Answer to RC #2

Page 1

Line 10-12: It may be rewritten, although the sentence is understood, I think it is not the best way. Try to avoid double parentheses. Besides, the manuscript is focused on VMM, so this model should be first. E.g.: 'The ⋯ (VMM) is compared with three models, one physical based model, the MIKE SHE, and two conceptual models, the XAJ and the Shanbei' (this is just a suggestion, feel free to write your way).

Answer: We accept the reviewer's suggestion.

Change: The vertically mixed runoff model (VMM) is compared with three models, one physical-based model, MIKE SHE, and two conceptual models, the Xinanjiang model (XAJ) and Shanbei model (SBM).

Line 16: "its figure is only 41% in four semiarid catchments", could you explain better this sentence?

Answer: We apologize for our unclear expression, which has been revised.

Change: Our results show that the VMM has a better flood estimation performance than the other models, and under the multicriteria assessment framework, the average acceptance of flood events is improved. In addition, the framework can provide reasonable flood early warning information for decision-makers.

Line 24: "and", but

Answer: This change has been made.

Change: but .

Page 2

Line 9-10: Maybe it is better "to be used on a large-scale semiarid area".

Answer: Thank you for this good suggestion. We have changed this phrasing.

Change: The radar costs are too high to be used on a large-scale semiarid area.

Line 14: "severe", what is the meaning of severe here? It should be better explained.

Answer: "Severe rainstorms" means the rainstorms have produced the largest peak flows according to Michaud and Sorooshian (1994). We have added more explanation.

Change: Michaud and Sorooshian (1994) used 24 severe rainstorms that produced the largest peak flows during 1957-1977 to compare three hydrologic models, i.e., the lumped SCS model, simple distributed SCS model and distributed KINEROS model, in the Walnut Gulch catchment.

Line 20-21: "Therefore, it is urgent to search for useful information based on the limited accuracy of modeling results to serve as flood warnings and to improve decision making.". I do not understand this sentence. Please, rewrite the sentence.

Answer: We apologize for the unclear phrasing. We have rewritten the sentence.

Change: Therefore, determining how to use modeling results with limited accuracy to provide guidance for flood early warning is important.

Line 24-25: “Four hydrologic models (the vertically mixed runoff model (VMM), MIKE SHE, Xinanjiang Model (XAJ) and Shanbei model (SBM)) ...”. To much parenthesis; Try to make a continuous sentence: "Four hydrological models: the vertical ⋯, the MIKE SHE model, ⋯, are compared ...

Answer: Thank you for this suggestion. We have rewritten the sentence.

Change: Four hydrological models: the vertically mixed runoff model (VMM), the MIKE SHE model, Xinanjiang model (XAJ) and Shanbei model (SBM), are compared on the….

Line 27: “remainder”. Please, use another word.

Answer: We have rewritten the sentence.

Change: The rest of the paper is organized as follows.

Line 27: “section below”. Specify to which section corresponds exactly, so it is necessary to use capital letters: e.g. "The Methodology Section", or use instead of the name, just the number of the section "Section 3 describes⋯".

Answer: This is a good suggestion. We have changed this phrasing.

Change: Section 2 describes the study area and the data set used.

Line 28-31: “The VMM model...of the study.”. Please rewrite the whole paragraph, it is difficult to follow it.

Answer: We have rewritten the paragraph to clarify our intent.

Change: Section 3 presents the VMM model methodology, initial condition set, model calibration and validation, multicriteria assessment framework and parameter sensitivity analysis. Section 4 describes the results and discussion of model comparison, sensitivity analysis and analysis of the multicriteria assessment framework for the VMM model. The final section presents the conclusions of the study.

Page 3

Line 1: Is there any reference for all these data? I mean the temperature, rainfall and seasonality of the rainfall data.

Answer: The temperature data are obtained from previous papers, and we have added references. In addition, all rainfall data are collected from rain gauging stations, and we have added an explanation.

Change: This change has been made.

Line 3: “is”, between.

Answer: This change has been made.

Change: between.

Line 4:” is”, between.

Answer: This change has been made.

Change: between.

Line4: “65 – 80%”, more less the 65 to 80%.
Answer: This change has been made.
Change: more less the 65 to 80%.

Line 5: “is”, between.
Answer: This change has been made.
Change: between.

Line 8: “soil erosion”, how much?
Answer: Thank you for this good question. According to Li et al. (2019), the average sediment concentration reaches 126 kg m^{-3} in these regions. We have added this information to the text.
Change: The lack of vegetation in these catchments leads to serious soil erosion, and the average sediment concentration reaches 126 kg m^{-3} according to Li et al. (2019).

Line 9: “few attempts have been applied to model hourly flood flows”. Please rewrite the sentence. It has poor writing.
Answer: We apologize for the poor phrasing. We have rewritten this sentence.
Change: Only a few studies have modeled hourly floods.

Line 11: “Hence, modeling floods and providing a useful method for decision makers in charge of flood defense are essential and urgent.”. Please rewrite the whole sentence, I cannot follow it.
Answer: We apologize for the poor phrasing. We have rewritten this sentence.
Change: Hence, it is important for decision-makers to know how to evaluate the flood risk when a flood is approaching.

Line 12: “Streamflow and rainfall data are from 1983 to 2009.” Please rewrite the sentence. It has poor writing.
Answer: We apologize for the poor phrasing. We have rewritten this sentence.
Change: The period used in the modeling is from 1983 to 2009. Streamflow and rainfall data are collected from streamflow gauging stations and rain gauging stations at an hourly time-step.

Line 12: “stations”, which type of stations?
Answer: They are streamflow gauging stations. We have corrected this phrasing.
Change: Streamflow and rainfall data are collected from streamflow gauging stations and rain gauging stations at an hourly time-step, respectively.

Line 14-15: Thiessen polygon method was...
Answer: This change has been made.
Change: Thiessen polygon method was used to interpolate the rainfall data for each catchment.

Line 15: “...interpolate the rainfall data.”. For each basin? I guess.
Answer: We agree with the reviewer. We have added this information.
Change: Thiessen polygon method was used to interpolate the rainfall data for each catchment.

Line 17: Here you should characterize the model in its whole, I mean, explain here also that it is lumped, continuous or event based, and so on.
Answer: Thank you for this suggestion. We have added this information.
Change: The VMM is a lumped continuous hydrologic model developed by Bao and Wang (1997).

Line 18: "...conceptual hydrologic model...", is continuous or event based?
Answer: It a continuous model; we have added this information.
Change: The VMM is a lumped continuous hydrologic model and has been used in many areas in China, especially in semiarid and subhumid catchments.

Line 21: "etc", better "and others".
Answer: Thank you for raising this good point. We have corrected this phrasing.
Change: This change has been made.

Page 4
Line 24: "in", at.
Answer: This change has been made.
Change: at.

Line 25-27: "Hence, tension water storage *W*... on daily values of *W* and *S*". Please, explain it better, because I do not understand what type of variables you used in the case you just simulate a flood event.
Answer: We are apologized for our unclear phrasing. We have rewritten this paragraph.
Change: The initial condition has important effects in modeling flood events. The VMM model was run continuously from 1983 to 2009 for each catchment. Two initial values are the initial tension water storage (*W0*) and the initial free water storage (*S0*) should be determined. Both of them represent the moisture content of the soil and were assumed to be zero due to the dry conditions at 00:00:00 on January 1, 1983. Rainfall data were available only at an hourly time step over the periods of flood events, and for other periods, they were available at a daily time step. Hence, the time step of simulation was daily between flood events and hourly within flood events.

Page 5
Line 1: "Because...". The sentence has poor writing. Maybe you it is better to write "Due to the fact that only ⋯ ". But I recommend rewriting the sentence.
Answer: Thank you for this good suggestion. We have rewritten the sentence.
Change: Due to the fact that only one streamflow gauging station is available for each catchment, the spatial variation in each catchment's parameters cannot be determined by calibration.

Line 1: "gauge", gauging.
Answer: This change has been made.
Change: gauging .

Line 4: "Li et al., 2018", 2018a or 2018b?
Answer: We have checked the reference.

Change: Li et al., 2018.

Line 8: ”...full fitness...”. I guess you mean "fullfillnes", but I think you should better rewrite the whole sentence in order to get a continuity in the writing.
Answer: Thank you for raising this good point. We have rewritten this sentence.
Change: however, it may not be suitable for semiarid catchments because a good fit is not required between the simulated and observed streamflows.

Line 9: “(McIntyre and Al-Qurashi, 2009: SHARMA and MURTHY, 1998)”. Remove parenthesis, write it properly. "McIntyre and Al-Qurashi (2008) and Sharma and Murphy (1998) used ⋯".
Answer: This change has been made.
Change: McIntyre and Al-Qurashi (2009) and Sharma and Murphy (1998) used….

Line 12: “Q_p^i,”. It is difficult to distinguish the apostrophe. Please, try another mark in order to make out the different variables
Answer: Thank you for this suggestion. We have changed this phrasing.
Change: Q_{pm}^i.

Line 13: The same as previous.
Answer: Thank you for this suggestion. We have changed this phrasing.
Change: ${Q^i}_{vm}$.

Line 15: Please, try another mark in order to make out the different variables.
Answer: This change has been made.
Change； Q_{pm}.

Line 16: The same as previous.
Answer: This change has been made.
Change: $Q_{vm.}$

Line 20: “We...”. It is up to you, but you should try to get coherence throughout the text. Consider using the impersonal form, as you have done in the rest of the manuscript.
Answer: Thank you for raising this good point. We have rewritten the sentence.
Change: The number of iterations was set to 2000 in the calibration process.

Line 21: “...step”. Please, use another Word, as step can be misunderstood by iteration.
Answer: Thank you for raising this good point. We have changed this phrasing.
Change: The number of iterations was set to 2000 in the calibration process.

Line 26: ”Wei-jian et al., 2016”. It is not in the reference list.
Answer: We apologize. We have replaced it with another paper due to the reference being in Chinese.
Change: Cheng, C. T., Zhao, M. Y., Chau, K., and Wu, X. Y.: Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure, J. Hydrol., 316, 129-140, 2006.

Line 26-27: “We test the performance... the middle Yellow River.” This sentence is relevant? I mean, you should restructure the whole paragraph in order to get coherence, instead of specifying one by one each model.
Answer: Thank you for this good suggestion. We have made some changes.
Change: Please refer to the revised version of the manuscript.

Line 27: “Zhao (1983)”. It is not in the reference list.
Answer: We have added this information.
Change: Zhao, R.: Watershed Hydrological Model: Xinanjiang Model and Shanbei Model, Water and Power Press, Beijing, China, 1983.

Line 28: “(Bao et al., 2017)”. Could you provide more references?
Answer: We have added some references.
Change:
Li, Z. J., and Zhang, K.: Comparison of three GIS-based hydrological models, J. Hydrol. Eng., 13, 364-370, 2008.
Zhao, L., Xia, J., Xu, C. Y., Wang, Z., Sobkowiak, L., and Long, C.: Evapotranspiration estimation methods in hydrological models, J. Geogr. Sci., 23, 359-369, 10.1007/s11442-013-1015-9, 2013.

Line 28: “MIKE SHE...”. Could you provide more characteristics of the model?
Answer: We have added more characteristics of the model.
Change: MIKE SHE originated from the Système Hydrologique Européen (SHE) program, and it is a deterministic, physically based distributed hydrologic model that can simulate surface water flow, unsaturated flow and saturated flow (Jayatilaka et al., 1998). MIKE SHE has been used to solve water resources and environment problems at different spatiotemporal scales (Li et al., 2018; Rujner et al.,2018; Samaras et al., 2016).

Page 6
line 3: There are many grammatical errors. Please, revise the whole point.
Answer: Thank you for pointing out these errors. We have corrected them.
Change: Please refer to the revised version of manuscript.

Line 3: “...of...”. "for" instead of "of”?
Answer: This change has been made.
Change: for.

Line 4: “...characterizes...”, characteristics.
Answer: This change has been made.
Change: characteristics.

Line 4: “and lack of enough rain gauges,”, and also dispersion.
Answer: This change has been made.
Change: and also dispersion.

Line 5: “flood simulation”, flood simulations.
Answer: This change has been made.
Change: flood simulations.

Line 6: “...to...”, for.
Answer: This change has been made.
Change: for.

Line 7: “...flood feature...”. I would remove "flood" as It is obvious.
Answer: This change has been made.
Change: This change has been made.

Line 9: “calculation...”. I would remove this.
Answer: This change has been made.
Change: This word has been removed.

Line 11: “...and Bayesian method...”. Which one? could you give some more information?
Answer: We have added this information.
Change: and Bayesian method with Markov chain Monte Carlo sampling.

