Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Journal of Hydrology 286 (2004) 19–35

Journal of Hydrology
www.elsevier.com/locate/jhydrol

Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models

Vazken Andréassian*, Charles Perrin, Claude Michel

Unité de Recherche Qualité et Fonctionnement Hydrologique des Systèmes Aquatiques, Cemagref, BP 44, Antony cedex 92163, France

Received 5 March 2002; revised 1 August 2003; accepted 12 September 2003

Abstract

This paper attempts to assess the impact of improved estimates of areal potential evapotranspiration (PE) on the results of two rainfall–runoff models. A network of 42 PE stations was used for a sample of 62 watersheds and two watershed models of different complexity (the four-parameter GR4J model and an eight-parameter modified version of TOPMODEL), to test how sensitive rainfall–runoff models were to watershed PE estimated with the Penman equation.

First, Penman PE estimates were regionalized in the Massif Central highlands of France, a mountainous area where PE is known to vary greatly with elevation, latitude, and longitude. The two watershed models were then used to assess changes in model efficiency with the improved PE input. Finally, the behavior of one of the model's parameters was analyzed, to understand how watershed models cope with systematic errors in the estimated PE input.

In terms of model efficiency, in both models it was found that very simple assumptions on watershed PE input (the same average input for all watersheds) yield the same results as more accurate input obtained from regionalization. The detailed evaluation of the GR4J model calibrated with different PE input scenarios showed that the model is clearly sensitive to PE input, but that it uses its two production parameters to adapt to the various PE scenarios.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Rainfall–runoff modeling; Sensitivity analysis; Potential evapotranspiration; Parameter uncertainty; Parsimony

1. Introduction

What knowledge on the evaporative demand is necessary to model the rainfall–runoff relationship? This question is crucial for watershed modelers since, at the watershed scale, atmospheric loss through evaporation and transpiration is an obvious and very often an important component of the water balance. Until now, studies on the sensitivity of rainfall–runoff (or watershed[1]) models to the uncertainty of their inputs have focused quite exclusively on rainfall. Although the attention devoted to rainfall is understandable, it is striking that so few studies have focused on the sensitivity of watershed models to potential evapotranspiration (PE) estimation. As a possible explanation, we can list the following specificities of PE estimates

* Corresponding author.
E-mail address: vazken.andreassian@cemagref.fr (V. Andréassian).

[1] Hereafter, the term 'watershed model' is used to describe continuous models, which include a conceptualization of rainfall storage and release at the watershed scale.

0022-1694/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2003.09.030

1. Watershed models need watershed-scale (or grid-scale) evaporation estimates, and obtaining such estimates is far from simple. As Brutsaert (1982, p. 2) points out, 'the regional estimation of the other phases in the (hydrological) cycle, such as precipitation or streamflow, involves formidable sampling problems. But in the case of evaporation beside sampling, there is also the problem of simply determining it at a point location'.
2. With PE formulations such as Penman's, there often is a problem of data availability. The Penman formula requires data on air temperature, total radiation, relative humidity, and wind speed. As such data are relatively scarce, the sensitivity of rainfall–runoff models to PE data is difficult to investigate.
3. Most watershed models can cope numerically with imperfect PE estimates (note that this ability exists to a much lesser extent with rainfall estimates, e.g. Andréassian et al., 2001). The quite surprising adaptability of watershed models may have numbed modelers and led them into believing that estimating the evaporative demand was not important.

As PE plays a major role in the long-term watershed water balance, it was feared[2] that imperfect estimates could either impede the calibration of watershed model parameters or modify their optimal values, and have a detrimental influence on model simulations. Therefore, a sensitivity study, whose aim was to determine the impact of imperfect PE knowledge on the efficiency and the parameters of watershed models, was done. It was limited to Penman estimates, since this formulation is considered the most physically satisfactory by many hydrologists (Beven, 2001, p. 60; Shuttleworth, 1993), as well as being the formulation selected by most meteorological services internationally. After a review of relevant literature in Section 2 and a presentation of the study area in Section 3, a three-step approach will be developed to test the sensitivity of rainfall–runoff models to PE

(a) First, improved Penman PE estimates are sought in the Massif Central highlands of France, a mountainous area where PE is known to vary rapidly with elevation, latitude, longitude, and aspect (Section 4).
(b) Two models of different complexity (described in Section 5) are used to assess changes in model efficiency with improved PE input quality (Section 6).
(c) Finally, parameter behavior in one of the models is analyzed to understand how it copes with biased estimations of PE input (Section 7).

The above analysis involves a network of 42 meteorological stations, a sample of 62 watersheds, and two watershed models of different complexity (the four-parameter GR4J model and an eight-parameter modified version of TOPMODEL, Table 1).

2. Relevant literature

2.1. Model sensitivity to PE: a neglected area of watershed model research

One of the first studies of watershed model sensitivity to errors in PE input was published by Parmele (1972). The author used three models and a sample of nine watersheds to assess the impact of PE errors on model efficiency. He compared streamflows simulated with erroneous PE to 'perfect' synthetic streamflow, obtained with 'true' PE. He used several rainfall–runoff models tested on a quite large watershed sample. However, this methodology was constrained by the available computing power (models were most often calibrated manually) and by the excessive optimism prevalent at the time as to the ability of the largely overparametrized watershed models used in this study to represent the hydrological cycle with unique representative parameters. Today, this approach would not be acceptable, since much more is known about the detrimental impact of

[2] For example, it seems that calibrating a model on a watershed whose mean elevation is 1000 m a.s.l., with PE data from a meteorological station located at an airport at an altitude of 300 m, might introduce a water deficit in the modeling phase that would only be an artifact. It is believed that, at least in France, this is not an uncommon engineering practice: most automatic meteorological stations used for PE computations are located at airports, and some hydrologists consider that French PE data are representative of an airport microclimate and not really of their surroundings (C. Scherer, personal communication).

Table 1
Characteristics of the 42 weather stations used to regionalize Penman PE in the Massif Central highlands

Station		Easting (km)	Northing (km)	Elevation (m)	Mean annual Penman PE (mm)
Code	Name				
03060001	CHARMEIL	682	2130	249	771
03155003	LURCY-LEVIS	647	2191	225	722
03185006	MONTLUCON	621	2151	207	809
03200001	NEUVY	673	2173	250	764
03248001	SAINT-NICOLAS-DES-BIEFS	714	2118	1022	578
07068001	COLOMBIER-LE-JEUNE	784	2005	585	937
07131001	LANAS	762	1951	280	1168
11004001	ALAIGNE	581	1790	293	991
11069001	CARCASSONNE	598	1801	126	1080
12145001	MILLAU	655	1902	715	961
12145011	MILLAU-BARRY	657	1900	409	892
12194002	QUINS	604	1915	632	820
12208004	SAINT-AFFRIQUE	641	1885	365	867
12300004	VILLEFRANCHE-DE-ROUER	575	1930	345	804
15014004	AURILLAC	607	1989	639	746
15187006	SAINT-FLOUR	658	2003	909	666
19031008	BRIVE-LA-GAILLARDE	532	2017	111	832
26198001	MONTELIMAR	791	1956	73	1119
30132003	LA GRAND-COMBE	734	1915	288	1144
30189001	NIMES	767	1875	59	1289
30339001	VALLER-MONT AIGOUAL	700	1903	1567	628
30341003	VAUVERT	759	1853	50	1294
34151005	MARSILLARGUES	748	1849	2	1001
34154001	MAUGUIO	731	1843	3	1157
42005001	ANDREZIEUX-BOUTHEON	753	2061	400	850
43046001	CHADRAC	723	2007	714	806
43062001	CHASPUZAC	709	2010	833	715
43096001	FONTANNES	685	2034	435	838
43111002	LANDOS	719	1986	1148	629
46127001	GOURDON	525	1972	259	776
48030001	BRENOUX	695	1946	1019	711
48095005	MENDE	689	1949	932	789
63098001	CHASTREIX	634	2059	1385	625
63113001	CLERMONT-FERRAND	664	2088	329	897
63319002	SAINT-ANTHEME	720	2064	1260	543
63354004	SAINT-GERVAIS-D'AUV	637	2115	705	734
63399001	SAINT-SULPICE	621	2072	850	672
69029001	BRON	802	2084	198	901
69174001	TARARE	759	2107	720	714
69299001	COLOMBIER-SAUGNIEU	814	2087	235	943
82039001	CAYRAC	533	1902	132	718
87085001	LIMOGES	510	2093	402	719

