On evaluating the robustness of spatial-proximity-based regionalization methods.

Laure Lebecherel, Vazken Andréassian, Charles Perrin

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

Technical note

On evaluating the robustness of spatial-proximity-based regionalization methods

Laure Lebecherel, Vazken Andréassian* and Charles Perrin
Irstea, Antony, France

** Corresponding author, vazken.andreassian@irstea.fr*

Abstract

This technical note deals with a question of importance for regionalization methods based on spatial proximity. These methods transfer hydrological information (typically calibrated parameter sets) from neighbor gaged stations to the target ungaged stations. The robustness of these regionalization methods (i.e., how their performance degrades when the hydrometric network becomes sparser) must be assessed. Here, we evaluate and compare two options for assessing this robustness: the random hydrometrical reduction (HRand) method, which consists in sub-sampling the existing gaging network around the target ungaged station, and the hydrometrical desert method (HDes), which consists in ignoring the closest gaged stations. Our tests show that the HDes method is a more conservative testing method than the HRand method.

Keywords

Rainfall-runoff modeling; ungaged basins; regionalization; spatial-proximity; robustness assessment;

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

1. Why is it important to assess the sensitivity of regionalization methods to the density of the hydrometric network?

Hydrological models with parameters that cannot be directly derived from physical catchment characteristics require calibration for parameter identification. Calibration is mostly based on observed flow series. Therefore, ungaged catchments, where no observed runoff data are available, require specific treatment. Much work has been done since the 1970s to handle the absence of runoff data (see e.g. James (1972); Magette et al. (1976)), and the corresponding approaches are usually referred to as regionalization approaches (Gottschalk et al., 1979). Recent advances on regionalization within the framework of the IAHS Prediction on Ungauged Basin (PUB) decade have been reviewed by Hrachowitz et al. (2013), showing how information can be transferred from gaged to ungaged catchments.

Among the commonly used regionalization approaches, some use the principle of physical similarity, based on the hypothesis that basins with similar physical characteristics have hydrologically similar responses (see Oudin et al. (2010)). Other approaches use information from the catchment's spatial neighborhood, based on the hypothesis that surrounding physical conditions are similar. In this paper, we will specifically focus on this second type of approach, the efficiency of which strongly depends on the density of the hydrometric network.

One of the important expected properties for a regionalization method is robustness. Two regionalization methods could perform very similarly in a data-rich environment and perform much differently under conditions of more limited data availability: assessing the sensitivity of any regionalization method to the level of information availability (typically the density of the surrounding hydrometric network in case of proximity-based approaches) is a good way to avoid disappointments when comparing academic methods to real-world data (Andréassian et al., 2010). Operational networks are rarely as dense as we hydrologists wish they were…

Surprisingly, the impact of hydrometric data density on regionalization efficiency does not seem to be a matter of concern in the literature. We addressed this issue in a previous study on the regionalization of the Turc-Mezentsev regionalization formula (Lebecherel et al., 2013). Here we would like to defend the point of view that this sensitivity test is not a mere matter of "hydrological comfort" but rather a scientific necessity, a kind of essential "crash test" to ensure credibility before practical use (Andréassian et al., 2009).

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

This note explores spatial-proximity-based regionalization methods. It proposes and compares two methods to evaluate the impact of hydrometric network density on regionalization performance. We start with a description of the material in section 2: a data set of 609 French catchments, a rainfall-runoff model and a spatial proximity-based regionalization method. Section 3 presents the two alternative methods to evaluate the robustness of a regionalization method: the hydrometrical random reduction (HRand) method and the hydrometrical desert method (HDes). Finally, the two methods are compared.

2. Material

2.1 Catchment set

The two evaluation methods of regionalization robustness were tested on a data set consisting of 609 French catchments (Figure 1), where daily streamflow, rainfall and potential evaporation time series were available over the 1997–2006 period. These catchments are spread over France and encompass a variety of climatic conditions (oceanic, Mediterranean, continental).

Potential evaporation (PE) was computed using the Oudin formula (Oudin et al., 2005) and precipitation data come from a countrywide interpolation(SAFRAN reanalysis) of rain-gage data (Quintana-Segui et al., 2008; Vidal et al., 2010).

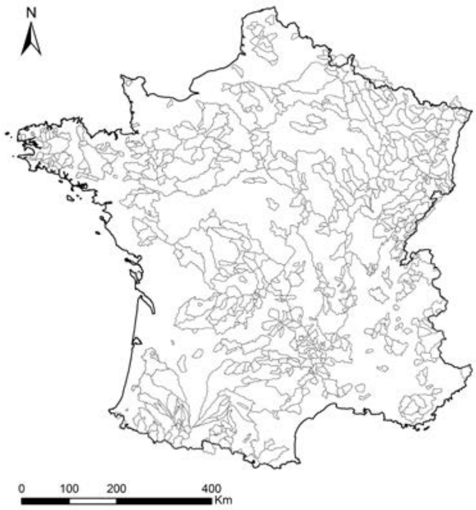

Figure 1. Location of the 609 French catchments used in this study.

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

Table 1 gives the main characteristics of the data set in terms of catchment area, mean annual streamflow, precipitation and potential evaporation.

