Responses to referee comment [RC2]

Interactive comment on "Chilling accumulation in temperate fruit trees in Spain under climate change" by Alfredo Rodríguez et al.

We thank the reviewer for her thoughtful comments. Our answers are highlighted in green italics.

Alfredo Rodríguez et al. did an extensive and rigorous job on trying to quantify future developments on chilling accumulations for Peninsular Spain and the Balearic islands. They did a major effort in modelling and validation of input data and consider a highly relevant aspect of local fruit production that is vulnerable to climate change (Campoy et al., 2011; Luedeling, 2012). In this sense, and in my opinion, this regional study has its relevance and its place in this journal. This study does also contribute to a better understanding in this domain, by improving the methodology with regards to previous studies through the use of state of the art climate models and scenarios, although it does not stand out for the novelty of the used approaches. To increase the value, that the paper brings to the scientific community as well as to end users, a couple of revisions are suggested below, which, if taken into account, would make this paper more suitable for publication.

We appreciate the referee's comments.

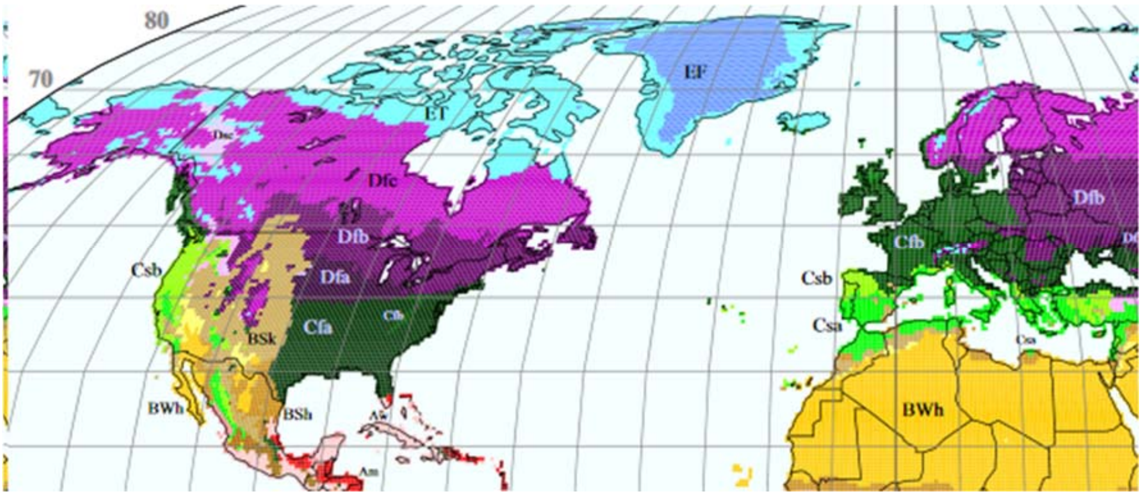
Major remarks regarding the content

With regards to the methodology and scope of the paper, I agree in most points with Eike Luedelings review comment (RC1):

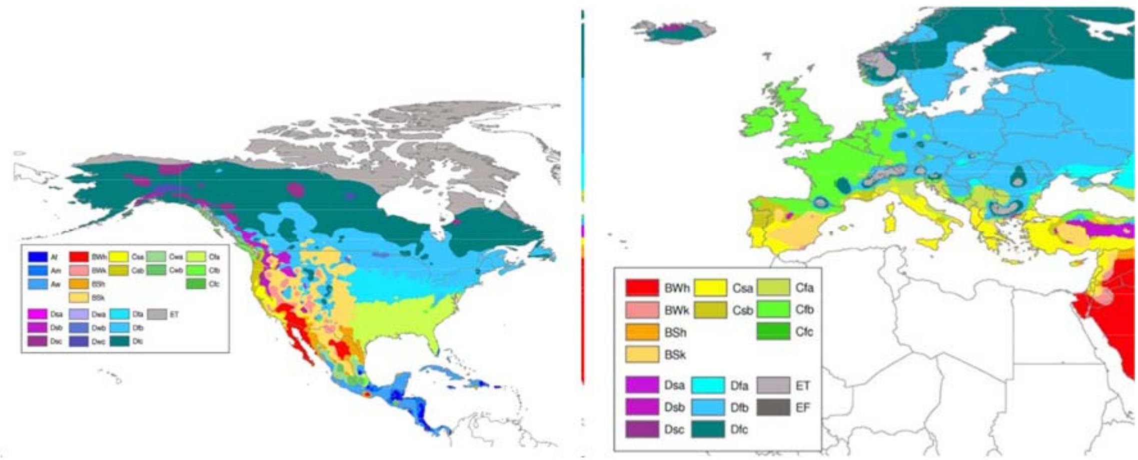
(1) First of all, combining models that have been found to be inadequate (Luedeling, 2012) is not innovative, and the fact that the models were apparently applied without calibration to local conditions is in my eyes the biggest shortcoming of the paper. To my knowledge, there is no evidence that a model that was tested for North Carolina (Latitude range 36.5N-33.8N, Köppen-Geiger classification 'Warm temperate with hot summer climate' (Peel et al., 2007)), can be transferred to Spain (Latitude range 43.5, 36.0, major Köppen-Geiger classification 'Arid steppe cold' climate (Peel et al., 2007)); nor can be safely assumed, that the cultivars in all regions have the same physiology, which is implied by using the same model, despite the mention of this fact on p.3, l. 11.

