Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment–MOPEX. IAHS Publ. 307, 2006.

Regionalization of a monthly rainfall–runoff model for the southern half of France based on a sample of 880 gauged catchments

NATHALIE FOLTON & JACQUES LAVABRE
Cemagref, BP 31, 13612 Aix-en-Provence, France
nathalie.folton@cemagref.fr

Abstract The aim of this study was to provide a means of determining reference low flows at any point in the hydrographic network in the southern half of France, using a continuous monthly rainfall–runoff modelling approach. The model adopted here has two parameters that were calibrated on 880 catchments. The accuracy of the calibrated model was excellent for assessing the mean annual discharge, and limited errors were observed in predicting low flows. A regionalization procedure was performed for the two parameters, using an upward approach and adopting a grid of set values based on environmental data. Based on comparisons with optimum parameters, a regional map of corrected parameters was drawn up. This model satisfactorily simulates mean annual discharge, and the level of uncertainty observed on low flow values is acceptable, given the difficulty in estimating extremely low flows.

Key words low flows; rainfall–runoff model; regionalization

INTRODUCTION AND SCOPE

The French Water Law of 1992, enforced by decree no. 93-742 in March 1993, sets the 5-year return period mean monthly flow (QMNA5) as the standard legal basis to deliver authorizations for wastewater discharge and freshwater withdrawal in rivers. The minimum discharge of rivers downstream of dams is defined by the article L 232-5 of the French “Code Rural” (Law no. 84-512 of 2 June 1984) as a percentage of the mean annual discharge.

However, these legal provisions raise problems as to how to determine these reference flows at any point of the national hydrographic network. Of course, the discharge of some rivers has been recorded at gauging stations. However, in most cases flows are still unknown. The French Environment Agencies (DIREN) and State water services (MISE) responsible for implementation of the water policy are confronted by this lack of knowledge on flows. In the absence of any overall studies, agencies have to rely on their own empirical field knowledge to extrapolate data from the information collected by the available recording networks.

The development of streamflow models based on rainfall data may help to solve these problems. One of the advantages of this approach is its broad scope, since it is not restricted to a single hydrological variable. The regionalization must be made here on the parameters of the rainfall–runoff model. It provides a means to simulate historical flow time series at any point of the drainage network. From these simulated flow series, it is possible to deduce the hydrological variables of interest. The GR2M model (Makhlouf & Michel, 1994) was used here. It simulates the transformation of rainfall into runoff at a monthly time step. This model was tested on catchments ranging in size from a few km^2 to more than 1000 km^2. Its main advantage is its simplicity and the limited number of parameters to optimize (only two free parameters). An improved version of the model was proposed by Folton & Lavabre (2006) to cope with the uncertainties found on some mountainous catchments. A snowmelt module was introduced in the model, without adding new free parameters, to ease the regionalization procedure. All the runoff variability is therefore accounted for by only two parameters.

To apply the model to an ungauged catchment, it is necessary to first relate model parameters to the physical and/or climatic characteristics of the catchment. The method classically used to regionalize the model parameters consists of establishing correlations of this kind. The aim here is to identify the physical variables that affect the rainfall–runoff transformation, and then to look for a regression relationship between model parameters and physical characteristics. This regionalization approach has been widely used by hydrologists. Many authors have adopted it using various models under various climatic conditions: Servat & Dezeter (1992), Sefton & Howarth (1998), Seibert (1999), Post & Jakeman (1999), Perrin (2000) and Mwakalila (2003), to name but a few. Other authors such as Vandewiele & Atlabachew (1995) applied a spatial interpolation technique to the parameters. Fernandez & Vogel (2000) used simultaneous parametric calibration methods and expressions relating the parameters to physical descriptors. The success rates obtained with all these methods depend on the model used and the regional context, but the outcome was often not very satisfactory. The best results were often obtained in specific contexts. More recently, solutions based on neighbouring regions were proposed, especially for flood quantile estimation. To determine the similarity between two stations, physiographic and hydrological data are used in methods known as the "regions of influence" method developed by Burn (1990), and the canonical correlation analysis method (Ouarda *et al*., 2001). Hydrological homogeneous regions are often identified for one or two flood quantiles, and the analysis has to be repeated to deal with other variables.

In the present study, a classical parametric regionalization approach was adopted. However, the method tested here also involved the use of an original upward procedure. A physical grid was set to assess model parameters. Then a regional grid based on the comparison with optimum calibration parameters was used to correct parameter values.

DATA AND MODEL

Study area

The study area includes 43 French administrative departments and is approximately 256 000 km^2. The contrasts of the climate in this region are extreme, ranging from the Mediterranean coastal climate to the high Alpine mountainous climate. The mean annual rainfall ranges from 500 to 2500 mm. This yields a strong hydrological variability:

– the mean annual discharge ranges from only a few L s^{-1} km^2 near the coastline to 70 L s^{-1} km^2 in some mountain catchments;
– the 5-year return period mean monthly flow (QMNA5) value ranges from 0 to more than 15 L s^{-1} km^2.

