SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Reply to reviewers of paper hess-2019-334

On the representation of water reservoir storage and operations in large-scale hydrological models: implications on model parameterization and climate change impact assessments

Thanh Duc Dang, AFM Kamal Chowdhury, Stefano Galelli

Resilient Water Systems Group
Pillar of Engineering Systems and Design
SUTD, 8 Somapah Road
Singapore 487372
T. +65 6303 6600
http://people.sutd.edu.sg/~stefano_galelli/

Editor

Dear authors,

your contribution seems to fit well into the scope of HESS and the topic is timely and well presented. The two reviewers have suggested some improvements and I strongly encourage you to follow those suggestions carefully and try to improve your work accordingly. Otherwise, please explain why you disagree with some suggestion of the reviewers:

We would like to thank the Editor for the positive response and opportunity to revise our work. Following the reviewers' suggestions, we:

- *Included a sensitivity analysis of the model (VIC) output with respect to the six soil parameters used in the calibration process. This analysis (described in the Supplement) helped us strengthen the results regarding the parameterization of models with / without reservoirs (Section 4.1);*
- *Used data from an upstream gauging station to (1) validate the selected models and (2) further study the impact of a flawed parameterization on climate change impact assessment. These new results are included in the revised version of the manuscript (Figures 9 and 11);*
- *Clarified a few minor aspects in Sections 2, 3, and 4.*

Finally, please note that in our reply-to-reviewers line numbers correspond to the marked-up version of the manuscript.

I have an additional request: You state that you can calibrate your model satisfactorily, both with and without reservoirs. Of course, this is somehow surprising. For instance, I do not think that one can model the Yangtze River downstream the 3 Gorges Dam without explicitly taking care of this (huge) reservoir. Thus, your findings might be dependent on the size and total storage volume of the reservoirs. Please elaborate on that. In addition, your model-parameters with considering reservoirs should result in more realistic soil parameters, I think. Can you also comment on that?

Thanks for raising these two points, on which we elaborated in Section 5. Please refer to line 29-34 (page 15) and 4-8 (page 16, marked-up manuscript).

Reply to reviewer #1

General Comments:

This paper presents a computational framework based on the Variable Infiltration Capacity (VIC) model and a Multi-Objective Evolutionary Algorithm that enables to analyze the effects of water reservoir representation on the parameterization of hydrological models. The modelling approach was applied to the upper Mekong river basin, upstream the Chiang Saen gauging station, considering two configuration schemes with and without water reservoirs. The authors exposed the theme in a clear, logical sequence, which resulted in a well-written comprehensive text. In my opinion, this work should be accepted for publication at the Hydrology and Earth System Sciences (HESS) journal after minor revisions. Attached you find some comments/suggestions.

We thank the reviewer for the positive comments and practical suggestions for improving the manuscript.

Introduction: A comprehensive literature review was done concerning hydrological modelling applied to large river basins, highlighting the limited number of model approaches that enables the direct representation of reservoir water storage with target operating rules to achieve pre-defined objectives. Furthermore, the scientific contribution of this manuscript is clearly defined from line 21 to 34 (Page 03).

Thanks for the positive feedback.

Study area: The study area description looks poor. Please include more information concerning soil properties and classes, land use, geology and geomorphology, just to enable a better comprehension of main driving forces related to hydrological processes in the Upper Mekong river basin.

We included more details regarding soil properties, land use, geology and geo-morphology in Section 2. Please refer to line 32-33 (page 4) and 1-3 (page 5) of the marked-up manuscript.

In Section 3.2.1 you present the way the input variables (Land use and land cover data, soil data, and flow direction) are obtained/processed, but a discussion about those data is lacking.

We included a discussion about these data in the second paragraph of Section 3.2.1. Please refer to line 1-5 (page 8) of the marked-up manuscript.

Materials and methods: The way used to achieve the objectives of the manuscript is clearly presented and well-organized. Nevertheless, I would suggest including a sensitivity analysis of the model results to the parameters controlling the rainfall-runoff process (Ds, Dmax, Ws, b, d1 and d2) in the Variable Infiltration Capacity model.

Thanks for the suggestion. We indeed explored the sensitivity of the model output with respect to the values of these six parameters (for both model configurations). Specifically, we linked VIC (with and without reservoirs) with the SAFE toolbox (Pianosi et al., 2015), which provides a number of global sensitivity algorithms—among which we chose the extended Fourier Amplitude Sensitivity Test, or eFAST.

The results of the sensitivity analysis show that all parameters are indeed important with respect to the model output. The only exception is the depth of the second soil layer (d2), which does not appear to depend on the absence (or presence) of water reservoirs (see the description of Figure 6 provided in Section 4.1). Since the description of the sensitivity analysis is a bit unwieldy, we preferred to keep it in the Supplement, to which we refer in Section 4.1.

Results: The questions raised in the introduction section were answered. The results showed that a flawed model parametrization by disregarding anthropogenic interventions (such as hydraulic infrastructure) led to the overestimation of baseflow and runoff during the dry season to compensate water release related to hydropower production. Despite this, I would suggest discussing more deeply the model parameterization and results. According to the authors, the VIC model has been previously used by other researchers. Could your results be compared with them? Or other models with similar purpose.

Literature offers a few studies that implemented VIC model to the Mekong River basin (e.g., Zhong et al. (2019), Chang et al. (2019)). Unfortunately, these studies do not report details about the parameterization—they tend to focus on the modelling accuracy. The same comment applies to studies that applied other hydrological models to the Mekong (e.g., Lauri et al. (2012), Rasanen et al. (2012), Hoang et al. (2016)). For this reason, we could not carry out a direct comparison between our parameterization and the one obtained by previous studies.

Specific Comments:

1) Page 04, line 07: How large is the Upper Mekong river basin? Please include the catchment area in km2.

The Upper Mekong river basin has an area of about 167,400 km^2; that's about 24% of the Mekong's basin area. We included this information in Section 2 (Study area).

2) Page 04, lines 07-08: The information of elevation ranging from 362 m to 6,494 m should be included.

We added this information.

3) Page 04, line 30: Please include a complementary information explaining why the hydrological alterations became more evident since 1992. The largest reservoir (Xi'er He 1) was built before in 1989.

Previous works focusing on the hydrological alterations caused by hydropower dams in the Lancang basin generally consider two periods, namely pre and post 1991 (e.g.,

1960-1991 and 1992-2013) (Cochrane et al. (2014), Lu et al. (2014), Dang et al. (2016)). The reason for this choice is due to Manwan dam, which received significant public attention because of its location—it is indeed the first dam built on the main stem of the river. We understand that our sentence is potentially misleading, so we clarified this point in the revised version of the manuscript (please refer to line 21-23, page 4, marked-up manuscript).

4) Page 05, line 19 and Table 1: How the feasible ranges of the two soil layers thickness (d1 and d2) were defined?

The definition of the feasible ranges is based on the ranges that are adopted in other studies applying VIC model to large catchments (e.g., Dan et al. (2012), Park and Markus (2014), Xue et al. (2015)). Note that the same ranges are also recommended by Wi et al. (2017), who developed a general-purpose user-friendly software package for VIC. We clarified this point in the caption of Table 1.

5) Page 06, line 29: How the target water level is defined for each reservoir?

The target water level is defined as follows. First, we determine the minimum and maximum water levels that a reservoir should reach within a year. In our case, we use the minimum and maximum elevation levels, which are given in the design specifications of each reservoir. Then, we set the time at which the minimum and maximum water levels should be reached. In our case, we use the months of May and November, which correspond to the beginning and end of the wet season. Finally, we connect these points with a piecewise linear function that gives us the daily target level for each calendar day. We further clarified this point in the second paragraph of Section 3.1.2.

6) Page 08 lines 02-17: What processes are considered in the reservoir water balance and operation? Are infiltration loss and groundwater inflow disregarded?

We considered three main processes, namely inflow, release, and evaporation. As noted by the reviewer, there are two other processes that could be considered, that is, infiltration and seepage (via dam body, abutment, and foundation). The reason for which we disregarded them is twofold. First, the Upper Mekong basin is a mountainous region, with mostly rocks and a shallow Quaternary alluvium (Gupta, 2009; Carling, 2009), so the infiltration losses are to some extent marginal as compared to inflow, release, and evaporation. Second, the dams considered in our study are built with concrete (and with rocky abutments and foundations), so seepage is indeed limited. We touched upon this point in Section 3.2.2 (last paragraph).

7) Page 09, line 11: "parellelized" should be changed to "parallelized".

Thanks for spotting this typo.

8) Page 11, lines 05-24: Please discuss the pattern observed in Figure 6, with different ranges of the model parameters for the simulation with and without reservoirs. Some parameters presented a more spread pattern for the scenario with reservoirs and a more uniform one for the scenario without reservoir and vice versa.

Thanks for the suggestion. We discuss about this pattern in both Section 4.1 (line 7-9, page 12, marked-up manuscript) and Supplement.

9) Fig. 2: Please use the term "modelling approach" instead "model" (Figure legend).

We think that the term "model" is more appropriate, since the figure depicts two specific numerical models (i.e., VIC's rainfall-runoff and routing modules) and an optimization algorithm (i.e., epsilon-NSGA-II)—and not a family of similar models that could be denoted with the term "approach".

10) Fig. 3: Please include a legend describing the four levels highlighted in Figure 3c.

Modified as suggested

11) Fig 8: I would suggest to reduce the number of baseflow intervals to enable a better visualization of model results related to the configuration scheme with and without reservoirs.

Thanks for the suggestion. We indeed modified the number of baseflow and runoff intervals.

Reply to reviewer #2

General Comments:

This work contributes to literature on the representation of human interactions in hydrological model by demonstrating the importance of modeling reservoirs. Using the high-profile case of the Lancang / Mekong river basin, they calibrate a large-scale hydrological model (VIC), both with and without reservoirs. They show that while there exist parameter sets for which the model can be calibrated in both cases, the without- reservoir model compensate the absence of human-operated storage by artificially creating soil storage. Then they go on to investigate the reactions of both models to climate change.

This is an interesting, timely and well-written piece of work that fits well within the scope of HESS. My comments are mainly at how the results are handled and interpreted.

We thank the reviewer for the positive feedback.

1) While authors explain very well the consequences of the no-reservoir calibration in terms of structural model behavior, they do not show what that means for water resources appraisals. Indeed, both models are calibrated to have the same behavior downstream of 1 gauge. The crucial difference is that the model including reservoirs can be 1) validated at other (upstream) sites, and 2) validated with post 2005 dams (or the results can be extrapolated with new dams or new operating rules). This point is not really demonstrated.

Thanks for this suggestion. We totally agree that we can strengthen our message by implementing one of these two validations. In particular, we believe that the first option is the most feasible, although it presents some minor challenges on which we elaborate below. (Our thoughts regarding the second option are outlined as a response to comment no. 4.)

Discharge data in the Chinese section of the Mekong basin are not available to the international research community. More specifically, only water level data for the flood season (June-October) at Yunjinghong station (located downstream of Jinghong reservoir; see Figure 1 in the manuscript) are available through the Mekong River Commission (http://www.mrcmekong.org/news-and-events/news/mrc-and-china-renew-pact-on-water-data-provision-and-other-cooperation-initiatives/). Therefore, we opted for retrieving data from three previous studies (He et al. (2009), Wang et al. (2018), and Tang et al. (2019)) that published monthly discharge values at Jiuzhou station (located near Gonguoqiao reservoir; see Figure 1) for the period 1996-2005. When retrieving (digitalizing) the data, we found a maximum discrepancy between the three time series of about 10 m^3/s; a negligible value if we consider that the minimum discharge is roughly 400 m^3/s.

With this 10-year time series, we carried out a thorough validation exercise, whose results are illustrated in the figure below. We can note that the model calibrated without reservoirs largely overestimates the dry season flow and slightly

underestimates the wet season flow. This point is further discussed in Section 4.1 (line 27-31, page 12, marked-up manuscript).

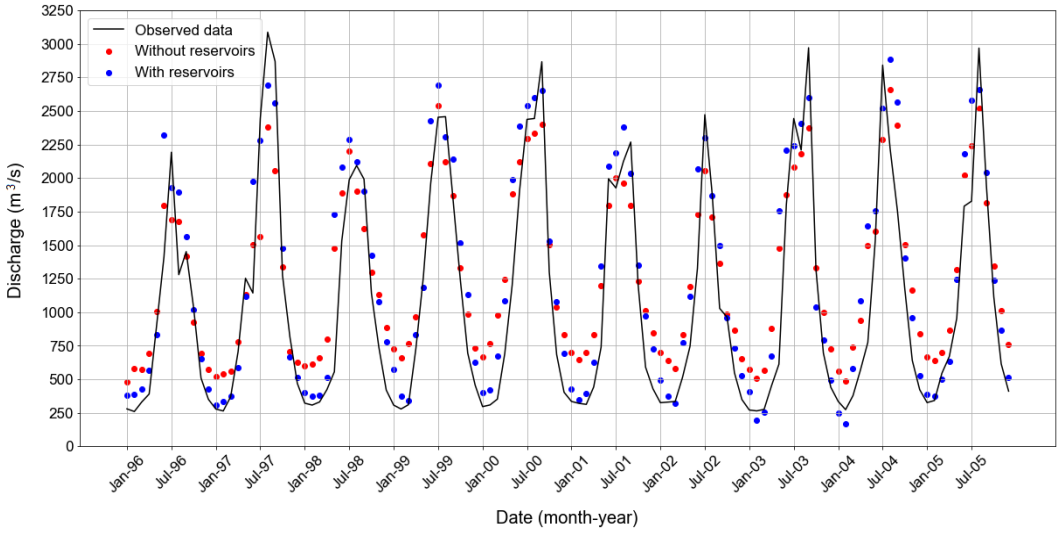

Figure 1. Comparison between observed and simulated monthly discharges at Jiuzhou station over the period 1996–2005. Simulated data are produced by the two selected models with and without reservoirs (blue and red dots, respectively).

2) Likewise, I disagree that the results from climate models are very different. Yes, there are some small difference, and authors do a good job of explaining them, but these differences are arguably small compared with the uncertainty surrounding downscaled climate projections. In fact, it is remarkable how robust the no-reservoir model results are with climate change.

We agree with the reviewer that the uncertainty surrounding the climate change projections is larger than the difference (in discharge) due to the representation of water reservoirs in VIC. This said, we also note that such difference is non-negligible and consistent across both RCPs. We thus proceeded by slightly toning down this result in the revised version of the manuscript (please refer to line 28-34, page 13, marked-up manuscript).

3) I would advise authors to try and show how both calibrations differ at upstream gauges (where only the with-reservoir calibration would yield sensible results) where there is data, then carry on that comparison with climate change.

Yes, we included an additional validation based on data at an upstream gauge, as explained in our reply to comment no. 1. We also included the same comparison under the climate change scenarios (please refer to line 16-22, page 14, marked-up manuscript).

4) Alternatively (but this is probably more work), authors can focus on post-2005 years with the addition of new dams and look at the differences between both model at the outlet. And again look at how the two models differ with climate change then.

The validation based on discharge data collected in the upstream reaches of the basin is indeed simpler to implement, so we preferred to adopt such option. This said, we

understand the potential of this second suggested validation, so we did elaborate on this point in the concluding remarks.