- *About the comment on combining models, please see Answer#2 to referee 1*
- *About the comment on applying models to other countries without additional adjustment, please see Answer# 4 (and also 2 and to 3) to referee 1, and the references included there.*

Additionally, we have checked the locations used to develop North Carolina model (Shaltout and Unrath, 1983). The paper says "A chill unit model was developed for 'Starkrimson Delicious' (Malus domestica Borkh.) apples grown under the ***wide range of temperature and elevations in North Carolina****", and the locations considered were Wake, Cleveland, Wilkes, Mitchell, Henderson. Looking at the climate at these locations we can see that parts of Spain (northern Spain) actually share the temperature regime with them, which can be checked by looking at the second subindex of the Köppen-Geiger climate classification (a or b, denoting temperature regime, "hot summer" and "warm summer" respectively; see images below from http://koeppen-geiger.vu-wien.ac.at/pdf/kottek_et_al_2006_A4.pdf, and Peel et al., 2007), which are the same for North Carolina and the northern Spain.*



Source: [*http://koeppen-geiger.vu-wien.ac.at/pdf/kottek_et_al_2006_A4.pdf*](http://koeppen-geiger.vu-wien.ac.at/pdf/kottek_et_al_2006_A4.pdf)



Source: Peel et al., (2007)

That is why in answer#4 to referee 1 we proposed: In our case, the main driver is temperature regime; and actually, in the case of North Carolina model for *apples, main production area is North Spain, with climatological characteristics (temperature) more similar to North Carolina than the Spanish average. Accordingly, we will delimit more the concrete area of the apple tree production in the introduction section.*

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633-1644, 10.5194/hess-11-1633-2007, 2007.

Shaltout, A. D., and Unrath, C. R.: Rest completion prediction model for 'Starkrimson Delicious' apples, J. Amer. Soc. Hort. Sci., 108, 957-961, 1983.

(2) At this point of the introduction, a better contextualization and reference for the obtained values would be highly appreciated. Only on p.3, ll. 14-17, an exemplary chilling requirement is given, and this for apricot which is not considered in this study. Without a knowledge of local requirements of apple, olive and vineyard, the severity of the change in chilling units is hard to grasp. Also, with the quoted requirements at hand ("631 chill units [Utah model, 'Palsteyn' variety), the observed difference between models ("less than 500 chill units", p.7 l. 25) can be substantial, and the outcome of Figures 7-8 more alarming than described in the paper. Later, on p.10, ll. 11- 17, exemplary requirements for an apple and an olive variety is given, which are at risk of not being fulfilled according to the 'far future' predictions. For better understanding of the key findings of the paper, more such values should be given.

- *To follow the referee's comment, we will add a table with values for different species showing the range of chilling requirement exhibited by the main varieties. Also, we will remove any comparative comments between methods, following referee's 1 indications.*
- *About the message from Figure 7-8: Even if the impacts are high as pointed by the referee, we wanted to stress that the wide range of chilling requirements exhibited by the varieties of a given species will facilitate adaptation. In most locations, variety change will be enough, and crop change will not be required. We will modify the text to clarify this, using the new table to illustrate it.*

(3) In my opinion, estimations of concrete, crop or variety related shortcomings in chilling have highest relevance for planning applications and various end users, so if this is possible, it would be very interesting to find in this paper indications which zones under cultivation of a given crop will become unsuitable in terms of chilling for major varieties.