Line 12: “ but the way may...”. I do not understand that phrase.
Answer: We have improved the sentence.
Change: although these methods may not lead to clear decisions.

Line 13: “...acquire...”. Consider to change the Word.
Answer: This change has been made.
Change: obtain.

Line 13: “...utility...”, useful.
Answer: This change has been made.
Change: useful.

Line 14: “...Yellow Rivers.”. Please be careful. Remove "s".
Answer: This change has been made.
Change: …Yellow River.

Line 16: “modeling...”, modeled.
Answer: This change has been made.
Change: modeled.

Line 16: “peak flows”. Remove "s".
Answer: This change has been made.
Change: peak flow.

Line 21: “modeling...”, modeled.
Answer: This change has been made.
Change: modeled.

Line 22: “(detailed ...)”. I would not use the parenthesis here. Just use the semicolon.
Answer: Thank you for this good suggestion. We have improved the sentence.
Change: one component of the Bayesian forecasting system is detailed in Krzysztofowicz (1999) and Biondi et al. (2010).

Line 23: “(Krzysztofowicz, 1999; Biondi et al., 2010)”. Remove parenthesis and write them in a correct way.
Answer: Thank you for noting these errors. We have corrected this phrasing.
Change: ⋯Krzysztofowicz (1999) and Biondi et al. (2010).
Line 25: “modelling peak flow...”, the modelling peak.
Answer: This change has been made.
Change: the modelling peak.

Page 7
Line 4: “(Pianosi and Wagener, 2015)”. Please write the parenthesis in the correct way. "pianosi and Wagener (2015) proposed ⋯".
Answer: Thank you for pointing out these errors. We have corrected this phrasing.
Change: Pianosi and Wagener (2015).

line 4: “PAWN”, What is the meaning?. Explain the method.
Answer: PAWN is derived from the authors names according to Pianosi and Wagener (2015). We have added more explanation.
Change: Pianosi and Wagener (2015) proposed the novel GSA method PAWN (derived from the authors’ names) based on the cumulative density function.

Line 5-6: “(Khorashadi Zadeh et al., 2017)”. Please, write the reference in the correct way.
Answer: We have corrected this phrasing.
Change: Khorashadi et al. (2017).

Line 21: “5 events...”. "five" instead of 5. Correct the rest of remaining ones throughout the paragraph.
Answer: We have corrected them.
Change: Five.

Line 27: Results Section is good, but you have to rewrite it better, and expand it. Please, try to be more clear, and specify every result you have.
Answer: Thank you for these suggestions. We have rewritten Section 4.1.
Change: Please refer to the revised version of the manuscript.

Page 8

Line 1: "...VMM performs better...". In both calibration and validation? Please, expand your explanation.

Answer: We apologize for our unclear expression. We have rewritten the sentence.

Change: In terms of the median and average of the absolute relative errors for peak flows, except for the validation period in the Kuye River catchment shown in Figure 3 (h), Figures 3 (a)–(g) reveal that the VMM has lowest values for both calibration and validation.

Line 2: "...for both median and average peak flows." I understand what you are saying, but try to work out your explanation.

Answer: We have rewritten the sentence.

Change: In terms of the median and average of the absolute relative errors for peak flows, except for the validation period in the Kuye River catchment shown in Figure 3 (h), Figures 3 (a)–(g) reveal that the VMM has lowest values for both calibration and validation.

Line 11: "Overall...". Remove that. Rewrite the sentence.

Answer: We have rewritten the sentence.

Change: The analysis of Figure 3, Table 3 and Table 4 shows that the VMM has the best performance for flood modeling in the four studied catchments of the middle Yellow River.…

Line 11-12: "in the middle Yellow River". " of the" instead of "in".

Answer: This change has been made.

Change: of the.

Line 12: "MIKE SHE...". There is a lack of connectors in some parts of the text. Here, you should use: "Besides", "In addition", etc.

Answer: This change has been made.

Change: In addition.

Line 21: "Eq.(9) and Eq.(11)". Substitute the Equation by the variable, or write both of them. "Ep (Eq. 9) and Epv (Eq. 11)".

Answer: Thank you for this good suggestion. We have rewritten them.

Change: E_p(Eq. 9) and E_{pv}(Eq. 11).

Line 21-22: "The most sensitive ...objective functions.". Try to explain better what you want to mean. Rewrite the sentence.

Answer: We have rewritten the sentence.

Change: The higher the ranking is, the more sensitive the parameters. We can find that the parameters *CS*, *IM* and *KE* have the highest rankings whether the objective function of the VMM model is E_p or E_{pv}.

Line 29: "...must meet...". There are two "must" in the sentence. Please, use synonym or rewrite the sentence.

Answer: We have rewritten the sentence.
Change: The framework requires that an accepted flood event should meet one of the requirements of C1 and C2; in addition, C3 needs to be satisfied simultaneously.

Line 30: “Flood events conforming to conditions...”. Is that right?
Answer: We apologize for the unclear expression. We have rewritten the sentence.
Change: The observed peak flows and the modeled peak flows under the conditions C1, C2 or C3 are shown in Figure 5.

Page 9
Line 15: “...VMM...”. In the conclusions part, you should rewrite the abbreviations. Please, check the rest of references and abbreviations.
Answer: We have checked them and added an explanation for clarification.
Change: Please refer to the revised version of the manuscript.

Line 19-20: “In the four catchments, by...sensitive parameter.”. The same as previous.
Answer: We have rewritten the sentence.
Change: In the four catchments, the parameters confluence coefficient of surface flow (*CS*), impermeable area (*IM*), and residence time of Muskingum (*KE*) are the most sensitive based on an analysis by the global sensitivity method PAWN; in addition, the sensitivity ranking of the parameter *WM* related with the soil moisture capacity is the most affected by the objective functions.

Line 23: “The condition C2... ”. Explain that. You should bear in mind Conclusions could be read by anyone, so they should not contain references to the rest of the article unless it is completely necessary.
Answer: We have added more explanation.
Change: The condition C2, which divides peak flows into three flow zones, will be affected….

Line 24: “... enrich...”. Please, use another word.
Answer: This change has been made.
Change: The framework… can provide guidance for decision making.

Page 10
Line 5: “Reference:”. There are several errors in the reference list. Please, read the journal's rules, and correct every mistake. I just write down some of them. Besides, order the references according to the standards of the journal.
Answer: We have checked the references and corrected them.
Change: Please refer to the revised version of the manuscript.

Line 6: “Andersen....”. This is a PhD Thesis. Please, read journal's rules.
Answer: This change has been made.
Change: Andersen, F. H.: Hydrological modeling in a semi-arid area using remote sensing data, Ph.D. thesis, University of Copenhagen, Copenhagen, Denmark, 2008.

Line 8-9: “Burnash...”. I found the paper, but there is a lack of information here. Please complete the reference.
Answer: This change has been made.
Change: Burnash, R. J., Ferral, R. L., and McGuire, R. A.: A generalized streamflow simulation system, conceptual modeling for digital computers, Report by the Joliet Federal State River Forecasts Center, Sacramento, CA, 204 pp., 1973.

Line 11-12: “Bao, W. and Wang, C....”. Unless, it is completely necessary, please, try to avoid references that are hard to find out. I could not find those references in Chinese.
Answer: Thank you. We believe that this suggestion is very constructive. We have replaced this reference with another paper. However, some references in Chinese are completely necessary, so we have kept them.
Change: Wang, G., and Ren, L.: A Contrastive Study of Simulation Results between GWSC-VMR and Hybrid Runoff Model in Dianzi Basin, in: International Conference on Environmental Science and Information Application Technology, Wuhan, China, 4 – 5 July, 583-588, 2009.

Line 22: “Beven, K.:...”. It is needed more info about these references.
Answer: This change has been made.
Change: Beven, K. J.: Environmental modelling: An uncertain future?, CRC press, London, UK, 328 pp., 2007.

Line 23: “Beven, K.:...”. It is needed more info about these references.
Answer: This change has been made.
Change: Beven, K. J.: Rainfall-runoff modelling: the primer, John Wiley & Sons, UK, 488 pp., 2011.

Page 11
line 4: “3-23, 2007.” DOI?
Answer: This change has been made.
Change: Collier, C. G.: Flash flood forecasting: What are the limits of predictability?, Q. J. Roy. Meteor. Soc., 133(622), 3 – 23, https://doi.org/10.1002/qj.29, 2007.

Line 31-32: “Li, D: ...”. Rewrite according to the journal's rules.
Answer: This change has been made.
Change: Li, D.: Hydrologic model: the vertically mixed runoff model (vmm), HydroShare, https://doi.org/10.4211/hs.c523228 7d5c04bfb8cac5ce4e391ea 0f, 2018.

Page 12
Line 2: “...Qiushui River, Yellow River, 06, 24-28, 2018a.”. This is a journal?? Please, again try to avoid these references, but also try to make easy to find them in case you want to add them to the list.
Answer: We have deleted it.
Change: The reference has been deleted.

Line 29: “1-32”. Pages range is wrong.
Answer: Thank you for noting these errors. We have corrected this phrasing.
Change: Sharma, K. D., and Murthy, J. S. R.: A practical approach to rainfall-runoff modelling in arid zone drainage basins, Hydrolog. Sci. J., 43(3), 331 – 348, 1998.

Line 32: “United Nations Environment...”. More information about the Book.
Answer: This change has been made.
Change: United Nations Environment Programme (UNEP): World Atlas of Desertification, Edward Arnold, London, 69 pp., 1992.

Page 13
Line 3: “...Yellow River.”. The same as previous.
Answer: This change has been made.
Change: Please refer to the revised manuscript.

Line 8: “Mathematical modelling and computational experiments.”, abbreviation?
Answer: This change has been made.
Change: Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp., 1, 407 – 414, 1993.

Line 13: “44(5), 2008.”, pages range?
Answer: We have added this information.
Change: Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M. and Stewart, A.: Understanding uncertainty in distributed flash flood forecasting for semiarid regions, Water Resour. Res., 44(5), 61 – 74, 2008.

Line 15: “J.GEOPHYS. RES-ATMOS.”. Why in capital letters?
Answer: Thank you for noting these errors. We have corrected this phrasing.
Change: Young, C.B., Nelson, B.R., Bradley, A.A., Smith, J.A., Peters-Lidard, C.D., Kruger, A. and Baeck, M.L.: An evaluation of NEXRAD precipitation estimates in complex terrain, J. Geophys. Res.-Atmos., 104, 19691 – 19703, 1999.

Page 18
Figure 1: What is the meaning of the big R? Please, include other references in the map of the Yellow River Basin, maybe the border of the see or the cities included inside the basin. Also, the rain gaunges stations are not clear. Please, change the mark. Finally, the drainage basin is too gross. Please, try to add some more detail.
Answer: Thank you for these suggestions. The big R may be a display error. It should be the North Arrow. In addition, we have checked all figures in case other errors occurred. We have improved Figure 1 for clarity.
Change:

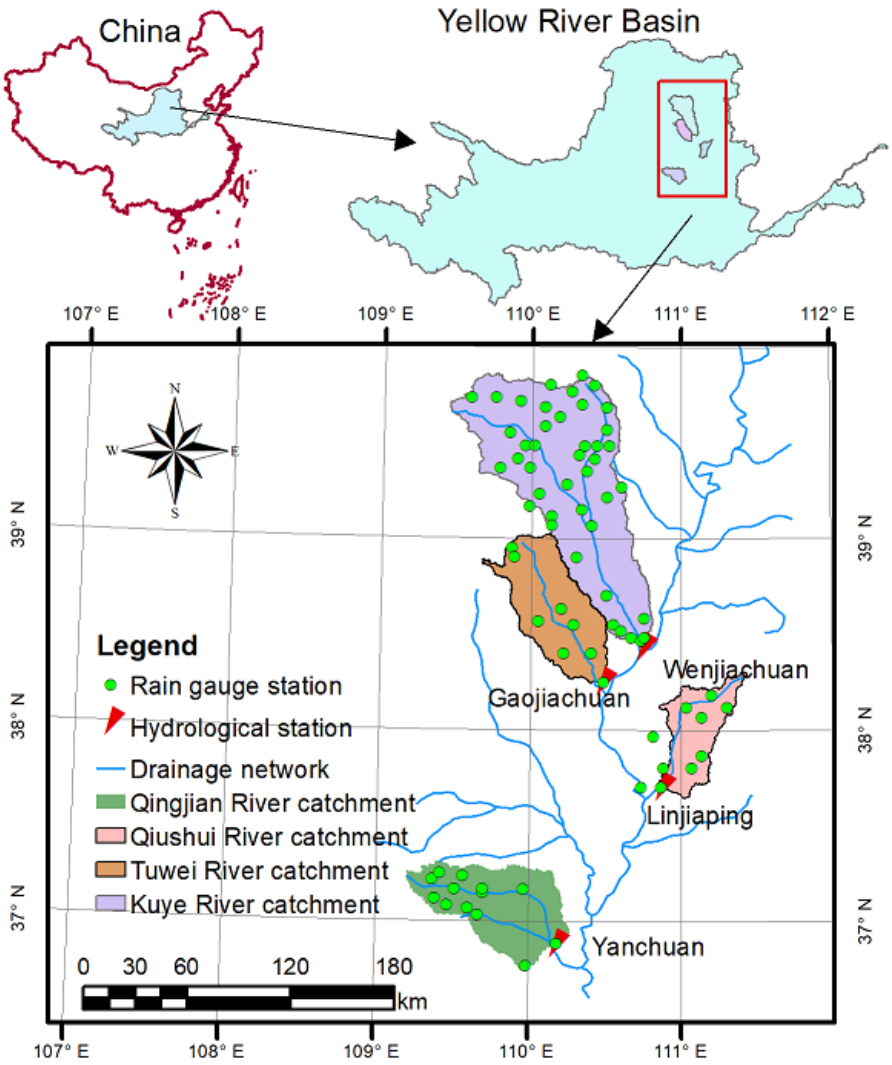


Figure 1: Location of the Qingjian River catchment, Qiushui River catchment, Tuwei River catchment and Kuye River catchment.

