Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-508-AC1, 2017

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "A Hydrological Prediction System Based on the SVS Land–Surface Scheme: Implementation and Evaluation of the GEM-Hydro platform on the watershed of Lake Ontario" *by* Étienne Gaborit et al.

Étienne Gaborit et al.

etienne.gaborit.5@gmail.com

Received and published: 8 February 2017

Answers to the reviewer follow the "AR" letters.

Hydrol. Earth Syst. Sci. Discuss. Manuscript Number: doi:10.5194/hess-2016-508, 2016 Title: A Hydrological Prediction System Based on the SVS Land-Surface Scheme: Implementation and Evaluation of the GEM-Hydro platform on the watershed of Lake Ontario General comments This is an interesting paper describing the devel-

Printer-friendly version

Discussion paper

opment of a modelling system to estimate net basin supply to a strategic Canada/USA water body: Lake Ontario. From a technological point of view, the work is state of the art and clearly highlights recent efforts undertaken by Environment and Climate Change Canada to develop a robust hydrological modelling system. More specifically, the objectives of this paper are to: (i) ''propose a methodology for calibrating the distributed GEM-Hydro platform developed by ECCC in order to improve streamflow simulations for Lake Ontario, which we expect would ultimately propagate into improved simulations of Lake Ontario Net Basin Supplies (or NBS, the sum of lake tributary runoff, overlake precipitation, and overlake evaporation: Brinkmann 1983); (ii) compare GEM-Hydro with two other distributed models (inter-comparison study) in order to identify avenues to further improve GEM-Hydro; and (iii) propose and evaluate a method for estimating runoff for the ungauged parts of the watershed."

Objective (i) As far as I can get, the paper never actually demonstrated what is highlighted in yellow. So unless the reader is an insider, there is no evidence that currently applied modelling systems do not provide satisfactorily the sought-after streamflow simulations. I understand with the interest of developing a state-of-the-art modelling system, but the paper does not provide a strong motivation. The authors need to convince the readers here.

AR: after the answers to the first reviewer, the sentence highlighted in yellow was removed, and the first objective was reformulated as follows: " this study mainly aims at finding a methodology to implement the distributed GEM-Hydro model over the whole Lake Ontario watershed, including its ungauged parts, in an efficient manner. Distributed models are more complicated to implement and more computationally-intensive than lumped ones, but have a broader range of applications." However, it is true that this work did improve streamflow simulations in comparison to former studies, which mainly consist of those of Croley and Haghnegahdar, as mentioned in the introduction of the first GRIP-O paper (Gaborit et al., 2016 a): "In the Lake Ontario watershed, Croley (1983) and Haghnegahdar et al. (2014) calibrated hydrologic models with runoff observations to optimize simulations of this variable. The former demonstrated that the LBRM worked well for simulating weekly flows, while the second evaluated the MESH model on 15 Great Lakes subbasins, including two Lake Ontario subbasins. However, even after calibration, the MESH model did not perform particularly well for validation catchments (Nash-Sutcliffe values below 0.6)." Therefore, an additional sentence was added in results section, close to p.12, l.10: "Therefore, GRIP-O allowed to improve streamflow simulations for the Lake Ontario basin, in comparison to the studies of Croley (1983) and Haghnegahdar et al. (2014), which are the main former studies who proposed the implementation of hydrologic models over this area."

Objective (ii) After reading the paper a couple of times, I feel the paper has yet to actually and clearly identify the avenues to further improve GEM-Hydro. The authors mentioned that SVS would benefit from a soil heat balance equation based on one missed spring peakflow; that is farfetched as it could have resulted from multiple sources, so I believe the paper does not make a strong case here. Now, I am not sure the reader will get anything out of this objective and, at best, I think the model intercomparison should be considered as supplemental material. Furthermore, neither the abstract, nor the conclusion, provide any tangible answers to this objective.