Table 1. Main characteristics of the data set

Quantiles	**0.05**	**0.25**	**0.5**	**0.75**	**0.95**
Area (km^2)	33	109	270	833	4515
Mean annual precipitation, P (mm/yr)	714	863	1003	1209	1688
Mean annual runoff, Q (mm/yr)	159	272	411	643	1308
Mean annual potential evaporation, PE (mm/yr)	533	616	655	687	782

2.2 Rainfall-runoff model and calibration procedure

The GR4J hydrological model (Perrin et al., 2003), a daily lumped continuous model with four free parameters, was used. The GR4J model parameters need to be calibrated (on gaged catchments) or transferred from neighbors (on ungaged catchments). The model has two stores: a production store, which computes effective rainfall, and a routing store combined with a unit hydrograph for water transfer. The input model data are rainfall (*P*) and potential evaporation (*PE*) data. We use on top of GR4J, an altitude-distributed snow accounting routine, Cemaneige (Valéry et al., 2014). Here, the two parameters of the routine are not calibrated and we use regionalized values.

The objective function used for GR4J optimization is a transformation of the Nash and Sutcliffe (1970) criterion (NS) computed on root square-transformed flows. The transformation proposed by Mathevet et al. (2006) (C2M) is chosen to compute meaningful mean model performance values over the entire catchment set and to avoid the bias introduced by highly negative NS values:

$$C2M = \frac{NS}{2 - NS} \qquad \text{Eq. 1}$$

Note that this criterion keeps the same zero reference as the Nash-Sutcliffe criterion (C2M = 0 when NS = 0), has the same optimum (1 means perfect simulation for both criteria), but yields lower positive values compared to the NS criterion (e.g., C2M = 0.67 when NS = 0.80).

The model was calibrated on each catchment over the available data using a steepest descent search algorithm that proved efficient for this model (Edijatno et al., 1999). Hence, for each catchment considered ungaged, 608 parameter sets were available as possible donors. Model performance in calibration will be later used as a reference to evaluate efficiency loss due to regionalization.

Obviously, the evaluation methods could be applied with other rainfall-runoff models, calibration procedures and objective functions.

2.3 Spatial-proximity-based regionalization method

Since the aim here is not to develop new regionalization methods but only to evaluate them, we used the existing method proposed by Oudin et al. (2008b). This spatial-proximity approach with the output averaging pooling option consists in transferring the GR4J parameter sets calibrated at the *n* closest neighbor catchments to the target ungaged catchment. Then *n* daily runoff series are simulated on the ungaged catchment successively using rainfall and PE data available for this catchment and each parameter set from the neighboring catchments. The final simulated runoff time series for the ungaged catchment is computed as the weighted average of the *n* runoff time series. Here, weights consist in the inverse squared distance between the ungaged catchment and the gaged catchment:

$$Q_{ungaged\ catchment} = \frac{\sum_{i=1}^{n}\left(Q_{\theta i} \times \frac{1}{d_i^2}\right)}{\sum_{i=1}^{n}\frac{1}{d_i^2}} \qquad \text{Eq. 2}$$

with $Q_{\theta i}$ the runoff of the ungaged catchment simulated with the parameter set θ_i of the neighboring catchment *i* and d_i the distance between the ungaged and the neighboring catchment *i*.

McIntyre et al. (2005) and Oudin et al. (2008b) showed that this method is more coherent than the parameter averaging method since it transfers whole parameter sets from the gaged to the ungaged catchments.

Based on preliminary tests, we chose the following modalities of the regionalization method:

- *n*=10: we limit the parameter transfer to the 10 closest neighbors;
- the distance between catchments combines the distance to the outlet and the distance to the centroid (necessary to transfer information between catchments of different sizes), the distance considered in Eq. 2 is thus defined as:

  $$\mathrm{d} = 0.2 \times \mathrm{d}_{\mathrm{outlet}} + 0.8 \times \mathrm{d}_{\mathrm{centroid}} \qquad \text{Eq. 3}$$

- inverse distance weighting based on a squared distance.

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

3. Two alternatives for sensitivity analysis of a spatial-proximity-based regionalization method

3.1 Sensitivity analysis based on a random thinning of the hydrometrical network (HRand)

The random hydrometrical reduction (HRand) method consists in randomly removing 10%, 20%, 30%,…, 90% of the available hydrological network. The number of donor catchments remains the same (10), but they are located on average further from the receiver catchment (though the sampling may keep close neighbors).

The thinning procedure proceeds as follows: for each receiver catchment, the neighboring catchments are ranked from the closest to the most distant. For each neighboring catchment, a number x between 0 and 1 is drawn randomly and a threshold of acceptance x_{thresh} is considered: if x is larger than the threshold, the neighboring catchment is not used for regionalization. The sensitivity analysis starts with the full network (x_{thresh}=1) and moves progressively to a reduced network corresponding to 90% (x_{thresh}=0.9), 80%, … and 10% (x_{thresh}=0.1) of the initial network. The selection always yields the 10 closest catchments among the remaining donors, but for the reduced network, these donors are located farther on average than for the complete network. The random drawing was done once per catchment. Since the number of catchments is large, we believe that this does not preclude obtaining robust results.

