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 and markep-up manuscript version

Comments and suggestions made by the two referees are gratefully acknowledged. We modified the text in response to the main criticisms. In the following, we list the referee comments (in *italic and blue*), we provide specific responses to these comments (in black) and finally we present a marked-up manuscript version.

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 has ben added in the section 3 (*Methodology*), with a focus on the inputs needed and on the timestep of the model:

The structure of the CemaNeigeGR4J model is presented in the Figure 3. GR4J is based on two non-linear stores (production and routing stores) and a unit-hydrograph, while CemaNeige is a degree-day snow accounting routine, which divides the studied catchment into five elevation bands. CemaNeigeGR4J uses as inputs daily series of precipitation, minimal and maximal air temperatures and a daily potential evapotranspiration series, calculated using Oudin et al. (2005) formula, designed for rainfall-runoff modelling. CemaNeigeGR4J produces daily streamflow series.

A table has also been added in this section (Table 1), giving a description of each of the six calibrated parameters, their unit and their final calibrated values.

Although there is an entire result subsection devoted to the rainfall-runoff model calibration performances (*4.1 Rainfall-runoff model calibration performances*), we added the calibration metric values obtained after calibration (Kling and Gupta Efficiency score (Gupta *et al.*, 2009), KGE = 0.93).

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 is now clearly stated in the manuscript (see previous answer).

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 article goal, which is to compare the new streamflow reconstruction with two other reconstructions (using tree-rings) available at the annual resolution. Nevertheless, as detailed in the previous answers, outputs of the reconstruction methodology are available at the daily resolution. We now evaluate the performance of the climatic reconstruction at the daily timestep and present it alongside the monthly and annual performances (see for example Figure 6).

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 is now clearly defined in the manuscript, in section 2.1 (*Datasets used for the climatic reconstructions*):

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 (example from the NOAA's National Weather Service). Note that for pressure levels close to sea level (typically 1000 hPa), the geopotential height can sometimes be negative. The analysis of geopotential height *fields over a given domain describes the spatial distribution of high/low pressure systems upon which similarity in between days can be measured.*

Also, we now 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 now explicitly state that we compare the streamflow series reconstructed in this article with two streamflow series obtained with tree ring data by other authors:

In this paper, we applied a new hydro-climatic reconstruction method on the Caniapiscau Reservoir and compare the obtained streamflow time series against time series derived from dendrohydrology by other authors on the same catchment and study the natural streamflow variability over the 1881-2011 period in that region.

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 (now 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 now 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 was unclear, and we thus rephrased 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.

*Of the 56 ensemble members constituting the 20CR reanalysis, the members 1 to 5 were extracted and used over this region (see section 3.1.1**Erreur ! Source du renvoi introuvable.** for more details).*

8. Lines 203-204: Keep the greater than sign (>) with the numbers (i.e., >100)

Agreed, we now use 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 were missing in this sentence and we thus reworded it:

For the air temperature, the Berkeley Earth Surface Temperature (hereafter denoted as BEST) analysis has been used, taken from the http://berkeleyearth.org/ Web site (Rohde et al., 2013).

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.

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 changed 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 rephrased this sentence in order to be clearer:

The ANATEM method (Kuentz et al., 2015) is built on the combination of two approaches: (i) the ANA (which stands for "ANAlogue") approach, that aims to find, for a given day, a given number of analogue days, based on the similarity of synoptic circulation (Obled et al., 2002) and (ii) the TEM (which stands for "TEMoin", the French word for "witness") approach, which is a basic regression model that uses a continuous and long-term reference (the witness) climatic series to reconstruct past climate.

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 added some details on how the Teweles and Wobus (1954) distance is calculated (the formula and an example of the calculation) in a new Appendix part of the paper (Appendix B). The "thanks to" has been deleted. Finally, we now explicitly state that 20 time series are created for each considered members:

For each day, the 100 climatic values are obtained based on the 20 "closest" analogue days for each of the 5 20CR members considered.

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

Yes, we added 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 has been 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 moved 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 now explicitly state that a perfect KGE value is 1:

The KGE criterion ranges between -∞ and 1 (perfect simulation).

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 corrected 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 corrected 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 are clearly defined in the text (in section 3.5 *Comparison of reconstructed series against observations*) and are now distinguishable in the figure, to avoid any confusion:

In order to compare the reconstructed streamflow time series against observations, the reconstructed ensembles were first aggregated: a daily series was generated for each of the five 20CR members considered by averaging the 20 daily series constituting each ensemble. The five daily mean series are denoted as $\overline{ANA}$ *or* $\overline{ANATEM}$*, depending on the method used to produce them.*

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 now use the "resolution" term in the entire manuscript.

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

We reworded this sentence:

Averaging each ensemble of the considered 20CR members (blue points) results in better temporal correlations at the daily and yearly resolutions, but at the expense of lower variability reproduction performance.

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 changed 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 changed 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 now explicitly state so in the manuscript:

All the rainfall-runoff model outputs presented in the manuscript have been produced at the daily resolution by using both GR4J rainfall-runoff model and its snowmelt routine CemaNeige.

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.

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 was 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 changed this descriptive term in the manuscript:

Another significant difference exists over the 1950-1960 period, seen as an common decade by the tree ring reconstruction (reconstructed spring flood ranging from 47 to 87 [mm/m]), while being seen as a highly variable hydrological decade for the ANATEM reconstruction, with high values for the first five years (around 110 [mm/m] for the 1950-1955 period), and then two very low values (around 20 [mm/m] for the 1956-1957 period), finally followed by three high value years (around 110 [mm/m] for the 1958-1960 period).

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 added 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):

Finally, the reconstructed climatic time series are transformed into streamflow time series thanks to a daily rainfall-runoff model, previously calibrated over the relatively short observation period (with really good calibration performances). The use of one model, one objective function and one parameter set is questionable. Quantifying the sensitivity of the obtained reconstruction to the hydrological modeling assumptions made was out of the scope of this paper, but definitively deserves further research, especially considering the issue of uncertainty due to rainfall-runoff model parameters in a changing climate

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:

The inability of the analogue approach to reproduce the interannual precipitation variability - already highlighted by Kuentz et al. (2015) over 22 French catchments – is due to the absence of a local reference climatic time series, unlike for temperature reconstruction, where a local temperature time series is used, and ensures that the simulated interannual temperature variability is reproduced efficiently. Finding an additional series which significantly improves the precipitation reconstruction is a major perspective of this work. The use of variables produced by the available reanalyses (e.g., relative humidity, precipitable water content) for finding analogue dates will be investigated, along with the testing of time series of local pressure measurements. For example, Caillouet et al. (2016) showed that adding the sea surface temperature variable to the temperature, geopotential, vertical velocity and humidity for finding analogue dates significantly improves the reconstruction of air temperature and precipitation over France.

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) are now presented in a new Appendix part added to the manuscript (Appendix A). 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 (now Figure 9): 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 now explicitly point the reader to this figure in the manuscript:

Interestingly, the uncertainty due to the use of five members of the 20CR reanalysis appears to be limited, and even null from 1940 onward. See for example Figure 9 which presents the centennial ANATEM streamflow reconstructions: it is impossible to distinguish the five ANATEM average series after 1940, highlighting that considering five different members of the 20CR reanalysis as inputs of the reconstruction method has a negligible impact on the reconstruction of the mean annual streamflow.

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

Yes, we added "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 reworded this paragraph to be clear:

Finally, the reconstructed climatic time series are transformed into streamflow time series thanks to a daily rainfall-runoff model, previously calibrated over the relatively short observation period. The use of one model, one objective function and one parameter set is questionable. Quantifying the sensitivity of the obtained reconstruction to the hydrological modeling assumptions made was out of the scope of this paper, but definitively deserves further research, especially considering the issue of uncertainty due to rainfall-runoff model parameters in a changing climate. Thus, numerous authors highlighted that calibrated parameters of rainfall-runoff models are dependent on the climate of the calibration period and that performance decreases when applied over periods where the climate differs from that of calibration period (e.g., Merz et al. 2011; Coron et al. 2012 and Brigode et al. 2013b). Thus, testing different calibration strategies (e.g., bootstrap calibration used by Brigode et al. 2015), testing particular objective functions especially devoted to the final study objective (e.g., studying mean annual streamflow), and adapting the time step of the rainfall-runoff model to the objective would be interesting for future works.

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 changed the colors between ANA and ANATEM (in Figures 5, 6 and 7) 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 corrected 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 are now presented in a new Appendix part of the manuscript (Appendix A).

We also added 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 general comments).

Moreover, wed added a flowchart summarizing the reconstruction methodology applied (Figure 3).

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 was hidden by an intermediate sub-catchment. The Caniapiscau catchment is now highlighted in the Figure 1b with shading lines.

Page 3 line 4 add reference after "dendrohydrology".

We added 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 added 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 thus corrected the Figure 9 (now Figure 10) legend.

3 CITED 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

4 MARKED-UP MANUSCRIPT VERSION

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

Brigode, P.[1,2,*]; Brissette, F.[1] ; Nicault, A.[3]; Perreault, L.[4],; Kuentz, A.[5]; Mathevet, T.[6] & Gailhard, J.[6]

[1] École de technologie supérieure de Montréal, Montreal, Canada.

[2] Ouranos, Montreal, Canada.

[3] ECCOREV, Aix-en-Provence, France.

[4] IREQ, Varennes, Canada.

[5] SMHI, Norrköping, Sweden.

[6] EDF, DTG, Grenoble, France.

* Now in: Université Côte d'Azur, CNRS, OCA, IRD, Géoazur

Correspondence to: pierre.brigode@unice.fr

Abstract:

Over the last decades, different methods have been used by hydrologists to extend observed hydro-climatic time series, based on other data sources, such as tree rings or sedimentological datasets. For example, tree ring multi-proxies have been studied for the Caniapiscau Reservoir in northern Quebec (Canada), leading to the reconstruction of flow time series for the last 150 years. In this paper, we applied a new hydro-climatic reconstruction method on the Caniapiscau Reservoir ~~to~~ and compare the obtained streamflow time series against time series derived from dendrohydrology by other authors on the same catchment ~~obtained streamflow series~~ and study the natural streamflow variability over the 1881-2011 period in that region. This new reconstruction is based, not on natural proxies, but on a historical reanalysis of global geopotential height fields, and aims firstly to produce daily climatic time series, which are then used as inputs to a rainfall-runoff model in order to obtain daily streamflow time series. The performances of the hydro-climatic reconstruction were quantified over the observed period, and showed good performances, both in terms of monthly regimes and interannual variability. The streamflow reconstructions were then compared to two different reconstructions performed on the same catchment by using tree ring data series, one being focused on mean annual flows, and the other one on spring floods. In terms of mean annual flows, the interannual variability of the reconstructed flows were similar (except for the 1930-1940 decade), with ~~significant~~ noteworthy changes seen in wetter and drier years. For spring floods, the reconstructed interannual variabilities ~~reconstructed~~ were quite similar for the 1955-2011 period, but ~~significantly~~ strongly different between 1880 and 1940. The results emphasize the need to apply different reconstruction methods on the same

catchments. Indeed, comparisons such as those above highlight potential differences between available reconstructions, and finally, allow a retrospective analysis of the proposed reconstructions of past hydro-climatological variabilities.

1 INTRODUCTION

1.1 Challenge of decadal hydrological variability

Time series of ~~S~~streamflow ~~series~~ observations, which constitute the basis for all hydrological analyses, are generally characterized by a relatively short record period, typically ranging from several years to several decades. ~~Thus~~In fact, the average length of 6945 daily streamflow series collected by the Global Runoff Data Center, and available worldwide, is 44 years (GRDC, 2015). The information extracted by hydrologists from these time series (in the form of statistical indices, calibration of model parameters, etc.) are generally used for water resource management, ~~in the form~~for instance for ~~of the~~ hydropower generation mid- to long-term planning~~, for instance~~. The short record period is a major issue for hydrologists since it may be insufficient to capture and provide a~~n~~ clear understanding of the decadal variability of hydrological processes. For example, after studying a 90-year long daily streamflow series of the Po River (Italy), and highlighting significant natural variability at the decadal scale, Montanari (2012) stated that “more research efforts are needed to improve the interpretation of such long-term fluctuations”.

Studying natural variability requires long instrumental records (typically longer than 100 years), but such long time series are non-existent in remote regions such as northern Quebec (Canada).~~ : t~~The length (number of years) of 221 observed streamflow time series from Quebec - ~~(~~extracted from the ~~(~~cQ~~)~~2 (*Impact des Changements Climatiques sur l’hydrologie (Q) au Québec*) database~~,~~ (Guay et al., 2015) - is shown in Figure 1~~Figure 1~~b and c, highlighting that very few series have more than 50 years of data. Hydrological decadal variability is crucial in this region, since it is home to some of the largest hydropower systems in the world; as well, significant inter-annual inflow variability has been recorded in several Quebec catchments (e.g.~~, for annual flows, by~~ Perreault et al., 2000 and 2007; Jandhyala et al., 2009). The few decades of observations available for this region are not sufficient to allow a robust analysis of multi-decadal hydrological variability, and thus, raise the issue of the reconstruction of past hydrology, i.e., occurring before the systematic recording of streamflows.

1.2 Reconstruction of past hydrology

Over the past decades, different methods have been used by hydrologists to reconstruct natural flows on catchments of interest, depending on available data. These methods may be classified into two groups, according to the ~~time steps~~temporal resolution of the reconstructed series.

The first group brings together the methods based on long and continuous hydro-climatic series constructed with daily or sub-daily observations, and consequently, allowing the reconstruction of streamflow time series at a fine temporal scale (e.g., daily ~~time step~~resolution). When long streamflow series are available for other catchments close to the one under study, classical statistical regressions or other regionalization methods could be applied for the reconstruction (e.g., Hirsch (1982), Hernández-Henríquez et al. (2010), and Arsenault & Brissette (2014)). The paired catchment approach - consisting ~~in~~ of calibrating and then using a streamflow-streamflow model - could also be used (e.g., Andréassian et al., 2012). When long climatic series (typically covering precipitation and temperature) are available in the studied region, the reconstruction could be done by using a rainfall-runoff model, in order to transform the climatic series into streamflow series (e.g., simulation of 124 years of streamflow for the Thames River (UK) by Crooks & Kay, 2015).