5) Figure 8 is not very clear and could be replaced with curves comparing average monthly basinwide runoff for the two calibrations (and same for baseflow). This would show explicitly how the no-reservoir calibration compensates for the reservoirs. Alternatively, authors could map the dry / wet season cell-by-cell differences in baseflow / runoff between the two calibrations.

Yes, we agree that Figure 8 is not that clear. The option of replacing it with curves comparing average monthly basin-wide runoff for the two calibrations is sound. Yet, such figure would not represent explicitly the spatial variability or runoff and baseflow; something that we deem important. We thus preferred to keep the current 'skeleton' of Figure 8, which we improved by modifying the number of baseflow intervals (as suggested by reviewer #1).

6) Figures 9 and 10 don't bring much and could be relegated to supplementary material to make place for new figures that can show the consequences of omitting reservoirs, e.g., as suggested in 3) and 4)

Yes, we moved both Figure 9 and 10 to the Supplement.

Detailed Comments:

Figure 1.b: the scale, useful in 1.a, is missing.

Modified as suggested.

Section 3.1.2: a short description of the version 4.2 of VIC's routing module would be helpful here. Recall that this research is useful for all large-scale hydrological models, not just for VIC experts.

We provided a brief description of VIC's routing module. However, we thought it's better to include such description in Section 3.1.1, which focuses entirely on VIC (section 3.1.2 is about the modification we introduced). Please refer to line 21-25 (page 5, marked-up manuscript).

Page 8, lines 2-9: there is a recent publication presenting a database of reservoir storage-area-depth relationships in Yigzaw et al (2018) there https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017WR022040 could that affect your results in significant ways?

Thanks for pointing us to this paper. We noticed that only Manwan reservoir is in this new database, so we believe this cannot impact our model implementation / results. This said, we believe that the adoption of Liebe's method should provide robust results—the method has been tested and adopted by several state-of-the-art studies in large-scale hydrology.

Page 9, lines 8-12: how many parallel processors are running? (l. 12) by runtime, do you mean the wall clock time or the total computational time used by all the processors (i.e. wallclock time times # of processors)? How do we know that 20 seeds are enough? And that the algorithm has converged after 250 function evaluations? Finally, what is the runtime for one run of VIC for that basin?

For each (of the 20) seed(s), we used four cores. As for the runtime, we reported the wall-clock time per core. The runtime of VIC for the basin is about 50 mins (on one core). We clarified these details in the revised version of the manuscript (line 29-30, page 9, marked-up manuscript).

There are no specific guidelines in the literature on the number of seeds that should be used (this also depends on the specific MOEA that one adopts). In general, the idea is to use at least 10 seeds, so as to contribute diversity to a given MOEA's search results. Naturally, the number of seeds that one can use depends on the computational requirements of the simulation model as well as the available computational power.

To measure the convergence of the optimization algorithm, we used the hypervolume indicator, which captures both convergence and diversity of a Pareto front (see Reed et al., (2013)). Results indicated that the algorithm reached convergence after ~150 iterations.

Page 14, lines 14-17: I disagree this is a real limitation, since proposing a universal rule system for reservoir operations is not an aim of this paper. Instead, a consequence of your work is that hydrological model calibration with reservoirs and a bespoke release rule means that the model still captures key hydrological processes once the release rules change (because there are new reservoirs / because reservoirs' purposes evolve). In contrast, if reservoirs are not represented to begin with, once the number and / or operations of reservoirs evolves, the model's hydrological parameters have to be recalibrated every time.

Thanks for this comment. We agree that the sentence was structured in an awkward—and potentially misleading—way, so we removed it. We also elaborated on the reviewer's comment in the latter part of the Conclusions (line 4-12, page 16, marked-up manuscript).

References

Carling, P. A. (2009). The geology of the lower Mekong River. In The Mekong (pp. 13-28). Academic Press.

Chang, C.H., Lee, H., Hossain, F., Basnayake, S., Jayasinghe, S., Chishtie, F., Saah, D., Yu, H., Sothea, K., Du Bui, D. (2019). A model-aided satellite-altimetry-based flood forecasting system for the Mekong River. Environmental Modelling & Software, 1:112-27.

Cochrane, T., Arias, M., Piman, T. (2014). Historical impact of water infrastructure on water levels of the Mekong River and the Tonle Sap system, Hydrology and Earth System Sciences, 18, 4529–4541.

Dan, L., Ji, J., Xie, Z., Chen, F., Wen, G., Richey, J. E. (2012). Hydrological projections of climate change scenarios over the 3H region of China: A VIC model assessment, Journal of Geophysical Research: Atmospheres, 117.

Dang, T. D., Cochrane, T. A., Arias, M. E., Van, P. D. T., de Vries, T. T. (2016). Hydrological alterations from water infrastructure development in the Mekong floodplains, Hydrological Processes, 30, 3824–3838.

Gupta, A. (2009). Geology and landforms of the Mekong Basin. In The Mekong (pp. 29-51). Academic Press.

He, D., Lu, Y., Li, Z., Li, S. (2009). Watercourse environmental change in Upper Mekong. In The Mekong (pp. 335-362). Academic Press.

Hoang, L.P., Lauri, H., Kummu, M., Koponen, J., van Vliet M., Supit, I., Leemans, R., Kabat, P., Ludwig, F. (2016). Mekong River flow and hydrological extremes under climate change. Hydrology and Earth System Sciences, 20:3027-41.

Lauri, H., Moel, H. D., Ward, P. J., Räsänen, T. A., Keskinen, M., Kummu, M. S. (2012). Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge. Hydrology and Earth System Sciences, 16: 4603-4619.

Lu, X., Li, S., Kummu, M., Padawangi, R., Wang, J. 2014. Observed changes in the water flow at Chiang Saen in the lower Mekong: Impacts of Chinese dams?, Quaternary International, 336, 145–157.

Park, D., Markus, M. (2014) Analysis of a changing hydrologic flood regime using the Variable Infiltration Capacity model, Journal of Hydrology, 515, 267–280.

Pianosi, F., Sarrazin, F., Wagener, T. (2015). A Matlab toolbox for global sensitivity analysis. Environmental Modelling & Software, 70, 80-85.

Räsänen, T. A., Koponen, J., Lauri, H., Kummu, M. (2012). Downstream hydrological impacts of hydropower development in the Upper Mekong Basin. Water Resources Management, 26(12), 3495-3513.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in Water Resources, 51, 438-456.

Tang, X., Zhang, J., Wang, G., Yang, Q., Yang, Y., Guan, T., Liu, C., Jin, J., Liu, Y., Bao, Z. (2019). Evaluating Suitability of Multiple Precipitation Products for the Lancang River Basin. Chinese Geographical Science, 29(1), 37-57.

Wang, Z., Chen, J., Lai, C., Zhong, R., Chen, X., Yu, H. (2018). Hydrologic assessment of the TMPA 3B42-V7 product in a typical alpine and gorge region: the Lancang River basin, China. Hydrology Research, 49(6), 2002-2015.

Wi, S., Ray, P., Demaria, E. M., Steinschneider, S., Brown, C. (2017). A user-friendly software package for VIC hydrologic model development. Environmental Modelling & Software, 98, 35-53.

Xue, X., Zhang, K., Hong, Y., Gourley, J. J., Kellogg, W., McPherson, R. A., Wan, Z., Austin, B. N. (2015). New multisite cascading calibration approach for hydrological models: Case study in the red river basin using the VIC model, Journal of Hydrologic Engineering, 21, 05015 019.

Zhong, R., Zhao, T., He, Y., Chen, X. (2019). Hydropower change of the water tower of Asia in 21st century: A case of the Lancang River hydropower base, upper Mekong. Energy, 179, 685-696.

On the representation of water reservoir storage and operations in large-scale hydrological models: implications on model parameterization and climate change impact assessments

Thanh Duc Dang[1], AFM Kamal Chowdhury[1], and Stefano Galelli[1]

[1]Pillar of Engineering Systems and Design, Singapore University of Technology and Design, Singapore 487372

Correspondence: Stefano Galelli (stefano_galelli@sutd.edu.sg)

Abstract. During the past decades, the increased impact of anthropogenic interventions on river basins has prompted hydrologists to develop various approaches for representing human-water interactions in large-scale hydrological and land surface models. The simulation of water reservoir storage and operations has received particular attention, owing to the ubiquitous presence of dams. Yet, little is known about (1) the effect of the representation of water reservoirs on the parameterization of hydrological models, and, therefore, (2) the risks associated to potential flaws in the calibration process. To fill in this gap, we contribute a computational framework based on the Variable Infiltration Capacity (VIC) model and a Multi-Objective Evolutionary Algorithm, which we use to calibrate VIC's parameters. An important feature of our framework is a novel variant of VIC's routing module that allows us to simulate the storage dynamics of water reservoirs. Using the upper Mekong river basin as a case study, we calibrate two instances of VIC—with and without reservoirs. We show that both model instances have the same accuracy in reproducing daily discharges (over the period 1996–2005); a result attained by the model without reservoirs by adopting a parameterization that compensates for the absence of these infrastructures. The first implication of this flawed parameter estimation stands in a poor representation of key hydrological processes, such as surface runoff, infiltration, and baseflow. To further demonstrate the risks associated to the use of such model, we carry out a climate change impact assessment (for the period 2050–2060), for which we use precipitation and temperature data retrieved from five Global Circulation Models (GCMs) and two Representative Concentration Pathways (RCPs 4.5 and 8.5). Results show that the two model instances (with and without reservoirs) provide different projections of the minimum, maximum, and average monthly discharges. These results are consistent across both RCPs. Overall, our study reinforces the message about the correct representation of human-water interactions in large-scale hydrological models.

1 Introduction

Hydrological systems consist of multiple physical, chemical, and biological processes, most of which are profoundly altered by anthropogenic interventions (Nazemi and Wheater, 2015a, b). Land cover modifications or hydraulic infrastructures, for

instance, affect both surface and sub-surface hydrological processes by redistributing water over time and space (Haddeland et al., 2006; Bierkens, 2015). Such alterations are expected to amplify in the near future, owing to the increase in water and energy consumption (Abbaspour et al., 2015). In this context, hydrological models play a key role, as they help plan the use of water resources in a sustainable way, so as to avoid adverse impacts on ecosystems and livelihoods (Bunn and Arthington, 2002; Yassin et al., 2019). A detailed and accurate representation of the anthropogenic interventions within hydrologic models is thus of paramount importance: successful water management plans must necessarily build on reliable models.

Water reservoirs are arguably one of the most common infrastructures altering hydrological processes at the catchment scale; yet, their representation in hydrological and land surface models is challenged by multiple factors. First, the vast majority of the models currently available was initially conceived to study and understand the behaviour of natural systems, so the added representation of water reservoirs entails the partial modification of the model structure. Second, the existing databases (e.g., GRanD; Lehner et al. (2011)) provide details on dam design specifications, but no information on the management aspects, such as the operating rules or flood contingency plans. Third, the installation of dams is generally combined with impoundment (or filling) strategies, which may largely differ from the steady-state operating rules and last from a few months to several years (Gao et al., 2010; Zhang et al., 2016). Although the complexity of these factors varies with the study site at hand, one might imagine that the representation of water reservoir storage and operations is particularly challenging for large-scale models, simply because of the number of dams deployed over time in large river basins. It is perhaps not surprising to observe that water reservoirs—and their corresponding operations—have not been consistently accounted for across the broad number of large-scale hydrological modelling studies available in literature.

A simple and popular approach stands in the exclusion of large impounds from the streamflow routing modules; a modelling choice that has been adopted in many regions across the globe (Maurer et al., 2002; Jayawardena and Mahanama, 2002; Akter and Babel, 2012; de Paiva et al., 2013; Leng et al., 2016). Such approach can support the investigation of various physical processes (e.g., emergence of new hydrological regimes, generation of land surface fluxes), but obviously prevents the application of the hydrological models to downstream water management problems, such as investigating the impact of regime shifts on hydropower production. Another potential issue with this approach stands in the model parameterization, which might be affected by a calibration process carried out with hydrological time series altered by anthropogenic interventions. de Paiva et al. (2013), for instance, implemented the MGB-IPH hydrologic/hydraulic model to the Amazon River basin—a region characterized by the presence of hydroelectric dams (Finer and Jenkins, 2012)—and yet obtained reliable calibration performance at multiple gauging stations. A similar example is represented by Abbaspour et al. (2015), who simulated hydrological and water quality processes for the entire European continent. Despite neglecting the presence of hydraulic infrastructures, the model yielded acceptable values for the goodness of fit statistics. One may thus wonder whether the calibration process somehow compensates for a deficiency in the model structure.

With the goal of striking a balance between an accurate representation of reservoirs and the 'costs' due to the modification of the model structure, several researchers have adopted an hybrid approach, in which the output of hydrologic/hydraulic models (e.g., runoff or streamflow at multiple locations) is post-processed with the aid of water management (or reservoir operation) models. The very first efforts employed data on water uses to correct the output of global models, such as WaterGAP (Alcamo

et al., 1997) or WBM (Vörösmarty et al., 1998). Using a similar concept, Hanasaki et al. (2006) accounted for 452 reservoirs in a global river routing model. More sophisticated post-processing techniques are based on optimization algorithms, which are used to design either reservoir operating rules or sequences of reservoir discharges that meet pre-defined objectives (e.g., hydropower production). Lauri et al. (2012) and Hoang et al. (2019), for example, first calibrated the distributed hydrological model VMod for the Mekong river basin, and then post-processed its output using a linear programming algorithm that designed the discharge time series for 126 dams over a given simulation scenario. Similarly, Turner et al. (2017) and Ng et al. (2017) examined the vulnerability of global hydropower production to climate changes and El Niño Southern Oscillation by correcting the discharge simulated by WaterGAP. In this case, the correction entailed designing bespoke reservoir operating rules through the use of a stochastic dynamic programming algorithm (Turner and Galelli, 2016). Other recent applications of post-processing techniques were adopted in Masaki et al. (2017); Veldkamp et al. (2018); Zhou et al. (2018).

Naturally, the most suitable approach stands in the direct representation of water storage and operations within a large-scale hydrological model (Bellin et al., 2016). This approach requires not only to modify the model structure (or to develop a new one), but also to gather information on the design specifications and operating rules of the water reservoirs. Because of these challenges, the number of large-scale hydrological modelling studies adopting such approach is limited. A first attempt was carried out by Pokhrel et al. (2012), who incorporated a water regulation module into the MATSIRO model to reproduce the dynamics of heavily regulated global river basins. More recently, Shin et al. (2019) integrated a reservoir storage dynamics and release scheme into the continental hydrological model LEAF-Hydro-Flood to simulate $\sim$1,900 reservoirs within the contiguous United States. In both studies, the authors gave particular emphasis to the calibration of the reservoir operating scheme, and demonstrated that the hydrological model accurately represents some processes altered by human interventions, such as the reservoir-floodplain inundation.

