Review of "Seasonal streamflow forecasts for Europe – II. Explanation of the skill" by W. Greuell et al.

Reviewed: December 2016

Recommendation: The manuscript is acceptable with major revisions.

In this paper, the authors present the sources of skill of a model-based seasonal hydrological forecasting system, which produces hydrological forecasts for up to seven months of lead time over Europe. Seasonal hydrological forecast systems over Europe are scarce, as well as the analysis of the sources of the skill over this region, which makes this work relevant to HESS and to the wider hydro-meteorological community.

The authors analyse the sources of skill in the seasonal meteorological, discharge, runoff and evapotranspiration forecasts, using a variety of skill metrics and experiments. This complete and very interesting analysis enables to disentangle the relative contributions of initial hydrological conditions of soil moisture and snow and of meteorological forcing on the skill of seasonal hydrological forecasts up to several months of lead time. The results would however largely benefit from being more concise and structured, in order to guide the readers throughout the paper. To this end and for reproducibility, some results would also need more explaining in the methods section of the paper (i.e., climate change).

I would like to raise two major comments about the methods of this paper, which should be addressed by the authors. These are:

1. Presenting the skill of the meteorological forecasts is a great idea as it highlights the importance of meteorological forecasts for seasonal hydrological forecasting. Furthermore, it shows the differences between the skill in seasonal meteorological and hydrological forecasting, which is a prime motivation for producing hydrological forecasts at seasonal time scales. Nevertheless, I do not fully understand the choice of presenting the skill of the raw S4 forecasts rather than the bias-corrected S4 forecasts, ultimately used to produce the hydrological forecasts in WUSHP. Therefore, I believe that it would make more sense if the skill of the bias-corrected S4 forecasts was presented in the results of the paper, with appropriate skill metrics. Otherwise, readers might question your choice to bias-correct the S4 forecasts in WUSHP.
2. The ESP-experiments carried out for this paper are very original! However, I am not sure if I agree with the way that you designed your ESP-experiments (ESPall, ESPsoilm, ESPsnow). The use of an identical meteorological forcing resampled from the S4 hindcasts for each year might in the end produce an artificial signal, which might lead to a biased analysis of the skill in this paper and is not the aim of the experiment design. Please read and address the detailed comments I have made about this in the rest of the revision. Additionally, the reasons you give for designing all of the experiments differently from the widely used standard ESP and reverse-ESP are not sufficient to argue your choice. In the methods, you state that it is in order for these experiments to be closer to the Full Hindcasts, which would not be the case when using the standard ESP and reverse-ESP formulations. However, I do not see the need for the experiments to be close to the Full Hindcasts, at least in the context of this paper. After reading the detailed comments I have made in this revision referring to the ESP-experiments, please consider either arguing more thoroughly why you have decided

to take a different approach from the standard ESP and reverse-ESP or changing the experiment designs.

The paper is overall written in a generally fluent and precise language. As a whole, I thought that this paper provided valuable results and I would therefore be pleased to see it published in HESS, after major revisions. Below are comments which will hopefully help the authors to improve the paper and highlight the value of the results it contains.

Title: The title is pertinent with regards to the contents of the paper. However, the formulation "Explanation of the skill" could be rephrased in order to sound more scientific. You could consider rephrasing it to for example "Sources of the skill".

Abstract: Overall, the abstract provides a complete summary of the paper. It could however benefit from being more concise; here are a few suggestions in that direction:

- Page 1, line 12: could you please consider rephrasing "hindcast simulations […] were carried out" to "hindcasts […] were generated"? The term simulations could be confusing here as you are referring to forecasts.
- Page 1, line 17: please change "simulations" to "hindcasts" for the two instances.
- Page 1, line 20-21: this sentence could be removed from the abstract, not hindering the content of the abstract and making it more concise overall. Please consider doing so.
- Page 1, line 23-24: the sentence could be shortened and thus made more concise by removing "to all potential sources of skill but".

