Reviewer 1

General comments:
This paper compared several kinds of regionalization strategies with two kinds of rainfall-runoff models. Generally speaking, this paper is well written and provides convincing evidence to support the conclusion. The title, abstract, introduction and discussion should be substantially revised. Here are some detailed comments.

Specific Comments:
1. The title is inconsistent with the numerical experiments of this paper. In this paper, 5 kinds of regionalization approaches (KR, NN-1, MS-1, NN-OA, MS-OA) were comprehensively compared with test data from Austria. The effects of nested catchments and gauging station density were also evaluated. Two rainfall-runoff models (TUW and GR6J) were used in the evaluation and similar results can be obtained from different models, indicating the conclusion of this work can potentially be extended to various models. However, the key words 'Importance of the information content' has not been intensively discussed in the discussion part (please search 'information'). The concept 'information content' has not been well defined in this paper, and it seems not very necessary to have it defined. This paper is more likely to be a 'inter-comparison' paper instead of a 'concept promoting' paper. Consequently, I suggest the authors to revise the title in order to sharply and concisely show what have been done in this paper in a few words.
Reply: We thank the referee for suggesting more clarity in highlighting what the paper presents: the main focus and novelty of the work is the analysis of the role of the available data set, that is which and how many gauged catchments are available (to be used as donors) for the regionalisation (this is what we mean with information content of the data set: we will explain it better). And in order to do so, we compare different regionalisation approaches, to understand which approaches are more or less impacted by a change in the donors' data set (and, as you wrote, we applied two well-known rainfall-runoff models, for generalisation purposes). We will modify the introduction and the conclusions sections to further clarify our main focus.

2. The abstract is too long! Please briefly introduce your work in a single paragraph. A wordy 'abstract' is not friendly to your potential readers. It is uncomfortable to locate useful information from a 'one-page, 5 paragraphs abstract'.
Reply: As suggested, we will substantially cut and revise the abstract, in order to have a more useful and agile summary (and also here focussing more on our main objective).

3. line 43-53: In order to clarify the relationships between different regionalization approaches, please add a 'mind map' about: regression-based, distance-based, output-averaging, parameter-averaging.
Reply: We will try to clarify it, reorganising the sentence in bullet points (so that is more 'graphical') in order to make the difference between the approaches easier to understand.

4. line 67: the best the best?
Reply: We will correct this typo.

5. line 155: routing routine?
Reply: Yes, it is the programming routine used for computing the flow response and routing (we agree that it is an alliteration and it does not "sound" well, but it is the same terminology used in the TUW model manual). However, we will consider replacing it with "routing module".

6. Discussion and conclusion part.
In this paper the hydrographs of simulated and observed streamflow were not presented. Although this is not necessary, a hydrograph may possibly intuitively show a lot of detailed mismatching patterns which are helpful for model diagnose. I suggest to add hydrographs from a typical catchment to compare not only the KGE and NSE criterions but also the details of systematic errors of different regionalization approaches. KGE and NSE

