Dear Editor

Dear Reviewer

Authors answers to the anonymous reviewer_1 comments

We thank the reviewer for his comments. We hereafter reply and/or clarify the reviewer comments point by point.

Reviewer Comment: This paper presents an approach to transfer model parameters from the gauged catchments to ungauged catchments based on the similarity between donor and receptor catchments. This approach was implemented to a SWAT application in the Mediterranean catchments. Parameter uncertainty and prediction uncertainty were studied and discussed. This paper is suggested to go through a major change before accepted based on the major comments below.

Major comments:

1. The English is too descriptive (not scientific) and could be shortened

a. Highlight the approach. It took me a while to understand the procedure

b. Some texts which are not the main focus could be removed: e.g., too much sentences on GLUE (e.g., comments), attributing wide uncertainty in baseflow to Karst (actually the main reason is the objective function NS which is favorite of the High flows) while this paper has nothing to do with the Karst, etc.

Authors answers:

1. The English is too descriptive (not scientific) and could be shortened.

- Authors answer: English will be revised. A shorter and more scientifically English will be used.

a. Highlight the approach. It took me a while to understand the procedure

- Authors answer: Thank you for this constructive comment. We have more clearly presented and highlighted the approach. The text from Line 25, Page 18 to Line 26, Page 19 was changed as follow:

"*In this section we propose a more objective method for selecting the appropriate Mps to be transferred from the gauged to the ungauged catchment. First, the similarity measure ($R^2_{(d,r)}$) between all possible attributes of gauged and ungauged catchment in the CAs dimension, is calculated (data not shown) and clusters with similar catchments are constructed. Then, the SWAT model is implemented and parameterized at each catchment based on the SWAT pre-processing procedure with the available data while model parameter calibration and uncertainty analysis are conducted simultaneously only at the gauged catchments (donor catchments) using the GLUE approach. At this stage, only Mps sets that led to positive NS values between observations and simulations at the gauged catchment are retained. However, this threshold value is updated in Eq.(3) based on the similarity measure between the donor and the receptor catchment and the candidates Mps to be transferred are identified accordingly:*

$$Thresh_{(d,r)} = R^2_{(d,r)} \times \max NS_d \qquad (3)$$

where $R^2_{(d,r)}$ *is the similarity measure between the donor catchment (d) and the receptor catchment (r) and scaled between 0 and 1 and* NS_d *is the highest likelihood value reached in the model simulations at the donor catchment (d).*

By applying Eq.(3) the number of the candidate Mps will increase linearly as the dissimilarity between the donor(s) and the receptor(s) catchment(s) increases. Furthermore, besides parameter uncertainty, additional uncertainty related to the regionalization schemes is explicitly accounted in the final model prediction uncertainty at the ungauged catchment(s) by introducing the similarity measure in Eq.(3). As the dissimilarity between the donor(s) and the target catchment(s) increases, model prediction uncertainty in the target catchment(s) intuitively increases and vice versa. Another advantage of using Eq.(3) is that the selection of the threshold value to define the number of the candidate Mps is based on the similarity metric rather than on a subjective choice of the modeler which may reduce this additional uncertainty component in the final regionalization procedure.

Because updating manually the parameter values in the text SWAT file is a time consuming and tedious task, a sampling and rewriting program in the MATLAB® computing language was developed that provides the Mps from the donor catchment to the receptor catchment in the SWAT model format".

Reviewer Comment

> b. Some texts which are not the main focus could be removed: e.g., too much sentences on GLUE (e.g., comments), attributing wide uncertainty in baseflow to Karst (actually the main reason is the objective function NS which is favorite of the High flows) while this paper has nothing to do with the Karst, etc.

Authors answer

We agree with the reviewer that due to the selection of the NS as objective function, the parameter estimation algorithm will lead to smaller prediction uncertainties for peak flow as compared to prediction uncertainties for base flow. We would like to add also that the presence of the karst features in the study area and the low performances of the SWAT model reported in the literature when applied in karst-fed catchments (Sellami et al., 2013; Spruill et al., 2000; Coffey et al., 2004; Benham et al., 2006) contribute to the baseflow prediction uncertainty. However, we also agree with the reviewer that the karst features in the catchment are of little relevance for the overall uncertainty assessment "this paper has nothing to do with the Karst") and, therefore, the text related to the discussion of the karst feature effect on the predicted baseflow uncertainty (from Line 24, Page 22 to Line 24, page 23) is removed.