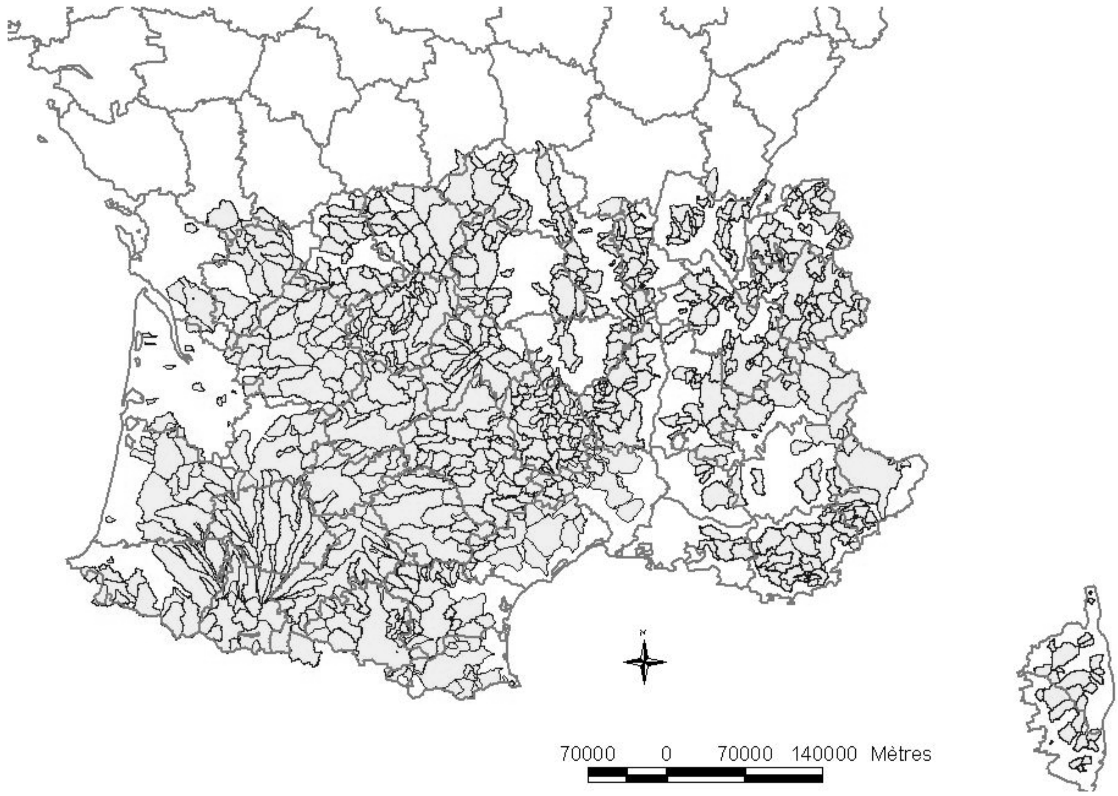


Fig. 1 Map of the study catchments.

All the available hydrological information was used in this study. Data on 880 catchments were used. 36% are smaller than 100 km^2, and 88% smaller than 1 000 km^2 (Fig. 1). Data collected at approximately 3200 rainfall recording stations and 68 potential evapotranspiration (PET) measuring stations were used.

We also used 1 km^2 pixel maps of mean monthly temperatures and mean monthly rainfall, which were drawn up using the AURHELY's method (Benichou & Le Breton, 1987). Temperature data were used to draw up a PET map (Folton & Lavabre, 2001, 2004). Topographic information was used to calculate catchment contours and derive monthly catchment rainfall maps (Folton & Lavabre, 2001).

A large variability of the mean annual discharge was observed over the study area. For more than 68% of the catchments, the mean annual runoff was below 750 mm, i.e. approximately 25 L s^{-1} km^2, but for a limited number of catchments (less than 10%), it was larger than 1250 mm, i.e., approximately 40 L s^{-1} km^2. Table 1 shows that mean annual discharge logically increases with mean annual rainfall.

Table 1 Distribution of catchments in classes of mean annual runoff and mean annual discharge. (scale: 1 mm = 0.0317 L s^{-1} km^2).

		Mean annual rainfall classes (mm)								
		0–250	250–500	500–750	750–1000	1000–1250	1250–1500	1500–1750	1750–2500	
Mean annual runoff classes (mm)	0–250	–	–	2.6%	12.9%	0.6%	0.1%	–	–	
	250–500	–	–	0%	19.8%	12.6%	0.5%	–	–	68%
	500–750	–	–	–	1.2%	13.3%	4.7%	0.1%	0.2%	
	750–1000	–	–	–	0.4%	3.1%	5.6%	2.5%	0.1%	
	1000–1250	–	–	–	0.1%	1.7%	3.8%	4.7%	0.2%	
	1250–1500	–	–	–	0.1%	0.4%	0.7%	2.1%	1.0%	
	1500–1750	–	–	–	–	0.2%	0.4%	1.1%	1.5%	9%
	1750–2500	–	–	–	–	–	–	0.4%	1.2%	

The GR2M rainfall–runoff model

The GR2M model used here is a monthly lumped rainfall–runoff model. Its structure includes two reservoirs. Figure 2 shows a diagram of the model. To account for snow or ice influences in mountainous catchments, a monthly snowmelt module was added. Each of the two reservoirs included in the model depends on a parameter.

XV_1 is the parameter of the production function. This parameter governs the catchment water balance via a multiplication factor applied to the monthly discharges. In the model structure, this parameter corrects only the output flow. XV_1 is positive without upper bound. Values approaching 0 indicate a low yielding catchment, whereas values greater than 2 correspond to mountainous catchments, where the rainfall and snowfall inputs are generally underestimated by the gauging network.

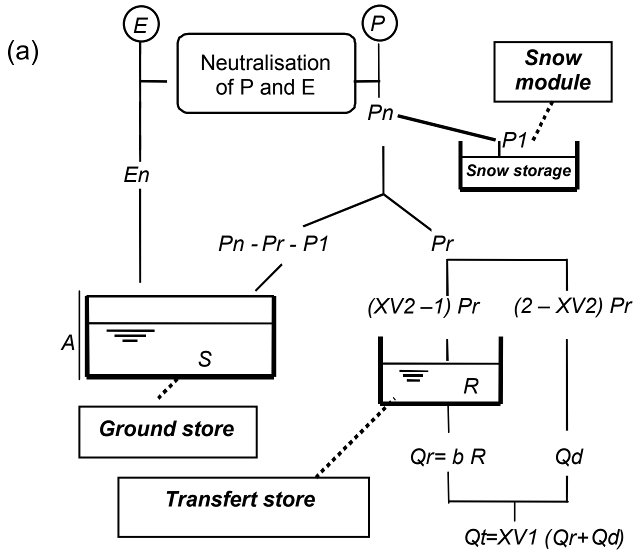
XV_2 is the parameter of the transfer function. It determines the temporal distribution of output discharge and accounts for the time lag between rainfall and runoff values. It actually determines the part of direct runoff that does not enter the transfer store. The value of XV_2 ranges between 1 and 2. A catchment responds more quickly as XV_2 approaches 1. In fact, if XV_2 is equal to 1, all the net rainfall for a given month is directly transferred to the outlet during this month. In this case, it will not be stored in the transfer store. Such values are characteristic of highly responsive catchments. If XV_2 is equal to 2, all the net rainfall enters the transfer reservoirs. This occurs in the case of catchments showing a high level of inertia.