Introduction: The introduction is interesting and introducing this paper's content with a summary of the previous paper's main findings is a great idea. The introduction however contains a lot of overlap in what is being said. Here are a few suggestions that could maybe help to make the introduction more concise and structured:

- Page 1, line 26: the word "may" sounds like society may also not benefit from such forecasts. It would therefore be interesting to refer to papers tackling this topic, such as: Viel et al. (2016), Soares and Dessai (2016), Crochemore et al. (2016), among others.
- Page 1, line 28: it would be good to add references for other applications of the seasonal predictions, as done for the energy generation sector.
- Page 1, lines 30-31: please consider rephrasing the beginning of the sentence to "WUSHP produces hydrological simulations and forecasts from the Variable Infiltration Capacity […]".
- Page 2, line 1: could you please rewrite "[…] in runoff is fading […]" to "[…] in runoff was found to be fading […]"?
- Page 2, lines 1-2: please consider rephrasing this to "[…], but some significant skill remained up to 7 months of lead time".
- Page 2, line 3: could you please change the word "causes" to "sources"? which is widely used in this context.
- Page 2, line 3: please consider changing "[…] along two lines" to "[…] and is structured in two main parts". Then in the following paragraphs introduce the two parts by saying something similar to: "First, an analysis of the skill […] is carried out" and "In a second part, sources of predictability are analysed […]" (page 2, line 12).
- Page 2, line 4: could you please specify here what variables of the S4 meteorological forcing are analysed in this paper?
- Page 2, line 7: "starting date" or "initialisation month" is better here than "start period".

- Page 2, lines 8-11: these few lines sound too much like results, it should sound more like a literature review. Please consider rephrasing these sentences to sound more like an introduction material. Referring to specific maps of the paper is for instance not adequate here.
- Page 2, lines 14-17: could you please consider combining this part of the introduction with page 2, lines 23-24? which is essentially a repetition of the former.
- Page 2, line 17: please add the word “contributions” after “soil moisture and snow initial conditions”.
- Page 2, line 18: the ESP refers to a forecasting technique rather than modelling.
- Page 2, line 19: instead of “as realistic as possible and vary from year to year”, a clearer formulation would be for example “our best estimates of the current initial conditions for this specific forecast starting date”.
- Page 2, lines 26, 30, 38 and 39: please change “simulations” to “forecasts” or “hindcasts”.
- Page 2, line 26: consider rephrasing “is as realistic as possible and is” to “it is our best estimate of the current meteorological conditions,”.
- Page 2, lines 32-33: I would rather use the term “climatological information” instead of “no information at all”, which is not accurate, and then “, which is the case when they have a climatological distribution” could be removed.
- Page 2, lines 35-36: the sentence “All of these studies basically looked at uncertainty in seasonal forecasts.” could be removed as it is not necessary and does not sound so good.
- Page 2, lines 37-38: rephrase this to “(2010), we will first look at the skill of the ESP hindcasts which we will then compare to the standard […]”.
- Page 3, line 1: could you please specify here the total skill of what hydrological variables will be quantified by removing one or more sources of skill?
- Page 3, lines 1-3: the sentence starting with “It should be noted” is a repetition of what was already said earlier in the introduction. Could you please consider removing it here?
- Page 3, line 4: could you please rephrase this to “the sources of skill for seasonal hydrological forecasting over Europe”?
- Page 3, line 4: it would be good to specify what is dominated by initial conditions.
- Page 3, lines 13-14: it would be nice if you could add references for this use of evapotranspiration predictions.
- Page 3, line 15: please change “of evapotranspiration” to “in evapotranspiration forecasts”.
- Page 3, lines 18-22: could you please specify the sections for each of these different analysis parts? As was done on page 7, lines 10-14.
- Page 3, lines 19 and 20: it is not clear from the introduction what is meant here by “the various ESP experiments”. It becomes clearer after reading the methods section though. Could you thus rephrase this here to “the ESP and reverse-ESP experiments”?
- Page 3, line 22: please change “evaporation” to “evapotranspiration” here.
- Page 3, line 22: the sentence about additional figures is not appropriate here. Please consider moving it to the methods or results section of this paper.

Section 2.1:

- Page 3, lines 25-29: in this description section, it would be nice if the time step of the simulations, as well as downscaling of the meteorological forcing for the hydrological simulations was mentioned.
- Page 3, line 26: please specify that the bias correction is for the meteorological forcing.

