Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-710-SC1, 2018

Hydrology and Earth System Sciences
Discussions
Open Access

HESSD

Interactive comment

Interactive comment on "Using a multi-hypothesis framework to improve the understanding of flow dynamics during flash floods" *by* Audrey Douinot et al.

A. Douinot

audreydouinot@gmail.com

Received and published: 11 February 2018

We wish to thank the referee for their careful evaluation of the manuscript which would be hopefully more intelligible thanks to their constructive comments. Below are our detailed responses (in blue) to the comments (in black).

Sincerely yours,

This paper proposes a methodology for the analysis of catchment hydrological behaviors during flash floods, based on the introduction and comparison of several hypotheses in a distributed hydrological model. This topic is of broad interest for the hydrological scientific community, and is fully relevant in my opinion for a publication in HESS. However, in its current form the paper suffers from a lack of detail and explanations on several aspects (calibration procedure, explanations related to some figures, ..), causing difficulties for a detailed understanding of the research content. The presentation of sections 4 and 5 should particularly be improved in my opinion (and maybe organized in a slightly different way) to facilitate the overall understanding of the results and related analyses. The paper well illustrates the difficulties in the interpretation of modelling results, due to equifinality issues and lack of internal observations to confirm the nature of the main hydrological processes. Therefore, even if some solutions to cope with these difficulties are proposed here, I think the conclusions relative to the catchments behaviors (section 5.6) should finally be relativized and presented in the discussion section as the most reasonable assumptions, provided the modelling results obtained here.

We are grateful for the constructive comments. Most of them shift toward quite deep modifications of the sections 4 and 5. Following the reviewer's suggestions, the organization of section 5 has been changed, and many efforts were made to describe the methodology (section 4). The answers to the comments are listed below. The modifications made in the paper are quoted in italics. Line references correspond to the marked manuscript (attached file).

Specific comments:
The abstract is very short and could be slightly more detailed. The abstract has been reworded, giving now more details about the results (page 1, lines 5 - 10 :). Below, the reworded abstract :

"A method of multiple working hypotheses was applied to a range of catchments in the Mediterranean area to analyse different types of possible flow dynamics in soils during flash flood events. The distributed, process-oriented model, MARINE, was used to test several representations of subsurface flows, including flows at depth in fractured bedrock, and flows through preferential pathways in macropores. Results revealed different hydrological behaviours along the catchment set, giving advances in characterising the flash flood processing over the Mediterranean area. Those results are supported by their consistency with the rare available in-situ measurements and the priori knowledge of several catchments. The characterisation is nevertheless uncompleted owing to arising equifinality issues. The descriptive potential of the distributed model was then used to spot counterbalancing effects between internal flow processes and to finally propose new insights into strategical monitoring and calibration constraints setting up."

References would be welcome in section 1.2. References to relevant publications are added page 2, lines 16 - 24.

Comments about the section 4:

The description of the calibration procedure (section 4.1) and of the metrics for evaluation (section 4.2) are not sufficiently clear in my opinion, and should be improved:

- Please indicate how the "confidence intervals" are obtained for observations ($y_i \pm 2\sigma_i$?) and also for modelling results. This should clarify why the uncertainties ranges mentioned in the text (respectively 20% for observations and 10% for modelling errors) are consistent with eq.(6) and eq.(7). The definition of the confidence interval of the observed flows is now explicitly written page 15, lines 1 - 2 : *"The envelop* $\left((\hat{y}_i \pm 2\sigma_{\hat{y}_i}),\ i = 1...n\right)$ *consequently defines the 95 %*

confidence interval of the observed flows."