Latitude and longitude are expressed in the extended Lambert II coordinate system.

overparametrization on the nonuniqueness of the optimum parameter set, and on the limited but nonetheless real ability of models to cope with input errors (Andréassian et al., 2001; Paturel et al., 1995). However, Parmele was aware of some of these problems. He recognizes 'the nonuniqueness of the model parameters', and agrees that confronted with what he calls 'biased PE data', a hydrologist can change model parameters to cope with them. For Parmele, however, this would lead to 'accepting a set of parameters that are not representative of the hydrologic process being modelled'.

Andersson (1992) used the HBV watershed model (Bergström, 1995) and compared seven different methods of computing PE input. He used the same set of calibrated parameters in each case, allowing only the precipitation correction factor to compensate for PE over or underestimates to obtain, over the calibration period, the same total runoff amount for all the formulae. Expressed in terms of model efficiency, the differences between methods were very small. In terms of model sensitivity, however, we consider that the actual result of this study is that the HBV model can probably compensate for PE estimation biases by adapting its precipitation correction factor without noticeable efficiency loss. It is interesting to note here how the author dealt with model adaptability and with the need to compare the effect of differences in PET time modulations rather than differences in PET mean values.

Joukainen (2000) also used the HBV model and modified the routine for computing actual evapotranspiration (AE) from PE to achieve a better representation of rainfall interception by trees, adding eight new parameters to the model. She found a very slight improvement in the calibration results, although she nearly doubled the degrees of freedom of the model. Clearly, the HBV model is not sensitive to such refinements of its AE computation routine.

Paturel et al. (1995) assessed the sensitivity of the GR2M monthly watershed model to systematic PE errors. Initially using the same approach as Parmele (1972), without parameter recalibration, they found that, compared to errors in rainfall, systematic PE errors induced much smaller output errors. Then they studied the ability of the model to compensate for errors by calibration. They concluded that watershed models have a certain capacity to 'absorb systematic input errors'.

Nandakumar and Mein (1997) studied the effects of random systematic errors in pan coefficients and model parameters on the predictions of a rainfall–runoff model. They used a model with 13 optimized parameters, into which they introduced a 14th parameter to adjust potential evaporation from pan measurements. They found that a bias in the potential evaporation estimate does not have as great an effect as one in the rainfall estimate, but that it remains significant (10% bias in potential evaporation can cause up to 10% bias in runoff predictions). Incidentally, the authors noted that the 14th parameter could take 'exceptionally high' values, and suggested that it was because of 'moisture leakage from the catchment (i.e. unknown water losses)'. This illustrates how a parameter initially meant for PE adjustment can be used in a different way by the calibration process.

In an exercise aimed at demonstrating the adaptive ability of the IHACRES conceptual watershed model, Kokkonen and Jakeman (2001) modified the formulation used to compute AE from PE. The modification resulted in an increase in the contrast in evapotranspiration losses as computed by the model: values were much higher in summer and much lower in winter. However, this modification did not affect the ability of the conceptual rainfall–runoff model to adequately represent the rainfall–runoff relationship. This is another example of the adaptability of watershed models: they can use some of their internal degrees of freedom to balance the inflated amplitude of evaporative losses and produce acceptable streamflow simulations.

Vázquez and Feyen (2003) tested three different PE formulations as input to the MIKE-SHE model and calibrated this model with each of the formulae. The authors report large differences, not only in control mode (Nash = 76, 44, 39%), but also in calibration mode (Nash = 73, 66 and 63%). These substantial differences seem rather surprising, since the model should have enough degrees of freedom to adapt to differences in PE estimation. These results on the sensitivity of a watershed model to PE input would tend to contradict the rest of the literature on the subject.

To study the possible advantage of using real time series instead of long-term averages for PE, different solutions have been tested with the GR4J watershed model. Edijatno (1991) used a sample of ten watersheds, where he compared long-term averages of 10-day PE and actual 10-day PE as input in the rainfall–runoff model. He found that using actual PE could result in a higher as well as a lower efficiency. Moreover, the mean absolute change in model efficiency was very slight. With a different sample of three watersheds and daily PE data, Kribèche (1994) obtained similar results.

Fowler (2002) used a daily soil water balance model, for which he compared simulations made using actual PE and long-term average PE. He found that the substitution of actual values with long-term averages produced a soil water regime very similar to that derived using actual PE values, including over relatively extreme periods. All the above-mentioned results support Burnash's (1995) conclusion: 'in many areas an average annual evapotranspiration curve appears to be as meaningful as any readily available discrete information'. Therefore, in this paper, the focus is put on the evaluation of the long-term averages of PE in watershed models.

2.2. Two different approaches to the sensitivity analysis of watershed models

Let us start this discussion by defining what we call Sensitivity Analysis (SA). SA has several different meanings: in its broader sense, it studies how the variation in the output of a model can be apportioned, qualitatively or quantitatively, to different sources of variation. SA aims to ascertain how the model depends on the information fed into it, its structure, and the framing assumptions made to build it (Saltelli et al., 2000).

So as not to confound different entities, we propose to classify the input-related sensitivity analyses of watershed models into two categories: (1) static and (2) dynamic.

1. *Static sensitivity studies* are those that explore model sensitivity to PE estimates by first obtaining a calibration considered to be optimal and then leaving it unchanged. Model sensitivity is assessed by comparing flows simulated with 'erroneous' PE and flows simulated with 'perfect' PE.
2. *Dynamic sensitivity studies* involve a reference calibration (and a corresponding reference streamflow simulation), using a reference PE. But model recalibration is allowed with erroneous PE, and the reference simulation is then compared to the flow simulated with the recalibrated watershed model.

Many of the studies described in the literature follow approach (1) Parmele (1972), Andersson (1992), Nandakumar and Mein (1997) and Joukainen (2000). The study by Paturel et al. (1995) uses a mixed approach. The studies by Edijatno (1991), Kribèche (1994), Fowler (2002) and Vázquez and Feyen (2003) adopt exclusively approach (2), which also was chosen in this paper to assess the impact of imperfect PE knowledge on watershed model efficiency.

Several authors identified shortcomings in their static approach to SA: by adapting the precipitation correction factor of the HBV model, Andersson (1992) carried out a sort of rainfall re-adjustment to the change in PE. Similarly, Joukainen (2000) stated that her approach 'assumes that other parameters of the model are not dependent on the calculation of evapotranspiration'. Since her conceptual model required calibration, she acknowledged that 'the changes in any calculation routines may have an impact on the optimal parameter values of other routines'.

The proposed classification might also reveal a more fundamental difference in modeling philosophy: with static sensitivity studies, modelers may assume implicitly that the true parameters are watershed-specific (with no decisive influence from the climatic input data), while with dynamic sensitivity studies, modelers acknowledge explicitly that the calibrated watershed parameters are dependent on climatic input data.