AR: We do not agree with the reviewer that this objective is useless despite we have only results for 3 subbasins for the intercomparison of the models. As clearly stated in the text, this inter-comparison allows to emphasize these important facts: "Even if the intercomparison is obviously limited in the number of available test cases, it allows highlighting the mandatory need of calibrating hydrologic models, that models have unique behaviors that translate in substantial differences in hydrographs, and that each of the models could benefit from some strengths of its competitors". It does also show that GEM-Hydro is a priori competitive with the two other main distributed models used in this region. In regard of the avenues for improving GEM-Hydro, we do not claim that the difference between the model simulations is due to the lack of soil freezing and melting processes in SVS, as can be seen in the following sentence with the use "which may be due to": "Peak flow events associated to the spring freshet are generally better represented by MESH, which may be due to a better representation of the soil freezing and melting processes occurring in CLASS (MESH LSS)." Moreover, we do not make this assumption based on one spring peak flow event, but we have three subbasins with 7 years of simulations, which does begin to be a significant number of events. Finally, it is true that SVS should benefit from the implementation of the soil freezing and melting processes, because it is the aim of this LSS to represent physical processes. The potential benefit of a more complex snow module is also straightforward in comparison to the current force-restore approach (which only relies on an average snowpack temperature to estimate snowmelt), given the complexity of the physical processes related to snow. Therefore, we do believe that these two avenues consist in some of the main ones to further improve SVS. We did add another avenue of improvement for SVS with the following sentence (p. 15, l.18, in the conclusion): "Finally, work is under way to represent a surface of variable area of ponded water in each SVS grid cell, in order to represent subgrid-scale lakes, wetlands, and to better represent the delay associated with surface runoff transfer into streams." 2 The above comments might seem harsh, but they are factual. Do not get me wrong I think the paper reports relevant technological information to the hydrometeorological community. The paper shows promises and I believe the authors can streamline the content to the essentials, that is the demonstration of the advantage of substituting a unit hydrograph for WATROUTE during calibration can actually reduce the computational time required for model calibration and illustrate that GEM-Hydro can benefit from a local/global calibration strategy to provide ‘’ good streamflow predictions ”. Should the authors decide to follow this suggestion, they should substantiate their rationale from a scientific point of view rather than a technological one.

For example, they should provide more fundamental information between the computational time scales of the LSS and those of WATROUTE and UH. Furthermore, they should discuss the relationship between the computational time scales and the dimension of the computational elements used in WATROUTE and the UH versus those used in the LSS.

AR: Information was added on the computational times required by the system. (see specific comment on p.5, lines 5-34).

It is noteworthy, Section 1.4. is confusing. In fact, although I am somewhat familiar with the work of the first author on local and global calibration, but I needed to read the section more than once to comprehend; and even then, I am still not sure what was actually done.

AR: we know this section is hard to follow but we did prefer explain clearly what was done rather than take the risk to leave the reader with unanswered questions regarding the methodology. Moreover, the methodology proposed also illustrates that different possibilities do exist to fulfill this kind of work, and that it is important to envision the different ones before starting the work, which is still an important message to remind to modelers.

I do not quite follow the use of GR4J in the calibration of GEM-Hydro-UH (see my comment below related to the content of P.10, lines 15-17) because the paper does not provide the underlying hypothesis.

AR: see answer to specific comment on page.10, line 16.

I strongly recommend the authors to provide a diagram or sketch describing the steps taken to achieve the calibration strategy introduced in Section 1.4. At this point, I doubt that most readers can appreciate what the authors actually did.

AR: we will let the publishing authority decide on this but we do believe that the text of section 1.4 accompanied with Fig. 1 does allow to understand.

I have made suggestions in the following list of specific comments below on how to fulfill what I perceive as shortcomings. I have also added a few editorial comments to improve the paper. As a side note, I found a bit difficult the exercise of going back and forth between the content of the introductory sections to remind myself what the acronyms meant. The authors should provide a list of acronyms to facilitate the reading of the manuscript.

