



HAL
open science

How should a rainfall-runoff model be parameterized in an almost ungauged catchment? A methodology tested on 609 catchments

Claudia Rojas-Serna, Laure Lebecherel, Charles Perrin, Vazken Andréassian, Ludovic Oudin

► To cite this version:

Claudia Rojas-Serna, Laure Lebecherel, Charles Perrin, Vazken Andréassian, Ludovic Oudin. How should a rainfall-runoff model be parameterized in an almost ungauged catchment? A methodology tested on 609 catchments. *Water Resources Research*, 2016, 52 (6), pp.4765-4784. 10.1002/2015WR018549 . hal-01400511

HAL Id: hal-01400511

<https://hal.science/hal-01400511v1>

Submitted on 22 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **How should a rainfall-runoff model be parameterized in an almost ungaged** 2 **catchment? A methodology tested on 609 catchments**

3 Claudia Rojas-Serna¹, Laure Lebecherel², Charles Perrin^{2*}, Vazken Andréassian², Ludovic
4 Oudin³

5 ¹ Universidad Autónoma Metropolitana, División de Ciencias Básicas e Ingeniería, México, Mexico

6 ² Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), Antony, France

7 ³ UMR Metis, Université Pierre et Marie Curie, Paris, France

8 * Corresponding author, e-mail: charles.perrin@irstea.fr; Fax: +33 1 40 96 61 99

9 **Key points**

- 10 • A methodology to use point flow measurements for parameter estimation is presented
- 11 • Tests were made on a set of catchments with various flow conditions
- 12 • Better efficiency than classical regionalization approaches with only a few flow
13 measurements

14 **Abstract**

15 This paper examines catchments that are *almost* ungaged, i.e. catchments for which only a
16 small number of point flow measurements are available. In these catchments, hydrologists
17 may still need to simulate continuous streamflow time series using a rainfall-runoff model,
18 and the methodology presented here allows using few point measurements for model
19 parameterization. The method combines regional information (parameter sets of neighboring
20 gaged stations) and local information (contributed by the point measurements) within a
21 framework where the relative weight of each source of information is made dependent on the

22 number of point measurements available. This approach is tested with two different
23 hydrological models on a set of 609 catchments in France. The results show that on average a
24 few flow measurements can significantly improve the simulation efficiency, and that ten
25 measurements can reduce the performance gap between the gaged and ungaged situations by
26 more than 50%. Model parameters tend to come closer to the values obtained by calibration in
27 fully gaged conditions as the number of point flow measurements increases.

28 **Keywords:** Rainfall-runoff modeling; Regionalization; Point flow measurements; Ungaged
29 catchment; Neighborhood; Parameter estimation

30 **1 Introduction**

31 *1.1 Parameter estimation on entirely ungaged catchments*

32 Prediction in ungaged basins (PUB) has been one major focus of the hydrological community
33 in the past decade [Sivapalan, 2003] but still remains a great challenge. A full review of
34 ungaged basin research is not within the scope of this article and readers are referred to the
35 recent reviews by Blöschl *et al.* [2013] and Hrachowitz *et al.* [2013].

36 Here, we only mention that the most common parameter estimation techniques to transfer
37 information from gaged (donor) to ungaged (target) catchments are based on: (1) regression
38 equations linking parameters to physical and climate catchment characteristics (regressions to
39 be calibrated on gaged catchments), (2) transfer of parameter sets obtained in gaged
40 catchments that are similar in terms of physical and climate characteristics to the target
41 catchment, (3) transfer of parameter sets obtained in geographically close catchments. Many
42 variants exist, including procedures of regional model calibration [Castiglioni *et al.*, 2010;
43 Fernandez *et al.*, 2000; Hundecha *et al.*, 2008; Lombardi *et al.*, 2012; Parajka *et al.*, 2007].

44 These different approaches have been compared in various contexts, sometimes producing
45 contradictory results. *Parajka et al.* [2013] made a cross-analysis of 34 past regionalization
46 studies, showing that climate conditions or network density can explain differences between
47 results. In dense network conditions, spatial proximity approaches are often those that
48 perform the best.

49 ***1.2 Point streamflow measurements are ubiquitous***

50 To cope with the difficulties of estimating parameters in ungaged catchments, using
51 complementary or soft data, i.e. additional measurements or information on the catchment,
52 was advocated by some authors [*Fenicia et al.*, 2008; *Seibert and McDonnell*, 2002;
53 *Winsemius et al.*, 2009]. Among these complementary data, short time series or point
54 streamflow measurements are increasingly recognized as a valuable source of information for
55 model parameterization [see e.g. *Tada and Beven*, 2012]. There are indeed many locations in
56 the world where it is difficult to maintain long-term flow gauging stations and where only
57 short series are available [see e.g. *Bhatt and Tiwari*, 2008; *Kim and Kaluarachchi*, 2009; *Konz*
58 *et al.*, 2007]. Also, when a hydrological question is raised for an ungaged river, practicing
59 hydrologists may not have the time and resources to install a perennial gauging station, but
60 they will generally have the opportunity to take a small number of flow measurements.

61 One can mention a few examples of such situations reported in the literature. *Hughes et al.*
62 [2014] made weekly flow measurements on a small stream in South Africa over an 18-month
63 period and investigated how these data can help better constrain the parameters of the Pitman
64 hydrological model. *Temnerud et al.* [2007] carried out point flow measurements at 66 sites
65 within a 78-km² catchment in northern Sweden during a low-flow period to investigate the
66 role of spatial patterns in water quality issues. In France, systematic point flow measurement
67 campaigns have been coordinated by the Rhine-Meuse Water Agency over the last two

68 decades to improve the knowledge of low flows [Corbonnois *et al.*, 1999; Decloux and Sary,
69 1991; Drogue and Plasse, 2014; François and Sary, 1990; 1994; Plasse *et al.*, 2014]. These
70 point flow measurements may be useful for a number of objectives, such as low-flow
71 estimation [Catalogne *et al.*, 2014; Chopart and Sauquet, 2008; Eng and Milly, 2007;
72 Goodwin and Young, 2007; Laaha and Blöschl, 2005; Oberlin *et al.*, 1973].

73 Although the hydrologist needing to calibrate a model will not consider these catchments with
74 short times series or point flow measurements as *properly* gaged, they are not strictly
75 speaking ungaged. In this paper, we will refer to them as *almost ungaged*.

76 **1.3 Sensitivity of model performance to flow data availability**

77 For gaged or almost ungaged catchments, the robustness of the parameter sets identified
78 clearly depends on the information content of flows available for calibration [Wagener *et al.*,
79 2003]. When information is lacking, the *mathematical optimum* identified during calibration
80 will be potentially different from the *hydrological optimum* (i.e. the parameter set that can
81 reproduce catchment behavior over the long term), with possible problems of over-calibration
82 on the data at hand [see Andréassian *et al.*, 2012, for a more complete discussion on this
83 issue]. Hence, increasing the length of the flow series available for calibration generally tends
84 to produce more robust parameter sets, and several authors advised using series covering 5–8
85 years to calibrate models [see e.g. Anctil *et al.*, 2004; Yapo *et al.*, 1996]. Similarly, Gill *et al.*
86 [2007] showed that increasing the percentage of missing data in a calibration series tends to
87 produce less robust models. Indeed, longer time series generally encompass a larger variety of
88 hydrological conditions, which makes the series more informative for calibration.

89 However, a number of studies tend to indicate that shorter time series may also provide
90 valuable information. Brath *et al.* [2004] reported tests using calibration periods ranging from
91 1.5 to 12 months. The best results in validation were obtained using the parameters calibrated

92 in the 12-month period, but acceptable results could be obtained with 3 months of data.
93 Similarly, *Melsen et al.* [2014] concluded that 5 months of data are sufficient for parameter
94 optimization to obtain good results on the full observation period on a Swiss catchment when
95 applying a two-parameter model. Using discontinuous series, *Kim and Kaluarachchi* [2009]
96 tested the sensitivity of a water balance model to decreasing data availability. They showed
97 that short data series could produce parameter sets which can be effectively transposed in time
98 as long as some parts of the hydrograph (especially high and low flows) are included in the
99 series.

100 The sensitivity of model performance to flow availability was also assessed by *Perrin et al.*
101 [2007]. They used from 10 to 1,000 flow data randomly sampled out of long series to
102 calibrate two hydrological models on 12 US catchments. They showed that the optimized
103 parameter values became stable for the two models when 350 flow data were available for
104 calibration (with a significant drop in performance when fewer data were used).

105 ***1.4 Making the most of limited flow data***

106 The problem of parameter estimation in almost ungaged catchments has received increasing
107 attention over the last few years. Several authors suggested making explicit use of point flow
108 measurements. *Montanari and Toth* [2007] proposed a practical approach to calibrate rainfall-
109 runoff models with sparse data, using the Whittle estimator as a likelihood function and
110 calibrating the model in the spectral domain. *Seibert and Beven* [2009] used point flow
111 measurements to constrain the choice of model parameters. They concluded that "surprisingly
112 little runoff data was necessary to identify model parameterizations that provided good results
113 for the 'ungaged' test periods. These results indicated that a few runoff measurements can
114 contain much of the information content of continuous runoff time series." *Tada and Beven*
115 [2012] tested various optimization options of the TOPMODEL parameters on three Japanese

116 catchments using short continuous periods of 4, 8, ..., 512 days selected within a 10-year
117 period. They showed that considering an ensemble of acceptable parameter sets could
118 improve the results obtained by classic optimization when calibration time series are shorter
119 than 1 year. *Singh and Bardossy* [2012] also proposed an approach to identify robust
120 parameter sets when only short calibration periods are available. Using critical events
121 identified in a series with a depth function, they showed that calibration on events
122 representing 6–7% of a 10-year time series provided similar results to calibration on the
123 whole time series and better results than random selection of events.

124 ***1.5 Using regional information in parameter estimation on almost ungaged catchments***

125 Although the previous approaches concentrate on extracting information from the available
126 flow data only, the knowledge on parameter values gained from gaged catchments was
127 recognized early as valuable information even in the context of calibration procedures. For
128 example, *Koren et al.* [2003] found it beneficial to use the regionalization relationships
129 established for ungaged catchments as initial estimates of parameter values in the calibration
130 process. *Kuzmin et al.* [2008] also underlined the advantage of starting calibration with
131 already informative initial values and then improving these a priori estimates during the
132 calibration process. These authors mention that searching in the vicinity of the starting values
133 using a local search algorithm yields better results than using a global search algorithm,
134 especially where calibration data are lacking.

135 These results encouraged hydrologists to develop methods for exploiting regional information
136 in the case of limited flow information, to better constrain parameter estimation.

137 One solution is to limit the exploration of the parameter space during calibration. *Perrin et al.*
138 [2008] showed that the drop in performance caused by shortening calibration time series was
139 effectively attenuated by searching for the optimal parameter set in pre-sampled parameter

140 space, i.e. within a collection of parameter sets previously obtained in other gaged
141 catchments. The authors showed that this approach outperforms the classic calibration
142 approaches when only short-flow time series are available. The results of *Andréassian et al.*
143 [2014] corroborate these findings: they found that choosing parameters within a short-list of
144 27 parameter sets was more robust than making full calibration of the GR4J model when less
145 than 1 year of flow data is available.

146 From a different perspective, *Seibert and McDonnell* [2013] proposed an approach combining
147 the use of point flow measurements and soft data (user-defined fuzzy rules of acceptance of
148 groundwater contributions). They used data collected over a 3-month period and showed that
149 a single event or ten observations during high flows provided the same information as the
150 continuous 3 months. *Winsemius et al.* [2009] also proposed a framework to integrate hard
151 and soft data to constrain the estimation of parameters. Their results confirm the potential of
152 their method to be used for almost ungaged catchments.