Page 20

Figure 3: Please, add letters to identify each graph, in order to be able to talk about them in the manuscript. Also, you should add some space among graphs. Regarding the axes, specify clearly what is the variable they are measuring, I guess they talk about errors.

Answer: Thank you for these suggestions. We have improved Figure 3. In addition, the Y axis label should be “Absolute relative errors of peak flows (%)”. We have corrected this phrasing.

Change:

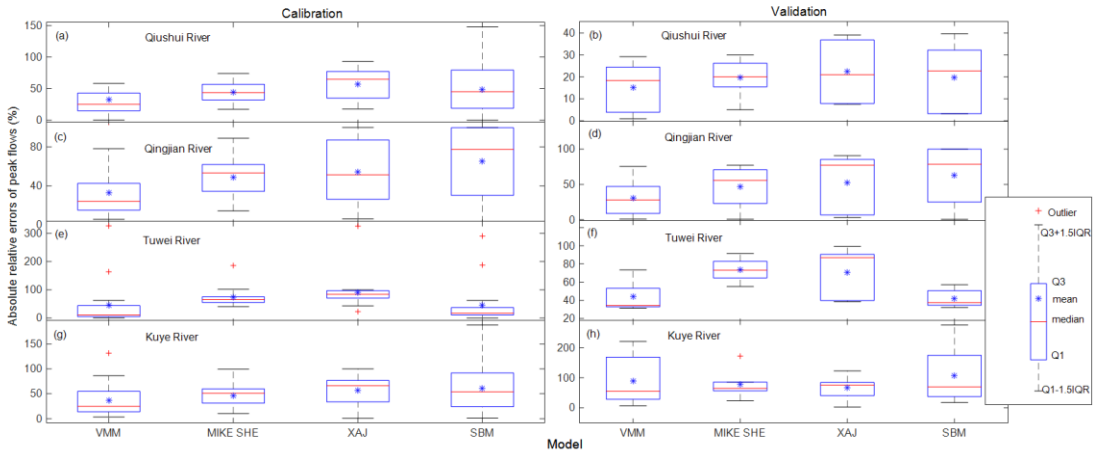


Figure 3: Boxplots of the absolute relative errors of the peak flows in the four catchments; Q1 and Q3 mean the first quantile and third quantile, respectively; interquartile range (IQR)$= \mathrm{Q3} - \mathrm{Q1}$; and an outlier is defined as an extreme value that exceeds the IQR.

Page 21

Figure 4: Please, rewrite the whole figure caption. It is very difficult to understand what you try to say.

Answer: We apologize for the lack of clarity. We have rewritten the figure caption.

Change: Figure 4: Sensitivity rankings of the VMM parameters based on the global sensitivity analysis method PAWN for different objective functions: (a) E_{pv} as the objective function, and (b) E_p as the objective function. The value P is used to assess the sensitivity degree of the parameter with PAWN method, and a larger value corresponds to greater sensitivity. The numbers on the ordinate represent the sensitivity rankings.

Page 22

figure 5: Please, add letters at each graph, in order to be able to talk about them in the text. In addition, clarify the legend, it is difficult to identify each variable. Regarding the caption, please give more info, more detail about the graph. Explain more about we are watching.

Answer: Thank you for this good suggestion. We have added letters at each graph and additional explanations in the figure caption. In addition, additional explanations have been added to Section 4.3.

Change:

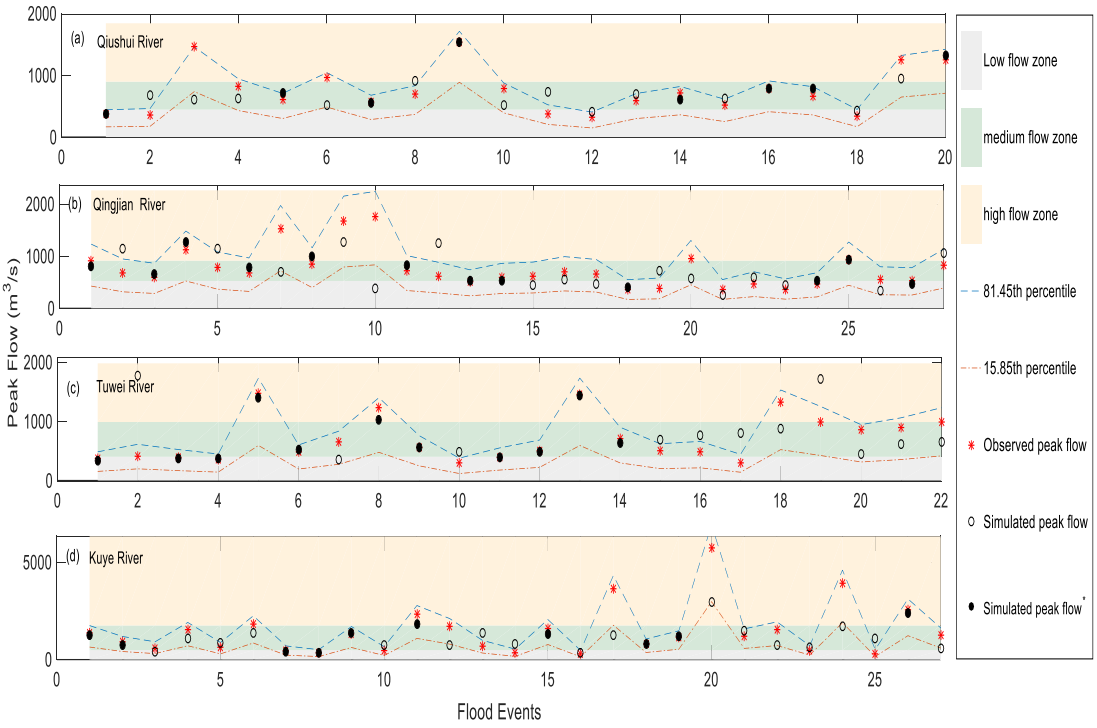


Figure 5: Observed peak flows (red asterisk) and simulated peak flows (solid ball and circle) with the VMM for each catchment under the conditions C1, C2 and C3. Flood peaks conforming to the condition C1 is represented by a solid ball; the three flow zones (low, medium and high flow zones) divided by the condition C2 are shown in gray, green and off-white, respectively; 68.3% confidence interval of peak flows estimated by the condition C3 is between the blue dashed line (81.45th percentile) and the red dash-dotted line (15.85th percentile).

Multicriteria assessment framework of flood events simulated with the vertically mixed runoff model in semiarid catchments in the middle Yellow River

Dayang Li, Zhongmin Liang, Yan Zhou, Binquan Li, Yupeng Fu

College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

Correspondence to: Binquan Li (libinquan@hhu.edu.cn)

Abstract. Flood forecasting and simulation in semiarid regions are always poor, and a single criterion assessment provides limited information for decision making. Here, we propose a multicriteria assessment framework that combines the absolute relative error, flow partitioning and confidence interval estimated by the Hydrologic Uncertainty Processor (HUP) to assess the most striking feature of an event-based flood: the peak flow. The vertically mixed runoff model (VMM) is compared with three models, one physical-based model, the MIKE SHE and two conceptual models, the Xinanjiang model (XAJ) and the Shanbei model (SBM). The 100 flood events in the four catchments of the Yellow River are modeled over the period of 1983–2009. Our results show that the VMM has a better flood estimation performance than the other models, and under the multicriteria assessment framework, the average acceptance of flood events is improved. In addition, the framework can provide reasonable flood early warning information for decision-makers.

1 Introduction

Arid and semiarid regions account for approximately one-third of the global land surface and half of China. A trend towards a warmer climate has increased the global incidence of intense precipitation events. Arid and semiarid regions, i.e., areas where the annual rain is less than 250 and 250–500 mm/a, respectively, are particularly vulnerable to this change in climate (Khomsi et al., 2016; Yatheendradas et al., 2008). More than 50% of flood-related casualties occur in these regions worldwide (Brito and Evers, 2016).

Hydrological models play an important role in flood simulation and forecasting (Devia et al., 2015). Many studies have focused on the improvement and estimation of hydrologic models in humid catchments, although similar work for semiarid catchments is relatively few (Jiang et al., 2015). The runoff generation mechanism of semiarid catchments is complex and may be simultaneously dominated by infiltration excess and saturation excess mechanisms (Beven, 1983; Beven and Freer, 2001).

Modeling semiarid catchments is a difficult task due to the strong spatial variability in rainfall and complexity of

Deleted: on

Formatted: Font: 10 pt, Not Bold

Deleted: ¶

Deleted: on

Deleted: combining

Deleted: the

Deleted: the

Deleted: –

Deleted: –

Deleted:

Deleted: modell

Deleted: -

Deleted: The physically based model MIKE SHE and three conceptual models (two models with a single runoff generation mechanism, the Xianjiang model (XAJ) and the Shanbei model (SBM), and one model with the mixed runoff generation mechanism, the vertically mixed runoff model (VMM)) are compared in terms of flood modeling performance in four semiarid catchments (Qiushui River, Qingjian River, Tuwei River and Kuye River) in the middle Yellow River.

Deleted: on

Deleted: with a single criterion(), and the isthe average acceptance of flood events accounts for 58%, but when absolute relative error 20% is used as the performance criterion, its figure is only 41% in four semiarid catchments.

Deleted: s

Deleted: The

Deleted: h

Deleted: s

Deleted: on the

Deleted: but

Deleted: and

Deleted: simultaneously

landscape characteristics (vegetation, soil, etc.) (Pilgrim et al., 1988). Compared with humid catchments, the rainfall of semiarid catchments is characterized by a high intensity and short duration (Andersen, 2008). In certain areas with developing economies and small populations, the network of rain gauges is generally sparse. Rainfall data are important inputs for hydrologic models, and the high temporal-spatial rainfall variability combined with sparse rain gauges makes modeling runoff more difficult (Hao et al., 2018; Li and Huang, 2017; Mwakalila et al., 2001).

Satellite technology has the possibility to solve the issue of low rain gauge densities, although the low spatial and temporal resolutions of the products limit their applicability to subdaily rainstorms (Dinku et al., 2007). Weather radar has high spatial resolution (1 km) and temporal resolution (15 min). However, the radar costs are too high to be used for large scale semiarid areas (Young et al., 1999).

Literature on the subdaily modeling of rainfall runoff is limited in semiarid catchments. Due to quick times-to-peak and scarce rainfall data, capturing rainstorm flood responses is more difficult than estimating daily, monthly or annual runoff (Andersen, 2008; McMichael et al., 2006). Flood simulation results in semiarid catchments are often poor. Michaud and Sorooshian (1994) used 24 severe rainstorms that produced the largest peak flows from 1957–1977 to compare three hydrologic models, i.e., the lumped SCS model, simple distributed SCS model, and distributed KINEROS model, in the Walnut Gulch catchment, and none of them were able to accurately simulate flood events. McIntyre and Al-Qurashi (2009) analyzed 27 flood events with three hydrologic models (the lumped IHACRES model, distributed IHACRES model, and a 2-parameter regression model) in a catchment in Oman. The average absolute relative errors in the flow peak and flow volume were 53% and 36%, respectively, for the best performing models. Under current technical conditions, it seems difficult to achieve an acceptable simulation/forecasting result for flood events in semiarid catchments. Therefore, determining how to use modeling results with limited accuracy to provide guidance for flood early warning is important.