MODEL RESULTS

Model performance

The analysis of mean annual discharge and QMNA5 values simulated by the model indicates its ability to produce accurate reference low flow estimates (when calibrated).

The mean annual discharge values were predicted extremely accurately by the model (Fig. 3(a)). The coefficient of determination R^2 (coefficient of correlation measure), has a value of 99.8%. The simulation of the QMNA5 variable were less accurate, but still acceptable (Fig. 3(b)). The overall R^2 was 89.7%.



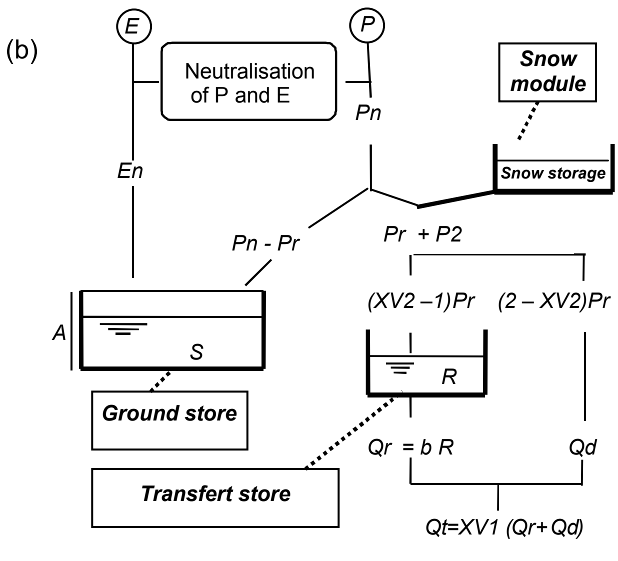


Fig. 2 Model structure with (a): storage of a part of net rainfall (*Pn*) in "snow module"; and (b): snowmelt. *P*, rainfall; *E*, evapotranspiration; *Pn*, net rainfall after evaporation; *Pr*, runoff-effective rainfall; *En*, net evapotranspiration; *A*, maximum capacity of the ground store; *S*, level of this store; Qd, direct flow; Qr, discharge from the transfert store with b = 0.4; *R*, level of this store; *Qt*, total flow; *XV*1, *XV*2: parameters.

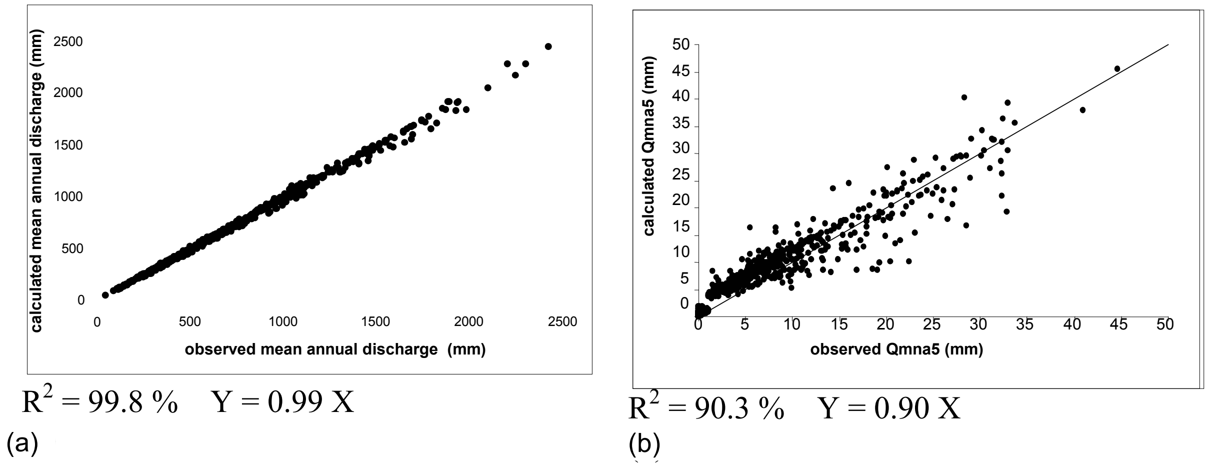


Fig. 3 Comparison of the observed and calculated values (using parameters calibrates for the GR2M model).

Parameter variability

No link could be established between the two model parameters and catchment area (the coefficients of determination were found to be not significant). There is also no significant correlation between model parameters and mean annual rainfall. The lack of significant correlations observed in both cases shows that model parameters mainly reflect the hydrological behaviour of the catchments. It is also important to note that the production and transfer parameters are independent, which guarantees better efficiency.

XV_1 is the production parameter. For similar rainfall, runoff is greater as the value of XV_1 increases. As shown in Fig. 4, for 75 % of the catchments, the value of XV_1 ranges between 0.75 and 1.25. For about 1% of the catchments, the value of this parameter was very low, since it was below 0.5. Conversely, for 8% of the catchments XV_1 showed high values.