AR: A list of acronyms will be put in supplementary material.

I strongly encourage the authors to address these comments as I feel the paper could certainly be a good technological contribution to the hydrometeorological community. Specific comments âĂć P. 2, line 10: the following sentence: o Âń Going beyond anomaly forecasts (which are bias corrected based on a model climatology) to obtain unbiased short-term streamflow forecasts is more challenging due to limitations of operational Land-Surface Schemes (LSS), which are generally geared towards improving weather forecasts, sometimes at the cost of not representing (or misrepresenting) surface and subsurface hydrological processes that are critical to hydrological simulation. Âż ïĆğïĂă. . .should be modified as follows:

âĂć Going beyond anomaly forecasts, which are bias-corrected based on a modeled climatology to obtain unbiased short-term streamflow forecasts, is more challenging. This is due to limitations of operational Land-Surface Schemes (LSS), which are generally geared towards improving weather forecasts, sometimes at the cost of not representing (or misrepresenting) critical surface and subsurface hydrological processes.

AR: this sentence was removed from the text due to reviewer 1 comments who suggested to shorten the introduction.

âĂć P.2, line 15: the following sentences: o Hydrological processes in land-surface models used for NWP are improving quickly (Balsamo et al., 2009; Masson et al., 2013; Alavi et al., 2016; Wagner et al., 2016), as soil water content and snow are recognized as important sources of their predictability that remain to be fully tapped into (Koster et al., 2004; Entekhabi et al., 2010). Environment and Climate Change Canada (ECCC), the Canadian department that provides operational weather and environmental forecasts, is in the process of implementing a major upgrade to the LSS used by its NWP model, the Global Environmental Multi-scale model (GEM) ïĆğïĂă. . .should be modified as follows: âĂć Hydrological processes simulated by land-surface schemes (LSS) used for NWP are improving quickly (Balsamo et al., 2009; Masson et al., 2013; Alavi et al., 2016; Wagner et al., 2016), as soil water content and snow water equivalent are recognized as key state variables for streamflow forecasting (Koster et al., 2004; Entekhabi et al., 2010). Environment and Climate Change Canada (ECCC), which provides operational weather and environmental forecasts within its boundary, is in the process of implementing a major upgrade to the LSS of the Global Environmental Multi-scale model (GEM), the national model.

AR: this was done.

âĂć P. 2, line 20: please delete Âń . . .in order. . . Âż

AR: done

âĂć P.3, line 6: please modify as follows: Âń . . .thermodynamics, as reported by Wiley. . . Âż

AR: done

âĂć P.5, line 6: please correct me if I am wrong, but WATFLOOD has no LSS, just a simple potential evapotranspiration equation, unless WATCLASS was used. So WATFLOOD is more along the line of GR4J with that respect.

AR: Yes, but it is still distributed and in this sense is very different from GR4J. However, the sentence below was added close to page 5, line 8: "It relies on the GRUs concept and on many empirical equations."

âĂć P.5, lines 5 through 34, I think there is room here to provide more fundamental information between the computational time scales of the LSS and those of WATROUTE and UH. o Furthermore, discuss the relationship between the computational time scales and the dimension of the computational elements used in WATROUTE and the UH.

AR: the sentences below were elongated at page 5, line 13 and page 5, line 23: " Sensitivity tests (Gaborit et al., 2016 b) revealed that 2 and 10 arcmin resolutions for SVS lead to quite similar performance in terms of streamflow at the outlet, while a substantial amount of computational time is saved when running the coarser resolution (almost proportionally if using the same number of nodes)." "As the GEM-Hydro suite (including WATROUTE) is quite demanding in terms of computational time, it was decided to test a stand-alone configuration of GEM-Hydro relying on text files only and in which WATROUTE is replaced by a Unit Hydrograph (UH).This version is here forth referred to GEM-Hydro-UH"