criterions are good objective functions for optimization, but hydrographs can tell more useful information about the physical processes.
Consequently, the discussion part of this paper is mainly about explaining the results, but lack of physical reasons. This paper has too much statistics but not enough hydrological mechanics on explaining the performance of different regionalization approaches. A comparison of regionalized parameters is strongly suggested to be added in the discussion part in order to provide fundamental physical reasoning of regionalization. For an example, please compare the parameter values from KR, NN and MS methods in a typical catchment, and discuss the influence of parameter values to corresponding hydrographs.
Reply: We understand your concern about the robustness of the regionalisation processes and their effect on the simulated hydrographs.
The analysis that you propose of comparing the regionalised parameters would be very interesting (and to our knowledge it would be novel) when using a regression-based method, where the parameters are estimated as a regression function of catchment characteristics and such function might allow to discuss the influence of regionalisation on the corresponding hydrographs: in such case we might expect to identify patterns of how the regionalised parameters vary across the study area, having a direct relationship with the attributes, which may contribute to systematic errors in the simulations.
But such comparison is not possible in our work because the regionalisation methods that we tested all belong to the "distance-based" class of approaches, which identify the model simulation at the target section based on the repetition of the simulation with one or multiple sets of parameters found for donor catchments (spatially closer in NN or most similar in MS) or based on the spatial correlation of the parameters (KR) and there is not a univocal influence of the regionalisation (or of the exclusion of nested donors or of the reduction of the density) on the parameters and therefore on the hydrographs.
We fully agree that reporting only goodness-of-fit indexes do not fully diagnose the capability of the model to accurately reproduce observations, but we did not show any hydrograph in the manuscript because we think that focusing on a single catchment could be misleading: the effect of the regionalisation, and the effect of the exclusion of nested catchments or of the density reduction on the approaches (that is our main focus) is strictly related to the single watershed and its donors, it varies based on the approaches and the specific topological relationships and characteristics of the catchment location.
As an example, we report below the analysis on the effect of the exclusion of nested donor basins when regionalising model parameters for one of the catchments in the Austrian dataset, and in particular we chose one target where excluding a nested donor resulted in a very strong deterioration for one of the methods, so that it is well visible on the hydrographs.
For the sake of brevity, the results regarding the only TUW model and the NN-1, MS-1 and KR regionalisation methods are reported. Moreover, just Criterion 2 is used for nested exclusion.
Figure R1.1 shows the position of the example target catchment *T* (yellow) and the selected donor basins according to the NN-1 and MS-1 method: basin *A* is the basin selected as donor for both methods NN-1 and MS-1 and it is nested to *T*. If we imagine to remove the availability of nested catchments, *A* would be no longer available and the methods would select respectively basin *B* (closest for NN-1) and *C* (most similar for MS-1). Of course, this representation is not feasible for the kriging method KR since it regionalises each model parameter independently, based on the parameters of all donors, and it is not based on the selection of one or more specific donors.

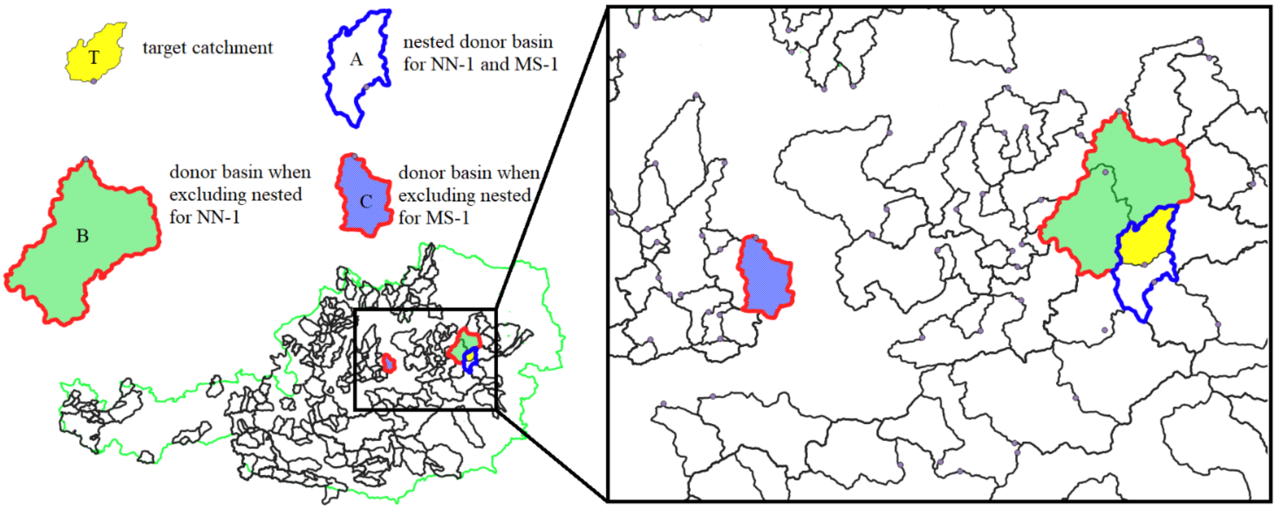

Figure R1.1. Location of the donor catchments for an example target basin *T*: nested basin *A* is the selected donor catchment for both methods NN-1 and MS-1, while if nested are not considered (following Criterion 2) basin *B* and *C* are selected as donor respectively from NN-1 and MS-1 approaches.

