Regionalization of rainfall for broad-scale modeling: An inverse approach

J.-L. Pinault[1] and D. Allier[1]

Received 23 October 2006; revised 2 May 2007; accepted 7 June 2007; published 29 September 2007.

[1] A spatially consistent approach is used for the representation of rainfall at catchment scale for continuous rainfall-streamflow simulation by using inverse modeling. Representing rainfall data at every location as the product of the mean rainfall by the rainfall series reduced to unit average, it is shown that the regionalization of both terms should follow different ways for broad-scale modeling. Whereas the regionalization of the mean rainfall is based on its spatial continuity, it is demonstrated from the study of three French basins that are subject to different climates that the reduced rainfall data should be represented from a weighted sum of a small number of observed rainfall data (three to five) located both inside and outside the catchments to be as representative as possible at catchment scale. The reliability of peak flow modeling increases with the basin size as well as the return period of flood events provided that the rainfall is correctly regionalized, which is particularly important for real-time forecasting of rainfall and flow. This contradicts the widespread assumption that for the distributed rainfall-runoff model, the denser the network of rain samplers used as input in the models, the more accurate the broad-scale flood model will be.

Citation: Pinault, J.-L., and D. Allier (2007), Regionalization of rainfall for broad-scale modeling: An inverse approach, *Water Resour. Res.*, *43*, W09422, doi:10.1029/2006WR005642.

1. Introduction

[2] Recent floods in Europe have raised public, political and scientific awareness and flooding is now widely recognized as being of major strategic importance, with significant economic and social implications. Complex and difficult questions have been identified, concerning, for example, the actual and potential future effects of climate variability and climate change on flood risk and the effects of change in land use or land management. These have far-reaching scientific, technical and socioeconomic implications. Rainfall-runoff models provide flood discharge hydrographs. The impact of flooding is related to water level. Conventionally, hydrological models are used to simulate the former and provide input to hydraulic models of channel and floodplain flow.

[3] The regionalization of rainfall data is a current issue in flood modeling. There has been rapid progress in stochastic modeling of single-site rainfall [*Cox and Isham*, 1980, 1988; *Rodriguez-Iturbe et al.*, 1987, 1988], in which storms arrive according to a Poisson process in time. This type of rainfall has been regionalized and applied in different contexts by *Onof and Wheater* [1993, 1994, 1995], *Onof et al.* [1996], *Wheater et al.* [1999, 2005], *Cowpertwait* [1994], *Velghe et al.* [1994], *Khaliq and Cunnane* [1996], *Verhoest et al.* [1997], *Gyasi-Agyei and Willgoose* [1997], *Foufoula-Georgiou* [1998], *Cameron et al.* [2000, 2001], *Fowler et al.* [2000], and *Cowpertwait et al.* [2002]. The identification of homogeneity at regional scale is a basic step in the inference of the estimation of flood probabilities. This operation is traditionally carried out using statistical methods with large uncertainties. This deficiency has led to the introduction of the concept of scale invariance of annual maximum flood to identify homogeneous regions in flood frequency regionalisation [*Gupta and Waymire*, 1990; *Gupta and Dawdy*, 1994; *Gupta et al.*, 1994; *De Michele and Rosso*, 1995; *Robinson and Sivapalan*, 1997a]. The analysis of the mechanisms of flood production can provide useful information on the clustering of river basins. *Robinson and Sivapalan* [1997b] investigated the influence of the different timescales on the hydrological regimes and the implications on the flood frequency analysis. *Burn* [1997] introduced seasonality measures as catchment similarity indices for the regional flood frequency analysis. *De Michele and Rosso* [2002] used a multilevel approach to flood frequency regionalisation that combines physical and statistical criteria to cluster homogeneous groups in a geographical area.

[4] The importance of spatial rainfall for flood management will vary as a function of the spatial scale of the catchment (which will determine the spatial and temporal scale of rainfall input), catchment properties and rainfall type [*Wheater et al.*, 2006]. Two different approaches are used therefore to model floods depending on the catchment size and the concentration time of rainwater to the outlet:

[5] 1. Flash floods are characterized by a high specific discharge, typically higher than 5 m^3 s km^2. In exceptional cases, specific discharge can exceed 20 m^3 s km^2. Flash flood modeling generally applies to small basins (<500 km^2). It requires a high spatial and temporal resolution of rainfall, which needs a dense network of rain

[1]Water Research Division, Bureau de Recherches Géologiques et Minières, Orléans, France.

Copyright 2007 by the American Geophysical Union.
0043-1397/07/2006WR005642

samplers. Meteorological radars are extensively used to anticipate rainfall events while representing the spatial and the temporal structures of precipitations [*Delrieu et al.*, 2005].

[6] 2. On the other hand, modeling medium and large catchments (500 to 25,000 km^2) needs a comprehensive modeling strategy for a wide range of water management issues. The determination of the level of detail appropriate to broad-scale modeling is a crucial issue. At best, a comprehensive framework should enable to use nested models of different levels of complexity and resolution for a variety of purposes within the context of flood management and flood design. This should, for example, allow evaluation of joint probabilities of fluvial flooding from a combination of contributory effects. Broad-scale modeling for fluvial flooding must focus on the simulation of river basin response (i.e., flood discharge and associated water levels) to precipitation input [*Hosking and Wallis*, 1993, 1997; *Hay et al.*, 2002].

[7] The purpose of this paper is the regionalization of rainfall for broad-scale modeling by using an inverse approach that has been developed previously for hydrograph separation. Representing rainfall data at every location as the product of the mean rainfall by the rainfall series reduced to unit average, it is shown that the regionalization of both terms of rainfall data should follow different ways for broad-scale modeling. Whereas the regionalization of the mean rainfall is based on its spatial continuity (interpolation is commonly used to represent the isohyets), it will be demonstrated from the study of three French basins that are subject to different climates that the reduced rainfall data should be represented from a weighted sum of a few numbers of observed rainfall data to be as representative as possible at catchment scale: Only a few rain samplers (3 to 5) located both inside and outside the catchments are required to optimize rainfall representativeness. The low-noise weighted sum of observed rainfall data exhibits a long correlation length.

[8] The three basins used to develop a methodological approach for the regionalization of the reduced rainfall data are: the Allier and the Maine, which are subcatchments of the Loire basin, and the Somme. The Somme River is supplied mainly by groundwater and therefore has a very long concentration time, up to 100 days, whereas some subcatchments of the Loire basin are very rapid, with rainwater concentration times of around one day. The Somme and Maine floods are generated by rainfall of oceanic origin, whereas the Allier basin floods are subject to convective rainfall on the Cevennes (Southern Massif Central) (Figure 1).

[9] Whereas rain gauges are used as inputs of process-based models the purpose of which is the best representation of streamflow, the gauged catchments are used as pluviometers in inverse models. Indeed, inverse modeling attempts to represent the measurements recorded at pluviometers used as the output of the models, i.e., the streamflows observed at the outlet of catchments, by using information given by rain samplers used as inputs.

[10] Contrary to the process-based models that attempt to reproduce hydrological and hydrogeological processes as accurately as possible, the priority in inverse modeling is to reduce the number of parameters to be calibrated. So, the method does not attempt to reduce the systematic errors of the model as much as possible, but the discrepancies between the estimated and the observed outputs are minimized by using a few numbers of degree of freedom as well as a regularization technique to make the systems invertible. More precisely, the systematic errors of the output and the random errors resulting from the propagation of sampling errors of input data are separated. In other words, the relevancy of the sampled rainfalls used as inputs of the models may be checked objectively without significantly altering the parameters to be calibrated. So, the robustness of the inverse models enables to test coherence between the sampled rainfall and discharge at catchment scale. In this way, the contribution of every rain sampler to the streamflow can be estimated and the most representative weighted sum of rainfall data can be deduced for rainfall regionalization.

2. Methodology

2.1. General Approach

[11] Process-based models and inverse models have opposite requirements:

[12] 1. Process-based models attempt to reproduce streamflow from background data as vegetation, topography, soil, geology and rainfall recorded at rain samplers. Systematic errors of the computed streamflow are minimized by simulating subtle processes. Most of the random errors of the modeled streamflow are originating from rainfall sampling.

[13] 2. Sampling errors of rainfall data used in inverse models are removed since the pluviometer is the gauged catchment and rainfall is not sampled. Nevertheless catchments used as pluviometers are imperfect rain gauges, because of three reasons: (1) The response of the catchment is delayed, all the more as groundwater contributes to streamflow, (2) a part of rainfall is evaporated and does not contribute to the streamflow, and (3) transfer processes may be altered when extensive rainfalls occur, because of water saturation of soils that enhances runoff or because of piston flow effect that may occur both in the unsaturated and in the saturated zones.

[14] Hence inverse modeling of rainfall consists in removing most of the biases introduced by the rainfall measurement at the catchment scale and in overcoming the three mentioned drawbacks, the purpose being to select the most representative rain samplers.

[15] It is convenient to represent the recorded rainfall at a particular location by the product of two terms: (1) the mean rainfall height that is influenced by the local environment of the rain gauge, represented by the isohyets and (2) the rainfall series reduced to unit average that is generally subject to large-scale correlations.

[16] Whereas isohyets are generally accurately known since they are deduced from mean values of rainfall data recorded at rain gauges, the regionalization of reduced rainfall is much more difficult to perform. Indeed, reduced rainfall at a given location cannot be related to the nearest rain samplers and interpolation techniques are inappropriate. Nevertheless, short-term rainfall events determine the reliability of transfer models, whether intended for flood or low water conditions. The rainfall sampling rate depends on the concentration time of rainwater to catchment outlets,

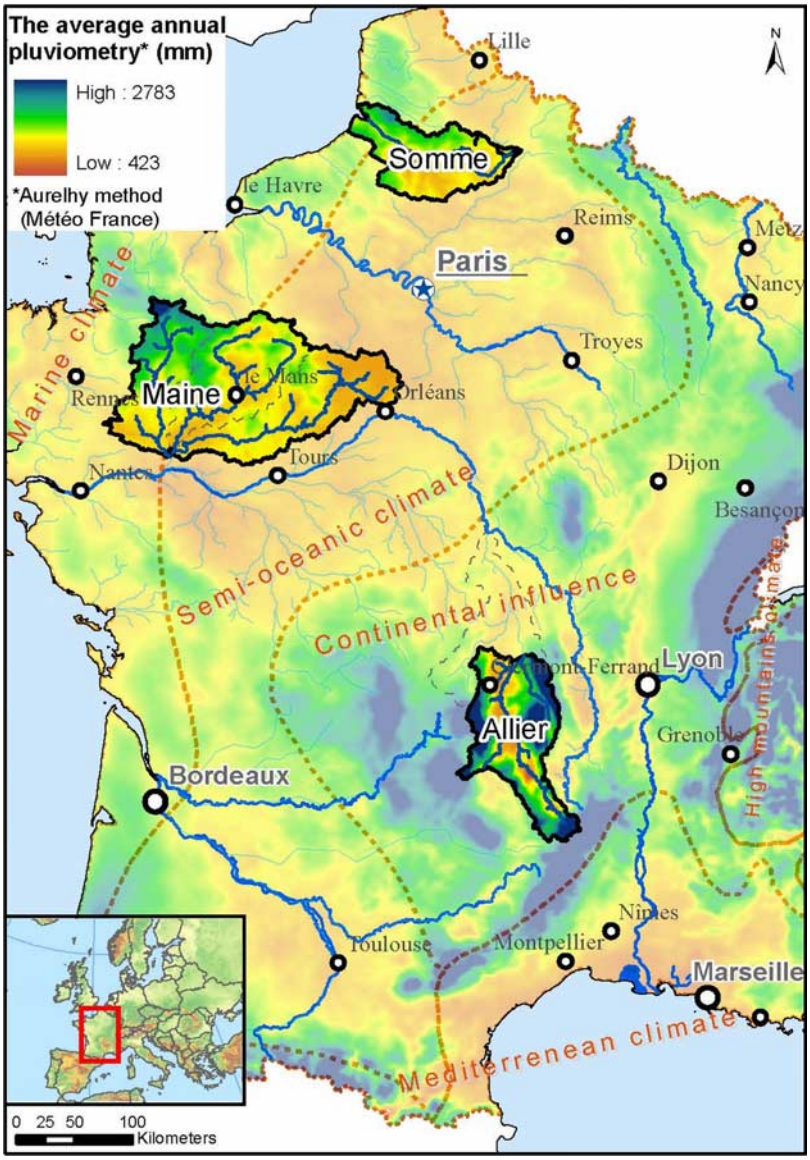


Figure 1. Localization of the three study areas in France: The Somme, Maine, and Allier basins.

which is rarely less than one day in the case of medium size and large catchments.