This method aims at randomly thinning the network of donor catchments. Figure 2 illustrates an example of the impact of network thinning on the selection of neighboring catchments. It shows that although the neighboring catchments are located farther away on average as the thinning becomes stronger, some close neighbors can still remain in the selection, even at high thinning rates.

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

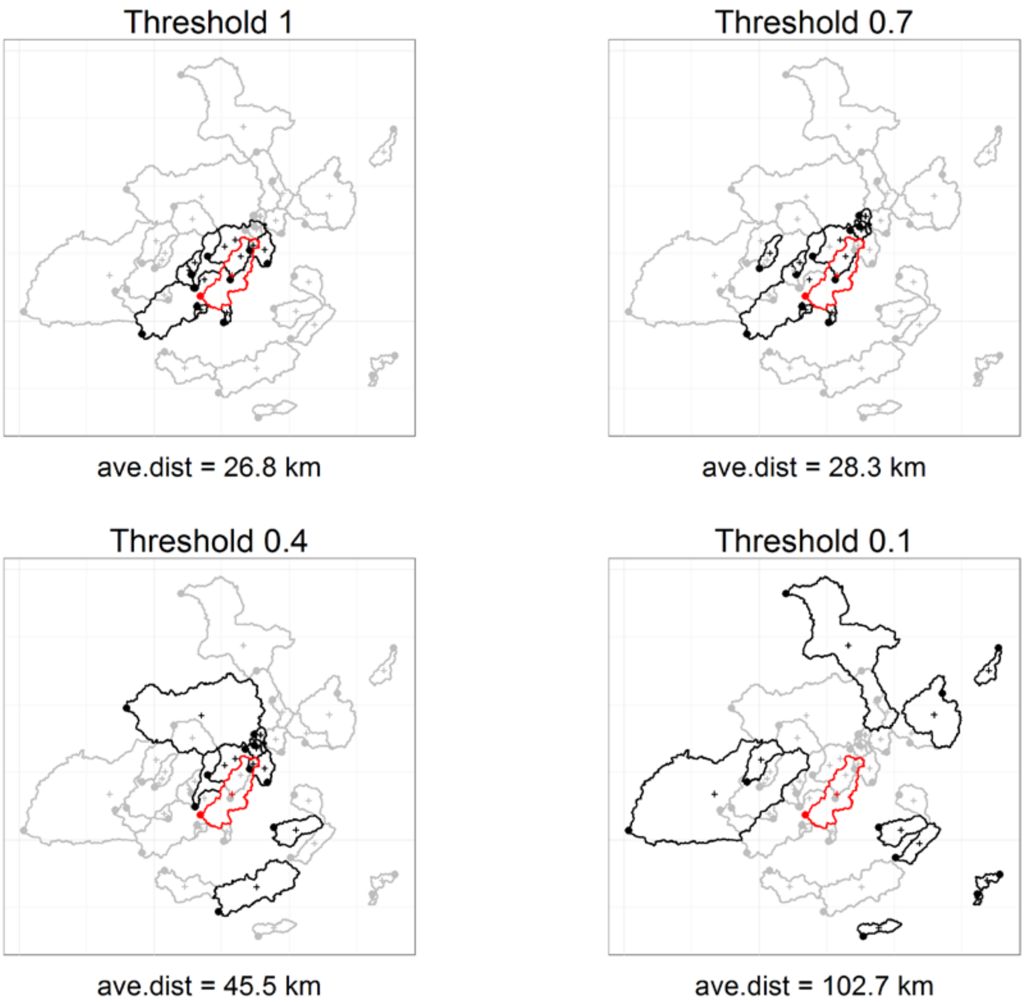

Figure 2. Illustration of the selection of neighboring catchments with the full network (Threshold = 1) and when applying the procedure of random reduction of the hydrometric network (HRand) for different thinning levels (Thresholds = 0.7, 0.4 and 0.1). In red, the ungaged (receiver) catchment, in black the selected neighboring (donor) catchments, and in grey the complete set of neighboring catchments. *ave.dist* is the average distance of the 10 donor catchments.

3.2 Sensitivity analysis based on the hydrometrical desert (HDes) method

The hydrometrical desert method (HDes) first proposed by Boldetti (2012) consists in progressively excluding the closest donor catchments: the parameters are transferred from neighbors that are increasingly distant from the ungaged target catchment by setting a lower limit below which neighbors are ignored (Figure 3). Here, we propose a test to analyze the sensitivity of the regionalization method

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

when information has to be transferred from progressively farther distances: we tested the following threshold distances: 0 (no distance limit), 10, 20, 30, …, 100, 150 and 200 km.

To measure the potential impact of such exclusion thresholds, we can compare them to the distribution of distances of the ten closest catchments for the 609 catchments of the dataset. Figure 4 shows that quantiles 10% and 90% of this distribution are of the order of 20 and 50 km, and the minimum value is around 15 km. It confirms the relatively high density of the French hydrometrical network. Then, the choice of the threshold distances for the nearest neighbor will clearly impact the choice of donor catchments.