While the relevance and needs for the description of human-water interactions in hydrological models are now well acknowledged (Nazemi and Wheater, 2015a), less is know about the risks associated to a poor representation of such interactions. For example, can the estimation of some hydrological parameters be flawed by an inaccurate representation of water reservoir storage? What are the implications for the downstream applications of a flawed model? To answer these questions, we take the upper Mekong river basin as a case study, for which we develop a computational framework based on the Variable Infiltration Capacity (VIC) model (Liang et al., 1994) and a Multi-Objective Evolutionary Algorithm (MOEA) tasked with the problem of calibrating the model. A key feature of the framework is a novel variant of VIC—named VIC-Res—that allows us to represent the reservoir storage dynamics and operating rules within the streamflow routing module. In a first experiment, we use this framework to calibrate two instances of VIC—with and without reservoirs. As we shall see, both model instances attain the same accuracy; a result obtained by the model instance without reservoirs by adopting a parameterization that compensates for the absence of these infrastructures. In turn, this leads to a poor representation of key hydrological processes, such as infiltration or baseflow. In our second experiment, we demonstrate the potential implications of these unintended consequences by applying two selected model instances (with and without reservoirs) to a climate change impact assessment, for which we obtain partially-diverging expectations on the hydrological alterations caused by global warming.

In the remainder of the manuscript, we first describe the study area (Section 2) and then proceed by illustrating the computational framework (Section 3), including the data on dams and operating rules. In Section 4, we provide a detailed description of the results obtained for the aforementioned experiments, whose implications are further discussed in Section 5.

2 Study area

The Mekong is a trans-boundary river that flows through China, Myanmar, Thailand, and Laos before pouring into one of the world's largest delta located in Cambodia and Vietnam. The catchment area of about 795,000 km^2 can be divided into two parts, namely the upper Mekong, or Lancang, and the lower Mekong basins (Figure 1a). The upper Mekong stretches in a North-to-South direction and drains an area of 167,400 km^2. ~~, and~~ As shown in Figure 1b, the region is characterized by a complex orography, with high mountains and deep valleys ~~(Figure 1b)~~ (the elevation ranges from 362 to 6,494 m). Because of these orographical conditions, the spatio-temporal variability of rainfall and temperature is remarkable. The average annual precipitation across the basin ranges from 752 to 1,025 mm, 70% of which is concentrated in the monsoon season (May to November). The precipitation in the Northwestern part of the basin is sometimes lower than 250 mm/year, making it dryer than the Southeastern part, which receives an average of 1,600 mm/year (Han et al., 2019). The average annual temperature across the basin varies narrowly (from 12.3 to 14.3 °C), but the latitudinal temporal gradient is much larger—about 2.2 °C/100 km (Wang et al., 2014). Climate changes are expected to modify both rainfall and temperature patterns, making the region warmer, wetter, and more susceptible to extreme weather events (Tang et al., 2015).

The favourable orography and abundant water availability have attracted massive investments in the hydropower sector (see the location of the dams in Figure 1b), with consequent impacts on the riverine ecosystems (Lauri et al., 2012; Dang et al., 2018; Hoang et al., 2019). The impact of these dams goes beyond the upper Mekong basin (Zhao et al., 2012; Han et al., 2019): the analysis of historical data shows that dams have already modified many indicators of hydrological alterations in the entire basin, including the Cambodian lowlands and river delta (Hecht et al., 2018). These alterations appear to be more evident since ~~1992~~ the early 1990s, when Xi'er He 1 and ~~the~~ Manwan dams started storing water (Cochrane et al., 2014; Lu et al., 2014; Dang et al., 2016). Overall, the upper Mekong basin offers two desirable features for investigating the effect of water reservoir storage and operations on the parameterization of hydrological models. First, the catchment is heavily regulated (Hecht et al., 2018). Second, the catchment area is about 24% of the whole Mekong River basin, so this helps reduce the computational requirements of the optimization-based calibration process. The location of the gauging station (Chiang Saen) used ~~in our work~~ for the calibration process is illustrated in Figure 1a. This station provides a long and reliable ~~streamflow~~ daily time series, which has been adopted by several studies on the Mekong basin (e.g., Lauri et al. (2012); Cochrane et al. (2014); Lauri et al. (2014); Hoang et al. (2016)). To validate the model, we use monthly discharge values at Jiuzhou station (see Figure 1a), retrieved from He et al. (2009); Wang et al. (2018); Tang et al. (2019). For both stations, we used data belonging to the period 1996–2005. ~~In this study, we use daily discharges measured in the period 1996–2005 for the model calibration.~~

The aforementioned orography and climate conditions are not particularly suitable for agricultural activities, which are indeed limited. The basin is mountainous, with mostly rocks and a shallow Quaternary alluvium (Carling, 2009; Gupta, 2009).

Due to the impermeability of bedrock underneath isolated valleys, only a very small fraction of water leaks into the ground through karst aquifer units (Lee et al., 2017). As a result, subsurface water is mostly generated in the shallow loam layer in the form of baseflow.

3 Materials and methods

The first goal of our study is to investigate the role of water reservoir storage and operations on the parameterization of large-scale hydrological models. To this purpose, we adopt the computational framework illustrated in Figure 2, which consists of VIC's rainfall-runoff and routing modules and the $\varepsilon-$NSGAII MOEA. In Section 3.1 we provide a detailed description of VIC's modules, including the proposed variant for representing reservoir storage dynamics. The data and experimental setup of the framework are outlined in Section 3.2 and 3.3. In Section 3.4, we describe the climate change data used for our second goal, that is, to demonstrate that different model parameterizations caused by the absence (presence) of water reservoirs can affect the results of a climate change impact assessment.

3.1 Hydrological-water resources management model

3.1.1 Variable Infiltration Capacity model

VIC is a large-scale, semi-distributed land hydrological model maintained and developed by the University of Washington (http://www.hydro.washington.edu). The model consists of two core components, namely a rainfall-runoff and routing module (Figure 2), which can be applied to multiple spatial scales and implemented with different temporal resolutions—daily, in our case. The rainfall-runoff module simulates the water and energy fluxes that govern the terrestrial hydrological cycle (Liang et al., 1994). To this purposes, it takes as input climate forcings (precipitation and temperature), land use and soil maps, Leaf Area Index and albedo, and a Digital Elevation Model (DEM). For each computational cell, the module uses one vegetation layer and two (or three) soil layers: the upper soil layer controls evaporation, infiltration, and runoff, while the lower layer controls the baseflow generation. These gridded variables are then used by the routing module (Lohmann et al., 1996, 1998), which ~~calculates the streamflow using the Unit Hydrograph approach and linearized de Saint-Venant equations.~~ simulates discharge throughout the river network using a linearized version of the Saint-Venant equations. Specifically, the module first creates the impulse response functions for each grid cell, and then simulates the flow convolution by aggregating the flow contribution from all upstream cells at each time step lagged according the response functions (ibid).

Following the approach adopted in previous works on the calibration of VIC (e.g., Dan et al. (2012); Park and Markus (2014); Xue et al. (2015)), we focus our attention on six main parameters that control the rainfall-runoff process (Table 1). These parameters are the thickness of the two soil layers (d_1 and d_2), the infiltration parameter (b), and three baseflow parameters (D_s, D_{max}, and W_s). The parameter b characterizes the shape of the Variable Infiltration Capacity curve, and therefore influences the available infiltration capacity and quantity of runoff generated by each cell (for additional details, please refer to Ren-Jun (1992) and Todini (1996)). A higher value of b leads to a lower infiltration rate and higher surface runoff. The three parameters

D_s, D_{max}, and W_s determine the shape of the Baseflow curve (Franchini and Pacciani, 1991), which relates the soil moisture in the lower layer to the amount of baseflow. More specifically, D_{max} is the maximum baseflow that can occur in the lower layer, while D_s is the fraction of D_{max} associated to the transition from linear to non-linear (rapidly increasing) baseflow generation. W_s is the fraction of the maximum soil moisture (in the lower layer) where non-linear baseflow occurs. Hence, higher values of W_s increase the water content needed for rapidly increasing baseflow. The thickness of the two soil layers affects several processes. In general, thicker layers delay the seasonal peak flow and increase the evaporation losses (since they increase the water storage capacity).

3.1.2 Water reservoir storage and operations

To represent the storage dynamics of water reservoirs, we modified VIC's routing module (version 4.2) using the following steps. First, we determine the location of all dams within the basin, and directly add them to the model using a dam cell (Figure 3a-b). To avoid allocating multiple dams within the same cell, we adopt a high-spatial resolution of 0.0625 degree (approximately 6.9 km). Then, we aggregate the reservoir storage in the dam cell, where we calculate the daily mass balance. ~~from which~~ From the dam cell, water is discharged using the rule curves described in the following paragraph. Since the construction of a dam is likely to create an impoundment with surface area larger than the dam cell, we proceed by estimating the maximum reservoir extent; an information used to determine the so-called reservoir cells, namely cells that are at least half-covered by water (see Figure 3b). Although these cells do not contain the reservoir storage, they can affect the evaporation processes, so their number and location must be determined accurately. The flow routing in these cells follows the information provided in the flow direction map (described in Section 3.2.1). We note that a more realistic way of representing a reservoir within a hydrological model is to spread the reservoir storage over multiple upstream cells from the dam location (Shin et al., 2019). Yet, a successful implementation of this method requires a detailed bathymetry of all reservoirs within the basin (an information that may not always be available) and a 2D model of the reservoir, so as to accurately calculate the water fluxes between the different reservoir cells.

As for the reservoir operations, we adopt an approach similar to that of Piman et al. (2012), which relies on rule curves conceived to maximize the hydropower production—an assumption justified by the fact that all dams within the upper Mekong are operated for hydropower supply (Räsänen et al., 2017). Determining the rule curve for a given reservoir means determining the daily target water levels. For the case of hydropower production in the Mekong basin, such rule should allow to (1) drawdown the reservoir storage during the drier months (e.g., December to May) to maximize the production of electricity, (2) recharge the depleted storage during the monsoon season, and (3) avoid the risks of spilling water at the end of the monsoon season (see the illustration in Figure 3c). ~~Such rule can be~~ Rule curves are tailored to each reservoir within the basin by determining the time at which the minimum and maximum water levels are reached (May and November, in the Mekong; Piman et al. (2012)), ~~and~~ setting the value of the minimum and maximum water levels~~. In our case, we use the minimum and maximum elevation levels of each reservoir.~~ (the minimum and maximum elevation levels of each reservoir, in our case), and finally connecting these points with a piecewise linear function that gives us the daily target level for each calendar day.

As shown in Figure 3c, there are three water levels that divide the storage into four zones. These levels are the dead water (or minimum elevation) level, the target water level, and the full (or maximum elevation) level. If the water level falls below the dead water level (Zone 1), the turbines are not operated. If the level is between the dead water and target level (Zone 2), the model first uses the information on the incoming daily inflow to solve a mass balance equation, in which the discharge from the dam is kept at zero. This is aimed to understand whether the water level is expected to go beyond the target at the end of the day. If that is the case, the model discharges through the turbines the amount of water needed to keep the level close to the target. Otherwise, the turbines are not activated. In Zone 3 (between the target and full level), the turbines are used at their maximum capacity, until the water reaches the target level. In Zone 4 (i.e., level above the maximum elevation), both turbines and spillways are used. The key advantage of the rule curves adopted here is that they do not require the calibration of any parameter. Naturally, such approach is less applicable when the information on the operating objectives is not available, or when dealing with multi-purpose water systems.

3.2 Data and preprocessing

3.2.1 Climate forcings and other input variables

Climate forcings are represented by precipitation and air temperature (maximum and minimum), which must be provided at a daily time step. As far as precipitation is concerned, we use the APHRODITE dataset (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation), developed by the University of Tsukuba, Japan, using rain-gauge data (Yatagai et al., 2012). APHRODITE is available with a spatial resolution of 0.25 degree, and has been shown by Lauri et al. (2014) to be the most suitable precipitation dataset available for the Mekong basin. A similar observation applies to the CFSR (Climate Forecast System Reanalysis) maximum and minimum temperature dataset (Saha et al., 2014). These data are then interpolated to meet the spatial resolution of 0.0625 degrees adopted in our implementation. More specifically, we use the bilinear interpolation method, which has found successful application in some recent studies (e.g., Hoang et al. (2016); Shin et al. (2019)). We also bias correct APHRODITE dataset (using a multiplying factor of 1.26), as recommended by Lauri et al. (2014).

The monthly Leaf Area Index and albedo are derived from the Moderate Resolution Imaging Spectroradiometer (Terra MODIS) satellite images, which represent changes in canopy and snow coverage over time. (It is worth noting that snowmelt only marginally contributes to the streamflow of the Mekong River; Räsänen et al. (2016).) Land use and land cover data are obtained from the Global Land Cover Characterization (GLCC) dataset, developed by the United States Geological Survey. We choose this product because it was completed in 1993, close to the simulation period adopted in our study (1995–2005). With such choice, we make sure that the influence of land use dynamics on the model parameterization is minimized. Soil data are extracted from the Harmonized World Soil Database (HWSD), developed by the International Institute for Applied System Analysis and Food and Agriculture Organization, and last updated in 2013. Both land use and soil maps are generated with the majority resampling technique, since their original spatial resolution is 30 arcsecond (approximately 1 km). This technique assigns the most common values found from the group of involved pixels to the new cell. The resulting maps are illustrated

in Figure 4a-b. The land use map shows that the upper reaches of the basin are characterized by the presence of grassland, while the lower reaches—with complex terrain and large altitudinal variations—present mixed coniferous forest ecoregions. Soil characteristics are also heterogeneous: in the central and northern part of the basin, soil is characterized by a shallow layer consisting of loam, sandy loam, and clay. At the border between China, Myanmar, and Laos (near Chiang Saen station), soil characteristics are dominated by the presence of sandy clay loam. ~~The monthly Leaf Area Index and albedo are derived from the Moderate Resolution Imaging Spectroradiometer (Terra MODIS) satellite images, which represent changes in canopy and snow coverage over time. (It is worth noting that snowmelt only marginally contributes to the streamflow of the Mekong River; Räsänen et al. (2016).)~~

To estimate the flow directions, we use the Global 30 Arc-Second Elevation (GTOPO30) DEM, which has been adopted in several studies (e.g., Kite (2001); Wu et al. (2012); Li et al. (2013)). First, we mask this DEM with the shape of the upper Mekong basin. Since GTOPO30 has a spatial resolution of 30 arcsecond, we then resample the DEM to the resolution of our VIC model using the average resampling technique (Hoang et al., 2019). Finally, we manually correct the flow direction map generated by ArcGIS by comparing it to a detailed river network provided by the Mekong River Commission. Such correction is necessary, since errors are to some extent unavoidable when automatically generating a flow direction map—because overland runoff and interflow directions depend on the relation between hillslope characteristics and adopted spatial resolution. The resulting flow direction map is illustrated in Figure 4c.

3.2.2 Dams and reservoir informations

Our model requires detailed information on the reservoirs, namely location, storage capacity, dam height, dead storage, turbine design discharge, and maximum and minimum elevation levels. Such information (summarized in Table 2) was retrieved by cross-checking the databases provided by the Mekong River Commission, the International Commission On Large Dams, and the Global Reservoir and Dam Database. Since data on reservoir bathymetry are not available, we modelled the storage-depth relationship with Liebe's method, which assumes that the reservoir is shaped like a top-down pyramid cut diagonally in half (Liebe et al., 2005). In other words, the relation between reservoir volume (V) and depth (or level, h) is equal to $V = ah^3$, where a is a shape factor equal to V_{cap}/h^3_{max} (V_{cap} is the live storage capacity and h_{max} the maximum water depth). This method has been adopted for regional and global studies (see Ng et al. (2017); Shin et al. (2019)).