The second method is based on continuous or discrete series of paleo-indicators, generally producing reconstructed series at seasonal or annual ~~time steps~~resolutions (Bradley, 1999). The most natural proxies used for hydrological reconstructions are sediment stratigraphy (e.g., Thorndycraft et al., 2005) and tree ring series (see reviews by Loaiciga et al. (1993) and Meko & Woodhouse (2011)). This latter proxy for streamflow reconstruction, referenced as dendrohydrology (Loaiciga et al., 1993), is analyzed in a bid to reconstruct past hydro-climatological variations of a given catchment by studying tree ring width variations among different trees sampled in the same region. Reconstructed streamflow series are obtained by applying either direct or indirect methods. The direct methods aim to link tree ring series with streamflow series through statistical models calibrated over an observation period (e.g., in Tasmania (Australia) by Allen et al., 2015 and in the southeastern United States by Patskoski et al., 2015). The indirect methods aim firstly to reconstruct climatic series, such as temperature or precipitation, and secondly, to transform these climatic series into streamflow series through rainfall-runoff models (e.g., in the Western US by Gray & McCabe (2010) and Saito et al. (2015)). These methods allow the continuous reconstruction of the annual or seasonal water balance of a given region, over long time periods. Additionally, other information could be extracted following tree ring analysis and used to reconstruct discrete chronologies of extreme hydrological events. For example, George & Nielsen (2003) used anatomical tree ring signatures to reconstruct paleofloods of the Red River in Manitoba (Canada).

Recently, dendrohydrological methods have been successfully applied in boreal environments, characterized by a rarity of long hydro-climatological series. For example, Nicault et al. (2014) used tree ring multi-proxies (tree ring widths, tree ring densities and tree ring stable isotope ratios) to produce spring, summer and annual flow series of the Caniapiscau Reservoir in northern Quebec (Canada) for the 1800-2000 period. On the same catchment, Boucher et al. (2011) used both continuous series (tree ring minimal density measurements) and discrete series (with ice-scars due to ice abrasion during floods) to produce spring flood series for the 1850-1980 period. These two reconstructions revealed significant flow variability in this region, both in terms of annual flows and flood frequency. It should be noted that the Caniapiscau Reservoir is the most upstream and one of the largest reservoirs of the La Grande complex, which is one of the biggest hydro-power generation complexes in the world, with a total~~n~~ installed generating capacity of 17,418 ~~15,240~~ megawatts. Decadal hydro-climatological variability in this region thus provides important information concerning the long-term planning of hydro-power generation.

1.3 Scope of paper

Although the above-mentioned hydrological reconstructions were associated with good verification statistics on the calibration period, the lack of observed streamflow data did not allow a rigorous independent verification of those reconstructions. An alternative solution involved carrying out new reconstructions based on different proxies and different methods, and then, as an additional verification step, analyzing the consistency between the different reconstructions. Comparisons of streamflow reconstruction methods are rare in the literature, and the Caniapiscau Reservoir catchment offers an interesting case study since various tree ring reconstructions have been performed there. Thus, our objective is to apply a new reconstruction method on the Caniapiscau Reservoir, in order to compare the obtained streamflow series with series obtained by dendrohydrology and to study the observed streamflow variability over the 1881-2011 period. This new reconstruction is based, not on natural proxies, but on a historical reanalysis of geopotential height fields. A climatic ensemble was ~~reconstituted~~ reconstructed at the daily ~~time step~~resolution using the ANATEM methodology (Kuentz et al., 2015), a resampling method based on synoptic situation similarities between days (found by looking at the geopotential height reanalysis), with a sampling of observed climatic series for a given time period (the *observation period*) over a longer time period (the *reconstruction period*). ~~combining large-scale atmospheric information (geopotential height reanalysis) with local climatic observations (reference climatic series).~~ Then, a rainfall-runoff model - previously calibrated on the observed period - was used to transform this climatic ensemble into a streamflow ensemble. The performances of the hydro-climatic reconstructions and of the rainfall-runoff model calibration were firstly evaluated over the observed period, by comparing the reconstructions and the simulations with the observations. Secondly, the tree ring based on the ANATEM centennial reconstructions were compared, and finally, the long-term hydrological variability of the Caniapiscau Reservoir was discussed.

2 DATA

2.1 Datasets used for the climatic reconstructions

2.1.1 Geopotential height reanalysis

The climatic reconstruction method applied in this study (fully detailed in the following section) is based on finding similarity between days at the synoptic scale. The similarity is based on geopotential height fields over a given spatial domain. 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 (example from the NOAA's National Weather Service). Note that for pressure levels close to sea level (typically 1000 hPa), the geopotential height can sometimes be negative. The analysis of geopotential height fields over a given domain describes the spatial distribution of high/low pressure systems upon which similarity in between days can be measured. Several long-term geopotential height reanalysis have been produced during the last decade, in order to study climate variability and climate change over the last century.

The geopotential height reanalysis used in this study was drawn from the 20th Century Reanalysis V2c data, provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, available from their Web site at http://www.esrl.noaa.gov/psd/ (Compo et al., 2011). This global reanalysis (hereafter denoted as 20CR), assimilating only surface observations of synoptic pressure, monthly sea surface temperature, and sea ice distribution, spans the period of 1851 to 2011, with a six-hourly temporal resolution and a 2° spatial resolution. For each day, two levels were considered here, 1000 hPa at 0h and 500 hPa at 0h. The geopotential height fields were extracted over an area covering the entire province of Quebec, with 221 grid points, as shown in Figure 1~~Figure 1~~a. Of the 56 ensemble members constituting the 20CR reanalysis, ~~the 5 first~~the members 1 to 5 were extracted and used over this region (see section 3.2 for more details).

2.1.2 The quest for centennial climatic series in northern Canada

Centennial and continuous climatic series are rare in Canada, and almost non-existent in remote high-latitude regions, such as northern Quebec (Cowtan & Way, 2014). In this ~~our~~ study, there is a need for both consistent and very long (> 100 years) climatic series. Vincent et al. (2012) and Mekis & Vincent (2011) built two databases of “adjusted and homogenized” air temperature and precipitation series, respectively, both available at monthly and daily ~~time steps~~resolutions for all of Canada. These databases were specifically created for use as references in climate change impact studies. During their creation, care was taken to correct any errors that may surface, and to account for any shifts that may occur as a result of stations being moved or of changes in measurement instruments that may be present in the climatic series observed. Nevertheless, the average length of such series in northern Quebec is 50 years, which is considered too short for this work or for any study concerning natural climatic variability.

In Quebec, the few long climatic series (> 100 years) available are generally for large cities, which are all located in the southern part of the province. These series are rarely continuous at the daily time scale, and are derived from different sources; as a result, producing good quality continuous series therefore requires a lot of work. For example, Slonosky (2014) compiled data from numerous sources (mainly from the cities of Québec and Montreal) to produce continuous daily temperature series for the St. Lawrence Valley region for the 1798-2010 period. In northeastern Canada, two sources of such historical data exist. First, the Moravian missionaries, who have been living among the Inuit in the Labrador coastal region since 1771, have measured and recorded climatic variables (Demarée & Ogilvie, 2008). Secondly, interesting qualitative information for the Hudson Bay and the James Bay (northwestern Quebec) 19th century climate are present in the Hudson’s Bay Company trade post journals. Wilson (1983) compiled these data and produced summer temperature series and a wetness index for this region, and the series was then used by Bégin et al. (2015) as a reference series for comparisons with their climate reconstruction of the Canadian northeastern boreal forest. Unfortunately, no such data sources are present in the interior part of northern Quebec.

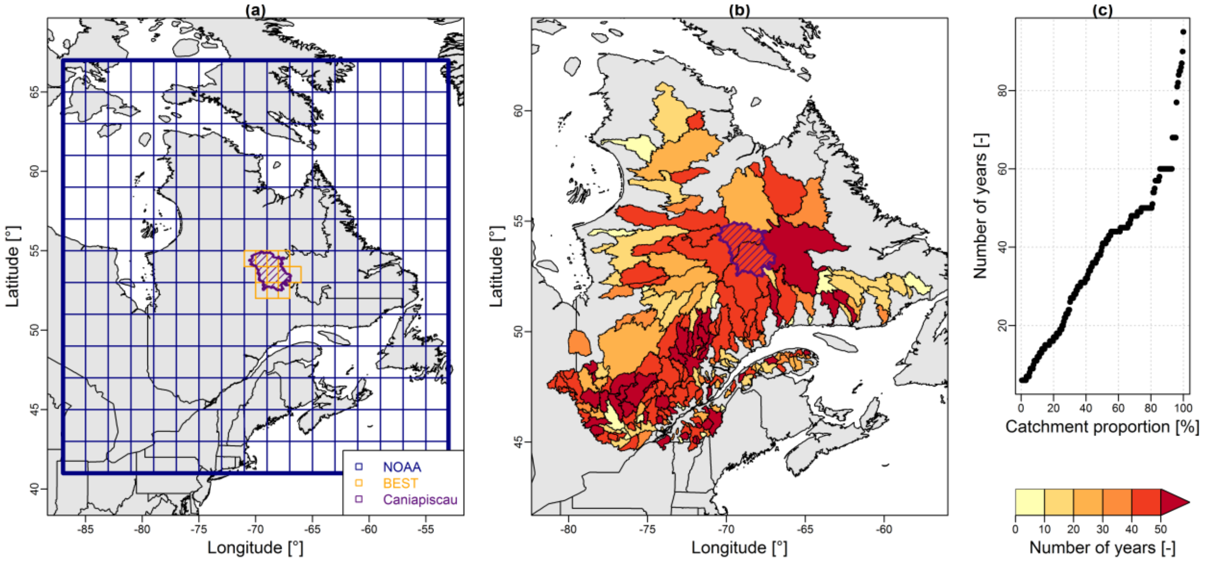

Figure 1. (a) Datasets used for the hydro-climatic reconstruction: the extension of the 20CR grid points used is shown in blue, while the BEST grid points used are highlighted in purple~~orange~~. The Caniapiscau reservoir catchment is plotted in red. (b) Spatial distribution and (c) distribution of the length (number of years) of the observed streamflow series for 211 catchments in Quebec, extracted from the cQ2 database, Guay et al. (2015).

2.1.3 A reanalysis as local reference temperature series

For the air temperature, the Berkeley Earth Surface Temperature (hereafter denoted as BEST) analysis has been used, taken from the http://berkeleyearth.org/ Web site (Rohde et al., 2013)~~; the site provides ~~. BEST is a gridded air temperature reanalysis for lands, starting in 1753 at the monthly ~~time step~~resolution, and in 1880 at the daily ~~time step~~resolution, with a 1° spatial resolution. A daily catchment series has been assembled for the 1880-2011 period by averaging the 11 BEST grid points covering the Caniapiscau reservoir catchment, highlighted in Figure 1~~Figure 1~~. Note that this reanalysis was recently used in northeastern Canada by Way and Viau (2014), in their study of past air temperature variability in ~~New-Brunswick~~Labrador.

2.2 Caniapiscau reservoir catchment

In Quebec, ~~97~~99% of the produced electricity comes~~is coming~~ from hydropower generation systems. The La Grande ~~operational chain~~water resources system, located in northern Quebec and operated by Hydro-Québec (HQ), is one of the most important hydropower systems in the world, with an installed capacity of 17,418 megawatts (the Three Gorges Dam is the most important hydropower system in the world with a total installed capacity of around 22 000 megawatts).~~ and~~ This system produces 50% of the total energy generated by HQ. The Caniapiscau hydroelectric reservoir catchment is the first dam of the La Grande operational chain (the Brisay power plant installed at the outlet of the Caniapiscau reservoir is ranked as the 9th with an installed capacity of around 500 megawatts) and is a 37,328 km² snowmelt-dominated catchment. Figure 2~~Figure 2~~ illustrates the hydro-climatic context of the Caniapiscau reservoir catchment. The catchment elevation (SRTM data, Jarvis et al., 2008) ranges from a~~b~~round 500 to 900 m a.s.l., with the highest elevation areas located in the southern parts of the catchment.

The daily streamflow series (a) and the monthly regimes (c) show the strong snow-dominated signature of the catchment, with an annual flood observed due to snowmelt during the month June. On average, the mean annual precipitation and runoff are around 800 mm (with around 300 mm falling as snow~~; the snow mean annual series is plotted in light blue in Figure 2b~~) and 650 mm, respectively, on the Caniapiscau reservoir, and the mean annual temperature is around -3.6 °C.

Catchment climatic data used in this study consists of daily series of minimum, mean and maximum air temperature and of total precipitation, available for the 1950 to 2011 period. This dataset was produced by ~~Hydro-Québec~~HQ, using kriging methods (Tapsoba et al. 2005). Daily streamflow series are available from 1962 to 2011. Note that only the 1962-1979 period was considered for the rainfall-runoff model calibration here, since the Caniapiscau Dam was built during the 1980-1982 period, and streamflow series available for 1982 to 2011 are naturalized flows produced by ~~Hydro-Québec~~HQ. Nevertheless, this second period (1982-2011, mean annual values are plotted in grey in Figure 2~~Figure 2~~b) will be used as a validation period for the reconstruction.

2.3 Reconstructed yearly streamflow series from tree rings

Two yearly time series of Caniapiscau Reservoir flows have been used here for comparison at the centennial scale: (i) the series of annual flows proposed by Nicault et al. (2014) for the 1800-2000 period, and (ii) the series of spring floods proposed by Boucher et al. (2011) for the 1850-1980 period. The first yearly time series was processed from continuous tree ring series derived from 20 black spruce (*Picea mariana* [Mill.] BSP) sites located within 200 km around the Caniapiscau reservoir. Two reconstruction methods were used (Partial Least Square regression (PLS) and Best Analogue Methods), and the reconstructions obtained were combined in a single composite reconstruction. The second yearly series was processed from ice-scar time series derived from a small lake located next to the Caniapiscau reservoir and using tree ring densities obtained from 12 black spruce sites. A new transfer model technique based on Generalized Additive Model (GAM) theory was used to process spring flood reconstructions.