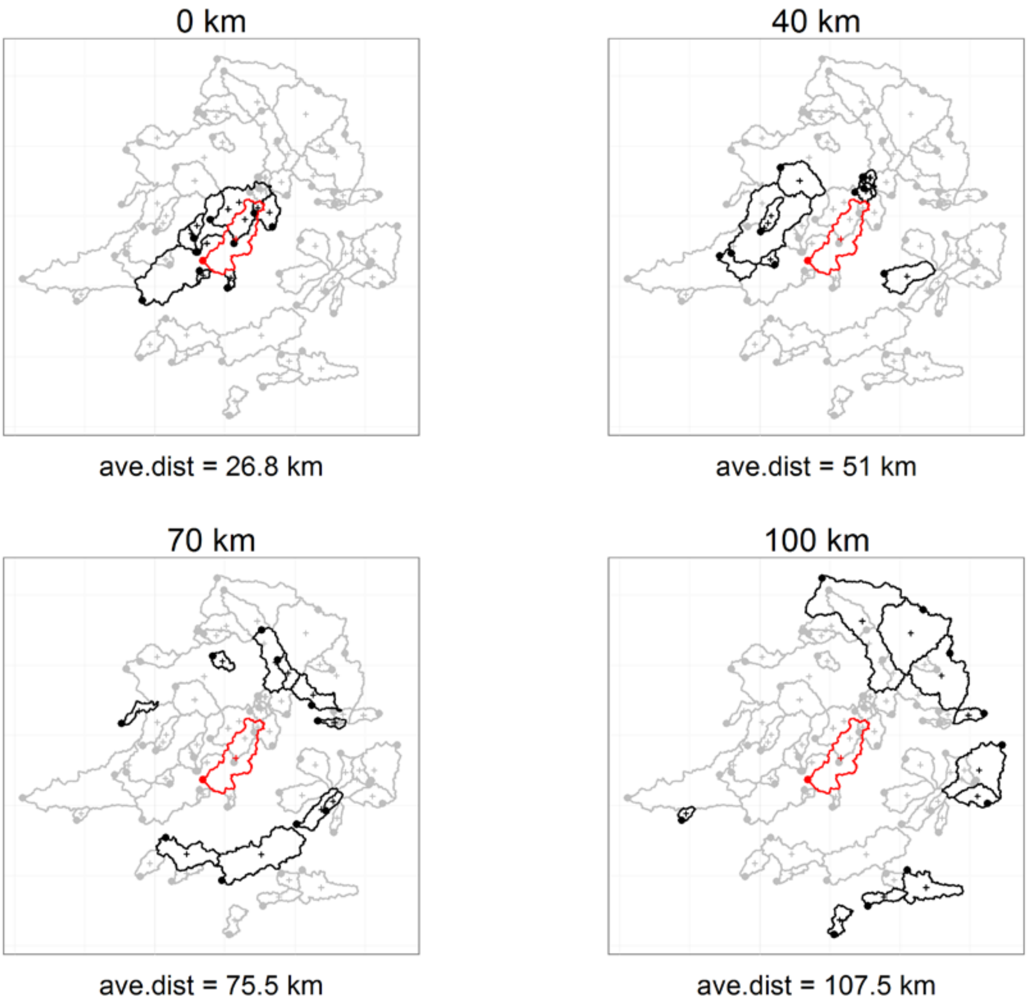

Figure 3. Illustration of the selection of neighboring catchments with the full network (distance = 0 km) and when applying the procedure of the hydrometrical desert (HDes) method for different limit distances (40, 70, 100 km). In red, the ungaged (receiver) catchment, in black the selected neighboring (donor) catchments, and in grey the complete set of neighboring catchments. ave.dist is the average distance of the 10 donor catchments.

Author-produced version of the article published in Journal of Hydrology (2016), vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