2.3. Which way to go?

It is remarkable that none of the studies reviewed in this section identified a line of research to improve the treatment of evapotranspiration in watershed models. Several explanations are possible

- First, with some notable exceptions (Paturel et al., 1995; Fowler, 2002), hydrologists interested in model sensitivity to PE seem to have underestimated the adaptive capacity of watershed models (Bras and Rodriguez-Iturbe, 1976; Bras, 1979; Storm et al., 1989; Andréassian et al., 2001). They focused on static sensitivity studies, involving calibration on one type of PE estimate, and evaluation of several types of alternative PE formulations without recalibration. Therefore, a possible explanation for the relatively uninspiring results of most previous studies is the static type of analysis.

- Second, another possible explanation for the relative insensitivity of watershed models to the type (actual time series as opposed to long-term average time series) of PE data is that they vary little from year to year. In the Massif Central highlands, for example, the coefficient of variation of annual PE time series is very low (mean value of 0.08, on a sample of 15 locations).
- Third, it seems that given the data requirements of the Penman formula, appropriate data are available only at a few locations. This problem is particularly acute in mountainous areas, where PE may vary rapidly in space. If models are fed with input that does not represent the reality of climatic forcing at the watershed scale, there would be little difference in relevance between actual and average input data.

The investigation reported in this paper focused on the possibilities of improving watershed modeling through a better regionalization of mean interannual PE, and the case of the Massif Central highlands in France was studied. This region is considered crucial for water resources, since it covers the headwaters of several large French rivers such as the Loire, the Dordogne, and the Tarn, with large hydroelectric reservoirs as well as important low-flow issues downstream.

In Section 3, the network of meteorological stations whose data were used to establish regional relationships is described, as well as the watershed sample used to assess the impact of improved regional PE knowledge on model efficiency.

3. Study area

3.1. Meteorological stations for PE computation

The study area covers the central mountainous part of France (Massif Central highlands). Elevation ranges between 100 and 1900 m a.s.l. (median: 700 m). Since the beginning of the 1990 s, the French Meteorological Service (Météo France) has installed new automatic weather stations, and the number of locations where Penman PE is available rose from 14 to more than 50, with an improved representation of higher elevations. The data from 42 of these stations, where at least 5 years of data were available (Fig. 1), were used in this study (Table 1). The elevation of the stations ranges between 2 and 1567 m (median: 400 m). From daily Penman PE values, annual and monthly long-term PE averages were produced for each station and used in the regionalization (Section 4).

3.2. Test watersheds

A sample of 62 watersheds was used, all of them situated in the Massif Central area (Fig. 1). The watershed size varies between 5 and 89 km^2 (median: 46 km^2). The mean watershed elevation ranges between 300 and 1430 m a.s.l. (median: 880 m). The watershed land cover is mixed (mainly forest and pasture). For each watershed, daily rainfall and runoff over a period of 6 years were used (3 years for model calibration and 3 years for model validation).

In Section 4, the initial step of the sensitivity study, i.e. the regionalization approach by which an improved watershed-scale estimate of Penman PE can be obtained and used as input in the rainfall–runoff models, is described.

4. Improved estimation of potential evapotranspiration (PE) in the massif central highlands

Since we suspected that the lack of sensitivity to the PE input could be attributable to its lack of accuracy, we initially tried to improve the PE data fed into rainfall–runoff models for watersheds where PE estimates could be deemed unsatisfactory because of the great variation in elevation. First, the results of regression equations are discussed at annual and monthly time steps. Then a Fourier series development for the monthly time step is presented, and it is shown how, with a simple time-scale change, the monthly equation can be adapted to yield long-term daily averages.

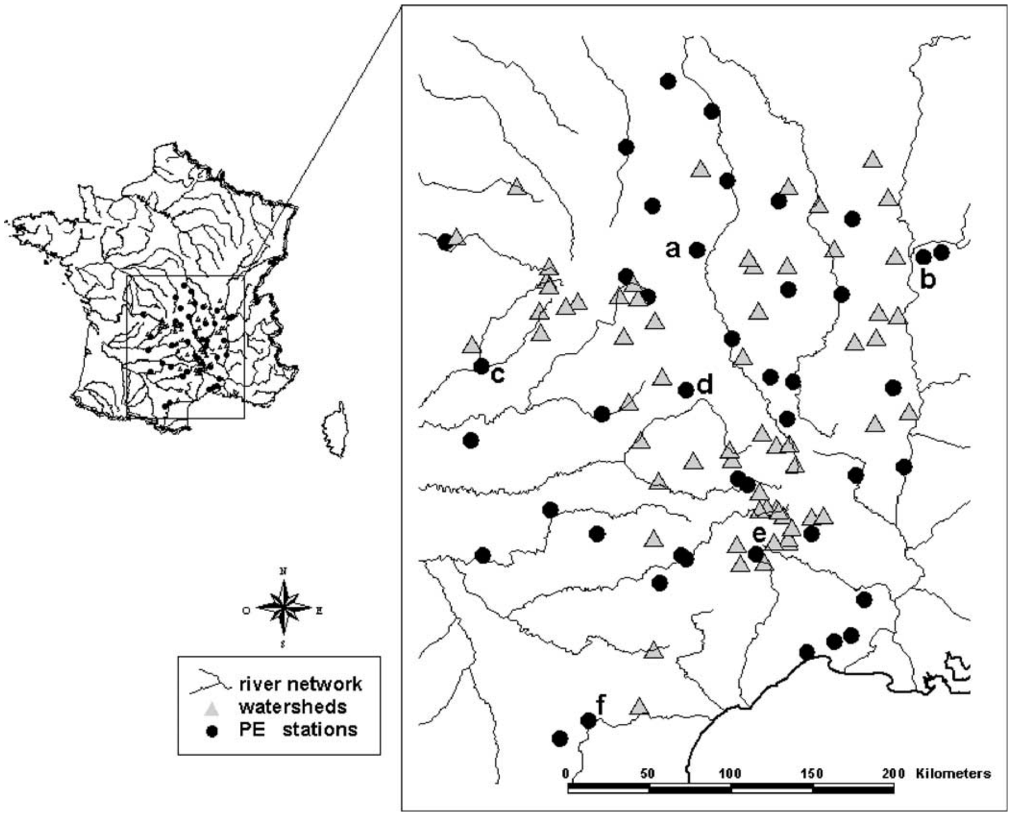

Fig. 1. Location of the 62 test watersheds and the 42 weather stations. The weather stations for which data presented are noted by a letter: Clermont-Ferrand (a), Lyon (b), Brives (c), Saint Flour (d), Mont Aigoual (e), and Carcassonne (f).

4.1. Regionalization of annual PE and its climatic interpretation

The investigation was started with a standard multiple regression on annual long-term averages, to see if, at least at that time step, PE totals could be explained by longitude (X), latitude (Y) and elevation (Z), as proposed in Eq. (1)

$$\mathrm{PE} = a + bX + cY + dZ + \varepsilon \tag{1}$$

The results of the regression with an annual time step proved encouraging: the three independent variables were significant (Table 2), and thus explain 84% of the PE variability in the Massif Central highlands. The results show that in the study area, PE decreases by 29 mm/100 m of elevation, decreases by 82 mm/100 km northwards, and increases by 92 mm/100 km eastwards. This last trend can be interpreted as an effect of continentality, as the main moisture source is the Atlantic Ocean, to the west of the area.

