICCSA 2014), University of Minho, Guimaraes, Portugal.***
- ***Marvuglia et al. 2015. Random Forest for toxicity of chemical emissions: features selection and uncertainty quantification. Journal of Environmental Accounting and Management 3(3): 229-241;***

- Song et al. 2017. Rapid Life-Cycle Impact Screening Using Artificial Neural Networks. Environ. Sci. Technol. 2017, 51, 10777−10785.
- Wu and Wang 2018. Machine Learning Based Toxicity Prediction: From Chemical Structural Description to Transcriptome Analysis. Int. J. Mol. Sci. 2018, 19, 2358; doi:10.3390/ijms19082358.
- Lysenko et al 2018. An integrative machine learning approach for prediction of toxicity-related drug safety. https://doi.org/10.26508/lsa.201800098.

- Song et al. 2021. Accelerating the pace of ecotoxicological assessment using artificial intelligence. Ambio. https://doi.org/10.1007/s13280-021-01598-8

We thank the referee for these references. All the references were added to the manuscript. The corresponding paragraph is now: "*Recently, machine learning algorithms have been used to predict some midpoints based on molecular descriptors and environmental parameters (Marvuglia et al., 2014 and 2015; Song et al., 2017; Lysenko et al 2018) and a first review on this subject could be found in Wu and Wang (2018). After these first works, predictions of hazardous concentration 50% (HC50) based on 14 physicochemical characteristics (Hou et al., 2020a) or on 691 more various variables (Hou et al., 2020b) were carried out. Nevertheless, their input variables need some experiments and could be difficult to collect. This problem was tackled by Song et al. (2021) who predicted Lethal Concentration 50 (LC50) based on 2000 easy-to-obtain molecular descriptors.*"

At page 11, when the clustering protocol is described, it is not clear to me how the clustering is chosen. The authors mention that the whole algorithm is repeated 200 times. However, this is not a deterministic procedure and at each iteration a (slightly or not) different partitioning can come up. Therefore, a criterion of cluster quality is needed. For example, in hierarchical clustering, not always the cut height that determines how many clusters to choose, is clear. If I understand correctly, the error criterion that the authors use, pertain only to the evaluation of the forecasting capacity of the models to determine the two factors CF_{ET} and CF_{HT}, but nothing is said on how to chose the best clustering partition. There are many cluster validity measures (see e.g. Vazirgiannis M. (2009) Clustering Validity. In: LIU L., ÖZSU M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_616).

We agree with the referee that the concatenation of the sentences could be confusing. As already explained in the previous Section 2.4, the clustering is made only one time on the whole TyPol database. It allows us to determine the clusters for the following analyses. The clustering is not repeated during the comparison procedure. It could have been done and would have some benefits (explore a wider range of models) and some drawbacks (difficulty to assess the performances in the clusters as the clusters are changing from one iteration to another ; clustering performed only on the common molecules, not on the whole TyPol database). To clarify that we have removed the sentences on the Typol clustering after the definition of the 5 steps of the procedure (as all is already explained in Section 2.4) and add the sentence on cluster 5 in Section 2.4.

At page 11, line 18, the term NA appears but it is not explained in the paper. It is only explained in the caption of table S2 in supporting information. I think it should also be explained in the text of the paper.

The explanation has been added: « NA (Not Available, *i.e.* missing) ».

At page 21, line 2-3 read as follows: "We could see in this Table that the important molecular descriptors strongly differ from one cluster to another, highlighting the

usefulness of the cluster-then-predict approaches". This is true, but the important molecular descriptors (and the ranking of the descriptors overall) differ not only because we change from one cluster to another, but also because the best model changes from one cluster to the other. Therefore, how can we say that the important descriptors change only because of the cluster? To estimate how much of this change in ranking depends on the cluster and how much on the model used, the authors should provide the full ranking in each cluster for each model. Then one could calculate for example the change in raking position for each variable within the same cluster when passing from one model to the other.

In table 2, it is not clear how the descriptors are selected. Is it possible to add the % of variance of the output explained by each descriptor?

We agree with the remarks of the referee. As the deep study of all the best models is not the core of the paper and would require a lot of development, we thought that this table could give a quick overview of the important molecular descriptors and highlight that they are different between the clusters/models. Now, we think that a deep analysis of all the models (for example comparing the ranking of the molecular descriptors of each model in each cluster, so 12 cluster-then-predict models and 3 globals for each CF) would be of course very interesting but could blur the main messages of the paper and, thus, is considered out of the scope of this paper. So, and as requested by the previous reviewer, the two tables were moved in Supplemental Material and we let two sentences for CF_{ET} (« *To compare the different models in each cluster and give an idea of what are the important molecular descriptors we provide the five most important molecular descriptors for each cluster in the Table S4. We could see in this Table that the important molecular descriptors strongly differ from one cluster to another.* ») and one for CF_{HT} in this subject (« *The most important descriptors of these two models are gathered in the Table S7 and, as for CF_{ET}, are strongly different between the different best models.* »).