In this study, a multicriteria assessment framework that combines the absolute relative error, flow zone partitioning and the confidence interval estimated by Hydrologic Uncertainty Processor (HUP) is proposed to provide information for engineers' decision making. Four hydrological models: the vertically mixed runoff model (VMM), the MIKE SHE model, Xinanjiang model (XAJ) and Shanbei model (SBM), are compared based on the performance of the modeling results in four catchments in the middle Yellow River. The global sensitive analysis (GSA) method PAWN is used to analyze the parametric sensitivity of the VMM. The rest of the paper is organized as follows. Section 2 describes the study area and the data set used. Section 3 presents the VMM model methodology, initial condition set, model calibration and validation, multicriteria assessment framework and parameter sensitivity analysis. Section 4 describes the results and discussion of model comparison, sensitivity analysis and analysis of the multicriteria assessment framework for the VMM model. The final section presents the conclusions of the study.

Deleted: but
Deleted: on a
Deleted:
Deleted: .
Deleted: in semiarid areas
Deleted: -
Deleted: which
Deleted: during
Deleted: (
Deleted:)
Deleted: Therefore,
Deleted: the
Deleted: it is urgent to search for useful information based on the limited accuracy of modeling results to serve as flood warnings and to improve decision making.
Deleted: on
Deleted: ing
Deleted: partitioning
Deleted: s
Deleted: ,
Deleted: Four hydrologic models (the vertically mixed runoff model (VMM), MIKE SHE, Xinanjiang model (XAJ) and Shanbei ...
Deleted: on the basis of
Deleted: of VM
Deleted: restremainder
Deleted: mainder
Deleted: The
Deleted: s
Deleted: below
Deleted: presents
Deleted: a description of
Deleted: methodology of VM
Deleted: the
Deleted: multicriterion
Deleted: multicriterion
Deleted: The VMM model, model calibration, initial conditions of ...

2 Study area and data

The 4 selected study catchments are all key tributaries located in the middle Yellow River, China (Fig. 1). The maximum and minimum areas of catchments are 1989 km^2 and 8706 km^2, respectively. The average annual temperature ranges from 6–14°C. The average annual precipitation ranges from 1010–1150 mm, and 65 to 80% is concentrated in summer (Li et al., 2019; Li and Huang, 2017). The rainfall is generally characterized by high intensity and short duration. The average annual evaporation ranges from 1010–1150 mm. All selected catchments are semiarid due to an aridity index between 2.31 and 2.78 (UNEP, 1992). This catchment information is listed in Table 1.

The lack of vegetation in these catchments leads to serious soil erosion, and the average sediment concentration reaches 126 kg m^{-3} according to Li et al. (2019). Some hydrologists have studied daily and monthly rainfall runoff, although few studies have modeled hourly floods. With the rapid increase in population and economic development, flood disasters have received increasing attention. Hence, it is important for decision-makers to know how to evaluate the flood risk when a flood is approaching.

The period used in the modeling is from 1983 to 2009. Continuous streamflow and rainfall data are collected from streamflow gauging stations and rain gauging stations at a daily time step, respectively; streamflow and rainfall data for each of the flood events are collected at an hourly time step. Nine rainfall gauging stations in the Qiushui River catchment, 15 rainfall gauging stations in the Qingjian River catchment, 12 rainfall gauging stations in the Tuwei River and 41 rainfall gauging stations in the Kuye River were selected. Thiessen polygon method was used to interpolate the rainfall data for each catchment.

3 Methodology

3.1 Vertically mixed runoff model

The VMM is a lumped, continuous hydrologic model and has been used in many areas in China, especially in semiarid and subhumid catchments (Bao and Zhao, 2014; Li, 2018; Wang and Ren, 2009). Compared with other conceptual models, such as the XAJ model (Zhao, 1992), Sacramento Soil Moisture Accounting Model (SSMA) (Burnash et al. 1973), among others, the VMM is able to simulate the saturation excess and infiltration excess runoff generation mechanisms simultaneously. As shown in Fig. 2, the VMM combines the infiltration capacity curve and tension water content storage capacity curve in the vertical direction. Net rainfall (observed rainfall after removal of evaporation, *PE*) is partitioned into surface runoff (*RS*) and infiltration flow (*FA*) by the infiltration capacity curve in the VMM. *FA* is regulated by the tension water storage capacity curve, part of which supplements the tension water storage (*W*), with the rest forming the groundwater flow (*RB*) (including unsaturated flow and saturated flow). Here, the calculation of runoff generation is described briefly. More detailed information about the VMM is contained in Bao and Zhao (2014).

Deleted: Fig.

Deleted: ure

Deleted: is betweenis…anges from 6–14°C. The average annual precipitation ranges from is between …010–1150 mm, of which more less the…

Deleted: –…0% is concentrated in summer .

Deleted: is between …010–1150 mm. All selected catchments are semiarid due to an aridity index between 2.31 and 2.78 (UNEP, 1992). Thisese…

Deleted: More

Deleted: information about the …atchment informations

Formatted: Superscript

Deleted: The study catchments have poor vegetation coverage and serious soil erosion.

Deleted: but

Deleted: and …few studies have modeled hourly floods. few attempts have been applied to model hourly flood flows (Cheng et al., 2009; Wang et al., 2006).

Deleted:

Deleted: Hence, modeling floods and providing a useful method for decision makers in charge of flood defense are essential and urgent.…

Deleted: -step… respectively; streamflow and rainfall data for each of the flood events are collected at an hourly time step-step

Deleted: Streamflow and rainfall data are from 1983 to 2009. Hourly streamflow data came from hydrological stations. …ine rainfall gauginge…stations in the Qiushui River catchment, 1515…rainfall gauginge…stations in the Qingjian River catchment, 12 …2 rainfall gauginge…stations in the Tuwei River and 41 41 …ainfall gauginge…stations in the Kuye River were selected.

Deleted: conceptual hydrologic model…developed by Bao and

Deleted: Li et al., 2018a; Wen and Cai, 2015

Formatted

Deleted: and others

Deleted: etc

Deleted: …. the VMM is able to simulate the saturation excess and

Deleted: ure

Deleted: and…the rest of which…formings

Deleted: below-ground runoff…roundwater flow (*RB*) (including

Deleted: For more…ore detailed information about the VMM, you

The improved Green-Ampt infiltration curve (Bao, 1993) is applied in the VMM as the infiltration capacity curve, and the equation is as follows:

$$FM = FC\left(1 + K\frac{WM - W}{WM}\right) \quad (1)$$

where FM is the average point infiltration capacity of the catchment, and the descriptions of WM, K, and FC are shown in Table 2.

FA is calculated by Eq. (2):

$$FA = \begin{cases} FM - FM\left(1 - \frac{PE}{(FMM)^{1+BF}}\right) & PE < FMM \\ FM & PE \geq FMM \end{cases} \quad (2)$$

where

$$FMM = FM(1 + BF) \quad (3)$$

in which FMM is the maximum point infiltration capacity of the catchment and BF is shown in Table 2.

The part that exceeds the average point infiltration capacity of the catchment FM forms RS. RS can be calculated by Eq. (4).

$$RS = PE - FA \quad (4)$$

RB can be calculated by Eq. (5):

$$RB = \begin{cases} FA - WM + W + WM\left(1 - \frac{W^* + FA^{B+1}}{WMM}\right) & FA + W^* < WMM \\ FM - WM + W & FA + W^* \geq WMM \end{cases} \quad (5)$$

where

$$W^* = WMM\left[1 - \left(1 - \frac{W}{WM}\right)\right]^{\left(\frac{1}{B}+1\right)} \quad (6)$$

$$WMM = WM(1 + B) \quad (7)$$

in which WMM is the maximum point tension water storage capacity of the catchment, W^* is the ordinate of Fig. 2 (b), which represents the point tension water content capacity in the catchment, and B is shown in Table 2.

The outlet runoff R can be calculated as follows:

$$R = RS + RB \quad (8)$$

3.2 Initial condition of the VMM

The initial condition has important effects in modeling flood events. The VMM model was run continuously from 1983 to 2009 for each catchment. Two initial values are the initial tension water storage ($W0$) and the initial free water storage ($S0$) should be determined. Both of them represent the moisture content of the soil and were assumed to be zero due to the dry conditions at 00:00:00 on January 1, 1983. Rainfall data were available only at an hourly time step over the periods of flood

Deleted: .
Deleted: .
Deleted: Fig.
Deleted: ure
Deleted: .
Deleted: event-based
Formatted: Font: Times New Roman, 10 pt
Formatted: Normal, Space Before: 0 pt, After: 0 pt, Line spacing: Double
Formatted: Font: Times New Roman, 10 pt, Italic
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt, Italic
Formatted: Font: Times New Roman, 10 pt
Deleted: only
Formatted: Font: Times New Roman, 10 pt
Deleted: Due to the fact that the
Deleted: r
Formatted: Font: Times New Roman, 10 pt
Deleted: as
Formatted: Font: Times New Roman, 10 pt
Deleted: only
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Deleted: -step

events, and for other periods, they were available at a daily time step. Hence, the time step of simulation was daily between flood events and hourly within flood events.

3.3 Model calibration

To consider the spatial variation in rainfall, the subcatchments are divided and the VMM model is applied to each subcatchment. Due to the fact that only one streamflow gauging station is available for each catchment, the spatial variation in each catchment's parameters cannot be determined by calibration. Thus, the parameters are set uniformly in all subcatchments. The fourteen parameters (Table 2) of the VMM are calibrated by the global optimization algorithm SCE-UA (Duan et al., 1993). The ranges of parameters are determined based on previous literature and prior knowledge (Bao and Zhao, 2014; Li et al., 2018).

In semiarid catchments, due to the rapid rise and fall of floods (usually less than 24 h), accurate simulations of the full hydrograph are not needed and cannot be realized. The Nash-Sutcliffe efficiency (NSE; (Nash and Sutcliffe, 1970) is widely used as an objective function of calibration in humid catchments; however, it may not be suitable for semiarid catchments because a good fit is not required between the simulated and observed streamflows. McIntyre and Al-Qurashi (2009) and Sharma and Murphy (1998) used the absolute relative error to evaluate model outputs (flow peak and flow volume) for semiarid areas, and the calibrated results indicated that the flow peak results are more accurate than suggested based on the NSE. Thus, the simulated hydrograph is reasonable for the majority of flood events. The equations are as follows:

$$E_p = \frac{1}{n}\sum_{i=1}^{n}\frac{|Q_p^i - Q_{pm}^i|}{Q_{pm}^i} \quad (9)$$

$$E_v = \frac{1}{n}\sum_{i=1}^{n}\frac{|Q_v^i - Q_{vm}^i|}{Q_{vm}^i} \quad (10)$$

where E_p and E_v are the average performances (in terms of absolute relative error) for peak flows and flow volumes in each catchment, respectively; n is the number of events; the index i denotes each event; Q_p and Q_{pm} are the simulated and measured values of peak flow per event, respectively; and Q_v and Q_{vm} are the simulated and measured values of flow volume per event, respectively.

Constraining the model output with peak flows and flow volumes can be expressed as follows:

$$E_{pv} = \frac{E_p + E_v}{2} \quad (11)$$

where E_{pv} is the objective value. The closer E_{pv} is to 0, the better the model outputs. The number of iterations was set to 2000 in the calibration process.

Deleted: otherwise they were
Deleted: -step
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Formatted: Font: Times New Roman, 10 pt
Deleted: The initial condition has important effects in modeling flood events. Therefore, initial tension water storage (*W0*) and initial free water storage (*S0*) should be determined at the beginning of each flood event calculation. We use daily rainfall data over the period of 1983–2009 to simulate the daily streamflow with daily based VMM in each catchment's outlet. The simulation results are accepted when achieving water balance compared with observed streamflow. Hence, tension water storage *W* and free water storage *S* per day during 1983–2009 can be achieved, and the initial conditions (*W0* and *S0*) of each flood event are determined based on daily values of *W* and *S*.¶
Deleted: ,
Deleted: lumped model
Deleted: Due to the fact tha
Deleted: Because
Deleted: Because only one streamflow gauge station is present in each catchment, the spatial variation in model parameters cannot be recognized by calibration.
Deleted: the
Deleted: of VM
Deleted: hours
Deleted: is
Deleted: there is no need for good fitness of
Deleted: however, it may not be suitable for semiarid catchments because there is no need for full fitness of simulated and observed streamflows. (…
Deleted: ,
Deleted: ;
Deleted: SHARMA and MURTHY, 1998)
Deleted: Q_p
Deleted: Simultaneously, c
Deleted: are
Deleted: We set the number of iterations to 2000 in the model calibration step.…

3.4 Model comparison

To achieve a better performance in rainstorm flood simulations, three hydrologic models, including two conceptual models, XAJ and SBM, and one distributed model, MIKE SHE, are used for comparison with the VMM model. XAJ was developed by (Zhao, 1992) and has a single saturation excess runoff generation mechanism. XAJ has been successfully applied in humid and subhumid catchments (Cheng et al., 2006; Lü et al., 2013;). SBM was developed by Zhao (1983) and has a single infiltration excess runoff generation mechanism. SBM is generally used in semiarid or arid catchments in China (Bao et al., 2017; Li and Zhang, 2008; Zhao et al., 2013). In addition, MIKE SHE originated from the Système Hydrologique Européen (SHE) program, and it is a deterministic, physically based distributed hydrologic model that can simulate surface water flow, unsaturated flow and saturated flow (Jayatilaka et al., 1998). MIKE SHE has been used to solve water resources and environment problems at different spatiotemporal scales (Li et al., 2018; Rujner et al.,2018; Samaras et al., 2016).

