Hydrology and Earth System Sciences, 4(1), 79–93 (2000) © EGS

Hydrology & Earth System Sciences

Mapping mean monthly runoff pattern using EOF analysis

Eric Sauquet[1], Irina Krasovskaia[2] and Etienne Leblois[3]

[1]Cemagref, Hydrology – Hydraulics Research Unit, 3 bis quai Chauveau CP220, F-69336 Lyon cedex 09, France
[2]Hydroconsult AB, Lergravsvägen 33, SE-26352 Höganäs, Sweden
[3]Cemagref, Hydrology – Hydraulics Research Unit, 3 bis quai Chauveau CP220, F-69336 Lyon cedex 09, France
e-mail for corresponding author: sauquet@lyon.cemagref.fr

Abstract

Interpolation of flow patterns, as defined by the variation over the year of the twelve mean monthly discharges, should account both for the auto-correlation in time series and consider the river network, as discharge records are relative to the outlet of the upstream basin. An objective approach of flow pattern mapping along the river implemented within a Geographical Information System (GIS) is presented. Preprocessing of the digital elevation model enables extraction of the main river network (water paths) and basin descriptors, and offers a convenient background for the interpolation scheme. Monthly values are standardised to overcome scale dependence with basin size. Expansion into Empirical Orthogonal Function (EOF) allows interpreting each standardised time series as a linear combination of amplitude functions defined at the regional scale with the weights that need to be interpolated. Empirical dependencies relating these weights to basin characteristics are used for this purpose. Finally, kriging is applied for interpolating the residuals of the weights. The approach has been used for interpolation along a river net of flow patterns of an extensive monthly runoff sample for south-eastern France and has been validated successfully on a sample of eleven stations. Predicted flow patterns have been also introduced into a simple regime classification scheme to derive a map of flow regime in a formal and objective manner.

Introduction

Mapping runoff characteristics is one of the basic tasks in hydrology. Map representation of spatial and temporal variability of water resources has manifold application fields. A broad variety of professionals from environmentalists and planners to lawyers and politicians requires summarised information about spatial and temporal distribution of water resources presented in a map form. Combined with precipitation maps, the runoff maps may be used for determining water budgets at scales from individual basins to whole continents, and in this way contribute to an understanding of regional and global water fluxes. Maps of runoff characteristics provide an important validation tool for the output produced by atmospheric circulation models and macro-scale hydrological models.

In addition to the annual variation of hydrological characteristics in time and space, a description of their seasonal variation is important for many applications. Flow regime maps offer an insight into seasonal fluctuations of flow observed at a grouping basins with similar runoff formation mechanism. While the mean monthly runoff values were commonly used for description of temporal runoff variation, the techniques for representing variation in space are quite numerous.

The simplest way to present the seasonal behaviour of runoff on the flow regime map is to show the monthly hydrographs at the location of a set of gauging stations (Korzun, 1978, Arnell *et al.*, 1993b). Because no interpolation scheme is considered, such a map actually describes the monthly variability of runoff at a point but is often interpreted as a description for a region.

Information on climate and physiography has been used for interpolating runoff information in space. On the basis of topographic, geological and vegetative descriptors and local climatic conditions, Gottschalk *et al.* (1979) created flow regime maps outlining representative areas for each gauging station within a physiographic region. The resulting map shows the flow regime type determined according to the flow regime classification adopted.

The methodology proposed by Weingartner and Aschwanden (1992) allocates at each point, considered as an outlet, the flow pattern of the most similar representative basin (i.e. one that demonstrates strong physiographic and climatic similarities) and then assigns a regime type among 16 referenced ones. The regime map is restricted to the main river network defined by the outlets of basins larger than 10 km^2. The procedure contains subjective elements and is difficult to adapt for automated mapping.

Since the 1960s, different square grid techniques have been used by hydrologists to map water balance elements. Krasovskaia and Gottschalk (1992) presented a 0.5°

resolution grid map of Scandinavian river flow regimes, bringing thus further the approach of Gottschalk *et al.* (1979). A grid cell containing an observation station as well as its representative area was assigned to the flow regime of this station. A similar procedure has been used by Krasovskaia *et al.* (1994) to map river flow regimes across Northern and North-western Europe. Each cell has been assigned to the flow regime type characteristic of the majority of the stations within it. Arnell *et al.* (1993a) averaged the ratios of monthly discharges to the annual runoff of all the basins within each $2° \times 2°$ grid-cell and allocated the representative hydrographs to that cell. The lack of observation stations within cells hindered the creation of the maps in the two studies referred to above.

Mapping river flow regimes is a complex problem with three central topics *viz.* classification, interpolation and stability of river flow regimes (i.e. the regularity of the seasonal flow patterns from year to year). This study will address the problem of objective interpolation of mean monthly runoff values, in order to develop accurate procedures to produce maps avoiding expertise and subjective considerations. A traditional classification scheme is used as a reference when presenting the results of interpolation. The last remaining topic is open for further studies.