- please clarify the reason why the metric used for evaluation (Qmed_INT) is different from the one used for calibration (DEC) ? The Qmed_INT is here used as the meaning of this criteria is easier to interpret and understand, compared to the DEC value that is a standard error. This is clarified, page 16, lines 19-21: *"The DEC score has actually provided a standard assessment of the modelling errors enabling a reasonable weighting of the simulations. But in order to analyse the results, the Qmed_INT [%] score is preferred for the easy understanding it provides through it meaningful definition."*

- the definition of the "acceptability zone" should be provided ($y_i \pm 2\sigma_{y_i} \pm 2\sigma_{mod,i}$?). A definition of the acceptability zone is added page 15, lines 7-9 : *"Finally, the overall overarching envelop* $\left((\hat{y}_i \pm 2\sigma_{\hat{y}_i} \pm 2\sigma_{mod,i}),\ i = 1...n\right)$ *defines hereafter the acceptability zone, that is to say the interval into which any simulated flow would be considered as acceptable, according to the modelling and measurement uncertainty definitions."*

- the "a priori" and "a posteriori" modelling errors are not defined. This clearly limits the interpretation of figure 11 (see hereafter). Page 17, lines 5-20 : In order to clarify the variables used in the figure 11 and the related comments, 3 paragraphs were added in the section 4.2 (Metrics and key points in model evaluation and comparison):

The evaluation was then completed through the description of the modelling errors (section 5.2). The objective was to identify those that were inherent in the choice of model structure, regardless of the calibration methodology adopted. In that respect, attention was paid on the a priori and a posteriori confidence interval of the model simulations respectively defined by $\left([y_i^{prior-5th}, y_i^{prior-95th}],\ i = 1...n\right)$ *and* $\left([y_i^{DEC-5th}, y_i^{DEC-95th}],\ i = 1...n\right)$ *where* $y_i^{prior-5th}$ *and* $y_i^{prior-95th}$ *are the* 5^{th} *and*

the 95th percentile of the 5000 model simulation values at time i, and where $y_i^{DEC-5th}$ and $y_i^{DEC-95th}$ are the 5th and the 95th percentile of the same but weighted series according to the DEC calibration criterion.
Those confidence intervals were standardized according to the DEC modelling error definition (equation 10), respectively defining the a priori and a posteriori confidence intervals of the modelling errors:

$$\epsilon_i^{\alpha-xth} = \{ 0 \text{ if } \quad | y_i^{\alpha-xth} | \leq 2 \cdot \sigma_{\hat{y}_i} \frac{y_i^{\alpha-xth} \pm 2 \cdot \sigma_{\hat{y}_i}}{2 \cdot \sigma_{mod_i}} \text{ otherwise } \quad (- \text{ if } y_i^{\alpha-xth} > 0 \,; + \text{ if } y_i^{\alpha-xth} \leq 0\,) \tag{1}$$

with $\epsilon_i^{\alpha-xth}$ is the x^{th} percentile of the α modelling errors distribution at time i.
The latter definition allows for an informative translation of the prior and posterior confidence intervals (Douinot et al. 2017): a value of $\epsilon_i^{\alpha-xth}$ equal to 0 indicates that the $y_i^{\alpha-xth}$ bound lies within the discharge confidence interval; if $0 < \epsilon_i^{\alpha-xth} \leq 1$, the $y_i^{\alpha-xth}$ bound lies within the acceptability zone; and if $\epsilon_i^{\alpha-xth}$ is larger than 1 then errors of modelling is detected or remained. In addition, the benchmark of both a priori and a posteriori confidence intervals allows for highlighting which were the remaining modelling errors that were induced by the model's assumptions. For those reasons, $\epsilon_i^{\alpha-xth}$ were used as the baseline of the modelling errors analysis.

Comments about the section 5 (results)

According to the comments, the presentation of the results has been reorganized as follow:

- section 5.1: Performance of the models. In these section are exclusively presented the assessment of the models through the metric scores (that are defined in the section 4.2).

Printer-friendly version

Discussion paper

  - section 5.1.1: Overall performances of the models. It merges the paragraph that has been written into the previous section 5.1.1 and 5.1.2
  - section 5.1.2: Detailed performances of the models: assessment of the models when simulating the different stages of an hydrograph. It contains the previous section 5.1.3 and 5.1.4

- section 5.1.3: Summary of the assessment: This part has been added, in order to present a global overview of the results after detailed comments in the aforementioned sections.

- section 5.2: Modelling errors inherent in the models'structures: It contain the previous 5.1.5 section.

- section 5.3: Analysis of relevance of the internal hydrological processes simulated: As suggested, the previous sections 5.2 and 5.3 have been merged into one element.