3.5 Multicriteria assessment framework for flood events

Due to strong spatially variability of rainfall, complex landscape characteristics, insufficient rain gauges, and dispersion, flood event simulation and forecasting in semiarid catchments are very difficult. Although some hydrologists improve flood simulations and forecasting by improving hydrologic models, the improvements are always limited or are suitable for only certain regions (Collier, 2007). The flood peak is the most significant feature in semiarid regions. Determining the extent to which the calculation of flood peaks can be accepted is crucial. Generally, the absolute relative error is used to measure the calculation of flood peak accuracy, for example, 20%, 30% or similar values are acceptable (Li et al., 2014; McIntyre and Al-Qurashi, 2009). To provide more information for flood defense management, the generalized likelihood uncertainty estimation (GLUE) and Bayesian method with the Markov chain Monte Carlo sampling are used to provide probabilistic forecasting, such as the 95% confidence interval (Christiaens and Feyen, 2002; Li et al., 2017), although these methods may not lead to clear decisions (Beven, 2007).

In this study, to obtain a useful method for decision maker, we propose a multicriteria assessment framework for flood forecasting in the catchments of middle Yellow River. This framework can be described as follows:

(C1) the absolute relative error of peak flow should be less than 20%.

(C2) modeled and observation of peak flows should be in the same flow zone: the observed peak flow Q_p of all flood events in a catchment are divided into three zones (low flow zone, medium flow zone, high flow zone), with 25th percentiles Q_{p25} and 75th percentiles Q_{p75} as the boundary points; if $Q_p \leq Q_{p25}$, then the peak flow Q_p belongs to the low flow zone; and if $Q_p \geq Q_{p75}$, then the peak flow Q_p belongs to the high flow zone; the rest flow peaks belongs to medium flow zone. Both the 25th percentile and 75th percentile are commonly used to distinguish zones.

(C3) The modeled peak flows should fall within one standard deviation (σ) of the mean (approximately 68.3% confidence interval) peak flow estimated by the HUP, one component of the Bayesian forecasting system detailed in Krzysztofowicz (1999)

Deleted: event-based…rainstorm rainstorm …lood simulations
Deleted: We test the performance of XAJ in the semiarid
Formatted: Highlight
Deleted: programme of …ystème Hydrologique Européen (SHE)
Deleted: and
Deleted: spatio-temporal
Deleted: MIKE SHE is one of the most widely used physically
Deleted: A m…lticriteriaon
Deleted: of
Deleted: ity of
Deleted: zes
Deleted:
Deleted: and…
Deleted: lack of enough…rain gauges, and also
Deleted: event-based
Deleted: s
Deleted: s
Deleted: only
Deleted: only
Deleted: to
Deleted: some specific
Deleted: F
Deleted: flood
Deleted: other of that is…imilar values are acceptable (Li et al.,
Deleted: Bayesian method
Deleted: d…probabilistic forecasting, such as the like…95%
Deleted: the way
Deleted: acquire…a usefulutility
Deleted: multiit-criterion
Deleted: s
Deleted: ling…and observation of peak flows should be in the
Deleted: respectively… if each
Deleted: of …he 25th percentile and 75th percentile is
Deleted: t
Deleted: ing
Deleted: about …pproximately 68.3% confidence interval)
Deleted: hydrologic uncertainty processor (…UP)
Deleted: approach…(
Deleted: detailed can be found in
Deleted: (…rzysztofowicz (,

and Biondi et al. (2010).

The key of the framework is C2, and C1 is used to avoid errors caused by flow zone boundaries. For example, when $Q_{p75} = 200$ m^3/s, the modeled peak flow equals 198 m^3/s, and the observed peak flow equals 201 m^3/s. However, using only condition C2 may lead to inappropriate model results; adding C1 can help address the problem. C3 is used to test the confidence level of modeled peak flows. A modeled peak flow that can be accepted should satisfy condition C1 or condition C2 and then condition C3.

3.6 Parameter sensitivity analysis

To assess the effects of inputs on the model output, a sensitivity analysis (SA) was proposed (Saltelli et al., 1989). The SA can be classified into a GSA and local sensitivity analysis (LSA). Compared with the LSA, the GSA is able to analyze the effects of inputs within the entire input domain. The Fourier amplitude sensitivity test (Cukier et al., 1973), Sobol method (Sobol, 1993) and Morris screening method (Morris, 1991) are the most widely used GSA methods in the assessment of parameter sensitivity in hydrologic models. Pianosi and Wagener (2015) proposed the novel GSA method PAWN (derived from the authors' names), which is based on the cumulative density function. PAWN has advantages over the parameter ranking and time-consuming nature of other GSA methods (Khorashadi et al., 2017). In this study, we use the PAWN method to perform a GSA on the VMM model.

Considering $x_{i,j}$ $(i, j = 1, 2, \cdots$, where i and j represent the i-th input parameters and the j-th sampling, respectively) as sensitivity inputs, then the sensitivity of $x_{i,j}$ can be measured by the distance between $F_{(y_i|x_{i,j})}(y_i)$ (the cumulative probability distribution function of y_i when $x_{i,j}$ changes between the upper bound and lower bound) and $F_{y_i}(y_i)$ (the cumulative probability distribution function of y_i; when $x_i = \frac{1}{n}\sum_{j=1}^{n} x_{i,j}$, where n is the number of samplings per input parameter). The Kolmogorov–Smirnov statistic (Simard and Ecuyer, 2011) is used to measure the distance between $F_{(y_i|x_i)}(y_i)$ and $F_{y_i}(y_i)$:

$$KS(x_{i,j}) = \max_{1 \le j \le n} \left| F_{y_i}(y_i) - F_{(y_i|x_{i,j})}(y_i) \right| \quad (12)$$

As KS varies with $x_{i,j}$, the maximum of all possible KS is included in the PAWN index P_i:

$$P_i = \max_{1 \le j \le n} KS(x_{i,j}) \quad (13)$$

P_i ranges from 0 to 1. The closer P_i is to 1, the more sensitive x_i is. A P_i equal to 1 indicates that x_i has no effect on the model. For more information about PAWN, please refer to Pianosi and Wagener (2015). In this study, as Pianosi and Wagener (2018) suggested, the number of evaluations is set to 500.

3.7 Model validation

The modeling time step was hourly, and the modeling period was between 1983 and 2009. In the Qiushui River, 20 flood

Deleted: ;
Deleted: ,
Deleted:))
Deleted: ling
Deleted: only
Deleted: , only
Deleted:
Deleted: the
Deleted: the
Deleted: ed
Deleted: ing
Deleted:
Deleted: not to be accepted
Deleted: .
Deleted: So
Deleted: However,
Deleted: deal with
Deleted: ing
Deleted: ing
Deleted: '
Deleted: (
Deleted: ,
Deleted:)
Deleted: Zadeh
Deleted:
Deleted: seen as

events were selected, with the first 15 events used for calibration and the remaining five events used for validation. Similarly, in the Qingjian River, 29 flood events were selected, with 24 events used for calibration and the remaining five events used for validation. In the Tuwei River, 23 flood events were selected, with 18 events used for calibration and the remaining five events used for validation. Finally, in the Kuye River, 28 flood events were selected, with 23 events used for calibration and the remaining five events used for validation.

4 Results and discussion

4.1 Comparison of model results

Boxplots of the absolute relative errors of the peak flows for each model in the four catchments are shown in Fig. 3. In terms of the median and average of the absolute relative errors for peak flows, except for the validation period in the Kuye River catchment shown in Fig. 3 (h), Figures 3 (a)–(g) reveal that the VMM has lowest values for both calibration and validation; in most cases, MIKE SHE performs better than the XAJ and SBM, i.e., Figures 3 (a), (b), (c), (d), (g), (h). Except for the good performance in the Tuwei River catchment, the SBM is as poor as the XAJ in other catchments. In terms of ranges of the absolute relative errors for peak flows, the VMM and MIKE SHE have relatively small ranges (Figs. 3(a), (c), (d), (g)); and the SBM and XAJ have large ranges in most cases (Figs. 3(a), (b), (c), (d), (g)).

Tables 3 and 4 show the average performance in terms of the absolute relative error for flow volume E_v and the lag time for the four models in each catchment, respectively. The VMM has the minimum average E_v and lag time, with values of 39.01% and 3.05 h, respectively (Tables 3 and 4). In contrast, the XAJ has the maximum average E_v and lag time, with values of 58.93% and 4.51 h, respectively. MIKE SHE and SBM have similar performances in terms of average E_v and lag time.

The analysis of Fig. 3, Table 3 and Table 4 shows that the VMM has the best performance for flood modeling in the four studied catchments of the middle Yellow River, and the XAJ has the worst performance. In addition, MIKE SHE is slightly superior to the SBM. Although MIKE SHE is a distributed hydrologic model with more complex structures and more explicit physical meaning than the conceptual model VMM, it does not necessarily achieve better results than conceptual models because distributed models lack sufficiently high-resolution data, and this finding is consistent with other studies (Beven, 2002, 2011; Michaud and Sorooshian, 1994; Seyfried and Wilcox, 1995). Both infiltration excess and saturation excess can be simulated via the VMM, which may be why it performs better than the other two conceptual models (XAJ and SBM), which have single runoff generation mechanisms (saturation excess and infiltration excess, respectively).

4.2 Sensitivity analysis of the VMM

The GSA method PAWN is applied to estimate the influence of parameter uncertainty on the model output results. Figure 4 (a) and Figure 4 (b) show the SA results of all study catchments for the objective function E_p (Eq. 9) or E_{pv} (Eq. 11), respectively. The higher the ranking is, the more sensitive the parameters. Parameters *CS*, *IM* and *KE* have the highest rankings

Deleted: and …ith the first 15 events were …sed for calibration …
Deleted: 5
Deleted: were …sed for validation. Similarly, in the Qingjian …
Deleted: 5
Deleted: were …sed for validation. In the Tuwei River, 23 flood …
Deleted: 5
Deleted: were …sed for validation. Finally, in the Kuye River, 28 …
Deleted: 5
Deleted: were
Deleted: For flow peaks,
Deleted: The …b …
Deleted: is
Deleted: Fig.
Deleted: ure
Deleted: Except
Deleted: Fig.
Deleted: ure
Deleted: s
Deleted: s
Deleted: Figs.…igures 3 (a), (b), (c), (d), (g), (h).;…Ee…cept fo…
Deleted: ure
Deleted: , i.e., Figs.
Deleted: ure
Deleted: it is obvious that VMM performs better than the other …
Deleted: Table
Deleted: the average performance for
Deleted: s of
Deleted: Table
Deleted: s of
Deleted: It can be concluded from the
Deleted: An…he analysis of Fig. 3, Table 3 and Table 4 …
Deleted: ,…and the XAJ has the worst performance. In addition …
Deleted: of VM
Deleted: ….4 (a) and Figure. …
Deleted: (…))…orand …
Deleted: (…1)) …
Deleted: We can find that the p

whether the objective function of the VMM model is E_p or E_{pv}. The rankings of other parameters are influenced slightly by different objective functions, such as *CG*, except for *WM*. *WM* ranks sixth when E_{pv} is the objective function and 12th when E_p is the objective function. *WM* controls the tension water content in the soil, which determines the amount of rainfall stored in the soil and the generation of runoff. Therefore, it is reasonable to conclude that when the weight of the flow volume is added in E_p, which can be expressed as E_{pv}, the ranking of *WM* increases, in this case to sixth place.