And the whole following paragraph was added close to page 5, line 23:

" Indeed, the computational time for the experiment setup described here and when splitting the domain in four on an ECCC supercomputer is about 1.5 min per day for the LSS part of GEM-Hydro (SVS), provided that the pre-processing of the atmospheric variables was already done (which is the case in calibration: the pre-processing is done only once). The WATROUTE code is not yet parallelized, each grid point being processed from upstream to downstream, but requires only 25s per day for the setup described here when running on a local machine. However, the WATROUTE pre-processing (i.e., preparation of the WATROUTE input files from the SVS outputs) is very long - about 30s per day. Therefore, WATROUTE computational time was still lower than SVS one for this setup and was not the limiting factor. One simulation run over the GRIP-O period (4.5 years) therefore requires about 1.7 days with GEM-Hydro and prevents from performing any automatic calibration (which requires at least 400 runs, see below). Instead of using GEM-Hydro to run SVS, a stand-alone SVS version, coded in Fortran, was used. This executable saves a tremendous amount of computation time compared to GEM-Hydro mainly because of the Input/Output processing time: the stand-alone version makes use of text files which are kept open during the simulation and requires only 4.5s per day on a local machine for this setup (2 h for the 4.5 years GRIP-O period or 30 days of calibration with 400 runs if running the

Printer-friendly version

Discussion paper

whole domain). However, the computational time required by WATROUTE still had to be bypassed to perform automatic calibrations, which was done with the UH concept."

âĂć P.6, line 10: what does SA mean in SA-MESH?

AR: Stand-alone, but it was removed for consistency throughout the paper.

âĂć P.6, line 21: please replace Âń... the outlet of Lake Ontario.Âż by Âń..the Lake outlet. Âż AR: done 4 âĂć P.7, line 10: please replace Âń . . . that it is higher than 1 m.. Âż by Âń . . . that it is greater than 1 m.. Âż

AR: done

âĂć P.7, line 20: is there any spin-up for the calibration period?

AR: p. 7 line 24, it was specified that yes, we do use spin-up for the calibration. " Validation is from June 1st, 2005 to June 1st, 2007 (2 years, last one being used as spin-up for calibration)" and right after: " Note that during the automatic calibrations, the spin-up year was simulated only once and for all subsequent runs."

âĂć P.7, equation (1): why presenting the PBIAS expression and not the NS...the latter being more complex than the former...

AR: but way more used in the hydrologic community than this "percent" normalized BIAS criteria.

âĂć P.8, line 3: o Please specify those GEM-Hydro-UH GRIP-O sub-basins that were locally calibrated out of all sub-basins? ïĆğïĂăWhat is the the percentage of the Lake Ontario basin that had local model calibration? Is it 88.5% (P.9, line 15) ?

AR: it was emphasized that local calibration was not achieved for all sub-basins "but only those shown on Fig. 3" at this location.

âĂć P.9, line 1: Âń...some subbasins in Fig. 1 have several gauge stations. Âż o It would help if these gauge stations could be displayed, but I assume it might not be feasible given the coarse resolution of this figure.

AR: they are, with the blue or red circles.

âĂć P.10, lines 1-9: the equifinality problem still exists for the global calibration, please discuss?

AR: the following sentence was added to emphasize this fact: " Despite global calibration may not be exempt of equifinality, the attention paid to the parameter ranges used (Table 3) allows to be confident in the physical relevance of the final parameter values."

âĂć P.10, line 10: What does a unique implementation mean? It is not clear. I assume GR4J was first calibrated on each gauged sub-basin, then the global calibration took place and a single parameter set was found. Please define unique.

AR: it is already defined at page 9, line 22, as the synonym of "single". It cannot be clearer. One model for one area, either the total gauged area or the whole area, including ungauged portions. However, it was specified at p.10, line 10, to "(see above for the justification)" of this methodology, because as explained above, the unique model leads to similar performances than those obtained with several local models.