[17] The inverse approach proposed here consists in optimizing a weighted sum $\mathbf{R}_\Sigma$ of reduced rainfall data observed at rain samplers to explain, as well as possible, the rainfall observed at the catchment scale. Indeed, considering rainfall data reduced to unit average, the weighted sum of rainfall series observed at different locations may be considered as representative as possible of the mean rainfall at the catchment scale by reducing the random errors of the output, i.e., by minimizing the sampling errors of rainfall data. Consequently the data of the inverse models are dimensionless, whether it is question of sampled rainfall or streamflow: The time series are divided by their average (their expected value).

2.2. Inverse Modeling

[18] The inverse approach used into the procedure for the regionalizing of rainfall data (Figure 2) is outlined from (1) to (6). The concepts as well as the inversion and the regularization techniques were developed by *Pinault et al.* [2001a]. Some methodological developments of the inverse models were presented in previous studies to characterize denitrification processes in groundwater [*Pinault et al.*, 2001a], to determine water renewal in karsts [*Pinault et al.*, 2001b], to separate deep and shallow waters of a submarine karst aquifer system [*Pinault et al.*, 2004], to elucidate switching behavior of groundwater recharge from matrix flow to macropore flow in groundwater-induced flooding [*Pinault et al.*, 2005], to separate pressure head variations of the French Alsace water table into three components, recharge from rainfall and exchanges with surface water in the Rhine River and in Vosgean tributaries [*Pinault and Schomburgk*, 2006] and to characterize the resilience of aquatic ecosystems with respect to the evolution of environmental parameters [*Pinault and Berthier*, 2007].

[19] A new methodological approach is presented to regionalize rainfall data. The concepts that were developed in the previous studies are used: (1) to calculate the effective

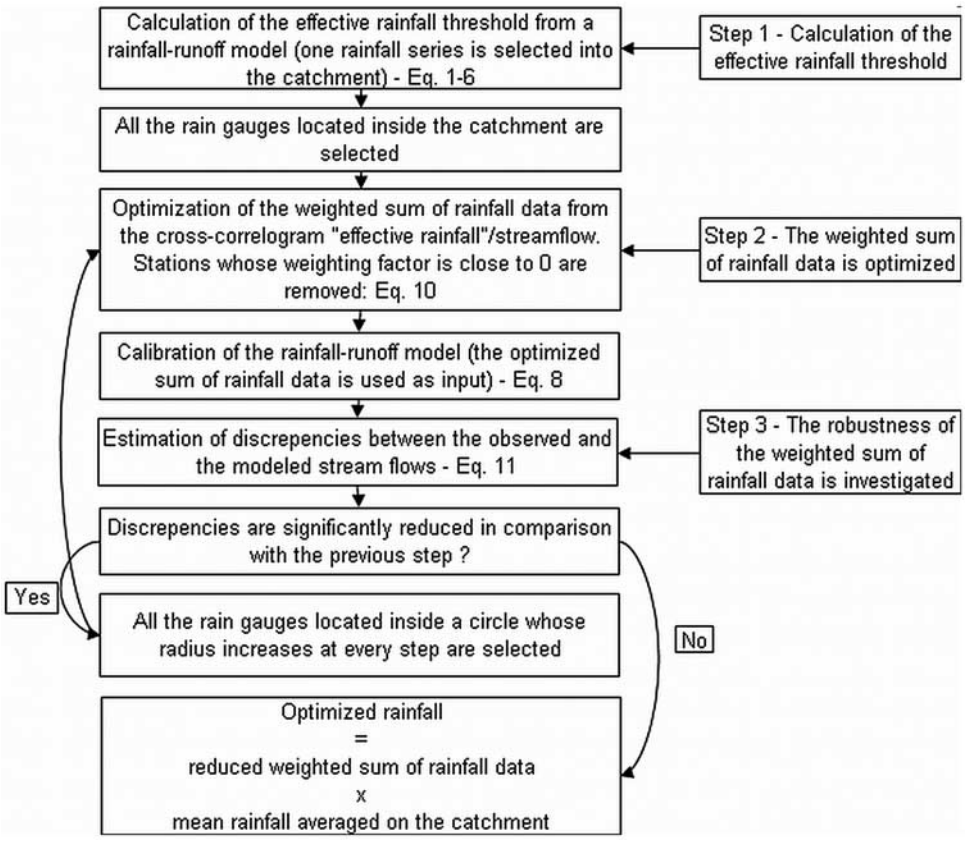


Figure 2. Procedure used to optimize the rainfall data at a catchment scale: Rainfall is represented as the product of the optimized sum of rainfall data reduced to unit average by the rainfall height averaged over the catchment.

rainfall threshold to express the reduced rainfall data into effective rainfall and (2) to calibrate the rainfall-runoff models in order to quantify the impact of rainfall sampling onto the computed streamflows.

[20] The inverse approach may be synthesized in the following way. Considering the streamflow $\mathbf{Q}$ as the output of a transfer model, a general formulation of the problem can be written:

$$Q(t_i)/\overline{Q} = \mathbf{\Gamma}_s * \mathbf{R}_s/\overline{R}_{eff} + \mathbf{\Gamma}_q * \mathbf{R}_q/\overline{R}_{eff} + \varepsilon \qquad (1)$$

where * is the discrete convolution product. Effective rainfall $R_{eff}(t_i)$ is composed of two components $R_{eff}(t_i) = R_s(t_i) + R_q(t_i)$, $R_s(t_i)$ and $R_q(t_i)$, representing those parts of the rainfall that induce the slow and quick components of the streamflow at the outlet, respectively ($\overline{X}$ represents the mean of the series X).

[21] Such a model requires the solutions $\mathbf{\Gamma_s}$ and $\mathbf{\Gamma}_q$, the slow and quick transfer functions, to the transport equation. The random part ε represents erratic, complex, and usually short-term variability of the streamflow that is not explained by the model.

[22] Effective rainfall $\mathbf{R}_{eff}$ is calculated from rainfall $\mathbf{R}$ and the effective rainfall threshold $\mathbf{\Omega}$, such that

$$R_{eff}(t_i) = \begin{cases} R(t_i) - \Omega(t_i) & \text{if } R(t_i) \geq \Omega(t_i) \\ 0 & \text{if } R(t_i) < \Omega(t_i) \end{cases}. \qquad (2)$$

The effective rainfall threshold $\Omega(t_i)$ is related to both rainfall and potential evapotranspiration, such that

$$\mathbf{\Omega} = \mathbf{\Gamma}_{\Omega,PET} * \mathbf{PET} + \mathbf{\Gamma}_{\Omega,R} * \mathbf{R} + \mathrm{C^{st}} \qquad (3)$$

where $\mathbf{\Gamma}_{\Omega,PET}$ and $\mathbf{\Gamma}_{\Omega,R}$ are impulse responses of $\mathbf{\Omega}$ to potential evapotranspiration **PET** and to rainfall $\mathbf{R}$, respectively ($\mathrm{C^{st}}$ is a constant).

[23] Quick transfer occurs only after exceptional rainfall, when a water-saturated soil enhances runoff or when a hydraulic continuity is established between matrix and macropores in the unsaturated zone, which can lead to groundwater discharge into the stream. The component $R_q(t_i)$ can be expressed according to effective rainfall $R_{eff}(t_i)$ so that

$$R_q(t_i) = R_{eff}(t_i) \times \alpha(t_i) \qquad (4)$$

where the function $\alpha(t_i)$ is the fraction of effective rainfall participating in quick transfer; $\alpha(t_i)$ reflects either the nonlinear switching of groundwater recharge from matrix flow to macropore flow or the water saturation index of soils. It is related to the effective rainfall of previous events such that

$$\alpha = \mathbf{\Gamma}_{\alpha,R_{eff}} * \mathbf{R}_{eff} \qquad (5)$$

where $\mathbf{\Gamma}_{\alpha,R_{eff}}$ is the impulse response of $\alpha(t_i)$ to effective rainfall $\mathbf{R}_{eff}$. The function $\alpha(t_i)$ increases as a result of the

stacking effect when several successive intense rainfall events occur. This function may express the filling up of fractures, which occurs when successive rainfall events occur rapidly enough to fill up the fractures while the drainage of macropores toward the surface water network occurs. Concerning runoff, the function $\alpha(t_i)$ is related to the water saturation index of the soils, which increases considerably when successive rainfall events are close enough in comparison with the duration of soil drainage.

[24] By corollary, the component $R_s(t_i)$ for rainfall inducing the slow component of the transfer function may be written as

$$R_s(t_i) = R_{eff}(t_i) \times (1 - \alpha(t_i)) \tag{6}$$

[25] The inverse method that is parsimonious (few parameters) and computationally efficient aims to calculate the normalized impulse responses $\mathbf{\Gamma}_s$ and $\mathbf{\Gamma}_q$, the impulse responses $\mathbf{\Gamma}_{\Omega,PET}$, $\mathbf{\Gamma}_{\Omega,R}$ and the constant C^{st} used to calculate the effective rainfall threshold $\Omega(t_i)$ in (3) and the impulse response $\mathbf{\Gamma}_{\alpha,R_{eff}}$ used to calculate the fraction of effective rainfall $\alpha(t_i)$ involved in quick flow in (5). Actual implementation of (1)–(6) was carried out by means of the Tempo code [*Pinault et al.*, 2001a].

[26] The calibration of the transfer model (1) is illustrated in Figure 3. The effective rainfall threshold that represents the available water deficit in the soil (Figure 3e) exhibits a pseudo periodicity so that the effective rainfall deduced from (2) is 60% of the total rainfall. The proportion of quick transfer of effective rainfall that quantifies the contribution of runoff to streamflow given by (4) depends strongly on the temporal structure of rainfall height (Figure 3f). The comparison of the observed flow and the model shows some discrepancies that are interpreted as systematic errors, i.e., errors resulting from subtle processes that are not taken into account in the inverse model (Figures 3a and 3b). Although the distinction between transfer processes might be arbitrary for small lags, the quick and slow impulse responses $\mathbf{\Gamma}_q$ and $\mathbf{\Gamma}_s$ represent mainly rainwater transfer to the outlet via runoff and groundwater, respectively. The quick transfer reaches its maximum one day after an effective rainfall event and vanishes four days after, whereas the slow transfer decreases slowly for 20 days after reaching its maximum 2 days after the rainfall event (Figure 3g and 3h). The recession lasts 200 days. The separation of the hydrograph into slow and quick components (Figure 3c and 3d) shows that floods are caused mainly by runoff whereas the groundwater discharge supports low water. The mean contribution of quick transfer is 9.2% of the total flow. The impulse response $\mathbf{\Gamma}_{\Omega,PET}$ of the effective rainfall threshold $\mathbf{\Omega}$ to rainfall and to PET (3) indicates water dynamics in the soils (Figure 3i). The duration of the available water content response is very similar for both inputs, 16 and 17 days respectively, whereas it responds immediately after a rainfall event or after a sudden increase in evapotranspiration. The impulse response $\mathbf{\Gamma}_{\alpha,R_{eff}}$ of the effective rainfall contribution to quick flow $\alpha(t_i)$ (5) is represented in Figure 3j. There must be at least 16 days between two rainfall events to enable the contribution to quick transfer $\alpha(t_i)$ to decrease, which corresponds to the duration of soil drainage.