- Page 3, line 27: could you specify hindcasts of which variables are used for this paper? Namely runoff, discharge and evapotranspiration.
- Page 3, line 33: “each of” can be removed here.

Section 2.2:

- Page 4, line 6: are those terciles of the observations or of the forecasts? It would be good to specify here.
- Page 4, lines 6-9: I would explain here that for that reason this paper presents results only in terms of the correlation coefficient.
- Page 4, line 16: please specify what is the maximum area that a basin can reach in order to be called a small basin in this paper.
- Page 4, lines 16-18: I would move these results earlier, after the sentence about the comparison between theoretical and actual skill on page 4, lines 12-13, where it fits better.
- Page 4, lines 19-23: I do not understand why you decided to analyse the skill of the non-bias-corrected (raw) S4 forecasts here as you are using the bias-corrected S4 forecasts to produce your hydrological forecasts for this paper. It would thus make more sense to present the skill analysis of the bias-corrected forecasts here. Also, the fact that there are only negligible differences between the bias-corrected and the raw S4 forecasts is, as you mention it, due to your choice of the metrics to compare them. I would thus suggest to use different skill metrics for this specific comparison analysis. If there are still only negligible differences with appropriate skill metrics, it would be ideal to mention that, as previously shown by Wood et al. (2016), negligible meteorological forcing skill improvements can lead to large seasonal streamflow skill improvements (as you mention it in your discussion section), which is why you decide to bias-correct the S4 forecasts here.
- Page 4, line 25: please add “in the scores overview” after “high temporal resolution”.
- Page 4, lines 27-28: “lead month zero” is present in many results in this paper, I would thus remove this sentence which is not accurate.
- Page 4, line 29: please specify what will be analysed at the level of the entire domain, the skill?
- Page 4, lines 30-32: please remove the example, sit does not fit here.

Section 2.3: for this section, it would be good to make a figure of the various “ESP experiments”, this would help the readers understand exactly what was done here.

- Page 4, line 34: you could also refer to the ESP experiments with “ESP” and the reverse-ESP experiment with “reverse-ESP” or “revESP”, which would be much clearer. Also, please make sure that you use either the term “ESP experiments” or “ESP-experiments” if you decide to keep this terminology.
- Page 4, line 36-page 5, line 3: ESPall
    - It is not clear to me how the ESPall can have 15 members, since there are 28 years of hindcasts from which the members can be selected. Are some years not used? This should be clarified here.
    - It is mentioned in the results that the same meteorological forcing is used each year (this should be made clearer in the methods). However, I am not sure if this is a good resampling strategy. It could indeed be that the members selected lead to an artificial and persistent skill/signal in certain regions and for some initialisation dates. I would suggest to resample the members for the ESP in a random way for each year individually, the forcing would thus vary for each year of the forecasts.

Alternatively, you could show here that using the same meteorological forcing each year, or using a different meteorological forcing by randomly resampling the members for each year, leads to the same results and that you thus decided to use the former and simpler resampling method.

- It is in theory a nice idea to resample from the S4 hindcasts instead of the observed meteorological conditions. However, the wider reason for using the standard ESP as a reference to analyse the skill of a seasonal hydrological forecasting system is because it is a computationally cheap method (ideal for operational forecasting) invented when seasonal meteorological forecasts were not skilful enough and based on the assumption that previous years' meteorological conditions are a good indication of future meteorological conditions for the same time of the year. The standard ESP is furthermore known to be a skilful reference and having a more skilful seasonal hydrological forecasting system (here called Full Hindcasts) would guarantee that it is skilful. Here, resampling from the S4 hindcasts is not computationally cheap since you first have to produce those meteorological hindcasts. I also do not entirely understand why you would want this ESP experiment to be as close to the Full Hindcasts as possible. Also, avoiding to reproduce the reference simulation is not a good argument here as this can also be avoided in the standard ESP by simply not selecting the current year. You thus have to argue the choice for this alternative ESP method better in order to use it for your paper. Otherwise, please consider redesigning the experiments. This will impact all other "ESP experiments" of this paper (including the revESP).

