

Hydrol. Earth Syst. Sci. Discuss., author comment AC3
<https://doi.org/10.5194/hess-2021-71-AC3>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Emmanuel Dubois et al.

Author comment on "Simulation of long-term spatiotemporal variations in regional-scale groundwater recharge: contributions of a water budget approach in cold and humid climates" by Emmanuel Dubois et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-71-AC3>, 2021

Dear Referee,

We would like to thank you for your feedbacks and propositions of improvement for our manuscript. We will take into account all your comments in our modifications. Hereafter we describe the main modifications that will be made to the document based on your comments. The general comments are addressed first (labelled C#), followed by the minor remarks (labelled mC#).

C1: In response to reviewer #1, comment 5, you provide a helpful table of how calibration results change with different weights in the objective function. You say that the calibration is "moderately sensitive to the weights" – but what implication does this have for the results? Do these results all fall within the sensitivity analyses? The key question here is does changing the objective function change the GWR estimates and trends? Would your interpretations and conclusions change?

A1: We would like to thank the Referee for this helpful comment. To demonstrate that these weights have only a "moderate" impact on the results, we are planning to further amend the supplementary Table A1 presented in our previous answer to include the interannual potential GWR. As well, we computed the trends on the potential GWR (Mann-Kendall test, significant if the p-value < 0.05) and since they were systematically positive but not significant, we did not include them in the table. The L203-205 will be rephrased as follows: "The weights attributed to each objective function in KGE_{mean} were arbitrarily chosen to select the calibrated parameter set that maximizes the quality of simulated baseflows, considered as a proxy for GWR (KGE_{qbase}), without losing the benefits of the multi-objective optimization. Supplementary Table A1 shows that the objective functions, the calibrated parameters, and the interannual potential GWR values were only moderately sensitive to the weights, with a variation of potential GWR between 176 mm/yr and 194 mm/yr when for the three tested weights (40%-60%, 50%-50% and 60%-40% for KGE_{qtot} and KGE_{qbase} respectively). No significant trends in annual potential GWR (Mann-Kendall test, significant if the p-value < 0.05) were found within these 75 model runs (all positive but not significant)."

Table A. 1: Best compromises, interannual potential GWR, and optimized parameters obtained with the best compromises for the gauging stations of W6 and ranges obtained with the 25 best compromises (before regionalization) using different calibration weights and different baseflow separation methods.

	Calibration	Validation	Interannual pot. GWR (mm/yr)
	KGE_{mean}	KGE_{qtot}	KGE_{qbase}
Lyne and Hollick - KGE _{mean} = 0.4 x KGE _{qtot} + 0.6 x KGE _{qbase}			
Best compromise	0.76 ₁	0.80 ₇	0.72 ₉
25 best compromises	0.75 ₈ - 0.76 ₁	0.78 ₁ - 0.82 ₀	0.71 ₈ - 0.74 ₅

Lyne and Hollick -
KGE_{mean} = 0.5 x
KGE_{qtot} + 0.5 x
KGE_{qbase}

Best compromise 0.77₂ 0.81₉ 0.72₅

25 best compromises 0.77₀ - 0.77₂ 0.80₅ - 0.82₅ 0.71₈ - 0.73₇

Lyne and Hollick -
 $KGE_{\text{mean}} = 0.6 \times$
 $KGE_{\text{qtot}} + 0.4 \times$
 KGE_{qbase}

Best compromise 0.78₂ 0.82₅ 0.71₉

25 best compromises 0.78₀ - 0.78₂ 0.81₄ - 0.83₅ 0.69₉ - 0.73₀

Eckhardt - $KGE_{\text{mean}} =$
 $0.4 \times KGE_{\text{qtot}} + 0.6 \times$
 KGE_{qbase}

Best compromise 0.81₁ 0.87₀ 0.77₂

25 best compromises 0.81₀ - 0.81₁

0.86₇ - 0.87₂

0.76₉ - 0.77₂

Chapman- $KGE_{\text{mean}} =$
 $0.4 \times KGE_{\text{qtot}} + 0.6 \times$
 KGE_{qbase}