4.2. Regionalization of monthly PE

The same kind of regional relationship can be established at a monthly time step, as shown in

Table 2
Parameter estimates for the linear regression in Eq. (1)

Parameter	a	$b(X)$	$c(Y)$	$d(Z)$
Unit	mm	mm/km	mm/km	mm/m
Estimated value	2015	0.92	−0.82	−0.294
Student ratio (estimated value/standard deviation)	7.9	5.9	−6.9	−9.4
Significance level	>99	>99	>99	>99

Eq. (2) below

$$PE_i = a_i + b_i X + c_i Y + d_i Z + \varepsilon_i \quad \text{with } i = 1 \text{ to } 12 \tag{2}$$

Parameters and regression results are given in Table 3, which shows that geographic variables provide a good explanation of the spatial variations in monthly Penman PE from February to October, but a less satisfactory explanation during the three winter months (November through January). Explanatory variables, however, remain highly significant throughout the year. Note that elevation is the variable whose weight varies the most during the year (in a ratio of 1:7), while the weights of X and Y vary in a 1:3 ratio only.

4.3. Fourier series development

The regularity of the evolution of monthly parameters in the course of a year (Table 3) suggested that a Fourier series development might be used to reduce the number of parameter. We followed Fennessey and Vogel (1996), who proposed a development of the second order for the northeast of the United States. For x_i, successively equal to a_i, b_i, c_i, and d_i (from Eq. (2)), this development can be written

$$x_i \approx \frac{A_0}{2} + \sum_{n=1}^{2} \left(A_n \cos\frac{ni\pi}{6} + B_n \sin\frac{ni\pi}{6} \right) \tag{3}$$

with

$$A_n = \frac{1}{6} \sum_{i=1}^{12} x_i \cos\frac{ni\pi}{6} \tag{4}$$

$$B_n = \frac{1}{6} \sum_{i=1}^{12} x_i \sin\frac{ni\pi}{6} \tag{5}$$

$$A_0 = \frac{1}{6} \sum_{i=1}^{12} x_i \tag{6}$$

The Fourier series development parameters are given in Table 4.

The Fourier series development presents two advantages

- First, it reduces the number of explanatory parameters (20 instead of 48);
- Second, it makes interpolation (Eqs. (7) and (8)) easy and suggests a usable formula at a daily time step, i.e. the time step used in our watershed models.

$$PE_j = a_j + b_j X + c_j Y + d_j Z + \varepsilon_j \quad \text{with } j = 1, 2, \ldots, 365 \tag{7}$$

Table 3
Parameter estimates for the linear regression in Eq. (2)

Month	a (mm)	$b(X)$ (mm/km)	$c(Y)$ (mm/km)	$d(Z)$ (mm/m)	R^2
1	83	0.05	−0.04	−0.008	0.57
	4.0	*3.9*	*−4.6*	*−3.2*	
2	96	0.05	−0.05	−0.013	0.69
	4.9	*4.2*	*−5.2*	*−5.6*	
3	123	0.06	−0.05	−0.023	0.79
	5.9	*4.7*	*−4.9*	*−9.0*	
4	151	0.08	−0.06	−0.032	0.81
	5.7	*4.7*	*−4.5*	*−10.0*	
5	209	0.07	−0.07	−0.036	0.82
	7.4	*4.2*	*−5.2*	*−10.4*	
6	281	0.10	−0.10	−0.044	0.85
	8.4	*5.0*	*−6.6*	*−10.7*	
7	304	0.15	−0.12	−0.043	0.84
	8.0	*6.5*	*−6.7*	*−9.3*	
8	261	0.13	−0.10	−0.037	0.79
	6.9	*5.7*	*−5.8*	*−8.0*	
9	227	0.08	−0.10	−0.030	0.77
	6.9	*4.1*	*−6.2*	*−7.5*	
10	123	0.05	−0.05	−0.013	0.69
	6.2	*4.0*	*−5.4*	*−5.5*	
11	86	0.05	−0.04	−0.008	0.56
	4.1	*3.9*	*−4.6*	*−3.0*	
12	72	0.05	−0.04	−0.006	0.59
	4.1	*4.5*	*−4.8*	*−2.9*	

The student ratio—estimated value/standard deviation—is given in italics.

Table 4
Parameters for the Fourier series development (Eq. (3)) to compute monthly PE in the Massif Central highlands

Fourier series parameter	a	$b(X)$	$c(Y)$	$d(Z)$
A0	335.8	0.153	−0.137	−0.0490
A1	−100.4	−0.036	0.030	0.0190
B1	−45.2	−0.018	0.019	0.0023
A2	4.7	0.003	−0.002	0.0006
B2	25.3	0.019	−0.013	−0.0017

With x_j—successively equal to a_j, b_j, c_j, and d_j—computed as

$$x_j \approx \frac{12}{365.25}\left[\frac{A_0}{2} + \sum_{n=1}^{2}\left(A_n \cos\frac{(j+15)2\pi n}{365.25} + B_n \sin\frac{(j+15)2\pi n}{365.25}\right)\right] \quad (8)$$

In Fig. 2, examples of fits for six sites spread throughout the Massif Central area are presented.

4.4. Conclusion from the results of PE regionalization in the Massif Central highlands

With a satisfactory regionalization formula for Penman PE, the watershed models can be provided with a PE estimate that takes into account the position and mean altitude of each watershed in the area. This a priori truer input should improve the ability to represent the rainfall–runoff relationship on the study watersheds: this is what will be checked in Section 6, after a short presentation of the watershed models (Section 5).

5. Models and methods

5.1. Models

Two simple, continuous lumped watershed models were used at a daily time step

- The conceptual four-parameter GR4J model (belonging to a family of models developed at Cemagref, beginning in the early 1980 s and widely used in France). Perrin et al. (2003) present the model structure in detail.
- An eight-parameter modified version of TOPMODEL (here, called TOPMO), which uses a parameterized analytical expression of the soil-topographic index distribution (Beven and Kirkby, 1979; Beven et al., 1995). Details of the model structure can be found in Michel et al. (2003).

Here, both model structures were used in a lumped mode and fed with the same data, i.e. rainfall–runoff time-series and PE estimates. A detailed discussion of model structures is not within the scope of this paper; only their structures are shown in Fig. 3.

The models selected are fairly different: they use their own parametrization of the rainfall–runoff transformation, and have distinct flexibility levels. We expect that by using these two different models, more general results will be obtained.

In calibration, the commonly used Nash and Sutcliffe (1970) goodness-of-fit criterion was used as the objective function. A description of the automatic, maximum-gradient optimization procedure used here is given in Edijatno et al. (1999).

5.2. Method of comparison and expected results

To study the impact of a better regional knowledge of PE on watershed model efficiency, the split-sample test procedure recommended by Klemeš (1986) was applied such that for each watershed, the model was calibrated on a 3-year period, and the results obtained in simulation on a different 3-year period were used as a measurement of model efficiency. As each of the 3-year periods can also be a calibration or a validation period, a total of $2 \times 62 = 124$ efficiency measurements were obtained for each model, in the validation mode. From these 124 efficiency values, an experimental cumulative probability function of efficiencies can be defined that characterizes each model.

What results can be expected? In principle, the model efficiency should improve because better knowledge of the meteorological demand by the watershed (through better spatial interpolation) is obtained. If no model efficiency improvement is achieved, then the problems probably lie within the watershed models.