- Page 5, lines 4-9: the ESPsoilm and ESPsnow are really clever!

Section 3.1:

- Page 5, lines 26-27: please rephrase to "significant skill approaches 5%, the no skill line. Hence […]".
- Page 5, line 28: rephrase to "there is more skill in January, February […] than during the other months".
- Page 5, line 32: please add "(see Fig. 1a)" after "coastal regions".
- Page 5, lines 34-36: this climate change analysis comes as a surprise here, it should be explained in the methods part of the paper to guide the readers throughout the paper. If it is an analysis done in another paper, this paper should refer to it.
- Page 5, line 38: please rephrase to "the theoretical no skill limit".
- Page 6, line 10: specify here that the three summer months are JJA.
- Page 6, line 12: please add "for the summer months" after "function of lead time".
- Page 6, lines 12-13: this is however a completely different area, compared to Europe. Can you really compare the two?
- Page 6, line 30: "mix" is not appropriate here, maybe using "are a combination of" would be better?
- Page 6, lines 36-38: since there is not much to show in the figure and this paper already contains many figures, I would suggest to move the figure to the supplementary material and say that it is not shown here in the text. You could in the text then say what fraction of the domain has skill for "lead month 0".

Section 3.2:

- Page 7, line 2: could you please specify here what other sources is referring to? Initial conditions?
- Page 7, lines 3-6: this climate change analysis was not mentioned in the methods, please mention it there.
- Page 7, lines 10-14: I find that this whole paragraph describing the content of the following results breaks the results section. I would remove it or remind the readers of the results structure in the methods rather.

Section 3.2.1:

- Page 7, line 18: this is however not entirely true, in the companion paper, some key differences were highlighted between runoff and discharge. Please consider rephrasing this to say that they show a high degree of similarity in terms of magnitude and spatial patterns of skill, or remove the sentence as a whole.
- Page 7, line 20: please specify that the reverse occurs beyond "lead month 1" for most target months, because it does not occur for all.
- Page 7, lines 21-23: there are quite a few differences between the large and the small basins plots (Fig. 4c and Fig. 4d respectively). The differences are however not highlighted here are it is not the main focus of this paragraph. This questions the existence of the two figures. I would either merge small and large basins in one figure, since there is not distinction in the text, or raise the differences (even briefly) in the text.
- Page 7, lines 26-34: this is quite a nice explanation for the reversal of the skill! However, this is hence due to the ESPall experiment design: the use of the same S4 forcing for each year of forecast produced. As discussed in the methods, this should probably be changed to using different random forcing for each year. Because it could be that this specific selection of S4 forcing made here leads to non-random weird skill patterns, in other words to some random signal. This is however not the goal of the method, which is as you said to assess the importance of initial conditions for seasonal discharge and runoff forecasting and the impact of losing the knowledge about the future meteorological forcing, and using random meteorological forcing from previous years of hindcasts as proxy for future meteorological forcing, on the seasonal runoff/discharge forecasting skill.
- Page 7, lines 35-36: it would be interesting to know whether for specific regions in Europe the revESP is more or as skilful as the ESPall for certain target months-lead times combinations. Could this be done and added here?
- Page 7, lines 38-39: the parentheses content is not needed here, the readers can go back to this part of the results if they want to read the specific numbers and it breaks this part of the results. These numbers were however not stated in section 3.1 and should be moved there.
- Page 8, lines 1-4: this is a very interesting observation!

Section 3.2.2:

- Page 8, lines 6-7: this is true compared to the ESPsnow experiment and should be specified.
- Page 8, lines 8-9: is this however true for all lead times?
- Page 8, line 14: please, first say what the figure shows in general.
- Page 8, lines 14-18: could you please specify which ESP experiment (ESPsoilm, ESPsnow or ESPall) you are referring to when you write those results, it will help the readers to understand them faster.
- Page 8, line 17-18: I am not sure what is meant by "combined initialisation map". Please rephrase.