Reviewer Comments:

2. Validity of the proposed technique

> a. The threshold of objective function NS. The authors chose NS >0. I doubt about this. In the literature, suggested "NS"s are greater than 0.5 or 0.6 for daily flow otherwise the model should be improved. When NS = 0, it means the simulation is no better than the average observed value. Low NS leads to wrong explanation of model behavior and uncertainty analysis.

Authors answers:

We agree with the reviewer that in literature the suggested NS value for separating "behavioral" from "non-behavioral" GLUE simulations is usually set to NS $\geq$ 0.5 or 0.6. However, it remains a subjective choice and may be the most criticized point when applying GLUE for assessing the modeling uncertainty. This critical point in the use of the standard GLUE methodology was clearly mentioned in the paper (Line 10 to Line 25, Page 15). In contrast to the standard GLUE approach, our methodology aims to provide a more objective method for selecting the model parameter sets to be transferred from gauged to ungauged catchment. The threshold value NS is set > 0 only at the gauged catchment to have enough parameter sets to be transferred. This implies, that we relax the condition for transferring Mps only in the initial Mps transfer stage; hence we transfer behavioural and less behavioural Mps, and we do not use an arbitrarily crisp threshold to demarcate behavioural and non-behavioural Mps. Then, the equation (3) is used to update this threshold value which is function of the similarity value between the donor and the receptor catchment and the highest NS value obtained from the GLUE simulations at the donor catchment. Therefore, the actual threshold value used to select the model parameter sets to be transferred to the ungauged catchment is selected objectively and is in our case generally higher than 0.5 or 0.6 (Please see Table 2).

Reviewer Comments:

b. There is no validation process of this technique. The validity is not sure.

Authors answer

Literature on model validation is very abundant in the hydrological science community, and it is widely accepted that different definition or levels of validation exist. We do not claim that the approach presented in this paper is strongly validated by comparison with independent observations. Indeed, for the considered case study, due to the lack of the observation data, a strong experimental validation could not be performed. We stated in the paper "*In the current work, catchments have very scarce streamflow records. Therefore any available observation data, field knowledge and/or previous work conducted in the area of interest can be precious and helpful to check the performance of the adopted regionalization method. Performance assessment of the regionalization procedure is based on three evaluation criteria*" (Line 1 to Line 5, Page 21). We preferred to use "performance evaluation of the regionalization technique" rather than "validation of the regionalization technique" to emphasize that the experimental validation level through independent observations is rather low. Nevertheless, we have conducted a performance evaluation of the approach based on three evaluation criteria, referred to as "fit to observations", "fit to reality" and "fit to geography". In this performance assessment we optimally used all available information in the scarce data catchment to make a performance assessment. We are aware that the approach and the assumptions behind the performance assessment can be improved if more and better data becomes available. We clearly cited this in the text "*However, this assumption is far to be validated in this work and needs to be further investigated with larger number of similar catchments or by simply gauging the catchments*" (Line 20 to Line 22 page 28).

Reviewer Comments:

c. A comparison could be made to the following approach:
1) Parameterize the SWAT with available DEM, landuse, soil and climate data for all the catchments based on SWAT pre-process procedure.
2) Apply GLUE with SWAT runs on all the catchments at the same time.
Parameters which are "behavioral" for the two gauged catchments are behavioral to other ungauged catchments
3) compare the result of this approach with proposed approach by author.

Authors answer We thank the reviewer for his suggestion.

1) The suggested approach of the reviewer starts by "Parameterize the SWAT with available DEM, landuse, soil and climate data for all the catchments based on SWAT pre-process procedure". This is actually what has been done in our approach. We have added in the text: "*… the SWAT model is implemented and parameterized at each catchment based on the SWAT pre-processing procedure with*

the available data while model parameter…" (Please see author's response to reviewer comment 1. a. highlight the approach).