We will add some examples in the discussion for the highest vulnerable varieties or those with strongest market influence. We think that going into a deep analysis is beyond the scope of this paper because 1) We would need to set thresholds. As *illustrated by the new table, for a given crop and even just considering the most important varieties we should consider very different thresholds to estimate suitability area; 2) we would need to consider other measurements than the mean, (as the Safe Winter by Luedeling et al. (2009) and use a confidence/robustness index that allow to support any recommendation or conclusion on suitability area.*

Luedeling, E., Zhang, M., and Girvetz, E. H.: Climatic Changes Lead to Declining Winter Chill for Fruit and Nut Trees in California during 1950–2099, PLOS ONE, 4, e6166, 10.1371/journal.pone.0006166, 2009.

(4) Obviously, the diversity of species cannot be fully covered in this paper, but, joining the suggestion of RC1, with open source code and output maps, interested parties could quickly assess these zones following an example. It might be a subject of discussion in this stage of the paper, if these findings would be improved or not by considering the agreement of different chilling models. A priori, there is a major concern with this methodology, that I share with the author of RC1, because of the unjustified comparison of chill units among models and the mentioned inadequacy of some of them.

- *About including open source code in supplementary material, we have agreed to do so, see Answer#8 to referee 1.*
- *About the comment on combining or comparing models, please see Answer#2 to referee 1. In further work, we propose a methodology to assess robustness of the individual model outcomes (as they cannot be put together due to have different units), the EOA index (Rodríguez et al., 2019). Then, the agreement between chill models on the suitable zones for a given variety (as suggested by this referee in her question 2) could be compared by using the EOA values. This development, however, is out of the scope of this paper. We will clarify this in the text.*

Rodríguez, A., Ruiz-Ramos, M., Palosuo, T., Carter, T. R., Fronzek, S., Lorite, I. J., Ferrise, R., Pirttioja, N., Bindi, M., Baranowski, P., Buis, S., Cammarano, D., Chen, Y., Dumont, B., Ewert, F., Gaiser, T., Hlavinka, P., Hoffmann, H., Höhn, J. G., Jurecka, F., Kersebaum, K. C., Krzyszczak, J., Lana, M., Mechiche-Alami, A., Minet, J., Montesino, M., Nendel, C., Porter, J. R., Ruget, F., Semenov, M. A., Steinmetz, Z., Stratonovitch, P., Supit, I., Tao, F., Trnka, M., de Wit, A., and Rötter, R. P.: Implications of crop model ensemble size and composition for estimates of adaptation effects and agreement of recommendations, Agric. For. Meteorol., 264, 351-362, https://doi.org/10.1016/j.agrformet.2018.09.018, 2019.

(5) Potentially reducing the number of models and increasing the documentation (equations, parameters) of the models should help overcome the, in my view, given uncertainty about how the different models can be understood considering the three studied fruit crops mentioned in the paper. In p.5 ll.22-24, the North Carolina model is introduced as being developed for apple trees, the De Melo-Abreu method for olive trees and the Dynamic method for peach trees. I would wish for more elaboration on how these choices have been justified and on how to make use of the findings presented in the Figures 3-8. The codes should be open access, too, since I totally agree with RC1, a research should be reproducible and with the given information this is not of application.

- *About the models used in the study and dealing with uncertainty, please see Answer#2 and 3 to referee 1. Also, a brief history of each model origin, development and applications (references suggested in the Answers#2 and 3 will be included).*
- *About including open source code in supplementary material, we have agreed to do so, see Answer#8 to referee 1.*

Remarks regarding the form

Title

In p.2, ll.1-22, the authors state "Vineyard, apricot trees, olive trees and almond trees could be also included in this last subgroup [of temperate fruits], although some of their climatic requirements are nearer the subtropical fruit trees" p.2,ll 6-7). Bearing this in mind, the mention 'temperate fruit trees' in the title of the paper is in my opinion a bit misleading, although reference handbooks do classify olives and grape as temperate (Schaffer, 2018).

We will remove the word "temperate" from the title, as follows:

"Chilling accumulation in fruit trees in Spain under climate change"

Abstract

The abstract could be more concise and feature more detail about the findings of this study than the context.

We will modify it making it more focused on results.

