Streamflow variability over 1881–2011 period in northern Quebec: Comparison of hydrological reconstructions based on tree rings and on geopotential height field reanalysis

Paper *cp-2016-5*

by Brigode, P.; Brissette, F.; Nicault, A.; Perreault, L.; Kuentz, A.; Mathevet, T. & Gailhard, J.

Answer to the referee comments

Comments and suggestions made by the two referees are gratefully acknowledged. We intend to modify the text in response to the main criticisms. In the following, we list the referee comments (in *italic and blue*) and we provide specific responses to these comments (in black).

Furthermore, several modifications will be made in the manuscript related to minor errors detected:

- **Wrong co-authors affiliations:**
  We will correct the wrong co-authors affiliations.
  P. Brigode[1,2,], F. Brissette[1], A. Nicault[3], L. Perreault[4], A. Kuentz[5], T. Mathevet[6], and J. Gailhard[6]*
  [1] Ecole de Technologie Supérieure de Montréal, Montréal, Canada
  [2] Ouranos, Montréal, Canada
  [3] ECCOREV, Aix-en-Provence, France
  [4] IREQ, Varennes, Canada
  [5] SMHI, Norrkoping, Sweden
  [6] DTG, DMM, Electricité de France, Grenoble, France
  [*]Now in: Université de Nice Sophia Antipolis, CNRS, IRD, OCA, Géoazur UMR 7329
- **Wrong reference of the work of Way and Viau (2014) in Labrador as being in New-Brunswick**
  We will correct this mistake.

1 REFEREE #1

1.1 General comments

Overall, the questions addressed by the modeling effort are interesting and the results presented are also interesting. However, I do not feel enough information has been provided to substantiate the findings of the paper due to the lack of detail on the rainfall-runoff modeling. The authors refer to several citations about the model, but the application of the model to this study should be justified. To do this, information on how the model was calibrated needs to be described to show that such calibration was appropriate for the current use. Performance metrics of the calibration should be included. In addition, the model itself needs to be described regarding what inputs are needed, what the "4 and 2 free parameters to calibrate" are (that wording was very confusing to me; line 429). There should also be a description of what those calibrated parameters were and whether their values are appropriate. Their influence on the model results for the study described in this manuscript would also be helpful, given that streamflow reconstruction with the model had discrepancies.

We agree with the referee #1 that the description of the rainfall-runoff model was lacking important information in the present form. A complete description of the GR4J model and its snowmelt routine CemaNeige will be added, with a focus on the inputs needed and on the timestep of the model. A table will also be added, giving a description of each of the six calibrated parameters, their unit and their final calibrated values. We will clearly define what a "free parameter to calibrate" is (terminology classically used in the rainfall-runoff modeling community).

Although there is an entire result subsection devoted to the rainfall-runoff model calibration performances (*4.2 Rainfall-runoff model calibration performances*), we will add the calibration metric values obtained after calibration (Kling and Gupta Efficiency score (Gupta *et al.*, 2009) and its three components).

Finally, quantifying the influence of each rainfall-runoff model parameter on the final streamflow reconstruction is out of the scope of this paper and is definitively an open question (and thus an interesting perspectives of this work). Here, the idea was to apply a classical rainfall-runoff model calibration strategy and then used the obtained parameter values in order to have a model able to transform an ensemble of daily climatic series into an ensemble of daily streamflow series. Nevertheless, our expert (and thus biased) judgement, as hydrologists, is that the rainfall-runoff transformation is not a "significant issue" on this catchment, mainly due to its topographic (topography relatively flat) and hydro-climatic context (catchment hydrology strongly influenced by snowmelt, with slow flow dynamics and none sudden events) and its (very) large size.

I suggest that the authors need to make clearer the inputs needed for the reconstructed streamflows – I assumed it was time series of air temperatures and precipitation only, but that never clearly stated. The timestep necessary for these inputs also should be clear.

We agree that the timestep of the model was not clearly stated. The climatic reconstruction described in this paper is done at the daily timestep and the rainfall-runoff model used is also operating at the daily timestep. Thus, input and output series are all at the daily timestep. It will be clearly stated in the manuscript.

This relates to another thing that was unclear to me regarding why the authors used daily data if all of the comparisons/results shown were monthly. I am guessing the reason is possibly because the rainfall-runoff model only operated at the monthly timestep (relates to the lack of detail on the rainfall-runoff model). Alternatively, perhaps the reservoir operations would like daily data and hence, the approach needs to produce daily data. If this latter is the case, then the authors should present daily results and model performance as well, even if they do not perform as strongly as the monthly summaries of results. Regardless, there needs to be some explanation regarding why daily inputs are needed, but only monthly and annual results are reported.