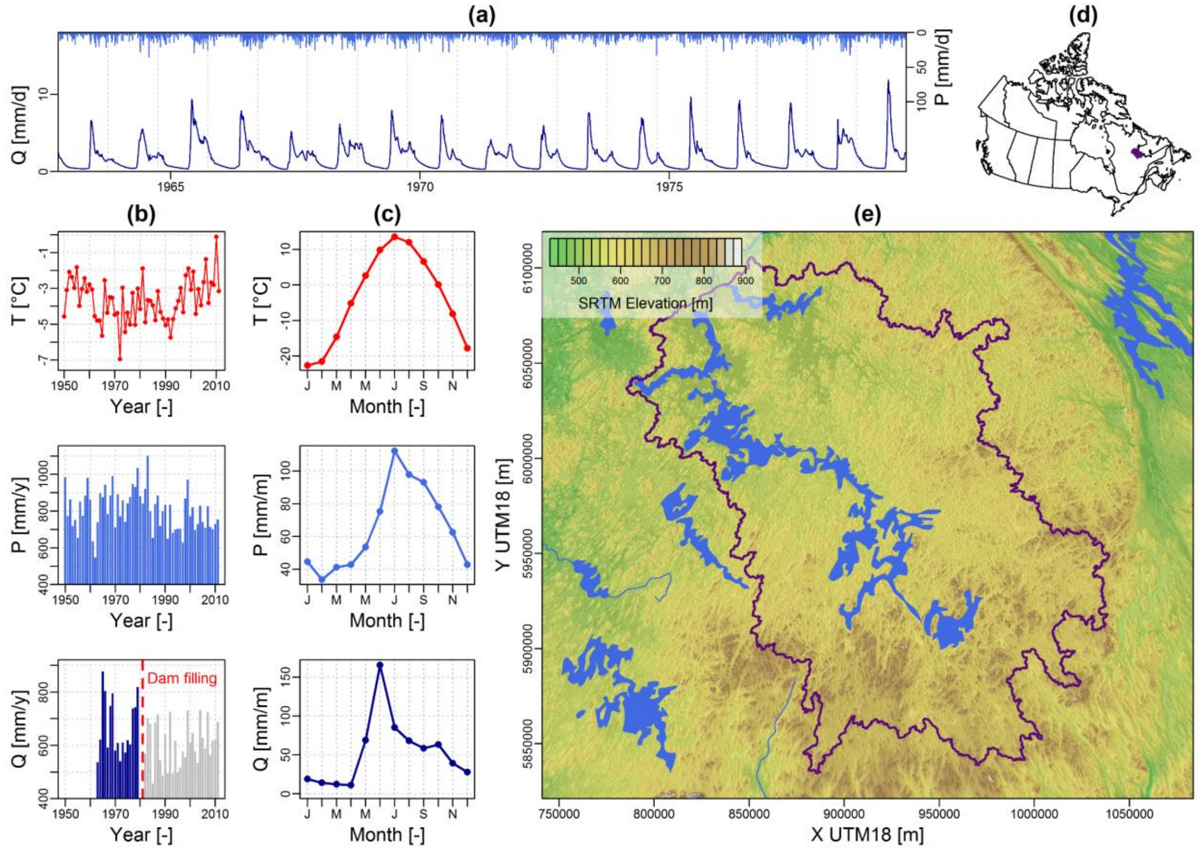

Figure 2. Hydro-climatic context of the Caniapiscau reservoir catchment: (a) observed daily streamflow and precipitation time series used for the rainfall-runoff model calibration (1962-1979), (b) temperature, precipitation and streamflow mean annual series, (c) temperature, precipitation and streamflow monthly regimes, (d) catchment location within Canada, and (e) SRTM elevation data. Monthly regimes were calculated for the 1950-2011 period for temperature and precipitation, while for the 1962-1979 period, the calculations were for streamflow.

3 METHODOLOGY

3.1 General streamflow reconstruction methodology

The general methodology consists in the reconstruction of an ensemble of daily climatic time series (with the ANATEM method) and of the transformation of this daily climatic ensemble into a daily streamflow ensemble, using a rainfall-runoff model.

The ANATEM method (Kuentz et al., 2015) is built on the combination of two approaches: (i) the ANA (which stands for “ANAlogue”) approach, that aims to find, for a given day, a given number of analogue days, based on the similarity of synoptic circulation (Obled et al., 2002) and (ii) the TEM (which stands for “TEMoin”, the French word for “witness”) approach, which is a basic regression model that uses a continuous and long-term reference (the witness) climatic series to reconstruct past climate. The ANATEM method thus allows the reconstruction of the climate of the past by combining synoptic information (ANA approach) with local climatic observations (TEM approach). Finally, this method allows the production of an ensemble of daily climatic time series by the selection of several analogues for any given day. For a complete description of the ANATEM method and an evaluation of its performance at the regional scale (French Alps), see Kuentz et al. (2015).

The rainfall-runoff transformation is done here with GR4J (Perrin et al., 2003), a daily lumped continuous rainfall-runoff model and its snowmelt routine, CemaNeige (Valéry et al., 2014). GR4J and CemaNeige have 4 and 2 free parameters to calibrate, respectively, using the observed streamflow data available on the studied catchment.

The whole streamflow reconstruction methodology - performed in the R-project environment (2014, http://www.r-project.org/) - is carried out in four steps (see Figure 3~~Figure 3~~):

- **Step 1: calibration of the rainfall-runoff (R-R) model.** The rainfall-runoff model is calibrated on the observed streamflow data.
- **Step 2: finding analogue dates (ANA)**. Synoptic states are compared in order to find analogue days for each day of the reconstruction period, amongst the days of the observation period.
- **Step 3: reconstruction of a daily climatic (P and T) ensemble (ANATEM)**. The best analogue obtained at step 2 are stochastically resampled and long-term reference climatic series are used (if available) to improve the resampled series.
- **Step 4: reconstruction of a daily streamflow ensemble.** The climatic ensemble is transformed into a streamflow ensemble using the rainfall-model parameter set obtained at step 1.

These four steps are further detailed hereafter.

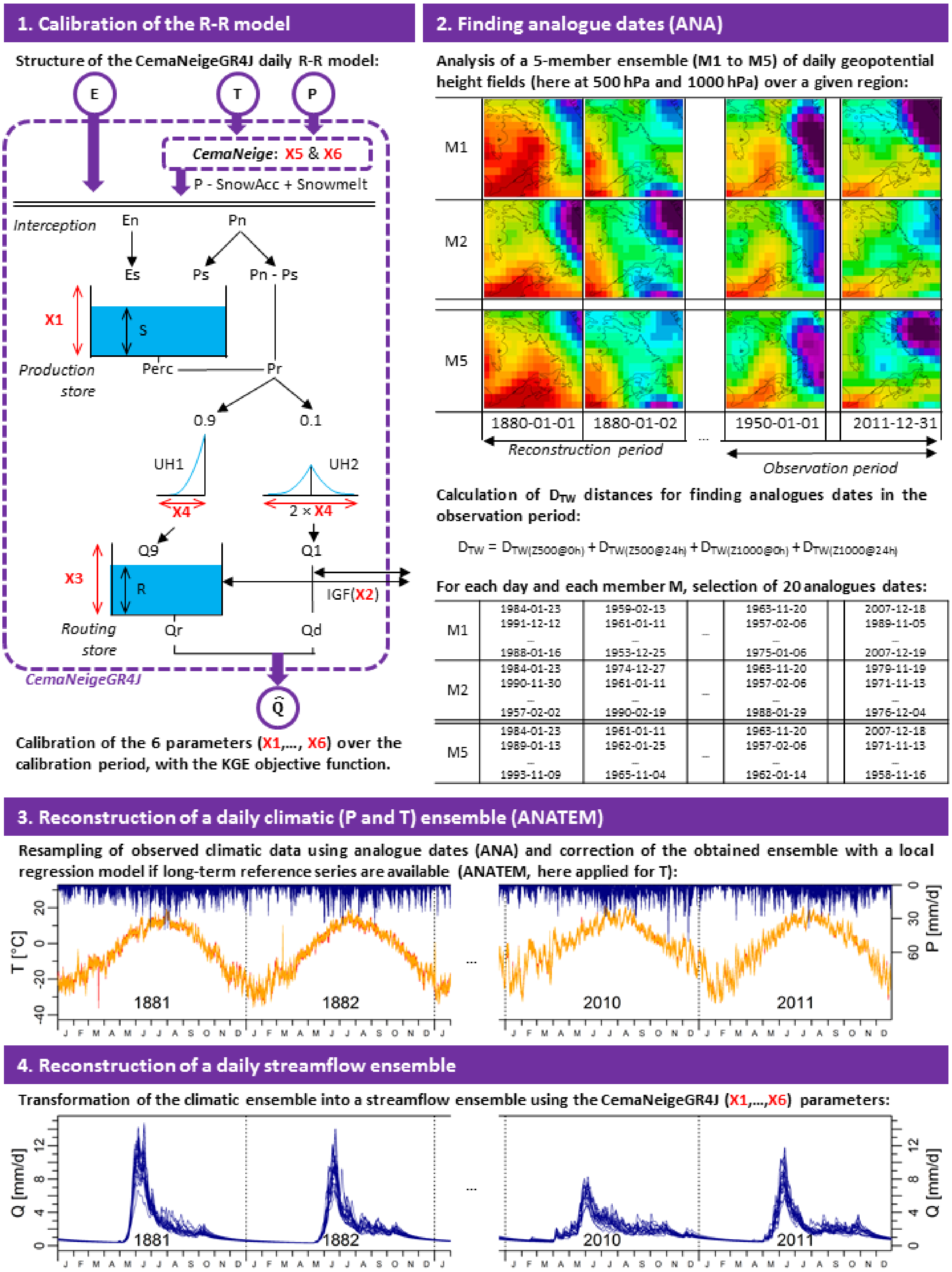

Figure 3 : Illustration of the four-step methodology used for the reconstruction of a daily streamflow ensemble (R-R stands for rainfall-runoff, E for potential evapotranspiration, T for air temperature, P for precipitation, Q for streamflow).

3.1 ~~Rainfall-runoff modelling~~Step 1: calibration of the rainfall-runoff model

The GR4J (Perrin et al., 2003) rainfall-runoff model was used to transform the climatic ensemble into an ensemble of streamflow time series. GR4J is an efficient and parsimonious (only four free parameters to be calibrated) daily lumped and continuous model, which, when it is combined with its snow accumulation and melt ~~snowmelt~~ routine, CemaNeige (Valéry et al., 2014), is well suited for the hydrological modeling of snow-dominated catchments. GR4J and CemaNeige (model-pair hereafter denoted as CemaNeigeGR4J) were recently evaluated over several catchments located in Quebec (e.g., Seiller et al., 2012; Valéry et al., 2014) and showed good modelling performances.

The structure of the CemaNeigeGR4J model is presented in the Figure 3~~Figure 3~~. GR4J is based on two non-linear stores (production and routing stores) and a unit-hydrograph, while CemaNeige is a degree-day snow accounting routine, which divides the studied catchment into five elevation bands. CemaNeigeGR4J uses as inputs daily series of precipitation, minimal and maximal air temperatures and a daily potential evapotranspiration series, calculated using Oudin et al. (2005) formula, designed for rainfall-runoff modelling. CemaNeigeGR4J produces daily streamflow series.

GR4J and CemaNeige have 4 and 2 free parameters to calibrate, respectively. These 6 parameters - highlighted in Figure 3~~Figure 3~~ and described in Table 1~~1~~ - ~~These 6 parameters~~ were calibrated ~~together~~ conjointly over the same calibration period ~~(1962-1979, cf. section 2.1),~~ using a local gradient search procedure, applied in combination with pre-screening of the parameter space (Perrin et al., 2008). ~~t~~The Kling and Gupta Efficiency criterion (Gupta et al., 2009, hereafter denoted as KGE~~; see section 3.3 for more details~~) was used as objective function. The KGE criterion ranges between -∞ and 1 (perfect simulation) and is calculated as follows:

$$\mathrm{KGE} = 1 - \sqrt{(\beta - 1)^2 + (\alpha - 1)^2 + (r - 1)^2} \qquad (1)$$

With:

- β: ratio between the means of the simulated and observed streamflow time series; this quantifies the simulation bias, and ranges between 0 and +∞ (values > 1 indicate a model overestimation).
- α: ratio between the standard deviations of the simulated and observed streamflow time series; this quantifies the ability of the simulation to reproduce the variability of the considered variable, and ranges between 0 and +∞ (values > 1 indicate a model overdispersion).
- r: coefficient of correlation between the simulated and the observed streamflow time series; this quantifies the ability of the simulation to reproduce the observed temporal variations of the considered variable, and ranges between -1 and 1 (perfect correlation).

Using KGE limits the biases of both water balance and variability, while keeping a good temporal correlation.

Note that ~~F~~for each model simulation~~ (calibration and reconstruction)~~, the first simulated year was used as an initialization period, and was not considered for the final performance evaluation. All the rainfall-runoff model outputs presented in the manuscript have been produced at the daily resolution by using both GR4J rainfall-runoff model and its snowmelt routine CemaNeige.

Table 1 : Description and final values of the 6 free parameters of the CemaNeigeGR4J model after being calibrated over the observed streamflow series of the Caniapiscau catchment.

Parameter	Description (and unit)	Calibrated values
X1 (GR4J)	Capacity of the production store (mm)	405
X2 (GR4J)	Water exchange coefficient (mm/day)	3.06
X3 (GR4J)	Capacity of the nonlinear routing store (mm)	326
X4 (GR4J)	Unit hydrograph time base (day)	3.50
X5 (CemaNeige)	Cold content factor (-)	0.004
X6 (CemaNeige)	Snowmelt factor (mm/day/°C)	3.66

3.2 Step 2: ~~Finding~~ finding analogue ~~days~~ dates (ANA)

The ANA approach is a resampling method based on ~~pressure field~~synoptic circulation similarities between days, with a sampling of observed climatic series for a given time period (here, 1950-2011, the *observation period*) over a longer time period (here, the 1880-2011 period, the *reconstruction period*). The synoptic information considered for the analogy is *geopotential height fields*. Here, each day is described by four geopotential height fields: (i) 1000 hPa at 0h, (ii) 1000 hPa at 24h, (iii) 500 hPa at 0h, and (iv) 500 hPa at 24h. The geopotential height fields are extracted over a large domain covering the studied area (cf. sub-section 2.1.1, see Appendix A). The metric used to rank the days in terms of analogy is the Teweles-Wobus (1954) distance (see Appendix B), which highlights similarities in terms of geopotential field shapes (Obled et al., 2002), and has been shown to provide better outcomes than what is obtained by using classical Euclidean distances in this framework (Wetterhall et al., 2005). Note that a seasonal constraint is imposed for the identification of analogue days: the potential analogue days of a given day are the ones included in a 60-day period centered on the calendar studied day. Thus, analogues of a winter day are themselves winter days: for example, the potential analogue days for January 1, 1880 are all of the available days within the December 1st to January 30th period of the ~~observed time series~~observation period (here 1950-2011~~ period~~). Another constraint is also imposed for the identification of analogues in which no analogue can be selected if they are closer than 15 days from the chosen date. For example, the potential analogue days for January 1, 2000, are all of the available days within the December 1st to January 30th period of the observation period except the December 15, 1999 to January 15, 2000 period. ~~Finally, t~~The ranking of analogue days - ~~thanks to~~ is based on the Teweles-Wobus distance (see Appendix B). For each day, ~~– allows~~ a given number of *n* analogues ~~to be~~is considered for each studied day, ~~and ~~thus generating~~es~~ a climatic ensemble of *n* series. Here, the 20 nearest analogue days were selected for each studied day and each 20CR member considered (n = 20). Table 2~~Table 2~~ illustrates the generation of this climatic ensemble by giving several analogue days obtained for three particular dates (1880-01-01, 1880-01-02 and 2011-12-30). For example, when considering member 1 of the 20CR (M1), the first analogue day of 1880-01-01 is 1984-01-23, the second analogue day is 1991-12-12, and the 20th analogue day is 1988-01-16. Finally, 20*5 (5 members of the 20CR considered) daily climatic series were generated over the 1880-2011 period. ~~Note that a similar approach was tested over France in the framework of precipitation downscaling (Chardon et al., 2014).~~

3.3 Step 3: ~~Addition of local information to improve reconstructions (ANATEM)~~reconstruction of a daily climatic (P and T) ensemble

Using ANA outputs, ANATEM aims to exploit the available long-term reference time series (hereafter denoted as TEM) to improve the climatic reconstruction, by applying a classical regression between ANA outputs and the reference series. In this study, the ANA approach was directly applied for the precipitation reconstruction (since no precipitation “witness” series was available), while the ANATEM approach was applied for the reconstruction of daily temperature (using the BEST daily temperature series). As in Kuentz et al. (2015), the local regression model (hereafter denoted as LM), applied here for the temperature reconstruction, is based on an additive correction, modeled by a daily harmonic function. The parameters of this regression~~is~~ function were~~as~~ ~~calibrated ~~estimated over the ~~observed ~~observation period (here, 1950-2011) on the interannual mean monthly residuals of the differences between the catchment temperature series and the TEM series, and has the following expression:

$$\hat{\mathrm{T}}_{\mathrm{LM}(\mathrm{d})} = \mathrm{T}_{\mathrm{TEM}(\mathrm{d})} + \beta(\mathrm{d}) + \varepsilon(\mathrm{d}) \tag{\cancel{1}2}$$

where $\hat{\mathrm{T}}_{\mathrm{LM}(\mathrm{d})}$ is the estimate of the air temperature for the day d, $\mathrm{T}_{\mathrm{TEM}(\mathrm{d})}$ is the value of the witness series temperature for the same day, $\beta(\mathrm{d})$ is the correction, depending on the calendar day of the year, and $\varepsilon(\mathrm{d})$ is a residual assumed to have zero mean.