At page 24, the lines from 6 to 11 of the Conclusions are more fit for the introduction, rather than for the conclusions. I suggest moving this part there.

The sentences and the reference to the study of Aemig et al. (2021) were moved in the Introduction as requested.

Suggested changes to the text:

- ***Page 3, line 11: begin the sentence with "therefore" rather than with "so".***

- ***Page 3, lines 23-24 from "To best" to "case-by-case basis": this sounds like a repetition of something already mentioned above.***

- ***Page 5, line 8: change "That's why" with "That is why".***

- ***Page 6, line 28: change "that are" with "that is".***

- ***Page 9, line 26: add a comma after "performs well".***

- ***Page 10, line 18: correct "cluster-the-SVM" in "cluster-then-SVM".***

- ***Page 17, line 11: change "in each cluster" to "from one cluster to another". The meaning changes, and I think my suggestion reflects better what you want to say.***

- ***Page 17, line 13: begin the sentence with "therefore" rather than with "so".***

- ***Page 20, line 16: change "the more difficult" with "the most difficult".***

- ***Page 21, line 8: change "lonely" with "single".***

- ***Page 21, line 10: change "the more important" with "the most important".***

- ***Page 23, line 4: although also the cited paper (Lesnoff et al., 2020) uses the term "explicative", I believe a more common term in statistics and machine learning is "explanatory".***

We thank the reviewer for his/her very careful reading. All these errors were corrected.

Reviewed by anonymous reviewer :

The first impression reading the paper is that it contains some naïf considerations. The authors insist on the novelty of using non linear methods; those methods are in use since about 20 years, both in QSAR and many more modeling tasks. Using a non-linear method is the good practice today when simple linear methods fail.

We agree with the reviewer that non-linear methods are more common nowadays (in many applications) and our point was not to explain that we were the first to apply non-linear methods. Nevertheless, many models (QSAR or not) remain linear that is why we chose to compare linear and non-linear methods and their performances and assess if a linear model could be enough in this modeling task. We also hope that the incorporation of the references suggested by the previous reviewer on other machine learning uses helps to clarify this point.

So the novelty of the paper is not in choosing tools that are already accepted in QSAR ; it can be in the idea of computing the characterization factors (CFs) using molecular descriptors instead of relying on the traditional LCA methods that depend on data (chemical, toxicological, etc.) not easily available for every chemical.

Yes, we agree, that it is the main idea of our paper.

The authors compute 40 molecular descriptors (including some quantum chemical descriptors), selected since they appear relevant to describe the behavior of organic compounds in the environment. Then they apply both classifiers (using 3 modeling methods) and clustering, defining different local models for the 5 different clusters. A point that should need more attention is the descriptor selection. In any modeling method (machine learning included) the features are important and a wider exploration of the features and their number is missing in the paper.

We agree with the reviewer that the choice of the molecular descriptors is very important. The choice of the 40 descriptors included in this study has been made previously based on a literature review (Mamy L**,** Patureau D, Barriuso E, Bedos C, Bessac F, Louchart X, Martin-Laurent F, Miège C, Benoit P, 2015. Prediction of the fate of organic compounds in the environment from their molecular properties: A review, Critical Reviews in Environmental Science and Technology, 45:1277-1377. http://doi.org/10.1080/10643389.2014.955627) and also detailed in Servien et al. (2014). This review allowed the determination of the molecular descriptors that were best correlated to seven environmental parameters (water solubility (Sw) and octanol-water partition coefficient (Kow) were selected to describe dissolution; vapor pressure (Pvap) and Henry's law constant (KH) for volatilization; adsorption coefficient normalized to soil carbon organic content (Koc) for adsorption; half-life (DT50) for degradation; and bioconcentration factor (BCF) for ecotoxicity). Then, this choice was proved to be consistent to explain several complex environmental processes (Benoit et al. 2017, Traoré et

al. 2018). By consequence we keep these 40 molecular descriptors as inputs for our models to predict the CFs in this paper and, to provide more details for an interested reader on this choice, we add the reference of the review that was indeed missing.
Note that no variable selection procedure is applied in any of our models so each model combines the same number of features (40) with the more important provided in Table S4 and S7. Another approach could have been to take as input several thousands of molecular descriptors and to apply sparse modeling approach to have a selection on the inputs. This strategy is now studied by a PhD student in a slightly different context and the first tests seem to assess that the choice of the 40 molecular descriptors remains consistent.

The combination of the classifiers with clustering is interesting in that the results can be more accepted by the users, which often like to consider also the compounds similar to the one under investigation.

We thank the reviewer for this remark and we add this argument in the conclusion.

As the authors report, USEtox® is commonly used; it provides in one single CF the chemical fate, the exposure, and the effect for each compound in a set of several thousands chemicals. Then the CF can be extended to other endpoints, both human and environmental (DALY and PDF). The observation that the computation of those final endpoints can be done in one model using directly the chemical information is the advantage of the proposed method over the traditional one. In conclusion, even though the methods applied are quite common in QSAR, and the machine learning methods should be better applied, the paper proposes something new in the LCA domain.

We thank the reviewer for his careful reading and this positive comment.