Monthly and annual values are showed in the paper because of the main goal, which is to compare the new streamflow reconstruction with two other reconstructions (using tree-rings) available at the annual timestep. Nevertheless, as detailed in the previous answers, outputs of the reconstruction methodology are available at the daily timestep. We will evaluated the performance of the climatic and hydrologic performances at the daily timestep and present it alongside the monthly and annual performances (in figures 4 and 7).

I also had some difficulty following the terms used by the authors. This may be because I am not an atmospheric scientist and if the journal feels that its audience is most likely to follow the terminology used then these comments may not be valid. In particular, I was not familiar with "geopotential height," which therefore made discussion of one of the primary datasets used for the reconstructions to be very difficult for me to follow. I recommend if the audience for this article is likely to be interdisciplinary, that the authors provide more description of what geopotential height is and how that relates to the data they used in their study. Also, the authors use "reconstitute" or "reconstitution" quite a bit in the manuscript. I think a more appropriate word is "reconstruct" or "reconstruction." The meaning of "reconstitute" is different from "reconstruct" and I think it is inappropriate here.

Geopotential height fields will be clearly defined in the manuscript, based on the NOAA's National Weather Service Glossary. A geopotential height is the height above sea level of a given pressure level. For example, if a station reports that the 500 [hPa] height at its location is 5600 [meters], it means that the level of the atmosphere over that station at which the atmospheric pressure is 500 [hPa] is 5600 [meters] above sea level. Note that for pressure levels close from sea level pressure (typically 1000 [hPa]), the geopotential height could be negative.

Also, we will only use the words "reconstruct" and "reconstruction" in the manuscript.

1.2 Specific comments

1. Abstract: Suggest rewording line 9 "to compare the obtained streamflow series" to something like "to compare streamflow series obtained with the new method" to be more clear (but also, compare to what?)
Agreed, and we will explicitly state that we compare the streamflow series reconstructed in this article with two streamflow series obtained with tree ring data by other authors.

2. Line 58: The colon (:) after "Canada" seems inappropriate. I suggest just starting a new sentence with "The length (number of years): : :"
Agreed.

3. Line 59: What is "(cQ)2"? Is this an abbreviation for something? If it is a publically available database, should a website be given?
(cQ)² is the abbreviation for "*Impact des Changements Climatiques sur l'hydrologie (Q) au Québec*". The cQ2 database is not publically available.

4. Line 87: Suggest changing "consisting in cal-" to "consisting of cal-"
Agreed.

5. Line 143: Is 15,240 megawatts for the whole complex or just for Caniapiscau Reservoir?
It is for the whole complex. The revised total installed capacity is in fact 17 418 megawatts (to be corrected in the manuscript). The installed capacity for Brisay power plant at Caniapiscau is 469 megawatts.

6. Section 2.1.1: I am unfamiliar with geopotential height reanalysis and a couple of sentences here to define the approach would be useful.
We will introduce this sub-section by defining what a geopotential height reanalysis is and how it is generated (see answer to the general comments).

7. Lines 195-196: I did not understand what the "5 first" were that were extracted –what determines what are first and last in the 56 members?
We agree with the referee #1 that this sentence is unclear, and we will rephrase it. Since each member is equiprobable, selecting the members 1 to 5 (i.e. the "5 first") is equivalent to randomly selecting 5 members out of the 56 members available.

8. Lines 203-204: Keep the greater than sign (>) with the numbers (i.e., >100)
Agreed, we will used the sign here and in other equivalent sentences in the manuscript.

9. Lines 242-244: This is a fragment sentence – please reword
Yes, few words are missing in this sentence and we will reword it.

10. Lines 247-248: What is meant by "A daily catchment series" – do you mean a series of air temperatures for the catchment of Caniapiscau reservoir?
Yes, we meant that we used one and only daily series of air temperature for the entire catchment. We will rephrase this sentence in the manuscript.

11. Line 255: Change "is coming" to "comes"
Agreed.

12. Line 258: change "system" to "systems"
Agreed.

13. Lines 258-259: Why is the La Grande system one of the most important hydropower systems in the world?

The Three Gorges Dam is the most important hydropower system is the world with a total installed capacity of around 22 000 megawatts, the La Grande system has an installed capacity of 17 418 megawatts and is thus one of the most important hydropower system in the world. The Brisay power plant (at Caniapiscau) is ranked as the 9th with an installed capacity of around 500 megawatts.