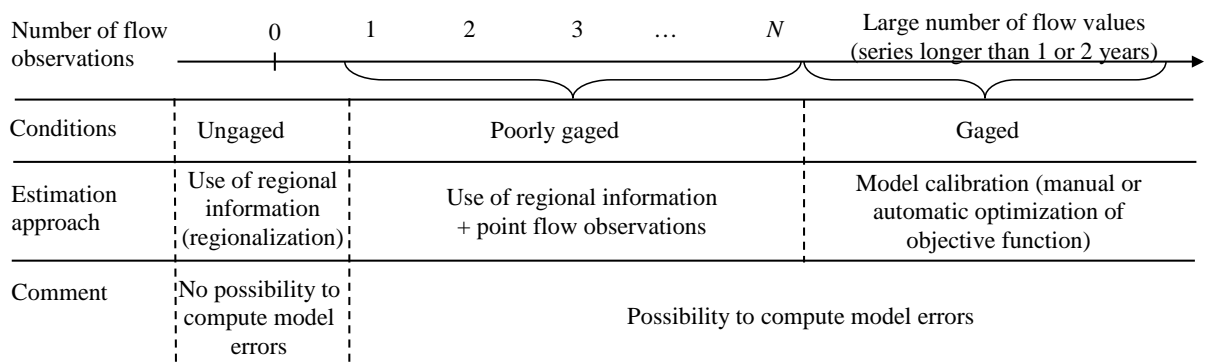
153 Another option is to combine parameters estimated by regionalization and point flow
154 measurement. Among the first attempts, *Rojas-Serna et al.* [2006] presented a method
155 merging the use of regionalized parameter values and optimization to parameterize a model
156 when only a few point measurements were available. More recently, *Viviroli and Seibert*
157 [2015] proposed a framework to improve parameter constraints with point flow
158 measurements. On a set of catchments in Switzerland, they showed that even a few flow
159 measurements help constrain parameters and improve model performance over purely
160 regionalized parameters.

161 The above studies suggest that regional sources of information can be useful for parameter
162 estimation even in the classic case of model calibration at a long-term gauging station. In
163 almost ungaged stations, on which this paper focuses, starting from regional hydrological
164 knowledge appears promising.

165 **1.6 Scope of the paper**

166 The above review suggests that parameters can be estimated using a limited number of flow
 167 observations (not necessarily contiguous but possibly spread out in time), but that classic
 168 optimization algorithms are insufficient for model parameterization when only a few flow
 169 data are available for calibration. Exploiting the prior knowledge gained at a regional level (in
 170 physically or spatially neighboring gauging catchments) is a valid alternative for parameter
 171 identification.

172 This paper presents a new parameter estimation approach for almost ungauged catchments,
 173 which specifically combines regional information transferred from neighboring gaged
 174 catchments with local information contributed by a limited number of flow measurements.
 175 The framework we present in this paper intends to make use of these measurements for the
 176 identification of hydrological model parameters, thus making a bridge between the fully
 177 ungauged and gaged cases, as illustrated in Figure 1.



178
 179 Figure 1. Comparison of the ungauged, almost ungauged and gaged conditions for the estimation of model
 180 parameters (N corresponds to the number of flow observations under which the application of classic estimation
 181 procedures are no longer robust)

182 After presenting the methodology (Section 2), we present the data set of 609 catchments and
 183 the two hydrological models used to evaluate the proposed approach (Section 3). The results
 184 are then presented and discussed in Section 4 and conclusions and perspectives are discussed
 185 in Section 5.

186 **2 Presentation of the parameter estimation method for almost ungaged** 187 **catchments**

188 This section outlines the proposed approach and describes how it blends regional knowledge
189 with point flow measurements in order to address the question of parameter estimation in
190 almost ungaged catchments.

191 **2.1 Origin of the method**

192 The proposed method builds on two existing approaches that we have merged:

- 193 1. The neighborhood approach [see e.g. *Oudin et al.*, 2008] used for entirely ungaged basins,
194 based on regional information: the parameter set for the target ungaged catchment is
195 chosen among existing parameter sets, previously calibrated on gaged catchments. A
196 distance between the ungaged catchment and its gaged neighbors can be defined either
197 geographically (spatial proximity) or in the space of catchment descriptors (physical
198 similarity).
- 199 2. The DIScrete Parameterization (DISP) approach [*Perrin et al.*, 2008] proposed for gaged
200 catchments: as in the neighborhood approach, the parameter set for the ungaged catchment
201 is selected within a collection of existing parameter sets, but here the criterion for
202 parameter set selection is the value of the model error obtained by running the model with
203 an existing parameter set on the target catchment. This method has been preferred over
204 classic optimization algorithms since *Perrin et al.* [2008] showed that they lead to
205 overfitting situations on short calibration time series.

206 These two approaches use the same prior information (a library of parameter sets previously
207 calibrated on gaged catchments) but differ in the way parameter sets are selected, i.e. the way
208 they define the distance between the donor gaged catchment and the target ungaged
209 catchment: in the first case, the distance is either defined in the space of physical descriptors

210 or geographically (regional information); in the second case, the distance is a function of the
211 difference between model simulations and flow observations (local information).

212 **2.2 Prerequisites**

213 In the presentation below, we consider that the lumped continuous rainfall-runoff model we
214 wish to apply has previously been calibrated on p gaged catchments, providing p parameter
215 sets that constitute a parameter library. For these p catchments and for the ungaged catchment
216 studied, a number of physical descriptors (e.g. catchment area, drainage density, mean slope,
217 vegetation cover, etc.) and geographic coordinates are also available. Based on this regional
218 information, the proposed approach can be applied using either physical similarity or spatial
219 proximity as a basis for hypothesizing catchment similarity.

220 In these neighborhood methods, a pool of M best parameter sets instead of a single best set
221 can be used, since this often gives better model results [Goswami *et al.*, 2007]. Pools of
222 parameter sets are generally applied by averaging the corresponding simulated model outputs:
223 the model is applied on the ungaged catchment with each of the M parameter sets before
224 averaging the M outputs. This option indeed performs better than averaging the M parameter
225 sets before applying the model, given the non-linearity of simulated processes [Oudin *et al.*,
226 2008].

227 In terms of data availability, we assume that the continuous time series of areal precipitation
228 and potential evapotranspiration (PE) are available in the ungaged target catchment over the
229 period where individual flow measurements are made, allowing the continuous rainfall-runoff
230 model to be run over this period.

231 **2.3 Proposed approach**

232 The method outlined below was found to be the most efficient among many variants that are
233 not presented here for the sake of brevity. *Rojas-Serna* [2005] provides full details of these
234 other attempts, e.g. the design of weighted objective functions based on model errors and the
235 departure of parameter values from their initial estimates obtained by regionalization. Note
236 that the framework proposed below is presented using flow as a source of information but
237 could also be generalized to other variables deemed useful for parameter estimation (soil
238 moisture, snow cover, etc.). The proposed approach involves three steps:

239 **Step 1: Ranking the library's p parameter sets**

240 The parameter sets of the p gaged catchments in the library are first ranked using the selected
241 neighborhood approach. The closest catchment is given rank 1, the most remote is given rank
242 p . We note r_j^{reg} the rank of the j^{th} parameter set in the library. It is up to the end-user to define
243 the neighborhood approach and the associated distance metric.

244 In parallel, the model is run on the ungaged catchment using each parameter set of the library
245 in turn and the available precipitation and PE time series. For each flow simulation produced,
246 a model error F is calculated on the N dates when flow observations are available. It is up to
247 the user to define the formulation of F , given his knowledge of the model and his modeling
248 objectives. The p parameter sets are ranked by increasing model error. The parameter set
249 providing the lowest model error is given rank 1, the one providing the largest error is given
250 rank p . We note r_j^{loc} the rank of the j^{th} parameter set in the library.

251 **Step 2: Combining ranks and selecting a pool of M parameter sets**

252 Two distances between the target catchment and the catchments in the library were defined in
253 the previous step: a distance in terms of neighborhood and a distance in terms of model error.
254 Both distances contain information on the relevance of the donor catchments and we wish to

255 combine them into a single index. As the distances are not expressed in the same unit and do
256 not vary over the same range in both cases, it is difficult to use their absolute values in the
257 combination. After testing various solutions [Rojas-Serna, 2005], the option to combine the
258 previously defined ranks was found to be a good compromise between simplicity and
259 efficiency. A linear combination of ranks is made:

$$r_j = \alpha.r_j^{reg} + (1 - \alpha).r_j^{loc} \quad \text{Eq. (1)}$$

260 where r_j is the combined rank of the j^{th} parameter set in the library and α is a weighting
261 coefficient (varying between 0 and 1) expressing the relative importance of the regional
262 information compared to the local information. Its value needs to be determined empirically
263 (see section 4.3). When α equals 1, the method comes down to the neighborhood approach.
264 When α equals 0, the method only uses point flow information. The combined rank r takes
265 values between 1 and p . Note that two parameter sets in the library may have the same
266 combined rank, which means that they will be considered equivalent in the proposed method.
267 This procedure merges the regional information with the local information gained from point
268 flow measurements and guides the selection of donor catchments that are eventually both
269 regionally and locally relevant. Using the combined rank r , a pool of the M closest parameter
270 sets can be selected (i.e. the M parameter sets ranked first). The choice of M will be discussed
271 in Section 4.2.

272 **Step 3: Determining the flow in the almost unged catchment using the pool of M** 273 **parameter sets**

274 Using the available precipitation and PE time series of the almost unged catchment, M flow
275 time series are simulated by the model on the almost unged catchment using each M

276 selected parameter set. Then the M simulated series are averaged (output averaging) to obtain
277 the flow simulation for the target catchment.

278 **2.4 *Setting rules for applying the method***

279 To practically apply the method when N flow measurements are available, two values must be
280 defined:

- 281 1. the number M of parameter sets selected in the pool to be applied on the target catchment;
- 282 2. the value of α , i.e. the weighting coefficient that defines the relative importance of
283 regional information.

284 We can expect the values of α to depend on the number N of flow measurements. Indeed, for
285 large values of N , there will be a lot of information in the flow data, so that we ought to give
286 more weight to the minimization of model error than to the neighborhood. Consequently, α
287 should take values closer to 0. The sensitivity of the proposed approach to M and α values is
288 analyzed in Section 4.

289 **2.5 *Options for case study application***

290 In the case study application described in Section 3, the following choices were made. We
291 used spatial proximity as the neighborhood approach, since it is the best performing method
292 on our data set [Oudin *et al.*, 2008]. To compute the distance separating the neighboring
293 catchment from the target ungaged catchment, a distance combining the horizontal Euclidean
294 distance between the outlets (d_{outlet}) and the horizontal Euclidean distance between the
295 centroids ($d_{centroid}$) was selected. Previous tests [Lebecherel, 2015] showed that using this
296 distance was beneficial to transfer information between catchments of different sizes. The
297 distance d considered is therefore defined as:

$$d = 0.2 \times d_{outlet} + 0.8 \times d_{centroid} \quad \text{Eq. (2)}$$

298 In terms of model error F , here we selected the root mean square error (RMSE) calculated on
299 all the available flow measurements:

$$F = \sqrt{\frac{1}{N} \sum_{i=1}^N (\sqrt{Q_i} - \sqrt{C_i})^2} \quad \text{Eq. (3)}$$

300 where Q_i and C_i are the observed discharge and the discharge calculated with the model,
301 respectively, for the date of the i^{th} flow measurement. The root square transformed flows were
302 used to compute F , because *Oudin et al.* [2006] showed that this formulation limits the
303 influence of high flows and provides a more general model.

304 Obviously, the method can be applied with other options according to the modeler's choice, in
305 terms of neighborhood, distance calculation [*Gottschalk et al.*, 2011] or model error
306 formulation [*Crochemore et al.*, 2015].