Figure R.1.2 reports observed and simulated hydrographs in year 1997 for the three approaches when including or excluding (Criterion 2) the nested donors. In addition, Table R.1.1 reports the parameter sets "at site" and regionalised. It can be noticed that, for this specific catchment, the method that deteriorates more when excluding the nested donor is MS-1 (panel c) whose simulated hydrograph (red line) differs substantially from observation (grey points) and from the regionalised simulation obtained when using the nested donor (blue line) as well, while for NN-1 and KR simulations are just slightly affected by the exclusion of A from the potential donors. The reason behind the strong degradation of MS-1 performances can be found looking at the change in the parameter set when excluding the nested donor *A*: the at-site parameter sets of the target and of its nested donor *A* are quite similar, while the parameters of donor *B* differ substantially; the smaller value of FC and k_1, along with a lower threshold L_{UZ}, lead to small storage capacity and thus to much more higher flow peak (red line in Figure R.1.2, panel c) in disagreement with observation. This does not happen, in this example, for the remaining two approaches where the regionalised parameter sets are still quite similar to the "at site" calibrated one, even when *A* is excluded.
The selected target catchment represents an unlucky case in which one of the method (MS-1) is deeply affected by the exclusion of the available nested donor, since the "most similar" among the not-nested donors, *C*, is evidently not so hydrologically similar to the target and its parameters do not allow a good simulation of the rainfall-runoff transformation in *T*. Anyway, we decided to display it to highlight the differences between the simulated hydrographs, otherwise not visible in less 'unlucky' cases.
A different catchment would include different multiple donors and different changes in the parameters and, consequently, in the hydrographs: the effect of the regionalisation and of the presence of nested donors is strictly related to the characteristic and the location of the single target catchment analysed. For these reasons, we would prefer to not include specific examples in the manuscript.
(And it should also be noted that this example shows the behaviour for the single donor methods, but if we analysed the output averaging methods, we would need to take into account other 3 (for NN-OA) plus 3 (for MS-OA) sets of parameters (and other 3+3 when excluding the nested donors), thus making this kind of representation even longer).