2.3. Procedure to Optimize the Reduced Weighted Sum of Rainfall Data

[27] The method consists in optimizing a weighted sum of reduced rainfall data observed at rain samplers to explain, as well as possible, a streamflow $\mathbf{Q}$ measured at the outlet of a catchment (or a subcatchment). The same analysis can be done using data from piezometers in unconfined aquifers.

[28] The best representation of the weighted sum of rainfall series

$$R_\Sigma = \sum_{k=1,p} \delta_k R_k \qquad \delta_k \geq 0 \, k = 1, \Lambda, p \qquad \sum_{k=1,p} \delta_k = 1 \tag{7}$$

(where p is the number of rain samplers) should consist in calculating the weighting factors δ_k so that the Euclidian distance between the observed streamflow $Q(t_i)$ and its estimator $\hat{Q}(t_i)$ for which the reduced rainfall $\mathbf{R}/\overline{R}$ is replaced by the reduced weighted sum $\mathbf{R}_\Sigma/\overline{R}_\Sigma$ in (1):

$$\hat{Q}(t_i)/\overline{Q} = \mathbf{\Gamma}_s * \mathbf{R}_{\Sigma,s}/\overline{R}_\Sigma + \mathbf{\Gamma}_q * \mathbf{R}_{\Sigma,q}/\overline{R}_\Sigma + \varepsilon \tag{8}$$

is minimum. From a numerical point of view, this is an ill-posed problem, and to give the weighting factors δ_k meaning the calculations have to be carried out in three steps using a linear approximation (Figure 2): (1) An effective rainfall threshold $\Omega(t_i)$ is calculated at every gauging station to estimate effective rainfall at rain samplers located inside the catchments. (2) The weighted sum $\mathbf{R}_\Sigma$ of effective rainfall series is defined so as to maximize the cross correlogram $Cor_{R_\Sigma,Q}(\tau)$ of $\mathbf{R}_\Sigma$ and the streamflow $\mathbf{Q}$ for small lags τ:

$$Cor_{R_\Sigma,Q}(\tau) = \frac{\sum_{i=1,N} [R_\Sigma(t_i) - 1] \times [Q(t_i + \tau) - \overline{Q}]}{\sqrt{\sum_{i=1,N} [R_\Sigma(t_i) - 1]^2} \sqrt{\sum_{i=1,N} [Q(t_i) - \overline{Q}]^2}} \tag{9}$$

where N is the number of sampling steps. For a given gauging station, the same threshold $\Omega(t_i)$ is used in (9) to estimate the effective rainfall at all weather stations. The weighting factors δ_k are calculated so that they maximize the objective function:

$$O_\Sigma = \sum_{l=1,L} Cor_{R_\Sigma,Q}(\tau_l) \tag{10}$$

defined from the increasing section of the cross correlogram $Cor_{R_\Sigma,Q}(\tau_l)$ (L is such that $Cor_{R_\Sigma,Q}(\tau_L)$ is maximum; that is, is the concentration time of the catchment). The optimization of the weighting factors, i.e., the maximization of the functional (10), is carried out in the hypercube $[0, 1]^P$ with $R_\Sigma = \sum_{k=1,p} \delta_k R_k / \sum_{k=1,p} \delta_k$. The functional (10) having no continuous derivatives, descent methods cannot be used; so the maximization of (10) is performed by dichotomy. At first, the cube $[0, 1]^P$ is split into 2^P cubes in which the functional (10) is evaluated. The one in which the functional (10) is maximum is split into 2^P cubes in which the functional (10) is evaluated again and so on. The accuracy of the weighting factors is $1/2^m$ after m iterations and the stations whose weighting factor is lower than $1/2^m$ are removed.

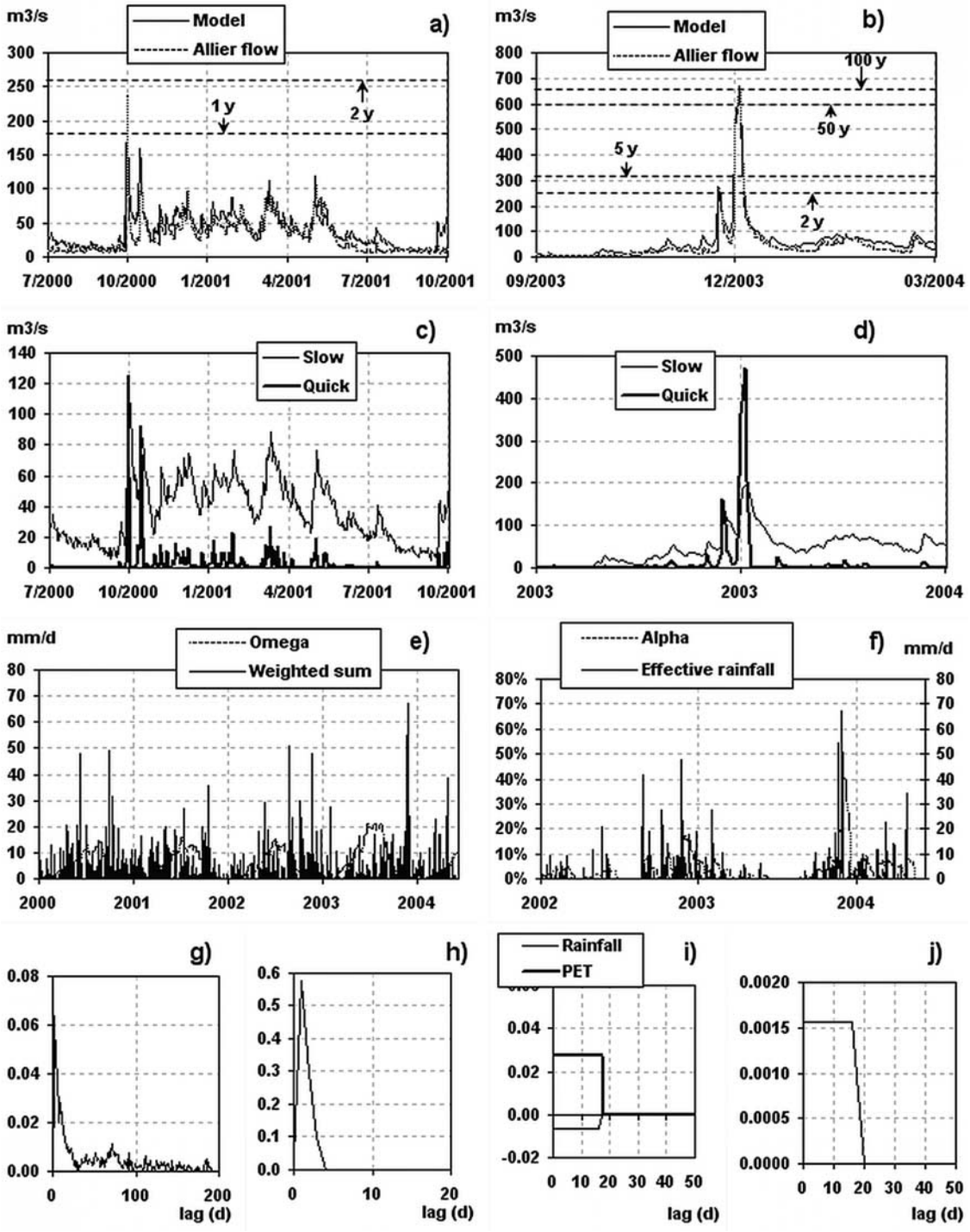


Figure 3. Inverse modeling of the Allier flow at Vieille Brioude: A 1-day sampling rate is used. (a and b) Comparison of model (1) and the observed flow. Return period of the flow is mentioned (wet years); (c and d) Hydrograph separation according to (1); (e) Effective rainfall threshold Ω deduced from (2) and (3) and the weighted rainfall expressed in Figure 7; (f) Fraction of quick transfer α of effective rainfall deduced from (4) and (5) and the effective rainfall; (g and h) slow and quick impulse responses Γ_s and Γ_q (both areas are 1); (i) impulse response of the effective rainfall threshold Ω to rainfall and to PET. Y axis and Y intersection are arbitrary; (j) impulse response $\Gamma_{\alpha,R_{eff}}$ for calculating the fraction of effective rainfall $\alpha(t_i)$ involved in quick flow.

[29] The recursive procedure reminds the stepwise method for the multiple regression to select the relevant variables. The radius of the circle containing rain samplers inside and outside the catchment is increased to include additional rain samplers until the cross correlogram $Cor_{R_\Sigma,Q}(\tau)$ is stabilized: Either the new rain samplers are irrelevant or they do not significantly improve the cross correlation. The optimization is achieved over periods of observation that include at least one 10-year return (or greater) event. At every step, the circle radius is decided so that a few numbers of new stations are integrated (less than 5), which approximately corresponds to a 10 km increase in the radius, depending on the density of the weather stations. The number of relevant stations oscillates between 3 and 5 as the radius increases while the sampling errors of rainfall decrease until the increase in the functional (10) is no longer significant.

[30] The robustness of the weighted rainfall series $\mathbf{R}_\Sigma/\overline{R}_\Sigma$ is investigated by comparing the model to the observed flow over the whole period of observation. At this step, the optimized rainfall is used as input of the transfer model and the normalized impulse responses, the effective rainfall threshold $\Omega(t_i)$ and the fraction $\alpha(t_i)$ of effective rainfall involved in quick flow are recalculated at every gauging station.

[31] In linear systems, streamflow $Q(t_i)$ depends linearly on rainfall $R(t_i)$ and the cross correlogram of $\mathbf{R}$ and $\mathbf{Q}$ $Cor_{R_\Sigma,Q}(\tau)$ is simply the impulse response of the system. This is not the general case because the rainfall $R(t_i)$ is not the effective rainfall, i.e., the fraction of rainfall that generates streamflow at the outlet. Moreover, streamflow $Q(t_i)$ results in quick and slow transfer of rainwater to the outlet.

[32] The nonlinearity of rainwater transfer enhances the contribution of exceptional events in the cross correlogram (9). So, because of the amplification of fast transfer, the contribution of exceptional rainfall events is strengthened. When the transfer is dominated by surface and the subsurface flow, nonlinearity can occur when soil is saturated, thus enhancing runoff while the infiltration of water through the unsaturated zone is considerably reduced, which lead to downward flow to the outlet. On the other hand, nonlinearity of transfer can occur when a stream is partly fed by groundwater during floods. Because of the dual porosity of some formations such as the chalk, nonlinear processes involve switching of groundwater recharge from matrix flow to macropore flow due to accumulated wetness.

[33] Owing to both the strong nonlinearities and the spatial variability of reduced rainfall data, the analysis of the cross correlogram at every step is not sufficient to ensure the optimum is reached or, in the contrary, to test the relevancy of the new weather stations that are integrated into the weighted sum $\mathbf{R}_\Sigma$. Only the quantification of discrepancies between the computed and the observed peak flows can be objectively used to end the recursive process. This criterion is based on the analysis of the relative standard deviations calculated from the estimation of peak flows that are expressed in comparison with a reference peak flow whose return period is 10 years.

3. Field Description

3.1. Allier Basin

[34] The Loire River is the longest river in France (1015 km) and drains a fifth of the country. Its management and development are a major issue both for the French and for bordering populations.