307 **3 Data, models and assessment methodology**

308 For a general evaluation of the method, we used a large set of catchments and two rainfall-
309 runoff models.

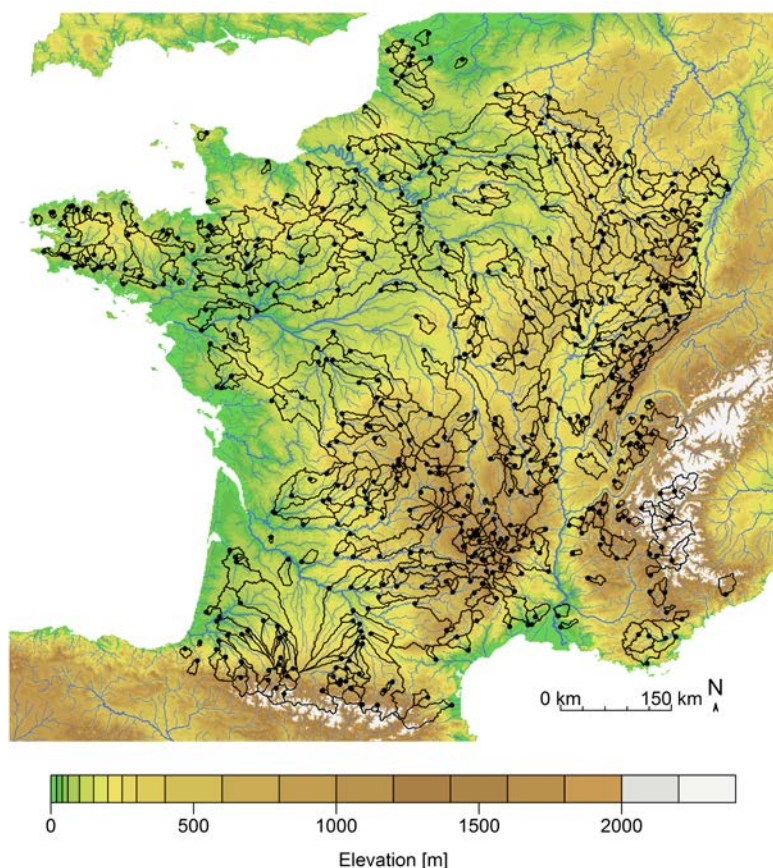
310 **3.1 Catchments and data set**

311 The proposed approach was tested on a large and varied catchment set for two reasons. First,
312 when a method is assessed on a large data set, we can have greater confidence in its
313 applicability and robustness [*Andréassian et al.*, 2006; *Gupta et al.*, 2014]. Second, the
314 method itself requires a library of parameter sets that should be large enough to be
315 representative of the conditions of the zone studied [see the discussion by *Perrin et al.*, 2008].

316 We used a set of 609 small to medium-size catchments in France (Figure 2). This set
317 represents a variety of hydrometeorological conditions, as shown in Table 1. Catchments were
318 selected to have limited gaps in flow series (less than 10% for every year of the series) and
319 limited snow influence. For the sake of brevity, physical catchment descriptors are not
320 detailed here since they were not used in the tests, but the catchment set includes various
321 physical conditions.

Quantiles	0.05	0.25	0.5	0.75	0.95
Catchment area (km ²)	34	109	270	833	4514
Mean elevation catchment (m)	87	180	375	781	1350
Annual rainfall, P (mm/yr)	714	863	1003	1209	1688
Annual potential evapotranspiration (PE) (mm/yr)	533	616	655	687	782
Annual discharge, Q (mm/yr)	159	272	411	643	1308

322 Table 1. Main characteristics of the 609 catchments used to test the approach.



323

324 Figure 2. Location of the 609 French catchments used in this study (dots indicate the gauging stations and solid
325 lines the catchment boundaries)

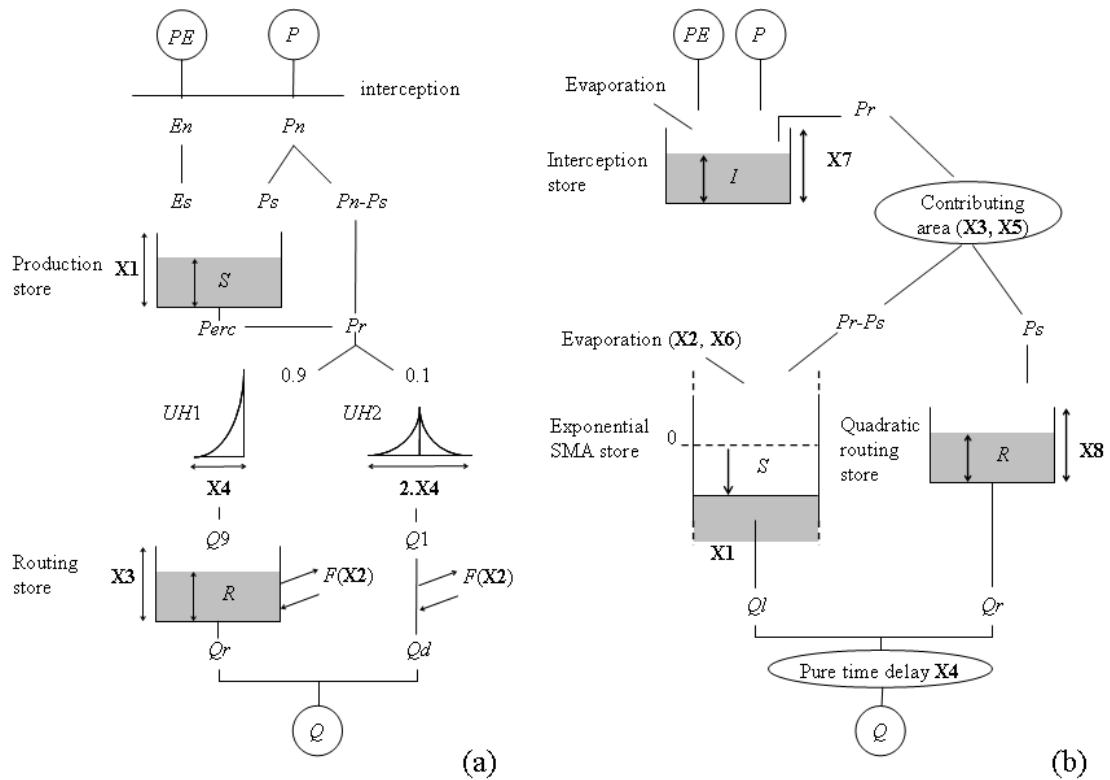
326 Daily data were available over the 1996–2005 period. Areal catchment rainfall was calculated
 327 using the SAFRAN gridded values provided by Météo-France [Vidal *et al.*, 2010]. Mean
 328 annual curves of potential evapotranspiration were computed using the formula provided by
 329 Oudin *et al.* [2005] based on air temperature and extra-terrestrial radiation. Streamflow time
 330 series were extracted from the HYDRO national archive (<http://hydro.eaufrance.fr>). These
 331 data are generally considered to be of good quality.

332 3.2 *Rainfall-runoff models and parameter library*

333 The method outlined in this paper can be applied with any lumped rainfall-runoff model. Here
 334 we used two models to reach more robust conclusions and possibly analyse differences
 335 between models. We applied the four-parameter GR4J model [Perrin *et al.*, 2003] and the
 336 eight-parameter TOPM model [Michel *et al.*, 2003]. A sketch of the model structures is
 337 shown in Figure 3 and the meaning of the parameters is given in Table 2. These two models
 338 were previously used in French catchments showing comparable levels of performance
 339 [Oudin *et al.*, 2008; Perrin *et al.*, 2008].

	Parameter	Meaning
GR4J	X1	Capacity of the production store (mm; positive)
	X2	Water exchange coefficient (mm; positive or negative)
	X3	Capacity of the nonlinear routing store (mm; positive)
	X4	Unit hydrograph time base (day; ≥ 0.5)
TOPM	X1	Parameter of the exponential store (mm; positive)
	X2	Evapotranspiration parameter (mm; positive or negative)
	X3	Topographic index distribution parameter (mm; positive)
	X4	Pure time delay (day; ≥ 1)
	X5	Topographic index distribution parameter (mm; positive or negative)
	X6	Evapotranspiration parameter (mm; positive)
	X7	Capacity of the interception store (mm; positive)
	X8	Capacity of the routing store (mm; positive)

340 Table 2. List of parameters of the GR4J and TOPM models



341

342 Figure 3. Schematic diagrams of the GR4J (a) and TOPM (b) rainfall-runoff models (PE: potential
 343 evapotranspiration; P: precipitation; Q: streamflow; Xi: model parameter i; other letters are internal state
 344 variables)

345 To build the library of model parameters, the parameters were calibrated on each catchment
 346 using the optimization algorithm applied by *Edijatno et al.* [1999]. The objective function
 347 used during optimization is the *Nash and Sutcliffe* [1970] criterion calculated on root square
 348 transformed flows, consistently with the formulation of F chosen here (see Section 2.5). Other
 349 objective functions [e.g. *Gupta et al.*, 2009] could be used consistently with the choices made
 350 in applying the method.

351 3.3 Assessment procedure

352 The proposed approach was successively applied to each catchment considered in turn as
 353 unengaged. Each time, the parameter set of the catchment under study was excluded from the
 354 library to test the approach.

355 We applied the split-sample test scheme advised by *Klemeš* [1986] by splitting the available
356 record into two periods (1995–2000 and 2000–2005) that were alternatively used for model
357 parameter identification and model assessment in validation (parameter identification in
358 period 1 and validation in period 2 and vice-versa). For each period, the first year (1995 and
359 2000, respectively) was used for model warm-up, which means that model performance was
360 actually computed in the 1996–2000 and 2001–2005 periods.

361 Here the flow measurements were randomly drawn in the flow series on each period. The
362 random option was chosen because it corresponds quite well to the case where one collects
363 point flow data without following a predefined acquisition strategy. This can be considered as
364 the “poor-man’s” option, i.e. a baseline strategy. The flow data were drawn incrementally: a
365 new flow measurement drawn in the flow series is added to the set of already selected flows,
366 mimicking what happens in operational conditions (i.e. a sample of $k+1$ measurements
367 includes the sample of k measurements already made). Hence it is considered that the flow
368 information is consistently increasing when N increases. The random selection was made
369 once per catchment. Since the number of catchments is large, this does not prevent obtaining
370 robust results.

371 More advanced sampling strategies could be adopted to improve modeling efficiency [*Viviroli*
372 *and Seibert*, 2015], but this was not within the scope of this article.

373 **3.4 Evaluation criteria**

374 The evaluation of the method was based on model performance obtained in validation mode
375 as measured by the *Nash and Sutcliffe* [1970] criterion (*NS*). *NS* varies within the interval $]-\infty,$
376 $1]$. The lack of a lower bound for this criterion is a problem when working on a large set of
377 catchments in ungaged conditions, because the criterion may take highly negative values in
378 some catchments where the model fails. This prevents making meaningful averages of

379 efficiency criteria over the catchment set. To circumvent this pitfall, we used the C criterion
380 proposed by *Mathevet et al.* [2006] and given by:

$$C = \frac{NS}{2 - NS} \quad \text{Eq. (3)}$$

381 This transformation is bounded in the interval $]-1, 1]$, which allows making meaningful
382 performance averages over the test set. Note that this criterion keeps the same baseline as the
383 Nash-Sutcliffe criterion ($C = 0$ when $NS = 0$), has the same optimum (1 means perfect
384 simulation for both criteria), but yields lower positive values compared to the NS criterion
385 (e.g. $C = 0.67$ when $NS = 0.80$).

386 The efficiency of the parameter estimation method will be assessed by mean C values
387 obtained by the hydrological model over all validation tests (here 1218, i.e. twice the number
388 of catchments). C was calculated on root square transformed flows (C_{RO}), like the objective
389 function, but also on flows (C_Q) and logarithm-transformed flows (C_{LQ}), to put more weight
390 on high and low flows than C_{RO} , respectively [see e.g. *Pushpalatha et al.*, 2012].

391 **3.5 Reference methods**

392 The model performance obtained by applying the proposed method was compared with the
393 model performance obtained:

- 394 1. in fully ungaged conditions: we applied the approach of spatial proximity with the
395 output averaging the pooling option [*Oudin et al.*, 2008];
- 396 2. in fully gaged conditions: we applied the optimization algorithm mentioned above
397 using all the data available in the calibration period.

398 Other benchmarks could be considered in model evaluation [see e.g. *Seibert*, 1999], but since
399 the intention here was not to compare various parameter estimation methods, we kept only
400 these two “extreme” benchmarks.