The ANATEM method was applied at the daily ~~time step~~resolution over the 1880-2011 period. The ensemble of temperature values reconstructed for the day d has the following expression:

$$\left[\hat{\mathrm{T}}^{\mathrm{k}}_{\mathrm{ANATEM}(\mathrm{d})}\right]_{\mathrm{k}=1,\dots,\mathrm{n}} = \hat{\mathrm{T}}_{\mathrm{LM}(\mathrm{d})} + \left[\mathrm{T}(\mathrm{d}_{\mathrm{k}}) - \hat{\mathrm{T}}_{\mathrm{LM}}(\mathrm{d}_{\mathrm{k}})\right]_{\mathrm{k}=1,\dots,\mathrm{n}} \tag{\cancel{2}3}$$

where $\left[\hat{\mathrm{T}}^{\mathrm{k}}_{\mathrm{ANATEM}(\mathrm{d})}\right]_{\mathrm{k}=1,\dots,\mathrm{n}}$ is the ensemble of n reconstructed temperature values for the target day d, $\hat{\mathrm{T}}_{\mathrm{LM}(\mathrm{d})}$ is the air temperature estimate obtained with the regression model for the day d, d_{k} is the kth analogue day selected for the day d, $\mathrm{T}(\mathrm{d}_{\mathrm{k}})$ is the observed temperature value for the kth analogue day, $\hat{\mathrm{T}}_{\mathrm{LM}}(\mathrm{d}_{\mathrm{k}})$ is the air temperature estimate obtained with the regression model for the kth analogue day, and n is the total number of analogue days (here n=20, see section 2.1.1).

The final climatic ensemble is built with 100 precipitation (ANA outputs) and air temperature (ANATEM outputs) daily series over the 1880-2011 period. For each day, the 100 climatic values are obtained based on the 20 “closest” analogue days for each of the 5 20CR members considered.

Table 2. Illustration of the analogue dates obtained with the ANA approach. Here, a sub-sample of the 20 analogue days of three particular dates (1880-01-01, 1880-01-02 and 2011-12-30) are given for each of the five 20CR members considered (M1 to M5). The ranking of analogue days is performed with Teweles-Wobus (1954) distances.

20CR MEMBER	ANA	1880-01-01	1880-01-02	...	2011-12-30
M1	**ANA1**	1984-01-23	1959-02-13	...	2007-12-18
	ANA2	1991-12-12	1961-01-11	...	1989-11-05
	...	...	...	...	...
	ANA20	1988-01-16	1953-12-25	...	2007-12-19
M2	**ANA1**	1984-01-23	1974-12-27	...	1979-11-19
	ANA2	1990-11-30	1961-01-11	...	1971-11-13
	...	...	...	...	...
	ANA20	1957-02-02	1990-02-19	...	1976-12-04
M3	**ANA1**	1950-02-03	1950-02-04	...	2007-12-18
	ANA2	1989-01-13	1971-12-24	...	1989-11-05
	...	...	...	...	...
	ANA20	1990-11-30	1957-02-07	...	2003-12-14
M4	**ANA1**	1986-12-15	1956-12-21	...	2007-12-18
	ANA2	2007-01-02	1974-01-19	...	1989-11-05
	...	...	...	...	...
	ANA20	2004-12-29	1971-12-24	...	1994-11-20
M5	**ANA1**	1984-01-23	1961-01-11	...	2007-12-18
	ANA2	1989-01-13	1962-01-25	...	1971-11-13
	...	...	...	...	...
	ANA20	1993-11-09	1965-11-04	...	1958-11-16

3.4 Step 4 : reconstruction of a daily streamflow ensemble

Using the rainfall-runoff model parameter set obtained after calibration (step 1), the reconstructed climatic ensemble is finally transformed into one streamflow ensemble, available over the 1881-2011 period (1880 being used as an initialization period) at the daily temporal resolution. The final streamflow ensemble thus consists of 100 daily streamflow series over the 1881-2011 period.

~~3.4~~3.5 ~~Evaluation~~ Comparison of reconstructed series against observations

In order to compare the reconstructed streamflow time series against observations, the reconstructed ensembles were first aggregated: a daily series was generated for each of the five 20CR members considered by averaging the 20 daily series constituting each ensemble. The five daily mean series are denoted as $\overline{\mathrm{ANA}}$ or $\overline{\mathrm{ANATEM}}$, depending on the method used to produce them.

The evaluation of the reconstruction performances was based on the three KGE components and its final values ~~(ranging between -∞ and 1),~~. For the reconstructed climatic time series, the computation of these four scores was carried out over the 1950-2011 period, at the daily time scale but also at the monthly time scale, in order to evaluate the intra-annual reconstruction performances, and at the yearly time scale, in order to evaluate interannual reconstruction performances. For the reconstructed streamflow ensemble, these scores were computed over mean annual flow values and mean May flow values over two time periods, 1963-1979 (rainfall-runoff model calibration period) and 1982 to 2011 (naturalized flows).

4 RESULTS

4.1 Rainfall-runoff model calibration performances (1963-1979)

Over the 1963-1979 calibration period, the CemaNeigeGR4J model performs really well with a KGE value of 0.93 (rainfall-runoff simulation with KGE > 0.8 are generally considered as good). The values of the 6 calibrated parameters are detailed in the Table 1~~1~~. Figure 4~~Figure 4~~ presents the performance of the CemaNeigeGR4J rainfall-runoff model over the calibration period (1963-1979). Simulated and observed quantiles of monthly streamflow show a strong correlation (Figure 4~~Figure 4~~a), with a limited~~n~~ overestimation of the lowest values by the rainfall-runoff model observed during the winter months (from January to April, Figure 4~~Figure 4~~b). The timing of the simulated regime is similar to the observed one. However, systematic limited biases are found, with an overestimation of the winter streamflow values (January to April) and of the spring flood values (June) and an underestimation of the streamflow values during the snowmelt period (July to October). The model is also able to simulate the general interannual variability of mean annual streamflow (Figure 4~~Figure 4~~c), with higher values for the 1964-1969 period and lower values for the 1970-1976 period, for example. Nevertheless, non-systematic biases are found for several years, with both underestimations (e.g., 1964 and 1969 years) and overestimations (e.g., 1972 and 1975 years) of mean annual streamflow values. Finally, the observed and modeled distributions of annual streamflow values are similar (Figure 4~~Figure 4~~d), with an overestimation of the lowest mean annual streamflow values.

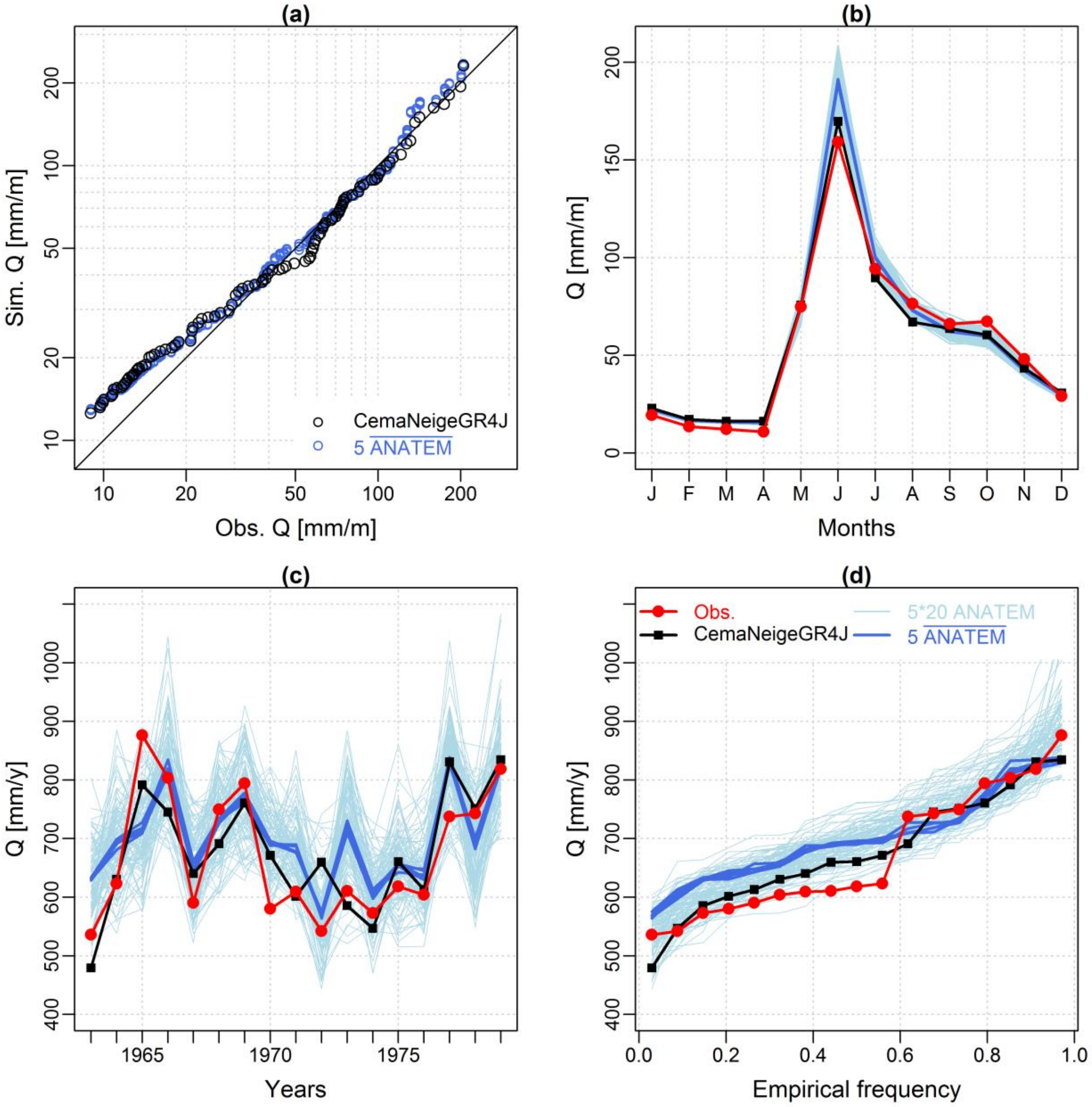

Figure 4. Performances of the CemaNeigeGR4J rainfall-runoff model (black) and of the ANATEM flow reconstruction (blue colors) evaluated over the calibration period of the rainfall-runoff model (1963-1979)~~)~~. (a) Monthly quantile-quantile plots (logarithmic scale), (b) observed and simulated monthly streamflow regime, (c) observed and simulated interannual streamflow variability, and (d) observed and simulated streamflow yearly mean distribution. The legend indicated on the (d) graph is also valid for the (b) and (c) graphs.

4.2 Climatic reconstructions (1950-2011 and 1880-2011)

In this section, the results of the climatic reconstruction are presented, first in terms of performance estimated over the observed period (1950-2011), and then in terms of centennial mean annual series (1880-2011).

4.2.1 Performance of the climatic reconstructions over the observed period (1950-2011)

Figure 5~~Figure 5~~ compares the temperature reconstruction (using ANA and ANATEM outputs) and precipitation reconstruction (using ANA outputs) to the observations for the 1950-2011 period, in terms of monthly regimes and yearly value distributions. For temperature, the ANATEM reconstruction is excellent, both in terms of monthly regime and ~~of~~ yearly mean value distribution. The ANA temperature reconstructions (in grey) show a limited performance for the coldest months (December and January) and for the warmest months (July and August), and thus highlight the importance of using the ~~TEM~~ BEST temperature series through ANATEM, which successfully corrects the ANA outputs. The intra-variability of the ANATEM temperature ensemble is very limited.

The precipitation ~~reconstitution~~ reconstruction is not as good as that of the temperatures. The timing of the monthly regime is well captured, with lowest monthly precipitations observed in February, and the highest in July. However, ~~A~~an overestimation of the ~~reconstituted~~ reconstructed precipitation is observed for all months, with the exception of January and September. Overall, a wet monthly bias of precipitation is found. This bias is also seen in the plot of the yearly value distributions (Figure 5~~Figure 5~~d), which show that a majority of the mean annual precipitation values are overestimated by the reconstruction. In terms of variability within the ensemble, the similarity of the five 20CR members ($\overline{\mathrm{ANA}}$, in blue) shows that the uncertainty of the geopotential height field (quantified here through the consideration of the five members) has a negligible impact on the precipitation reconstruction over this time period and at these ~~time steps~~resolutions (yearly and monthly). The relatively large width of the ANA ensembles (grey envelopes) indicates that the uncertainty due to the selection of 20 analogue days has an impact on the precipitation reconstruction.