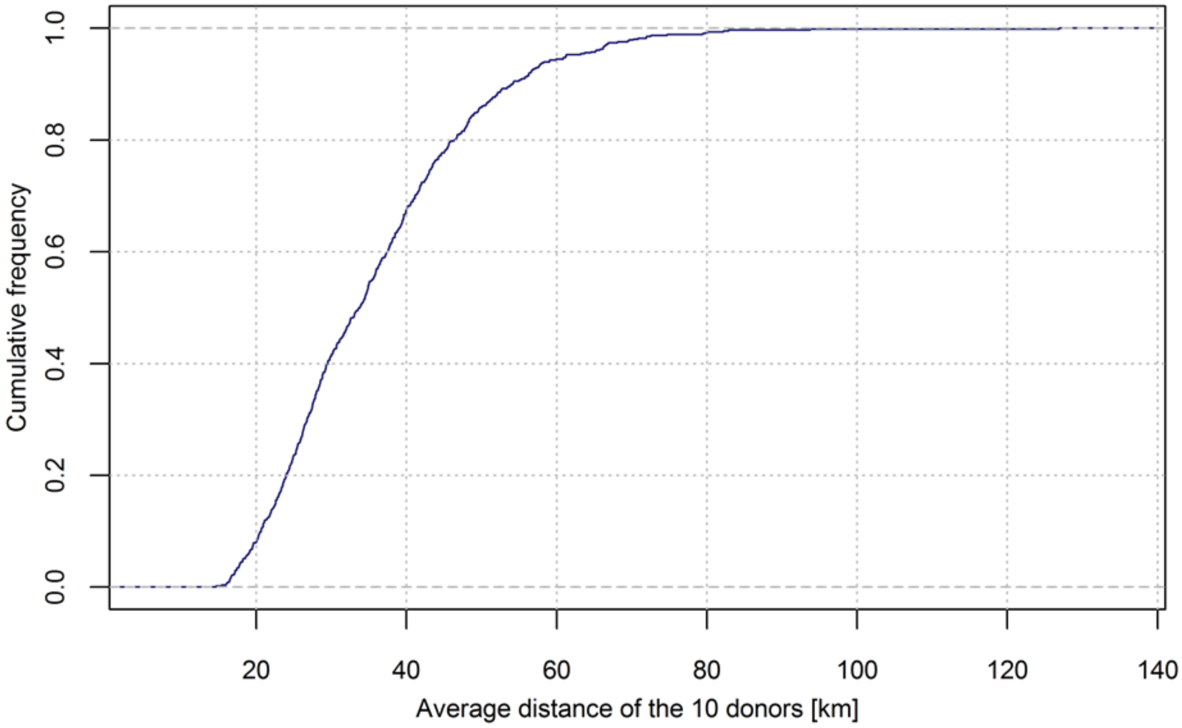

Figure 4. Distribution of the average distance of the 10 closest donors, for the 609 catchments.

3.3 Comparison of the two sensitivity analysis methods

The two sensitivity approaches HRand and HDes were applied to the regionalization of the GR4J model on the 609 catchments. Each catchment was successively considered ungauged. The regionalization approach was applied in each case, and model performance was evaluated using the C2M criterion. Figure 5 shows the performance distributions over the full catchment sample. As expected, the performance of the regionalized model, even with the full donor catchment set, is much lower than the calibrated model (see Oudin et al., 2008a).

The most obvious effect of applying the two sensitivity analysis methods is a clear decreasing trend in the efficiency of regionalization. The second important result is that the hydrometrical desert method (HDes) provides a more abrupt decrease than the random thinning (HRand) method.

The reason seems to lie in the fact that the HRand method allows situations where the first donor is quite close to the receiver (and potentially more similar and a better donor), as we can see in Figure 6. Because it forces the exclusion of the closest donors, the hydrological desert appears more

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

conservative.