âĂć P.10, line 16: What do you mean here: Âń . . .performances obtained with local Gr4J calibrations (Gaborit et al., in Press) were used when needed. . .Âż. o Do you mean that for those sub-basins not modelled by GEM-Hydro-UH, the performances of GR4J were substituted in the computation of the objective function (Eq. 2)? AR: Yes, and this sentence was modified as follows to be clearer: " However, as GEM-Hydro-UH was not locally calibrated for all of the 14 GRIP-O sub-basins (only those of Fig. 3 because of the computation cost), performances obtained with local GR4J calibrations (Gaborit et al., 2016 a) were used for the remaining ones to set the reference performance to be used in Eq. (2), justifying the use of that model in this study."

o The hypothesis behind this approach must be clearly stated in the paper; that is it is assumed that GEM-Hydro-UH would have a similar performance, am I right?

AR: the hypothesis is more that the local GR4J performance consists in the maximum performance which can be reached. Although mentioned under a more temperate statement, this was clarified in the Paper:

"This substitution does make sense considering that firstly, GR4J and GEM-Hydro-UH local performances are similar (Fig. 3), secondly that GR4J local performances were always very satisfactory (see Gaborit et al., 2016 a), and thirdly that the objective function still makes sense if global performance is higher than the local one (see above)." It seemed important to the authors' eyes to emphasize the last point of the above sentence by added, just above it, the following one: "It [the objective function] does rely on the hypothesis that global performance cannot be higher than local performance, but even if it was the case, this objective function would still make sense and the gain achieved with global over local performance would simply compensate for errors obtained on other catchments, possibly allowing to reach a perfect objective function value even with catchments having poorer performances with global calibration than with local calibration."

âĂć P.10, lines 15-17: Âń However, as GEM-Hydro-UH was not locally calibrated for all of the 14 GRIP-O subbasins, performances obtained with local GR4J calibrations (Gaborit et al., in Press) were used when needed (justifying the use of that model in this study). Âż o How was this done? Please provide a quick summary so the reader doesn't have to access the reference.

AR: given the clarifications brought in the previous answer to the reviewer, it is believed this point is clarified and does not need further explanation.

âĂć P.10, line 21: it is not arbitrary if it is based on prior work!

AR: this was corrected.

âĂć P.10, line 32 & P.11, lines 1 & 12: Watroute should be written with capital letters (WATROUTE).

AR: done

âĂć P.11, line 23: replace Âń . . .which. . .Âż by Âń . . .whose. . .Âż

AR: done

âĂć P.13, line 8: Please be consistent and replace watershed by basin.

AR: done

âĂć P.15, lines 20-21: Âń However, as a limited number of subbasins were used for the inter-comparison due to computational time limitations, no general model ranking can be derived from this study. Âż. o This means perhaps this paper is premature. Or as mentioned in the general comment section. Model intercomparison should be considered as supplemental information.

AR: see answer to the general comments

P.16, lines 4-5: I still do not get it, perhaps WATROUTE needs to be calibrated separately otherwise why calibrating with the UH? It is only valid to use WATROUTE if it can reproduce the UH at the chosen outlets used for the UH calibration. Unless there is a philosophical point I am not getting, which is perhaps possible, but doubtful. Please make a strong rebuttal to this statement.

AR: as clarified when answering the comment on P.5, lines 5 through 34 by giving more details on the computational time requirements, replacing WATROUTE with a simple UH was mandatory to proceed with the automatic calibrations; it is true that when the LSS is calibrated, we then could calibrate WATROUTE parameters using the calibrated SVS parameters in order to truly get the maximum possible performances. This was done in the GRIP-O report cited in the paper: Gaborit et al., 2016b. However, this requires a long calibration time for only 4 parameters. Moreover, experience has shown that a manual adjustment of these 4 parameter values allows to achieve simulation performances which are close to those obtained after automatic calibration. (The two LZS parameters are adjusted maximizing the recession and low-flow period simulations while the two Manning coefficients control the magnitude of peak flow events). So this manual work is quite straightforward with some training, and what is sure is that if the UH can do it, WATROUTE can do it. So the target when tuning WATROUTE parameters may even be the UH simulations themselves rather than the observations, because the calibrated UH simulations consist in the optimal simulations for a given methodology.