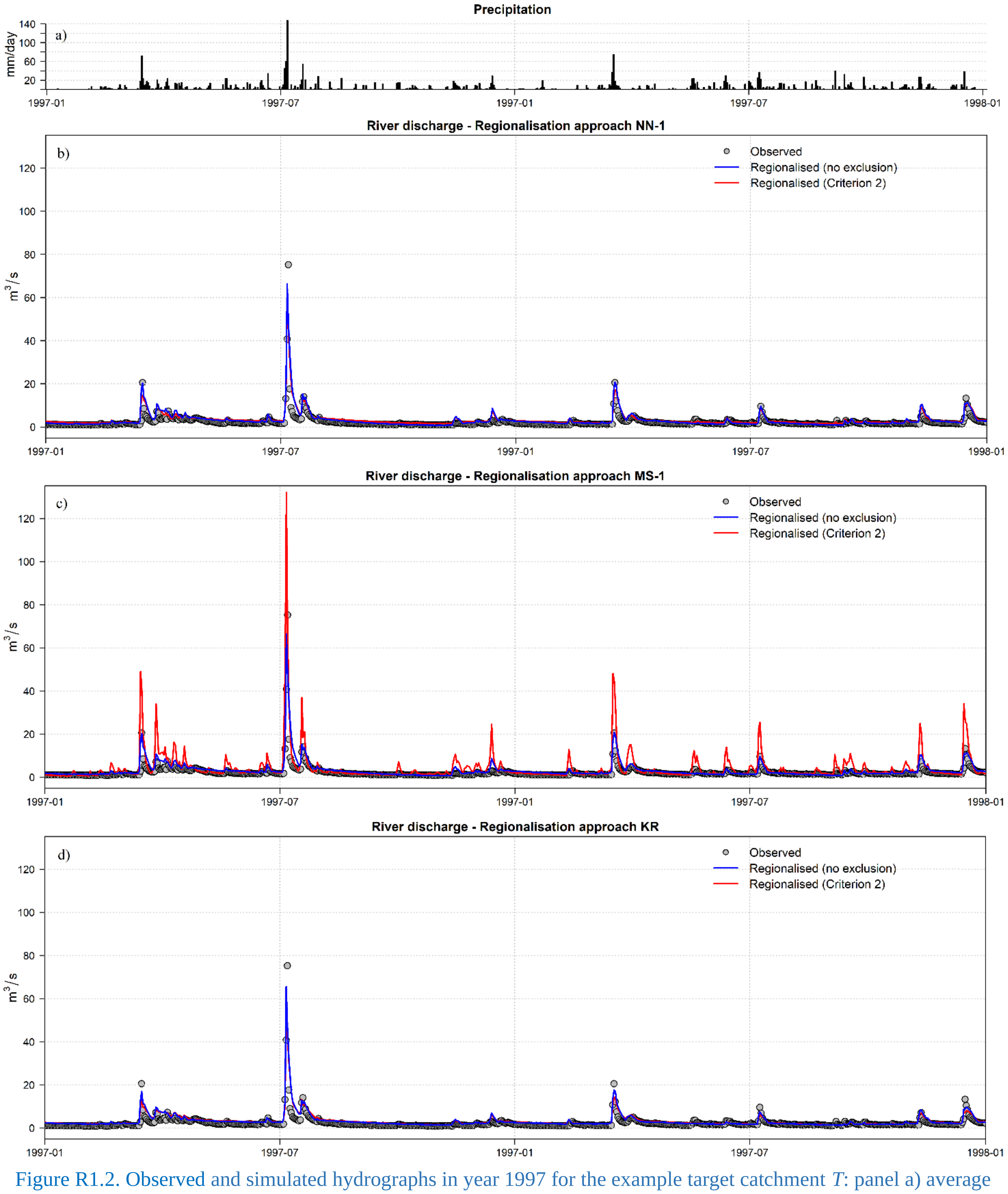

Figure R1.2. Observed and simulated hydrographs in year 1997 for the example target catchment *T*: panel a) average catchment precipitation, panel b c and d) observed streamflow (grey points) and simulated streamflow (respectively for approaches NN-1, MS-1 and KR) when nested catchments are available (blue line) or not (red line).

Table R1.1. Parameter sets for example target catchment *T*.

		NN-1		MS-1		KR	
	At Site	**Regionalized (no exclusion)**	**Regionalized (Criterion 2)**	**Regionalized (no exclusion)**	**Regionalized (Criterion 2)**	**Regionalized (no exclusion)**	**Regionalized (Criterion 2)**
SCF	0.9	0.9	0.9	0.9	1.2	0.9	0.9
DDF	1.9	2.4	3.1	2.4	3.4	2.5	2.5
LP	0.2	0.0	1.0	0.0	1.0	0.4	0.7
FC	285.1	114.4	150.7	114.4	33.5	177.7	137.1
β	1.0	20.0	20.0	20.0	1.3	15.7	13.2
k_0	0.3	0.3	0.4	0.3	1.9	0.4	0.4
k_1	16.1	6.3	8.0	6.3	2.3	9.0	12.5
k_2	86.8	51.0	73.6	51.0	30.0	88.8	84.9
L_{UZ}	42.8	65.0	87.1	65.0	18.8	71.4	76.3
C_{PERC}	1.2	4.7	5.8	4.7	1.9	4.6	4.7
C_{ROUTE}	49.7	19.3	43.5	19.3	16.0	32.6	43.9