[35] The upper Allier and the Loire basins can get rainfall from the Cevennes-Vivarais (southern France), characterized by convective cells, mesoscale convective systems and fronts, producing extensive precipitation events [*Rivrain*, 1998; *Miniscloux et al.*, 2001; *Ducrocq et al.*, 2002; *Anquetin et al.*, 2003; *Wobrock et al.*, 2003; *Kirshbaum and Durran*, 2005; *Delrieu et al.*, 2005]. This type of event is related to exceptionally intensive rainfall concentrated on very restricted sectors. The factors contributing to such heavy rainfall are (1) the Mediterranean Sea, which acts as a reservoir of energy and moisture for the lower layers of the atmosphere, especially at the end of summer and beginning of fall, when the sea is still warm; (2) upper level cold troughs generally extending from the United Kingdom to the Iberian Peninsula generate a southerly flow that generates advection of the warm and moist air masses from the Mediterranean Sea toward the coast and destabilizes these air masses; and (3) the pronounced relief of the Mediterranean region, with the Alps, Pyrenees and Massif Central, triggers convection and also channels the low-level flow inducing low-level convergence. Occurring in winter or in spring, the floods generated by rainfall of oceanic origin affect only the lower part of the basin. The floods of the Allier basin, whose catchment area at Limons is 7005 km^2 (Figure 4 and Table 1), can play a major role in those of the middle Loire when there is a conjunction of rainfall from both the Cevennes and the Atlantic Ocean; the mixed floods can be exceptional like the 1866 flood.

[36] The precipitation rate on the Allier catchment decreases as we move downstream: It is 952 mm/y at Langogne, 623 mm/y at Brioude and 619 mm/y at Issoire. Snow is frequent on the upper basin and the Allier can be supplied mainly by snowmelt in spring.

[37] The Allier River and its main tributaries, the Alagnon and the Senouire, flow across the Massif Central and drain granite, metamorphic and volcanic formations. The upper Allier basin is constituted of granite, volcanic rocks and migmatite. Downstream of Vieille Brioude, the Allier drains Oligocene sediments. The mean contribution of groundwater to the Allier flow increases from Langogne where it is about 20% to Limons where it exceeds 40%. Because of the low rainwater concentration time as a result of the steep slope of the upper part of the basin, the specific discharge may reach 1.5 m^3 s km^2.

3.2. Maine Basin

[38] The Maine basin (Figure 5 and Table 2) in western France extends over a surface area of 22,020 km^2 and represents the principal subcatchment of the Loire. It is made up of three main subcatchments, the Mayenne (4160 km^2), the Sarthe (7380 km^2) and the Loir (7920 km^2). The magnitude of the floods in the basin is directly related to atmospheric circulation from the Atlantic. The 1995 10-year return flood followed a series of rainy episodes corresponding to passages of cold fronts as a result of the persistence of atmospheric circulation [*Galley and Fleury*, 2001]. This type of situation is very frequent during the winter, from November to March. The wettest months are November and December with 21% of the annual precipitation. The rainfall rate on the Maine catchment is between 934 mm/y (Saint Cyr) and 605 mm/y (Chateaudun).

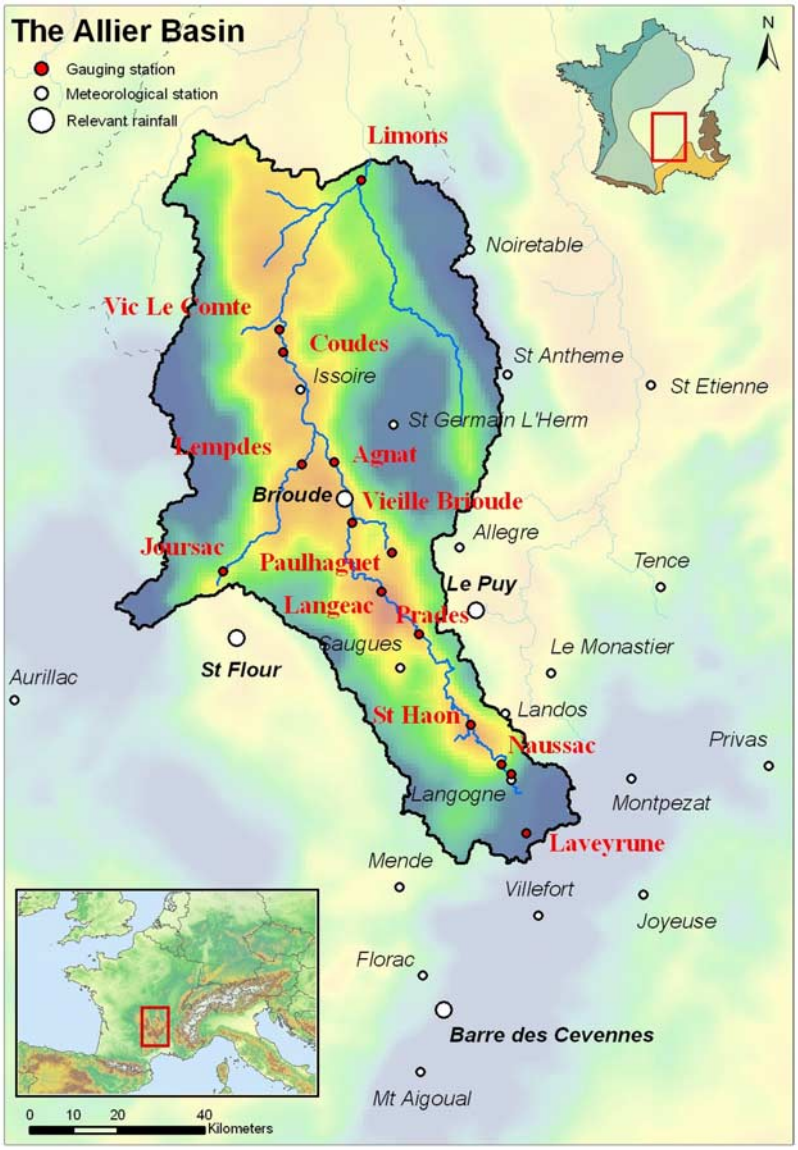


Figure 4. Allier basin.

It decreases as we move away from the sea, the steepest gradient being in the northwest to southeast direction. Snow is rare in the basin and snowmelt does not contribute significantly to water transfer processes.

[39] The Mayenne and the Oudon flow across the eastern part of the Armorican massif constituted mainly of granite. Upstream to downstream, the mean contribution of groundwater to the flow decreases from 50% to about 10% for the Oudon, whereas it decreases from 50% to almost 0% for the Mayenne. The Mayenne valley is narrow and causes rapid flood propagation. The Sarthe River crosses middle Jurassic limestone in the western part of the Paris basin. The mean contribution of groundwater reaches almost 40% at Saint Denis d'Anjou. The Loir River and its tributary, the Braye, cross sedimentary formations of the Paris basin, flinty clay and the Beauce limestone. About 50% of the Huisne flow, a tributary of the Sarthe, is supplied by Cenomanian groundwater, which is similar to the mean groundwater contribution to the Loir flow at Durthal. Groundwater contributes only 25% of the mean flow in the Braye, which flows mainly on flinty clay. Exceptionally, the specific discharge may reach 0.3 m^3 s km^2 locally. The 100-year flood at the outlet of the Maine basin corresponds to a specific discharge of 0.15 m^3 s km^2.

Table 1. Allier Basin

Tributary	Gauging Station	Area, km^2	Elevation, m	Concentration Time, d	Capacity, km^2 d
Allier	Laveyrune	48.8	993	1.1	54
	Langogne	324	931	1.1	356
	Naussac	396	877	1.1	436
	St Haon	514	740	1.1	565
	Prades	1350	540	1.2	1620
	Langeac	1781	495	1.2	2137
	Vieille Brioude	2269	434	1.6	3630
	Agnat	2950	360	1.9	5605
	Vic le Comte	5370	343	2.0	10,740
	Limons	7005	278	2.2	15,411
Senouire	Paulhaguet	155	538	1.4	217
Alagnon	Joursac	310	870	1.1	326
	Lempdes	984	450	1.1	1082

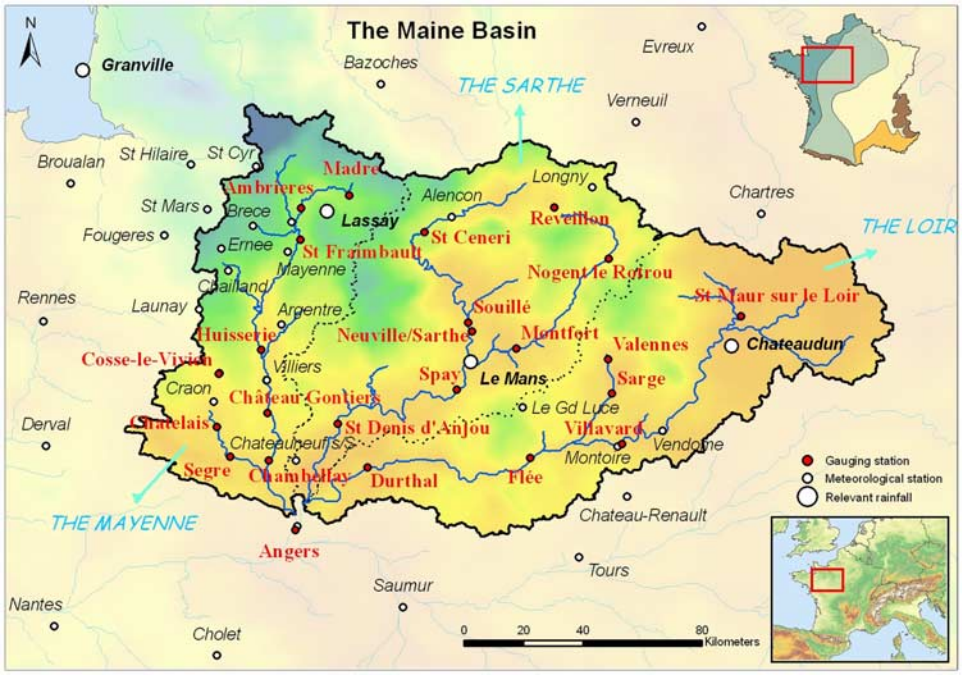


Figure 5. Maine basin.

3.3. Somme Basin

[40] The Somme basin (Figure 6 and Table 3) is located in northwestern France, near the Channel. It is exposed to an oceanic influence, generally mild in winter. The rainfall rate is governed mainly by northwest winds coming from the sea. The wettest months are November and December with 20.5% of the annual precipitation. The isohyets show significant variations in mean rainfall values over the catchment. It rains more on the plateaus (771 mm/y at Epehy) than in the valleys (657 mm/y at Amiens). The mean rainfall values increase near the sea (818 mm/y at Abbeville, 894 mm/y at Vron). The variations reach 42% between extreme values. Snow is rare in the basin and snowmelt does not contribute significantly to the water transfer processes.

[41] The catchment has a surface area of 5560 km^2 at the Abbeville gauging station. The hydrographic network of the basin is fed mainly by the chalk groundwater. A catastrophic flood occurred during the spring 2001, whereas no such event had been recorded for centuries [*Pinault et al.*, 2005]. Nonlinear processes, involving a hydraulic continuity between the macropores of the unsaturated zone and the chalk groundwater, govern water migration through the unsaturated zone. In spite of nonlinear transfer processes, the Somme basin is characterized by a long concentration time, owing to the contribution of groundwater to surface water flow. The 2001 flood, therefore, resulted from the basin's very low drainage capacity since the specific discharge never exceeded 0.02 m^3 s km^2.