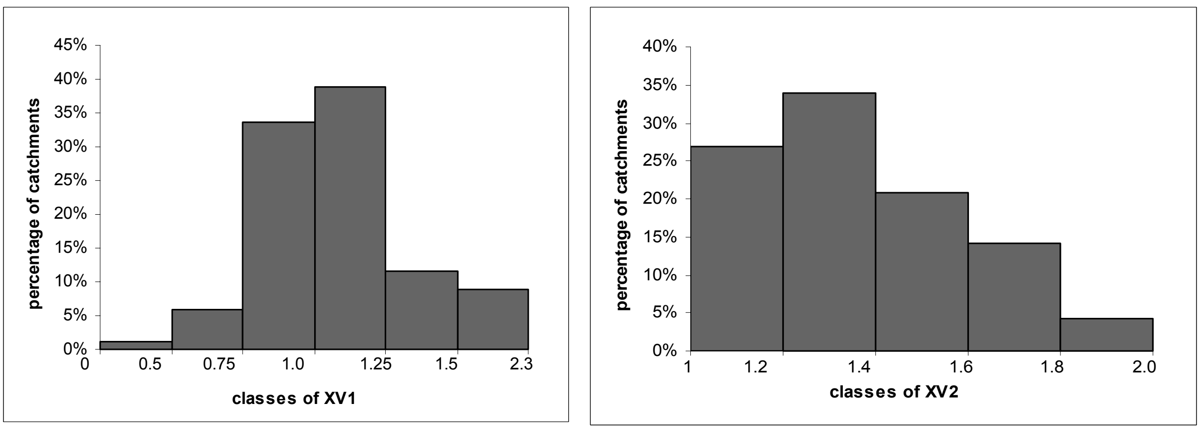


Fig. 4 Distribution of parameters XV_1 and XV_2 for the 880 catchments.

By its construction, the transfer parameter XV_2 ranges between 1 and 2. A value of around 1 reflects a highly reactive hydrological behaviour; whereas a value of around 2 means that all the rainfall is transferred by the routing reservoir and that the runoff will therefore be delayed. 55% of the catchments showed XV_2 values between 1.2 and 1.6, which corresponds to classical hydrological behaviour. Of the catchments (mountainous catchments), 4% had a greatly delayed runoff and 27%, an extremely fast runoff.

A parametric sensitivity analysis was performed to determine the influence of each parameter on the mean annual discharge and on the QMNA5. Figures 5 and 6 show the variation of these parameters with mean annual discharge and the QMNA5. As shown in Fig. 5, XV_1 was quite proportional to the mean annual discharge, whereas the effects of XV_2 on the mean annual discharge were negligible. Conversely, Fig. 6 shows the regular increase of XV_1 and XV_2 parameters when the QMNA5 values increase. On average, catchments with extremely low flows were associated with low XV_1 and XV_2 values, whereas catchments with consistently low flows were associated with high XV_1 and XV_2 values.

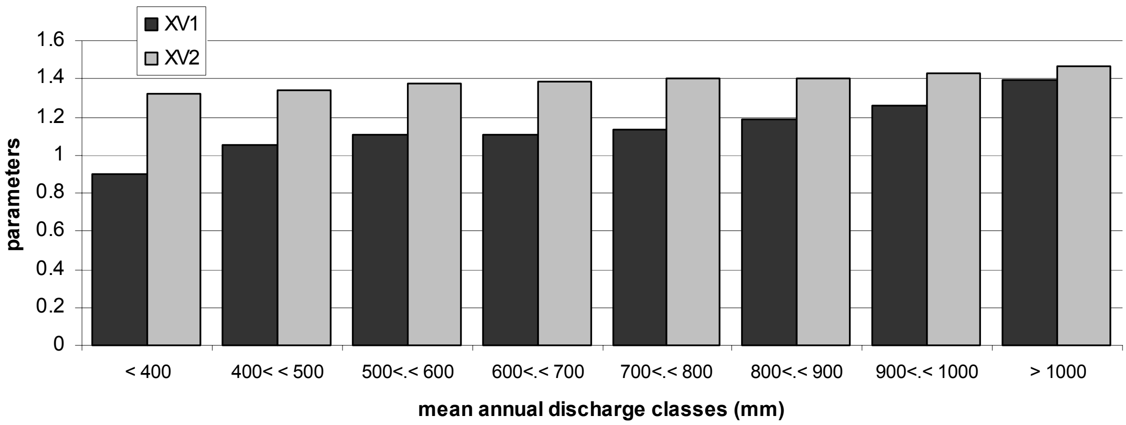


Fig. 5 Variation of XV_1 and XV_2 parameters with mean annual discharge.

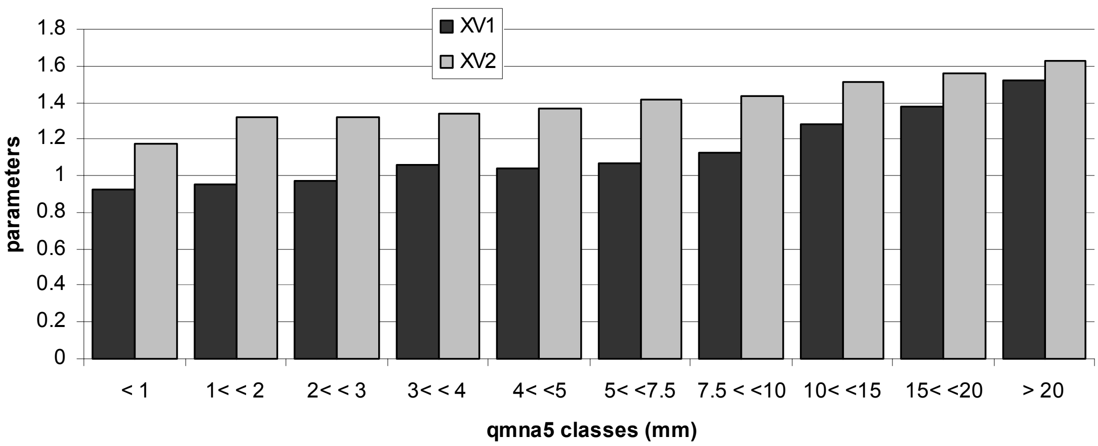


Fig. 6 Variation of XV_1 and XV_2 parameters with QMNA5.