Best compromise 0.80₁

0.87₀

0.75₆

25 best compromises 0.79₈ - 0.80₁

0.86₉ - 0.87₅

0.74₈ - 0.75₆

Optimized parameters

Lyne and Hollick - $KGE_{mean} = 0.4 \times KGE_{tot} + 0.6 \times KGE_{base}$

Best compromise

25 best compromises

Lyne and Hollick - $KGE_{mean} = 0.5 \times KGE_{tot} + 0.5 \times KGE_{base}$

Best compromise

25 best compromises

Lyne and Hollick - $KGE_{mean} = 0.6 \times KGE_{tot} + 0.4 \times KGE_{base}$

Best compromise

25 best compromises

Eckhardt - $KGE_{mean} = 0.4 \times KGE_{tot} + 0.6 \times KGE_{base}$

Best compromise

25 best compromises

Chapman - $KGE_{mean} = 0.4 \times KGE_{tot} + 0.6 \times KGE_{base}$

Best compromise

25 best compromises

C2: Further to #1, and in line with reviewer #1's comment 4, what effect does the

selection of baseflow separation method have on the results? As you point out in the discussion and in the response to reviewer #1 there can be significant variance in baseflow estimates between different methods. If you selected a different method, would the interpretation change? I acknowledge that these analyses may be time consuming, but I think they speak to the robustness of the approach. These baseflow estimates come with such high uncertainty that I feel they should almost be treated as another parameter – how would your results change if baseflow varied?

A2: We understand the Referee's concern about that the choice of the baseflow separation method and its impact on the simulation results. We chose to use the Lyne and Hollick filter (1979) with the proposition of Ladson et al. (2013 – cited in our manuscript) with a stochastic calibration and 30 passes of the filter. When compared to other options, this filter appeared to provide the "most likely" smoothed groundwater signal at a daily time step (Figure 1 – the figure will not be presented in our manuscript). When aggregated to monthly time steps, the higher baseflow peaks of the Eckhardt (2005) and Chapman (1991) filters translated in higher baseflows during the high flow periods (mainly during spring) (Figure 2 – the figure will not be presented in our manuscript). Nevertheless, we compared the effect of the choice of the baseflow separation method on the calibration and the GWR estimates and trends for the gauging stations from W6 in the same manner as we compared the effect of the calibration weights to the results. We performed an automatic calibration with the Eckhardt and Chapman filters for the baseflow estimates and extracted the 25 best compromises to compare the results. The comparison is reported in the supplementary Table A1 presented above. As the absolute value of baseflow changes with the filter method, the interannual GWR and calibration parameters vary in consequence (Figure 3 - the figure will not be presented in our manuscript). Similarly to the simulations from the model calibrated with the Lyne and Hollick filter, all trends in simulated GWR based on the Eckhardt and Chapman filters are positive but not significant (Mann-Kendall test, significant if the p-value < 0.05). The objective functions remain similar, between 0.7 and 0.8, except for KGE_{qbase} in validation which is slightly lower for Eckhardt and Chapman than for Lyne and Hollick.

To include these new results in the manuscript, the paragraph L178-182 will be rephrased into: "The daily flow rates at the outlets were those from the 51 gauging stations (Fig. 1). Baseflows were estimated from the flow rate time series following the proposition of Ladson et al. (2013) of a standard approach of the Lyne and Hollick filter (Lyne and Hollick, 1979), using a stochastic calibration and 30 filter passes. To compare the effects of the baseflow filters on GWR simulation, baseflows were also estimated with the Eckhardt filter (2005) and the Chapman filter (1991) for the gauging stations of W6. Total flows and baseflows were divided by the area of the given watershed to provide flow values in mm/yr, and thus facilitate the comparison of calibration results between watersheds of very different sizes."