2) The second step in the approach proposed by the reviewer consists of retaining the "behavioral" parameters at the gauged catchments and applying them at the ungauged catchments under the assumption that parameters which are "behavioral" for the two gauged catchments are behavioral to other ungauged catchments. In fact, we have already referred to this approach in our paper as the "traditional approach" (Line 15 to Line 18, Page 18). As we have already explained in the text, the traditional approach consists of selecting the "behavioral" parameters above a subjective cutoff threshold value (e.g. NS >0.5 or 0.6) and transfer them to all the ungauged catchment. However, as we have stated in our text (Line 18 to Line 24, Page 18) "*…doing this way all the receptor(s) catchment(s) will receive equal number of Mps despite that they are not equally similar to the donor(s) catchment(s). This may overestimate the prediction uncertainty at the closest receptor(s) catchment(s) and may underestimate it at catchments that are further from the donor(s) catchment(s). Furthermore, the selection of the "behavioral" Mps is based on an arbitrary and entirely subjective choice of a threshold value which may add to the uncertainty of the final regionalization results"*. Therefore, the advantage of our proposed approach in comparison to the one suggested by the reviewer is that the threshold value to select the parameter sets to be transferred from the gauged to the ungauged catchments is more objective since it is based on the similarity value between these two catchments. This may reduce the additional uncertainty related to the modeler subjectivity in selecting the cutoff threshold. Furthermore, it allows the propagation of the uncertainty of the parameters in function of the similarity measure between the donor and the receptor catchments. Nevertheless, because we found the reviewer proposed approach interesting, we applied it but only for the Pallas catchments group and we confronted the results to these derived from our approach. We have proceeded as follow:
1- From the GLUE results in the Pallas catchment (donor catchment), we have selected a threshold value of NS ≥ 0.5 to discriminate between "behavioral" and "non-behavioral'' parameter sets (Mps). Then, the SWAT model is run with all the "behavioral" Mps at the ungauged catchment of the Pallas group (4 ungauged catchments, See Table 3). Finally, we compared the uncertainty interval for each catchment to the one obtained from our approach.
Table 1 gives the threshold value applied for selecting the Mps transferred from the gauged (Pallas) to the ungauged catchments for the approach suggested by the Reviewer and for the approach developed in this paper (our approach). The results of the predicted flow duration curve (FDCs) uncertainty interval derived from both approach are shown in Fig.1.

From Fig.1 it's clearly seen that in the Nègues_Vacques catchment our approach, based on the similarity value, suggests to transfer only about 17% of the Mps corresponding to a threshold value ≥ 0.66 at the donor catchment while about 90% of the Mps corresponding to a threshold value ≥ 0.5 , chosen subjectively, are transferred according to the approach suggested by the Reviewer. This can explain the larger prediction uncertainty interval obtained from the former approach in comparison to our approach. However, for the Fontanilles catchment, the opposite is obtained; our approach suggest larger prediction uncertainty than the one obtained by the Reviewer approach since our calculated threshold value (threshold value ≥ 0.38), based on similarity between the Pallas and this catchment, is lower than the one chosen subjectively (threshold value ≥ 0.50) according the Reviewer approach. For the Soupié and Aygues_Vacques, very close uncertainty interval are obtained using both approaches since the threshold values obtained using both approaches are quite similar (Table 1).

The results of this exercise clearly show how uncertainty in the transferrable Mps can be propagated to the ungauged catchment function at the similarity distance that separates the donor and the receptor catchment and how this is different to the traditional approach suggested by the Reviewer. This comparison supports the assumptions behind our approach; model prediction uncertainty at the ungauged catchments increases as the dissimilarity between the donor and the receptor catchment increases. Our approach is appealing and reasonable and provides more objective prediction uncertainty at the ungauged catchment than the traditional approach. Although such comparison is interesting, we don't think that is necessary to include in the text since the aim of the paper is to develop more objective approach in selecting the transferrable model parameters for estimating the discharge at the ungauged catchments using regionalization technique rather than comparing the developed approach to the traditional one.

Minor comments

Reviewer Comments

1. The equation (1) is not correct

Authors answers. We thank the reviewer for his comments.