THE REGIONALIZATION METHOD

Spatial information used in the regionalization procedure

The calibrated value of the two parameters model being available for the 880 catchments, our aim was to draw up a map over the whole study area for these parameters. It was therefore proposed to link the values of the two model parameters with some physical descriptors of the catchments: geology, land cover and altitude.

The following spatial information digitised on a 1 km^2 grid was used:

– The INRA Europan Soil Geographical Database on the 1/1 000 000 scale was used. The parent material class was used to determine the material from which the soil was originally formed. The system of classification used is based on the types of rock rather than on the geological age. This system of classification was developed at the European level using the chemical composition, the grain size and the origin of materials. There are 10 classes of materials in France.
– Based on the CORINE land cover database all over France, several simplified typological patterns were identified. The CORINE data system is based on a heirarchicalized 3-level list of 44 items classified in terms of the main land cover patterns. The following five categories were selected here to characterize the impact of land cover on runoff: artificial landscapes, cultivated agricultural landscapes, natural landscapes, semi-natural landscapes and forests.
– A Digital Elevation Model was used to determine the mean altitude of the catchments.

The attempts made to link model parameters with these physical features turned out to be most disappointing. Due to the low correlation obtained in the multiple regression analyses, it was not possible to link model parameters with the selected catchment descriptors. This means that the differences in the catchments' hydrological behaviour reflected by the two model parameters involved uncertainties of several kinds: rainfall amounts, runoff values and the modelling approach. These cumulated errors have a level of impact similar to the variability of the runoff due to the specificity of the catchments. It was therefore very difficult to interpret the values of XV_1 and XV_2 directly derived from catchment descriptors. In addition, the explanatory variables used here are lumped variables that did not reflect the heterogeneity of the catchments that may also vary in time (Perrin, 2000). The only significant trend was observed between parameter XV_1 and the mean altitude of the Alpine catchments. But this correlation was due to the fact that parameter XV_1 accounts for the uncertainty introduced by the underestimation of rainfall. Many authors, such as Servat & Dezetter (1992), Sefton & Howarth (1998), Seibert (1999), Post & Jakeman (1999), Perrin (2000) and Merz & Blöschl (2004), have faced similar problems. In studies made at a regional level, more satisfactory correlations have often been obtained. For example Makhlouf (1994) obtained encouraging results in regionalizing parameters of a previous version of the GR2M model in Brittany (western France). Correlations are less difficult to obtain for a specific zone than for a whole country, as noticed by Johansson (1994) and Post (1999). These authors concluded that the significant correlations obtained in their study were apparently due to the more homogeneous behaviour of the catchments studied.

Given these problems, it was proposed to use an upward regionalization procedure involving the use of physical descriptors. The idea here was to use a physical assessment grid to evaluate the two model parameters, and then to compare the parameters calculated from these grids with those given by the model at each gauged catchment. This comparison yielded a map of corrective coefficients defined as the ratio between the parameters evaluated using the physical grid and those obtained by calibration. A map of corrective coefficients was then drawn up for each parameter. The final grid was obtained by combining the physical grid and the map of corrective coefficients.

Regionalization methods

To draw up the physical parameter evaluation grid, the previous spatial data were used. These included:

– 5 classes of vegetation cover
– 10 geological classes

These data were interpreted as substratum retention data. A multiplication coefficient corresponding to the median spatial value of each parameter was attributed to each of the classes. The coefficient to be applied to parameter XV_1 was such that the parameter increased with decreasing substrate retention values, and *vice versa* in the case of the coefficient applicable to XV_2. The significant correlation previously observed with the altitude was taken into account by including an altitude correction coefficient ($Ccor_{Altitude}$), to improve the estimation of model parameters in the Alpine catchments.

The final grids for the two parameters were eventually drawn up pixel by pixel, by multiplying the grids of physical corrective coefficients by the median value of the parameter.

$$XV_{physique} = Ccor_{Vegetation} * Ccor_{Geology} * Ccor_{Altitude} * \text{Median value of the parameter}$$

The Alpine region showed the highest XV_1 values, whereas the sandy Aquitaine catchments showed the lowest values.

These parameter values were integrated for each of the 880 catchments. It was thus possible to compare the XV_1 and XV_2 values obtained by model calibration with those obtained using the physical estimation grid.

A corrective coefficient defined as the ratio between the "physical" value and the calibrated value was therefore calculated for each catchment. The next step consisted of drawing up a grid (with pixels corresponding to 1 km^2) giving the corrective coefficient applicable to each of these parameters. The regional approach is based here on 2/3 of the catchments. The catchments with the smallest areas were selected to avoid giving too much weight to the large catchments. These larger catchments will be used for control purposes to check the results of the regionalization procedure. Figure 8 gives the corrective ratio applicable to XV_1 in the case of the calibrated sample. This ratio was obtained by spatially drawing ratio values, starting from the larger catchments and finishing with the smaller catchments. It was then necessary to fill the empty spaces in the grid, corresponding to the zones where there was no catchment data. We used an automatic procedure, in which the study area was automatically segmented successively and the pixels completed by calculating the mean corrective coefficient between the patterns of segmentation (i) and (i – 1). Eight successive segmentation steps were carried out in this way on our study area.