- Page 8, lines 19-22: I would move this section earlier, when you are talking about Figure 5, to make it more structured. You can then refer back to this feature when looking at the maps of Figure 6.
- Page 8, lines 25-26: it is hard to understand this point, please consider rephrasing.
- Page 8, lines 27-30: this is a very interesting observation!
- Page 9, lines 2-9: this is a very interesting observation, it would be interesting to show the ESPsnow for soil moisture for May with lead 0, for a comparison. So it is also proving that spin-up is important for hydrological modelling.
- Page 9, lines 6-9: this is a repetition of page 9, lines 2-6. Please consider removing this repetition or combining both explanations.
- Page 9, lines 10-12: this paragraph should be moved earlier, when Figure 5 is described, which would make it more structured and hence clearer to read.
- Page 9, line 12: this order rather depends on the month, not the season, because you are talking in this paper in terms of months. Please change.
- Page 9, lines 13-15: could you please state here what was done exactly with those maps? Or maybe say this in the methods section.
- Page 9, lines 13-19: so none of these hotspots have skill thanks to the meteorological forcing?

Section 3.3:

- Page 9, line 21: I would repeat here the intrinsic value of evapotranspiration hindcasts.
- Page 9, lines 21-22: "the power […] ESP experiments" is an odd phrase, rephrase or remove.
- Page 9, lines 23-24: I would remove the piece of the sentence about the April and July decomposition. It is not needed and distracts the readers from the first analysis.
- Page 9, line 25: please explain the overall Fig. 9a before entering into details.
- Page 9, line 25: specify that the levels of predictability in Fig. 9a are for the Full Hindcasts.
- Page 9, line 27: could you please remind the readers in between parentheses what the three ESP experiments are.
- Page 9, line 29: specify that you are talking about the evapotranspiration hindcasts.
- Page 9, lines 33-34: in this sentence, add in between parentheses the ESP experiment you are talking about (revESP, ESPsoilm, ESPsnow), to guide the readers through the results nicely.
- Page 10, lines 1-2: this is during the snowmelt season, please mention.
- Page 10, lines 3-10: this part of the results will benefit greatly from explaining the climate change analysis in the methods section of the paper.
- Page 10, line 11: remind the readers the skill of what variable they are currently looking at.
- Page 10, line 20-21: specify that this is for evapotranspiration hindcasts in April.
- Page 10, lines 21-26: the fact that both temperature and evapotranspiration hindcasts may have the same predictability source is a hypothesis here. Please rephrase the sentences to sound like one.
- Page 10, lines 29: please remind the readers once again what three ESP experiments you are referring to.
- Page 10, lines 30-31: please consider saying that this result can be drawn from the fact that the ESPsoilm shows a higher skill than the ESPsnow and revESP for the Mediterranean.
- Page 10, lines 34-35: please specify which ESP experiments those are.
- Page 11, line 4: is Figure 11f really needed? There are already a lot of figures so I would consider removing it.

Discussion:

- Page 11, line 8: please use the terms runoff or discharge rather than streamflow here, to be consistent with the rest of the paper.
- Page 11, line 13: it is not really clear what you are referring to when you say the "uncertainty strategy", could you please rephrase?
- Page 11, line 23: remind the readers that these hotspots regions and periods of skill were identified in the companion paper.
- Page 11, lines 24 and 27: the term "sources of skill" is preferred over "causes of skill".
- Page 11, line 28-page 12, line 8: this analysis is very interesting!
- Page 12, line 10: instead of "replaces" I would say "becomes less skilful than […]".
- Page 12, lines 11-13: you can however not exactly compare your results to results from other papers here as their ESP experiment was different from yours I suppose. You could still cite their results to compare to yours but make sure to highlight this difference.

Conclusions: the conclusion is overall too long. I would shorten it to keep only the main results of the paper. Here are some suggestions:

- Overall, please don't refer to figures here.
- Page 13, line 11: describe quickly the different ESP experiments.
- Page 13, lines 14-15: when you say "other ESP-experiments" mention that these were performed in this paper. Otherwise it sounds like you are talking about ESP experiments from other papers.
- Page 13, lines 18-19: I would remove the piece of sentence "Similar […] domain".
- Page 13, lines 22-29: I would summarise this whole paragraph in just a few sentences. Just to convey the main conclusion from these results. The detailed results that you are currently describing can be found in the results section of the paper.
- Page 13, lines 30-36: same as above.
- Page 13, line 37-page 14, line 5: I would move this paragraph in the discussion section of the paper.