6. Impact of improved potential evapotranspiration (PE) estimates on watershed model efficiency

In this section, the results are discussed in terms of model efficiency. The GR4J and TOPMO models, with five different PE inputs, were tested successively. The five scenarios—from (a) to (e)—are presented in Table 5. Note that for variants (a)–(c), all watersheds were fed with the same PE input, irrespective of their

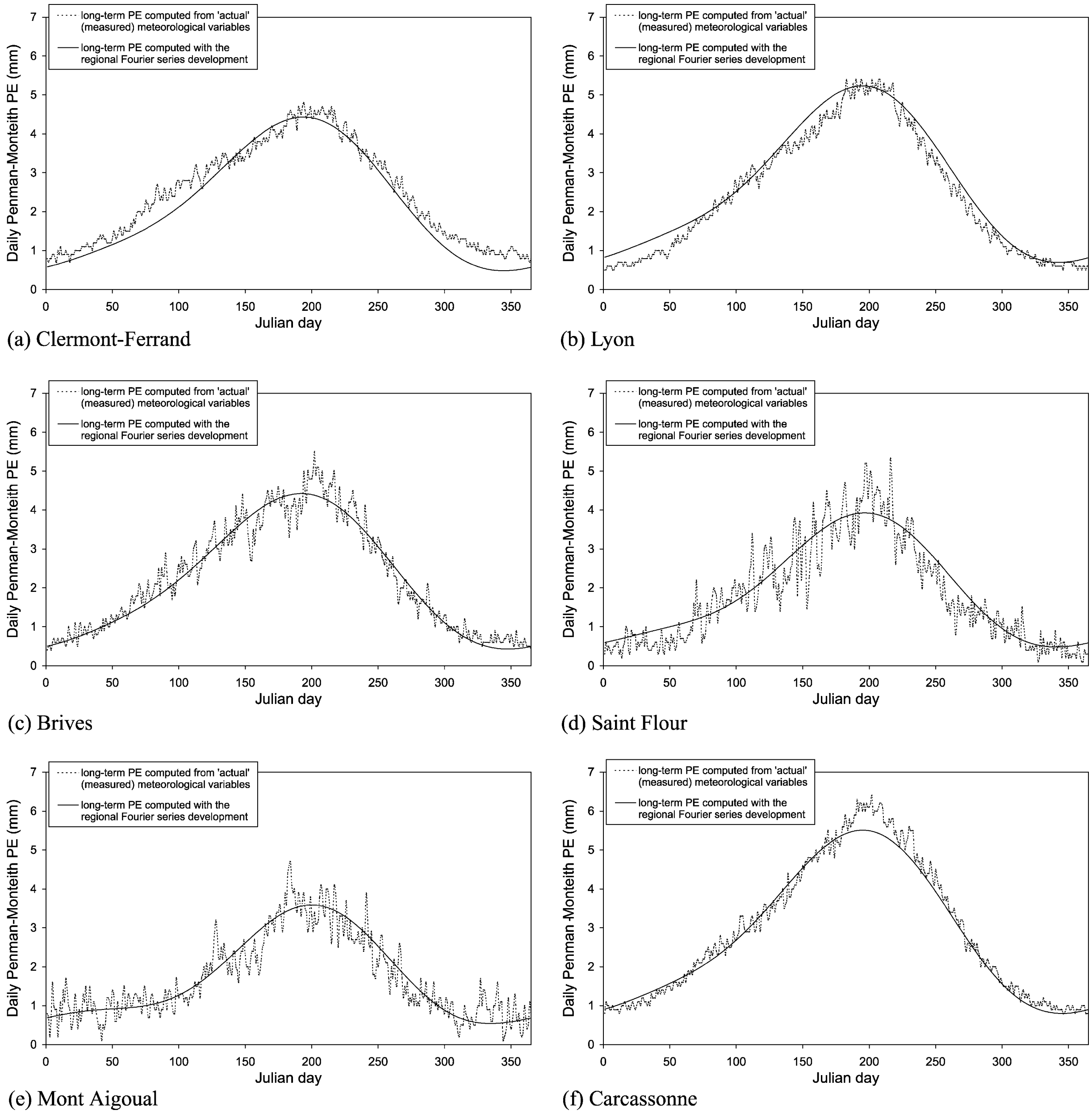

Fig. 2. Examples of the monthly variation in PE in the Massif Central highlands for six locations). The letter also refers to the location on the map.

location. For variants (d) and (e), PE input was watershed-specific.

Results are presented in Fig. 4 for GR4J and TOPMO. Both models show very similar behavior, i.e. overall, they seem remarkably insensitive to the differences in PE data, as shown by mingled (braided) efficiency distributions. It is concluded that

- The only PE scenario that clearly decreased model efficiency ratings is what we have termed the High scenario (a). Both GR4J and TOPMO suffer from

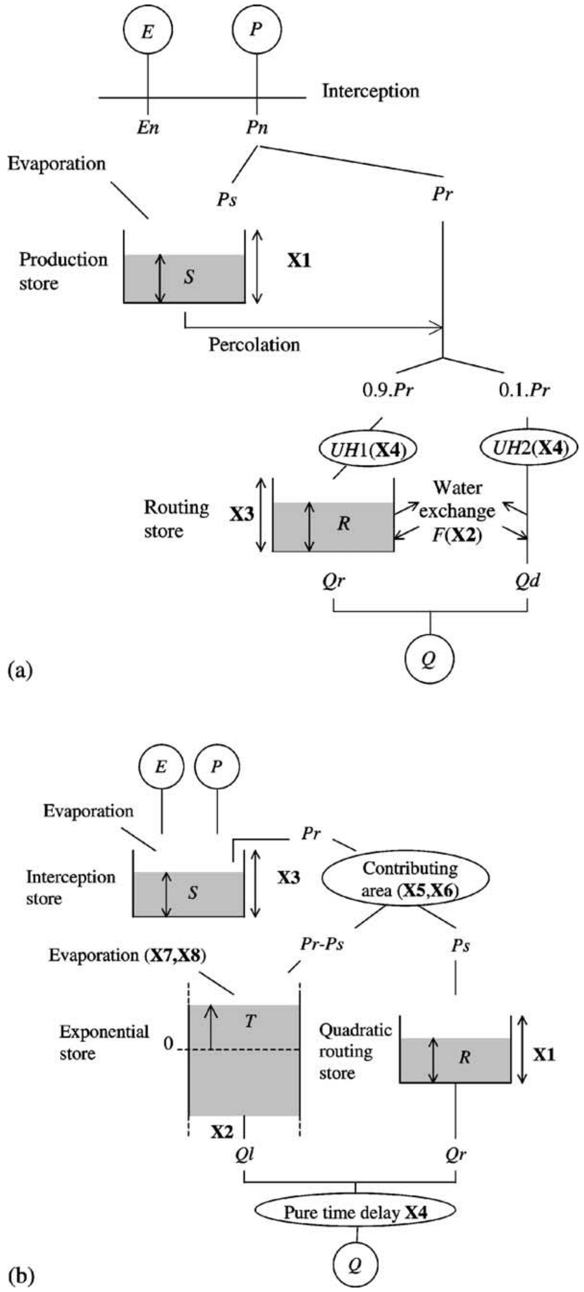

Fig. 3. Diagrams of GR4J (a), and TOPMO (b) models.

this likely overestimation in PE. Here, the model sensitivity to the PE overestimate does not depend on the level of model complexity: TOPMO, with eight calibrated parameters (two of which are specific to the function which transforms PE to

Table 5
Definition of the five scenarios tested to assess model sensitivity to PE input

PE scenario name	Input used	Explanation
(a) High	The PE of the Nîmes meteorological station is used for all watersheds	As Nîmes has the highest PE in the area, this scenario overestimates the PE input to all watersheds
(b) Low	The PE of the Mont Aigoual meteorological station is used for all watersheds	As Mont Aigoual has the lowest PE in the area, this scenario underestimates the PE input to all watersheds
(c) Average	The PE obtained from the regional formula (Eq. (8)), for a point located at the center of the Massif Central and at the mean elevation of the area, is used for all watersheds	This average scenario probably underestimates the PE input for some of the watersheds and overestimates it for the others
(d) Classical	The PE data for the closest synoptic meteorological station is used at each watershed site	This approach is classical in the sense that it is the common practice in France in the absence of regionalized information
(e) Regionalized	The PE obtained from the regional formula (Eq. (8)), for a point located at the center of the watershed and at the mean elevation of the watershed is used	This scenario is considered the best estimate of PE input in all of the watersheds studied

Note: all five scenarios use long-term averages; the same Julian day has the same PE amount every year.

actual evapotranspiration, AE) reacted as strongly as GR4J, with only four calibrated parameters.