As for the maximum reservoir extent (needed to determine the reservoir cells), the existing databases do not provide detailed information, such as the reservoir polygon, so we proceeded by analyzing remote sensed data. More specifically, we extracted surface water profiles from Landsat TM and ETM+ imagery. Landsat images are raster grids with seven layers corresponding to seven bands (excluding the panchromatic band). The Normalized Difference Water Index (NDWI) was calculated using the near-infrared (NIR, Band 4) and Short-Wave infrared (SWIR, Band 5) bands: NDWI=(NIR-SWIR)/(NIR+SWIR). Water bodies have NDWI values greater than 0.3 (McFeeters, 2013), so from the NDWI raster we can create a binary raster in which 1 denotes a reservoir cell (and 0 a non-reservoir cell). This process can yield an accurate estimation of the reservoir cells, since Landsat images have a spatial resolution of 30 x 30 m.

When calculating the daily mass balance for each reservoir, we consider three main processes, namely inflow, evaporation, and release. Infiltration and seepage (via dam body, abutment, and foundation) are neglected. That is because of two reasons. First, the Upper Mekong basin is a mountainous region, with mostly rocks and a shallow Quaternary alluvium (see Section 2), so infiltration losses are to some extent marginal as compared to inflow, release, and evaporation. Second, the dams considered in our study are built with concrete (and with rocky abutments and foundations), so seepage is indeed limited.

3.3 Experimental setup

To carry out the calibration exercise (with and without reservoirs), we couple VIC with the $\varepsilon-$NSGAII algorithm (Reed et al., 2013), which has found successful application in many water resources problems—including model calibration (ibidem). In our case, the decision variables are represented by the six parameters controlling the rainfall-runoff process in VIC (Section 3.1.1), and whose range of variability is reported in Table 1. As for the objective functions, we consider two goodness of fit statistics dependant upon the simulated streamflow, namely the Nash-Sutcliffe Efficiency (NSE) and Transformed Root Mean Square Error (TRMSE), which assess the model performance on high and low flows, respectively (Dawson et al., 2007). The NSE is defined as:

$$NSE = 1 - \frac{\sum_{t=1}^{n}(Q_s^t - Q_o^t)^2}{\sum_{t=1}^{n}(Q_o^t - \overline{Q_o})^2}, \tag{1}$$

where n is the number of time steps, Q_s^t the simulated streamflow (at time t), Q_o^t the observed streamflow (at Chiang Saen station), and $\overline{Q_o}$ the mean of the observed streamflow. The TRMSE is defined as:

$$TRMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(z_{s,t} - z_{o,t})^2}, \tag{2}$$

where $z_{s,t}$ and $z_{o,t}$ represent the value of the simulated and observed streamflow (at time t) transformed by the expression $z = \frac{(1+Q)^\lambda - 1}{\lambda}, (\lambda = 0.3)$. In other words, λ scales down the values of the streamflow, and TRMSE thus emphasizes the errors on the low flows. In this specific modelling problem, capturing both high and low flows is particularly important, since the riverine ecosystems are sensitive to both dry and wet conditions (Hoang et al., 2016).

Both objective functions are calculated for the period 1996–2005—after a one-year spin-up period, 1995—and scaled between 0 and 1, so we set only one value of ε (equal to 0.001). The other $\varepsilon-$NSGAII parameters to setup are the size of the initial population and number of function evaluations, which are equal to 10 and 250—a setting that strikes a reasonable balance between the computational requirements of the calibration exercise and the quality of the solutions. Each calibration exercise (with and without reservoirs) is solved with 20 different random seeds, so as to characterize the variability in the $\varepsilon-$NSGAII stochastic search process. The final set of Pareto-efficient solutions (i.e., alternative parameterizations of VIC) thus corresponds to the set of Pareto-efficient solutions identified across all 20 seeds. All experiments are carried out on an Intel (R) Xeon (R) W-2175 CPU 2.50 GHz with 128 GB RAM running Linux Ubuntu 16.04 (Xenial Xerus), using a Python implementation of various MOEAs (Platypus) that allows to ~~parellelize~~ parallelize the optimization experiments. ~~The average runtime (across the 20 seeds) is about 200 hours.~~ For each of the 20 seeds, we used four cores, taking approximately 200 hours per core (wall-clock time).

Since six (out of eleven) dams became operational during the study period (see Table 2), the VIC simulation with reservoirs is implemented in such a way to activate the reservoirs at the right time. In this specific implementation, we do not use filling strategies different from the rule curves described in Section 3.2.2, because all six dams reach a steady-state operation within a few months (data not shown).

3.4 Climate change data

For our second experiment, we used the CMIP5 climate projections to derive climate change scenarios for the period 2050–2060. Since the data provided by the Coordinated Regional Climate Downscaling Experiment only cover one GCM for our study site (Giorgi and Gutowski Jr, 2015), we followed the approach taken by previous studies (e.g., Hoang et al. (2016, 2019)), and proceeded by using GCM projections as basis for our scenarios. As far as the GCMs are concerned, we used ACCESS1-0, CCSM4, CSIRO Mk3.6, HadGEM2-ES, and MPI-ESM-LR, whose reliability for this region has been evaluated in a few previous studies (Sillmann et al., 2013; Huang et al., 2014; Ul Hasson et al., 2016; Hoang et al., 2016). The main characteristics of the GCMs are summarized in Table 3. As for the Representative Concentration Pathways (RCPs), we chose RCPs 4.5 and 8.5. The former is a medium-to-low scenario that assumes a stabilization of radiative forcing to 4.5 W m^{-2} by 2100, while the latter is a high emission scenario based on an increase of the radiative forcing to 8.5 W m^{-2} by 2100. These two RCPs should provide a broad range of climate variability for the region—and thus exclude RCP 2.6, which is characterized by the lowest radiative forcings.

To prepare the precipitation and temperature data used by VIC, we then re-gridded and bias-corrected the GCMs outputs. The first step is necessary to overcome the limited spatial resolution of the GCMs (our VIC implementation uses a resolution of $0.0625° \times 0.0625°$), and is carried with the bilinear interpolation method. The bias-correction is performed with the delta method (Diaz-Nieto and Wilby, 2005; Choi et al., 2009), which has already been applied to our study site (Lauri et al., 2012). With this method, we calculate correcting factors for precipitation and temperature using the following expressions:

$$\Delta_{PRE} = \frac{\bar{P}_{\text{series},i}}{\bar{P}_{\text{ref},i}}, \tag{3}$$

$$\Delta_{TEMP} = \frac{\bar{T}_{\text{series},i} - \bar{T}_{\text{ref},i}}{\sigma_{\text{ref},i}}, \tag{4}$$

where $\bar{P}_{\text{series},i}$ and $\bar{T}_{\text{series},i}$ are the (11 year) average precipitation and temperature for month i produced by the GCM in our control period (1995–2005), $\bar{P}_{\text{ref},i}$ and $\bar{T}_{\text{ref},i}$ the (11 year) average observed precipitation and temperature for month i in the period 1995–2005, and $\sigma_{\text{ref},i}$ the standard deviation of the monthly average temperature during the same period for month i. These factors were then used to correct the future climate projections for each time series (using the same factor for all daily data in a given month).

The impact of climate change on hydrological processes are often assessed by studying changes in the flow regime, and, in particular, changes in the monthly, seasonal, and annual river discharges (Lauri et al., 2012, 2014). More recently, some studies have focussed on hydrological extremes, such as high (Q_5) and low flows (Q_{95}) (Hoang et al., 2016). Since our goal is

to demonstrate that different model parameterizations caused by the absence (presence) of water reservoirs can largely impact the results of climate change assessments—and not to push forward the boundaries of climate change impact assessments—we chose a simple and established criterion, namely the annual and monthly river discharges at ~~the catchment outlet (Chiang Saen gauging station)~~ Chiang Saen and Jiuzhou stations.

4 Results

To discuss about the impact of water reservoirs on the parameterization of hydrological models, we first compare the results of the calibration exercise carried out with and without reservoirs, and then proceed by comparing the performance of two selected parameterizations on the climate change impact assessment.

4.1 Model parameterization

The optimization-based parameterization exercise yielded a total of 118 and 109 parameterizations (or Pareto-efficient solutions) for the VIC implementations with and without reservoirs, respectively. To prove our hypothesis that the calibration process may somehow compensate for a deficiency in the model structure—the absence of reservoirs, in our case—we begin by analyzing the values of the goodness of fit statistics, namely NSE and TRMSE. Figure 5 reports the probability plots of NSE and TRMSE values obtained for the two model setups: results show that the calibration exercise yields a reasonable modelling accuracy, with NSE and TRMSE varying in the ranges 0.68–0.79 and 8.10–16.69. More interestingly, these results show that the NSE and TRMSE values of both model setups belong to the same range of variability and follow an almost identical distribution. In addition, all NSE and TRMSE values of the models without reservoirs fall within the 95% confidence limits calculated using the NSE and TRMSE values attained by the models with reservoirs. To corroborate this finding, we carried out a Kolmogorov-Smirnov two sample test to reject the null hypothesis that the values of NSE (and TRMSE) produced by the two model setups come from the same distribution. For both goodness of fit statistics, the hypothesis cannot be rejected (with a 5% significance level). Overall, this confirms that the accuracy of the models is not affected by the presence (absence) of the reservoirs.

How does the parameterization compensate for the absence of water reservoirs? To answer this question, we visualize both goodness of fit statistics (NSE and TRMSE) and model parameters (D_s, D_{max}, W_s, b, d_1 and d_2) in a parallel-coordinate plot (Figure 6). These eight variables are shown in eight parallel axes, so each line connecting the axes represents a parameterization (i.e., a solution of the optimization problem) along with the corresponding value of the goodness of fit statistics (i.e., the objectives). Blue and red lines denote solutions obtained with and without reservoirs, respectively. First of all, one can notice that while NSE and TRMSE spread over the same ranges (results discussed in the previous paragraph), the presence/absence of reservoirs consistently yields different parameterizations. Let's analyze them. The value of b—characterizing the shape of the Variable Infiltration Capacity curve—belongs to two distinct ranges (0.319–0.495 and 0.002–0.195) for the model implementation with and without reservoirs, respectively, indicating that the model without reservoirs has higher infiltration and lower surface runoff than the model with reservoir (recall that a higher value of b leads to a lower infiltration rate and

higher surface runoff; Section 3.1.1). A similar observation applies to the parameters D_s, D_{max}, and W_s, which determine the shape of the Baseflow curve. In this case, the model without reservoirs has higher values of D_{max} (i.e., maximum baseflow) and lower values of D_s and W_s (i.e., fraction of D_{max} where rapidly increasing baseflow begins, and fraction of the maximum soil moisture in the lower layer where rapidly increasing baseflow occurs), suggesting that the absence of reservoirs leads to model paramaterizations that favour the generation of baseflow in the lower layer. Finally, we can note that d_1 (the thickness of the first layer) in the models without reservoirs tends to be larger, indicating that these model instances increase the water storage capacity of the top layer. The only parameter that does not appear to depend on the absence (or presence) of water reservoirs is d_2, the thickness of the second soil layer. This result is corroborated by a global sensitivity analysis, which shows that d_2 is indeed the parameter with the least influence on the model output (Figure S1). Overall, it appears that the calibration process compensates for the absence of water reservoirs by determining values of the soil parameters that can somehow 'mimic' the alterations caused by water reservoirs, namely an increase in the evaporation and delay in the peak flows—obtained by increasing infiltration, baseflow, and soil water storage capacity.

To further understand the unintended consequences of the absence of water reservoirs, we select two model parameterizations (with and without reservoirs) characterized by the same performance over the period 1996–2005. The values of NSE, TRMSE, and model parameters are illustrated in Figure 7a, while the simulated daily discharges produced by both models are compared in the scatter plot of Figure 7b. ~~(The two models were also tested on a independent validation period, 1985–1995, for which we obtained a value of NSE equal to 0.724 (with reservoirs) and 0.718 (without reservoirs), and a value of TRMSE of 11.735 and 12.103.)~~ In Figure 8, we contrast the average values of simulated baseflow and runoff during the dry (December–April) and wet (May–November) seasons of the period 1996–2005. Unsurprisingly, results show that during the dry season the model without reservoirs generates more baseflow and runoff than the model with reservoirs (left four panels of Figure 8): during the dry months, hydropower reservoirs release part of the water stored during the monsoon (recall the rule curves described in Section 3.1.2); a process simulated by the model without reservoirs by increasing both baseflow and runoff—and, therefore, the discharge at the catchment outlet. During the wet season, we find an opposite trend: in these months, hydropower reservoirs tend to store part of the water (thus reducing the discharge at the catchment outlet), so the model without reservoirs slightly decreases the discharge by reducing baseflow and runoff (right four panels of Figure 8). We also note that the difference between the two models is clearer during the dry season, when a larger amount of the water volumes is controlled by the hydropower reservoirs. The effect of such flawed representation of baseflow and runoff is further demonstrated by validating the simulated discharge at Jiuzhou station. Figure 9 shows a macroscopic difference between the models calibrated with and without reservoir. In particular we note that the model calibrated without reservoirs largely overestimates the dry season flow and slightly underestimates the wet season one; a result confirmed by the values of NSE (equal to 0.82 and 0.79 for the model with and without reservoirs) and TRMSE (equal to 21.48 and 28.95). One may also suspect that these unintended consequences could further propagate in downstream applications of the models, such as a climate change impact assessment.

4.2 Climate change impact assessment

To begin the climate change impact assessment, we compare the data produced by the GCMs over the reference and future period (1996–2005 and 2050–2060). In general, the total annual precipitation in the Lancang basin is projected to increase under almost all climate change scenarios—only the CSIRO MK3-RCP 8.5 scenario projects a -3.12% decrease in the total annual precipitation. Yet, we observe a large spatial variability in the total annual rainfall within each scenario (see Figure S2). For example, in ACCESS-RCP 4.5, rainfall changes vary between -2% in the central part of the basin to more than +10% in the southern part. All scenarios (but for CSIRO MK3-RCP 8.5) tend to share a similar spatial pattern, in which the lower part of the basin exhibits an increase in the projected precipitation. As for the temperature, we observe an increase in both minimum and maximum temperature across all scenarios (see Figure S3), with higher warming for the RCP 8.5. Also in this case, we can note some variability across the GCMs as well as the spatial domain. As discussed in Hoang et al. (2016), these precipitation and temperature scenarios represent an improvement with respect to the CMIP3 ones, which shown a broader variability. However, there still are some non-negligible differences across the scenarios that are likely to cause different projections of the annual and monthly river discharges.