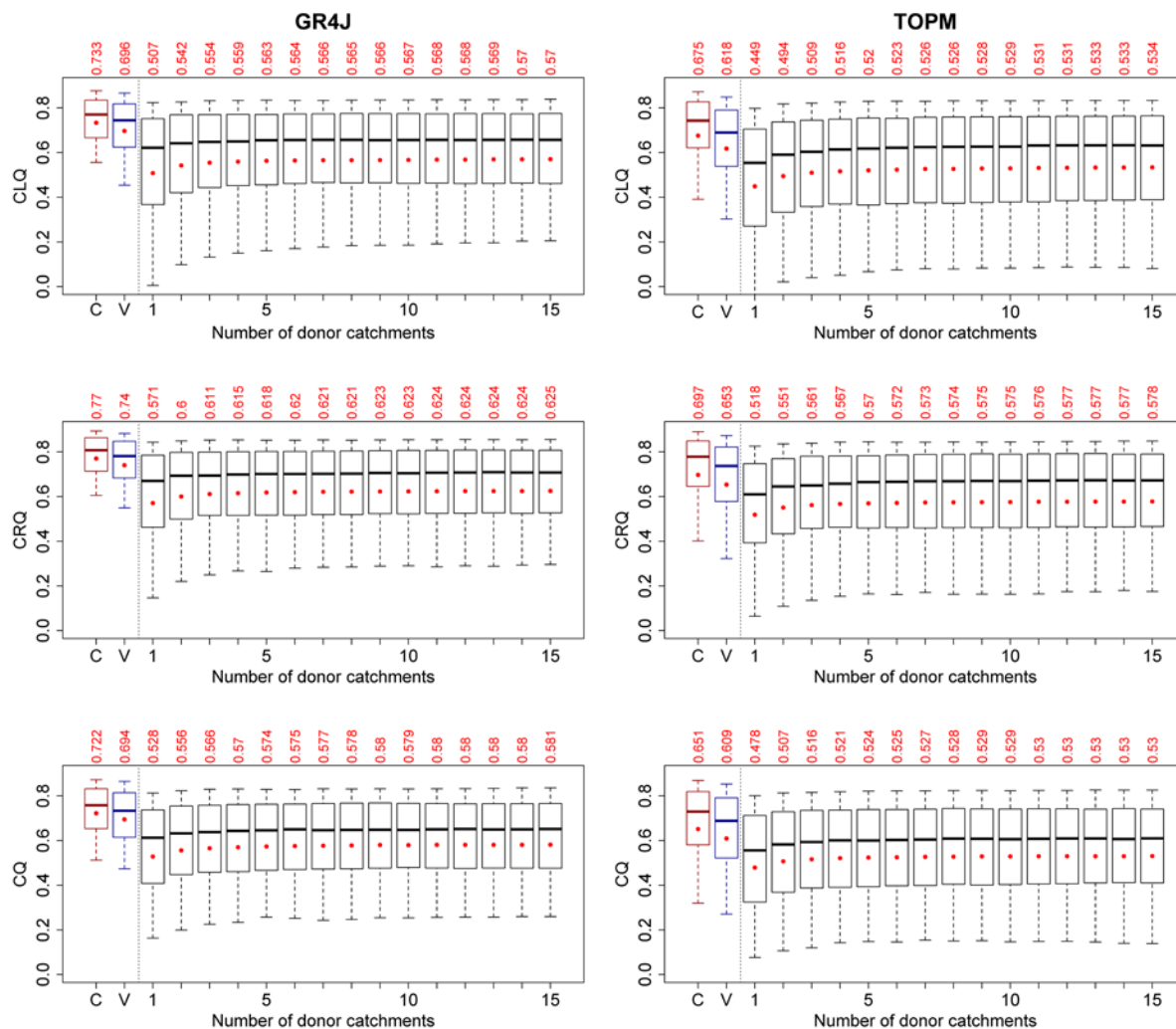
401 **4 Results**

402 In this section, we present the results of the proposed approach for almost unged
403 catchments. We discuss the sensitivity of the method to the value of the weighting factor of
404 the regional information (α) and to the number (M) of parameter sets in the pool applied to
405 the unged catchment. The two hydrological models were applied with an increasing number
406 of flow measurements (N). After determining the optimum values, we evaluate the
407 performance of the proposed approach, which we compare with the reference methods.

408 **4.1 Results of reference methods**

409 For the two models (GR4J and TOPM) and three evaluation criteria (C_{LQ} , C_{RQ} and C_Q),
410 Figure 4 shows the performance distributions obtained on the catchment set in calibration and
411 validation using the full flow data (fully gaged conditions) and when applying the
412 neighborhood approach (fully unged conditions) for an increasing number of donor
413 catchments. The best results in unged conditions are obtained with a small pool of
414 parameter sets: seven parameter sets for GR4J and nine parameter sets for TOPM. This
415 difference between the two models is consistent with the results found by *Oudin et al.* [2008]
416 on a similar data set. The smoothing effect of the output averaging option probably explains
417 why no significant improvement is found with more donors on average.

418 The efficiencies obtained by the two models are close in unged conditions. GR4J is slightly
419 better than TOPM in fully gaged conditions. Note that in gaged mode, the models are quite
420 efficient on average on the 609 catchments, since they reach C efficiencies of around 0.60–
421 0.65 (equivalent to 0.75–0.79 in terms of the NS criterion). As expected, the efficiency
422 obtained in fully gaged conditions is far better than that obtained in fully unged conditions.
423 This sets the range of improvements that can be obtained using flow information.

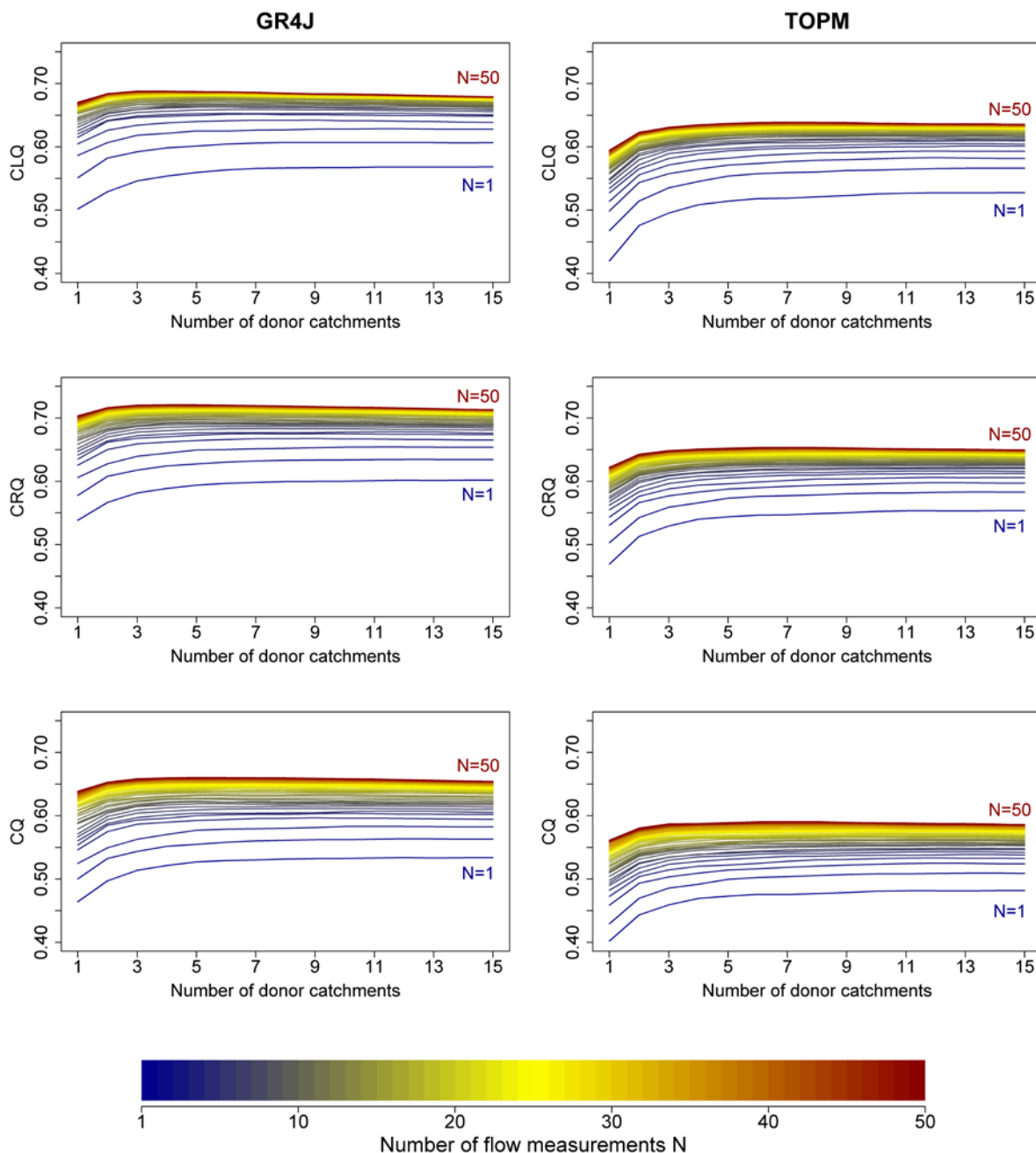


424

425 Figure 4. Distribution of model efficiency in calibration and validation in gaged conditions (red and blue
 426 boxplots, respectively) and in validation in ungaged conditions with the number of donor catchments (boxplots
 427 show the 0.1, 0.25, 0.50, 0.75 and 0.9 percentiles; the mean value is indicated on top of the frame and by the
 428 cross).

429 **4.2 Sensitivity to the number M of neighbors in almost ungaged conditions**

430 We evaluated the sensitivity of the proposed approach to the number M of parameter sets
 431 selected in the pool applied to the ungaged catchment. To simplify the presentation, we fixed
 432 the weight of regional information, α , at 0.5 (the results are very similar for other α values).
 433 We tested the two models using an increasing number of donors ($M = 1, 2, \dots, 9, 10, 15$) and
 434 an increasing number of flow measurements ($N = 1, 2, \dots, 6, 7, 10, 20, 50$).



435

436 Figure 5. Mean model efficiency (C_{LQ} , C_{RQ} and C_Q) in validation over the catchments with the number M of
 437 donor sets in almost ungaged conditions for the GR4J and TOPM models. Each line represents the application of
 438 the method with a given number of flow measurements N (from 1 to 50). Regional and local parameters are
 439 equally weighted ($\alpha = 0.5$ in Eq. 1).