- Also common to both TOPMO and GR4J is that no difference was found between the model efficiency distributions of the Average (c) and Regionalized (e) scenarios.
- There is a slight difference in the reaction to scenarios (b) and (d). GR4J does not differentiate between (d) and the two best scenarios (c and e), but TOPMO shows intermediate results with scenario (d). Concerning the Low estimate (b),

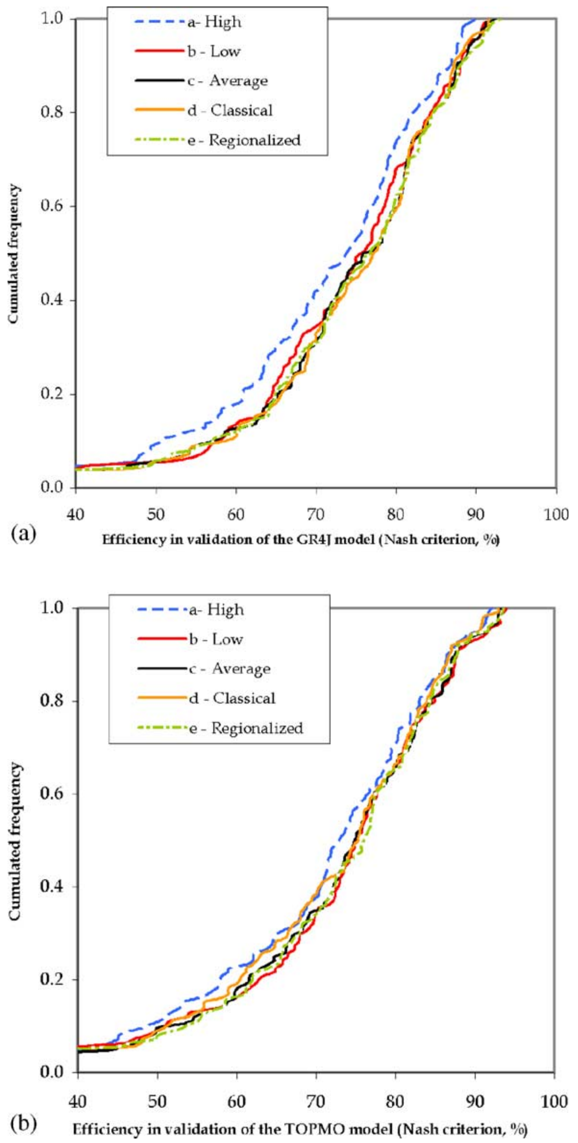

Fig. 4. Distribution of (a) GR4J and (b) TOPMO results over 62 watersheds for five different PE inputs (see Table 6 for a definition of the five PE scenarios).

GR4J appears slightly perturbed by systematic PE underestimates, whereas TOPMO sees no difference between this scenario and the two best ones ((c) and (e)).

At this point, a question arises: are watershed models simply insensitive to their PE inputs, or are they able to adapt to different PE scenarios? The answer is sought in Section 7.

7. How do watershed models adapt to imperfect PE input?

Surprisingly, in Section 6 an improvement in model efficiency distribution only was observed when comparing results of very crude PE estimates (the High (a) and sometimes the Low (b) scenario) with results of the regionalized approach. There was no noticeable difference in efficiency between a rainfall–runoff model fed with an Average PE (scenario c) and a model fed with a Regionalized PE.

In this section, the ability of the automatic calibration procedure to adapt to different PE scenarios (while maintaining the model's ability to represent the rainfall–runoff relationship) is studied. Here, only GR4J is considered because its parsimony makes it easier to analyze the role of each parameter separately (Table 6 lists the parameters of GR4J and their signification). First the distributions of parameters obtained with the High (a), Low (b) and Average (c) PE inputs are compared (Fig. 5), and then the distributions of parameters obtained with the Average (c), Classical (d) and Regionalized (e) PE inputs (Fig. 6) are interpreted.

7.1. Comparison of parameter distributions: high (a), low (b) and average (c) PE scenarios

7.1.1. Production parameters

We here comment on the first part of Fig. 5

- Capacity of the production reservoir ($X1$). The capacity of the GR4J production reservoir

Table 6
List of parameters of the GR4J model

Parameter		Parameter signification
Production module	$X1$	Capacity of the production reservoir (mm)
	$X2$	Water exchange coefficient (mm)
Transfer module	$X3$	Capacity of the nonlinear routing reservoir (mm)
	$X4$	Unit hydrograph time base (day)

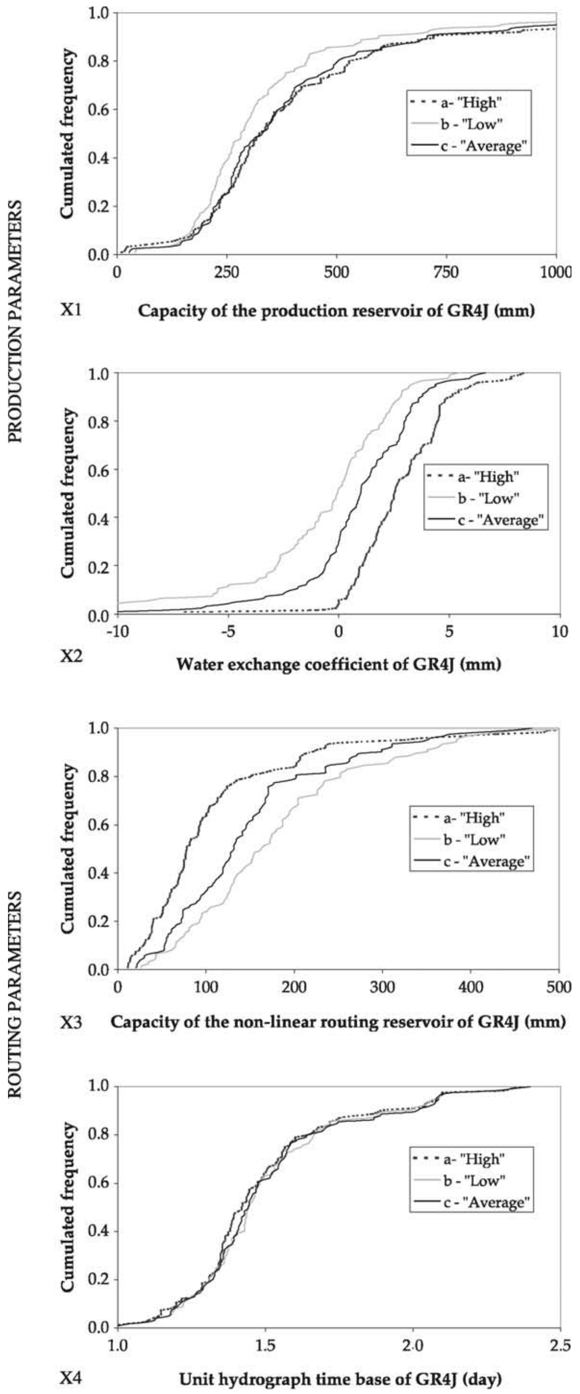

Fig. 5. Distribution of the parameters of GR4J over 62 watersheds for three extreme PE input scenarios: (a) High, (b) Low, (c) Average (all watersheds are fed with the same input). See Table 6 for a definition of the scenarios.