One straightforward way to estimate regimes at ungauged locations is, firstly, to interpolate the twelve monthly discharges independently and then to assign a flow regime type at each point of interest according to a chosen classification. However, such an approach neglects the natural intrinsic coherence of flow regimes and the dependence between consecutive monthly discharges imposed by the water cycle (groundwater storage, flow routing, etc.).

To overcome the bias introduced by the size of the basin, runoff is generally converted to water depths on a runoff map to eliminate scale effect within the data set. This conversion has been thought to provide a homogeneous sample for use in mapping techniques (Rochelle *et al.*, 1988, Bishop and Church, 1992). Unfortunately, such a conversion ignores the spatial heterogeneity of hydrological processes: the respective contribution of the components of the water balance depends on the size of the basin (Leblois and Sauquet, 1997). In addition, the river network which propagates runoff and discharge measured at a point also contains upstream information. Thus, an interpolation procedure should investigate the relationships between the basin characteristics and the topological river patterns and consider variables free from an effect of scale.

Mean monthly runoff values are the starting point for many flow regime classification schemes. Here, an interpolation technique for flow patterns in space that avoids some of the problems described above will be presented. The procedure suggested utilises the EOF-expansion, keeping the month-to-month dependence within the flow pattern, and empirical formulae, linking the spatial features identified by the EOF analysis to physiographic descriptors of a basin, in combination with kriging.

Data used

THE STUDY AREA

The study area covers the south-eastern part of France: the Rhône River and its tributaries from the outlet of Lake Geneva to its mouth and all the Mediterranean coastal rivers including those in Corsica (Fig. 1). It stretches from lowlands under Atlantic influences (Saône basin in the northern part of the study area) to high mountains (French Alps). The topographic information used in this study is a subset of the Global 30 Arc Second Elevation Data Set of the U.S. Geological Survey, available on the Internet. Some digitised drainage boundaries of major basins were available (Anonymous, 1996). The land-use, extracted from the CORINE Land Cover data base (IFEN, 1996), was generalised to 1×1 km cells according to the dominant land-use class.

STREAMFLOW DATABASE

The French hydrological database HYDRO offers series of observed daily discharges and reconstructed natural monthly discharges. To ensure temporal consistency, a common observation period of continuous records for over eight years was used (01/01/1977 to 31/12/1984). This eight-year period of observations has been selected as a compromise between a requirement for an observation network dense enough to identify correctly the spatial structure when interpolating, and a time-period long enough to provide reliable estimates of the mean monthly discharges. The selected time series have minimal human influence and represent a wide range of environmental conditions. The total data set consists of 212 gauged basins with drainage areas ranging from 50 to 96 500 km^2. Mean annual runoff varies from 82 to 2458 mm with an average of 761 mm per year. The lowest values are found in the southern part of the study area whereas the highest ones originate in the mountains.

For validation purposes, eleven stations were withdrawn from the data set before mapping. They are representative of the main hydrological flow regimes characteristic for the area under study. The other 201 gauging stations define the reference network.

Interpolation of mean monthly runoff patterns

The suggested approach for interpolation of mean monthly runoff patterns consists of two steps. Firstly, the spatial correlation structure in the data is summarised by EOF analysis. Secondly, regression dependencies between this structure and the physiographic descriptors of the basin are identified. Below there follows a brief presentation of the background to these steps.