14. Line 265: Should "abound" be "around"?

Yes, we will change this in the manuscript.

15. Line 314: What do pressure fields have to do with analogue days?

The term "pressure fields" is used here to describe the "geopotential height fields" (see answer to the general comments and to the specific point #6) which are used to find meteorological analogy between days: days with similar geopotential height fields are assumed to be meteorologically "analogue" and thus to produce similar temperature and precipitation pattern over a given region. We will rephrase this sentence in order to be clearer.

16. Line 314: change "fields" to "field"

Agreed.

17. Section 3.1.1: The authors made a good attempt to explain this complicated process of finding analogue days, and Table 1 was helpful. More detail on the Teweles and Wobus (1954) distance is needed – I was not familiar with it, so lines 359-362 were not helpful in describing how the ranking was done (I also suggest avoiding such colloquial phrasing as "thanks to" to be more clear). As I interpreted by reading between the lines, it looks like 20 time series were created for M1, 20 time series were created for M2, and so on. If so, could that also be explicitly stated?

We will add some details on how the Teweles and Wobus (1954) is calculated (the formula and an example of the calculation). The "thanks to" will be deleted. Finally, we will explicitly state that 20 time series are created for each considered members.

18. Line 404: I think a closing parenthesis is missing for "T(dk)"

Yes, we will add a closing parenthesis.

19. Line 410: Delete "In conclusion," – the paper is not finished yet.

Agreed.

20. Lines 414-419: I suggest deleting these two sentences as they are repetitive with statements in Section 3.1.2.

Agreed.

21. Section 3.2: Please see previous comments about needing more detail on the rainfall-runoff model.

Information about the rainfall-runoff model will be added in this section (see answer to general comment).

22. Lines 439-444: Description of the Kuentz et al. (2013) study belongs more in the discussion where the authors could compare their results with those of the previous (similar) study.

We will move this sentence to the discussion section.

23. Line 455: State what is a good value versus a bad value for KGE (i.e., is 1 best?)

We will explicitly state that a perfect KGE value is 1.

24. Lines 458-462: Wouldn't all values of beta be positive, thus what type of values would indicate an overestimation (perhaps values >1)?
The referee #1 is right, all beta values are positive and values greater than 1 indicate an overestimation while values lower than 1 indicate an underestimation. We will correct this mistake in the manuscript.

25. Lines 463-468: Wouldn't all values of alpha be positive, thus what type of values would indicate an overdispersion?
The referee #1 is right, all alpha values are positive and values greater than 1 indicate an overdispersion. We will correct this mistake in the manuscript.

26. Lines 469-473: It probably would be helpful to indicate what value is a better result (i.e., 1 is a perfect correlation)
Agreed.

27. Line 496: delete "of" before "yearly"
Agreed.

28. Lines 513-522: Isn't the ANA with the line over it representing the average of the five 20CR members? If so, isn't it expected that it would have less variability than the individual reconstructions? I do suggest that a definition of the terms with the lines over them (5 ANA with line over it and 5 ANATEM with line over it) be given in the text and in the figure captions
We agree with the referee #1: these lines and associated terms will be clearly defined in the text and in the figure captions, to avoid any confusion.

29. Lines 523-540: I think that the use of the term "time step" is incorrect here unless the modeling was truly done at different time steps (which should be clearly explained if so). Otherwise, "period" or "resolution" would be more appropriate.
Agreed, we will then use the "resolution" term.

30. Line 540: I suggest using "as expected" rather than "logically" or else explain what you are considering as logical.
We will use "as expected" rather than "logically".

31. Section 4.1.2: Is the TEM series referred to here the BEST series?
Yes, TEM referred here (and after) to the BEST series. In order to avoid any confusion, we will change here (and after) TEM to BEST.

32. Section 4.2: I was not clear about how this section was providing different information than Section 4.3.1. Perhaps those two sections could be combined?
These two sections are providing different information since the first one (section 4.2) details the rainfall-runoff model calibration performances (i.e. using observed air temperature and precipitation daily series for reproducing daily observed streamflow series), while the second one is giving detail on the ability of the reconstruction to reproduce observed streamflow (i.e. using reconstructed air temperature and precipitation series for reconstructing observed streamflow series).
We will consider to change the order of these subsections in the manuscript, by presenting them in this new order:

4.1 Rainfall-runoff model calibration performances (1963-1979);
4.2 Climatic reconstructions (1951-2010 and 1880-2011);
4.3 Streamflow reconstructions (1962-2011 and 1881-2011);

33. Lines 635-644: Is this paragraph and Figure 7 about output from CemaNeige model? If so, please state so.

All the rainfall-runoff model outputs presented in the manuscript have been produced by using both GR4J rainfall-runoff model and its snowmelt routine CemaNeige. We will explicitly state so in the manuscript.