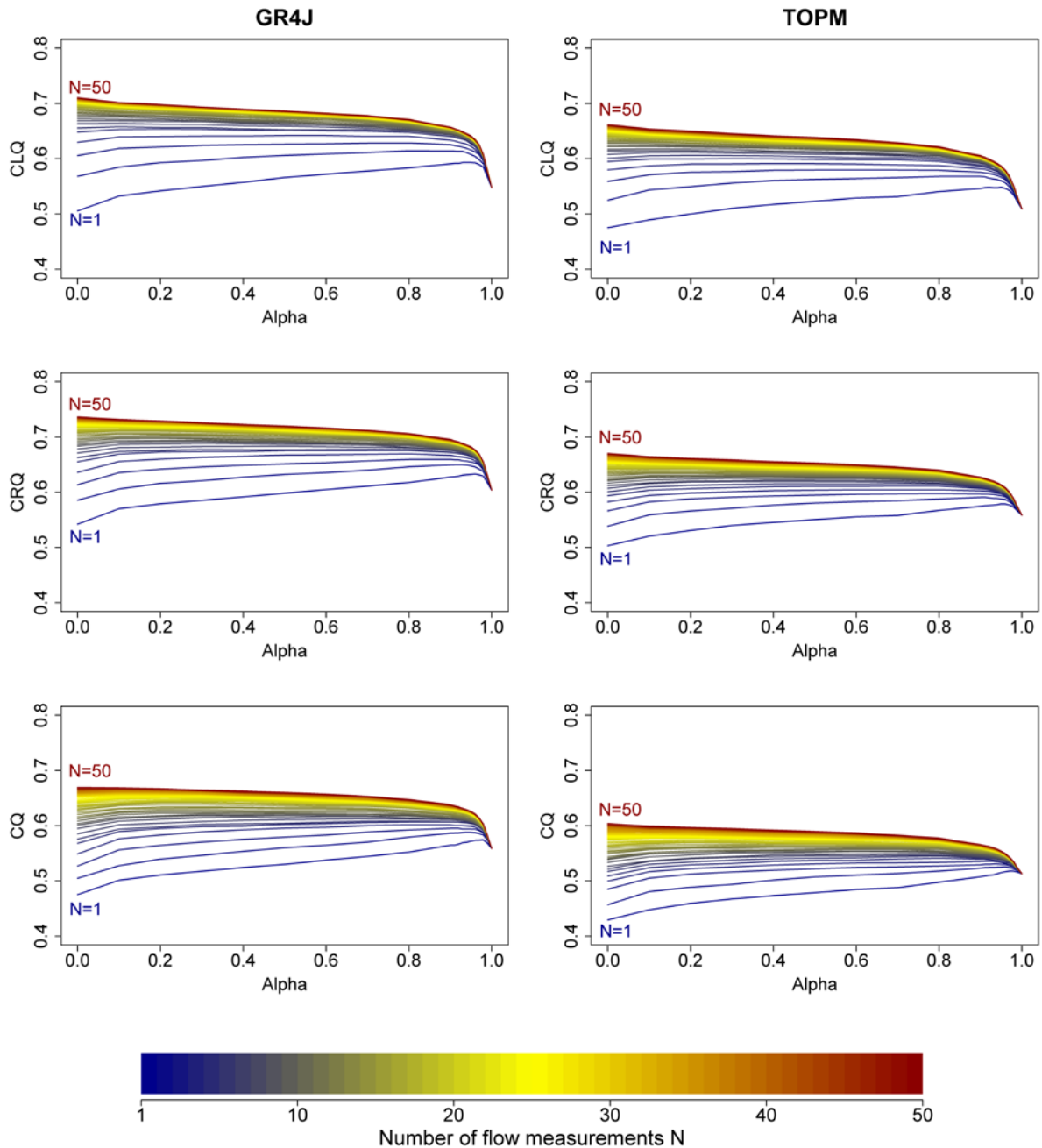
440 Figure 5 presents the mean model efficiency (C_Q , C_{LQ} and C_{RQ}) over the catchments with the
 441 number M of donor sets in almost ungaged conditions for the GR4J and TOPM models for an
 442 increasing number of flow measurements used for parameter estimation. It shows that as soon
 443 as a few donor sets are used, the results stabilize and are no longer sensitive to M . The same

444 behavior is observed for any number N of point flow measurements, the model or the
445 criterion. For a given model, the optimum number of donors does not seem to depend much
446 on N . Therefore we chose to set the number of donor sets to the same value identified in
447 ungaged conditions, i.e. $M = 7$ for GR4J and $M = 9$ for TOPM. These settings were used in
448 subsequent tests.

449 **4.3 Sensitivity to the weighting factor α**

450 We evaluated the sensitivity of the proposed approach to the value of weighting factor, α ,
451 which sets the weight of the regional information. We can expect that the larger the number of
452 point flow measurements, the lower the value of α , given that the flow measurements will
453 yield more information. We tested values of α from 0 (no use of regional information) to 1
454 (no use of local information) (see Eq. 1). α was varied between 0 and 1 by steps of 0.1 and
455 each value of α was tested using different numbers of flow measurements (from 1 to 50, as in
456 the previous section).

457 Figure 6 presents the mean model efficiency (C_Q , C_{LQ} and C_{RQ}) over the catchments with the
458 weighting factor α for an increasing number of flow measurements (from 1 to 50). When α
459 equals 1 (i.e. using only regional information), the results are the same as those found
460 previously for the strictly ungaged catchments. When α equals 0 (i.e. only considering local
461 flow information), model efficiency progressively increases when N increases. The
462 performance exceeds the performance obtained in ungaged conditions as soon as a few
463 (actually five at most) flow measurements are available for GR4J and TOPM.



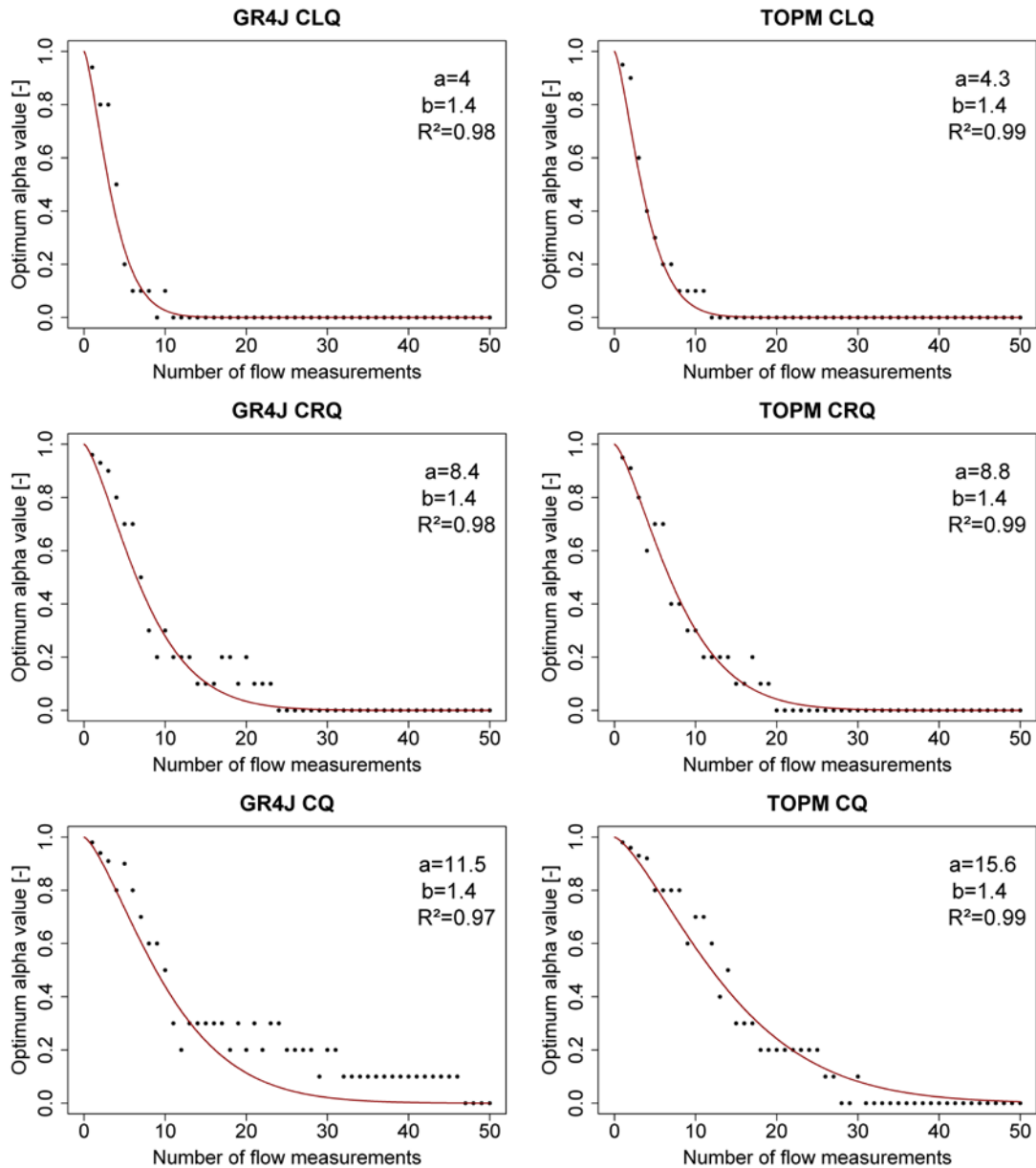
464

465 Figure 6. Mean model efficiency (C_{LQ} , C_{RQ} and C_Q) in validation over the catchments with the weighting factor
466 α of the regional information ($\alpha=0$: no regional information; $\alpha=1$: only regional information) in almost ungaged
467 conditions for the GR4J and TOPM models. Each line represents the application of the method with a given
468 number of flow measurements N (from 1 to 50). The number M of donor catchments is set to seven and nine for
469 GR4J and TOPM, respectively.

470 Figure 6 also shows that an optimum value of α can be identified in the efficiency curves
471 drawn for each value of N . This value is very close to 1 when a single flow measurement is
472 available, which means that it is better to mostly trust regional information in these
473 conditions. However, it does not strictly equal 1: this indicates that some information may be

474 gained even from a single flow measurement. Note that for all values of N , the increase in
475 model performance when α departs from 1 is very rapid. This means that flow measurements
476 yield very different and complementary information compared to solely regional information
477 (i.e. that the rankings are quite different). This can be linked to the finding of *Vrugt et al.*
478 [2002] and *Seibert and Beven* [2009], who found that even a few flow measurements already
479 contain valuable information.

480 When N increases, the optimum α value progressively decreases. This can be expected, since
481 the regional information becomes less and less relevant in comparison to the local (flow)
482 measurements. For each value of N , we identified this optimum value (based on the mean
483 results over the catchment set) and plotted it against N (see Figure 7). The patterns are quite
484 similar for both models, with slight differences, which may come from the fact that the
485 neighborhood approach does not give equally valuable information in each case. The
486 differences between criteria are greater than between models, but consistent between the two
487 models. When N increases, the weight of regional information drops faster for C_{LQ} than C_Q
488 (with intermediate results for C_{RQ}): approximately ten measurements are necessary to get α
489 close to 0 for C_{LQ} , while at least 30 values are necessary for C_Q . This may indicate that
490 regional information is worse and/or that fewer flow measurements are necessary when
491 focusing on low-flow conditions.



492

493 Figure 7. Optimum values of the weighting factor α of the regional information in almost ungaged conditions for
494 the GR4J and TOPM models for increasing numbers of flow measurements (from 1 to 50). The number N of
495 donor catchments is set to seven and nine for GR4J and TOPM, respectively. The solid line corresponds to the
496 curve defined in Eq. (4). α values were discretely tested with a 0.1 step.

497 It therefore seems that the optimal settings of the proposed approach in terms of the relative
498 importance of the regional information may partly depend on the model and, to a greater
499 extent, on the modeling objectives. However, given that the shapes of the relationship
500 between α and N are similar, a general formulation to determine the value of α for a given
501 number N of flow measurements available is proposed:

$$\alpha = \exp\left[-\left(\frac{N}{a}\right)^b\right] \quad \text{Eq. (4)}$$

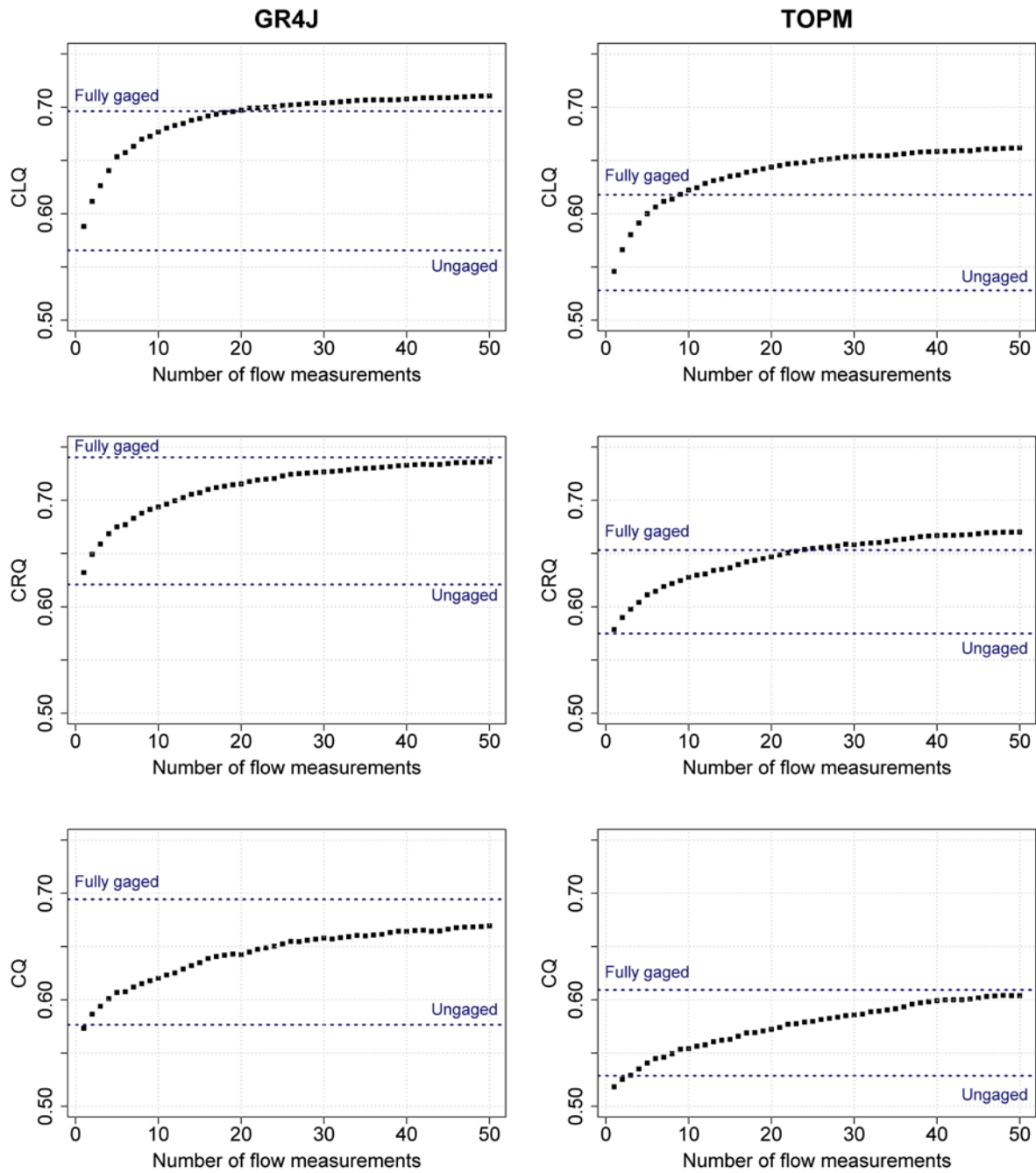
502 where a and b are two coefficients to be determined for the model and objectives selected.
503 The values of a and b and the corresponding curves are shown in Figure 7. Additional tests
504 (not detailed here) show that b is the less sensitive parameter of the two and the results
505 indicate that fixing it does not yield significant loss in modeling efficiency. In the subsequent
506 tests, we used the value $b=1.4$, which was found to provide the best results.

507 **4.4 Performance of the proposed approach**

508 Figure 8 shows the changes in the mean efficiency of the GR4J and TOPM models obtained
509 by applying the proposed approach for almost ungaged conditions to our data set, using
510 between 1 and 50 point measurements. It exploits both local flow measurements and regional
511 information, using the number M of neighbors determined previously and the optimum values
512 of α calculated by Eq. 4. It is compared with the two reference methods (the neighborhood for
513 the fully ungaged case and the optimization algorithm for the fully gaged case).

514 On average, the proposed approach efficiently uses the information provided by regional
515 information and local flow measurements, since it proves more efficient than the reference
516 method (i.e. the fully ungaged case) for all the efficiency criteria (C_Q , C_{LQ} and C_{RQ}).

517 These results show that the method proposed here can effectively exploit the two sources of
518 information. It is particularly valuable when only a few flow measurements are available.
519 Some improvement in model performance can be obtained in comparison with the sole use of
520 regional information when only one or two flow measurements are available. In most cases
521 (models and criteria), ten measurements (or a few more in the case of C_Q) can reduce the
522 performance gap between the gaged and ungaged situations by more than 50%.

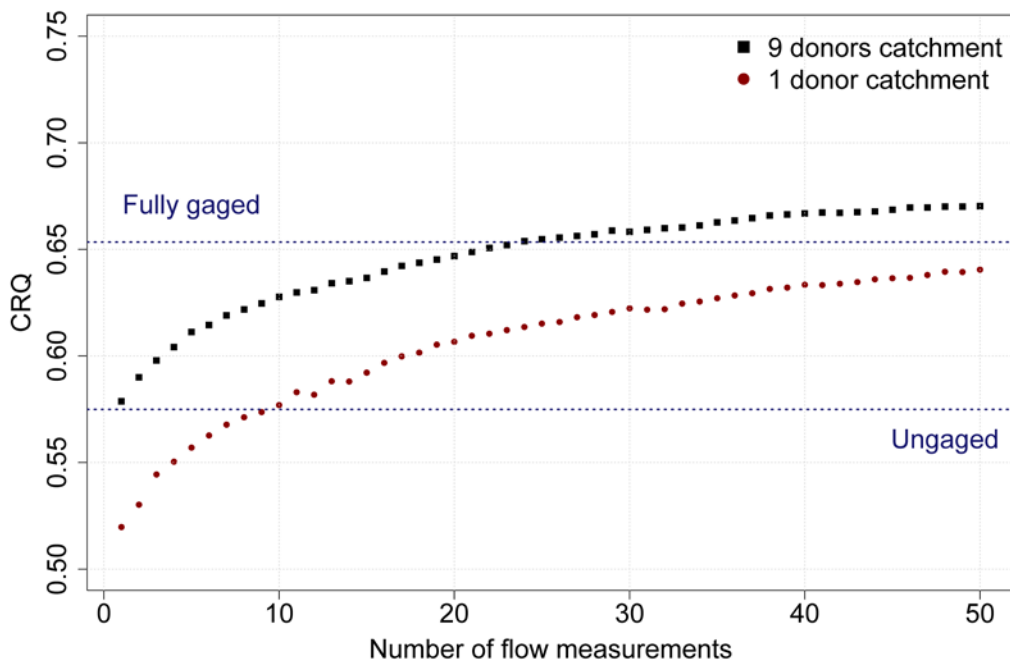


523

524 Figure 8. Efficiency of the proposed approach in validation for almost ungaged catchments compared to the
 525 ungaged and gaged conditions for the GR4J and TOPM models for an increasing number of flow measurements
 526 (from 1 to 50). The number N of donor catchments is set to seven and nine for GR4J and TOPM, respectively.

527 For the C_{LQ} criterion, the proposed approach provides more efficient results than the fully
 528 gaged case for both models when N exceeds 10–20. This may appear surprising, but can
 529 partly be explained by the fact that the evaluation criterion is different from the calibration
 530 objective function (C_{RQ}). Here, the fully gaged value is not the optimal validation value for

531 this criterion. More surprisingly, a similar behavior can be observed for the TOPM model and
532 the C_{RQ} evaluation criterion. These results may be explained by the use of multiple sets of
533 parameters (a pool of $M=9$ for TOPM) to simulate flows on almost ungaged catchments.
534 Figure 9 shows the efficiency of TOPM when applying the method for $M=9$ and $M=1$ (i.e.
535 using a single donor). Clearly, the multi-parameter approach outperforms the single-parameter
536 approach for TOPM, which remains below the fully gaged case. Note also that the results
537 shown here were obtained in validation mode, and the multi-parameter approach may be more
538 robust than the single-parameter approach. This behavior is observed to a larger extent for
539 TOPM than for GR4J, which may come from its higher number of parameters and
540 consequently lower robustness due to possible equifinality [as discussed by Perrin *et al.*,
541 2008].



542

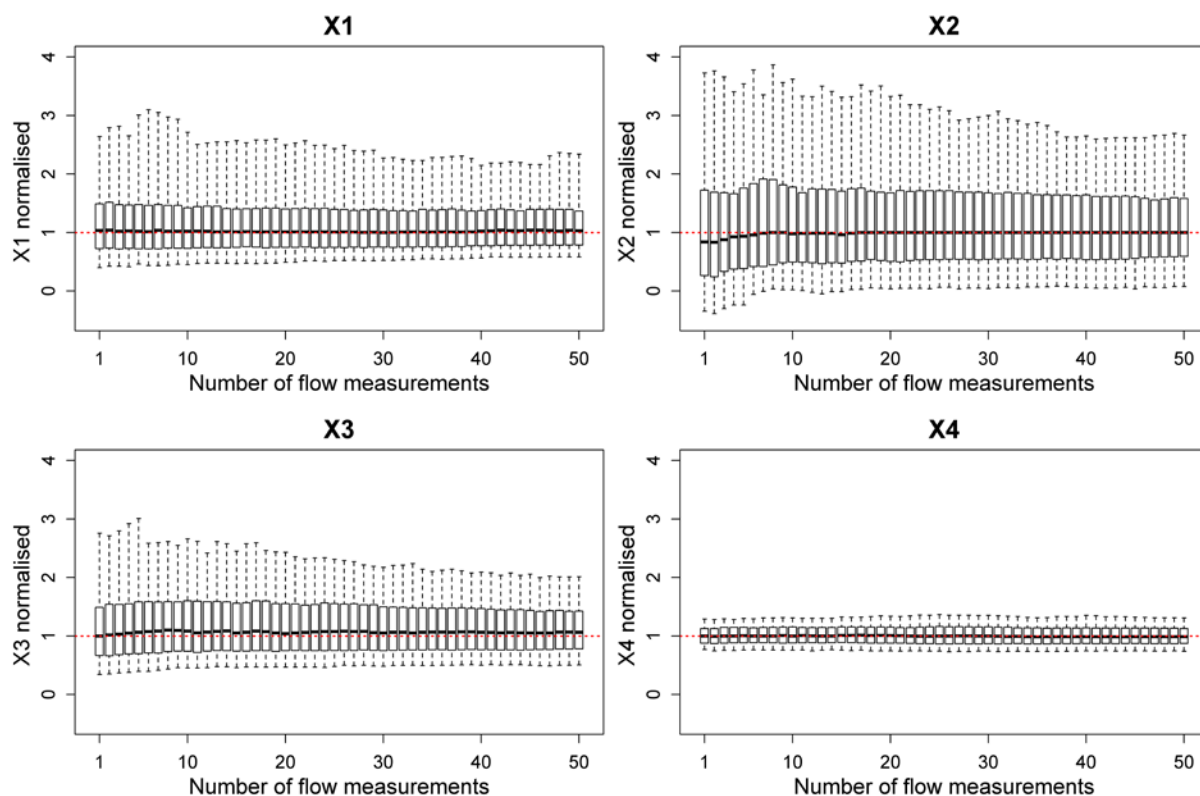
543 Figure 9. Efficiency of the proposed approach for the TOPM model and for the C_{RQ} evaluation criterion in
544 validation, for increasing numbers of flow measurements (from 1 to 50) and two donor catchment values (M)
545 (squares for nine donors and dots for one donor).

546 **4.5 Impact on parameter estimates**

547 We showed that progressively adding local information to existing regional information
548 yields better model performance in validation. This means that estimated parameters are more
549 transferable in time. Hence one could expect that they are closer to values estimated using full
550 flow information, which should be more representative of the catchment behavior over the
551 long term. Therefore, the proposed approach should help obtain more general parameters.

552 This was investigated by analyzing the parameter variability when the number of point flow
553 measurements increases. Because parameters can vary over a wide range, parameter values
554 were normalized in each case by the long-term optimum parameter set, i.e. the parameter set
555 obtained by calibration on the full data period. The distribution of these normalized
556 parameters on the sample of 609 catchments is shown in Figure 10 for N between 1 and 50.

557 To simplify the analysis, only the parameter set ranked first in our approach was considered
558 here (trends are similar when using the pool of parameter sets). For sake of brevity, we only
559 show the results for the GR4J model. Although the trend is smooth, one can see that the
560 variability decreases (the inter-quantile ranges decrease) for most parameters when N
561 increases. This means that on average, the parameter values tend to come closer to the long-
562 term optimum value. No real trend is observed for the X4 parameter (time base of the unit
563 hydrograph), but this parameter shows the least variability, i.e. it is probably better
564 determined by regionalization than the others. Note that median values of each boxplot are
565 very close to 1, indicating that the procedure does not tend to introduce any significant bias in
566 the estimated parameter values on average.



567

568 Figure 10. Distribution of normalized best parameters for the GR4J model over the catchment set with the
569 increase of flow measurements (the meaning of parameters is given in Table 2). Boxplots show the 0.1, 0.25,
570 0.50, 0.75 and 0.9 percentiles.