4.3 Multicriteria assessment framework of the VMM

The multicriteria assessment framework we propose is applied to assess the ability of the VMM to model flood peaks in four catchments. The framework requires that an accepted flood event should meet one of the requirements of C1 and C2; in addition, C3 needs to be satisfied simultaneously. The observed peak flows and the modeled peak flows under the conditions C1, C2 or C3 are shown in Fig. 5. We find that the majority of peak flows for the observations and modeling fall between the 15.85th percentile and the 81.45th percentile (68.3% confidence interval) estimated by HUP, which means that the VMM modeling results satisfy C3 fairly well. Under the premise of satisfying C3, the number of modeling events satisfying C2 is slightly greater than that satisfying C1. Under the multicriteria assessment framework, 15 of the 20 (75.0%) modeled flood events in the Qiushui River catchment, 12 of the 29 (41.4%) of that in the Qingjian River catchment, 15 of the 23 (65.2%) of that in the Tuwei River catchment, and 16 of the 28 (57.1%) of that in the Kuye River catchment can be accepted. The average acceptance rate for the four catchments is 58%, which is greater than the acceptance rate of 41% under the single criterion C1, a common assessment for peak flows.

The multicriteria assessment framework can provide more reasonable and reliable flood early warning information for decision-makers. Taking the 13th flood event of the Kuye River catchment as an example, the observed and modeled peak flows are 1230 m^3/s and 1510 m^3/s, respectively. As shown in Fig. 5 (d), the absolute relative error for peak flow is greater than 20%, and the peak flow does not fall in the 68.3% confidence interval; however, these parameters are in the same zone, i.e., the medium flow zone. For the Kuye River catchment, it is reasonable to believe that the peak flows 1230 m^3/s and 1510 m^3/s correspond to the same level according to the known flood peak data, which is the role played by C2. Although the dividing flow zone method of C2 is coarse, it is convenient and beneficial for flood defense.

5 Conclusions

In this study, a multicriteria assessment framework of flood peaks is proposed with the vertically mixed runoff model (VMM) in four catchments in the middle Yellow River. The main conclusions are as follows.

(1) Compared with the distributed model MIKE SHE and the two conceptual models Xinanjiang (XAJ) and Shanbei (SBM), the VMM has better performance for modeling flood events in semiarid catchments of the middle Yellow River.

(2) In the four catchments, the parameters confluence coefficient of surface flow (*CS*), impermeable area (*IM*), and residence

Deleted: The most sensitive parameters – *CS*, *IM* and *KE* – are not affected by different objective functions.
Deleted: slightly
Deleted: …xcp…pc
Formatted: Font: Italic
Deleted: Eq. (11)
Deleted: on…assessment framework of VM
Deleted: multicriterion…ulticriteria assessment framework we
Deleted: must…meet one of the requirements of C1 andor
Deleted: besides
Deleted: , and …3 must be
Deleted: Fig.
Deleted: ure…5. Flood events conforming to conditions for C1
Deleted: both of
Deleted: peak flows of most flood events
Deleted: more …reater than that satisfying C1. Under the
Deleted: flood events
Deleted: ;
Deleted: . In this case,
Deleted: t
Deleted: for…the single criterion C1, a common assessment for
Deleted: multicriterion…ulticriteria assessment framework can
Deleted: ure
Deleted: ,
Deleted: s
Deleted: , but…these parametersy…are in the same zone, i.e., the
Deleted: multicriterion
Deleted:
Deleted: vertically mixed runoff model (…MM))
Deleted: model
Deleted: :
Deleted: ,…and the two conceptual models Xinanjiang
Deleted: in terms of…or modeling event-based…flood
Deleted: by PAWN analysis of VMM, *CS*
Formatted: Font: Italic
Deleted: *IM*…mpermeable area (*IM*), and *KE*

time of Muskingum (*KE*) are the most sensitive based on an analysis by the global sensitivity method PAWN; in addition, the sensitivity ranking of the parameter *WM* related with the soil moisture capacity is the most affected by the objective functions.

(3) The multicriteria assessment framework can provide more reliable flood early warning information than single criterion (such as absolute relative error of peak flows) when engineers need to make decisions in semiarid catchments.

The condition C2, which divides peak flows into three flow zones, will be affected by the number of known peak flows when data availability is limited. The framework is suitable for semiarid regions with poor modeling results and can provide guidance for decision making.

Code availability

We have shared the MATLAB code of the VMM model at https://doi.org/10.4211/hs.c5232287d5c04bfb8cac5ce4e391ea0f.

Acknowledgments

This study was supported by the National Key Research and Development Program of China (grant no. 2016YFC0402706) and National Natural Science Foundation of China (grant nos. 41730750, 41877147). We would like to thank Francesca Pianosi (University of Bristol) for providing the program code of PAWN at https://www.safetoolbox.info/pawn-method/. We also thank the anonymous reviewers, whose comments have largely improved this work.

References:

Andersen, F. H.: Hydrological modeling in a semi-arid area using remote sensing data, Ph.D. thesis, University of Copenhagen, Copenhagen, Denmark, 2008.

Bao, H., Wang L., Zhang, K. and Li, Z.: Application of a developed distributed hydrological model based on the mixed runoff generation model and 2D kinematic wave flow routing model for better flood forecasting, Atmos. Sci. Lett., 18(7), 284–293, 2017.

Bao, W.: Improvement and application of the Green-Ampt infiltration curve, Yellow River, 9, 1–3, 1993. (In Chinese)

Bao, W. and Zhao, L.: Application of Linearized Calibration Method for Vertically Mixed Runoff Model Parameters, J. Hydrol. Eng., 33(4), 85–91, 2014.

Beven, K. J.: Surface water hydrology—runoff generation and basin structure, Rev. Geophys., 21(3), 721–730, 1983.

Beven, K. J.: Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, Hydrol. Process., 16(2), 189–206, 2002.

Beven, K. J.: Environmental modelling: An uncertain future?, CRC press, London, UK, 328 pp., 2007.

Deleted: analyzed by

Deleted: are

Formatted: Font: Italic

Deleted: parameters and

Deleted: not

Deleted: choice of

Deleted: , whereas *WM* is the most sensitive parameter.

Deleted: multicriterion

Deleted: reasonable

Deleted: criteria

Deleted: regarding

Deleted: if

Deleted: enrich

Deleted: people's

Deleted: - making

Deleted: of VM

Deleted: s

Deleted: and 2016YFC0402709

Deleted: the Major Program of the

Deleted:

Deleted: s as

Deleted: their

Deleted: .

Deleted: :

Deleted: ¶

Deleted: Burnash, R. J. C., Ferral, R. L., McGuire, R. A. A.: Generalized streamflow simulation system, conceptual modeling for digital computers, 1973.¶

Bao, W.: Improvement and application of the Green-Ampt infiltration curve, Yellow River, 09, 1–3, 1993. (In Chinese)¶

Bao, W. and Wang, C.: Application of the Vertically mixed runoff model, Journal of China Hydrology, 03, 19-22, 1997. (In Chinese)¶

Bao, W. and Zhao, L.: Application of Linearized Calibration Method for Vertically Mixed Runoff Model Parameters, J. Hydrol. Eng., ...

Formatted: Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: ¶

Deleted: Beven, K. and Freer, J.: A dynamic TOPMODEL, ...

Beven, K. J.: Rainfall-runoff modelling: the primer, John Wiley & Sons, UK, 488 pp., 2011.Beven, K. J., and Freer, J.: A dynamic TOPMODEL, Hydrol. Process., 15(10), 1993–2011, 2001.

Biondi, D., Versace, P., and Sirangelo, B.: Uncertainty assessment through a precipitation dependent hydrologic uncertainty processor: An application to a small catchment in southern Italy, J. Hydrol., 386(1–4), 38–54, 2010.

Brito, M. and Evers, M.: Multi-criteria decision-making for flood risk management: A survey of the current state of the art, Nat. Hazards Earth Syst. Sci., 16(4), 1019–1033, 2016.

Burnash, R. J., Ferral, R. L., and McGuire, R. A.: A generalized streamflow simulation system, conceptual modeling for digital computers, Report by the Joliet Federal State River Forecasts Center, Sacramento, CA, 204 pp., 1973.

Cheng, C. T., Zhao, M. Y., Chau, K., and Wu, X. Y.: Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure, J. Hydrol., 316, 129-140, 2006.

Christiaens, K., and Feyen, J.: Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework, Hydrol. Process., 16(2), 373–391, 2002.

Collier, C. G.: Flash flood forecasting: What are the limits of predictability?, Q. J. Roy. Meteor. Soc., 133(622), 3–23, https://doi.org/10.1002/qj.29, 2007.

Cukier. R., Fortuin, C., Shuler, K., and Petschek, A., and Schaibly. J.: Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory, J. Chem. Phys., 59(8), 3873–3878, 1973.

Devia, G. K., Ganasri, B. P., and Dwarakish, G. S.: A review on hydrological models. Aquatic Procedia, 4, 1001–1007, 2015.

Dinku, T., Ceccato, P., Kopec, E. G., Lemma, M., Connor, S. J., and Ropelewski, C. F.: Validation of satellite rainfall products over East Africa's complex topography, Int. J. Remote Sens., 28(7), 1503–1526, 2007.

Duan, Q. Y., Gupta, V. K., and Sorooshian, S.: Shuffled complex evolution approach for effective and efficient global minimization, J. Optimiz. Theory App., 76(3), 501–521, 1993.

Hao, G., Li, J., Song, L., Li, H., and Li, Z.: Comparison between the TOPMODEL and the Xin'anjiang model and their application to rainfall runoff simulation in semi-humid regions, Environ. Earth Sci., 77(7), 279, 2018.

Jayatilaka, C., Storm, B., and Mudgway, L.: Simulation of water flow on irrigation bay scale with MIKE-SHE, J. Hydrol., 208, 108-130, 1998.

Jiang, Y., Liu, C., Li, X., Liu, L., and Wang, H.: Rainfall-runoff modeling, parameter estimation and sensitivity analysis in a semiarid catchment, Environ. Modell. Softw., 67, 72–88, 2015.

Khomsi, K., Mahe, G., Tramblay, Y., Sinan, M., and Snoussi, M.: Regional impacts of global change: seasonal trends in extreme rainfall, run-off and temperature in two contrasting regions of Morocco, Nat. Hazards Earth Syst. Sci., 16(5), 1079–1090, 2016.

Khorashadi, Z. F., Nossent, J., Sarrazin F., Pianosi, F., Griensven, V. A., Wagener, T. and Bauwens, W.: Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model,

Deleted: Beven. K.: Environmental modelling: An uncertain future?, CRC press, 2007.¶

Deleted: Beven, K.: Rainfall-runoff modelling: the primer, John Wiley & Sons, 2011.¶

Deleted: Cukier. R., Fortuin, C., Shuler, K., and Petschek, A., and Schaibly. J.: Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, I Theory, J. Chem. Phys., 59(8), 3873–3878, 1973.

Deleted: Cheng, L., Xu, Z., Luo, R., and Mi, Y.: SWAT application in arid and semi-arid region: a case study in the Kuye River Basin, Geogr. Res., 01, 65–73, 2009. (In Chinese)¶

Formatted: Highlight

Deleted:

Deleted: ,

Deleted: ¶

Formatted: Indent: Left: 0 cm, First line: 0 ch

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, Water Resour. Res., 35(9), 2739-2750, 1999.¶

Environ. Modell. Softw., 91, 210–222, 2017.

Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, Water Resour. Res., 35(9), 2739-2750, 1999.

Li, B., Liang, Z., He, Y., Hu, L., Zhao, W. and Acharya K.: Comparison of parameter uncertainty analysis techniques for a TOPMODEL application, Stoch. Env. Res. Risk A., 31(5), 1045–1059, 2017.

Li, B., Liang, Z., Bao, Z., Wang, J., and Hu, Y.: Changes in streamflow and sediment for a planned large reservoir in the middle Yellow River, Land Degrad. Dev., https://doi.org/10.1002/ldr.3274, 2019.

Li, B., Yu., Z., Liang., Z., Song. K., Li, H., Wang, Y., Zhang, W. and Acharya. K.: Effects of Climate Variations and Human Activities on Runoff in the Zoige Alpine Wetland in the Eastern Edge of the Tibetan Plateau, J. Hydrol. Eng., 19(5), 1026-1035, 2014.

Li, D.: Hydrologic model: the vertically mixed runoff model (vmm), HydroShare, https://doi.org/10.4211/hs.c5232287d5c04bfb8cac5ce4e391ea0f, 2018.

Li, D., Liang, Z., Li, B., Lei X., and Zhou Y.: Multi-objective calibration of MIKE SHE with SMAP soil moisture datasets, Hydrol. Res., https://doi.org/10.2166/nh.2018.110, 2018.