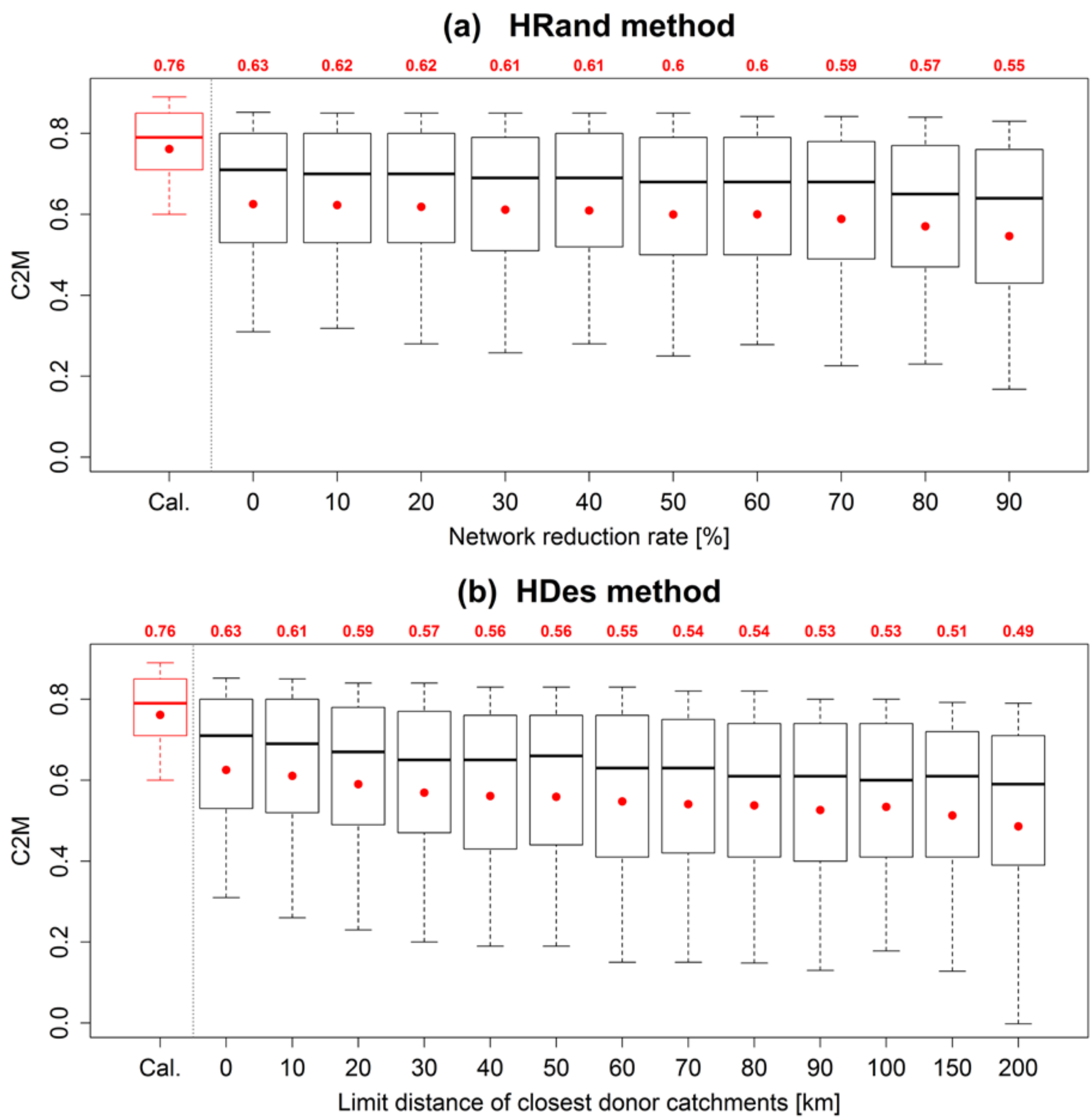

Figure 5. Distributions of performance of the calibrated and regionalized GR4J model on the 609 catchments, showing the impact of the two methods of robustness evaluation (a: random hydrometrical reduction method - HRand; b: the hydrometrical desert method - HDes). Calibration results (left) are used as a reference. Boxplots show the 10, 25, 50, 75 and 90 percentiles of the distribution from bottom to top. Mean performance is indicated by a dot and the corresponding value is shown on top of the graph.