~~The expected impact of climate changes on the annual river discharges is synthesized in Table 4, where we report the relative changes in discharge with respect to the period 1996–2005. These results show an expected increase in the river flow under climate change. The ensemble mean for RCP 4.5 and RCP 8.5 is rather similar (+13.56 and +13.83% for the model with reservoirs, +13.62 and +13.92% for the model without reservoirs), while the range of variability is different. All scenarios under RCP 4.5 are associated to an increase in the discharge, while the scenario predicted by the CSIRO GCM (under the RCP 8.5) is associated to a slight decrease in the river discharge. Overall, these results do not highlight any macroscopic difference between the projections issued by the two VIC implementations.~~

~~Figure 11a-b illustrates the projected monthly discharges at Chiang Saen for RCP 4.5 and RCP 8.5 (for the model without reservoirs). First of all, we can note that the ensemble range is broader during the monsoon months. This is an expected result, since the flow regime of the Lancang is driven by the monsoon. Results also show that the majority of the scenarios cause higher discharges (with respect to the baseline) during most of the months. This result is simply explained by the expected increase in precipitation described above. In terms of RCPs, it appears that the ensemble range of the RCP 8.5 is slightly broader than the RCP 4.5's one, and that RCP 8.5 is expected to cause higher flows during the months of September, October, and November.~~

The expected climate change impacts on the annual river discharge at Chiang Saen are synthesized in Table 4, where we report the relative changes in discharge with respect to the period 1996–2005. Interestingly, it appears that the projections are robust with respect to the representation of the water reservoirs. Indeed, the model with and without reservoirs yield comparable ensemble means and ranges for the two RCPs. Specifically, we find that the annual discharge is predicted to increase in the vast majority of the scenarios, in response to the increase in precipitation described above. Such similarity between the projections is arguably attributable to the calibration process, which generates models producing similar aggregate performance measures at Chiang Saen station.

What is perhaps more interesting is a comparison between the ~~projections yielded~~ monthly discharges at Chiang Saen predicted by the models with and without reservoirs. ~~Both~~ While both models produce similar ensemble ranges (see Figure 10a-d)~~; yet~~, a closer analysis of the data reveals a non-negligible difference in the minimum, maximum, and average monthly discharges (across the GCM scenarios) produced by the two models (Figure 10e-f). In particular, the model with reservoirs predicts higher discharges in the July–September period and lower discharges in October and November. Note that such difference is consistent across both RCPs. Since both models share the same rainfall and temperature scenarios, the only cause for this stark difference can stand in the unintended consequences of the parameterization process. As explained in Section 4.1, the model without reservoirs shows two 'artefacts' that help compensate for the absence of the hydropower reservoirs: first, it increases both baseflow and runoff during the dry season (to account for the water discharged to sustain hydropower production in the dry months); second, it decreases baseflow and runoff (to account for part of the water stored by the dams during the wet months). The latter artefacts is responsible for the macroscopic change in the hydrograph described above. In the wetter conditions depicted by the GCM-RCP scenarios, the hydropower reservoirs of the Lancang basin receive larger inflows, part of which is directly spilled into the downstream reaches (data not shown). This is an unprecedented situation for the model without reservoirs, which cannot simulate an increase in the use of the spillways. In fact, this model tends to reproduce the dynamics learned during the calibration process, that is, storing part of the water (in the lower soil layer) during the monsoon season and slowly discharging it in the following months. Naturally, the difference between the monthly discharges predicted by the two models becomes even more apparent when we consider Jiuzhou station, which was not used in the calibration process. As depicted in Figure 11, the model without reservoirs consistently yields higher discharges in the pre-monsoon season and lower discharges in the monsoon season. Note that, in some months, the difference between the average monthly discharges produced by the two models causes an uncertainty larger than the one surrounding the downscaled climate projections. For instance, the average monthly discharge in March (under both RCPs) predicted by two models is about 500 and 750 m^3/s, that is, a 50% difference.

5 Discussion and Conclusions

This work contributes to the existing literature on large-scale hydrological modelling by studying the effect of water reservoir storage and operations on the parameterization of process-based models. To this purpose, we developed a computational framework consisting of VIC and the Multi-Objective Evolutionary Algorithm $\varepsilon-$NSGAII, which we used to calibrate the model parameters through a simulation-optimization process. Our framework also includes a novel variant of VIC that simulates both storage dynamics and operations of water reservoirs. Using the Lancang river basin as a case study, we calibrated two implementations of VIC, with and without reservoirs. Inline with previous studies (e.g., de Paiva et al. (2013); Abbaspour et al. (2015)), we found that the model without reservoirs attains a reasonable modelling accuracy. In fact, we found that the calibration process of both model implementations yields de facto the same values of the goodness of fit statistics (NSE and TRMSE), suggesting that the model parameterization helps compensates for a structural error, namely the absence of the water

reservoirs. More specifically, this effect is achieved by determining the values of six soil parameters (D_s, D_{max}, W_s, b, d_1 and d_2) that let this model implementation emulate the presence of water reservoirs.

The first implication of a flawed parameter estimation stands in a poor representation of key hydrological processes, such as surface runoff, infiltration, and baseflow. In our case, we found that, during the dry months, the models calibrated without water reservoirs generate a higher amount of baseflow and runoff than the models with reservoirs. This is an artefact needed to reproduce the higher discharges of hydropower dams that sustain the production of hydro-electricity in the dry season. Vice versa, baseflow and runoff are reduced during the wet months, so as to account for the decrease in peak flows caused by the fact that dams store part of the water for the following dry season. A poor parameter estimation is also likely to affect several downstream applications of a hydrological model. In our second experiment we exemplify this concept through a climate change impact assessment, in which we contrasted the annual and monthly discharges projected by two selected models (with and without reservoirs). Both models show a similar trend in the flow regime—i.e., increased monthly discharges during the monsoon season, caused by the projected increase in precipitation—a results found in previous studies (Lauri et al., 2012; Hoang et al., 2016, 2019). Yet, one cannot neglect the different nuances of the flow regime alterations predicted by the two models. In particular, the model with reservoirs presents higher discharges at the peak of the monsoon season than the model without reservoirs. These nuances may impact some of the conclusions of a climate change impact assessment as well as other model-based studies depending on a reliable estimation of the flow regime.

~~Naturally, the framework adopted in this study has a few limitations.~~ Like any hydrological modelling study, also this work builds on a few modelling assumptions that should be properly discussed. First, our model calibration focuses solely on six main parameters controlling the rainfall-runoff process, and assumes that they are homogeneously distributed across the basin. As explained in Section 3.1.1, the choice of these ~~six~~ parameters is rather established in the literature (Dan et al., 2012; Park and Markus, 2014; Xue et al., 2015); yet, it is reasonable to expect that the use of more parameters could further improve the model accuracy. As for the use of homogeneously-distributed parameters, our modelling choice is justified by the fact that the use of heterogeneously-distributed parameters would largely impact the computational requirements of the calibration process. We also note that there are no reasons to believe that the use of more (or spatially-distributed) parameters would deeply alter the main findings of this work. ~~Second, our study relies on rule curves conceived to maximize the hydropower production. This is not a limitation of our study (reservoirs in the Lancang are indeed operated for hydropower supply; Räsänen et al. (2017)), but it certainly constraints the applicability of our modelling framework to regions in which the reservoir operating rules are either available or well understood. It is also a further testimony of the importance of studies aimed to deduct reservoir operating rules from discharge and remote sensed data nsed (Bonnema and Hossain, 2017; Coerver et al., 2018).~~ Second, the large spatial domain—and associated soil water retention capacity—might be a factor controlling the capability of the calibration process to compensate for the absence of water reservoirs. In other words, such capability might be dependant on the relation between soil water retention capacity and total storage volume of the reservoirs. In a small basin regulated by a large dam, a modified representation of runoff, infiltration, and baseflow may not be sufficient to compensate fully for a poor representation of reservoir storage and operations. Third, we focussed our attention on water reservoirs, which are indeed the infrastructures affecting the flow regime in the Lancang. In the lower Mekong basin (not considered in our spatial domain), the flow regime has

been modified not only by hydropower reservoirs, but also by withdrawals for irrigation supply (Hoang et al., 2019). Looking forward, It would thus be interesting to extend the spatial domain of our model and study how these withdrawals could affect its parameterization.

As the pervasiveness of water resources management in earth system models expands, so too does the need for a deeper understanding of the mechanisms regulating the calibration process. The explicit representation of water reservoirs—and other infrastructure—is indeed likely to result in more realistic soil parameters; an hypothesis whose verification depends on the availability of observations about soil physical properties for large spatial domains. In turn, this highlights the importance of studies aimed to infer such properties from remotely sensed images (Chang and Islam, 2000; Chabrillat et al., 2019). A related topic that may also deserve future research is the robustness of these models with respect to changes in the operations or physical characteristics of dams. Variability in water and energy demand is a key driver for multiple management and planning interventions (e.g., modifications of the operating rules, construction on new storage), so it is paramount to know the extent to which models can still capture key hydrological processes once these modifications are in place.

Overall, the findings of this study reinforce the message that water infrastructures—and their operational settings—play a key role on the reliability of a hydrological modelling exercise, like the quality of the hydro-meteorological data, the model structure, or the calibration process (Francés et al., 2007; Madsen, 2000). These findings gain further prominence if one considers the expected increase in hydropower development in several regions of the world (Zarfl et al., 2015).

Author contributions. TDD and SG conceptualized the paper and its scope. Data collection, model implementation, and experiments were carried out by TDD, with inputs from AFMKC and SG. All authors contributed to the manuscript preparation.

Code and data availability. Precipitation and air temperature data were retrieved from APHRODITE and CFSR datasets, available at http://www.chikyu.ac.jp/precip/english/ and https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr. Land use and land cover data were obtained from the GLCC dataset (https://www.usgs.gov/centers/eros/science/), while the soil data were extracted from the HWSD database (http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/). The Terra MODIS satellite images (used to calculate the monthly Leaf Area Index and albedo) are available at https://modis.gsfc.nasa.gov. The Landsat TM and ETM+ imagery are available at https://earth.esa.int/web/sppa/mission-performance/esa-3rd-party-missions/landsat-1-7/tm-etm/sensor-description. The global Digital Elevation Model (GTOPO30) is available at http://www.temis.nl/data/gtopo30.html. The GCMs projections were retrieved from https://esgf-node.llnl.gov/projects/esgf-llnl/. All these data are publicly available. The daily discharge data at Chiang Saen and the design specifications of all dams were obtained by the authors from the Mekong River Commission and the International Commission On Large Dams, so they cannot be shared without their consent. Additional data about the dams were retrieved from the Global Reservoir and Dam Database, available at http://globaldamwatch.org/grand/. The variant of VIC used in this study—named VIC-Res—is available at https://github.com/thanhiwer/VICRes.

Competing interests. The authors declare that they do not have individual or collective conflicts of interests.

Acknowledgements. This research is supported by Singapore's Ministry of Education (MoE) through the Tier 2 project 'Linking water availability to hydropower supply—an engineering systems approach' (Award No. MOE2017-T2-1-143).

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, Journal of Hydrology, 524, 733–752, 2015.

Akter, A. and Babel, M. S.: Hydrological modeling of the Mun River basin in Thailand, Journal of Hydrology, 452, 232–246, 2012.

Alcamo, J., Döll, P., Kaspar, F., and Siebert, S.: Global change and global scenarios of water use and availability: an application of WaterGAP 1.0, Center for Environmental Systems Research (CESR), University of Kassel, Germany, 1720, 1997.

Bellin, A., Majone, B., Cainelli, O., Alberici, D., and Villa, F.: A continuous coupled hydrological and water resources management model, Environmental Modelling & Software, 75, 176–192, 2016.

Bierkens, M. F.: Global hydrology 2015: State, trends, and directions, Water Resources Research, 51, 4923–4947, 2015.

Bunn, S. E. and Arthington, A. H.: Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity, Environmental management, 30, 492–507, 2002.

Carling, P. A.: The geology of the lower Mekong River, in: The Mekong, pp. 13–28, Elsevier, 2009.

Chabrillat, S., Ben-Dor, E., Cierniewski, J., Gomez, C., Schmid, T., and Van Wesemael, B.: Imaging spectroscopy for soil mapping and monitoring, Surveys in Geophysics, 40, 361–399, 2019.

Chang, D.-H. and Islam, S.: Estimation of soil physical properties using remote sensing and artificial neural network, Remote Sensing of Environment, 74, 534–544, 2000.

Choi, W., Rasmussen, P. F., Moore, A. R., and Kim, S. J.: Simulating streamflow response to climate scenarios in central Canada using a simple statistical downscaling method, Climate Research, 40, 89–102, 2009.

Cochrane, T., Arias, M., and Piman, T.: Historical impact of water infrastructure on water levels of the Mekong River and the Tonle Sap system, Hydrology and Earth System Sciences, 18, 4529–4541, 2014.

Dan, L., Ji, J., Xie, Z., Chen, F., Wen, G., and Richey, J. E.: Hydrological projections of climate change scenarios over the 3H region of China: A VIC model assessment, Journal of Geophysical Research: Atmospheres, 117, 2012.

Dang, T. D., Cochrane, T. A., Arias, M. E., Van, P. D. T., and de Vries, T. T.: Hydrological alterations from water infrastructure development in the Mekong floodplains, Hydrological Processes, 30, 3824–3838, 2016.

Dang, T. D., Cochrane, T. A., Arias, M. E., et al.: Future hydrological alterations in the Mekong Delta under the impact of water resources development, land subsidence and sea level rise, Journal of Hydrology: Regional Studies, 15, 119–133, 2018.

Dawson, C. W., Abrahart, R. J., and See, L. M.: HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts, Environmental Modelling & Software, 22, 1034–1052, 2007.

de Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M.-P., Frappart, F., Calmant, S., and Mendes, C. A. B.: Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin, Water Resources Research, 49, 1226–1243, 2013.

Diaz-Nieto, J. and Wilby, R. L.: A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom, Climatic Change, 69, 245–268, 2005.

Finer, M. and Jenkins, C. N.: Proliferation of hydroelectric dams in the Andean Amazon and implications for Andes-Amazon connectivity, Plos one, 7, e35 126, 2012.

Francés, F., Vélez, J. I., and Vélez, J. J.: Split-parameter structure for the automatic calibration of distributed hydrological models, Journal of Hydrology, 332, 226–240, 2007.

Franchini, M. and Pacciani, M.: Comparative analysis of several conceptual rainfall-runoff models, Journal of Hydrology, 122, 161–219, 1991.

Gao, X., Zeng, Y., Wang, J., and Liu, H.: Immediate impacts of the second impoundment on fish communities in the Three Gorges Reservoir, Environmental Biology of Fishes, 87, 163–173, 2010.

Giorgi, F. and Gutowski Jr, W. J.: Regional dynamical downscaling and the CORDEX initiative, Annual Review of Environment and Resources, 40, 467–490, 2015.

Gupta, A.: Geology and landforms of the Mekong Basin, in: The Mekong, pp. 29–51, Elsevier, 2009.

Haddeland, I., Lettenmaier, D. P., and Skaugen, T.: Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins, Journal of Hydrology, 324, 210–223, 2006.

Han, Z., Long, D., Fang, Y., Hou, A., and Hong, Y.: Impacts of climate change and human activities on the flow regime of the dammed Lancang River in Southwest China, Journal of Hydrology, 2019.

Hanasaki, N., Kanae, S., and Oki, T.: A reservoir operation scheme for global river routing models, Journal of Hydrology, 327, 22–41, 2006.

He, D., Lu, Y., Li, Z., and Li, S.: Watercourse environmental change in Upper Mekong, in: The Mekong, pp. 335–362, Elsevier, 2009.

Hecht, J. S., Lacombe, G., Arias, M. E., Dang, T. D., and Piman, T.: Hydropower dams of the Mekong River basin: a review of their hydrological impacts, Journal of Hydrology, 2018.