4. Results

[42] Rainfall and PET data were obtained from the Meteo-France data bank (http://climatheque.meteo.fr) and the streamflows from the French ministry of environment (http://hydro.eaufrance.fr).

4.1. Allier Basin

[43] The different steps of inverse modeling are illustrated for the Vieille Brioude flow of the Allier basin for which a 1-day sampling rate is used in calculations (Figure 3). The area of the Vieille Brioude catchment is 2269 km^2 and its capacity, i.e., the product of the catchment area by the concentration time of streamflow, is 3630 km^2 d. Floods at the outlet can result from either Mediterranean or Atlantic rainfall.

Table 2. Maine Basin

Tributary	Gauging Station	Area, km^2	Elevation, m	Concentration Time, d	Capacity, km^2 d
Mayenne	Madre	335	129	1.1	369
	Ambrieres	828	100	1.8	1490
	St Fraimbault	1851	85	1.9	3517
	Huisserie	2890	41	1.9	5491
	Chateau-Gontiers	3910	26	2.0	7625
	Chambellay	4160	20	2.1	8528
Sarthe	St Ceneri	908	121	2.1	1907
	Souillé	2700	49	2.2	5940
	Neuville sur Sarthe	2716	47	2.3	6247
	Spay	5285	38	2.3	12,156
	St Denis d'Anjou	7380	22	2.3	16,974
Loir	Villavard	4545	65	2.9	13,181
	Flée	5940	46	3.0	17,820
	Durthal	7920	21	4.1	32,472
Oudon	Cosse le Vivien	133	55	1.2	160
	Chatelais	734	35	1.6	1174
	Segré	1310	21	1.8	2358
Huisne	Reveillon	78.3		1.1	82
	Nogent le Rotrou	827	102	1.9	1571
	Montfort	1890	56	2.2	4064
Braye	Valennes	270	101	1.1	297
	Sarge	497	82	1.2	572
Maine	Angers	22,020	13	2.8	61,656



Figure 6. Somme basin.

[44] The weighted rainfall $\mathbf{R}_\Sigma$ maximizes the objective function (10) (Figure 7). Although 23 rain samplers are used, only four can be considered to be relevant since they contribute to the weighted rainfall $\mathbf{R}_\Sigma$ by reducing sampling errors: Barre des Cévennes, Brioude, Le Puy, and St Flour. Although the rain sampler at Langogne is well correlated with the Allier flow, it is not integrated into the weighted rainfall $\mathbf{R}_\Sigma$ because of the redundancy of information.

[45] The weighted rainfall $\mathbf{R}_\Sigma$ is built from four rainfall series that show the double influence of both Mediterranean rainfall in the NW-SE direction, which coincides with the Allier valley, and Atlantic rainfall in the W-E direction.

[46] Rainfall-streamflow models that are parsimonious and regularized (the l_2 norms $(\sum\Gamma_i^2)^{1/2}$ of the quick and slow impulse responses $\mathbf{\Gamma}_q$ and $\mathbf{\Gamma}_s$ are bounded) enable a realistic estimation of peak flow accuracy. Indeed, once the nonlinear effects affecting rainwater transport to the outlet of catchments have been taken into account, the main discrepancies observed between the observed and computed peak flows result from the heterogeneity of rainfall at the catchment scale, which is illustrated in Figure 8. In Figures 8a and 8b the weighted rainfall is optimized by using rain samplers inside and outside the catchment, whereas in Figures 8c and 8d it is optimized by using rain samplers inside the catchment and outside but close to the watershed. The discrepancies between the observed flow and the model are approximately the same in Figures 8a and 8c, which confirms they are resulting from systematic errors, the transfer processes being altered after the flood on 30 September 2000, whereas the peaks are accurately explained. On the other hand, the doublet that is represented in Figures 8b and 8d shows that the first peak on 24–25 November 2003 can be properly explained only if the rainfall measured at Barre des Cévennes is introduced into the optimization process of the weighted sum $\mathbf{R}_\Sigma$ although the rain sampler is not located on the Allier basin.

[47] Inverse modeling, therefore, enables us to estimate the random errors of peak flow once the principal systematic errors have been removed. These random errors are closely linked to catchment capacity and the return period of the events (Table 4). Relative standard deviations are calculated from the estimation of peak flows that are expressed in comparison with a reference peak flow whose return period is 10 years $Q^{(10y)}$:

$$S = 1/Q^{(10y)}\left(1/n\sum_{i=1,n}(Q_{obs}-Q_{mod})^2\right)^{1/2} \qquad (11)$$

where n is the number of peak flows, Q_{abs} and Q_{mod} are the observed and the computed peak flows, respectively. In this

Table 3. Somme Basin

Tributary or Observation Well	Station	Area, km^2	Elevation, m	Concentration Time, d	Capacity, km^2 d
Somme	Abbeville	5560	7	100	556,000
	Peronne	1294	48	10	12,940
Nievre	Etoile	269	10	90	24,210
Hallue	Bavelincourt	115	46	60	6900
Selle	Plachy	524	39	10	5240
Avre	Moreuil	642	36	10	6420
Obs. well	00323X0080		45	100	
	00332X0007		71	80	
	00341X0012		127	100	
	00342X0025		129	150	
	0471X0010H1		77	110	
	00347X0002		143	110	
	0463X0036H1		87	70	
	00465X0007		107	190	
	0478X0002P		79	190	
	00487X0015		102	210	
	00624X0085		47.5	20	

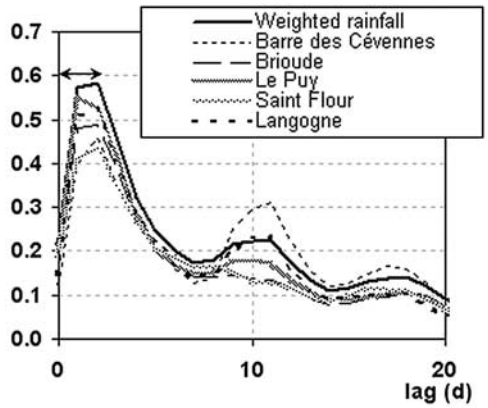


Figure 7. Cross correlograms between the Allier flow at Vieille Brioude and effective rainfall. Arrow indicates the lags from which the increasing part of the cross correlogram is maximized (2 days). Weighted rainfall:

$$\mathbf{R}_\Sigma = 0.07\,Barre\ des\ Cévennes + 0.26\,Brioude + 0.52\ Le\ Puy + 0.15\,St\ Flour$$

way, all the catchments of the Allier basin contribute to the estimation of the deviations, regardless of their capacity.

[48] Catchment capacity has a strong influence on the estimation errors of peak flow regardless of the return period (Table 4). Modeling of floods in small catchments whose capacity is less than 500 km^2 d requires much denser information than for rainfall regionalization at medium and large catchment scale. On the other hand, a sparse network of suitable rain samplers allows accurate coarse-scale modeling of peak flow whose return period is at least 5 years. The standard deviation of around 10% includes systematic errors resulting from the parsimony of degree of freedom of the inverse models as well as the regularization technique.

Table 4. Standard Deviation of Peak Flow Expressed in Comparison With a 10-Year Return Event (Allier Basin)[a]

	Capacity $\leq$500 km^2 d		Capacity >500 km^2 d	
Return Period	Number of Events	Standard Deviation	Number of Events	Standard Deviation
RP < 1 y	87	16.9%	57	10.7%
1 y < RP < 5 y	20	30.5%	31	17.3%
RP > 5 y	18	30.5%	11	9.9%

[a]The capacity of the piezometric surfaces is unknown because their area may vary significantly during the recharge period.

4.2. Maine Basin

[49] The calculations are very similar to those of the Allier basin. Here again the sampling rate is 1 day. 38 rainfall samplers are used. The weighted rainfall $\mathbf{R}_\Sigma$ shows five main independent influences (Figure 9). The rainfall rate on the Mayenne basin is clearly controlled by the oceanic influence and its inland penetration in the W-E direction.

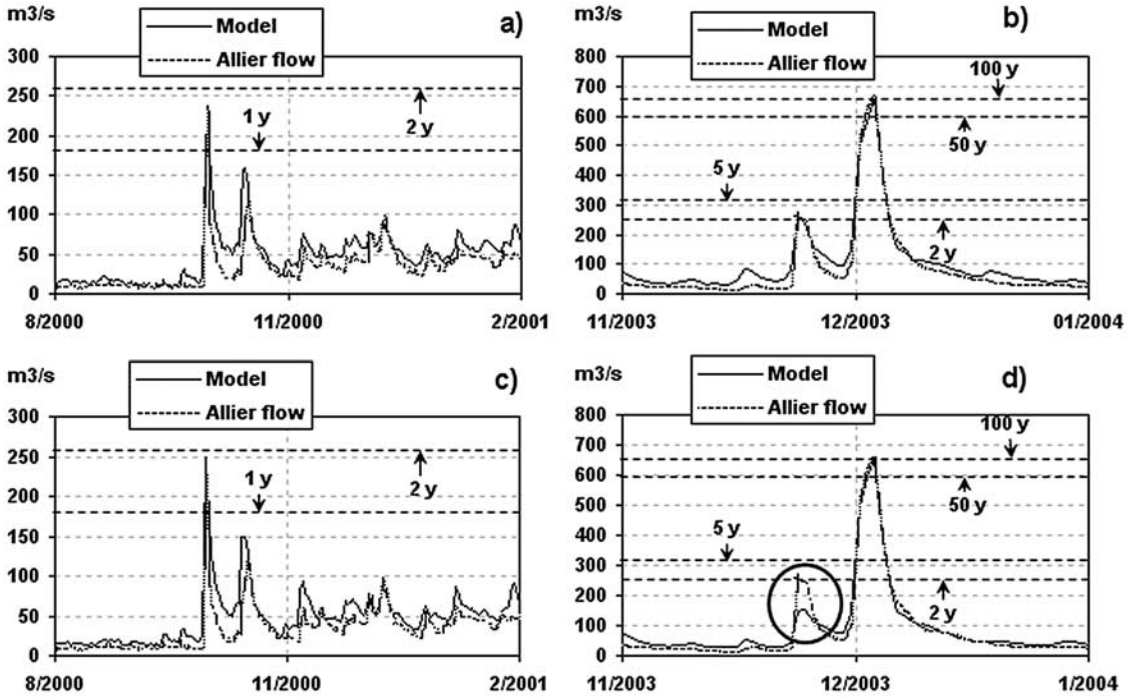


Figure 8. Inverse modeling of the Allier flow at Vieille Brioude: comparison of model (1) and the observed flow. Return period of the flow is mentioned (wet years). (a and b) Weighted rainfall is estimated from rain samplers that are located inside and outside the catchment; (c and d) weighted rainfall is estimated from rain samplers that are located inside the catchment and outside but close to the watershed:

$$\mathbf{R}_\Sigma = 0.25\,Brioude + 0.20\,Langogne + 0.32\ Le\ Puy + 0.23\,St\ Flour$$

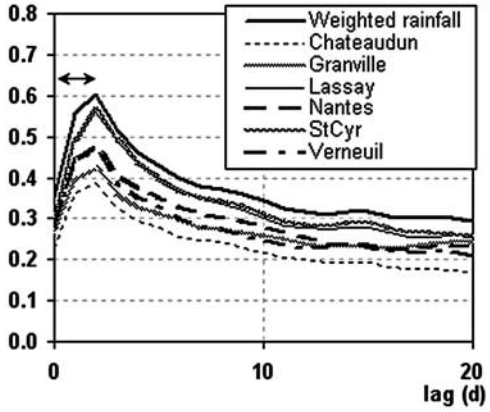


Figure 9. Cross correlograms between the Mayenne flow at Huisserie and effective rainfall. Arrow indicates the lags from which the increasing part of the cross correlogram is maximized (2 days). Weighted rainfall:

$$\mathbf{R}_{\Sigma} = 0.19\,Chateaudun + 0.21\,Granville + 0.16\,Nantes + 0.29\,StCyr + 0.15\,Verneuil$$

[50] The influence of rainfall sampling on the accuracy of the model is illustrated in Figure 10. The weighted rainfall $\mathbf{R}_{\Sigma}$ is optimized from rain samplers located inside and outside the catchment in Figures 10 and 10b while it is optimized from rain samplers located inside and in the vicinity of the catchment in Figures 10c and 10d. Peak flows higher than 300 m^3 s whose return period is 2 years are better estimated when Granville, Chateaudun, Nantes and Verneuil rain samplers are introduced into the optimization process although they are far from the Mayenne catchment at Huisserie. Although the 1995 flood is accurately reproduced by both models, the succession of floods that occurred in 2001 are poorly explained when rain samplers are limited to the catchment.