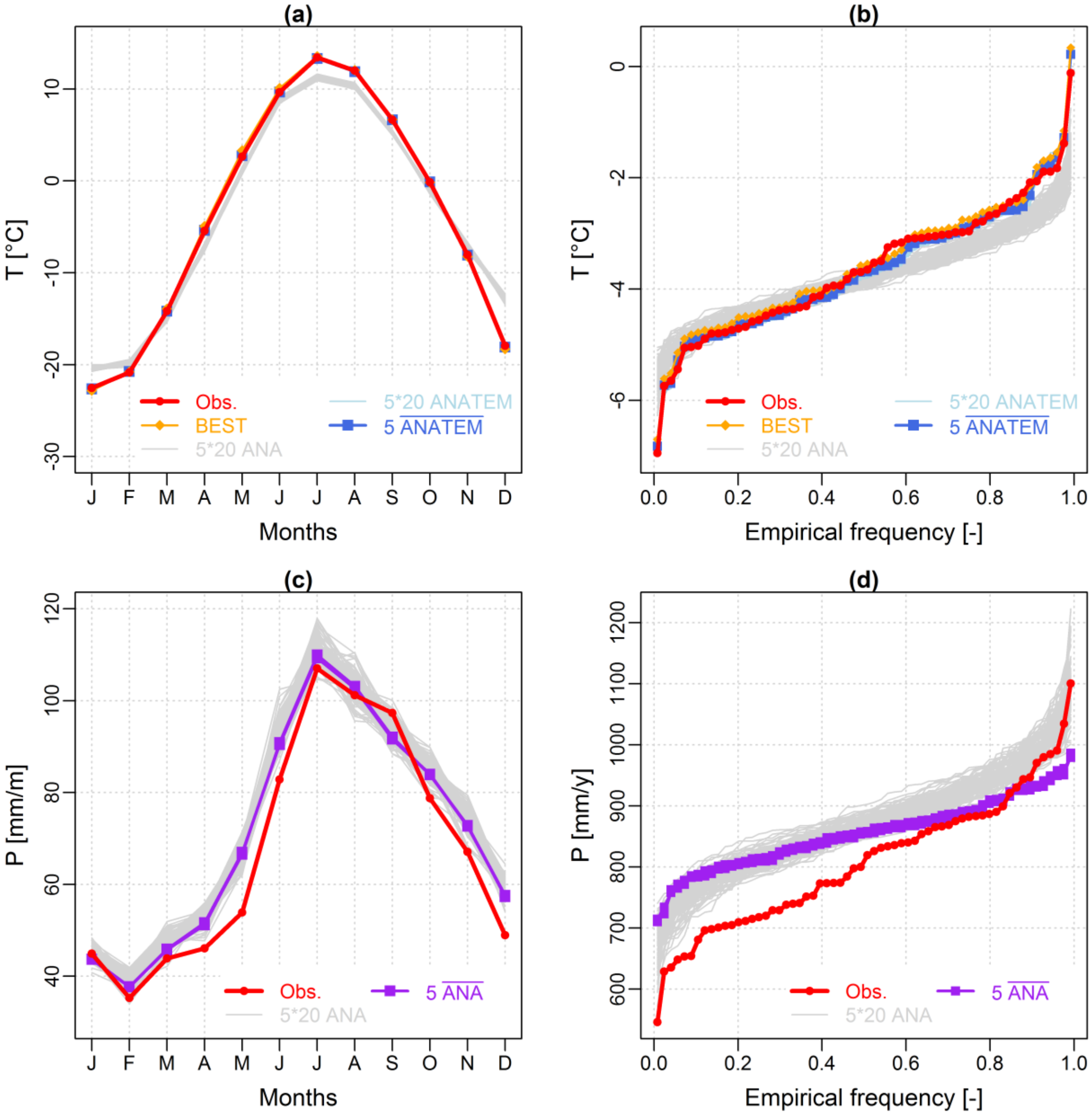

Figure 5. Monthly regimes (a and c) and yearly value distributions (b and d) for temperature (with ANA and ANATEM) and precipitation (with ANA) reconstructions and observations over the 195~~1~~0-2010 period. Note that for temperature monthly regime (a), the ANATEM simulations are similar to the observations, and thus, ANATEM curves (blue) are not visible since they are below the observation curve (red).

Figure 6~~Figure 6~~ summarizes the climatic reconstruction performances at the daily, monthly and yearly ~~time steps~~resolutions~~, both~~ over the 1950-2011 period. For air temperature (Figure 6~~Figure 6Figure 4~~a), and as previously indicated, the overall reconstruction performances are excellent for ANATEM outputs (KGE > 0.9), and limited for ANA outputs (KGE > 0.4). ~~At both time steps,~~ ANA outputs (grey points) are characterized by an overestimation ($\beta > 1$) tendency for the three resolutions and an underdispersion ($\alpha < 1$) tendency~~ies~~ for the monthly and annual resolutions. If the yearly temporal correlation is good at the daily and at the yearly ~~time step~~resolutions, the temporal correlation is excellent at the monthly ~~time step~~resolution ($r \approx 1$). For precipitation (Figure 6~~Figure 6Figure 4~~b), the overall reconstruction performance is better at the monthly resolution~~time step~~ (KGE > 0.6) than at the daily (KGE ranging between 0.3 and 0.5) and yearly resolution~~time step~~ (KGE ranging between 0.2 and 0.6). The reconstructed time series show a clear overestimation bias, an underdispersion problem, and a limited temporal correlation at ~~both time steps~~the three different resolutions. Averaging each ensemble of the considered 20CR members (blue points) results in better temporal correlations at the daily and yearly resolutions, but ~~logically~~at the expense of~~,~~ lower variability reproduction performance.

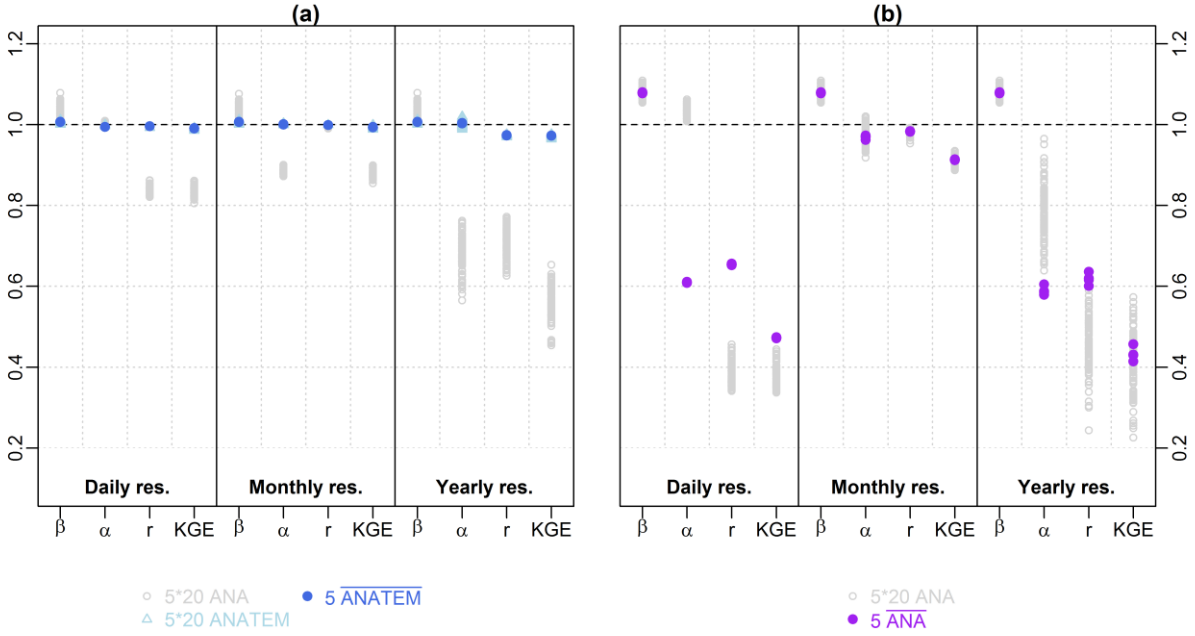

Figure 6. Daily, m~~M~~onthly and yearly performances of the air temperature ANA and ANATEM reconstructions (a) and the ANA precipitation reconstructions (b), for 1950-2011 period.

4.2.2 Centennial mean annual climatic series (1880-2011)

Figure 7~~Figure 7Figure 5~~ shows the reconstructed climatic series over the entire studied period (1880-2011), at the yearly ~~time step~~resolution. For temperature, the ANATEM reconstruction shows a very good fit to the observed series, with the exception of the first decade (1950-1960), when the reconstructed annual temperatures appear to be systematically lower than the observed annual temperature. ANA ensembles are larger than their ANATEM counterparts, and perform worse in terms of mean annual temperature variability. The good performance of the ANATEM reconstruction is ~~thanks~~ largely due to the ~~TEM~~ BEST series, which is strongly correlated with the observed series at the annual ~~time step~~resolution, except for the first observed decade. At the centennial scale, the reconstructed temperature time series are highly similar to the ~~TEM~~ BEST series, showing that the entire temperature signal reconstructed is driven here by the ~~TEM~~ BEST series. The ANATEM ensemble width is narrow at the annual time scale, as has already been seen for the monthly regime (Figure 5~~Figure 5~~a and b). The reconstruction shows an increase in the Caniapiscau catchment mean annual temperature over the last 130 years.

For mean annual precipitation, the ANA reconstruction does not perform as well, especially over the last two decades (1990-2010), where the reconstruction failed to reproduce the observed low values for the mean annual precipitation (compared to mean values over the entire observed period). A similar bias is found for the 1950-1965 period, while the variability of the mean annual precipitation values during the 1965-1985 period are well reproduced. Relatively, the precipitation reconstruction seems to be able to reproduce the wet-dry periods, but fails to match the observed values. Considering the reconstruction at the centennial time scale, no significant trend is found for mean annual precipitation. Several periods are interesting, such as the sequence of wet and dry years around 1920. Finally, variability due to consideration of five 20CR members is seen until 1940, and seems to be higher for several time periods, such as the 1880-1890 decade.

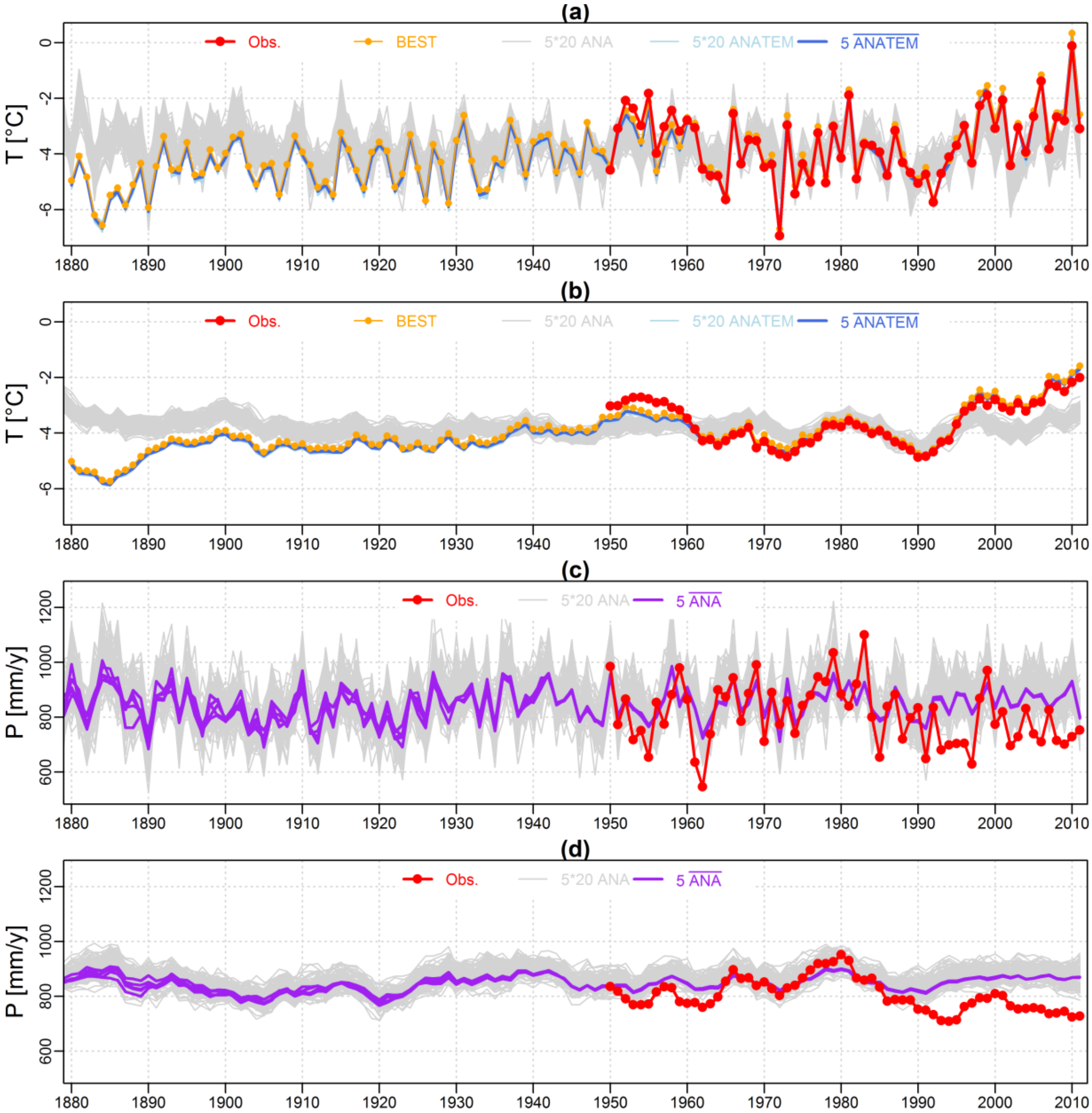

Figure 7. Interannual variability of reconstructed mean annual values of temperature (ANA and ANATEM outputs) and precipitation (ANA outputs) compared with observations over the 1880-2011 period. (a) and (c) are raw yearly values while (b) and (d) are 6-year running means of mean annual temperature and mean annual precipitation, respectively.

4.3 Streamflow reconstructions (1962-2011 and 1881-2011)

In this section, the results of the streamflow reconstructions are presented, first in terms of performance estimated over two time periods, and then in terms of centennial series (annual mean flows and spring flood values).