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

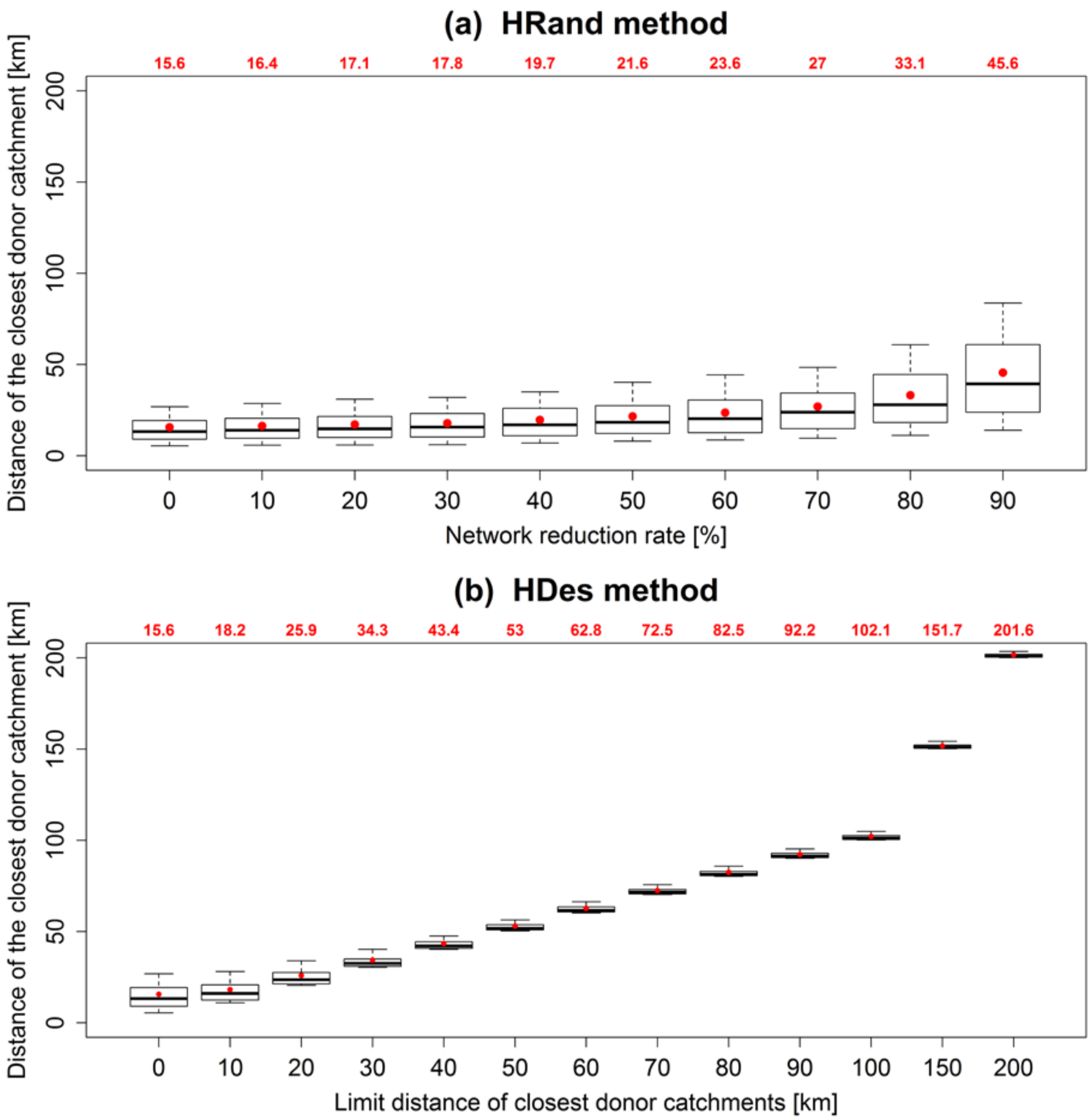

Figure 6. Distributions of the distance of the closest donor catchments for the two methods used to evaluate robustness of the regionalization method (a: random hydrometrical reduction method - HRand; b: the hydrometrical desert method - HDes). Boxplots show the 10, 25, 50, 75 and 90 percentiles of the distribution from bottom to top. Mean performance or mean distance is indicated by a dot and the corresponding value is shown on top of graphs a and b respectively.

However, the comparison between the two methods seems is not straightforward due to different x-axis. To make HRand and HDes methods comparable despite the different units (network reduction rate for HRand and limit distance of closest neighbors for HDes), we had to find a common x-axis for

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

both methods. We chose to take as common denominator the evolution of the average distance of the 10 donor catchments, which can be obtained for both methods. Then, we can take as a new x-axis (common to both methods) the average distance of 10 donor catchments. This will be associated with the values of the network reduction rate for HRand, and the limit distance of donor catchments for HDes. These results are presented in Figure 7. Thus, we can confirm with these graphics what was noted above. Actually, we observe a loss of performance, a little more abrupt for HDes method than for HRand method, particularly for the lowest average distances of the ten donors.

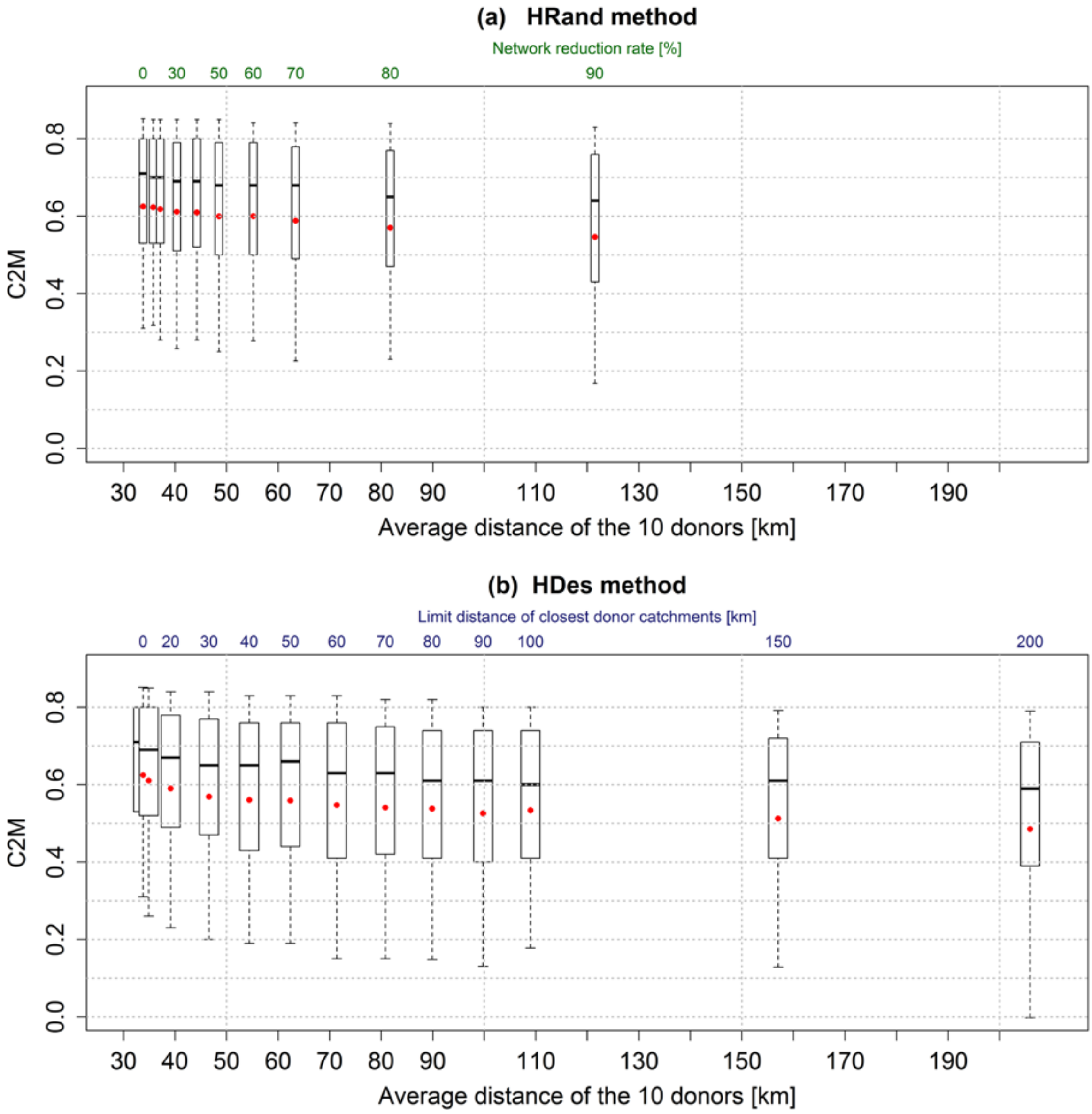

Figure 7. Distributions of performance of the regionalized GR4J model on the 609 catchments, according to the average distance of the 10 donor catchments for the two methods of robustness evaluation (a: random hydrometrical reduction method - HRand; b: the hydrometrical desert method - HDes). Corresponding values of network reduction rate for