Introduction

In p.2, ll.1-22: In line with RC1, I consider the description of the classification as too long and can be left out, especially in view of the ambiguity of the classification mentioned above. The section on bias adjustment (p.4 ll.1-9) could be slightly more elaborated, and precise how it is ensured that the change over time of the climate signal is not cancelled out, see also Michelangeli et al. (2009). The transition from this paragraph to the following is a bit sharp. At this point, an overview of similar (regional) studies on chilling requirements would be expected point.

We will remove the classification part. The bias adjustment section will be extended with some details on the methodology, according to the following guidelines:

Bias adjustment is based on a transfer function such that the marginal cumulative distribution function of the adjusted variable matches that of the observations. A complete discussion of the technique, including validation and effect on climate indices can be found in Piani et et al. (2010), Piani et al, (2010b), Dosio and Paruolo (2011), and Dosio et al. (2012), Ruiz-Ramos et al., (2016), Dosio and Fischer, (2018). Dosio (2016) showed that bias-adjustment largely improves the value of present and future threshold-based indices (e.g., the number of frost days): these indices are generally poorly simulated over the present climate, such that the projected climate change may not be reliable.

We will better link with next paragraph. Also, we will include the references of previous studies (see references included in the Answers to referee 1). Also, we can mention climate change differences when using bias correction methods and cancellation of climate change signal over time (Michelangeli et al, 2009; Casanueva et al., 2018).

Casanueva, A., Bedia, J., Herrera, S., Fernández, J., and Gutiérrez, J. M.: Direct and component-wise bias correction of multi-variate climate indices: the percentile adjustment function diagnostic tool, Clim. Change, 147, 411-425, 10.1007/s10584-018-2167-5, 2018.

Dosio, A., and Paruolo, P.: Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate, Journal of Geophysical Research Atmospheres, 116, 10.1029/2011JD015934, 2011.https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011JD015934

Dosio, A., Paruolo, P., and Rojas, R.: Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: Analysis of the climate change signal, Journal of Geophysical Research Atmospheres, 117, 10.1029/2012JD017968, 2012.

Dosio, A.: Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models, Journal of Geophysical Research: Atmospheres, 121, 5488-5511, doi:10.1002/2015JD024411, 2016.

Dosio, A., and Fischer, E. M.: Will Half a Degree Make a Difference? Robust Projections of Indices of Mean and Extreme Climate in Europe Under 1.5°C, 2°C, and 3°C Global Warming, Geophys. Res. Lett., 45, 935-944, 10.1002/2017GL076222, 2018.

Michelangeli, P. A., Vrac, M., and Loukos, H.: Probabilistic downscaling approaches: Application to wind cumulative distribution functions, Geophys. Res. Lett., 36, 10.1029/2009GL038401, 2009.

Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99, 187-192, 10.1007/s00704-009-0134-9, 2010.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, Journal of Hydrology, 395, 199-215, 10.1016/j.jhydrol.2010.10.024, 2010.

Ruiz-Ramos, M., Rodríguez, A., Dosio, A., Goodess, C. M., Harpham, C., Mínguez, M. I., and Sánchez, E.: Comparing correction methods of RCM outputs for improving crop impact projections in the Iberian Peninsula for 21st century, Clim. Change, 134, 283-297, 10.1007/s10584-015-1518-8, 2016.

Materials and methods

Regarding the selection of models and scenarios, although hardly done in literature, the choice of models could be better justified using methodologies as in (Mendlik and Gobiet, 2016), since there is evidence of high sensitivity of climate model selection (Wilcke and Bärring, 2016). However, the authors chose the two reasonable scenarios (RCP 4.5 and RCP 8.5), allowing for consistent comments on importance of mitigation in context of actual discussion. Key equations of the chilling models should be provided in the additional material. In the main text a comment on the validation of the models should be given, in the view of their applicability on future time series.

The ensemble of climate models contains 10 members, which was the whole set of models available. Additionally, Figure 3 is meant prove that our inputs are robust. To take into account the referee's concern we will include the following (bold) text in the corresponding section:

“..... This ensemble size is considered to be large enough by the agricultural impact community to retrieve robust results (Martre et al., 2015; Rodríguez et al., 2019). ***Due to the complex orography of the Iberian Peninsula and its remarkable climatic diversity (CLIVAR-Spain, 2010), no additional systematic selection was performed to reduce the number of RCM ensemble members (e.g., Mendlik and Gobiet, 2016). A thorough analysis in this sense would imply decomposing the Iberian Peninsula into several climatic sub-regions (Wilcke and Bärring, 2016) and would derive into a much more complex process to potentially improve already robust results.*** *The outputs of the EUR-11 ensemble for two RCPs were considered: 1) +4.5 W/m2 radiative forcing increase at the end of the 21st century relative to pre-industrial levels (RCP4.5) and 2) the same but for +8.5 W/m2 (RCP8.5).”*

CLIVAR-Spain: Climate in Spain: past, present and future. Regional climate change assessment report, Ministerio de Ciencia e Innovación (España), Ministerio de Medio Ambiente y Medio Rural y Marino (España) 978-84-614-8115-6, www.clivar.es, 2010.