L472-483 will also be rephrased into: "The simulated transient potential GWR was calibrated with baseflows computed using regressive filters from river flow rate time series. Although Partington et al. (2012) showed that the association between baseflows and GWR is not always satisfactory and different baseflow separation methods can lead to differences in volumes and timing (Gonzales et al., 2009; Zhang et al., 2017), baseflows are generally considered to be an acceptable proxy for GWR in cold and humid climates similar to that of Canada (Chemingui et al., 2015; Rivard et al., 2009). They are widely used for the calibration of GWR simulation (Batelaan and De Smedt, 2007; Croteau et al., 2010; Dripps and Bradbury, 2007; Gagné et al., 2018; Rivard et al., 2013). Using baseflow estimated based on the Lyne and Hollick (1979), Eckhardt (2005), and Chapman (1991) filters influenced the proportion of baseflow in total river flow, and consequently led to variations in simulated potential GWR (Supplementary Table A1). However, the

quality of the simulations were comparable due to parameter variations that remained in the possible range of the parameters (Table 2) and trends in the simulated potential GWR remained positive but not significant with the three methods. The structure of water budget models makes them compute GWR as the residual of the water budget, thus propagating the computational error from the other terms of the water budget, namely runoff, interception, evapotranspiration, subsurface runoff (depending on the model complexity) onto GWR estimation (Crosbie et al., 2015; Scanlon et al., 2002). Using baseflows as calibration data for GWR simulation limits the modeling error of GWR rates, but makes them dependant on the baseflow separation method.”

C3: I really appreciate how this work can take trends in baseflow (estimated from measurements at sparse gage stations) and use that to estimate the spatial distribution of potential groundwater recharge across a watershed. However, I think the manuscript would benefit from a more explicit discussion of the uncertainties that propagate through the workflow. Again, this kind of analysis and discussion really helps to make the results more robust and applicable.

A3: We thank the Referee for pointing out that an explicit discussion about uncertainty was missing. To quantify the uncertainty, potential GWR was simulated over the study area with the 100 best sets of regionalized parameters and the uncertainty was estimated with the standard deviation between the 100 model runs for each month. As well, to give an idea of the error comparison between the input data and the simulated variables, the mean bias will be compared (the first Referee suggested to add the mean bias of the input data).

This explanation about the uncertainty computation will be added after L205: “The 100 best compromises of each group of gauging stations were used to produce the 100 best regionalized parameter sets and the HB model was run with these parameters, estimating uncertainty from the standard deviation.”

As well, these sentences will be added after L245: “The mean bias on simulated river flow and potential GWR, computed over the entire period of observation data, and averaged by group of gauging stations, was comprised between -9 mm/month and 5 mm/month (Table 3). The uncertainty computed with the 100 best regionalized parameter sets was ≤ 10 mm/yr for the three simulated variables (Table 4).”

Finally, this sentence will be inserted after L440: “Furthermore, the uncertainty analysis showed that the calibration method allowed reaching acceptable uncertainty levels in the simulated variables and that the mean bias in the simulated river flow and potential GWR was similar to that of the mean bias in the input data.”

Table 3 will be modified as follows:

Table 3: Objective functions for the simulated outputs, for calibration and validation periods, and mean bias over the entire period of measurement

Gauging stations	Calibation	Validation	Mean bias (all period - mm/month)
-------------------------	-------------------	-------------------	--

Number	Measur. period	KGE_{qtot}	KGE_{qbase}
W1*	2	1980-2013	0.80
W2*	5	1973-2017	0.76
W3	14	1965-2017	0.76
W4*	8	1961-2009	0.81
W5	4	1961-2017	0.77
W6	8	1961-2017	0.84

W7	2	1993-2017	0.92
W8	8	1961-2015	0.80

**The presented values are for the stations of which the watershed are completely located in Quebec*

Table 4 will be modified as follows to include uncertainty in runoff, AET and potential GWR:

Table 4: Simulated runoff, actual evapotranspiration (AET), and potential groundwater recharge (GWR) and uncertainty for the study area between 1961 and 2017, in mm/year and in percentage for winter, spring, summer and fall, and for the eight watersheds (W1 to W8).