Hoang, L. P., van Vliet, M. T., Kummu, M., Lauri, H., Koponen, J., Supit, I., Leemans, R., Kabat, P., and Ludwig, F.: The Mekong's future flows under multiple drivers: How climate change, hydropower developments and irrigation expansions drive hydrological changes, Science of the Total Environment, 649, 601–609, 2019.

Hoang, P., Lauri, P., Kummu, M., Koponen, J., Van Vliet, M. T., Supit, I., Leemans, H., Kabat, P., and Ludwig, F.: Mekong River flow and hydrological extremes under climate change, Hydrology and Earth System Sciences Discussions, 20, 3027–3041, 2016.

Huang, Y., Wang, F., Li, Y., and Cai, T.: Multi-model ensemble simulation and projection in the climate change in the Mekong River Basin. Part I: temperature, Environmental Monitoring and Assessment, 186, 7513–7523, 2014.

Jayawardena, A. and Mahanama, S.: Meso-scale hydrological modeling: Application to Mekong and Chao Phraya basins, Journal of Hydrologic Engineering, 7, 12–26, 2002.

Kite, G.: Modelling the Mekong: hydrological simulation for environmental impact studies, Journal of Hydrology, 253, 1–13, 2001.

Lauri, H., de Moel, H., Ward, P. J., Räsänen, T. A., Keskinen, M., and Kummu, M.: Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge, Hydrology and Earth System Sciences, 16, 4603–4619, 2012.

Lauri, H., Räsänen, T., and Kummu, M.: Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong River case study, Journal of Hydrometeorology, 15, 1532–1545, 2014.

Lee, E., Ha, K., Ngoc, N. T. M., Surinkum, A., Jayakumar, R., Kim, Y., and Hassan, K. B.: Groundwater status and associated issues in the Mekong-Lancang River Basin: international collaborations to achieve sustainable groundwater resources, Journal of Groundwater Science and Engineering, 5, 1–13, 2017.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., et al.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9, 494–502, 2011.

Leng, G., Huang, M., Voisin, N., Zhang, X., Asrar, G. R., and Leung, L. R.: Emergence of new hydrologic regimes of surface water resources in the conterminous United States under future warming, Environmental Research Letters, 11, 114 003, 2016.

Li, L., Ngongondo, C. S., Xu, C.-Y., and Gong, L.: Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa, Hydrology Research, 44, 770–788, 2013.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research: Atmospheres, 99, 14 415–14 428, 1994.

Liebe, J., Van De Giesen, N., and Andreini, M.: Estimation of small reservoir storage capacities in a semi-arid environment: A case study in the Upper East Region of Ghana, Physics and Chemistry of the Earth, Parts A/B/C, 30, 448–454, 2005.

Lohmann, D., Nolte-Holube, R., and Raschke, E.: A large-scale horizontal routing model to be coupled to land surface parametrization schemes, Tellus A, 48, 708–721, 1996.

Lohmann, D., Raschke, E., Nijssen, B., and Lettenmaier, D.: Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, Hydrological Sciences Journal, 43, 131–141, 1998.

Lu, X., Li, S., Kummu, M., Padawangi, R., and Wang, J.: Observed changes in the water flow at Chiang Saen in the lower Mekong: Impacts of Chinese dams?, Quaternary International, 336, 145–157, 2014.

Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, Journal of Hydrology, 235, 276–288, 2000.

Masaki, Y., Hanasaki, N., Biemans, H., Schmied, H. M., Tang, Q., Wada, Y., Gosling, S. N., Takahashi, K., and Hijioka, Y.: Intercomparison of global river discharge simulations focusing on dam operation—multiple models analysis in two case-study river basins, Missouri–Mississippi and Green–Colorado, Environmental Research Letters, 12, 055 002, 2017.

Maurer, E. P., Wood, A., Adam, J., Lettenmaier, D. P., and Nijssen, B.: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, Journal of Climate, 15, 3237–3251, 2002.

McFeeters, S.: Using the normalized difference water index (NDWI) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach, Remote Sensing, 5, 3544–3561, 2013.

Nazemi, A. and Wheater, H. S.: On inclusion of water resource management in Earth system models–Part 1: Problem definition and representation of water demand, Hydrology and Earth System Sciences, 19, 33–61, 2015a.

Nazemi, A. and Wheater, H. S.: On inclusion of water resource management in Earth system models–Part 2: Representation of water supply and allocation and opportunities for improved modeling, Hydrology and Earth System Sciences, 19, 63–90, 2015b.

Ng, J. Y., Turner, S. W., and Galelli, S.: Influence of El Niño Southern Oscillation on global hydropower production, Environmental Research Letters, 12, 034 010, 2017.

Park, D. and Markus, M.: Analysis of a changing hydrologic flood regime using the Variable Infiltration Capacity model, Journal of Hydrology, 515, 267–280, 2014.

Piman, T., Cochrane, T., Arias, M., Green, A., and Dat, N.: Assessment of flow changes from hydropower development and operations in Sekong, Sesan, and Srepok rivers of the Mekong basin, Journal of Water Resources Planning and Management, 139, 723–732, 2012.

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J.-F., Kim, H., Kanae, S., and Oki, T.: Incorporating anthropogenic water regulation modules into a land surface model, Journal of Hydrometeorology, 13, 255–269, 2012.

Räsänen, T. A., Lindgren, V., Guillaume, J. H., Buckley, B. M., and Kummu, M.: On the spatial and temporal variability of ENSO precipitation and drought teleconnection in mainland Southeast Asia, Climate of the Past, 12, 1889–1905, 2016.

Räsänen, T. A., Someth, P., Lauri, H., Koponen, J., Sarkkula, J., and Kummu, M.: Observed river discharge changes due to hydropower operations in the Upper Mekong Basin, Journal of Hydrology, 545, 28–41, 2017.

Reed, P. M., Hadka, D., Herman, J. D., Kasprzyk, J. R., and Kollat, J. B.: Evolutionary multiobjective optimization in water resources: The past, present, and future, Advances in Water Resources, 51, 438–456, 2013.

Ren-Jun, Z.: The Xinanjiang model applied in China, Journal of Hydrology, 135, 371–381, 1992.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., et al.: The NCEP climate forecast system version 2, Journal of Climate, 27, 2185–2208, 2014.

Shin, S., Pokhrel, Y., and Miguez-Macho, G.: High-Resolution Modeling of Reservoir Release and Storage Dynamics at the Continental Scale, Water Resources Research, 55, 787–810, 2019.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, 2013.

Tang, J., Yin, X., Yang, P., and Yang, Z.: Climate-induced flow regime alterations and their implications for the Lancang river, China, River Research and Applications, 31, 422–432, 2015.

Tang, X., Zhang, J., Wang, G., Yang, Q., Yang, Y., Guan, T., Liu, C., Jin, J., Liu, Y., and Bao, Z.: Evaluating Suitability of Multiple Precipitation Products for the Lancang River Basin, Chinese Geographical Science, 29, 37–57, 2019.

Todini, E.: The ARNO rainfall—runoff model, Journal of hydrology, 175, 339–382, 1996.

Turner, S. W. and Galelli, S.: Water supply sensitivity to climate change: An R package for implementing reservoir storage analysis in global and regional impact studies, Environmental Modelling & Software, 76, 13–19, 2016.

Turner, S. W., Ng, J. Y., and Galelli, S.: Examining global electricity supply vulnerability to climate change using a high-fidelity hydropower dam model, Science of the Total Environment, 590, 663–675, 2017.

Ul Hasson, S., Pascale, S., Lucarini, V., and Böhner, J.: Seasonal cycle of precipitation over major river basins in South and Southeast Asia: a review of the CMIP5 climate models data for present climate and future climate projections, Atmospheric Research, 180, 42–63, 2016.

Veldkamp, T. I. E., Zhao, F., Ward, P. J., de Moel, H., Aerts, J. C., Schmied, H. M., Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., et al.: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study, Environmental Research Letters, 13, 055 008, 2018.

Vörösmarty, C. J., Federer, C. A., and Schloss, A. L.: Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling, Journal of Hydrology, 207, 147–169, 1998.

Wang, X., Liang, P., Li, C., and Wu, F.: Analysis of regional temperature variation characteristics in the Lancang River Basin in southwestern China, Quaternary international, 333, 198–206, 2014.

Wang, Z., Chen, J., Lai, C., Zhong, R., Chen, X., and Yu, H.: Hydrologic assessment of the TMPA 3B42-V7 product in a typical alpine and gorge region: the Lancang River basin, China, Hydrology Research, 49, 2002–2015, 2018.

Wi, S., Ray, P., Demaria, E. M., Steinschneider, S., and Brown, C.: A user-friendly software package for VIC hydrologic model development, Environmental Modelling & Software, 98, 35–53, 2017.

Wu, H., Kimball, J. S., Li, H., Huang, M., Leung, L. R., and Adler, R. F.: A new global river network database for macroscale hydrologic modeling, Water Resources Research, 48, 2012.

Xue, X., Zhang, K., Hong, Y., Gourley, J. J., Kellogg, W., McPherson, R. A., Wan, Z., and Austin, B. N.: New multisite cascading calibration approach for hydrological models: Case study in the red river basin using the VIC model, Journal of Hydrologic Engineering, 21, 05015 019, 2015.

Yassin, F., Razavi, S., Elshamy, M., Davison, B., Sapriza-Azuri, G., and Wheater, H.: Representation of Water Management in Hydrological and Land Surface Models, Hydrology and Earth System Sciences Discussions, https://doi.org/10.5194/hess-2019-7, 2019.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., and Kitoh, A.: APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Bulletin of the American Meteorological Society, 93, 1401–1415, 2012.

Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L., and Tockner, K.: A global boom in hydropower dam construction, Aquatic Sciences, 77, 161–170, 2015.

Zhang, Y., Erkyihum, S. T., and Block, P.: Filling the GERD: evaluating hydroclimatic variability and impoundment strategies for Blue Nile riparian countries, Water International, 41, 593–610, 2016.

Zhao, Q., Liu, S., Deng, L., Dong, S., Yang, J., and Wang, C.: The effects of dam construction and precipitation variability on hydrologic alteration in the Lancang River Basin of southwest China, Stochastic Environmental Research and Risk Assessment, 26, 993–1011, 2012.

Zhou, T., Voisin, N., and Fu, T.: Non-stationary hydropower generation projections constrained by environmental and electricity grid operations over the western United States, Environmental Research Letters, 13, 074 035, 2018.

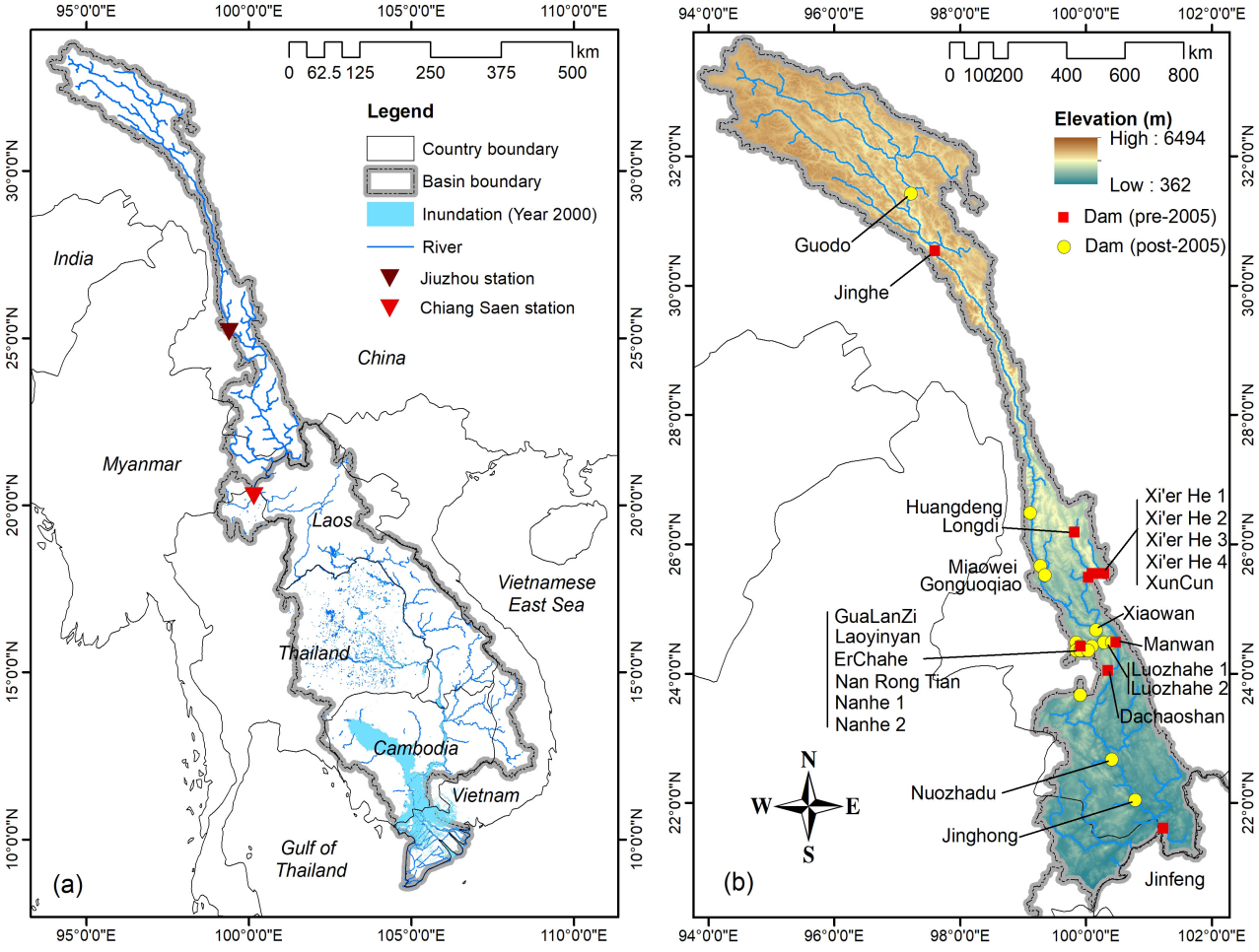

Figure 1. Mekong river basin (a); and elevation map and location of the hydropower dams in the upper Mekong basin (b). The red squares denote dams built before 2005 (and therefore included in our study), while the yellow circles indicate dams built after 2005.