Author-produced version of the article published in Journal of Hydrology (2016), vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

HRand method and corresponding values of limit distance for HDes method are indicated on top of the graphs. Boxplots show the 10, 25, 50, 75 and 90 percentiles of the distribution from bottom to top.

4. Conclusion

In this note, we proposed and compared two robustness evaluation methods: the random thinning method (HRand) and the hydrometrical desert method (HDes). Although both methods allow analyzing the sensitivity of regionalization methods to decreasing hydrometric data availability, we observe that the HDes method is the most conservative, since it produces the fastest decrease in model efficiency. From a hydrological "crash test" perspective, as discussed by Andréassian et al. (2009), we would recommend using the HDes method, which will provide a more realistic (although a more pessimistic view) of spatial-proximity-based regionalization efficiency. A further reason for this is that, for a practical application on a given ungaged catchment, it is always easier to compute the distance between this catchment and its closest neighbor than to assess the regional density of catchments. Once this distance is known, Figure 5-b can be used to give an indication of the expected value of GR4J regionalized efficiency, with an uncertainty interval on this expected value.

5. Acknowledgements

The authors wish to thank the different institutions that allowed us to build the large catchment sets on which this paper is based: MétéoFrance for meteorological data and SCHAPI for hydrometrical data. ONEMA provided financial support for the first author.

6. References

Andréassian, V. et al., 2009. HESS Opinions 'Crash tests for a standardized evaluation of hydrological models'. Hydrology and Earth System Sciences, 13(10): 1757-1764.

Andréassian, V., Perrin, C., Parent, E., Bardossy, A., 2010. The Court of Miracles of Hydrology: can failure stories contribute to hydrological science? Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 55(6): 849-856.

Boldetti, G., 2012. Estimation of the parameters of hydrological models on ungauged basins: a comparison of direct and indirect approaches. PhD Thesis, AgroParisTech - Irstea (Antony), 207 pp.

Edijatno, Nascimento, N.D., Yang, X.L., Makhlouf, Z., Michel, C., 1999. GR3J: a daily watershed model with three free parameters. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 44(2): 263-277.

Author-produced version of the article published in Journal of Hydrology (2016),
vol. 539, p.196-203. The original publication is available at http://www.sciencedirect.com/
doi : 10.1016/j.jhydrol.2016.05.031

Gottschalk, L., Jensen, J.L., Lundquist, D., Solantie, R., Tollan, A., 1979. Hydrologic regions in the Nordic countries. Nordic Hydrology, 10(5): 273-286.

Hrachowitz, M. et al., 2013. A decade of Predictions in Ungauged Basins (PUB)a review. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 58(6): 1198-1255.

James, L.D., 1972. Hydrologic modeling, parameter estimation, and watershed characteristics. Journal of Hydrology, 17(4): 283-307.

Lebecherel, L., Andreassian, V., Perrin, C., 2013. On regionalizing the Turc-Mezentsev water balance formula. Water Resources Research, 49(11): 7508-7517.

Magette, W.L., Shanholtz, V.O., Carr, J.C., 1976. Estimating selected parameters for Kentucky watershed model from waterched characteristics. Water Resources Research, 12(3): 472-476.

Mathevet, T., Michel, C., Andréassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS-AISH Publication(307): 211-219.

McIntyre, N., Lee, H., Wheater, H., Young, A., Wagener, T., 2005. Ensemble predictions of runoff in ungauged catchments. Water Resour. Res., 41: W12434, doi:10.1029/2005WR004289.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. Journal of Hydrology, 10(3): 282-290.

Oudin, L., Andreassian, V., Perrin, C., Michel, C., Le Moine, N., 2008a. Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. Water Resources Research, 44(3).

Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, N., 2008b. Spatial proximity, physical similarity, regression and ungaged catchments: a comparison of regionalization approaches based on 913 French catchments. Water Resources Research, 44: W03413, doi:10.1029/2007WR006240.

Oudin, L., Kay, A., Andréassian, V., Perrin, C., 2010. Are seemingly physically similar catchments truly hydrologically similar? Water Resources Research, 46(11): W11558.

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology, 279(1-4): 275-289.

Quintana-Segui, P. et al., 2008. Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France. Journal of Applied Meteorology and Climatology, 47(1): 92-107.

Valéry, A., Andréassian, V., Perrin, C., 2014. ‘As simple as possible but not simpler’: What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. Journal of Hydrology, 517: 1176-1187.

Vidal, J.P., Martin, E., Franchisteguy, L., Baillon, M., Soubeyroux, J.M., 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30(11): 1627-1644.