34. Lines 635-644: Why is there a focus on May values? Is this an important month or is it the month with the best fits?

There is a focus on the May values because Boucher *et al.* (2011) produced a May streamflow reconstruction, using both continuous series (tree ring minimal density measurements) and discrete series (with ice-scars due to ice abrasion during floods). This month is particularly important in this catchment since it is a month with a large increase of the streamflow and with the observation of the spring flood peak at the end of the month or in early June.

Nevertheless, we intend to change our definition of the spring flood in the manuscript, since it may produce some biases, for example for years for which the flood peak is observed in early June (and thus no more centered in May). Thus, we will produce several "annual spring flood series" from our daily streamflow series (e.g. mean May streamflow values, mean June streamflow values, mean May-June streamflow values, maximum of a moving 30-day window over the May-June period, etc.), and compare these series with the tree-ring reconstructions.

35. Section 4.3.2: Are the reconstructions described here using CemaNeige model?

Yes, see answer to the specific comment #33.

36. Lines 697-703: How did you determine that the 1950-60 period is an "average period" – was there a statistical analysis done to determine this, or are you arbitrarily deciding it is so?

The term "average" is arbitrary in this context, and is used here since the average of the May streamflow reconstructed using the tree-ring over this decade (1950-1960) is close from the overall May streamflow average value (1881-1980). We will change this descriptive term in the manuscript.

37. Section 5: I would like to see a discussion of the parameters and limitations of the rainfall-runoff model. Were assumptions made with the rainfall-runoff model reasonable for this application?

We will add a discussion about the rainfall-runoff transformation in this section, arguing that the assumptions made are reasonable regarding the performances obtained by the rainfall-runoff model over the calibration period (presented in the results section).

38. Line 779: change "representing" to "represent"

Agreed.

39. Lines 799-812: I do not follow the text here. What limited performances are being referred to? What did Kuentz et al. (2015) highlight? How does the work have a perspective of finding an additional series? Is that done and described (I don't think so, but I couldn't really tell what was being stated here)? Please elaborate more on how variables like relative humidity, precipitable water content (what is this?), and local pressure measurements would be used. Would they be used in the rainfall-runoff model? Would they be used to reconstruct precipitation or air temperature? Where would these variables come from? Are they something that you can get from geopotential height? When reconstructing into the past, how you do you estimate these variables? Or are you intending to just reconstruct back through the observational record rather than for centuries as would be done with paleoreconstructions using tree-ring data?

The limited performances referred here are the inability of the ANA approach to reproduce the long-term trend of climatic series (here temperature and precipitation), as already pointed out by Kuentz *et al.* (2015). Unfortunately, none long precipitation and temperature series are available in the studied region. The perspectives are thus to improve the current methodology and particularly testing variables available through the reanalysis for the analogy. Several authors used variables such as air temperature, vertical velocity and humidity at different atmospheric levels (variables produced by the

20CR reanalysis and thus available from 1851 to 2011) to find analogue dates and finally reconstruct daily air temperature and precipitation series. Trying to use such variables for the reconstruction and compare the obtained performances with and without these additional variables is an interesting perspective.

40. Lines 813-823: Although the sensitivity analyses results are not shown, it would be useful to know what variables or approaches were sensitive. I did not follow the last sentence – was this lack of uncertainty shown in the results, and if so, can the authors point the reader to what they are referring to?

(This comment is found also in the general comment of the Referee #2).

Several results of this sensitivity analysis (e.g. the spatial domain considered for the analogy) will be presented in a new Appendix part added to the manuscript. The last sentence was: "*Interestingly, the uncertainty due to the use of five members of the 20CR reanalysis appears to be limited, and even null from 1940 onward*". Yes, this "lack of uncertainty" is shown in results, see for example the Figure 8: it is impossible to distinguish the 5 ANATEM average series after 1940, highlighting that considering 5 different members of the 20CR reanalysis has a negligible impact on the reconstruction of the mean annual streamflow. We will explicitly point the reader to these figures in the manuscript.