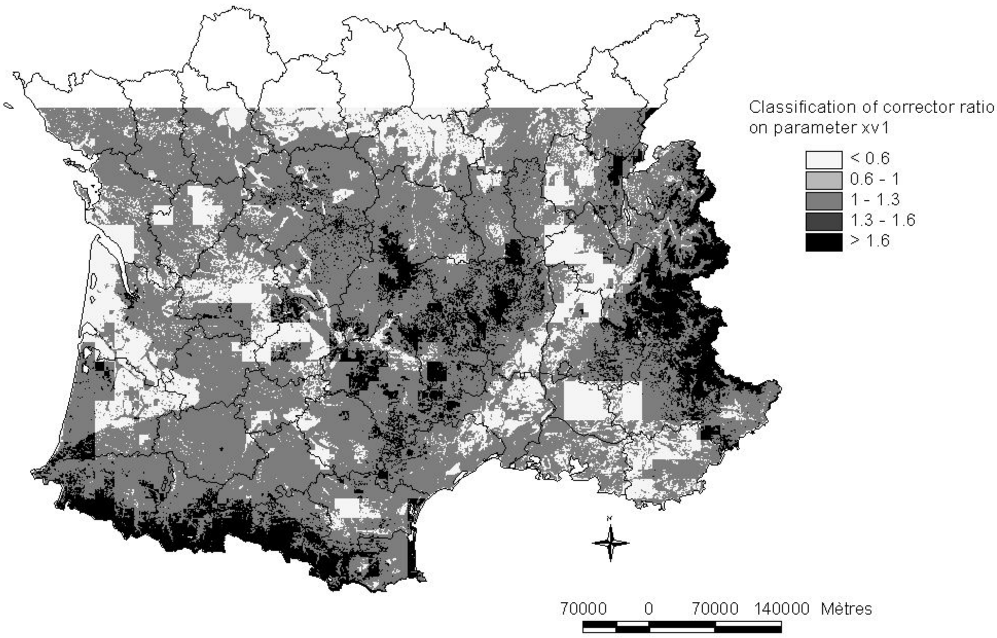


Fig. 7 Map of corrective ratios applicable to parameter XV_1 after the regionalization procedure.

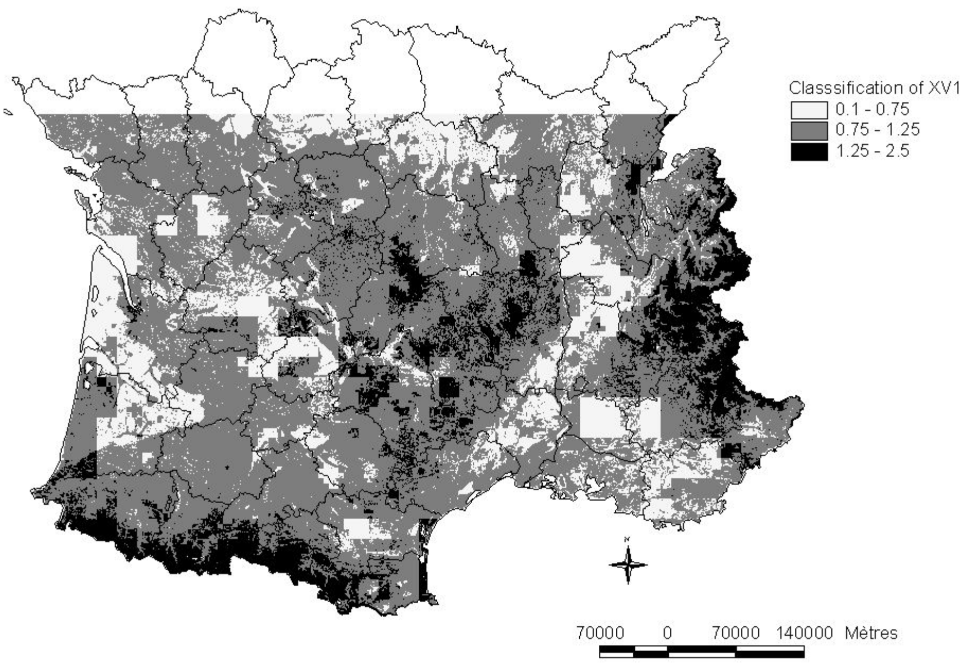


Fig. 8 Final map of regional parameter XV_1.

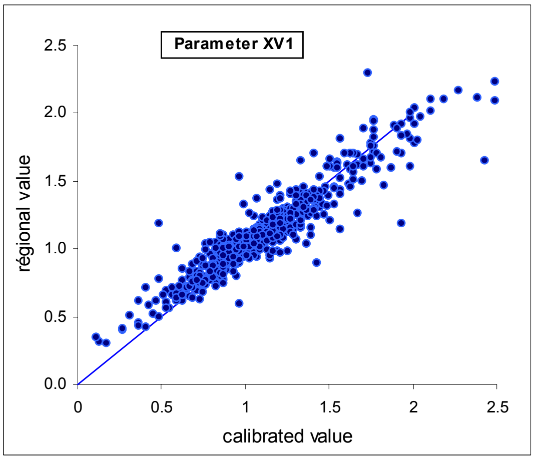
Let us take the following notations:

- D, the domain under investigation, which has an area of 1280 × 1280 km^2;
- M_D, the mean corrective coefficient calculated on domain D. This mean is weighted by the surface area of the studied catchments;
- n, the number of successive segmentation steps, $n = 8$.

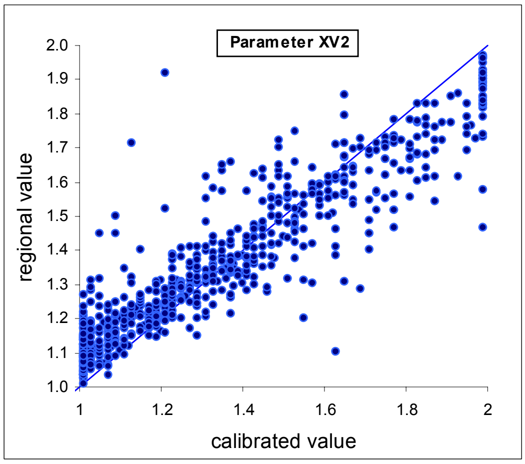
At each segmentation step, the value of the corrective coefficient ($Ccor$) for a pixel was obtained by applying the following formula:

$$Ccor = \frac{M_d}{2_n} + \sum_{i-1}^{n} \frac{M_{D/4^i}}{2^{n-i+1}} \quad \text{if } M_{D/4^i} = \text{no data, then } M_{D/4^i} = M_{D/4^{i-1}}$$