Figure 1:

- Please consider swapping Figure 1a and 1b as you are first describing result from Figure 1b in the results.
- Figure 1a: could you please add a label for the colour bar saying that this shows R?
- Figure 1b:
    - Please consider making the y-axis a log scale so that we can see what is happening around the 5% line?
    - Would making a colour scale for the initialisation months be possible? This would maybe make the plot more understandable.
- Caption:
    - Please specify that the legend which provides the percentage of cells with significant R values is in the top left corner of Figure 1a.
    - Could you also state that darker red colours signify a better skill?

Figure 2:

- Figures 2a and 2b:

    - These figures look quite messy for lead times 1 and 2. Would using a log scale for the y-axis help with that?
    - Also consider making a colour bar for the different starting months, as suggested for Figure 1a.
- Figure 2c: the labels of this figure are quite messy. Could you please put a legend outside the figure instead?
- Please remove the general title, it is already said in the caption.
- Specify that the colour bar is for R by adding a label next to it.
- Caption:
    - Instead of writing "As Fig. 1" I would mention here again what this figure is. Because it is easier to read directly under the figure than having to jump from a figure caption to the other figure.
    - Consider removing the exclamation mark after "not the trend itself".

Figure 3: I would suggest to remove this figure and put it in the supplementary material. In any case, all comments made to Figure 1b apply here, as well as the caption explanation instead of writing "As Fig. 1b".

Figure 4:

- These figures are too messy, please make a common legend to explain what the different lines are.
- Add an x-axis label specifying that these are target months.
- You could remove the x-axis tick labels for Figure 4b as it is shown in the Figure 4d below.
- You could remove the y-axis tick labels for Figures 4b and 4d as they are the same as in the Figures 4a and 4c.
- Caption:
    - Please explain what the figures show in the caption instead of writing "As Fig. 2c".
    - Instead of "first two panels" say "top two panels" and instead of "other two panels" say "bottom two panels".

Figure 5:

- Please remove the title as it is specified in the caption already.
- Could you please make a legend for the different lines, it is currently quite messy?
- Add an x-axis label specifying that these are target months.
- Explain what the figures show in the caption instead of writing "As Fig. 4".

Figure 6:

- Consider removing the main title.
- Please add a label for the colour bar.
- Caption:
    - Explain what the three ESP experiments are.
    - Could you also explain what the figures show in the caption instead of writing "For more explanation, see Fig. 1a"?

Figure 7: same comments as for Figure 6, except the comment regarding the different ESP experiments.

Figure 8: same comments as for Figure 7, except the comment about the main title. Additionally, the caption should not describe the results.

Figure 9:

- Consider removing the main title.
- Figure 9a: same comments as for Figures 1b, 2a and b.
- Figure 9b: same comments as for Figure 5.
- Figure 9c: same comments as for Figures 2c and 4a, b, c and d.
- Caption:
    - Explain what the figures show in the caption instead of writing “for more explanation, see Fig. 1b”.
    - Please specify what the different ESP-experiments are.

Figure 10:

- Please add a label for the colour bar.
- Specify what the figures show instead of saying “for more explanation, see Fig. 1a”.

Figure 11:

- Consider removing the main title.
- Add a label for the colour bar.
- Caption:
    - Could you specify what the figures show instead of saying “for more explanation, see Fig. 1a”?
    - “The final is panel f and depicts the skill […]”.
- Consider removing Figure 11f.

Figure 12:

- You do not need a colour bar for each sub figures: Figures 12a and b can share one, and Figures 12c and d as well.
- Add the labels for the two different colour bars.
- Caption:
    - The caption describes the results, it should not.
    - Please specify what the figures show instead of saying “for more explanation, see Fig.s 1”.