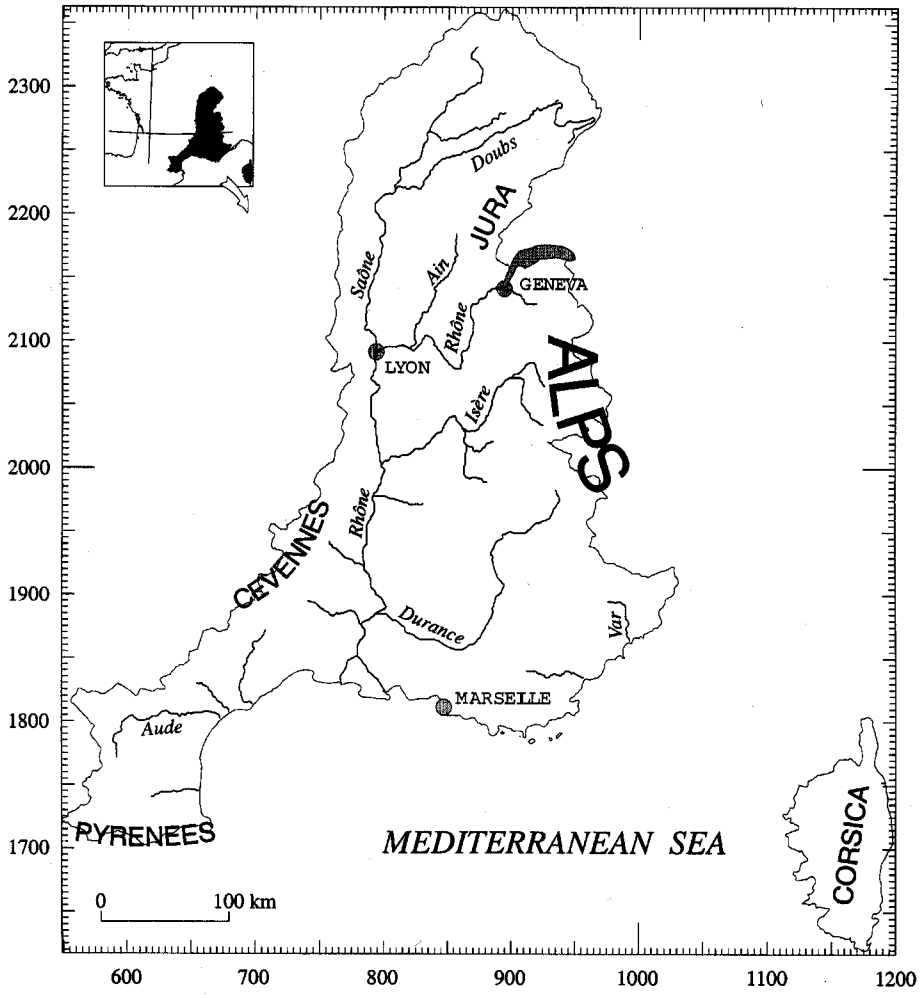

Fig. 1. *The study area.*

ANALYSIS BY MEANS OF EMPIRICAL ORTHOGONAL FUNCTIONS (EOF)

Empirical orthogonal functions (Karhunen-Loeve expansion) is a well-established method in hydrology and has been used for many purposes (Stidd, 1967; Bartlein, 1982; Gottschalk, 1985; Braud and Obled, 1991; Rao and Hsieh; 1991; Hisdal and Tveito, 1991, 1993; Krasovskaia and Gottschalk, 1995; Krasovskaia *et al.*, 1999). The general purpose of using the EOF is to carry out a linear transformation of the original data producing a new orthogonal set of functions, which simplifies and helps by excluding redundant information. The method is presented only briefly here. A detailed description of the method can be found in one of the earliest publications on this topic by Holmström (1963) or in a publication by Obled and Creutin (1986).

Consider a data set of M time series $F(\mathbf{x}_i,t)$, $i = 1,\ldots,M$, which corresponds to a set of values measured at station i with coordinate $\mathbf{x}_i$ at time t, $t = 1,\ldots N$. $f(\mathbf{x}_i,t)$ are the corresponding standardised time series with respect to the time averages μ_F and standard deviations σ_F. An expansion into EOF has the form:

$$f(\mathbf{x}_i,t) = \sum_{j=1}^{M} S_j(\mathbf{x}_i)\beta_j(t) \quad i = 1,\ldots,M \tag{1}$$

where $\beta_j(t)$ are time dependent amplitude functions and $S_j(\mathbf{x}_i)$ are the weight coefficients, which vary between the series (and thus spatial location) but are constant in time.

Requiring fastest possible convergence of the series expansion and normalising the weight coefficients orthogonal sets of weight coefficients and amplitude functions are obtained with the properties:

$$\sum_{i=1}^{M} S_n(\mathbf{x}_i)S_m(\mathbf{x}_i) = \delta_{nm} \tag{2}$$

and

$$\frac{1}{N}\sum_{t=1}^{N} \beta_n(t)\beta_m(t) = \delta_{nm}\lambda_n \tag{3}$$

where δ_{nm} is the Kronecker delta and λ_n are the eigenvalues of the correlation matrix of the original series $f(\mathbf{x}_i,t)$. The weight coefficients, $S_j(\mathbf{x}_i)$, are the elements of the eigenvectors of this correlation matrix.

The set of functions obtained is empirical in the sense that they are based only on the initial series themselves and not on any predetermined mathematical form. The eigenvalues are usually arranged in a descending order according to the proportion of the variance explained by each of them. Usually a small number of functions reproduce a great part of the total variance. The series expansion can accordingly be truncated at number $M^* < M$ of terms, depending on a requested level of acceptable explained variance:

$$f^*(\mathbf{x}_i,t) = \sum_{j=1}^{M^*} S_j(\mathbf{x}_i)\beta_j(t) \qquad (4)$$

In combination with interpolation techniques, the above approach was used, for example, by Hisdal and Tveito (1991) to generate hydrological series at ungauged locations. Indeed, regionalising the M^* spatial components, as well as the temporal mean μ_F^* and the standard deviation σ_F^*, it is possible to reconstruct, for any arbitrary point at coordinate $\mathbf{x}$, the standardised time series in a first step from Eqn. (4) and finally to compute the actual series as:

$$F^*(\mathbf{x},t) = \mu_F^*(\mathbf{x}) + \sigma_F^*(\mathbf{x})f^*(\mathbf{x},t) \qquad (5)$$

This approach can be generalised at each point within the study area to depict the spatial pattern of F.