4.3.1 Performance of streamflow reconstruction over two observed periods (1962-2011)

Using the five climatic ensembles produced by ANA (for precipitation) and ANATEM (for temperature) as inputs to the CemaNeigeGR4J rainfall-runoff model, five ensembles of 20 daily streamflow series were produced over the 1881-2011 period (the year 1880 is used as an initialization period for the rainfall-runoff model). ~~Figure 6~~ Figure 4~~Figure 4~~ presents the performance of the streamflow reconstructions over the rainfall-runoff model calibration period (1963-1979). The obtained reconstructions have, logically, the same qualities and defaults characterizing the climatic reconstructions (presented in section 4.2.1~~4.1.1~~) and the rainfall-runoff model performance (presented in section 4.1~~4.2~~). Figure 4~~Figure 4Figure 6~~a is a quantile-quantile plot between observed and simulated mean monthly streamflows. Monthly correlations between observations and simulations are good, but reveal a systematic overestimation of the lowest mean monthly streamflow values (winter months). A clear overestimation of the monthly flood peak (June) is also found (cf. Figure 4~~Figure 4Figure 6~~b), due both to the rainfall-runoff model performance on this catchment and a general overestimation of the precipitation by the climatic reconstruction, as already shown in Figure 5~~Figure 5Figure 3~~. Observed and simulated interannual variabilities are similar, but with an overestimation of the mean annual streamflow values by the reconstructions, especially for the years with relatively low mean annual streamflow values (1971-1976).

Figure 8~~Figure 8Figure 7~~ summarizes the performances of the streamflow reconstructions over two periods (1962-1979 and 1981-2011), in terms of mean annual streamflow values (Figure 8~~Figure 8Figure 7~~a) and May monthly flow values (Figure 8~~Figure 8Figure 7~~b). Overall KGE performances are limited to good for mean annual streamflow series and very good for the May monthly flow series. Again, an overestimation of mean annual flows is found for both periods. For May monthly flows, no specific trend is found for the first period, while a slight underestimation is observed for the second period.

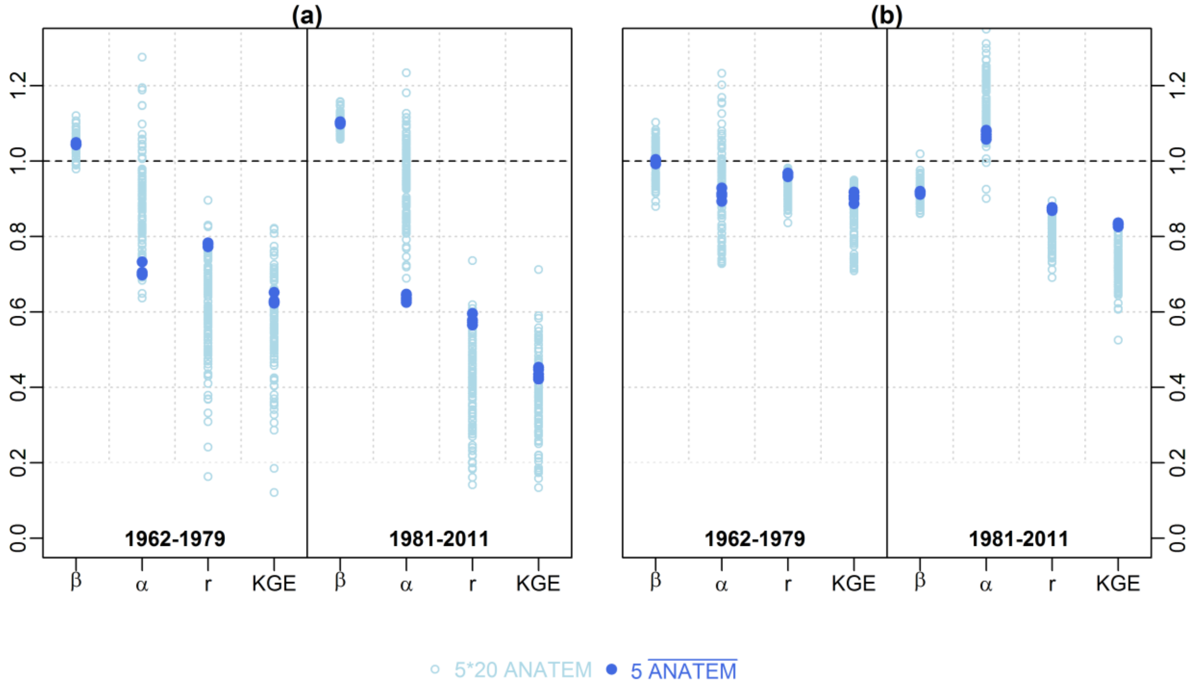

Figure 8. Streamflow reconstruction performances evaluated over two periods (1962-1979 and 1981-2011), (a) mean annual streamflow values ~~(a)~~ and (b) May monthly flow values. ~~(b)~~

4.3.2 Centennial mean annual flow reconstructions (1881-2011)

Figure 9~~Figure 9~~ presents the centennial ANATEM streamflow reconstruction and compares the reconstruction to observations and to the mean flow reconstruction proposed by Nicault et al. (2014) using tree rings. As shown in Figure 4**Erreur ! Source du renvoi introuvable.**, a good correlation is found between the ANATEM reconstruction and observations for the 1963-1979 period. Considering the other streamflow observation time period (naturalized flows of 1982-2011), the correlation is weaker, with a general overestimation of the mean annual streamflow. At the centennial scale, a comparison between ANATEM and tree ring mean flow series reveals that the two series are not statistically different, since the ANATEM ensemble is within the tree ring ~~interval~~ confidence interval (green envelopes), except for the 1930-1940 period. For this period, and especially around 1940, ANATEM mean flow reconstructed values are significantly higher than tree ring ones. A significant variability of mean annual streamflow is simulated for the 130 past years. The two reconstructions agree for the 1880-1910 period, simulated as a period of decreasing mean annual streamflows, followed by a 10-year increasing period. The 1920-1950 period shows differences between the two reconstructions, with ANATEM mean flows being larger than for tree rings. For the 1950-2011 period, the mean flow relative evolutions are similar, but the absolute values are different, with ANATEM values being systematically higher than tree ring values. This constant bias could be explained by the overestimation of precipitation over the record period. The 1912 year seems to be a "hydrologically interesting year", since it is simulated as a very wet year by tree rings, while simulated as a dry year by ANATEM. Finally, as for the ANA precipitation reconstruction, the variability due to consideration of five 20CR members is seen until the year 1940, and seems to be higher over the distant past.

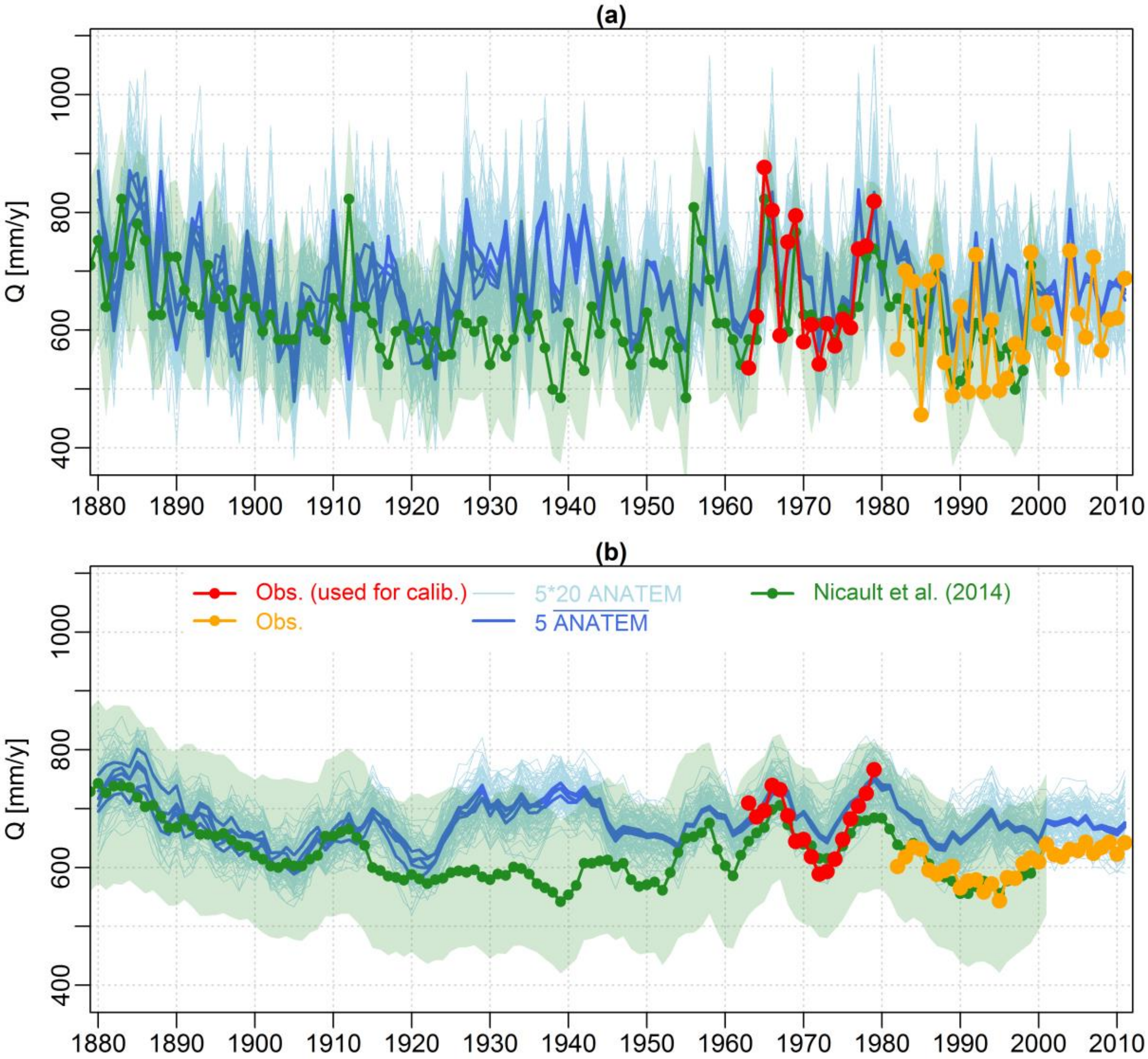

Figure 9. ANATEM mean flow reconstructions: comparison with observations and Nicault et al. *(2014) tree ring series, 1881-2011 period. (a) is raw yearly values while (b) is 6-year running means of mean flows.*

4.3.3 Centennial spring flood reconstruction (1881-2011)

Finally, Figure 10~~Figure 10~~ presents the ANATEM centennial spring flood reconstruction compared to observations and to the reconstruction proposed by Boucher et al. (2011) using tree rings. For ANATEM and for the observed streamflow series, these annual series were constituted by estimating, for each year, the May monthly flow, since Boucher et al. (2011) produced a May streamflow reconstruction. The correlations between the ANATEM reconstruction and the observed series (1963-1979 and 1982-2011) are excellent and very good, respectively, and thus reproduce the increase of spring floods during the 1970-1980 period, and then the decrease during the 1980-1990 period, finally followed by a slight increase and a stagnation over the two last decades. At the centennial scale, the two reconstructions appear to be significantly different for a long period of time, since the ANATEM ensemble is out of the tree ring confidence interval for the 1881-1920 period. Another significant difference exists over the 1950-1960 period, seen as an ~~“average~~ common decade~~”~~ by the tree ring reconstruction (reconstructed spring flood ranging from 47 to 87 [mm/m]), while being seen as a highly variable hydrological decade for the ANATEM reconstruction, with high values for the first five years (around ~~100~~ 110 [mm/m] for the 1950-1955 period), and then two very low values (around 20 [mm/m] for the 1956-1957 period), finally followed by three high value years (around 110 [mm/m] for the 1958-1960 period). Overall, the ANATEM reconstruction simulated an increasing trend of spring floods for the Caniapiscau catchment. This trend is related to the increasing temperature trend, as illustrated in Figure 7~~Figure 7Figure 5~~.

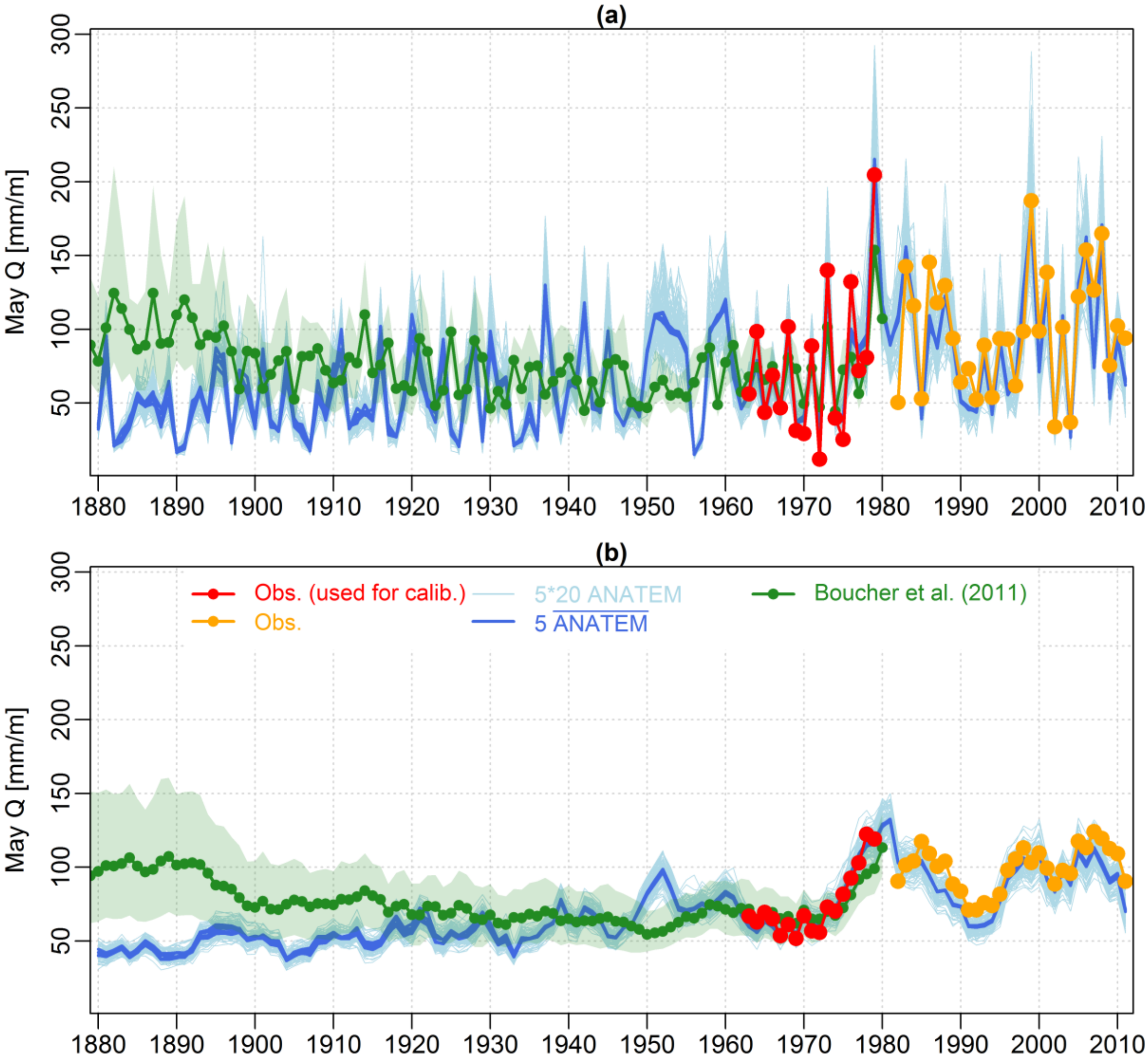

Figure 10. ANATEM spring flood reconstructions: comparison with observations and Boucher et al. *(2011) tree ring series, 1881-2011 period. (a) is raw yearly values while (b) is 6-year running means of spring flood values.*