41. Line 825: Should "model" be added after "rainfall-runoff"?

Yes, we will add "model" after "rainfall-runoff".

42. Lines 824-839: I do not follow what this paragraph is arguing. How (and why) would the parameter set change in changing climate? What parameter set are you talking about – the ones for the rainfall-runoff model, or perhaps the ones for Equation (1)? Please reword the entire paragraph to be more clear.

This paragraph is intend to reminding and discussing the assumptions made when using a (calibrated) rainfall-runoff model over a climatically-contrasted and long period of time. We thus talk about the parameter set of the rainfall-runoff model, obtained after a calibration over a short period (here 17 years). Numerous authors thus highlighted that calibrated parameter sets are dependent on the climate of the calibration period and that the rainfall-runoff models show limited performances when applied over periods that are climatically contrasted regarding to the climate of the calibration period. It is clearly out of the scope of this paper to quantify the sensitivity of the streamflow reconstruction to these "stationary" assumptions, but it is an interesting perspective of this work. We will reword this paragraph to be clear.

43. Line 859: change "focusing" to "focus"

Agreed.

*44. Figures 3, 4, 5, 6, 8, 9: I have a very difficult time making out the 5*20 ANATEM or 5*20 ANA data in these figures. I cannot distinguish 5*ANA from 5*20 ANATEM in Figure 3. I suggest the authors consider using some different colors for these lines or symbols.*

We agree that several lines are impossible to see or to distinguish on these figures. Even if this is a significant and interesting result (meaning that there is no dispersion between simulations or no difference between the observation and the simulation), we will change the colors and the point types and details the differences of the lines in the figure captions in order to distinguish the different simulations.

45. Figure 7: Suggest moving "(a)" before "mean annual streamflow" and "(b)" before "May monthly"

Agreed.

46. Figure 9: Should the reference to Nicault et al. (2014) in the caption actually be to Boucher et al. (2011)?

Yes, we will correct this mistake.