1. Equation (1) is changed and added to the text as follow:

$$\frac{\partial SW}{\partial t} = P_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw} \quad (1)$$

where SW is the soil water content (mm), P_{day} is precipitation rate (mm/day), Q_{surf} is the surface runoff rate (mm/day), E_a is evapotranspiration rate (mm/day), W_{seep} is the water percolation rate from the soil profile (mm/day), and Q_{gw} is the groundwater flow rate (mm/day).

Reviewer Comments

2. Line 7 of Page 4966: Should “a selection” be “the selection”?

Authors answers.

2. Line 7 of Page 4966: “a selection” is changed by “the selection”.

Reviewer Comments

3. Line 13 of Page 4966: what are the references in “in the literature”? Actually, weather data are the driving force

Authors answers

3. References corresponding to “in literature” in Line 13 Page 4966 are added and the text is changed as follow : “*These CAs are generally considered as the main drivers of the hydrological process in the literature (Merz and Blöschl, 2004; Heuvelmans et al., 2006; Wagener et al., 2007; Bastola et al., 2008) and are the …*”. The reviewer is correct to state that weather data are the driving forces. But as we already have mentioned in the text climatic descriptors are omitted since we are dealing with small and geographically close catchments located within a relatively small area under the same climate regime. (Please refer to Line 15 to Line 21 page 4967).

Reviewer Comments

4. More scientific explanation on “% clay, %silt, %sand” in line10 of page 4967

Authors answers

4. Line10 Page 4967 “The soil characteristics are based on the dominant soil texture (% clay, % silt, %

sand) within each catchment". The sentence is changed to "*Within each catchment, the dominant soil physical texture based on the relative proportion of sand, silt and clay is considered to identify the CA related to soil type"*.

Reviewer Comments

5. Rewrite lines 1-7 of page 4968. Lines are not well written.

Authors answers

5. Lines 1-7 Page 4968 were changed as follow: "*To identify similar catchments groups, each catchment is assigned to its own cluster and the similarity matrix between clusters, in the catchment attributes dimension, is calculated. Then, clusters with the largest similarity measure are linked together into binary clusters based on the average linkage method where the distance between two clusters is defined as the average distance between all objects belonging to these clusters. These steps are repeated and the similarity matrix between clusters is updated until all clusters are linked together in a hierarchical tree. The Pearson's correlation coefficient, denoted hereafter as R^2, is used as a similarity metric between catchment attributes; the higher the R^2 between the target and the donor catchments, the more similar they are."*

Reviewer Comments

6. Line 11 of page 4972, I doubt about it.

Authors answers

6. GLUE prediction uncertainty interval isn't able to bracket most of the observation data or the specified quantities of observations with a specific frequency (e.g. 95%) because it's assumed that all uncertainty sources (e.g. inputs, outputs, model errors) are mostly reproduced by parameter uncertainty. However, nor the inputs, outputs or model errors are explicitly considered to derive prediction uncertainty

Reviewer Comments

7. Expect explanation of lines 4 to 6 of page 4974

Authors answers

7. This is a relevant question from the reviewer. Some sensitive parameters depicted by the LH-OAT method turned out to be less sensitive by the GLUE approach. This is because the LH-OAT is based on variation of one single parameter at a time while within GLUE the sensitivity is assessed by means of a combination of parameters. The combined parameter set in GLUE is considered to produce "behavioral" model runs, given the selected likelihood function and the acceptance threshold value.

Reviewer Comments

8. Line 17 of 4975, should "receptor(s) catchment(s)" be "receptor catchments"?

Authors answers

8. Line 17 Page 4975. "Receptor catchment(s)" is changed to "receptor catchments".

Reviewer Comments

9. Equation (6), why there is no evapotranspiration?

Authors answers

9. In Equation (6) the Water yield (WYLD in mm) is computed as the sum of the surface runoff (Surf_Q), lateral flow (Lat_Q), groundwater flow (GW_Q) diminished by the losses (TLosses). The WYLD shows the contribution of the different flow types to the total water budget and it's actually equal to the Precipitation (P) diminished by Evapotranspiration (ET).

WYLD (mm)= P (mm) – ET (mm)

However, if the model is building up storage in the soil profile (ΔSW) and in the shallow aquifer (ΔGW), these terms should be added to the previous equation to close the hydrological water balance:

WYLD (mm)= P (mm) – ET (mm) – (ΔSW+ΔGW).