Technical corrections:

- General:
    - Could you please add the word “meteorological” in front of “forcing” when you refer to meteorological forcing. It will make it clearer to the readers what you are talking about.
    - Please consider changing “lead month” to “month of lead time” or “lead time”, which is more widely used, and will hence be clearer for the readers even without having read the methods section.
    - Could you please replace “panel” with Fig. figure# subfigure#? E.g., for Figure 5, panel c would be replaced by Fig. 5c.
    - Could you please consider renaming the terms “pseudo-observations” and “real observations”? I would for example use “analysis” (as done in meteorology) or “simulations”, for the pseudo-observations, and simply “observations” for the “real observations”.

    - Could you please change “North” to “Northern”, “South” to “Southern”, “West” to “Western” and “East” to “Eastern” when in front of a country’s name?
- Page 1, line 10: could you please add a comma after “In WUSHP”?
- Page 1, line 12: could you please change “To explain skill” to “To explain the skill”?
- Page 1, line 13: please consider using the term “analysed”, or something more scientific sounding instead of “looked at”.
- Page 1, line 13: please change “of the first […]” to “for the first […]”.
- Page 1, line 14: instead of “later”, consider using the word “subsequent”.
- Page 1, line 14: “Seasonal forecasts of temperature”.
- Page 1, line 30: consider removing “that was” and adding a comma before “built […]”.
- Page 1, line 33; page 2, line 3: please change the “[…] and lack of skill […]” to “[…] or lack thereof […]”.
- Page 2, line 4: please add a comma in “For S4, this was done […]”.
- Page 2, line 5: rephrase “with initialisation at the” to “initialisation on the”.
- Page 2, line 16: rephrase “and separated” to “to separate”.
- Page 2, line 17: remove the second dot.
- Page 2, line 30: remove the comma after “(2008)”.
- Page 2, line 35: please change the sentence to “changes in the information of the meteorological forcing and the initial conditions.”.
- Page 2, line 37: the word “However,” can be removed.
- Page 3, line 4: add “the” in front of “sources”.
- Page 3, lines 10-11: move “also” to before “play a role”.
- Page 3, line 12: add “the” in front of “skill”.
- Page 3, line 16: “Thus” is not needed here.
- Page 3, line 29: change to “so a total of 5400 simulations (30 years * 12 months * 15 members) was carried out.”.
- Page 3, lines 31-32: consider changing one “namely” to a synonym.
- Page 4, line 5: please add “the” in front of “Relative”.
- Page 4, line 7: change “are similar” to “were similar”.
- Page 4, line 9: if you put capital letters for Below Normal and Above Normal please also add the abbreviation in parentheses after the term is introduced.
- Page 4, line 15: is the hyphen needed in between “large” and “basins”?
- Page 4, line 16: the quotation marks are not needed around the word “observations” here.
- Page 4, line 24: add “a” in front of “relatively”.
- Page 4, line 19: please add a comma after “Here”.
- Page 4, line 26: please add a comma after “(2005)”.
- Page 4, line 29: please add a comma after “result sections”.
- Page 4, line 29: please consider changing the term “remarkable” to “outstanding” or “noteworthy”.
- Page 4, line 30: please add the word “intend to” in front of “provide”.
- Page 4, line 34: please add a comma after “total”.
- Page 5, line 23: please write “are used here as a reference”.
- Page 5, line 24: “A summary of the skill”.
- Page 5, lines 24-25: “in Fig. 1a with statistically”.
- Page 5, line 29: rephrase to “target months (not shown here) hot spots […]”.
- Page 7, lines 3-4: “the question of how much […]”.

- Page 7, line 33: remove the “at” in front of “some time”.
- Page 9, line 22: please add a comma after “First”.
- Page 9, line 27: add “the” in front of “three ESP”.
- Page 10, line 14: add a comma after “in revESP”.
- Page 10, line 14: add “the” before “revESP”, for both instances.
- Page 10, line 30: “From the ESP experiments, it can be concluded […]”.
- Page 11, line 4: there is a space missing between “July 1” and “(panel f)”.
- Page 11, line 22: “The same is probably true for the S4 hindcasts”.
- Page 12, line 8: replace “less than” with “lower than”.
- Page 12, line 19: add “the” in front of “ESP”.
- Page 12, line 33: please add “for” in front of “practical”.
- Page 12, line 36: add “also” in front of “demonstrates”.
- Page 13, line 17: please add a comma after “melt season”.