	Runoff	AET	Pot. GWR
mm/yr	Win.	Spr.	Sum.
W1*	368 ± 8	13%	47%
W2*	430 ± 9	12%	48%

W3	442 ± 10	12%	48%
W4*	485 ± 9	11%	50%
W5	438 ± 10	10%	50%
W6	465 ± 9	8%	53%
W7	453 ± 8	6%	56%
W8	470 ± 10	7%	53%

**Part of the watershed is located in the USA - the presented values are for the Quebec part only*

Winter: December, January, February; spring: March, April, May; summer: June, July,

August; fall: September, October, November

C4: *You compare results from this work to results from previous studies – a lot of them spatially-distributed numerical models. Can you provide some idea of how the spatial distribution of GWR compares between them, as opposed to just the mean/variance/ranges? Are you picking up the same high/low recharge patterns? The same temporal distributions? This is hinted in section 5.1 but I think a more robust comparison would improve the manuscript.*

A4: We understand the Referee's point of view and agree that such analysis would be interesting. However, following comments from the first Referee to make our paper more interesting for a wider audience and less focused on the study area, we will develop the comparison with other research studies on GWR in cold and humid climates. Therefore, we do not wish to elaborate further on the comparison with Quebec-based studies.

Minor comments:

mC1: *Line 60-61: It is bizarre that the reference that "acknowledges the lack of representiveness of the daily results...." Is from 2007, yet the reference for the model they are acknowledging is from 2010.*

Indeed, the paper from Dripps and Bradburry (2007 – full reference in our manuscript) was the reference for the SWB model. Therefore, the citation of Westenbroek et al., (2010) is not necessary and will be removed from the L61.

mC2: *Line 80-81: I think this is needed in a lot of places, not just southern QC!*

L80-81 will be rephrased as follows : "There is clearly a need for reliable regional-scale estimates of GWR that can be updated relatively easily using widely available data in cold and humid climates where groundwater is often the main source of drinking water in rural areas (Groupe Agéco, 2019; Kløve et al., 2017)."

mC3: *Line 304-305: So it is associated more with the precip trends than with the soil type?*

Clayey areas are mainly located in the St. Lawrence Lowlands, the flattest of the study area, that also happen to receive the less rainfall (mainly < 1 000 mm/yr) in the study area. The combination of the three (precipitation distribution, soil type, and flat topography) explains that besides having the smallest potential GWR rates of the study area, clayey areas are also associated with the smallest runoff rates and AET rates.

Therefore, we do not think that the text need to be modified.

mC4: *Section 4.4; 1st para: This is a rough paragraph to read. I think a table or graph would be more appropriate.*

The first paragraph of section 4.4 L310-321 will be rephrased as follows: "HydroBudget simulated the temporal evolution of the water budget in the study area since the 1960s, thus producing an exceptionally long simulated time series of runoff, AET, and potential GWR for the area (Fig. 8). The effect of interannual variability in precipitation was clear in the simulated runoff with low runoff rates (< 350 mm/yr) produced during the driest year and high runoff rates (> 550 mm/yr) produced during the wettest year (Fig. 8a, b). The simulated AET varied less than runoff, mainly between 450 mm/yr and 560 mm/yr (Fig. 8c). Variations of potential GWR were relatively high, comprised between 90 mm/yr and 200 mm/yr, and seemed more influenced by precipitation than by temperature variations (Fig. 8c)."

mC5: *Line 343-344: "somewhat higher" is almost double.*

L343-346 will be rephrased as follows: "This value is much higher than that obtained using HB in the same area (109 mm/yr), but the resulting preferential recharge areas located close to the Canada-USA border are similar with both approaches (i.e., 70 mm/yr to 250 mm/yr in Chemingui et al. (2015) and 70 mm/yr to 345 280 mm/yr with HB).

mC6: *Line 381-382: awkward sentence structure. Maybe provide range in main sentence text (between 89 and 198) and then average in brackets?*

L381-382 will be rephrased into: "The annual variability in potential GWR is closely linked to that of total precipitation, with maximum variation between 89 mm/yr (1968) and 198 mm/yr (1983)."

mC7: *Line 387: It is hard to support that this is the novel contribution when you follow it up with the fact that it is supported by a lot of previous literature. I would consider rephrasing this to highlight the specific contribution beyond what the existing literature provides.*

L387-390 will be rephrased into: "This spatiotemporal link between GWR and precipitation and temperature patterns is coherent with that reported in other studies, both in similar and different climate and geological environments (Abdollahi et al., 2017; Ashaolu et al., 2020; Chemingui et al., 2015; Fu et al., 2019; Hayashi and Farrow, 2014; Hu et al., 2019). However, one of the contributions of this work was to show that temperature

influenced GWR rates at the watershed scale and at the regional scale, both temporally (seasonality) and spatially (increase of GWR from warmer to colder climates).”

mC8: Line 399: This is a regional contribution. While significant to those interested in this region, authors should try to focus on the broader applicability and contributions.