Table 5. Standard Deviation of Ppeak Flow Expressed in Comparison With a 10-Year Return Event (Maine Basin)

	Capacity $\leq$500 km^2 d		Capacity >500 km^2 d	
Return Period	Number of Events	Standard Deviation	Number of Events	Standard Deviation
RP < 1 y	135	11.9%	333	8.7%
1 y < RP< 5 y	38	12.1%	190	11.2%
RP > 5 y	14	10.0%	50	10.0%

[51] The standard deviation deduced from the comparison of observed and computed peak flow values depends only a little on the catchment capacity and the return period (Table 5), which results from the homogeneity of stratified rainfall of oceanic origin. Nevertheless, as for the Allier basin where rainfall is both stratified and convective, the standard deviation of peak flow expressed in comparison with a 10-year return event is close to 10% for return periods equal to or greater than 5 years.

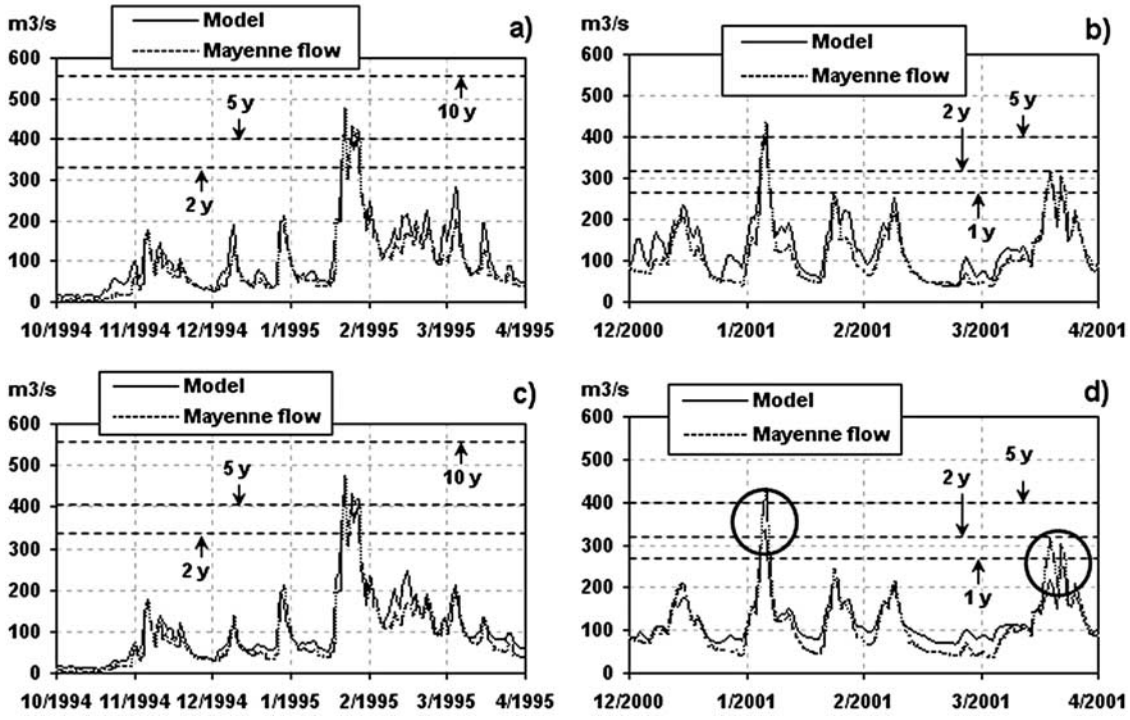


Figure 10. Inverse modeling of the Mayenne flow at Huisserie: comparison of model (1) and the observed flow. Return period of the flow is mentioned (wet years); (a and b) weighted rainfall is estimated from rain samplers that are located inside and outside the catchment; (c–d) weighted rainfall is estimated from rain samplers that are located inside the catchment and outside but close to the watershed:

$$\mathbf{R}_{\Sigma} = 0.05\,Brece + 0.40\,Lassay + 0.06\,Mayenne + 0.32\,StCyr + 0.17\,StMars$$

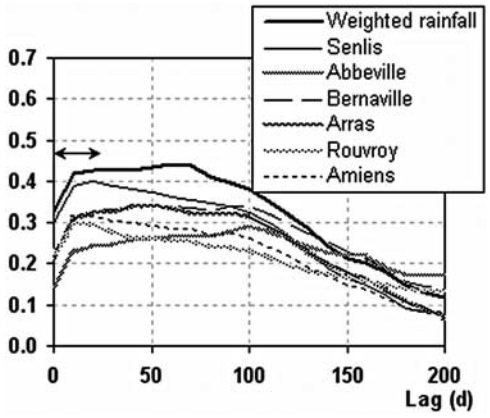


Figure 11. Cross correlograms between the Somme flow at Abbeville and effective rainfall. Arrow indicates the lags from which the increasing part of the cross correlogram is maximized (20 days). Weighted rainfall:

$$\mathbf{R}_\Sigma = 0.52\,Arras + 0.13\,Bernaville + 0.08\,Rouvroy + 0.27\,Senlis$$

4.3. Somme Basin

[52] The calculations are done with a 10-day sampling rate. 26 rainfall samplers are used. In Figure 11, the maximization of the objective function (10) is done for the Somme flow at Abbeville, near the outlet of the Somme basin. The closest stations that allow a significant improvement of the cross correlation are located on both sides of the catchment. The Arras station is located 56 km at the north of Amiens and the Senlis station is 76 km to the south. The weighted rainfall $\mathbf{R}_\Sigma$ therefore shows that the rainfall rate on the Somme basin reflects an N-S influence from the North Sea and the Channel.

[53] The different steps leading to the comparison between the observed and the computed flow values are shown in Figure 12. Not only is the model able to explain high water flow over the entire period of flooding but low water flow is also properly modeled (Figure 12a). The slow component represents the main contribution of the Somme flow (Figure 12b and 12e). This component expresses water migration through the matrix of the unsaturated zone down to the chalk aquifer. Such a process is the result of switching behavior of groundwater recharge from matrix flow to macropore flow due to accumulated wetness over several years, hence the important contribution of the quick component to the Somme flow (Figures 12b, 12c 12d and 12f). The relevancy of rain samplers is represented in Figure 13. The weighted rainfall that is optimized from rain samplers inside and outside the catchment enables to explain the 2001 flood properly (Figure 13a). The discrepancies that are observed at the end of the flood are resulting from systematic errors since the Somme flowed out of bank, which decreased drainage capabilities of the river. On the other hand, the peak flow cannot be explained properly when rain samplers are located inside the catchment only (Figure 13b). In particular the initiation of the flood cannot be explained without integrating the Arras rain sampler into the optimization process of the weighted rainfall $\mathbf{R}_\Sigma$.

[54] Thus the completeness of information for the decadal data of rainfall is supported mainly by two weather stations located at a distance close to the longest dimension of the catchment, which indicates the correlation distance of reduced rainfall data. Estimation error between observed and computed streamflow is deduced from several subcatchments of the Somme basin that are equipped with gauging stations, which allows the optimization of rainfall data at the subcatchment scale. Moreover, several observation wells located close to water table outlets are computed.

[55] Owing to the possible systematic errors related to the calibration curve of gauging stations, particularly for low water stage, standard deviations (Table 6) are higher for streamflow (8.6%) than for pressure head (5.6%) that is linearly related to groundwater recharge when the aquifer is homogeneous, which may be assumed in Somme basin chalk at broad scale.

5. Discussion

[56] Only a few suitable rain gauges are necessary to accurately model peak flow at medium and large catchment scales. Nevertheless, daily rainfall can be locally very erratic depending on the local environment and atmospheric turbulence. The apparent spatial homogeneity of short-term rainfall events, therefore, relies on statistical considerations. Let us consider local fluctuations whose correlation distance is l and regulation time is τ. If S and t_c denote the catchment area and the rainwater concentration time to the outlet, respectively, this apparent homogeneity results in the large number of elementary cells whose capacity is $l^2\tau$ that is aggregated into the cell associated with the catchment with a capacity of St_c. A global behavior results in the superposition of multiple influences that are globally represented from rainfall fluxes through the watersheds as a result of atmospheric circulation and advection. These fluxes, characterized from a sparse network of rain gauges, reflect (1) oceanic intrusion and its dampening as it moves over the continent where both evaporation and precipitation occur and (2) the condensation of warm wet air when it collides with cold air like what happens for storms on the Cevennes Mountains when warm Mediterranean air encounters cold mountain air. The greater the return period of observed events, the more accurate the flood estimation will be because of the coherence of the elementary cells used to produce such exceptional events at the catchment scale. At broad scale, the heterogeneity of elementary cells therefore appears mainly for events whose effect is weak, characterized by a small return period.

[57] The completeness of rainfall data to explain nonlinear transfer processes can be ascertained from a very small number of weather stations covering an area of several thousands of km^2. Everything happens as if most of the reduced data from rain gauges were not sensitive enough to reliably represent the rainfall fluxes generated by atmospheric circulation through the watersheds, and their response to short-term rainfall events is too noisy to provide relevant information for the transfer models. Moreover, the selected stations are always the same for each catchment unless two optimized weighted rainfalls $\mathbf{R}_\Sigma$ are equivalent, which results both on the capacity of the catchments and on the return period of flood events: So local storm events that may be due to the front passage or summer thunderstorm are excluded from the statistics.