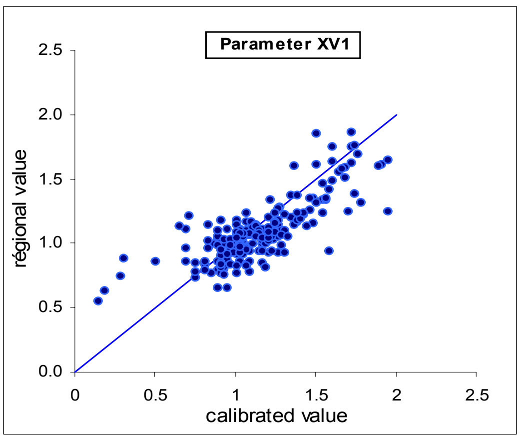
The final grid of the corrective ratios applicable to XV_1 was obtained after performing eight segmentation steps. Performing further segmentation steps did not significantly improve the results.



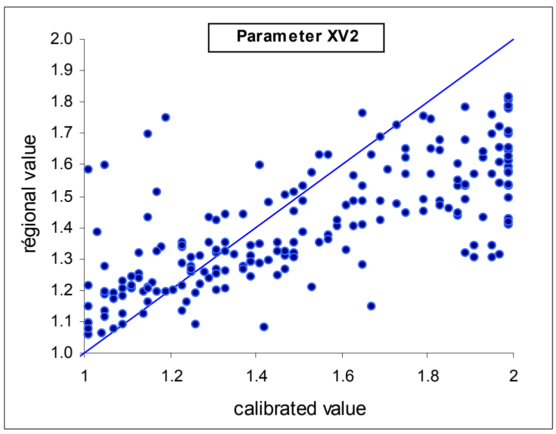
$R^2 = 87\%$ Y= 0.99 X
Calibration sample
(a)



$R^2 = 82$ % Y = 0.98 X
Calibration sample
(b)

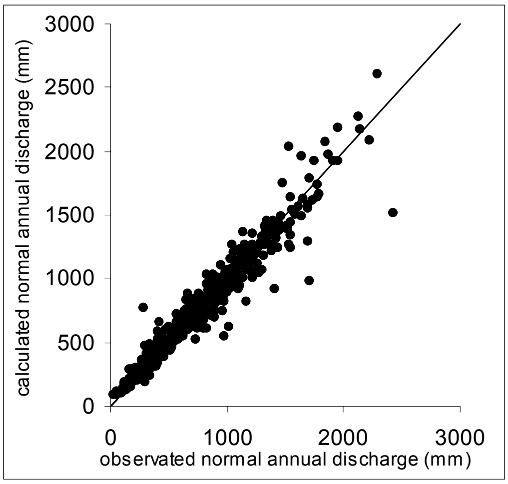


$R^2 = 56$ % Y= 1.06 X
Validation sample
(b)

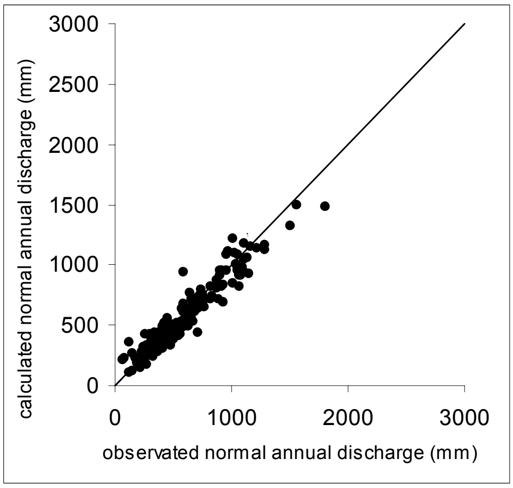


$R^2 = 41$ % Y = 1.08. X
Validation sample
(d)

Fig. 9 Comparisons of calibrated parameters and regionalized.



R^2 = 95.1 %   Y= 0.98 X
Calibration sample
(a)



R^2 = 91.1 %   Y = 0.93 X
Validation sample
(b)

Fig. 10 Mean annual discharge obtained with the regional model.

The final stage in the estimation of parameter XV_1 consisted of simply multiplying the map of physical parameters XV_1 by the corresponding corrective map. These two maps are given in Figs 7 and 8.

The procedure was repeated with parameter XV_2.

RESULTS

The regional parameter grids can be used to obtain mean values by integration over the study catchments. Fig. 9(a)–(d) compare the calibrated and regionalized parameters, for the calibration and validation catchments samples.

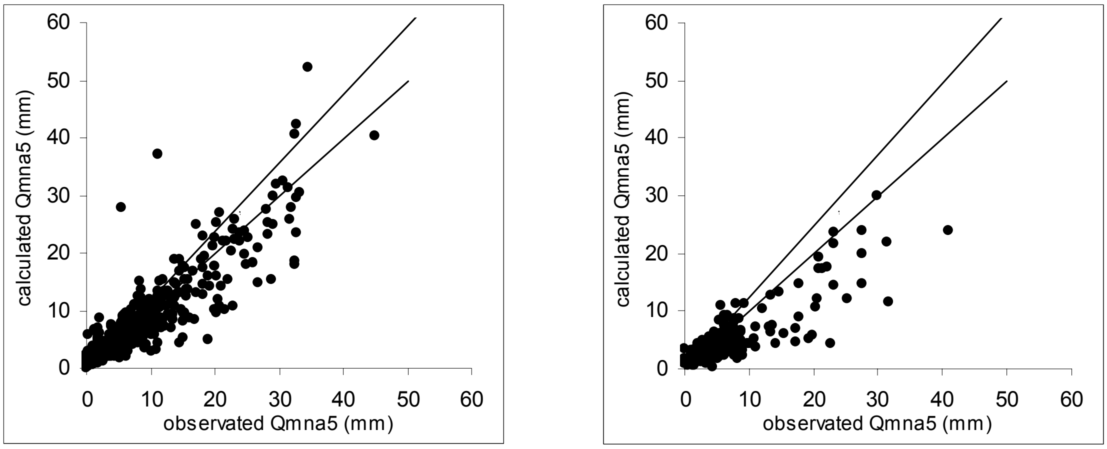
Using the regionalized parameters XV_1 and XV_2, the rainfall–runoff model can now be used to simulate the monthly flow time series.