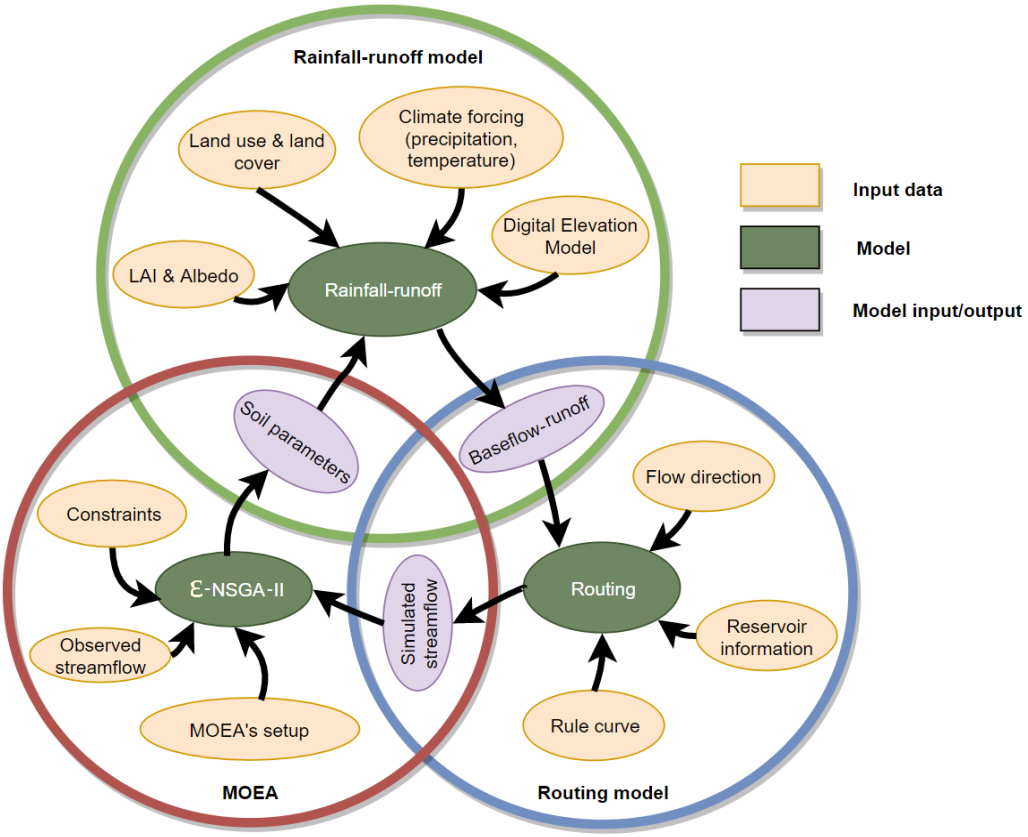

Figure 2. Computational framework adopted in the first part of this study. The framework consists of VIC's rainfall-runoff and routing modules and the MOEA $\varepsilon-$NSGAII. The output of the rainfall-runoff module (i.e., gridded baseflow and runoff) is used by the routing module, which simulates the streamflow at multiple locations within the upper Mekong basin. The simulated streamflow is then used to calculate goodness of fit statistics, whose value is optimized with $\varepsilon-$NSGAII by calibrating the parameters of the rainfall-runoff module. In other words, these parameters and goodness of fit statistics represent the decision variables and objective functions used by $\varepsilon-$NSGAII.

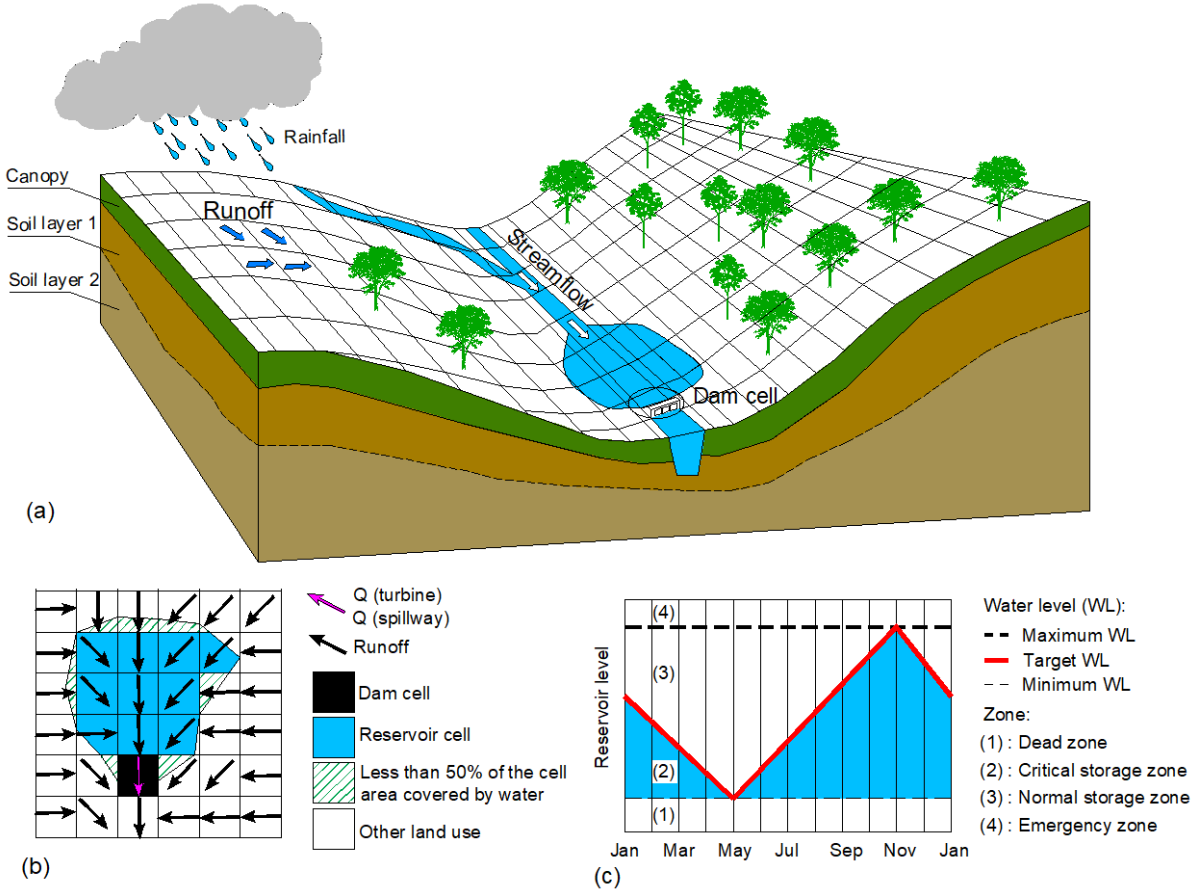

Figure 3. Graphical representation of VIC's spatial domain (adapted from http://www.hydro.washington.edu) (a), including the selection of dam cell (black), reservoir cells (blue), and cells with other land use (white and white with green lines). The black and pink arrows indicate the direction of the flow routing and discharge from the reservoir (b). Seasonal rule curve (c).

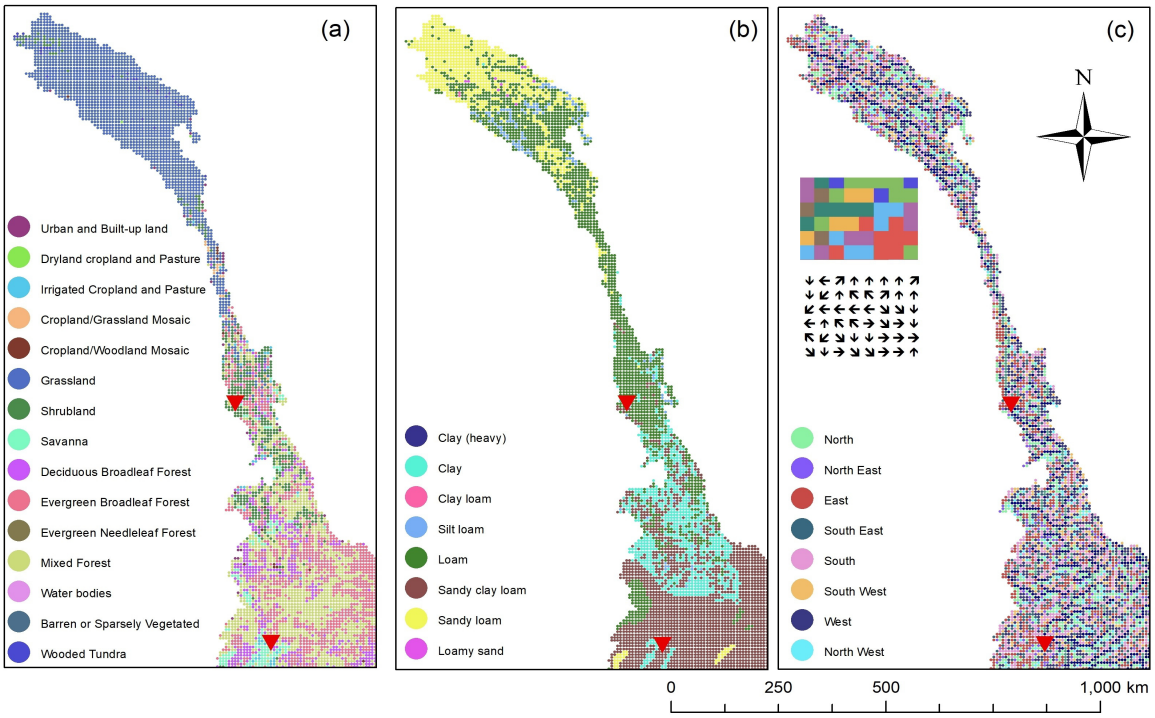

Figure 4. Land use map derived from the Global Land Cover Characterization dataset (a); soil map (for the top layer) retrieved from the Harmonized World Soil Database (b); flow direction map (c). The red triangles denote the position of the Chiang Saen and Jiuzhou gauging stations.

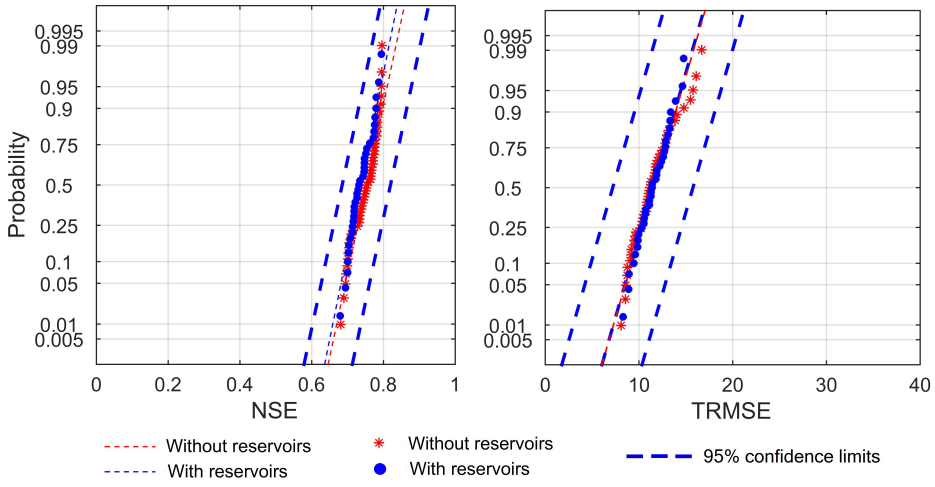

Figure 5. Probability plots for the NSE (left) and TRMSE (right) obtained in the model calibration process. The blue circles and red stars specify the results obtained by the models with and without reservoirs, respectively. The dashed blue and red lines represent the theoretical distributions. In both plots, we also report the 95% confidence limits for the models calibrated with reservoirs.

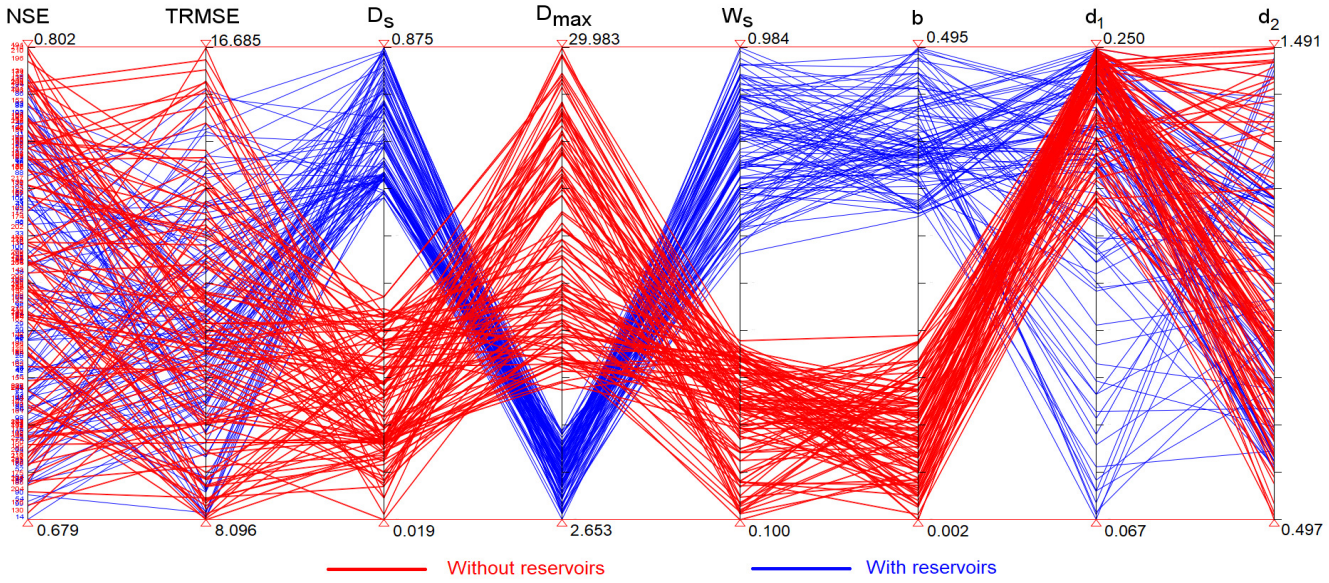

Figure 6. Parallel coordinate plot illustrating the values of the goodness of fit statistics (NSE and TRMSE) and model parameters (D_s, D_{max}, W_s, b, d_1 and d_2) obtained through the optimization-based parameterization exercise. Each line connecting the axes represents a parameterization, along with the corresponding model performance. Blue and red lines denote parameterizations obtained with and without reservoirs.

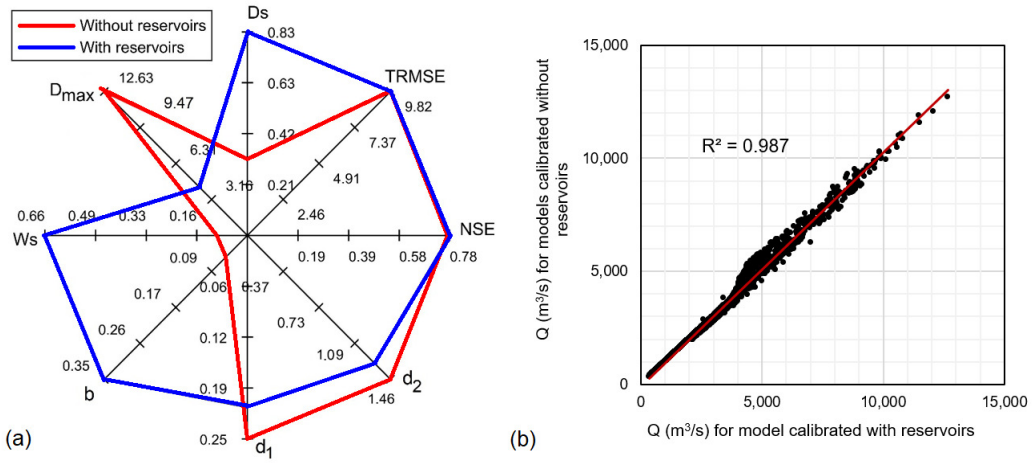

Figure 7. Radar chart illustrating the values of Nash-Sutcliffe Efficiency (NSE), Transformed Root Mean Square Error (TRMSE), and model parameters (D_s, D_{max}, W_s, b, d_1 and d_2) of the two selected models (a); scatter plot comparing the daily discharges at Chiang Sean station simulated by the two models over the period 1996–2005 (b).