571 **5 Conclusion and perspectives**

572 In this article, a combined approach was devised to simultaneously exploit regional
573 information and local measurements. The proposed approach weights the two sources of
574 information depending on the availability of point flow measurements. It provides simulations
575 intermediate between fully ungaged and fully gaged situations. In this sense, the method
576 intends to make the connection between these two cases. This approach opens ways to make
577 model parameter estimation more reliable in all the catchments where only a few flow
578 measurements are available. As more flow measurements become available, the model
579 becomes more robust, i.e. it obtains better results in validation. The parameters also tend to
580 progressively converge to the long-term values estimated in fully gaged conditions. Hence the

581 procedure progressively adapts the values of parameters as new flow information is added, in
582 some type of "learning process" [Buytaert and Beven, 2009].

583 Starting from a prior (the regional estimate) and progressively narrowing the parameter values
584 by adding new information may be seen as a typical Bayesian process [Thiemann *et al.*,
585 2001]. The advantage of the proposed approach is that it remains very simple without making
586 statistical assumptions, but it would be interesting to compare it with formal Bayesian
587 approaches in future work.

588 The method was tested using one specific regionalization approach, two rainfall-runoff
589 models and three evaluation criteria targeting various flow types (high or low). Interestingly,
590 the method's settings were shown to depend more on the modeling objective than on the
591 model used. However, the method is general and flexible enough to be efficiently applied to
592 other models and objectives and a simple way was proposed to weight the regional and local
593 information. The approach was tested on a large data set of 609 catchments and compared to
594 two simple benchmarks, which gives confidence on its generality, overall efficiency
595 [Andréassian *et al.*, 2007] and applicability at the regional level [Drogue and Plasse, 2014].

596 One major conclusion is that acquiring flow information by point measurements and
597 efficiently combining it with prior regional information can be very effective in improving
598 parameter estimation, as already suggested by other authors [Seibert and Beven, 2009].
599 Studies comparing various parameter estimation strategies (and possibly involving other types
600 of observations than flow) would be very useful to conduct in the context of almost ungaged
601 catchments.

602 One limit of these tests is that flows were sequentially drawn at random in the existing series,
603 irrespective of the flow magnitude, dynamics (rising limb or recession) or season. These
604 various conditions are known to provide different informative content for the estimation of

605 model parameters [Wagener *et al.*, 2003]. Several authors attempted to define strategies for
606 acquiring point flow data in a modeling perspective [see e.g. Clausen, 1995; Juston *et al.*,
607 2009; Konz and Seibert, 2010; Singh and Bardossy, 2012; Viviroli and Seibert, 2015]. This
608 could also be investigated in the context of the approach proposed here. Ultimately, this could
609 guide practicing hydrologists and operational gauging staff in defining gauging priorities and
610 making decisions in real time.

611 Last, we mentioned in the introduction the PUB decade, which gave way to the Panta Rhei
612 decade on “Change in hydrology and society” [Montanari *et al.*, 2013]. It is likely that the
613 notion of an *almost ungauged catchment* could be revisited and extended in this context.
614 Typically a sudden change on a catchment will strongly impact its behavior and make it
615 suddenly ungauged in the sense that no or not enough data are available to properly model it.
616 Hence the interactions between transposition in time and space could be further investigated
617 in this context.

618 **6 Acknowledgements**

619 The authors wish to thank Météo-France and SCHAPI for providing meteorological and
620 hydrological data sets, respectively. Conacyt (Mexico) and ONEMA (France) provided
621 financial support for the PhD research of the first and second authors, respectively, and are
622 gratefully acknowledged.

623 **7 References**

624 Anctil, F., C. Perrin, and V. Andréassian (2004), Impact of the length of observed records on
625 the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting
626 models, *Environmental Modelling & Software*, 19(4), 357-368.
627 Andréassian, V., A. Hall, N. Chahinian, and J. Schaake (2006), Introduction and Synthesis:
628 Why should hydrologists work on a large number of basin data sets?, paper presented at

- 629 Large sample basin experiments for hydrological model parameterization: Results of the
630 Model Parameter Experiment - MOPEX, Paris, IAHS Publication n°307, 1-5.
- 631 Andréassian, V., J. Lerat, C. Loumagne, T. Mathevet, C. Michel, L. Oudin, and C. Perrin
632 (2007), What is really undermining hydrologic science today?, *Hydrol. Processes*, 21(20),
633 2819-2822.
- 634 Andréassian, V., N. Le Moine, C. Perrin, M.-H. Ramos, L. Oudin, T. Mathevet, J. Lerat, and
635 L. Berthet (2012), All that glitters is not gold: the case of calibrating hydrological models,
636 *Hydrol. Processes*, 26(14), 2206-2210.
- 637 Andréassian, V., F. Bourgin, L. Oudin, T. Mathevet, C. Perrin, J. Lerat, L. Coron, and L.
638 Berthet (2014), Seeking genericity in the selection of parameter sets: Impact on
639 hydrological model efficiency, *Water Resour. Res.*, 50(10), 8356 - 8366.
- 640 Bhatt, V. K., and A. K. Tiwari (2008), Estimation of peak streamflows through channel
641 geometry, *Hydrol. Sci. J.-J. Sci. Hydrol.*, 53(2), 401-408.
- 642 Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (Eds.) (2013), *Runoff*
643 *Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales*, 465 pp.,
644 Cambridge University Press.
- 645 Brath, A., A. Montanari, and E. Toth (2004), Analysis of the effects of different scenarios of
646 historical data availability on the calibration of a spatially-distributed hydrological model,
647 *J. Hydrol.*, 291(3-4), 232-253.
- 648 Buytaert, W., and K. Beven (2009), Regionalization as a learning process, *Water Resour.*
649 *Res.*, 45, W11419.
- 650 Castiglioni, S., L. Lombardi, E. Toth, A. Castellarin, and A. Montanari (2010), Calibration of
651 rainfall-runoff models in ungauged basins: A regional maximum likelihood approach, *Adv.*
652 *Water Resour.*, 33(10), 1235-1242.
- 653 Catalogne, C., E. Sauquet, and M. Lang (2014), Using spot gauging data to estimate the
654 annual minimum monthly flow with a return period of 5 years, *La Houille Blanche*(4), 78-
655 87.
- 656 Chopart, S., and É. Sauquet (2008), Usage des jaugeages volants en régionalisation des débits
657 d'étiage, *Revue des Sciences de l'Eau*, 21(3), 267-281.
- 658 Clausen, B. (1995), Discharge Data-Collection and Analysis Strategies in Low-Flow Studies,
659 *Nordic Hydrol.*, 26(3), 191-204.
- 660 Corbonnois, J., D. François, I. Arts, M. Sary, E. Gille, P. Gamez, A. Devos, and J. P. Decloux
661 (1999), Répartition des écoulements en basses eaux. Aspects méthodologiques.
662 (Spatialization of low water flows. Methodology), paper presented at XXII IUGG General
663 Assembly, 18-30 July 1999, Birmingham (UK), 253-260.
- 664 Crochemore, L., C. Perrin, V. Andréassian, U. Ehret, S. P. Seibert, S. Grimaldi, H. Gupta, and
665 J. E. Paturel (2015), Comparing expert judgement and numerical criteria for hydrograph
666 evaluation, *Hydrol. Sci. J.*, 60(3), 402-423.
- 667 Decloux, J. P., and M. Sary (1991), Campagnes d'étiage. Objectifs, traitement et valorisation
668 des données, *Mosella*, XXIII, 121-134.
- 669 Drogue, G. P., and J. Plasse (2014), How can a few streamflow measurements help to predict
670 daily hydrographs at almost ungauged sites?, *Hydrol. Sci. J.*, 59(12), 2126-2142.
- 671 Edijatno, N. O. Nascimento, X. Yang, Z. Makhlof, and C. Michel (1999), GR3J: a daily
672 watershed model with three free parameters, *Hydrol. Sci. J.*, 44(2), 263-277, DOI:
673 210.1080/02626669909492221.
- 674 Eng, K., and P. C. D. Milly (2007), Relating low-flow characteristics to the base flow
675 recession time constant at partial record stream gauges, *Water Resour. Res.*, 43(1).

- 676 Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008), Learning from model
677 improvement: On the contribution of complementary data to process understanding, *Water*
678 *Resour. Res.*, 44(6), W06419.
- 679 Fernandez, W., R. M. Vogel, and A. Sankarasubramanian (2000), Regional calibration of a
680 watershed model, *Hydrol. Sci. J.*, 45(5), 689-707.
- 681 François, D., and M. Sary (1990), Etude méthodologique des basses eaux. estimation
682 minimale des points de mesure, 14 pp, Cegum, Université de Metz.
- 683 François, D., and M. Sary (1994), Etude méthodologique des débits d'étiage, 30 pp, Cegum,
684 Université de Metz.
- 685 Gill, M. K., T. Asefa, Y. Kaheil, and M. McKee (2007), Effect of missing data on
686 performance of learning algorithms for hydrologic predictions: Implications to an
687 imputation technique, *Water Resour. Res.*, 43(7), W07416.
- 688 Goodwin, T., and A. Young (2007), Reducing uncertainty for the estimation of hydropower
689 potential - the use of local data, paper presented at British Hydropower Association
690 Conference, 11 October 2007, www.british-hydro.org, Birnam, Australia.
- 691 Goswami, M., K. M. O'Connor, and K. P. Bhattarai (2007), Development of regionalisation
692 procedures using a multi-model approach for flow simulation in an ungauged catchment, *J.*
693 *Hydrol.*, 333(2-4), 517-531.
- 694 Gottschalk, L., E. Leblois, and J. O. Skoien (2011), Distance measures for hydrological data
695 having a support, *J. Hydrol.*, 402(3-4), 415-421.
- 696 Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean
697 squared error and NSE performance criteria: Implications for improving hydrological
698 modelling, *J. Hydrol.*, 377(1-2), 80-91.
- 699 Gupta, H. V., C. Perrin, G. Blöschl, A. Montanari, R. Kumar, M. Clark, and V. Andréassian
700 (2014), Large-sample hydrology: a need to balance depth with breadth, *Hydrol. Earth Syst.*
701 *Sci.*, 18(2), 463-477.
- 702 Hrachowitz, M., et al. (2013), A decade of Predictions in Ungauged Basins (PUB)—a review,
703 *Hydrol. Sci. J.*, 58(6), 1198-1255.
- 704 Hughes, D. A., M. Gush, J. Tanner, and P. Dye (2014), Using targeted short-term field
705 investigations to calibrate and evaluate the structure of a hydrological model, *Hydrol.*
706 *Processes*, 28(5), 2794-2809.
- 707 Hundecha, Y., T. Ouarda, and A. Bardossy (2008), Regional estimation of parameters of a
708 rainfall-runoff model at ungauged watersheds using the "spatial" structures of the
709 parameters within a canonical physiographic-climatic space, *Water Resour. Res.*, 44(1),
710 W01427.
- 711 Juston, J., J. Seibert, and P. O. Johansson (2009), Temporal sampling strategies and
712 uncertainty in calibrating a conceptual hydrological model for a small boreal catchment,
713 *Hydrol. Processes*, 23(21), 3093-3109.
- 714 Kim, U., and J. J. Kaluarachchi (2009), Hydrologic model calibration using discontinuous
715 data: an example from the upper Blue Nile River Basin of Ethiopia, *Hydrol. Processes*,
716 23(26), 3705-3717.
- 717 Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*,
718 31(1), 13-24.
- 719 Konz, M., and J. Seibert (2010), On the value of glacier mass balances for hydrological model
720 calibration, *J. Hydrol.*, 385, 238-246.