Figure 10(a),(b) shows the simulated mean annual discharges. The values are accurately simulated during the calibration model, in spite of a slight underestimation. The results obtained after the parameter regionalization were slightly less accurate than in calibration, but they generally remain very satisfactory. The R^2 coefficients were greater than 90% in both samples.

In terms of QMNA5 variable, the accuracy of the regional model was acceptable, although the points showed greater scatter (Fig. 11).

CONCLUSION

The rainfall–runoff modelling approach used in the present study was found to very accurately simulate the monthly discharge series. Using these simulated series, satisfactory estimates of the reference low flow discharge variables (the mean annual discharge and the QMNA5) were obtained.



$R^2 = 81.5\%$ $Y = 0.81\ X$
Calibration sample
(a)

$R^2 = 61.5\%$ $Y = 0.67\ X$
Validation sample
(b)

Fig. 11 Results obtained for the QMNA5 using the regional model.

The model was calibrated on 880 catchments located in the southern half of France. Maps of the two parameters used in the model were obtained. Using this information, a regionalization approach was used to evaluate model parameters at any point of the whole study region.

The regionalization procedure involved several phases:

– drawing up a grid for evaluating model parameters from physical characteristics,
– establishing a corrective coefficient, defined for each catchment as the ratio between the parameter value obtained from physical characteristics and the parameters obtained by calibration.
– automatically interpolating the corrective coefficients,
– overlaying the grid of corrective coefficients with the grid based on physical criteria.

We therefore ended up with maps of the two model parameters.

Using this information to estimate parameter values, the rainfall–runoff model could be applied to simulate monthly discharge time series, from which reference low flow variables could be deduced.

The comparisons between the values obtained by this procedure and the observed values show that the mean annual discharges were accurately simulated by the model. Despite the scatter observed, the accuracy of the 5-year return period monthly low flow (QMNA5) given by the model was satisfactory.

REFERENCES

Benichou, P.& Le Breton, O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie*, 7e série **19**, 23–34.

Burn, H. D. (1990) Evaluation of Regional Flood Frequency Analysis with a region of influence approach. *Water Resour. Res*. **26**(10), 2257–2265.

Fernandez, W., Vogel, R. M., Sankarasubramamian A. (2000) Regional calibration of a watershed model. *Hydrol. Sci. J.* **45**(5), 689–707.

Folton, N & Lavabre, J (2006) Amélioration d'un modèle pluie débit au pas de temps mensuel. Cemagref, France (Personal communication).

Folton, N. & Lavabre, J. (2004) Estimation des débits de référence d'étiage. Nouvelle estimation de l'ETP. Nouvelle optimisation des paramètres du modèle pluie-débit. Nouvelle approche régionale de la modélisation de la pluie en débit aux régions Languedoc-Roussillon, Provence-Alpes-Côte d'Azur, Corse, Auvergne, Limousin, Rhône-Alpes, Rapport d'étude Cemagref.

Folton, N. & Lavabre, J. (2001) Estimation des débits de référence d'étiage. Cartographie régionale de l'évapotranspiration potentielle. Calcul des pluies de bassin. Rapport d'étude Cemagref, France.

Johansson, B. (1994) The relationship between catchment characteristics and the parameters of a conceptual runoff model: a study in the south of Sweden. In: *FRIEND: Flow Regimes from International Experimental and Network* Data (ed. by P. Seuna, A. Gustard, N. W. Arnell & G. A. Cole), 475–482. IAHS Publ. 221. IAHS Press, Wallingford, UK.

Makhlouf, Z. & Michel, C. (1994) A two-parameter monthly water balance model for French watersheds. *J. Hydrol.* **162**(3–4), 299–318.

Makhlouf, Z. (1994) Compléments sur le modèle pluie-débit GR4J et essai d'estimation de ses paramètres. Thèse de Doctorat, Université Paris XI Orsay, France.

Merz, R. & Blöschl, G. (2004) Regionalization of catchment model parameters model. *J. Hydrol*. **170**, 277–291.

Mwakalila, S. (2003) Estimation of stream flows of ungauged catchments for river basin management. *Physics Chem. Earth* **28**, 935–942.

Ouarda, T. B. M. J., Girard, C., Cavadias, G. S. & Bobée, B. (2001) Regional flood frequency estimation with canonical correlation analysis. *J. Hydrol*. **254**, 157–173.

Paturel, J. E., Servat, E. & Vassiliadis, A. (1995) Sensitivity of conceptual rainfall- runoff algorithms to errors in input data—case of the GR2M model. *J. Hydrol*. **168**, 111–125.

Perrin, C. (2000) Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Thèse de Doctorat, Institut National Polytechnique de Grenoble, Grenoble, France.

Post, D. A. & Jakeman, A. J. (1999) Predicting the daily streamflow of ungauged catchments in SE Australia by regionalizing the parameters of a lumped conceptual rainfall–runoff model. *Ecol. Modelling* **123**, 91–104.

Sefton, C. E. M. & Howarth, S. M. (1998) Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *J. Hydrol*. **211**, 1–16.

Seibert, J. (1999) Regionalization of parameters for a conceptual rainfall–runoff model. *Agric. For. Met*. **98–99**, 279–293.

Servat, E. & Dezetter, A. (1992) Modélisation de la relation pluie-débit et estimation des apports en eau dans le nord-ouest de la Coôte d'Ivoire. *Hydrologie Continentale* **7**(2), 129–142.

Vandewiele, G. L. & Atlabachew, E. (1995) Monthly water balance of ungauged catchments obtained by geographical regionalization. *J. Hydrol*. **170**, 277–291.