- section 6: Discussion: We propose a novel section in oder to separate the strict description of the results (section 5), and the interpretation done from it (section 6). It finally contains the previous sections 5.1.6 and 5.4.

The presentation of modelling results (section 5.1) could also be enhanced : I think the separated presentation of each metric (overall hydrograph, rising limb, high-discharges, recession) does not help to give a synthetic overview of results. It seems that three main situations can be distinguished here: clear hierarchy (Gardons and Salz), contrasted hierarchy (Ardèche), or no clear hierarchy (equi- finality, Hérault). These three situations could be illustrated based on a common analysis of all the metrics. The comments about the performances of the models were mainly reworded. In the novel version, we separated the presentation of global metric (overall hydrograph, section 5.1.1) from those that are focused on a specific stage of the hydrograph (section 5.1.2). The objective of this organization is to highlight the differences between what we learned with a global point of view, and what we learn if we focus on the representation of one part of the hydrograph. In addition, the figures 8 and 10 were modified in order to support the new organization of the section. The figure 8 presents only the global performances while the figure 10 presents the detailed performances. Finally, for a sake of clarity, a summary is done in a last subsection (section 5.1.3). Those modifications can be find from page 18 line 1, to page 23 line 22.

Explanations in section 5.1.4 (now 5.1.2) are poorly supported by figure 10 in my opinion. Please try to clarify this section and figure. We modified the way to assess the good simulation of the flood recession, using another metric score. This is defined page 16-17 from lines 24 to line 4 :

"Conversely, Qmed_INT was not relevant for the evaluation of the capacity to reproduce recessions, because the calculation of this score during the recession interval strongly depends on performance at high discharges. Instead, we used the A_{slope} score defined in the equation 9. It calculates the average standard error in simulating the decreasing rate of the discharge during the flood recession interval. Through the consideration of the A_{slope} score here, it is assumed that the recession rate is a relevant feature of the catchment's hydrologic properties Troch et al., 2013; Kirchner, 2009. We therefore choose to make a visual comparison of the simulated and observed recession curves, $Q(t) = f\left(log(-\frac{dQ(t)}{dt})\right)$, which are characteristic of a catchment's hydraulic discharge properties.

$$A_{slope} = \frac{\sum_{i=k}^{l} |\frac{dy_i}{dt} - \frac{d\hat{y}_i}{dt}|}{\sum_{i=k}^{l} \frac{d\hat{y}_i}{dt}} \quad (2)$$

where $\frac{d\hat{y}_i}{dt}$ and $\frac{dy_i}{dt}$ are respectively the observed and the simulated recession rate at a time step i which belongs to the flood recession interval $(i = k...l)$.

The assessment focused on the simulation of the recession is then presented in a similar way than those on the simulation of the rising flood waters and the high flows. Consequently, the three assessments are presented in a same figure (figure 10).

The analysis proposed in section 5.1.5 (now 5.2) is also difficult to follow based on figure 11, which does not well illustrate in my opinion the differences in models behaviors. Figure 11 indeed is difficult to understand: $\sharp$ DDEC is not defined, the definition of prior and posterior errors is again missing. Moreover, is not clear why the width of the acceptability zone does not vary with y_i (not consistent with equation (7)). Please try to clarify this section and figure.
In order to clarify this figure, definition of the specific variables are written in the section 4.2, page 17, lines 7-22. In addition several comments were added to better link the underlying description of the figure with the possible interpretation of the models's performances (page 23-24).

Section 5.1.6: (now 6.1) I think this interpretation on catchment behaviors arrives too early here. I think it would be better to put this in the discussion section, and to present these analyses as plausible assumptions, according to the modelling results.
As suggested, the comment of the previous section 5.1.6 are now the basis of subsection 6,1 of the discussion section.

Section 5.2 (now 5.3) may be renamed in a more explicit way, such as: "Analysis of relevance of the internal hydrological processes simulated". It could include both considerations on proportion of surface runoff (current section 5.2), and detailed analysis of velocities and water contents in the case of Hérault (current section 5.3)
As said below, we incorporated the suggestion into the new organization of the result section.