5 DISCUSSION AND CONCLUSION

In this study, a daily hydro-climatic reconstruction is proposed for the Caniapiscau Reservoir (northern Quebec, Canada) for the 1881-2011 period. This reconstruction was generated by firstly applying the ANATEM method (Kuentz et al., 2015), combining large-scale atmospheric information (here the NOAA 20th Century geopotential height reanalysis, Compo et al., 2011) with local climatic observations – when such series are available – to produce a daily ensemble of climatic series (precipitation and air temperature). Secondly, this climatic ensemble was used as input to a rainfall-runoff model (here GR4J ~~, ~~(Perrin et al., 2003) and its snow accumulation and melt routine, CemaNeige (Valéry et al., 2014)) previously calibrated in order to obtain a streamflow ensemble, at the daily ~~time step~~resolution. The performances of the climatic reconstructions were quantified over the observed period (1950-2011) and showed very good performance for air temperature, both in terms of monthly regime and interannual variability. This excellent performance is due mainly to the use of a local reference temperature time series (here, a daily temperature time series extracted from the Berkeley Earth Surface Temperature analysis, Rohde et al., 2013). For precipitation, no local reference climatic time series was available and the precipitation ~~reconstitutions~~ reconstructions are thus only a function of geopotential height field analogy. The precipitation ~~reconstitutions~~ reconstructions present a good performance in terms of regime, but with a somewhat limited ability to reproduce the observed annual values and interannual variability, combined with a systematic wet bias. The performance of the streamflow reconstruction was then compared to streamflow observations. This comparison showed a good performance, both in terms of monthly regimes and interannual variability, with a systematic overestimation of the mean annual streamflow values, due mainly to the wet bias of the precipitation reconstruction by the ANATEM method.

These newly produced reconstructions were then compared to two different reconstructions performed on the same catchment by using tree ring data series, one being focused on mean annual flows (Nicault et al., 2014), and the other on spring floods (Boucher et al., 2011). In terms of mean annual flows, the interannual variability of flows reconstructed by tree rings and ANATEM were similar (except for the 1930-1940 decade), with significant changes seen in wetter and drier years. This variability seemed to be driven mainly by the variability of mean annual precipitation. In terms of spring floods, the interannual variabilities reconstructed by tree rings and by ANATEM were quite similar for the 1955-2011 period, but significantly different for the 1880-1940 period. The ANATEM spring flood reconstruction showed an increasing trend over time, and this variability seemed to be driven by the variability of the mean annual temperature.

These results emphasize the need to apply different reconstruction methods on the same catchments. Indeed, such comparisons highlight potential differences between available reconstructions, and finally, allow a retrospective analysis of the proposed reconstructions of past hydro-climatological variabilities. In this study, two very different reconstruction methods were applied on the same catchment, revealing several periods where the two reconstructed streamflow series differ considerably. Thus, in terms of mean annual flows, the year 1922 and the 1930-1940 decade appear to be particularly dry and wet, respectively, when reconstructed with the ANATEM method, while they are simulated as particularly wet and dry when ~~reconstituted~~ reconstructed using tree ring proxies. In terms of spring floods, the two reconstruction methods are in disagreement for the 1950-1960 decade, simulated as a decade with wide variabilities by ANATEM, with short sequences of alternating high and low spring flood values, compared to the tree ring reconstruction. Further investigation is needed in order to understand the differences for these specific periods. Finding indications of particular hydro-climatic conditions at the regional scale through the analysis of documents, reports or ad-hoc measurements could represent~~ing~~ a means of assessing the respective performances of each ~~reconstitution~~ reconstruction method. More generally, the long-term signals of the spring flood reconstructions~~reconstitutions~~ are different, with a clear increasing tendency for floods ~~reconstituted~~ reconstructed with ANATEM, related to the mean annual temperature rise in this region through the studied decades. Further work is needed to investigate this difference between the two reconstructions.

The evaluation of the analogue performance revealed two main limitations for the precipitation reconstruction. Firstly, a general wet bias was found when the reconstructed precipitation time series were compared to observations, and therefore, a similar bias was observed for streamflow reconstruction. A classical bias-correction method could be applied on the reconstructed precipitation time series in order to eliminate this bias. However, applying a bias correction method implies an additional error source which could be amplified when the streamflow is analyzed (Teng et al., 2015), and, even more importantly, raises the issue of the bias stationarity (e.g., Teutschbein & Seibert, 2013; Chen et al., 2015, and Velázquez et al., 2015). Secondly, the interannual variability of mean annual precipitation is reproduced with limited performances on the Caniapiscau reservoir catchment. Th~~is limitation~~ e inability of the analogue approach to reproduce the interannual precipitation variability - already highlighted by Kuentz et al. (2015) over 22 French catchments – is due to the absence of a local reference climatic time series, unlike for temperature ~~reconstitution~~reconstruction, where a local temperature time series is used, and ensures that the simulated interannual temperature variability is reproduced efficiently. Finding an additional series which significantly improves the precipitation reconstruction is a major perspective of this work. The use of variables produced by the available reanalyses (e.g., relative humidity, precipitable water content) for finding analogue dates will be investigated, along with the testing of time series of local pressure measurements. For example, Caillouet et al. (2016) showed that adding the sea surface temperature variable to the temperature, geopotential, vertical velocity and humidity for finding analogue dates significantly improves the reconstruction of air temperature and precipitation over France.

In this study, most of the ANA approach options used to find analogue days were defined by looking at previous applications of the same methodology (e.g., Horton et al. 2012 and Chardon et al. 2014) and by sensitivity analyses (results ~~not shown here~~partially shown in Appendix A). The sensitivity of the final ~~reconstitutions~~ reconstructions to these options (size of the geopotential height domain extension (see Appendix A), choice of the geopotential height levels studied, number of analogue days, etc.) could be further investigated in a future work. Interestingly, the uncertainty due to the use of five members of the 20CR reanalysis appears to be limited, and even null from 1940 onward. See for example Figure 9~~Figure 9~~ which presents the centennial ANATEM streamflow reconstructions: it is impossible to distinguish the five ANATEM average series after 1940, highlighting that considering five different members of the 20CR reanalysis as inputs of the reconstruction method has a negligible impact on the reconstruction of the mean annual streamflow.

Finally, the reconstructed climatic time series are transformed into streamflow time series thanks to a daily rainfall-runoff model, previously calibrated over the relatively short observation period (with really good calibration performances). The use of one model, one objective function and one parameter set is questionable. Quantifying the sensitivity of the obtained reconstruction to the hydrological modeling assumptions made was out of the scope of this paper, but definitively deserves further research, especially considering the issue of uncertainty due to rainfall-runoff model parameter~~-set~~ in a changing climate. Thus, numerous authors highlighted that calibrated parameters of rainfall-runoff models are dependent on the climate of the calibration period and that performance decreases when applied over periods where the climate differs from that of calibration period (e.g., Merz et al. 2011; Coron et al. 2012 and Brigode et al. 2013b). Thus, testing different calibration strategies (e.g., bootstrap calibration used by Brigode et al. 2015), testing particular objective functions especially devoted to the final study objective (e.g., studying mean annual streamflow), and adapting the time step of the rainfall-runoff model to the objective would be interesting for future works.

The combination of the ANATEM ~~reconstitution~~ reconstruction method with a rainfall-runoff model offers an interesting method for use in ~~reconstituting~~ reconstructing hydro-climatic series at a very fine ~~time step~~resolution (here daily), which is usually needed in applying impact models (such as dam management models), and finally, to discuss the climatic process, which significantly influences the hydrological decadal variability at the catchment scale. An interesting perspective would be to test this modeling approach on numerous other catchments, and focus~~ing~~ on regions where long and good quality hydro-climatic time series are available, thus giving the opportunity to quantitatively evaluate the ~~reconstitution~~ reconstruction methodology over long time periods. Kuentz et al. (2013) thus reconstructed 110-year streamflow time series for 22 French catchments with a combination of the ANATEM reconstruction method and a daily rainfall-runoff model, reconstitutions which allowed to discuss the hydro-climatic variability over the last century in the studied region (French Alps). Finally, these applications could also give interesting insights on regions where it is not sufficient to consider only climatic time series in explaining observed multi-decadal hydrological variability, and thus highlight other significant factors influencing hydrological variability that need to be quantified (e.g., changes in land use, urbanization or hydrogeology).

Another way to evaluate the two reconstruction methods would be to use the hydro-climatic time series reconstructed by ANATEM as inputs for a tree diameter growth model (e.g., models developed and applied for black spruces (*Picea mariana* [Mill.] BSP) in Canada by Subedi & Sharma 2013 and Huang et al. 2013), and to then compare the tree ring simulated through this growth model with the observed tree ring series.

6 APPENDIX A

Several tests have been performed for choosing the spatial domain to consider for the description of the geopotential height fields (see Brigode et al. (2013a) and Radanovics et al. (2013) for similar approaches). Here, eight different spatial domains have been tested (domain numbered from 1 to 8). These domains, illustrated on Figure A1a, are centered on the Caniapiscau catchment and are of progressively larger. For each domain, a climatic reconstruction has been performed with the ANA method for the Caniapiscau catchment but also for 211 other Quebec catchments of the cQ2 database (Guay et al., 2015). These reconstructions have been performed on the 1990-2010 period with only one member of the 20CR reanalysis. The performances of these different reconstructions have been evaluated by comparing observed series with reconstructed series looking at different precipitation and air temperature criterion.

Figure A1b presents the three criterion chosen to evaluate the precipitation reconstruction: (i) the correlation between observed and reconstructed annual precipitation series (first line, optimal value is 1), (ii) the correlation between observed and reconstructed daily precipitation series (second line, optimal value is 1) and (iii) the bias between observed and reconstructed precipitation series (last line, optimal value is 0). The boxplots summarize the performances obtained over the 211 catchments, while the purple point highlights the performance obtained specifically over the Caniapiscau catchment. Domain n°5 was finally chosen as a (subjective) compromise between having high correlation between reconstructed and observed precipitation series (at yearly and daily resolutions) and having low precipitation bias between reconstructed and observed series on both the studied catchment (Caniapiscau) and on other neighboring Quebec catchments. Thus, we believe that the methodology performed in this study could also be used for the reconstruction of streamflow series on other neighboring catchments.

Finally, Figure A1c presents the spatial distribution of the three criterion values obtained within domain n°5. These maps reveal interesting spatial patterns, highlighting for example higher performances in terms of daily precipitation correlation obtained for northern catchments compared to southern catchments. It is out of the scope of this paper to discuss the spatial variability and the spatial patterns of the climatic reconstruction performances, but this issue definitively deserves further research.

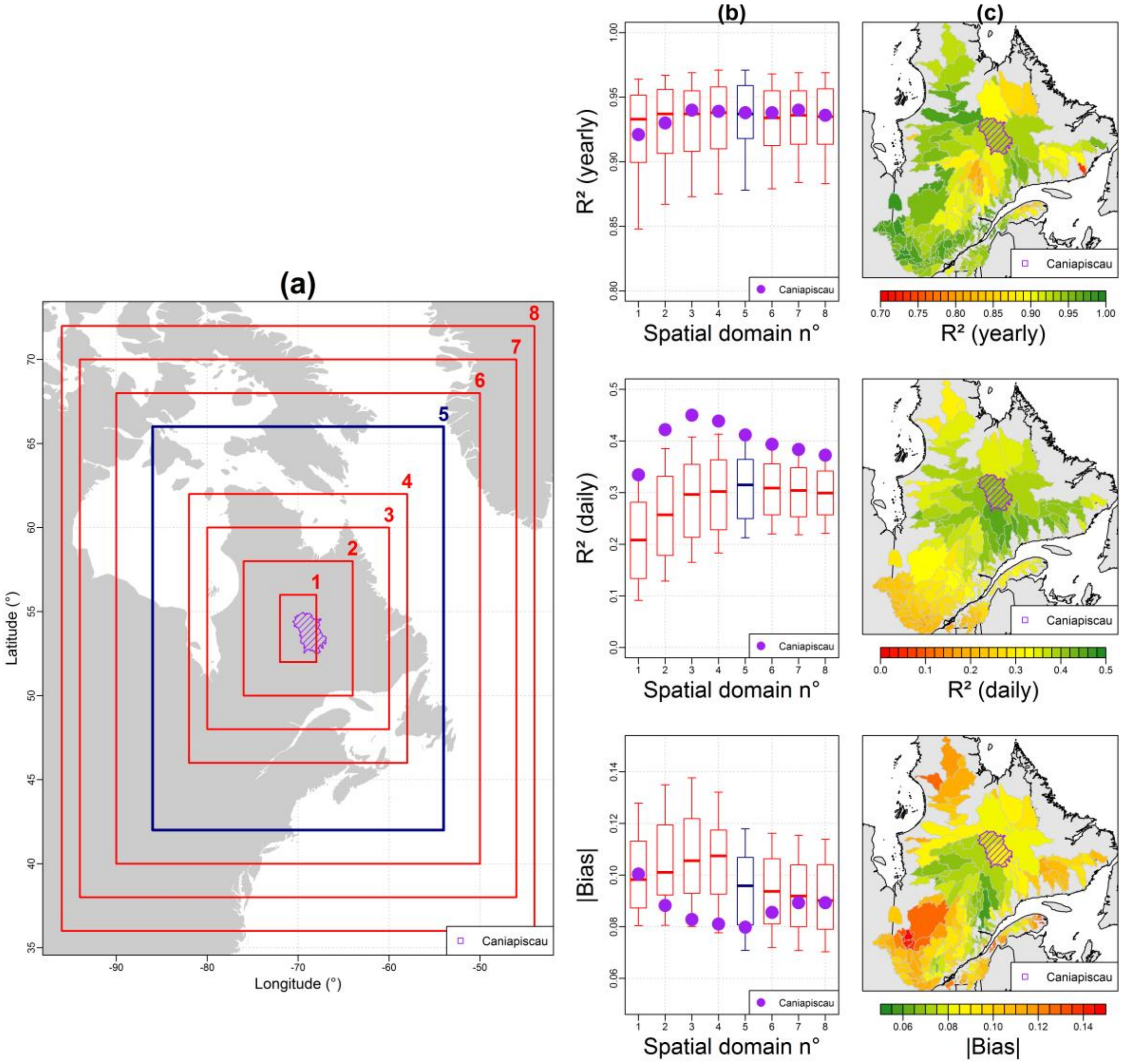

Figure A1 : (a) Spatial extension of the eight geopotential height domains considered. (b) Performances of the precipitation ANA reconstruction estimated over the 1990-2010 period for 211 catchments of the cQ2 database (the boxplots are constructed with the 0.10, 0.25, 0.50, 0.75 and 0.90 percentiles). (c) Spatial distribution of the performances obtained with the domain n°5 over the 211 catchments of the cQ2 database. The Caniapiscau catchment is highlighted with purple color.