Mendlik, T., and Gobiet, A.: Selecting climate simulations for impact studies based on multivariate patterns of climate change, Clim. Change, 135, 381-393, 10.1007/s10584-015-1582-0, 2016.

Wilcke, R. A. I., and Bärring, L.: Selecting regional climate scenarios for impact modelling studies, Environ. Modell. Softw., 78, 191-201, https://doi.org/10.1016/j.envsoft.2016.01.002, 2016.

Results

With regards to the CV, MAPE and IQR, the classes > 20, >0.4... are in my view not informative enough. Also, in section 3.1, the MAPE values are declared as problematic above 20% for few grid points, without mentioning until how high they stretch. Thus, no conclusion can be made if the computation for these grid points can be trusted at all.

High MAPE values can be related also to the low values of the chilling accumulation in those areas, and therefore it does not mean that necessarily the projections cannot be trusted at all. It means that we should be more careful when interpreting the results. Nevertheless, we will mark somehow the areas in the plots where values were greater than 20%. Also, we will choose a more understandable, representative classes for the figures, and the top end will be specified.

Discussion

The difference between the two researched scenarios could be expressed more clearly (p.10, ll.11-19).

We will introduce a sentence here discussing the main difference found between results obtained for each RCPs.

References

I join the request made in RC1 for indented references. In the text, the reference in p.3, l.29 should be revised.

Format change will be made, if the journal allows so.

The mentioned reference has already been checked and corrected.

Figures

As stated in RC1, all figures need to be presented with a scale bar, north arrow, and (due to inconsistency between figures) the reference system. Preferably all maps would be shown in the same projection (or the stretch of the figures should be revised). The layout of subfigures could be optimized so as to allow for bigger figures. If the decision will be taken to not report on all models, this could be of great improvement of the readability.

- *We will modify the figures as suggested. Layout would be revised to allow the figures to be as large as possible, although this kind of composite figures are quite common in climate and impact publications (e.g. see https://www.meteo.unican.es/es/view/publications)*

- *On the models to be reported: Please see our answers above. We are convinced that our arguments are correct and sound, but if the editor and both referees ask us to remove some of the chilling models considered, we would be willing to do so.*

Figure 1 shows a good overview of land use in Spain for the reader, exposing major growing areas for the considered crops. Values seem reasonable from my experience. However, the choice of the color map is unfortunate, <1%, which could be conceptually be negligible, is very hard to distinguish from the higher classes. I suggest to revise the classification to a lower number of classes, 5 being preferred. A clarification is needed whether the map shows the percentage from the total area or from area classified as cropland.

We will modify the figure 1 as suggested. We will specify that the percentage refers to the total area.

Figure 2 features a useful example output of the analysis, but it was not justified that this is a representative example. The most reliable model would have been preferred, the Dynamic model was judged as best performing (Luedeling, 2012). In subplot B, over the years, the chilling units decrease, a trend line could be interesting, next to the mean. Subplot C should highlight which model is used for subplots A and B. In subplot D, neighboring grid points expose substantial differences in this mountainous terrain. With regards to the shortcomings mentioned in mountains areas, a further study could envision a more focused analysis on those areas.

The example was considered representative for two reasons, 1) because it shows the general procedure followed for each cell to obtain an individual outcome from the climate ensembles, illustrating how the methodology aggregated yearly information and projections using different climate models, 2) to explain how the initial and final chilling accumulation dates are calculated, and this is particularly important for the Utah-based *methods considered in the study as warm temperatures sometimes negatively contribute to chilling accumulation. This is not the case of the Dynamic model where only positive increases of chilling accumulations are added up, being the purpose of the cited figure of illustrating one of the Utah-based methods, more complicated to explain in that sense. So Dynamic model is not as useful as the others to illustrate this.*

The text will be changed to further explain the need of this figure and the footnote will be modified to clarify that all subplots refer to the same model.