INTERPOLATION OF THE SPATIAL COMPONENTS

The weight coefficients $S_j(\mathbf{x})$ are by definition of EOF deterministic functions in space. Therefore, interpolation of these spatial components by kriging techniques (Matheron, 1971) is not appropriate. The method of georegression (Faulkner and Prudhomme, 1998) was considered instead. An empirical formula F_j is fitted to give the best estimation $S_j^*(\mathbf{x})$ of each spatial component $S_j(\mathbf{x})$, $j = 1,\ldots,M^*$:

$$\begin{cases} S_j^*(\mathbf{x}) = F_j(X_1, X_2, \ldots, X_k) \\ \varepsilon_j(\mathbf{x}) = S_j(\mathbf{x}) - S_j^*(\mathbf{x}) \end{cases}, \; j = 1,\ldots,M^* \qquad (6)$$

where X_i, $i = 1,\ldots,k$ are basin characteristics and ε_j, $j = 1,\ldots,M^*$ the residuals. Including integrated descriptors in the interpolation scheme is a simple and reliable way to account for the fact that streamflow data are related to the outlet of a river system and are the outcome of the processes operating over the whole basin. The choice of basin characteristics yields information to be propagated in a preferential direction (i.e. along the river in an upstream-downstream direction).

The residuals ε_j, $j = 1,\ldots,M^*$ of the fitted formulae represent local correction factors and can be assumed to be a realisation of a second-order stationarity random field. This assumption allows interpolating the residuals by ordinary kriging.

A map of $S_j^*(\mathbf{x})$ is then derived by applying the optimal linear model and using GIS tools. Thereafter, this map is combined with the map of the interpolated residuals to obtain the spatial pattern of the component $S_j(\mathbf{x})$.

Application to the French part of the Rhone Basin and the Mediterranean coastal rivers

The aim of this study is to interpolate the flow patterns along rivers, providing a basis for flow regime identification of the latter. The variables of interest are thus the twelve standardised mean monthly values (i.e. mean discharge value minus time average divided by the standard deviation). The use of standardised variables resolves scale dependence.

SELECTION OF THE SIGNIFICANT AMPLITUDE FUNCTIONS

The EOF analysis was applied to the mean monthly discharge series of 201 gauging stations.

The first amplitude function $\beta_1(t)$ (Fig. 2) shows the most frequently observed pattern and accounts for 65% of the explained variance. The highest values occur in winter and early spring and the lowest values in late summer, β_1 resembles temperate maritime regime. The second amplitude function (in terms of the variance explained) demonstrates the characteristics of a snowmelt-fed regime with the first main peak in May–June and a secondary one in autumn. The following amplitude functions reveal deviations from the temporal patterns above. They account numerically for the remaining variance within the dataset and attempting a hydrological interpretation of such additional amplitude functions would be spurious. To evaluate the influence of the number of amplitude functions used in constitution of each standardised monthly discharge, twelve Nash-Sutcliffe (1970) efficiency coefficients (E_t^0, $t = 1,\ldots,12$) shown in Table 1, were calculated for the reference network of 201 gauged basins. Each efficiency coefficient is computed as:

$$E_t^0 = 1 - \frac{\sum_{i=1}^{M}(f^*(i,t) - f(i,t))^2}{\sum_{i=1}^{M}(\bar{f}(t) - f(i,t))^2}, t = 1,\ldots,N \qquad (7)$$

where $\bar{f}(t)$ is the average of the observed standardised mean monthly value of the month t.

The two first amplitude functions, depicting high and low water periods that occur in summer or in winter within the study area, yield values of $E_t^0 > 0.70$. The first three