Comments about the conclusion

"For each catchment, the best performing models were those where results reflected the available knowledge and observations on the overall hydrological functioning of the catchments ...". Actually, it seems that very limited information is available on the real hydrological behavior, excepted maybe for the Gardon where detailed measurements were performed. Therefore, I would rather conclude that the modelling results help to draw consistent assumptions on hydrological behaviors, that can in some (rare) cases be confirmed by the existing knowledge and local observations.

As suggested this conclusion was reworded, taking into consideration that we actually have very limited information on those catchment (hence the interest of the study) page 35, lines 14-16 :

"The modelling results help to draw consistent assumptions on hydrological behaviours, which corroborate when available, the knowledge and observations on the overall hydrological functioning of the catchments, or the experimental estimations of flow processes."

"distinction in hydrological behaviours revealed between the catchment of the Gardons and the Ardèche may explain that taking into account the spatial nature of precipitation in a flash flood forecasting method results in an improvement only on the Gardon and not on the Ardèche ..." I think this conclusion is not really supported by the content of the paper. Moreover, another explanation could just be a difference in the rainfall spatial variability, which seems to be more pronounced in the Gardons catchment for climatic reasons.

This last statement is introduced as an open conclusion about the potential value of the study results facing to the flash flood forecasting issue. Revealing the contrasted hydrological behaviours of the Gardons and the Ardèche catchments - the first one clearly more reactive that the second one – it might shift towards different considerations when setting up a flash flood forecasting method over those contrasted catchments. We referred to the Douinot et al, 2016 study as it actually corroborates the fact that different considerations should be done, to develop a flood forecasting method. It shows contrasted sensitivities of the catchments to the rainfall spatial variability, which could either be a consequence of the contrasted hydrological behaviours of the catchments revealed here, or – we agree - be due to contrasted climatic forcing. We suggested to reword the statement as following (page 36, lines 1-12):

“Lastly, identifying the most pertinent hydrological models for each catchment enables the key elements in the generation of flash floods to be highlighted, which, in turn, could serve to further develop methods for forecasting flash floods. For example, distinctions in hydrological behaviour revealed between the catchments of the Gardon and the Ardèche - the first one appearing more reactive with important runoff and subsurface flows through preferential flowpaths - might shift towards different considerations when setting up a flash flood forecasting method. It corroborates the results of Douinot et al 2016, which highlighted contrasted impacts of taking into account the spatial variability of precipitation in a flash flood forecasting method. These contrasted impacts can indeed be explained by the more pronounced spatial variability of the rainfall over the Gardon catchment, but also by the local more pronounced dynamic of the soil water content in the Gardon catchment revealed in the present study."

Technical corrections

- Section 2.2: The reference Ministère de l’Ecologie (2015) just corresponds to an URL, which could be added directly in the text. The modification has been done (page 9, table 1 and page 10, line 2).

- “These measurements were calibrated by forecaters at the French Flood Forecasting service by monitoring a network of rain gauges ...". Sentence not clear, please reformulate. The sentence has been reworded, page 10, lines 13-15, as follow : *“The French Flood Forecasting Service used the CALAMAR software to produce the rainfall depth inputs of the model by combining these radar mea-*

Printer-friendly version

Discussion paper

surements with raingauge data."

- Figure 5: are θ_s and θ_i really the current and initial water contents respectively. Shouldn't rather θ_s be the saturation water content? This is true. The error has been corrected (figure 5, page 12).
- Section 3.2, description of the modelling principles: the equations (1) to (4) and description of variables should be placed in the text with reference to figures 6 and 7. The equations and the description of the variables were inserted into the text: page 13, lines 12-15 and page 14, lines 1-8; 16.
- Section 4.2: σ y and σ mod rather than Σ y and Σ mod. The modification has been done on page 16, line 26.
- Section 5.1.5: "the variation interval of the modelling errors": I don't really understand, please define this. The expression has been reworded, page 23, line 10 as follow: *"the confidence interval of the modelling errors"*.
- I finally suggest to check the overall quality of English. The text has been proofread by a professional translator.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-710/hess-2017-710-SC1-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-710, 2017.