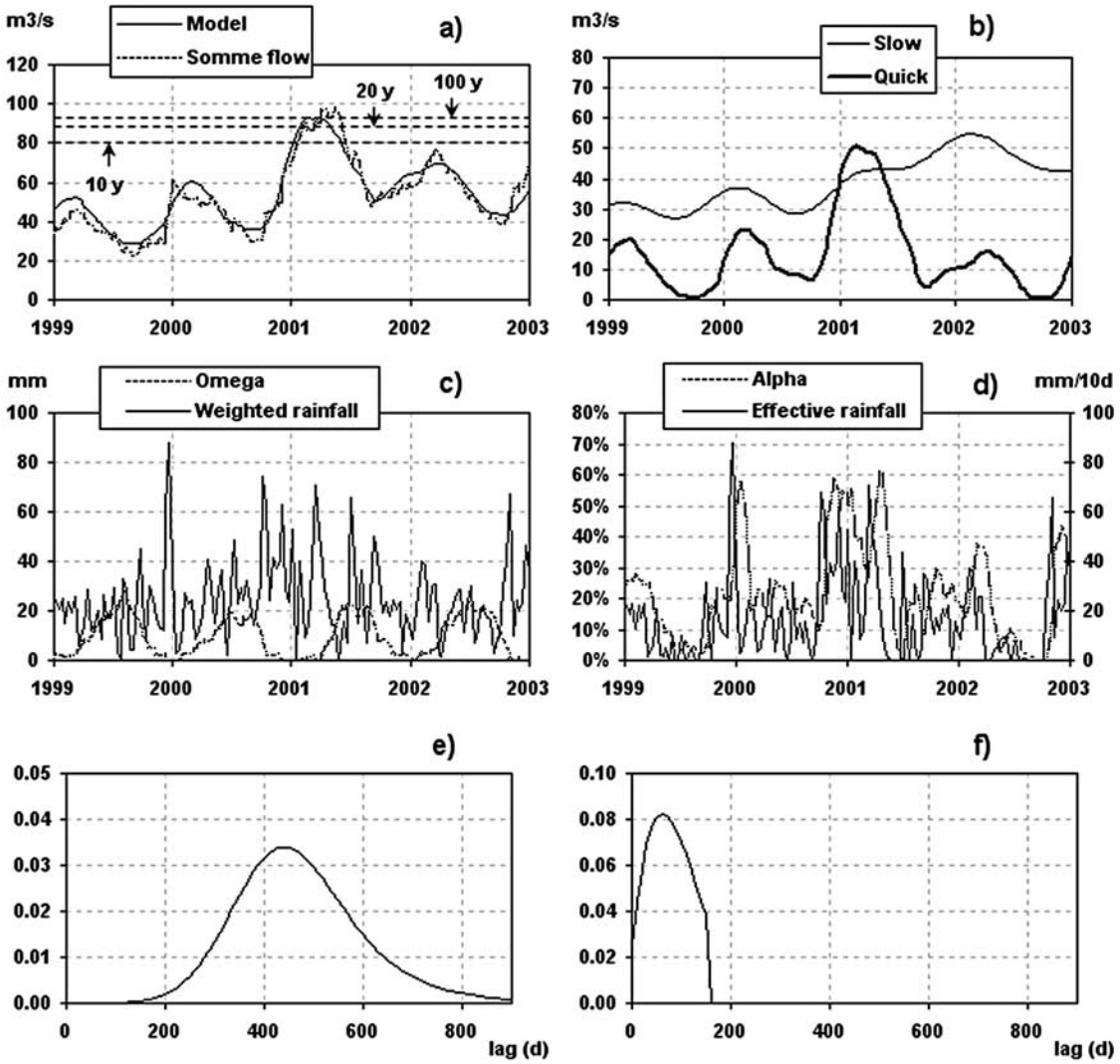


Figure 12. Different steps for optimizing the weighted rainfall of rainfall data: (a) comparison of model (1) and the observed flow. Return period of the flow is mentioned (wet years); (b) hydrograph separation according to (1); (c) effective rainfall threshold $\boldsymbol{\Omega}$ and the weighted rainfall; (d) fraction of quick transfer $\boldsymbol{\alpha}$ of effective rainfall and the effective rainfall; (e and f) slow and quick impulse responses $\boldsymbol{\Gamma}_s$ and $\boldsymbol{\Gamma}_q$.

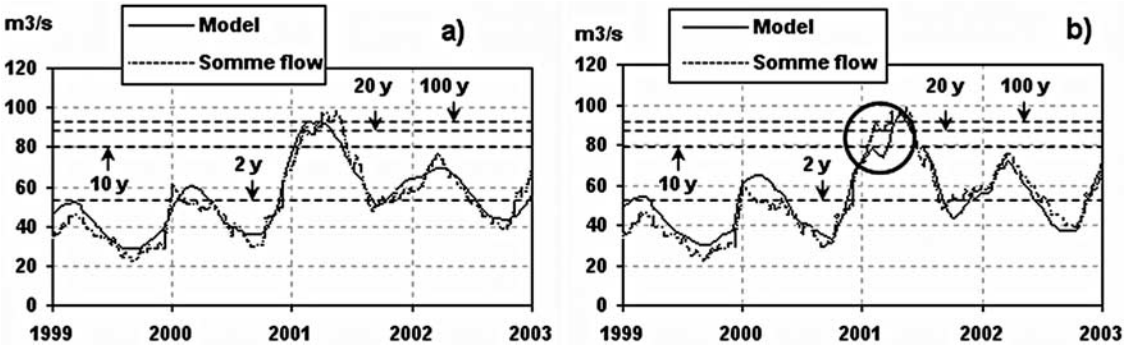


Figure 13. Inverse modeling of the Somme flow at Abbeville: comparison of model (1) and the observed flow. Return period of the flow is mentioned (wet years). (a) Weighted rainfall is estimated from rain samplers that are located inside and outside the catchment; (b) weighted rainfall is estimated from rain samplers that are located inside the catchment and outside but close to the watershed:

$$\mathbf{R}_{\Sigma} = 0.31\ \mathit{Epehy} + 0.07\ \mathit{Oisemont} + 0.52\ \mathit{Rouvroy} + 0.10\ \mathit{Vron}$$

Table 6. Standard Deviation of Peak Flow Expressed in Comparison With a 10-Year Return Event (Somme Basin)

	Number of Events	Standard Deviation
Streamflow	18	8.6%
Observation well	25	5.6%

[58] Thus upscaling at broad scale makes the notion of proximity of rain gauges to the catchment obsolete, only their representativeness of the fluxes through the watersheds as a result of atmospheric circulation is relevant. For example, the rain gauges of Arras (Somme basin), Granville (Maine basin) and Barre des Cevennes (Allier basin) fulfill such conditions. Located upstream in the direction of the dominant wind that generates heavy rain, their influence is significant over long distances of up to a hundred kilometers. In the opposite case, rain gauges located inside the catchment may not be representative of short-term rainfall events at the catchment scale, and rainfall data may increase random errors when they are used as inputs of hydrological models without selection on the basis of the analysis of the cross correlogram between rainfall data and streamflows.

6. Regionalization of Rainfall at Broad Scale

[59] The broad-scale regionalization of rainfall can be achieved by optimizing the weighted sum $\mathbf{R}_\Sigma$ on the mosaic of subcatchments included between the watersheds intersecting two successive gauging stations. Equation (1), therefore, has to be transformed into the following equation:

$$Q_{down}(t_i)/\overline{Q}_{down} = \mathbf{\Gamma}_s * \mathbf{R}_s/\overline{R}_{eff} + \mathbf{\Gamma}_q * \mathbf{R}_q/\overline{R}_{eff} + \mathbf{\Gamma}_{up} * \mathbf{Q}_{up}/\overline{Q}_{up} + \varepsilon \quad (12)$$

where $\mathbf{Q}_{up}$ and $\mathbf{Q}_{down}$ are the flows observed at the upstream and the downstream stations, respectively. Now, $\mathbf{\Gamma}_{up}$ is the impulse response of the upstream flow to the downstream flow, the area of which is the ratio of the mean upstream flow to the mean downstream flow:

$$\sum \mathbf{\Gamma}_{up} = \overline{Q}_{up}/\overline{Q}_{down} \quad (13)$$

The correlation distances of observed rainfall and of the weighted sum $\mathbf{R}_\Sigma$ are represented in Figure 14, Figure 15, and Figure 16 for the three basins.

[60] In the Allier basin, the weighted rainfall exhibits a weak correlation decrease from Langogne to Vieille Brioude, then decreases to Agnat before showing a plateau to Limons. The behavior of the correlation distance reflects the gradual attenuation of the Mediterranean influence as the distance increases from the watershed at the upper limit of the Allier basin to Vieille Brioude, then the step between Vieille Brioude and Agnat materializes the limit between the Mediterranean and the continental influences. The apparent spatial decorrelation observed for rain samplers results from noise, i.e., sampling errors.

[61] As opposed to what occurs in the Allier basin, the regionalization of reduced rainfall data on the Maine basin shows the strong correlation of daily rainfall at the basin scale regardless of the axis. The rainfall correlation decreases very slowly as the distance from the sea increases, as a result of inland penetration of the front.

[62] The Somme basin is characterized by an influence of the Channel at the edge of the basin, and then the inland penetration of the front shows a very long correlation distance of decadal rainfall from Bernaville.

7. Conclusion

[63] An inverse approach based on previous studies was demonstrated to evaluate the representativeness of rain samplers to explain flood events. Because of the small degree of freedom of the inverse models the regularized solutions enable to separate systematic errors resulting from subtle transfer processes not taken into account into the models and random errors produced by the propagation of sampling errors of rainfall. So, rainfall sampling can be optimized without interfering on the calibration of the inverse models, by minimizing the discrepancies between the observed and the computed streamflows.

[64] At catchment scale (500 to 25,000 km^2), the best representativeness of rainfall is obtained from the reduced rainfall $\mathbf{R}_\Sigma/\overline{R}_\Sigma$ that maximizes the cross correlation with the flow observed at the outlet of the catchment, this reduced rainfall being weighted by the mean rainfall height over the catchment (from the isohyets). The reduced weighted rainfall $\mathbf{R}_\Sigma/\overline{R}_\Sigma$ is characterized by the sparseness of the rain samplers from which it is defined and the long distance between them. As a corollary, the regionalization of rainfall data for broad-scale modeling shows a very long correlation length of daily (or decadal) rainfall. This correlation length becomes shorter only at the frontier between two climate systems. This contradicts the widespread assumption that for distributed rainfall-runoff model, the denser the network of rain samplers used as input in the models, the more accurate broad-scale flood model will be. Random errors of the models resulting from the propagation of sampling errors of rainfall could be substantially lessened by using the reduced rainfall $\mathbf{R}_\Sigma/\overline{R}_\Sigma$ weighted by the mean rainfall height over the catchment as input of the models instead of the row rainfall data. Flood events can be reproduced accurately only if some particular weather stations are introduced into the models, those stations being representative of the rainfall fluxes through the watersheds as a result of atmospheric circulation and advection. Such stations may be located outside the catchments and their

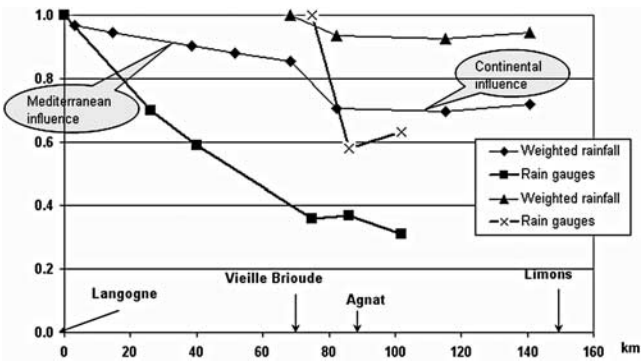


Figure 14. Correlation distance of rainfall in the Allier basin (sampling rate = 1 day).

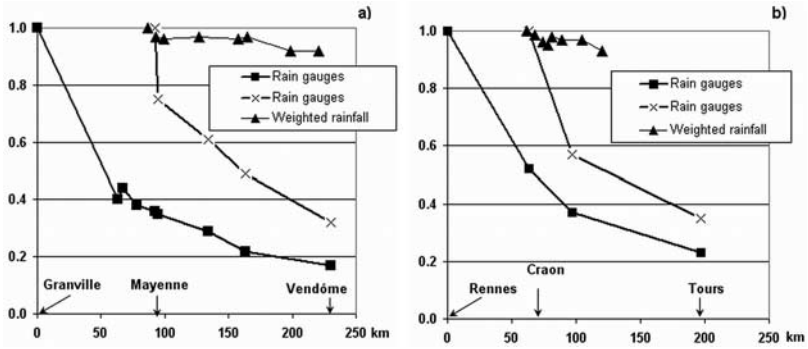


Figure 15. Correlation distance of rainfall in the Maine basin (sampling rate = 1 day) along two directions (a) Granville-Vendôme axis and (b) Rennes-Tours axis.

omission as input of the models may involve a lack of information, leading to a poor representation of exceptional flood events. Furthermore, the temporal and the spatial structure of the low-noise optimal combination $\mathbf{R}_\Sigma$ are much easier to decipher than those of raw rainfall data, which improves the reliability of both flood models and flood predictions. Precipitation events or sequences need to be generated with an appropriate spatial and temporal structure, and methods are capable of extension to represent climate change. In particular, the spatial resolution of regionalized rainfall events is compatible with that of the General Circulation Models and the temporal structure is well adapted to the integration of current and possible future effects of climate variability and change on flood risk. Modeling of large basins may require connected models linked from the cross correlation of the corresponding weighted rainfall $\mathbf{R}_\Sigma$. Considering the importance of flood forecasting in large basins requiring evaluation of joint probabilities of fluvial flooding from a combination of contributory effects, this work may be generalized to a large variety of issues of major strategic importance, with major economic and social implications.