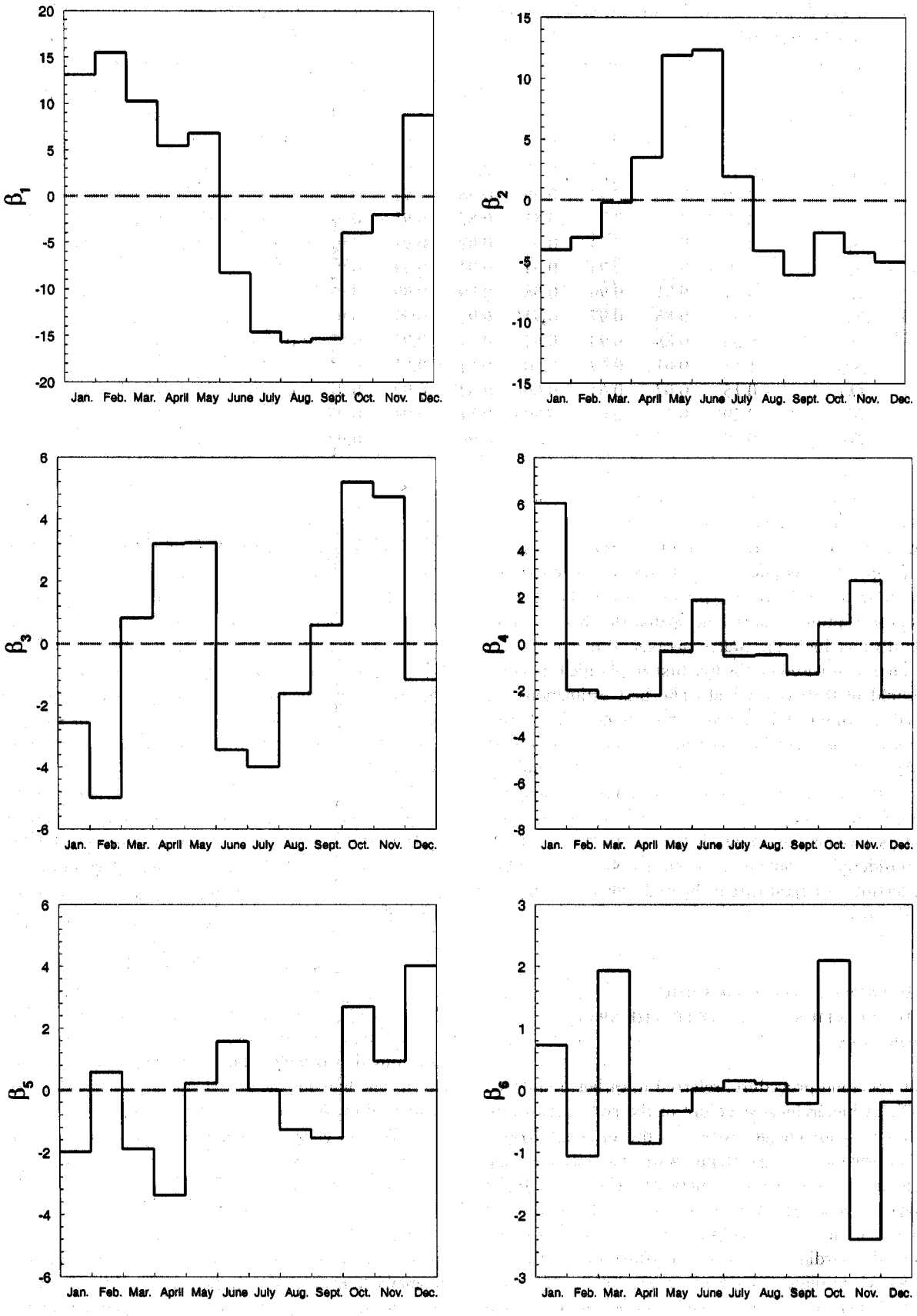

Fig. 2. *Six first amplitude functions deduced from the EOF analysis ranged in descending order according to their contribution to the explained variance.*

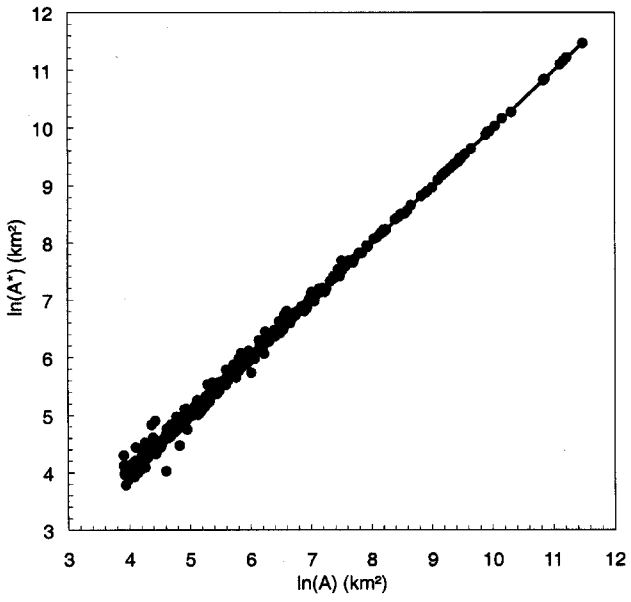

Fig. 3. *Regression of the logarithm of estimated area (km^2), ln(A*), vs. logarithm of actual measured area ln(A).*