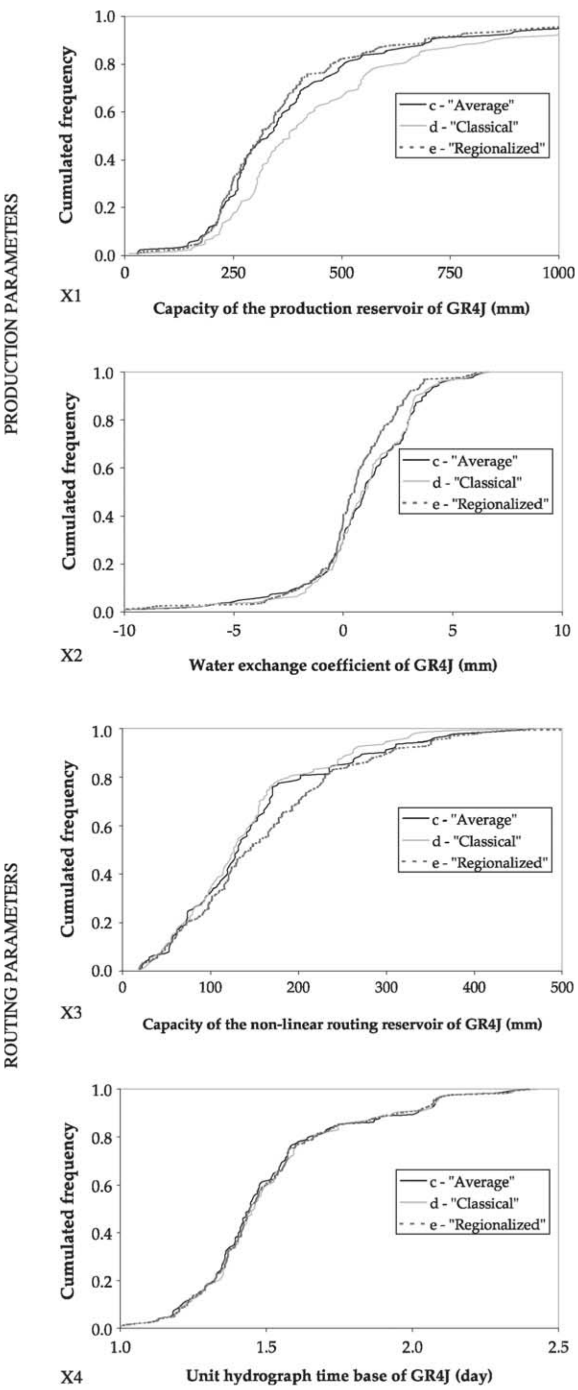

Fig. 6. Evolution of the distribution of GR4J parameters following an improvement in PE input: (d) classical: knowledge limited to data of the synoptic network, (e) regionalized: improved regionalized knowledge. The Average scenario (c) stands as reference. See Table 6 for a definition of the scenarios.

does not seem to be affected by PE over-estimates. However, the model reacts to under-estimates by reducing this capacity. The absence of symmetric behavior might indicate that the Penman PE itself is already overestimating watershed-scale PE: when PE is further over-estimated, no adaptation is possible.

- Water exchange coefficient ($X2$). The most obvious reaction by GR4J to PE over and underestimates is observed with the water exchange coefficient, which represents under-ground leakage from/to the watershed. $X2$ adapts logically: an overestimate produces quite exclusively positive values (i.e. underground inflow to the watershed), whereas an underestimate results in mostly negative values (i.e. underground outflow from the watershed). This interaction is not a surprise. It shows how it is difficult to extricate the various processes that are behind parameters, even in the simplest models. $X2$, and the function attached to it, had been initially incorporated to model intermittent basins.

7.1.2. Routing parameters

We here comment on the second part of Fig. 5

- Capacity of the non-linear routing reservoir ($X3$). The GR4J model adapts to PE overestimates by reducing the capacity of its nonlinear routing reservoir and to PE underestimates by increasing this capacity. It was not expected that PE under and overestimates would affect a parameter pertaining to the routing function, but this demonstrates the interdependence between all the constituent parts of a model. An explanation can only be found a posteriori. It may be assumed that PE underestimation induces an increased capacity to generate effective rainfall, because the soil moisture reservoir is less easily depleted. To avoid producing overly high flows, the automatic calibration increases the capacity of the routing reservoir: more effective rainfall can be stored and then released more regularly. Another possible explanation lies in the mathematical formulation of water exchanges: the observed behavior of the routing reservoir could be a consequence of its interaction with the exchanges.
- Unit hydrograph time base ($X4$). This routing parameter is unaffected by the choice of PE input scenario. This is easy to understand, since this parameter only helps to fit the lag between peaks of rainfall and runoff.

As a preliminary conclusion, it can be said that the GR4J watershed model, used here as an example, can rely on its two production parameters to adapt to imperfect PE input. The most sensitive parameter is a production parameter ($X2$, the water exchange coefficient). One of the routing parameters ($X4$, the unit hydrograph time base) remains unaffected by the differences in estimated PE input, while the second routing parameter ($X3$, the routing reservoir capacity) is sensitive. This behavior is rather surprising, and is discussed further in Section 7.2.

7.2. Comparison of parameter distributions: average (c), classical (d) and regionalized (e) PE scenarios

It was noted in Section 6 that PE estimation scenarios (c)–(e) produced very similar efficiency distributions for GR4J. In contrast, there is often a clear distinction between the distributions of the parameters calibrated within these scenarios. This means that the model definitely is sensitive to PE input, but that it is flexible enough to accommodate imperfect input and compensate for it, so that nothing convincing can be detected in terms of sheer efficiency as evidenced in Fig. 4.

Fig. 6 shows three different behaviors of GR4J parameters

1. In the case of the unit hydrograph time base ($X4$), all three distributions are similar. This is logical: as already noted, this parameter is not sensitive to PE input;
2. In the case of the first production parameter, the capacity of the production reservoir ($X1$), the parameter distribution is quite similar in the Regionalized (e) and the Average (c) scenarios, but differs from that in the Classical (d) scenario. This probably means that the distribution of this parameter only depends on the total PE amount. In fact, the Average scenario provides a PE amount very close to the mean PE input in our 62 watersheds, while the Classical scenario, which

uses low-elevation synoptic weather stations, overestimates the mean PE input to the watershed sample.
3. The second production parameter (the water exchange coefficient, $X2$) and one of the routing parameters (the capacity of the nonlinear routing reservoir, $X3$) behave similarly: their distributions are nearly identical in the Average (c) and the Classical (d) scenarios, but differ from the distribution in the Regionalized (e) scenario. This similarity in behavior argues for an interaction between these two parameters, and it can be considered that routing parameter $X3$ is contaminated by the sensitivity of $X2$ to PE input. The cause of this so-called contamination may be the fact that, in GR4J, water exchanges are computed (Eq. (9)) as a function of $X2$ and of the water level in the routing store (Perrin et al., 2003).

$$\text{exchange} = X2\left(\frac{\text{water level in the routing store}}{X3}\right)^{3.5} \tag{9}$$

8. Conclusions

8.1. Synthesis

The objective of this paper was to study the impact of using a priori improved estimates of areal PE as input in watershed models. In the Massif Central highlands of France, a successful regionalization of Penman PE was obtained, and the impact of these better estimates of PE on the efficiency of two watershed models, GR4J and TOPMO, was studied. No difference between improved and basic PE estimation strategies was found in terms of model efficiency in a sample of 62 watersheds. Furthermore, very simplistic assumptions on watershed PE input (the same average input for all watersheds) yielded similar efficiency distributions. The watershed models tested here were only sensitive to the crudest PE overestimation scenario.