7 APPENDIX B

The Teweles-Wobus (1954) distance (noted D_{TW} hereafter) is used to find analogues to the synoptic circulation of a given day and thus to quantify the (di)similarity between two synoptic spatial configurations, each characterized by by four geopotential height fields over a given spatial domain (see Appendix A): (i) 1000 hPa at 0h, (ii) 1000 hPa at 24h, (iii) 500 hPa at 0h, and (iv) 500 hPa at 24h. The final D_{TW} between a day A and another day B is the sum of four D_{TW} calculated for each of the four geopotential height fields. The distance between the geopotential height field Z (e.g. 1000 hPa at 0h) of the day A and the day B is calculated as follow:

$$D_{TW,Z} = 100 \times \frac{\sum_{i=1}^{I-1}\sum_{j=1}^{J}\left|\Delta_{i,j}^{i,A} - \Delta_{i,j}^{i,B}\right| + \sum_{i=1}^{I}\sum_{j=1}^{J-1}\left|\Delta_{i,j}^{j,A} - \Delta_{i,j}^{j,B}\right|}{\sum_{i=1}^{I-1}\sum_{j=1}^{J}\max\left(\left|\Delta_{i,j}^{i,A}\right|,\left|\Delta_{i,j}^{i,B}\right|\right) + \sum_{i=1}^{I}\sum_{j=1}^{J-1}\max\left(\left|\Delta_{i,j}^{j,A}\right|,\left|\Delta_{i,j}^{j,B}\right|\right)}$$

Where :

- $\Delta_{i,j}^{i,A} = Z_{i+1,j}^{A} - Z_{i,j}^{A}$ is the geopotential gradient of a west-east direction starting from a point (i,j) for the day A.
- $\Delta_{i,j}^{j,A} = Z_{i,j+1}^{A} - Z_{i,j}^{A}$ is the geopotential gradient of a south-north direction starting from a point (i,j) for the day A.

This distance is thus focused on the synoptic circulation gradients (south-north and west-east directions) and not on the absolute geopotential height values. D_{TW} ranges from 0 (for two identical fields) and 200 (for two opposite fields).

~~6~~8 ACKNOWLEDGMENTS

Support for the Twentieth Century Reanalysis Project, version 2c dataset, was provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. The authors thank the two reviewers who provided constructive comments on an earlier version of the manuscript, which helped clarify the text.

~~7~~9 REFERENCES

Allen KJ, Nichols SC, Evans R, et al (2015) Preliminary December-January inflow and streamflow reconstructions from tree-rings for western Tasmania, southeastern Australia. Water Resour Res. doi: 10.1002/2015WR017062

Andréassian V, Lerat J, Le Moine N, Perrin C (2012) Neighbors: Nature’s own hydrological models. J Hydrol 414–415:49–58. doi: 10.1016/j.jhydrol.2011.10.007

Arsenault R, Brissette F (2014) Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches. Water Resour Res. doi: 10.1002/2013WR014898

Bégin C, Gingras M, Savard MM, et al (2015) Assessing tree-ring carbon and oxygen stable isotopes for climate reconstruction in the Canadian northeastern boreal forest. Palaeogeogr Palaeoclimatol Palaeoecol 423:91–101. doi: 10.1016/j.palaeo.2015.01.021

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

Bradley RS (1999) Paleoclimatology: reconstructing climates of the Quaternary. Academic Press, San Francisco, USA

Brigode P, Bernardara P, Gailhard J, et al (2013a) Optimization of the geopotential heights information used in a rainfall-based weather patterns classification over Austria. Int J Climatol 33:1563–1573. doi: 10.1002/joc.3535

Brigode P, Oudin L, Perrin C (2013b) Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? J Hydrol 476:410–425. doi: 10.1016/j.jhydrol.2012.11.012

Brigode P, Paquet E, Bernardara P, et al (2015) Dependence of model-based extreme flood estimation on the calibration period: case study of the Kamp River (Austria). Hydrol Sci J 60:1424–1437. doi: 10.1080/02626667.2015.1006632

Caillouet L, Vidal J-P, Sauquet E, Graff B (2016) Probabilistic precipitation and temperature downscaling of the Twentieth Century Reanalysis over France. Clim Past 12:635–662. doi: 10.5194/cp-12-635-2016

Chardon J, Hingray B, Favre A-C, et al (2014) Spatial Similarity and Transferability of Analog Dates for Precipitation Downscaling over France. J Clim 27:5056–5074. doi: 10.1175/JCLI-D-13-00464.1

Chen J, Brissette FP, Lucas-Picher P (2015) Assessing the limits of bias-correcting climate model outputs for climate change impact studies. J Geophys Res Atmospheres 120:2014JD022635. doi: 10.1002/2014JD022635

Compo GP, Whitaker JS, Sardeshmukh PD, et al (2011) The Twentieth Century Reanalysis Project. Q J R Meteorol Soc 137:1–28. doi: 10.1002/qj.776

Coron L, Andréassian V, Perrin C, et al (2012) Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. Water Resour Res 48:W05552. doi: 10.1029/2011WR011721

Cowtan K, Way RG (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. Q J R Meteorol Soc 140:1935–1944. doi: 10.1002/qj.2297

Crooks SM, Kay AL (2015) Simulation of river flow in the Thames over 120 years: Evidence of change in rainfall-runoff response? J Hydrol Reg Stud 4, Part B:172–195. doi: 10.1016/j.ejrh.2015.05.014

Demarée GR, Ogilvie AEJ (2008) The Moravian missionaries at the Labrador coast and their centuries-long contribution to instrumental meteorological observations. Clim Change 91:423–450. doi: 10.1007/s10584-008-9420-2

Garçon R (1999) Modèle global pluie-débit pour la prévision et la prédétermination des crues. Houille Blanche 88–95. doi: 10.1051/lhb/1999088

George SS, Nielsen E (2003) Palaeoflood records for the Red River, Manitoba, Canada, derived from anatomical tree-ring signatures. The Holocene 13:547–555. doi: 10.1191/0959683603hl645rp

Gray ST, McCabe GJ (2010) A combined water balance and tree ring approach to understanding the potential hydrologic effects of climate change in the central Rocky Mountain region. Water Resour Res 46:W05513. doi: 10.1029/2008WR007650

GRDC (2015) Global Runoff Data Centre. Federal Institute of Hydrology (BfG), D-56002, Koblenz, Germany

Guay C, Minville M, Braun M (2015) A global portrait of hydrological changes at the 2050 horizon for the province of Québec. Can Water Resour J Rev Can Ressour Hydr 40:285–302. doi: 10.1080/07011784.2015.1043583

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J Hydrol 377:80–91. doi: 16/j.jhydrol.2009.08.003

Hernández-Henríquez MA, Mlynowski TJ, Déry SJ (2010) Reconstructing the Natural Streamflow of a Regulated River: A Case Study of La Grande Rivière, Québec, Canada. Can Water Resour J Rev Can Ressour Hydr 35:301–316. doi: 10.4296/cwrj3503301

Hirsch RM (1982) A comparison of four streamflow record extension techniques. Water Resour Res 18:1081–1088. doi: 10.1029/WR018i004p01081

Horton P, Jaboyedoff M, Metzger R, et al (2012) Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method. Nat Hazards Earth Syst Sci 12:777–784. doi: 10.5194/nhess-12-777-2012

Huang J-G, Bergeron Y, Berninger F, et al (2013) Impact of Future Climate on Radial Growth of Four Major Boreal Tree Species in the Eastern Canadian Boreal Forest. PLoS ONE 8:e56758. doi: 10.1371/journal.pone.0056758

Jandhyala VK, Liu P, Fotopoulos SB (2009) River stream flows in the northern Québec Labrador region: A multivariate change point analysis via maximum likelihood. Water Resour Res 45:W02408. doi: 10.1029/2007WR006499

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled SRTM for the globe Version 4.

Kuentz A, Mathevet T, Gailhard J, et al (2013) Over 100 years of climatic and hydrologic variability of a Mediterranean and mountainous watershed: the Durance River. In: Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections. IAHS Publications, Gothenburg, Sweden, pp 19–25

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 HA, Haston L, Michaelsen J (1993) Dendrohydrology and long-term hydrologic phenomena. Rev Geophys 31:151–171.

Mekis É, Vincent LA (2011) An Overview of the Second Generation Adjusted Daily Precipitation Dataset for Trend Analysis in Canada. Atmosphere-Ocean 49:163–177. doi: 10.1080/07055900.2011.583910

Meko DM, Woodhouse CA (2011) Application of Streamflow Reconstruction to Water Resources Management. In: Hughes MK, Swetnam TW, Diaz HF (eds) Dendroclimatology. Springer Netherlands, pp 231–261

Merz R, Parajka J, Blöschl G (2011) Time stability of catchment model parameters: Implications for climate impact analyses. Water Resour Res. doi: 10.1029/2010WR009505

Montanari A (2012) Hydrology of the Po River: looking for changing patterns in river discharge. Hydrol Earth Syst Sci 16:3739–3747. doi: 10.5194/hess-16-3739-2012

Nicault A, Boucher E, Bégin C, et al (2014) Hydrological reconstruction from tree-ring multi-proxies over the last two centuries at the Caniapiscau Reservoir, northern Québec, Canada. J Hydrol 513:435–445. doi: 10.1016/j.jhydrol.2014.03.054

Obled C, Bontron G, Garçon R (2002) Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach. Atmospheric Res 63:303–324. doi: 10.1016/S0169-8095(02)00038-8

Oudin L, Hervieu F, Michel C, et al (2005) Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. J Hydrol 303:290–306. doi: 10.1016/j.jhydrol.2004.08.026

Patskoski J, Sankarasubramanian A, Wang H (2015) Reconstructed streamflow using SST and tree-ring chronologies over the southeastern United States. J Hydrol 527:761–775. doi: 10.1016/j.jhydrol.2015.05.041

Perreault L, Garçon R, Gaudet J (2007) Analyse de séquences de variables aléatoires hydrologiques à l'aide de modèles de changement de régime exploitant des variables atmosphériques. Houille Blanche 111–123. doi: 10.1051/lhb:2007091

Perreault L, Parent É, Bernier J, et al (2000) Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. Stoch Environ Res Risk Assess 14:243–261. doi: 10.1007/s004770000051

Perrin C, Andréassian V, Rojas Serna C, et al (2008) Discrete parameterization of hydrological models: Evaluating the use of parameter sets libraries over 900 catchments. Water Resour Res 44:W08447. doi: 10.1029/2007WR006579

Perrin C, Michel C, Andréassian V (2003) Improvement of a parsimonious model for streamflow simulation. J Hydrol 279:275–289. doi: 10.1016/S0022-1694(03)00225-7

R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Radanovics S, Vidal J-P, Sauquet E, et al (2013) Optimising predictor domains for spatially coherent precipitation downscaling. Hydrol Earth Syst Sci 17:4189–4208. doi: 10.5194/hess-17-4189-2013

Rohde R, Muller RA, Jacobsen R, et al (2013) A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. Geoinformatics Geostat Overv 1:1–7. doi: 10.4172/2327-4581.1000101

Saito L, Biondi F, Devkota R, et al (2015) A water balance approach for reconstructing streamflow using tree-ring proxy records. J Hydrol 529, Part 2:535–547. doi: 10.1016/j.jhydrol.2014.11.022

Seiller G, Anctil F, Perrin C (2012) Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. Hydrol Earth Syst Sci 16:1171–1189. doi: 10.5194/hess-16-1171-2012

Slonosky VC (2014) Daily minimum and maximum temperature in the St-Lawrence Valley, Quebec: two centuries of climatic observations from Canada. Int J Climatol. doi: 10.1002/joc.4085

Subedi N, Sharma M (2013) Climate-diameter growth relationships of black spruce and jack pine trees in boreal Ontario, Canada. Glob Change Biol 19:505–516. doi: 10.1111/gcb.12033

Tapsoba D, Fortin V, Anctil F, Haché M (2005) Apport de la technique du krigeage avec dérive externe pour une cartographie raisonnée de l'équivalent en eau de la neige : Application aux bassins de la rivière Gatineau. Can J Civ Eng 32:289–297. doi: 10.1139/l04-110

Teng J, Potter NJ, Chiew FHS, et al (2015) How does bias correction of regional climate model precipitation affect modelled runoff? Hydrol Earth Syst Sci 19:711–728. doi: 10.5194/hess-19-711-2015

Teutschbein C, Seibert J (2013) Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? Hydrol Earth Syst Sci 17:5061–5077. doi: 10.5194/hess-17-5061-2013

Teweles J, Wobus H (1954) Verification of prognosis charts. Bull Am Meteorol Soc 35:455–463.

Thorndycraft VR, Benito G, Rico M, et al (2005) A long-term flood discharge record derived from slackwater flood deposits of the Llobregat River, NE Spain. J Hydrol 313:16–31. doi: 10.1016/j.jhydrol.2005.02.003

Valéry A, Andréassian V, Perrin C (2014) "As simple as possible but not simpler": what is useful in a temperature-based snow-accounting routine? Part 2 - Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. J Hydrol. doi: 10.1016/j.jhydrol.2014.04.058

Velázquez JA, Troin M, Caya D, Brissette F (2015) Evaluating the Time-Invariance Hypothesis of Climate Model Bias Correction: Implications for Hydrological Impact Studies. J Hydrometeorol 16:2013–2026. doi: 10.1175/JHM-D-14-0159.1

Vincent LA, Wang XL, Milewska EJ, et al (2012) A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis. J Geophys Res Atmospheres 117:D18110. doi: 10.1029/2012JD017859

Way RG, Viau AE (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

Wetterhall F, Halldin S, Xu C (2005) Statistical precipitation downscaling in central Sweden with the analogue method. J Hydrol 306:174–190. doi: 10.1016/j.jhydrol.2004.09.008

Wilson CV (1988) The summer season along the east coast of Hudson Bay during the nineteenth century. Part III: Summer thermal and wetness indices B. The indices, 1800-1900. Environment Canada, Downsview, Ontario, Canada