This was emphasized by slightly extending the following sentence (page 16 line 4): " The routing part of GEM-Hydro can be run afterwards, potentially adjusting the standard Manning values if needed (which can be done manually with a few runs)"

âĂć P.16, line 17: please replace No. XXXX

AR: this will be done right before the final publication.

âĂć P. 17- 20: In the References section, there are several references with Âń . . ., Âż, please fill them in.

AR: done

Figures âĂć Figure1 o The word Âń areas Âż should be replaced by sub-bassins, drainage areas or basins. o In the figure caption, replace sub-catchment by sub-basin, please be consistent.

AR: done for sub-catchment, but the term "areas" does not designate any subbasin but encompasses parts of several subbasins and therefore needs to remain different from the term subbasin.

âĂć Figure 2 o Moira river (CA) should be replaced by Moira River (CAN). CA usually stands for California. o Please remind the reader that the Moira River basin is subbasin 11.

AR: done

âĂć Figure 3 o Correct me if I am mistaken, but shouldn't the caption be as follows: Uncalibrated GEM-Hydro and GEM Hydro-UH performances. . .

AR: No, results only show GEM-Hydro-UH performances, before and after calibration.

o Wouldn’t be interesting to discuss the differences, at least for one or two sub-basins?

AR: It is not believed that more information is required for Figure 3 than its current description (page 11, line 24).

âĂć Figure 4 o Replace sub-catchment by sub-basin, please be consistent.

AR: done

âĂć Figure 5 o Please use upper case letters for Mesh (MESH), Watflood (WATFLOOD), please be consistent.

AR: done

âĂć Figure 6 o Why are not there any local calibration for sub-basins 13, 14 and 15 AR: because of computation time; the sentence close to p.10, l.15 was modified into: " However, as GEM-Hydro-UH was not locally calibrated for all of the 14 GRIP-O sub-basins (only those of Fig. 3 because of the computation cost)"

o What are the default parameter values when compared to those resulting from the calibration procedure, local and global? Wouldn’t be interesting to discuss the differences, at least for one or two sub-basins? AR: Table 7 gives some info. about parameter values; for the sake of brevity it is not believed that a more precise comparison between parameter values obtained with local and global calibrations is needed for a particular catchment, the main conclusion being that there is a strong variability in parameter values obtained with local calibration, which is shown in Table 7. However, the following sentence was added close to p.13, l. 11: " Moreover, it was noticed (not shown here) that parameter values were very different between local and global calibration procedures, even for catchments displaying very similar performances between the two strategies (such as subbasins 3, 5 and 8, see Fig. 6)."

o Please add the following precisions to the figure caption (at least that is my assump-

Printer-friendly version

Discussion paper

tion): Results are presented as NSE $\surd$ (left) and PBIAS (right), for many GRIP-O sub-basins.

AR: done

âĂć Figure 8 o Cumulative monthly NBS cannot by definition flow rate units, the units here should be cubic meters.

AR: done

o What are the numbers 7, 11 and 3 on the x-axis? Sub-basins number? I assume so as there are 14 tick marks between the occurrences of the number 7. Please provide this information in the figure caption.

AR: they are months. This was specified. 6 Tables âĂć Table 4 o The range for some parameter values defies the imagination, or any explanations? AR: No, the calibration was performed outside of ECCC.

âĂć Table 6 o To avoid any confusion please substitute CA for CAN, which is more often used - otherwise CA usually refers to California

AR: done

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-508/hess-2016-508-AC1-supplement.zip

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-508, 2016.