However, the parameter distributions for the GR4J model in different PE input scenarios indicated that the model reacted to the change in PE input through its parameters, which were consequently adjusted to different PE inputs.

The results reported here may seem evident to experienced watershed modelers, who are aware of the ability of the calibration process to compensate for biased input data, and of models to yield accurate results, as long as the application input is similarly biased. However, recent results such as those of Vázquez and Feyen (2003) seem to be contesting this ability of watershed models, and we thus believe that a debate needs to be opened between modelers on this topic.

8.2. Questioning the PE concept at the watershed scale

The above results are both reassuring and disconcerting

- Reassuring from a practical point of view, because if a perfect PE input was needed to run a watershed model successfully, such models would be rarely used in an engineering context. As such, the adaptability of watershed models is a good thing.
- Disconcerting from a modeling point of view, as it seems somewhat illogical that a priori improved knowledge of the evaporative demand does not translate into improvement in the efficiency of watershed models.

If average, simplistic assumptions on the evaporative demand are sufficient to represent the rainfall–runoff relationship, two questions arise

- First, are conceptual watershed models[3] efficient in their use of such an important boundary condition as PE?
- Second, is Penman PE relevant as climatic forcing to watershed models? This question was raised by Bouchet (1963), who proposed a complementary relationship between PE and AE at the scale of several square kilometers. This idea was further pursued by Brutsaert and Stricker (1979) and Morton (1983, 1994). The latter author even considered it ‘likely that the use of the Penman equation to estimate

[3] Only two models were tested in this paper, but the hydrological literature leads us to believe that the results are valid for most, if not all, watershed models used in hydrology.

evaporation from hydrologically significant areas has no real future, being merely an attempt to force reality to conform to preconceived concepts derived from small wet areas'. This question surely requires further investigation, on which we will report in due course.

Acknowledgements

The authors wish to thank Bruno Rambaldelli, of Météo France, for kindly providing PE data from the automatic weather station network in the Massif Central, as well as Christian Scherer, of the Hydro service at the French Department of the Environment, for interesting discussions on the theme of PE. Rainfall data for the test watersheds were obtained from the Météo-France Pluvio data bank, and runoff data from the Hydro data bank. We also extend our thanks to Jean-Louis Rosique for maps, and to Mrs de Marsily and to Mrs Northrup for their review of the English. Last, we wish to acknowledge the comments of an anonymous reviewer, which greatly contributed to improving the manuscript.

References

Andersson, L., 1992. Improvement of runoff models. What way to go? Nordic hydrology 23, 315–332.

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. Journal of Hydrology 250 (1–4), 206–223.

Bergström, S., 1995. The HBV model. In: Singh, V.P., (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, Highlands Ranch, CO, pp. 443–476.

Beven, K.J., 2001. Rainfall–runoff Modelling: the Primer. Wiley, Chichester, 372 pp.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin 24 (1), 43–69.

Beven, K.J., Lamb, R., Quinn, P., Romanovicz, R., Freer, J., 1995. TOPMODEL. In: Singh, V.P., (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, Highlands Ranch, CO, pp. 627–668.

Bouchet, R.J., 1963. Evapotranspiration réelle et potentielle—Signification climatique. IAHS Publication 62, 134–142.

Bras, R.L., 1979. Sampling of interrelated random fields: the rainfall–runoff case. Water Resources Research 15 (6), 1767–1780.

Bras, R.L., Rodriguez-Iturbe, I., 1976. Rainfall network design for runoff prediction. Water Resources Research 12 (6), 1197–1208.

Brutsaert, W., 1982. Evaporation into the Atmosphere. Kluwer, Dordrecht, 299 pp.

Brutsaert, W., Stricker, H., 1979. An advection-aridity approach to estimate actual regional evapotranspiration. Water Resources Research 15, 443–450.

Burnash, R.J.C., 1995. The NWS river forecast system-catchment modeling. In: Singh, V.P., (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, Highlands Ranch, CO, pp. 311–366.

Edijatno, 1991. Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier. PhD Thesis. Université Louis Pasteur, Strasbourg, 242 pp.

Edijatno, Nascimento, N., Yang, X., Makhlouf, Z., Michel, C., 1999. GR3J: a daily watershed model with three free parameters. Hydrological Sciences Journal 44 (2), 263–278.

Fennessey, N.M., Vogel, R.M., 1996. Regional models of potential evaporation and reference evapotranspiration for the northeast USA. Journal of Hydrology 184, 337–354.

Fowler, A., 2002. Assessment of the validity of using mean potential evaporation in computations of the long-term soil water balance. Journal of Hydrology 256, 248–263.

Joukainen, S., 2000. Improving the calculation of potential evapotranspiration of the HBV model: application to the Ounasjoki watershed. In: Nilsson, T., (Ed.), Nordic Hydrological Conference 2000. Nordic Hydrological Programme Report, Swedish Hydrological Council, Uppsala, pp. 347–354.

Klemeš, V., 1986. Operational testing of hydrologic simulation models. Hydrological Sciences Journal 31 (1), 13–24.

Kokkonen, T.S., Jakeman, A.J., 2001. A comparison of metric and conceptual approaches in rainfall–runoff modeling and its implications. Water Resources Research 37 (9), 2345–2352.

Kribèche, R., 1994. Etude de la sensibilité d'un modèle pluie-débit à l'exactitude de l'évaporation (modèle GR4J). DEA Thesis. Université Paris XII, Créteil, 42 pp.

Michel, C., Perrin, C., Andréassian, V., 2003. The exponential store: a correct formulation for rainfall–runoff modelling. Hydrological Sciences Journal 48 (1), 109–124.

Morton, F.I., 1983. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. Journal of Hydrology 66, 1–76.

Morton, F.I., 1994. Evaporation research—a critical review and its lessons for the environmental sciences. Critical reviews in environmental science and technology 24 (3), 237–280.

Nandakumar, N., Mein, R.G., 1997. Uncertainty in rainfall–runoff model simulations and the implications for predicting the hydrologic effects of land-use change. Journal of Hydrology 192, 211–232.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I—a discussion of principles. Journal of Hydrology 10, 282–290.

Parmele, L.H., 1972. Errors in output of hydrologic models due to errors in input potential evapotranspiration. Water Resources Research 3 (2), 348–359.

Paturel, J.E., Servat, E., Vassiliadis, A., 1995. Sensitivity of conceptual rainfall–runoff algorithms to errors in input datacase of the GR2M model. Journal of Hydrology 168, 111–125.

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology 279, 275–289.

Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity Analysis. Wiley, New York, 504 pp.

Shuttleworth, W.J., 1993. Evaporation. In: Maidment, D.R., (Ed.), Handbook of Hydrology, McGraw-Hill, New-York.

Storm, B., Hogh Jensen, K., Refsgaard, J.C., 1989. Estimation of catchment rainfall uncertainty and its influence on runoff prediction. Nordic Hydrology 19, 77–88.

Vázquez, R.F., Feyen, J., 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. Journal of Hydrology 270, 309–327.