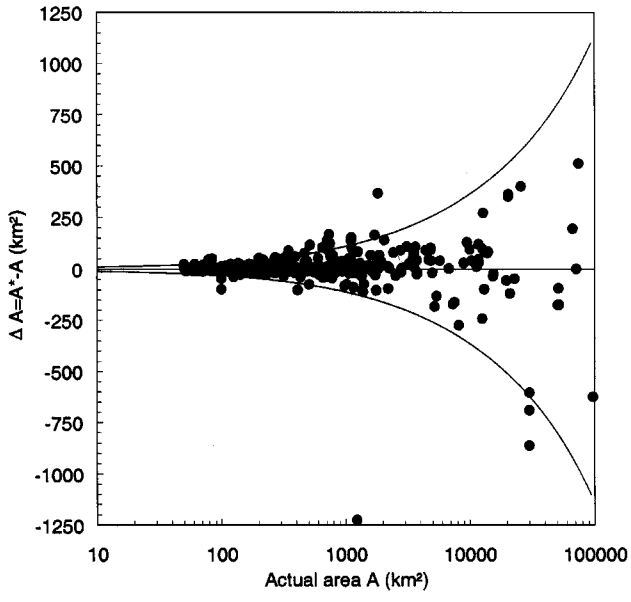

Fig. 4. *Differences between estimated and actual basin area ΔA vs. actual basin area A, enveloped by the curves delineating the interval defined by the tolerance criterion.*

assumed to be correctly reproduced if:

$$|\Delta A| = |A^* - A| \leq \chi \cdot d_0 \quad \text{with } \chi = 2\sqrt{A\pi} \qquad (9)$$

where χ is the perimeter of the round basin.

This criterion takes fully into account the grid structure of the DEM and the size of the basins to be identified, being thus more relevant than a subjective tolerance threshold. Only 19 basins violated the constraint imposed by the criterion (Fig. 4). Most of the significant deviations are found in chalky sectors, where the topographic watershed differs from the hydrological one. The results of calculations using Eqn. (9) show that an error of 50% can be observed for basins of about 50-km^2 area. This result makes it evident that the DEM resolution of 1×1 km is insufficient to outline the boundaries of smaller basins accurately.

Since a flow regime is related to a river, a channel network should support the mapping procedures. This is achieved by introducing a minimum accumulation area necessary to support a channel. This threshold is fixed here to 50 km^2 in accordance with the constraint imposed by the resolution of the DEM. Cells with an accumulated area of more than 50 km^2 define the main river network. The 212 stations used as basis for the mapping procedure and later for the validation are located along this main river network (Fig. 5).

In accordance with the threshold imposed by the resolution of the DEM, calculations are restricted to grid cells located along the main river network.

REGIONALISATION OF THE SPATIAL COMPONENTS (WEIGHT COEFFICIENTS)

To interpolate the weight coefficients using georegression, a number of physiographic characteristics were computed along the main river network. The following physiographic descriptors were derived combining the drainage network, the 1-km DEM and the CORINE Land Cover data base (IFEN, 1996):

- *POW*, the *P*roportion of basin area as *O*pen *W*ater (%),
- *PF*, the *P*roportion of basin area as *F*orest (%),
- *PSG*, the *P*roportion of basin area as *S*now or *G*lacier (%),
- *PIS*, the *P*ercentage of area with *I*mpervious *S*oil (%),
- *PCL*, the *P*ercentage of *C*ultivated *L*and (%),
- *PIL*, the *P*ercentage of potentially *I*rrigated *L*and (%),
- several characteristics Zp (p = 05, 10, 15, 20, 30, 40, 50, 70, 90) of the hypsometric curve (p refers to the percentage of the area with elevation above Zp) (m),
- *I*, an *I*ndex of catchment slope derived from the hypsometric curve (Roche, 1963) (m/km),
- *DMS*, the shortest *D*istance to the *M*editerranean *S*ea from the centroid of the basin (km),
- *DS*, the Distance minimal to the sea (between the Mediterranean Sea and the Atlantic Ocean) from the centroid (km),
- *DMB*, the *D*istance from the centroid of the basin to *M*ont-*B*lanc (km),
- *DMF*, the *D*istance from the centroid of the basin to *M*ont-*F*inniels, the highest point of the Cevennes mountains (km),
- *XG* and *YG*, the coordinates of the centroid of the basin in the French projection Lambert II (km).