Reviewer Comments

10. Tables 1 and 2, adjust numbers

Authors answers

10. Numbers in Table 1 and 2 were adjusted

Reviewer Comments

11. Figure 3, use normal dates instead of Julian dates

Authors answers

11. Figure 3, normal dates are used instead of Julian dates.

Reviewer Comments

12. Figure 10. Better texts for "WYLD", etc.

Authors answers

12. The reviewer suggests not to use the acronyms of WYLD, Surf_Q and GW_Q.
We think that the terms (WYLD, Surf_Q, GW_Q) used in the text to refer to the Water yield, surface runoff and groundwater flow, respectively, are the most appropriate ones because these terms are also used in the SWAT model output and in most of the published papers dealing with SWAT.

Table 1. Threshold value and % of the Mps transferred from the gauged to the ungauged catchments for the two approaches

	Donor catchment	Receptor catchment	Similarity	Threshold (*Thresh*)	% of Mps
Reviewer suggested approach		Nègues_Vacques	---	0.50	89.95
		Aygues_Vacques	---	0.50	89.95
	Pallas	Soupié	---	0.50	89.95
		Fontanilles	---	0.50	89.95
		Nègues_Vacques	0.88	0.66	16.60
Our approach	Pallas	Aygues_Vacques	0.71	0.54	85.16
		Soupié	0.70	0.53	86.47
		Fontanilles	0.50	0.38	96.52

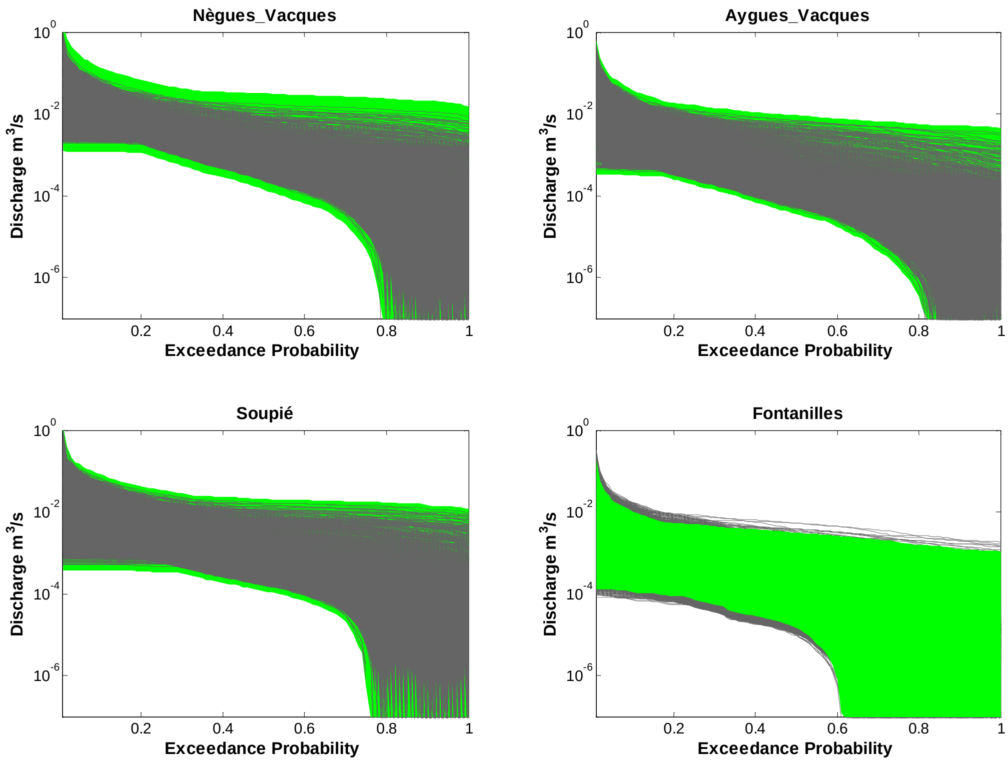

Fig.1. Predicted flow duration curve (FDCs) uncertainty interval derived from both approaches. The grey color refers to the uncertainty interval using our approach while green color refers to the uncertainty interval using the Reviewer suggested approach.