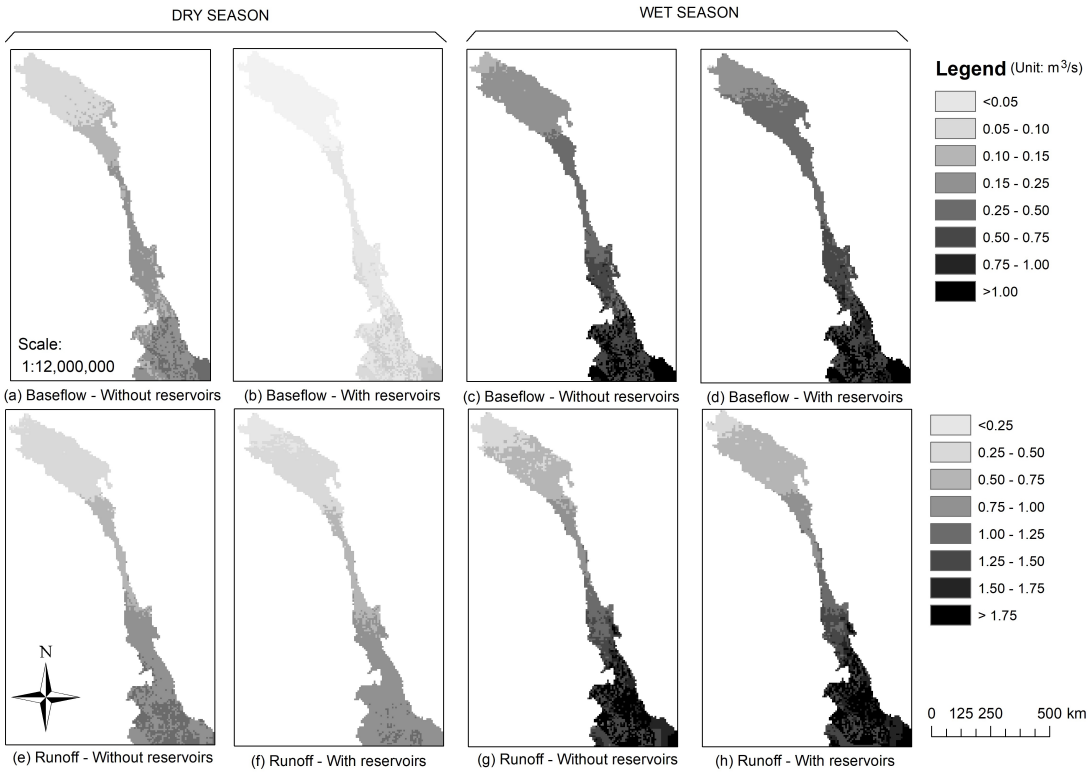

Figure 8. Average values of simulated baseflow (top panels) and runoff (bottom panels) simulated by the selected models (with and without reservoirs) during the dry (December–April) and wet (May–November) seasons of the period 1996–2005.

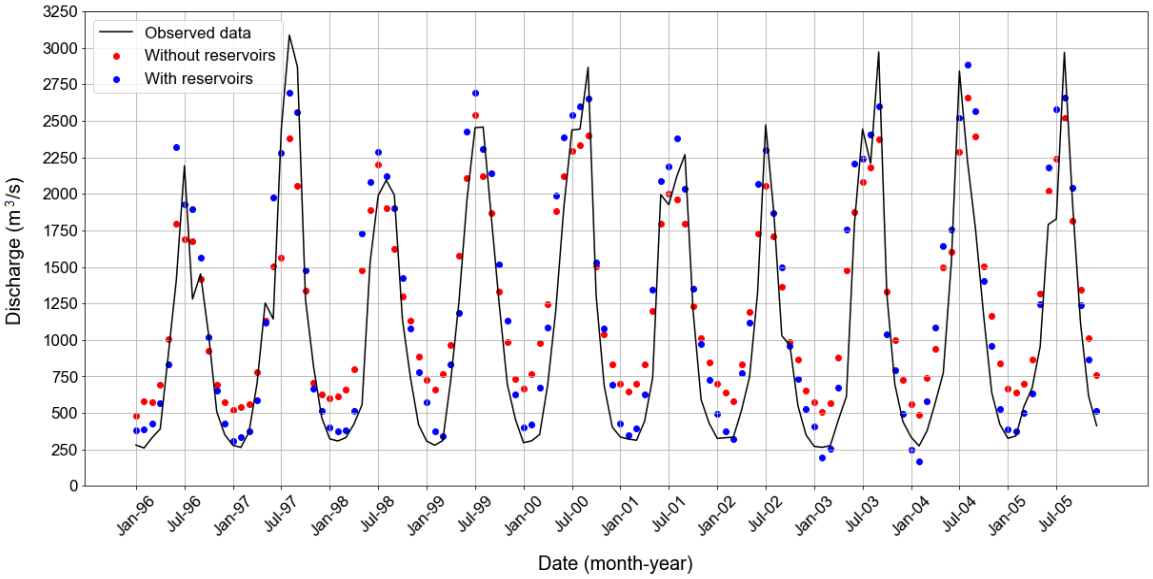

Figure 9. Comparison between observed and simulated monthly discharges at Jiuzhou station over the period 1996–2005. Simulated data are produced by the two selected models with and without reservoirs (blue and red dots, respectively).

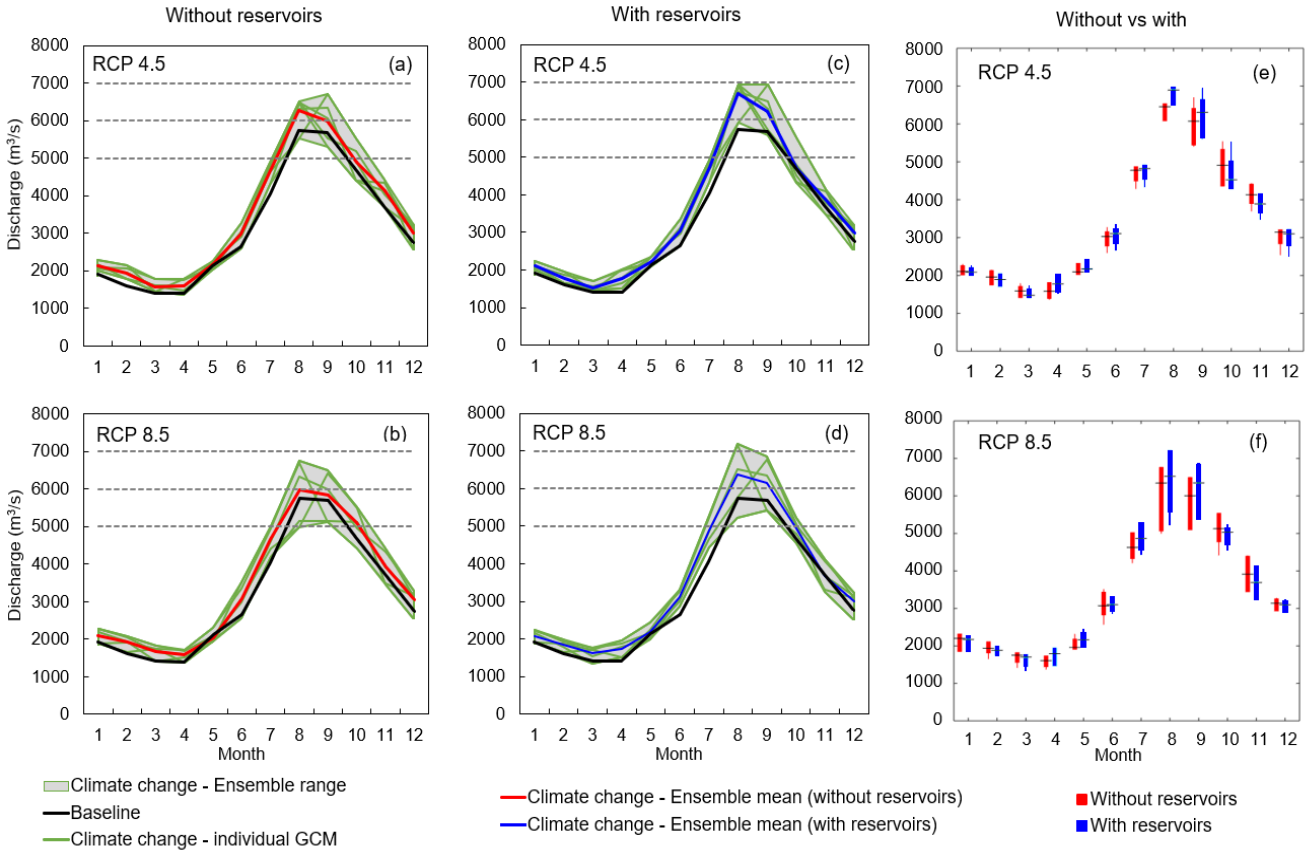

Figure 10. Projected monthly discharges at Chiang Saen under five GCMs and two RCPs for the two selected models calibrated without and with reservoirs (a-d). Box plots highlighting the variability in monthly discharges predicted by the two models under RCP 4.5 (e) and RCP 8.5 (f).

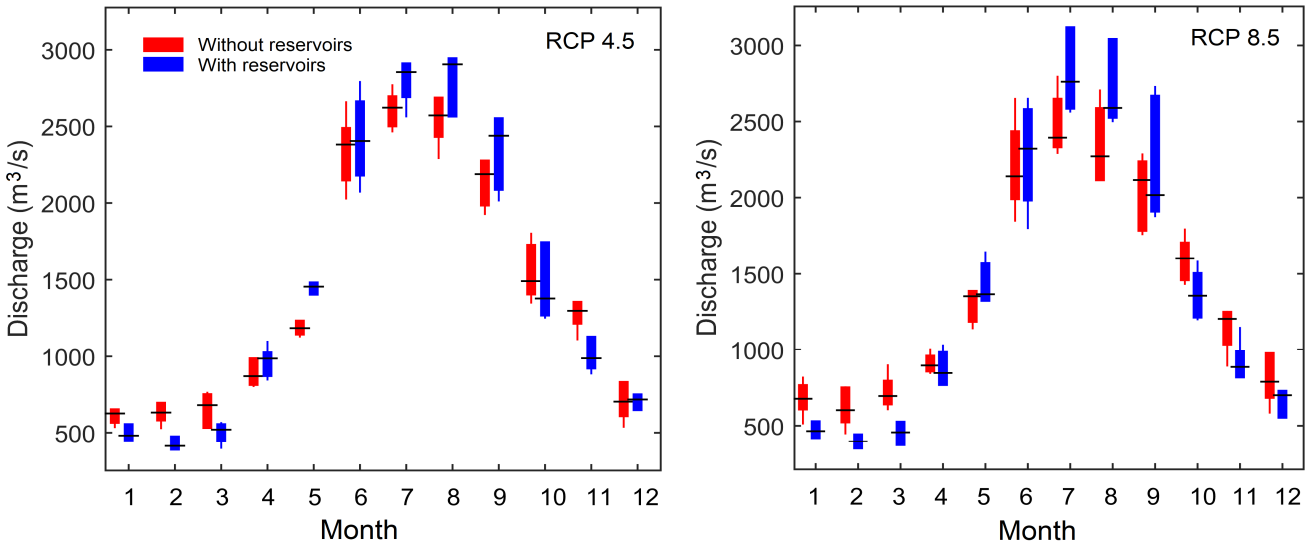

Figure 11. Variability in monthly discharges at Jiuzhou station predicted by the two selected models (with and without reservoirs) under RCP 4.5 (left) and RCP 8.5 (right).

Table 1. Main parameters controlling the rainfall-runoff process in VIC. The third column contains the range of each parameter value considered during the calibration process. Note that these are the same ranges typically adopted for the implementation of VIC to large basins (cfr., Dan et al. (2012); Xue et al. (2015); Wi et al. (2017)).

Name	Unit	Feasible range	Description
d_1	m	[0.05, 0.25]	Thickness of the upper soil layer
d_2	m	[0.3, 1.5]	Thickness of the lower soil layer
b	-	(0, 0.9]	Variable Infiltration Capacity curve parameter
D_{max}	mm/day	(0, 30]	Maximum baseflow
D_s	-	(0, 1)	Fraction of D_{max} where non-linear baseflow begins
W_s	-	(0, 1)	Fraction of maximum soil moisture where non-linear baseflow occurs

Table 2. Design specifications of the dams implemented in our VIC model (simulation period 1995–2005). The term Year denotes the time at which each reservoir became operational.

No.	Name	Year	Long. (°E)	Lat. (°N)	Height (m)	Storage (Mm^3)	Design discharge (m^3/s)	Inst. cap. (MW)
1	Xi'er He 4	1971	100.066	20.000	20	14	283	50
2	Xi'er He 1	1989	100.202	30.000	30	1,501	60	105
3	Xi'er He 2	1987	100.131	25.562	37.25	0.2	168	50
4	Xi'er He 3	1988	100.108	20.700	20.70	0.09	304	50
5	Manwan	1992	100.446	24.625	136	257	1,700	1,670
6	Longdi	1997	99.724	26.221	95	13.30	12.34	10
7	Laoyinyan	1997	99.818	24.469	4.31	10.92	9.3	16
8	XunCun	1999	99.993	25.422	67	73.74	146	78
9	Jinfeng	1998	101.225	21.592	45	19.48	45	16
10	Dachaoshan	2003	100.370	24.025	115	367	2,109	1,350
11	Jinhe	2004	97.333	34.000	34	4.27	222	60

Table 3. CMIP5 GCMs used for the climate change impact assessment.

GCM	Spatial resolution (long × lat)	Control baseline	Developer
ACCESS1-0	$1.875^\circ \times 1.25^\circ$	1850–2006	Commonwealth Scientific and Industrial Research Organization, Australia
CCSM4	$1.25^\circ \times 0.94^0$	1850–2005	National Center for Atmospheric Research, USA
CSIRO Mk3.6	$1.875^\circ \times 1.875^\circ$	1850–2005	Commonwealth Scientific and Industrial Research Organization and the Queensland Climate Change Centre of Excellence, Australia
HadGEM2 ES	$1.875^\circ \times 1.24^\circ$	1861–2010	Met Office, UK
MPI-ESM-LR	$1.875^\circ \times 1.875^\circ$	1850–2005	Max Planck Institute for Meteorology, Germany

Table 4. Relative changes in annual river discharges at the Chiang Saen station for the future period (2050–2060) relative to the reference one (1996–2005). The lowest and highest changes are presented with the corresponding scenarios. The results reported in the first and second rows were produced by the selected models without and with reservoirs.

Scenario	RCP 4.5		RCP 8.5	
	Ensemble mean (%)	Range (%)	Ensemble mean (%)	Range (%)
Without reservoirs	+13.62	+6.36 to +23.66 CSIRO–ACCESS	+13.92	-0.67 to +28.89 CSIRO–ACCESS
With reservoirs	+13.56	+6.28 to +23.56 CSIRO–ACCESS	+13.83	-0.63 to +28.68 CSIRO–ACCESS