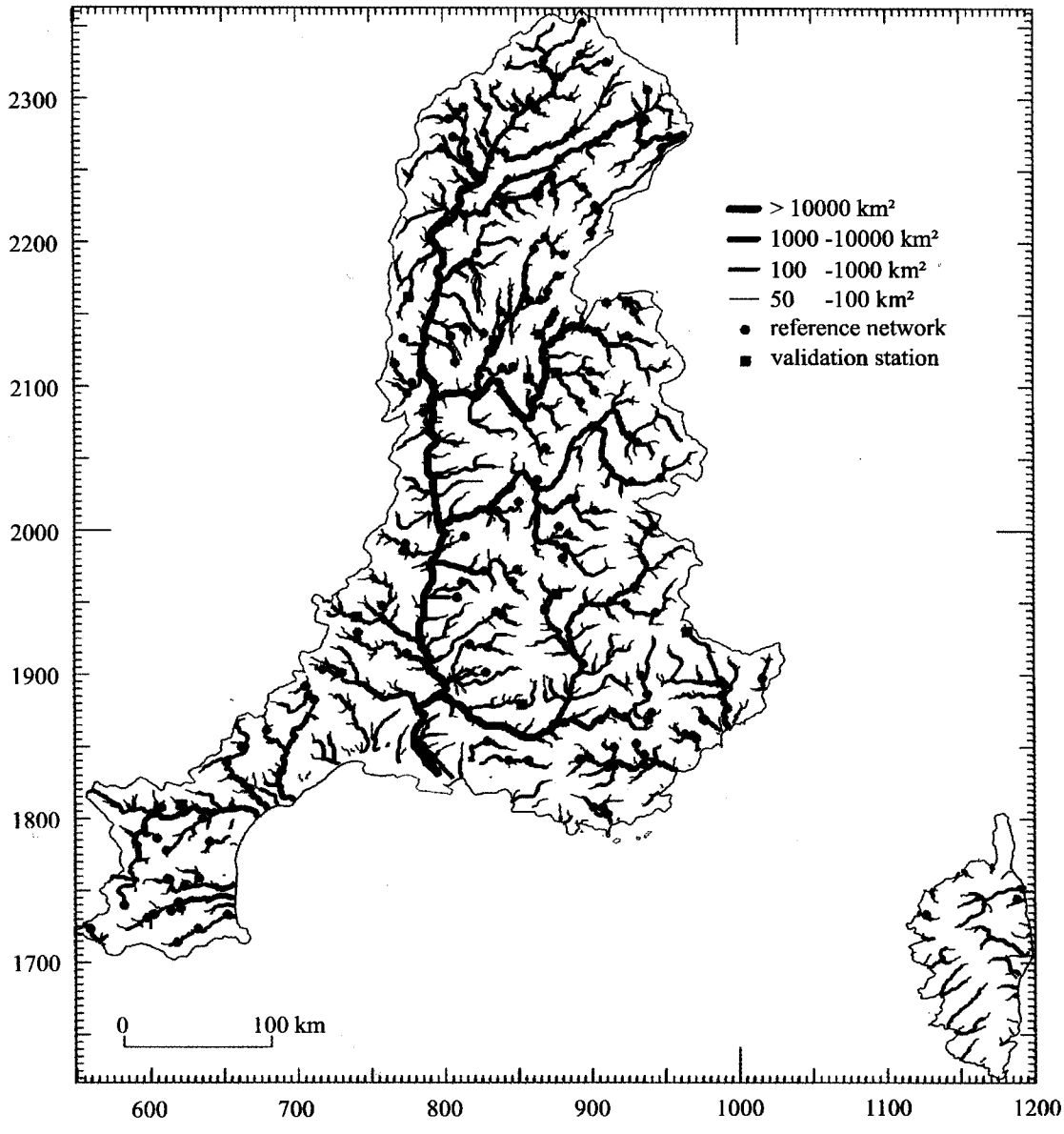

Fig. 5. *Main river network derived from DEM with 1 × 1 km resolution and the location of the gauging stations.*

The distances can be related to climate factors in the study area (e.g. *DS* measures the maritime influence, small *DMF* values are for the Cevennes-Vivarais region where conflicting atmospheric influences result in heavy rainfalls in autumn). To identify the most important basin characteristics for this particular application, a map of each of the weight coefficients $S_j(\mathbf{x})$, $j = 1,...5$ was drawn. Then, empirical formulae linking the spatial components to the characteristics of the basins were derived using stepwise multiple regression analysis. This procedure was applied to all the spatial components except $S_1(\mathbf{x})$ and $S_2(\mathbf{x})$; already a prior interpolation of these two weight coefficients showed an obvious relationship to the elevation of the gauged basin. A description of the regionalisation of the weight coefficient $S_1(\mathbf{x})$ below may illustrate the procedure.

Negative values of $S_1(\mathbf{x})$ were found in the headwaters of mountainous rivers whereas elsewhere $S_1(\mathbf{x})$ value was very close to 0.08. Hence, a relationship with descriptors extracted from the hypsometric curve could be expected. The best relationship was established using the variable Z15 (i.e. the elevation exceeded on over 15% of the basin area). As can be seen in Fig. 6, showing the relationship, a plateau is observed up to a break at 1400 m a.s.l.; then, $S_1(\mathbf{x})$ drops sharply. The dependence seems to be sound in a physical sense: the negative weight coefficient values are found in mountain areas. The slope of the linear part is fitted by least squares and the coefficient of determination is 0.95.