The paragraph L398-407 will be rephrased into: “The increasing trends in simulated winter potential GWR are most likely related to that of winter temperature and precipitation between 1961 and 2017. A contribution of this work is to show that the absence of decreasing trends in potential GWR despite the overall statistically significant increases in temperature and AET indicates that the increase in precipitation (vertical inflows in Table A. 2) is large enough to compensate for the increases in AET. These observations were also true when comparing two decades. Similarly, Levison et al. (2016) and Rivard et al. (2009) did not find significant trend in baseflow during the second half of the 20th century. Inversely, Nygren et al. (2020) found that groundwater levels significantly decreased in Sweden and Finland between 1980-1989 and 2001-2010 as a result of shifting from cold to temperate climate, inducing a change of the GWR dynamic from snowmelt-dominated to rain-dominated. In comparison, the temperature increase observed in southern Quebec over the 1961-2017 period was probably not large enough to produce such a change. An original contribution of this work was to show that the simulation of GWR with a water budget model allows associating changes in the climatic conditions to long-term trends in the simulated variables and their impacts on the regional hydrological dynamic.”

mC9: Line 439: I generally try to avoid using unquantifiable terms like “good” – how else can you describe this that can be backed up by the results?

The term “good” L439 will be changed for “satisfying”.

mC10: Line 490: ‘work’ instead of ‘word’

The term “word” L490 will be changed for “work”.

mC11: Line 515: Could this have been interpreted from trends in baseflow? Or does this work provide additional info that goes beyond that provided by baseflow results alone?

We do think that this work provides more information on trends in GWR than an analysis of trends in baseflow for several reasons: 1) GWR is simulated continuously over the study area for the 1961-2017 period while baseflows estimated from the river flow time series are only available for periods of time (gaps in the time series and abandonment of

gauging stations). Rivard et al (2009 – reference in our manuscript) did not find any consistent trend in Eastern Canada based on baseflow analysis. As well, computing trends on the simulated runoff and AET allows understanding better how the trends in the climate data (increase of precipitation and warming temperature) translated into the regional hydrological dynamic (increase of runoff and AET, but not GWR). 2) River flow measurements during winter are highly uncertain due to ice cover and ice flowing, therefore analysing trends from the raw data only during these periods could be misleading. This point has been added in our proposition of modification of paragraph L398-407.

Supplementary references:

Chapman, T. G.: Comment on "Evaluation of automated techniques for base flow and recession analyses" by R. J. Nathan and T. A. McMahon, 27, 1783–1784, <https://doi.org/10.1029/91WR01007>, 1991.

Eckhardt, K.: How to construct recursive digital filters for baseflow separation, 19, 507–515, <https://doi.org/10.1002/hyp.5675>, 2005.

Groupe Agéco: Recherche participative d'alternatives durables pour la gestion de l'eau en milieu agricole dans un contexte de changement climatique (RADEAU1) (participative research for sustainable options in water management in agricole region and within a climate change context), Ministère de l'agriculture, des pêcheries et de l'alimentation, Fonds Vert, Quebec City (Canada), 2019.

Kløve, B., Kvitsand, H. M. L., Pitkänen, T., Gunnarsdottir, M. J., Gaut, S., Gardarsson, S. M., Rossi, P. M., and Miettinen, I.: Overview of groundwater sources and water-supply systems, and associated microbial pollution, in Finland, Norway and Iceland, *Hydrogeol J*, 25, 1033–1044, <https://doi.org/10.1007/s10040-017-1552-x>, 2017.

Nygren, M., Giese, M., Kløve, B., Haaf, E., Rossi, P. M., and Barthel, R.: Changes in seasonality of groundwater level fluctuations in a temperate-cold climate transition zone, *Journal of Hydrology X*, 8, 100062, <https://doi.org/10.1016/j.hydroa.2020.100062>, 2020.

Please also note the supplement to this comment:

<https://hess.copernicus.org/preprints/hess-2021-71/hess-2021-71-AC3-supplement.pdf>