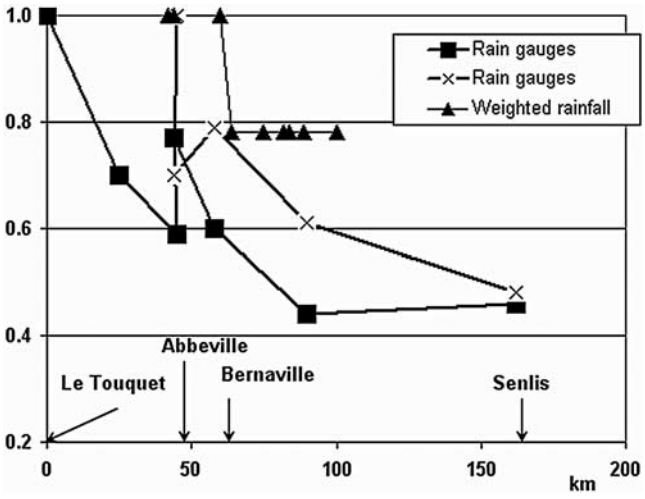


Figure 16. Correlation distance of rainfall in the Somme basin (sampling rate = 10 days).

[65] **Acknowledgments.** This work was supported by BRGM's research grant within the flood project. Thanks are due to three anonymous reviewers, the Associate Editor, and the Editor for valuable comments and advice during the review process.

References

Anquetin, S., F. Miniscloux, J. D. Creutin, and S. Cosma (2003), Numerical simulation of orographic rain, *J. Geophys. Res.*, *108*(D8), 8386, doi:10.1029/2002JD001593.

Burn, D. H. (1997), Catchments similarity for regional flood frequency analysis using seasonality measures, *J. Hydrol.*, *202*, 212–230.

Cameron, D., K. J. Beven, and J. Tawn (2000), An evaluation of three stochastic rainfall models, *J. Hydrol.*, *228*, 130–149.

Cameron, D., K. J. Beven, and J. Tawn (2001), Modelling extreme rainfalls using a modified random pulse Bartlett-Lewis stochastic rainfall model (with uncertainty), *Adv. Water Resour.*, *24*, 203–211.

Cowpertwait, P. S. P. (1994), A generalised point process model for rainfall, *Proc. R. Soc.*, *447*, 23–37.

Cowpertwait, P. S. P., C. G. Kilsby, and P. E. O'Connell (2002), A space-time Neyman-Scott model of rainfall: Empirical analysis of extremes, *Water Resour. Res.*, *38*(8), 1131, doi:10.1029/2001WR000709.

Cox, D. R., and V. Isham (1980), *Point Processes*, CRC Press, Boca Raton, Fla.

Cox, D. R., and V. Isham (1988), A simple spatial-temporal model of rainfall, *Proc. R. Soc., Ser. A*, *415*, 317–328.

Delrieu, G., et al. (2005), The catastrophic flash-flood event of 8–9 September 2002 in the Gard region, France: A first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory, *J. Hydrometeorol.*, *6*, 34–52.

De Michele, C., and R. Rosso (1995), Self-similarity as a physical basis for regionalisation of flood probabilities, paper presented at International Workshop on Hydrometeorology Impacts and Management of Extreme Floods, Univ. of Perugia, Perugia, Italy, 13–17 November.

De Michele, C., and R. Rosso (2002), A multi-level approach to flood frequency regionalisation, *Hydrol. Earth Syst. Sci.*, *6*(2), 185–194.

Ducrocq, V., D. Ricard, J. P. Lafore, and F. Orain (2002), Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field, *Weather Forecasting*, *17*, 1236–1256.

Foufoula-Georgiou, E. (1998), On scaling theories of space-time rainfall: Some recent results and open problems, in *Stochastic Methods in Hydrology*, edited by O. Barndorff-Nielsen et al., pp. 25–72, World Sci., Hackensack, N. J.

Fowler, H. J., C. G. Kilsby, and P. E. O'Connell (2000), A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions, *Hydrol. Earth Syst. Sci.*, *4*(2), 263–282.

Galley, R., and J. Fleury (2001), Rapport de la commission d'enquête sur les causes des inondations, *Rep. 3386*, French Natl. Assem., Paris.

Gupta, V. K., and D. R. Dawdy (1994), Regional analysis of flood peaks: Multiscaling theory and its physical basis, in *Advances in Distributed Hydrology*, edited by R. Rosso et al., pp. 149–168, Water Resour. Publ., Highlands Ranch, Colo.

Gupta, V. K., and E. Waymire (1990), Multiscaling properties of spatial rainfall and river flow distributions, *J. Geophys. Res.*, *95*, 1999–2009.

Gupta, V. K., O. J. Mesa, and D. R. Dawdy (1994), Multiscaling theory of flood peaks: Regional quantile analysis, *Water Resour. Res.*, *30*, 3405–3421.

Gyasi-Agyei, Y., and G. Willgoose (1997), A hybrid model for point rainfall modelling, *Water Resour. Res.*, *33*, 1699–1706.

Hay, L. E., M. P. Clark, R. L. Wilby, W. J. Gutowski, G. H. Leavesley, Z. Pan, R. W. Arritt, and E. S. Takle (2002), Use of regional climate model output for hydrologic simulations, *J. Hydrometeorol.*, *3*(5), 571–590.

Hosking, J. R. M., and J. R. Wallis (1993), Some statistics useful in regional frequency analysis, *Water Resour. Res.*, *29*, 271–281.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional Frequency Analysis: An Approach Based on L-moments*, Cambridge Univ. Press, Cambridge, UK.

Khaliq, M., and C. Cunnane (1996), Modelling point rainfall occurrences with the modified Bartlett-Lewis rectangular pulses model, *J. Hydrol.*, *180*, 109–138.

Kirshbaum, D. J., and D. R. Durran (2005), Observations and modeling of banded orographic convection, *J. Atmos. Sci.*, *62*(5), 1463–1479.

Miniscloux, F., J. D. Creutin, and S. Anquetin (2001), Geostatistical analysis of orographic rainbands, *J. Appl. Meteorol.*, *40*, 1835–1854.

Onof, C., and H. S. Wheater (1993), Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model, *J. Hydrol.*, *149*, 67–95.

Onof, C., and H. S. Wheater (1994), Improvements of the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model, *J. Hydrol.*, *157*, 177–195.

Onof, C., and H. S. Wheater (1995), Modelling of rainfall time-series using the Bartlett-Lewis model, *Proc. Inst. Civil Eng., Water Maritime Energy*, *112*, 362–374.

Onof, C., P. Northrop, H. S. Wheater, and V. Isham (1996), Spatiotemporal storm structure and scaling property analysis for modeling, *J. Geophys. Res.*, *101*, 26,415–26,425.

Pinault, J.-L., and F. Berthier (2007), A methodological approach to characterize the resilience of aquatic ecosystems with application to Lake Annecy, France, *Water Resour. Res.*, *43*, W01418, doi:10.1029/2006WR005125.

Pinault, J.-L., and S. Schomburgk (2006), Inverse modeling for characterizing surface water/groundwater exchanges, *Water Resour. Res.*, *42*, W08414, doi:10.1029/2005WR004587.

Pinault, J.-L., H. Pauwels, and C. Cann (2001a), Inverse modeling of the hydrological and the hydrochemical behavior of hydrosystems: Application to nitrate transport and denitrification, *Water Resour. Res.*, *37*(8), 2179–2190.

Pinault, J.-L., V. Plagnes, L. Aquilina, and M. Bakalowicz (2001b), Inverse modeling of the hydrological and the hydrochemical behavior of hydrosystems: Characterization of karst system functioning, *Water Resour. Res.*, *37*(8), 2191–2204.

Pinault, J.-L., N. Doerfliger, B. Ladouche, and M. Bakalowicz (2004), Characterizing a coastal karst aquifer using an inverse modeling approach: The saline springs of Thau, southern France, *Water Resour. Res.*, *40*, W08501, doi:10.1029/2003WR002553.

Pinault, J.-L., N. Amraoui, and C. Golaz (2005), Groundwater-induced flooding in macropore-dominated hydrological system in the context of climate changes, *Water Resour. Res.*, *41*, W05001, doi:10.1029/2004WR003169.

Rivrain, J. C. (1998), Les épisodes orageux à précipitations extrêmes dans les régions méditerranéennes du sud de la France (Storm-floods in the Mediterranean region of southern France), *Phénomènes remarquables*, vol. 4, 93 pp., Serv. cent. d'exploitation de la meteorol. (SCEM), Météo France, Toulouse, France.

Robinson, J. S., and M. Sivapalan (1997a), An investigation into the physical causes of scaling and heterogeneity of regional flood frequency, *Water Resour. Res.*, *33*, 1045–1059.

Robinson, J. S., and M. Sivapalan (1997b), Temporal scales and hydrological regimes: Implications for flood frequency scaling, *Water Resour. Res.*, *33*, 2981–2999.

Rodriguez-Iturbe, I., D. R. Cox, and V. Isham (1987), Some models for rainfall based on stochastic point processes, *Proc. R. Soc., Ser. A*, *410*, 269–288.

Rodriguez-Iturbe, I., D. R. Cox, and V. Isham (1988), A point process model for rainfall: Further developments, *Proc. R. Soc., Ser. A*, *417*, 283–298.

Velghe, T., P. Troch, F. de Troch, and J. Van de Velde (1994), Evaluation of cluster-based rectangular pulses point process models for rainfall, *Water Resour. Res.*, *30*, 2847–2857.

Verhoest, N., P. Troch, and F. de Troch (1997), On the applicability of Bartlett-Lewis rectangular pulses models for calculating design storms at a point, *J. Hydrol.*, *202*, 108–120.

Wheater, H. S., T. J. Jolley, C. Onof, N. Mackay, and R. E. Chandler (1999), Analysis of aggregation and disaggregation effects for grid-based hydrological models and the development of improved precipitation disaggregation procedures for GCMs, *Hydrol. Earth Syst. Sci.*, *3*, 95–108.

Wheater, H. S., R. E. Chandler, C. J. Onof, V. S. Isham, E. Bellone, C. Yang, D. Lekkas, G. Lourmas, and M.-L. Segond (2005), Spatial-temporal rainfall modelling for flood risk estimation, *Stoch. Environ. Res. Risk Assess.*, *19*, 403–416.

Wheater, H., et al. (2006), Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM, *RD Tech. Rep. F2105/TR*, Imp. Coll. of London and Univ. Coll. of London, London.

Wobrock, W., A. I. Flossmann, and R. D. Farley (2003), Comparison of observed and modelled hailstone spectra during a severe storm over the northern Pyrenean foothills, *Atmos. Res.*, *67–68*(1–4), 685–703.

D. Allier, Water Department, BRGM, 3 avenue Claude-Guillemin, BP 36009, F-45060 Orléans Cedex 2, France.

J.-L. Pinault, 96, rue du Port David, 45370 DRY, France.