Li, X. and Huang, C. C.: Holocene palaeoflood events recorded by slackwater deposits along the Jin-shan Gorges of the middle Yellow River, China, Quatern. Int., 453, 85–95, 2017.

Li, Z. J., and Zhang, K.: Comparison of three GIS-based hydrological models, J. Hydrol. Eng., 13, 364-370, 2008.

Lü, H., Hou, T., Horton, R., Zhu, Y., Chen, X., Jia, Y., Wang, W. and Fu, X.: The streamflow estimation using the Xinanjiang rainfall runoff model and dual state-parameter estimation method, J. Hydrol., 480, 102–114, 2013.

McIntyre, N., and Al-Qurashi, A.: Performance of ten rainfall–runoff models applied to an arid catchment in Oman, Environ. Modell. Softw., 24, 726–738, 2009.

McMichael, C. E., Hope, A. S., and Loaiciga, H. A.: Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation, J. Hydrol., 317, 307–324, 2006.

Michaud, J., and Sorooshian, S.: Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed, Water Resour. Res., 30, 593–605, 1994.

Morris, M. D.: Factorial sampling plans for preliminary computational experiments, Technometrics, 33, 161–174, 1991.

Mwakalila, S., Campling, P., Feyen, J., Wyseure, G. and Beven, K.: Application of a data-based mechanistic modelling (DBM) approach for predicting runoff generation in semi-arid regions, Hydrol. Process., 15, 2281–2295, 2001.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Pianosi, F., and Wagener, T.: A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ. Modell. Softw., 67, 1–11, 2015.

Deleted: ¶

Moved down [1]: Lü, H., Hou, T., Horton, R., Zhu, Y., Chen, X., Jia, Y., Wang, W. and Fu, X.: The streamflow estimation using the Xinanjiang rainfall runoff model and dual state-parameter estimation method, J. Hydrol., 480, 102–114, 2013.

Formatted: Indent: Left: 0 cm, Hanging: 2 ch, First line: -2 ch

Formatted: Font color: Black

Deleted: Li, B., Liang, Z., He, Y., Hu, L., Zhao, W. and Acharya K.: Comparison of parameter uncertainty analysis techniques for a TOPMODEL application, Stoch. Env. Res. Risk A., 31(5), 1045–1059, 2017.¶
Li, X. and Huang, C. C.: Holocene palaeoflood events recorded by slackwater deposits along the Jin-shan Gorges of the middle Yellow River, China, Quatern Int., 453, 85–95, 2017.¶

Deleted: ,

Deleted:

Deleted:

Deleted: ,

Formatted: Default Paragraph Font, Font color: Black

Deleted: Li, D., Liang, Z., Hou, B., Li, B., and Wang, J.: Application of vertical hybrid model on flood forecast and early-warning of Qiushui River, Yellow River, 06, 24–28, 2018a. (In Chinese)¶

Formatted: Highlight

Deleted: ogy

Deleted: earch

Deleted: b

Moved (insertion) [1]

Deleted: ¶

Pianosi, F., and Wagener, T.: Distribution-based sensitivity analysis from a generic input-output sample, Environ. Modell. Softw., 108, 197–207, 2018.

Pilgrim, D. H., Chapman, T. G. and Doran, D. G.: Problems of rainfall-runoff modelling in arid and semiarid regions, Hydrolog. Sci. J. 33, 379–400, 1988.

Rujner, H., Uuml, G., Leonhardt, N., Marsalek, J., and Viklander, M.: High-resolution modelling of the grass swale response to runoff inflows with Mike SHE, J. Hydrol., 562, 411–422, 2018.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto M.: Sensitivity Analysis in Practice. J. Am. Stat. Assoc., 101, 398–399, 1989.

Samaras, A. G., Gaeta, M. G., Moreno, M. A., and Archetti R.: High-resolution wave and hydrodynamics modelling in coastal areas: operational applications for coastal planning, decision support and assessment, Nat. Hazards Earth Syst. Sci., 16(6), 1499–1518, 2016.

Seyfried, M. S., and Wilcox, B. P.: Scale and the nature of spatial variability: Field examples having implications for hydrologic modeling, Water Resour. Res., 31, 173–184, 1995.

Sharma, K. D., and Murthy, J. S. R.: A practical approach to rainfall-runoff modelling in arid zone drainage basins, Hydrolog. Sci. J., 43(3), 331–348, 1998.

Simard, R., and Ecuyer, P. L.: Computing the two-sided Kolmogorov-Smirnov distribution, J. Stat. Softw., 39(11), 1–18, https://doi.org/10.18637/jss.v039.i11, 2011.

Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp., 1, 407–414, 1993.

United Nations Environment Programme (UNEP): World Atlas of Desertification, Edward Arnold, London, 69 pp., 1992.

Wang, G., and Ren, L.: A Contrastive Study of Simulation Results between GWSC-VMR and Hybrid Runoff Model in Dianzi Basin, in: International Conference on Environmental Science and Information Application Technology, Wuhan, China, 4–5 July, 583-588, 2009.

Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M. and Stewart, A.: Understanding uncertainty in distributed flash flood forecasting for semiarid regions, Water Resour. Res., 44(5), 61–74, 2008.

Young, C. B., Nelson, B. R., Bradley, A. A., Smith, J. A., Peters-Lidard, C. D., Kruger, A. and Baeck, M.L.: An evaluation of NEXRAD precipitation estimates in complex terrain, J. Geophys. Res.-Atmos., 104, 19691–19703, 1999.

Zhao, L., Xia, J., Xu, C. Y., Wang, Z., Sobkowiak, L., and Long, C.: Evapotranspiration estimation methods in hydrological models, J. Geogr. Sci., 23, 359-369, https://doi.org/10.1007/s11442-013-1015-9, 2013.

Zhao, R. J.: Watershed Hydrological Model: Xin'anjiang Model and Shanbei Model, Water and Power Press, Beijing, China, 1983. (In Chinese)

Zhao, R. J.: The Xinanjiang model applied in China, J. Hydrol., 135, 371–381, 1992.

Deleted: ical

Deleted: ences

Deleted: ournal

Deleted: Refsgaard, C. J.: Mike she, Computer models of catchment hydrology, 809–846, 1995.¶

Formatted: Highlight

Deleted: 32

Formatted: Indent: Left: 0 cm, Hanging: 2 ch, First line: -2 ch, Don't adjust right indent when grid is defined, Line spacing: 1.5 lines, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and

Deleted: ¶

United Nations Environment Programme (UNEP), World Atlas of Desertification, 1992.¶

Wen, S., and Cai, T.: Application of improved vertical mixed runoff model in Dalinghe river basin, Water Resources & Hydropower of Northeast, 33(4), 45–48, 2015. (In Chinese)¶

Wang, G., Wang, J., Jing, X., and Li, H.: Application of SIMHYD model in the Qinjiang River Basin, Yellow River, 03, 29–30, 2006. (In Chinese)¶

Sharma, K. D., and Murthy, J. S. R.: A practical approach to rainfall-runoff modelling in arid zone drainage basins, Hydrological Sciences Journal, 43(3), 331–348, 1998.¶

Simard, R., and Ecuyer, P. L.: Computing the two-sided Kolmogorov-Smirnov distribution, J Stat. Softw., 39, 1–18, 2011.¶

Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Mathematical modelling and computational experiments, 1, 407–414, 1993.¶

Guo, W., Wang, C. Ma, T., Zeng, X. and Yang, H.: A distributed Grid-Xinanjiang model with integration of subgrid variability of soil storage capacity. Water Science and Engineering, 9, 97–105, 2016.¶

Formatted: Highlight

Formatted: Highlight

Formatted: Font color: Red

Deleted: EOPHYS. RES-ATMOS.,

Formatted: Indent: Left: 0 cm, Hanging: 2 ch, First line: -2 ch, Don't adjust right indent when grid is defined, Line spacing: 1.5 lines, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and

Table 1. Characteristics of the four catchments

Catchment	Area (km^2)	Outlet station	Area* (km^2)	Mean annual precipitation (mm)	Mean evaporation (mm)	Aridity index
Qiushui River	1989	Linjiaping	1873	499	1150	2.31
Qingjian River	4080	Yanchuan	3468	451	1080	2.4
Tuwei River	3294	Gaojiachuan	2095	377	1050	2.78
Kuye River	8706	Wenjiachuan	8645	410	1010	2.46

* The area of the catchment controlled by the outlet station in the table.

Deleted: ¶

¶

¶

¶

¶

¶

¶

¶

¶

¶

Formatted: Indent: Left: 0 cm, First line: 0 ch

Formatted: Indent: Left: 0 cm

Deleted: a

Deleted: ¶

Table 2. Calibrated parameters of the VMM

Symbol	Definition	Range*
KC	Ratio of potential evapotranspiration to pan evaporation	[0.5, 1.5]
WM	Mean areal maximum possible soil moisture, mm	[50, 200]
FC	Stable infiltration capacity, mm/h	[5, 100]
K	Infiltration index related to soil permeability, /h	[0.05, 1]
BF	Index of the watershed infiltration capacity curve	[0, 0.5]
B	Index of the watershed water storage capacity curve	[1, 2]
KI	Outflow coefficient of interflow, d	[0.1, 0.5]
KG	Outflow coefficient of groundwater, d	[0.5, 2]
CS	Confluence coefficient of surface flow	[0.05, 0.9]
CI	Recession coefficient of interflow, d	[0.5, 0.95]
CG	Recession coefficient of groundflow, d	[0.90, 0.99]
KE	Residence time of Muskingum, h	[0.5, 5]
XE	Muskingum coefficient	[0.01, 0.49]
IM	Impermeable area	[0, 1]

*In [a, b], a and b represent the lower and upper bounds of the parameters, respectively.

Deleted: ¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶

Deleted: of VM

Deleted: Meaning

Deleted: ¶
¶
¶
¶
¶
¶
¶

Deleted: ¶

Table 3. Performance (in terms of absolute relative error) for peak flow E_v in each catchment in the four models

Unit: %

	Qiushui River	Qingjian River	Tuwei River	Kuye River	Average*
VMM	26.52	58.50	40.20	30.80	39.01
MIKE SHE	40.50	60.70	45.30	38.20	46.18
XAJ	56.60	66.61	60.20	52.30	58.93
SBM	38.14	55.82	35.50	45.2	43.15

*The average E_v of the four catchments for each model

Deleted: The

Deleted: p

Table 4 Lag time of the peak flow in the four catchments in the four models Unit: h

	Qiushui River	Qingjian River	Tuwei River	Kuye River	Average*
VMM	2.20	3.02	3.46	3.50	3.05
MIKE SHE	2.50	3.50	4.20	3.90	3.53
XAJ	4.10	3.81	5.62	4.50	4.51
SBM	4.00	2.95	3.46	4.20	3.65

*The average lag time in the four catchments for each model

Deleted: ¶

Deleted: ¶

¶

¶

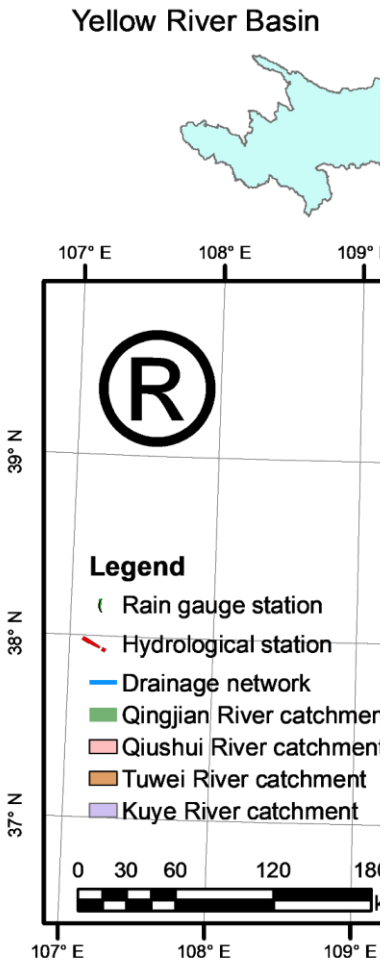
¶

¶

¶

Deleted: The

Deleted: l



Deleted:

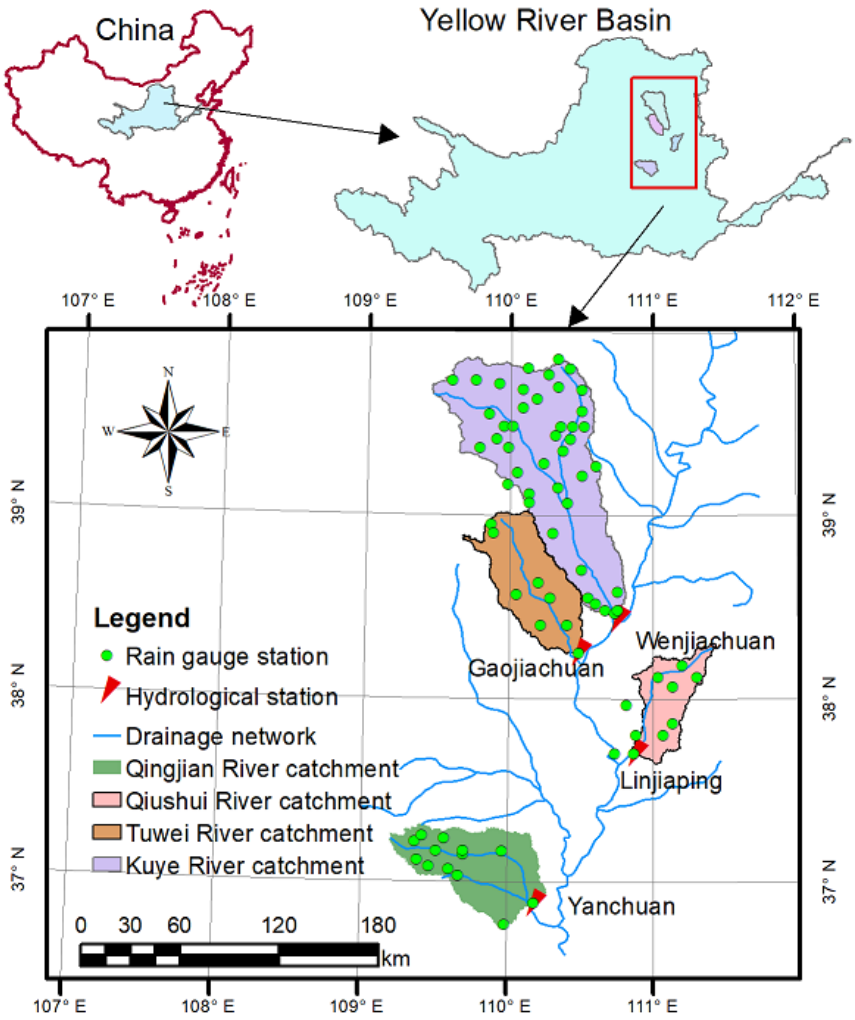


Figure 1: Location of the Qingjian River catchment, Qiushui River catchment, Tuwei River catchment and Kuye River catchment.

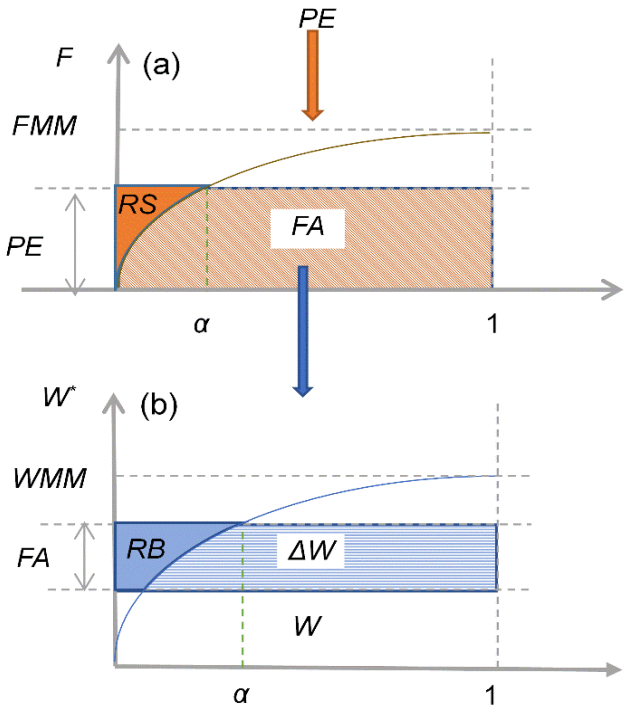


Figure 2: Runoff generation module in the VMM. (a) Infiltration capacity curve; and (b) tension water content storage capacity curve. α is the fracture area that is saturated, and F represents the point infiltration capacity.

Deleted: The i

Deleted: the

Deleted: ¶

¶

¶

Deleted: ¶

¶

¶

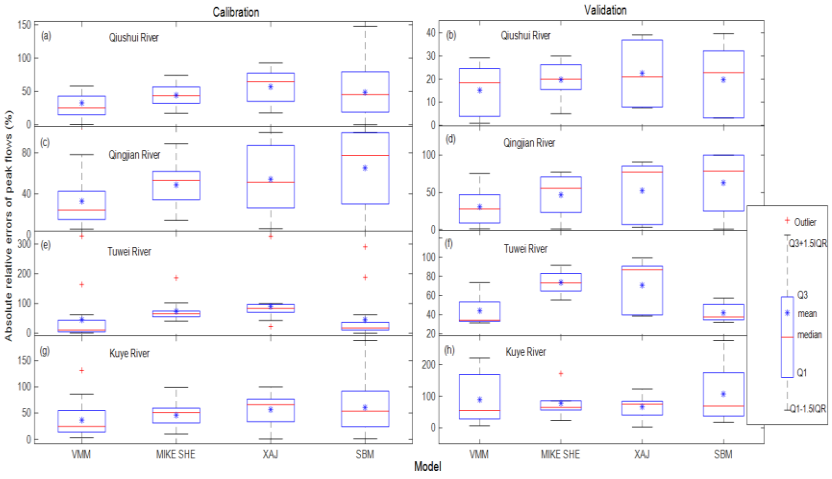
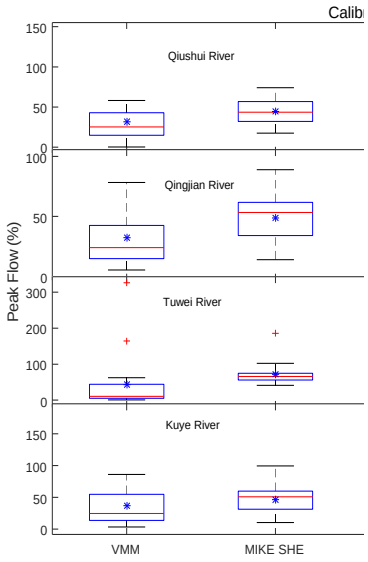


Figure 3: Boxplot of absolute relative errors of peak flows in the four catchments; Q1 and Q3 mean the first quantile and third quantile, respectively; interquartile range (IQR) = Q3 − Q1; and an outlier is defined as an extreme value that exceeds the IQR.

Formatted: Left: 3.17 cm, Right: 3.17 cm, Top: 2.54 cm, Bottom: 2.54 cm, Width: 21 cm, Height: 29.7 cm



Deleted:

Deleted: ¶

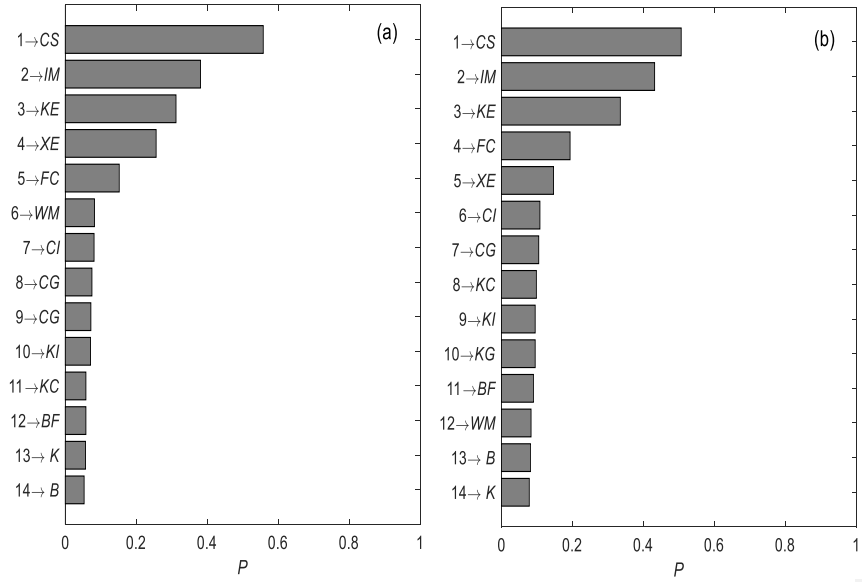


Figure 4: Sensitivity rankings of the VMM parameters based on the global sensitivity analysis method PAWN for different objective functions: (a) E_{pv} as the objective function, and (b) E_p as the objective function. The value P is used to assess the sensitivity degree of the parameter with PAWN method, and a larger value corresponds to greater sensitivity. The numbers on the ordinate represent the sensitivity rankings. E_{pv} E_{p_T}

Deleted: of VM

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Deleted: the

Deleted: the

Deleted: , the

Deleted: the

Deleted: ;

Deleted: t

Deleted: Figure 4: Sensitivity rankings of VMM parameters based on PAWN value *P* for different objective functions; (a)

Deleted: as the objective function; (b)

Deleted: as the objective function. The numbers on the ordinate represent the sensitivity ranking, where the larger the value, the greater the sensitivity.¶

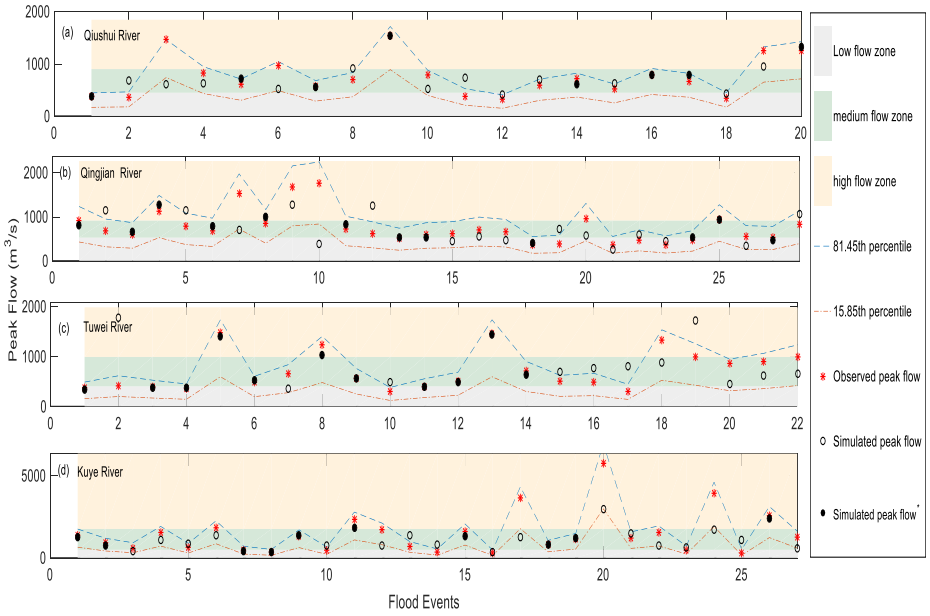
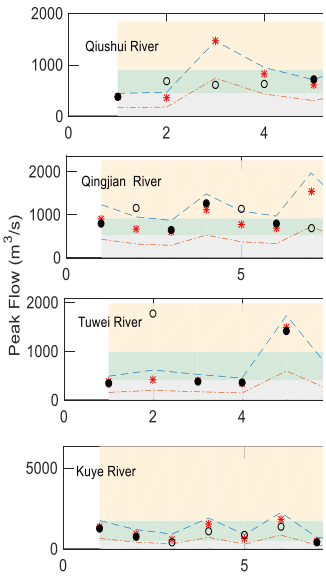


Figure 5: Observed peak flows (red asterisk) and simulated peak flows (solid ball and circle) with the VMM for each catchment under the conditions C1, C2 and C3. Flood peaks conforming to the condition C1 is represented by a solid ball; the three flow zones (low, medium and high flow zones) divided by the condition C2 are shown in gray, green and off-white, respectively; 68.3% confidence interval of peak flows estimated by the condition C3 is between the blue dashed line (81.45th percentile) and the red dashed-dotted line (15.85th percentile).

Formatted: Left: 1.65 cm, Right: 1.65 cm, Top: 1 cm, Bottom: 2.36 cm, Width: 21 cm, Height: 29.7 cm



Deleted:

Formatted: Centered

Deleted: T

Deleted: reseda

Deleted: zone

Deleted:

Deleted: are

Deleted: Figure 5: Multicriterion assessments of event-based floods with VMM in the four catchments. Simulated peak flow* means the value of the modeled peak flow meets¶ the condition C1.¶

Deleted: ¶