We agree with the referee's suggestion that a further study more focused on mountainous areas would be very much interesting from the scientific point of view. However, our priority was to focus in main productive areas that are usually at lower regions.

Regarding Figures 3 -8 and as mentioned above, classes such as >20 are little informative. In this line, it would be of great value if the maps could either exclude or highlight less reliable outcomes. This could be done by keeping grid points white, or, if readability is not compromised, with a hatched overlay. From visual comparison, there seems to be a substantial part of the apple cultivation shown in Figure 1 in coastal and mountainous areas, those reported as with comparatively high errors.

As answered above, we will choose a more understandable, representative classes for the figures, and we will highlight the areas >20 in these figures to facilitate interpretation.

Technical comments (additionally to those mentioned in RC1, to which I fully agree):

* P.2 l. 26 delete 'it

We will delete it (unless our professional English editor suggest otherwise).

* P.2 l.18, production, not productivity (if productivity is meant, the reference i.e. area should be specified, and I agree it is not relevant in this paper, rather give the importance of other fruits in Spain, ideally with national statistics rather than FAOSTAT)

Yes, the referee is right, we will change it in the revised version.

* P.3 l.34, add 'among other regions'

The referee's suggestion will be included in the revised version.

* P.5 l.23 inconsistent usage of Dynamic model / Dynamic method

The referee's suggestion will be included in the revised version, using Dynamic model throughout the paper.

* P.8 l.10, specify where the biggest change occurred

We will specify it in the revised version.

* P.9 l.16, Mediterranean','

The referee's suggestion will be included in the revised version (unless our professional English editor suggest otherwise).

* P.9 ll.17-18 reformulate

The referee's suggestion will be included in the revised version as follows:

"In light of the results, our hypothesis is that the stations in these areas are poorly represented by the interpolated Spain02 dataset."

* P.9.l.21 a warmer scenario

The referee's suggestion will be included in the revised version.

* P.9 ll.28-29 'Nonetheless, few tree crops are grown [...]' – have these areas also be found as potential new cropping areas?

Yes, at the lower part of the mountains, but the affected areas would be relatively small. That is why in our view improving the estimations for these areas would be interesting of course but not a priority.

* P.10 l.17 are you comparing this value (469 chilling units, according to the De Melo-Abreu method) with all outputs? It should only be compared to the output of the analysis using the same method, which, in the case of the far future under RCP8.5, where the map shows mainly values between 500-1000 chill units in the area coinciding with olive cropping.

The comparison is established only between results from the same models. We will stress this in the revised version as it is a key point. it seems that it was not clear enough. Also, we will specify more the region we are referring to (red areas in Figure 8, first column, third row).

Thank you for your thoughtful revision. We have tried to address all the issues you raised. We are convinced that our arguments are correct and sound, but if the editor and both referees ask us to remove some of the chilling models considered, we would be willing to do so.

References

Campoy, J.A., Ruiz, D., Egea, J., 2011. Dormancy in temperate fruit trees in a global warming context: A review. Sci. Hortic. 130, 357–372. https://doi.org/10.1016/j.scienta.2011.07.011

Luedeling, E., 2012. Climate change impacts on winter chill for temperate fruit and nut production: A review. Sci. Hortic. 144, 218–229. https://doi.org/10.1016/j.scienta.2012.07.011

Mendlik, T., Gobiet, A., 2016. Selecting climate simulations for impact studies based on multivariate patterns of climate change. Clim. Change 135, 381–393. https://doi.org/10.1007/s10584-015-1582-0

Michelangeli, P.-A., Vrac, M., Loukos, H., 2009. Probabilistic downscaling approaches: Application to wind cumulative distribution functions. Geophys. Res. Lett. 36. https://doi.org/10.1029/2009GL038401

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the KoÂ´lppen-Geiger climate classiïnˇA¸cation. Hydrol Earth Syst Sci 12.

Schaffer, B., 2018. Handbook of Environmental Physiology of Fruit Crops: Volume I: Temperate Crops, 1st ed. CRC Press. https://doi.org/10.1201/9780203719299

Wilcke, R.A.I., Bärring, L., 2016. Selecting regional climate scenarios for impact modelling studies. Environ. Model. Softw. 78, 191–201. https://doi.org/10.1016/j.envsoft.2016.01.002