- 721 Konz, M., S. Uhlenbrook, L. Braun, A. Shrestha, and S. Demuth (2007), Implementation of a
722 process-based catchment model in a poorly gauged, highly glacierized Himalayan
723 headwater, *Hydrology and Earth System Sciences*, 11(4), 1323-1339.
- 724 Koren, V., M. Smith, and Q. Duan (2003), Use of a priori parameter estimates in the
725 derivation of spatially consistent parameter sets of rainfall-runoff models, *Calibration of*
726 *Watershed Models, Water Science and Application*, 6, 239-254.
- 727 Kuzmin, V., D. J. Seo, and V. Koren (2008), Fast and efficient optimization of hydrologic
728 model parameters using a priori estimates and stepwise line search, *J. Hydrol.*, 353(1-2),
729 109-128.
- 730 Laaha, G., and G. Blöschl (2005), Low flow estimates from short stream flow records - a
731 comparison of methods, *J. Hydrol.*, 306(1-4), 264-286.
- 732 Lebecherel, L. (2015), Sensibilité des calculs hydrologiques à la densité des réseaux de
733 mesure hydrométrique et pluviométrique, PhD thesis, 280 pp, AgroParisTech - Irstea
734 (Antony).
- 735 Lombardi, L., E. Toth, A. Castellarin, A. Montanari, and A. Brath (2012), Calibration of a
736 rainfall-runoff model at regional scale by optimising river discharge statistics:
737 Performance analysis for the average/low flow regime, *Physics and Chemistry of the*
738 *Earth, Parts A/B/C*, 42-44, 77-84.
- 739 Mathevet, T., C. Michel, V. Andréassian, and C. Perrin (2006), A bounded version of the
740 Nash-Sutcliffe criterion for better model assessment on large sets of basins, in *Large*
741 *sample basin experiments for hydrological model parameterisation: Results of the Model*
742 *Parameter Experiment - MOPEX*, edited by V. Andréassian, A. Hall, N. Chahinian and J.
743 Schaake, pp. 211-219, IAHS Publ. 307.
- 744 Melsen, L. A., A. J. Teuling, S. W. van Berkum, P. J. J. F. Torfs, and R. Uijlenhoet (2014),
745 Catchments as simple dynamical systems: A case study on methods and data requirements
746 for parameter identification, *Water Resour. Res.*, 50(7), 5577-5596.
- 747 Michel, C., C. Perrin, and V. Andréassian (2003), The exponential store: a correct formulation
748 for rainfall-runoff modelling, *Hydrol. Sci. J.*, 48(1), 109-124, DOI:
749 110.1623/hysj.1648.1621.1109.43484.
- 750 Montanari, A., and E. Toth (2007), Calibration of hydrological models in the spectral domain:
751 an opportunity for scarcely gauged basins?, *Water Resour. Res.*, 43(5), W05434.
- 752 Montanari, A., et al. (2013), “Panta Rhei—Everything Flows”: Change in hydrology and
753 society—The IAHS Scientific Decade 2013–2022, *Hydrol. Sci. J.*, 58(6), 1256-1275.
- 754 Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models. Part
755 I - A discussion of principles., *J. Hydrol.*, 10(3), 282-290.
- 756 Oberlin, G., G. Galéa, and J. T. Toni (1973), Intérêt des jaugeages épisodiques pour
757 l'estimation des étiages de petits bassins non équipés, *Cahiers de l'ORSTOM, série*
758 *Hydrologie*, X, 349-368.
- 759 Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging
760 of rainfall-runoff model simulations from complementary model parameterizations, *Water*
761 *Resour. Res.*, 42(7), W07410, doi: 07410.01029/02005WR004636.
- 762 Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine (2008), Spatial proximity,
763 physical similarity and ungauged catchments: confrontation on 913 French catchments,
764 *Water Resour. Res.*, 44, W03413, doi:03410.01029/02007WR006240.
- 765 Oudin, L., F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, and C. Loumagne
766 (2005), Which potential evapotranspiration input for a rainfall-runoff model? Part 2 -

- 767 Towards a simple and efficient PE model for rainfall-runoff modelling, *J. Hydrol.*, 303(1-
768 4), 290-306, DOI: <http://dx.doi.org/210.1016/j.jhydrol.2004.1008.1026>.
- 769 Parajka, J., G. Blöschl, and R. Merz (2007), Regional calibration of catchment models:
770 potential for ungauged catchments, *Water Resour. Res.*, 43(6), W06406.
- 771 Parajka, J., A. Viglione, M. Rogger, J. L. Salinas, M. Sivapalan, and G. Blöschl (2013),
772 Comparative assessment of predictions in ungauged basins - Part 1: Runoff-hydrograph
773 studies, *Hydrology and Earth System Sciences*, 17(5), 1783-1795.
- 774 Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for
775 streamflow simulation, *J. Hydrol.*, 279(1-4), 275-289, DOI: 210.1016/S0022-
776 1694(1003)00225-00227.
- 777 Perrin, C., V. Andréassian, C. Rojas-Serna, T. Mathevet, and N. Le Moine (2008), Discrete
778 parameterization of hydrological models: evaluating the use of parameter sets libraries
779 over 900 catchments, *Water Resour. Res.*, 44, W08447,
780 doi:08410.01029/02007WR006579.
- 781 Perrin, C., L. Oudin, V. Andréassian, C. Rojas-Serna, C. Michel, and T. Mathevet (2007),
782 Impact of limited streamflow knowledge on the efficiency and the parameters of rainfall-
783 runoff models, *Hydrol. Sci. J.*, 52(1), 131-151, DOI: 110.1623/hysj.1652.1621.1131.
- 784 Plasse, J., G. Drogue, D. Francois, P. Battaglia, and C. Conan (2014), Using point flow
785 measurements for guided reconstruction of daily streamflow time series through
786 continuous flow simulation: a regional case study on the Rhine-Meuse district, *Houille*
787 *Blanche-Revue Internationale De L Eau*(1), 45-52.
- 788 Pushpalatha, R., C. Perrin, N. Le Moine, and V. Andréassian (2012), A review of efficiency
789 criteria suitable for evaluating low-flow simulations, *J. Hydrol.*, 420-421, 171-182.
- 790 Rojas-Serna, C. (2005), Quelle connaissance hydrométrique minimale pour définir les
791 paramètres d'un modèle pluie-débit ? (Which is the minimum hydrometric information to
792 estimate the parameters of a rainfall-runoff model?), PhD thesis, 319 pp, Cemagref
793 (Antony), ENGREF (Paris), France.
- 794 Rojas-Serna, C., C. Michel, C. Perrin, and V. Andréassian (2006), Ungauged catchments:
795 How to make the most of a few streamflow measurements?, *IAHS Publ.* 307, 230-236.
- 796 Seibert, J. (1999), Regionalisation of parameters for a conceptual rainfall-runoff model, *Agric.*
797 *For. Meteorol.*, 98-99, 279-293.
- 798 Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in
799 catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour.*
800 *Res.*, 38(11), 1241.
- 801 Seibert, J., and K. J. Beven (2009), Gauging the ungauged basin: how many discharge
802 measurements are needed?, *Hydrology and Earth System Sciences*, 13(6), 883-892.
- 803 Seibert, J., and J. McDonnell (2013), Gauging the Ungauged Basin: Relative Value of Soft
804 and Hard Data, *Journal of Hydrologic Engineering*, 20(1), A4014004.
- 805 Singh, S. K., and A. Bardossy (2012), Calibration of hydrological models on hydrologically
806 unusual events, *Adv. Water Resour.*, 38, 81-91.
- 807 Sivapalan, M. (2003), Prediction in ungauged basins: a grand challenge for theoretical
808 hydrology, *Hydrol. Processes*, 17(15), 3163-3170.
- 809 Tada, T., and K. J. Beven (2012), Hydrological model calibration using a short period of
810 observations, *Hydrol. Processes*, 26(6), 883-892.
- 811 Temnerud, J., J. Seibert, M. Jansson, and K. Bishop (2007), Spatial variation in discharge and
812 concentrations of organic carbon in a catchment network of boreal streams in northern
813 Sweden, *J. Hydrol.*, 342(1-2), 72-87.

- 814 Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter
815 estimation for hydrologic models, *Water Resour. Res.*, 37(10), 2521-2535.
- 816 Vidal, J.-P., E. Martin, L. Franchisteguy, M. Baillon, and J.-M. Soubeyroux (2010), A 50-year
817 high-resolution atmospheric reanalysis over France with the Safran system, *International*
818 *Journal of Climatology*, 30(11), 1627-1644.
- 819 Viviroli, D., and J. Seibert (2015), Can a regionalized model parameterisation be improved
820 with a limited number of runoff measurements?, *J. Hydrol.*, 529, Part 1, 49-61.
- 821 Vrugt, J. A., W. Bouten, H. V. Gupta, and S. Sorooshian (2002), Toward improved
822 identifiability of hydrologic model parameters: The information content of experimental
823 data, *Water Resour. Res.*, 38(12), 1312, doi: 1310.1029/2001WR001118.
- 824 Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards
825 reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability
826 analysis, *Hydrol. Processes*, 17(2), 455-476.
- 827 Winsemius, H. C., B. Schaefli, A. Montanari, and H. H. G. Savenije (2009), On the
828 calibration of hydrological models in ungauged basins: A framework for integrating hard
829 and soft hydrological information, *Water Resour. Res.*, 45, W12422.
- 830 Yapo, P. O., H. V. Gupta, and S. Sorooshian (1996), Automatic calibration of conceptual
831 rainfall-runoff models: sensitivity to calibration data, *J. Hydrol.*, 181, 23-48.
- 832
- 833

834

835

836 **List of Tables**

837 Table 1. Main characteristics of the 609 catchments used to test the approach..... 15

838 Table 2. List of parameters of the GR4J and TOPM models..... 16

839

840

841 **Figure captions**

842 Figure 1. Comparison of the ungaged, almost ungaged and gaged conditions for the
843 estimation of model parameters (N corresponds to the number of flow observations
844 under which the application of classic estimation procedures are no longer robust)..... 8
845 Figure 2. Location of the 609 French catchments used in this study (dots indicate the gauging
846 stations and solid lines the catchment boundaries)..... 15
847 Figure 3. Schematic diagrams of the GR4J (a) and TOPM (b) rainfall-runoff models (PE:
848 potential evapotranspiration; P: precipitation; Q: streamflow; X_i : model parameter i ;
849 other letters are internal state variables)..... 17
850 Figure 4. Distribution of model efficiency in calibration and validation in gaged conditions
851 (red and blue boxplots, respectively) and in validation in ungaged conditions with the
852 number of donor catchments (boxplots show the 0.1, 0.25, 0.50, 0.75 and 0.9
853 percentiles; the mean value is indicated on top of the frame and by the cross). 21
854 Figure 5. Mean model efficiency (C_{LQ} , C_{RQ} and C_Q) in validation over the catchments with
855 the number M of donor sets in almost ungaged conditions for the GR4J and TOPM
856 models. Each line represents the application of the method with a given number of
857 flow measurements N (from 1 to 50). Regional and local parameters are equally
858 weighted ($\alpha = 0.5$ in Eq. 1). 22
859 Figure 6. Mean model efficiency (C_{LQ} , C_{RQ} and C_Q) in validation over the catchments with
860 the weighting factor α of the regional information ($\alpha=0$: no regional information; $\alpha=1$:
861 only regional information) in almost ungaged conditions for the GR4J and TOPM
862 models. Each line represents the application of the method with a given number of
863 flow measurements N (from 1 to 50). The number M of donor catchments is set to
864 seven and nine for GR4J and TOPM, respectively. 24
865 Figure 7. Optimum values of the weighting factor α of the regional information in almost
866 ungaged conditions for the GR4J and TOPM models for increasing numbers of flow
867 measurements (from 1 to 50). The number N of donor catchments is set to seven and
868 nine for GR4J and TOPM, respectively. The solid line corresponds to the curve
869 defined in Eq. (4). α values were discretely tested with a 0.1 step. 26
870 Figure 8. Efficiency of the proposed approach in validation for almost ungaged catchments
871 compared to the ungaged and gaged conditions for the GR4J and TOPM models for an
872 increasing number of flow measurements (from 1 to 50). The number N of donor
873 catchments is set to seven and nine for GR4J and TOPM, respectively. 28

874	Figure 9. Efficiency of the proposed approach for the TOPM model and for the C_{RQ}	
875	evaluation criterion in validation, for increasing numbers of flow measurements (from	
876	1 to 50) and two donor catchment values (M) (squares for nine donors and dots for one	
877	donor).	29
878	Figure 10. Distribution of normalized best parameters for the GR4J model over the catchment	
879	set with the increase of flow measurements (the meaning of parameters is given in	
880	Table 2). Boxplots show the 0.1, 0.25, 0.50, 0.75 and 0.9 percentiles.....	31
881		