2 REFEREE #2

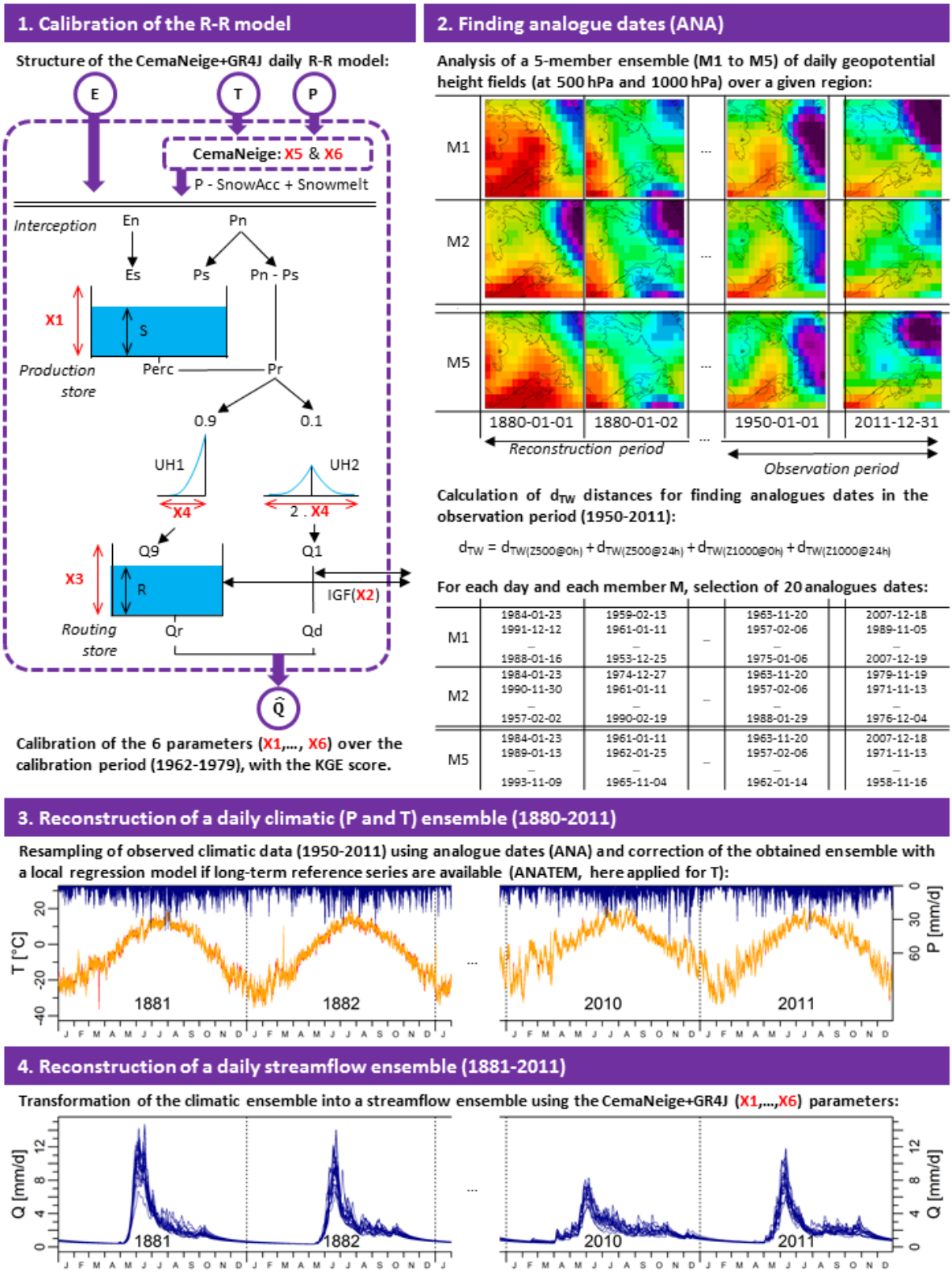
2.1 General comments

The paper is well written and is in a form very similar to other paper on the paleoclimate text. It is rather long, but using several methods, this is necessary to present everything. Nevertheless, there is not always justification of the choices. For instance the choice of the zone used for the geopotential is not justified. And some parameters for the different models are not explicit. If possible it would be nice to integrate them in a way, but I know it is an issue because the paper will be longer. Because it is a long paper using several concepts, I would recommend to the author to summarize in a flow-chart figure each step of their methodology to reach streamflow. It would make it easier for the reader to follow the whole text. If the author can take this remarks into account, the paper will be nearly ready for publication.

The tests performed for choosing the spatial domain considered for the geopotential height field will be presented in a new Appendix part of the manuscript.

We will also add several paragraphs in order to fully describe the rainfall-runoff model and its snowmelt routine and how are calibrated the parameters (cf. answers to Referee #1).

Moreover, adding a flowchart summarizing the entire methodology will definitively improve the manuscript, we would like to thank Referee #2 for this suggestion. We produced a new figure that we will add to the manuscript in order to summarize the reconstruction methodology applied:



2.2 Specific comments

Fig 1: I do not recognize the catchment on figure 1b? why?

The studied catchment is one of the 211 cQ2 catchments is thus plotted in the Figure 1b, but is hidden by an intermediate sub-catchment. The Caniapiscau catchment will be highlighted in the manuscript with shading lines.

Page 3 line 4 add reference after "dendrohydrology".
We will add the reference to the review of Loaiciga *et al.* (1993).

Legend figure 4: add "for" 1950??
Agreed.

Page 15 line 1: blank after the dot.
We will add a space.

Figure 9: I do not understand tree ring reference to Nicault and Boucher in b?
The "tree-ring series" presented in the Figure 9 is from Boucher *et al.* (2011), we will thus correct the Figure 9 legend.

3 REFERENCES

Boucher, É., Ouarda, T.B.M.J., Bégin, Y., Nicault, A., 2011. Spring flood reconstruction from continuous and discrete tree ring series. Water Resour. Res. 47, W07516. doi:10.1029/2010WR010131

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology 377, 80–91. doi:16/j.jhydrol.2009.08.003

Kuentz, A., Mathevet, T., Gailhard, J., Hingray, B., 2015. Building long-term and high spatio-temporal resolution precipitation and air temperature reanalyses by mixing local observations and global atmospheric reanalyses: the ANATEM model. Hydrol. Earth Syst. Sci. 19, 2717–2736. doi:10.5194/hess-19-2717-2015

Loaiciga, H.A., Haston, L., Michaelsen, J., 1993. Dendrohydrology and long-term hydrologic phenomena. Reviews of Geophysics 31, 151–171.

Teweles, J., Wobus, H., 1954. Verification of prognosis charts. Bulletin of the American Meteorological Society 35, 455–463.

Way, R.G., Viau, A.E., 2014. Natural and forced air temperature variability in the Labrador region of Canada during the past century. Theor Appl Climatol 121, 413–424. doi:10.1007/s00704-014-1248-2