The experimental semi-variogram of residuals $S_1(\mathbf{x}) - S_1^*(\mathbf{x})$, shown in Fig. 7, supports the evidence of the existence of a spatial structure, and an exponential

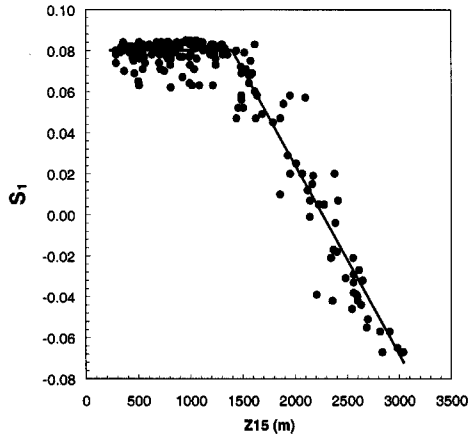

Fig. 6. *Relationship between the spatial component S_1 and the variable Z15 obtained from the hypsometric curve.*

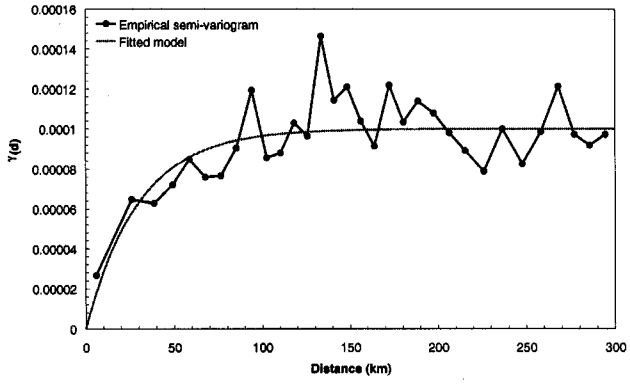

Fig. 7. *Empirical and fitted semi-variogram of residuals S_1-S_1*.*

model could be fitted for all the residuals:

$$\gamma(d) = \gamma_0(1 - \exp(-d/h)) \qquad (10)$$

The values of the practical range (i.e. $3h$) and the sill γ_0 are given in Table 2. The fitted sill is interpretable as variance of the residuals. Figure 8 illustrates the final map of $S_1(x)$, which is a combination of the map of residuals and the linear model application.

Table 2. Parameters of the exponential model fitted to the empirical semi-variogram of spatial components

Residual	Practical range (km)	Sill
S_1-S_1*	90	0.0001
S_2-S_2*	45	0.0006
S_3-S_3*	90	0.0018
S_4-S_4*	70	0.0015
S_5-S_5*	60	0.0023

Table 3 presents the empirical formulas used in interpolating the spatial components (weight coefficients). All the coefficients of the regression models are significant at the 5% level according to the Fisher-Snedecor test. All the linear models contain at least one characteristic of the hypsometric curve, which helps to differentiate the snowmelt-fed flow in the mountains and rainfall-fed flow of lowlands.

VALIDATION USING AN EFFICIENCY CRITERION

The observed and predicted mean monthly flow patterns for the 11 gauged basins, initially removed from the data set, were compared first. Figure 9 shows the observed and estimated standardised monthly discharges for six of these basins.

This validation is both temporal and spatial. In general, the estimated standardised monthly discharge is in good agreement with the observed patterns, especially for high and low flows. The Nash-Sutcliffe efficiency coefficient calculated over the 12 months ranged from 0.70 to 0.98 (Table 4). However, slight fluctuations in spring and/or autumn seem not to be captured by the EOF analysis. Noticeable differences between the observed and estimated patterns were found for April-May and October-November for the Séran River at Bavosière (station No V1414010). For this basin, however, inaccuracy in the linear approximation was not the only reason. Karst aquifers affect the natural flows of the Séran River, and neither the topographic descriptors nor the land-use variables used in the georegression reflect this geological peculiarity.

The Nash-Sutcliffe efficiency coefficients (E_t, t = 1,...,12) defined previously in Eqn. (9) were calculated also for each month for the 11 validation series. The coefficient values vary from −0.03 to 0.93. The results for September (−0.03) and April (0.29) are poor, but estimates tend to be better in summer and winter. On the total, the standardised mean monthly discharges are reproduced correctly for seven months ($E_t > 0.80$). When the series for the Séran River at Bavosière is removed, the value of the Nash-Sutcliffe criterion increases to 0.72 for April but for September it remains almost unchanged (0.03).

Based on the results above, it can be concluded that, overall, the interpolation technique gives reliable standardised mean monthly values, except for the basins where a specific geological environment is important in flow genesis.

VALIDATION AGAINST REGIME PATTERNS

A variety of river flow regime classifications is referenced in the hydrological literature, starting from the "classical" manual ones like Pardé's (1955), Zaikov's (1944) and Lvovich's (1973), to the totally computerised one by Krasovskaia and Gottschalk (1992) and Krasovskaia *et al.* (1994), all based on the origin of the river flows and the

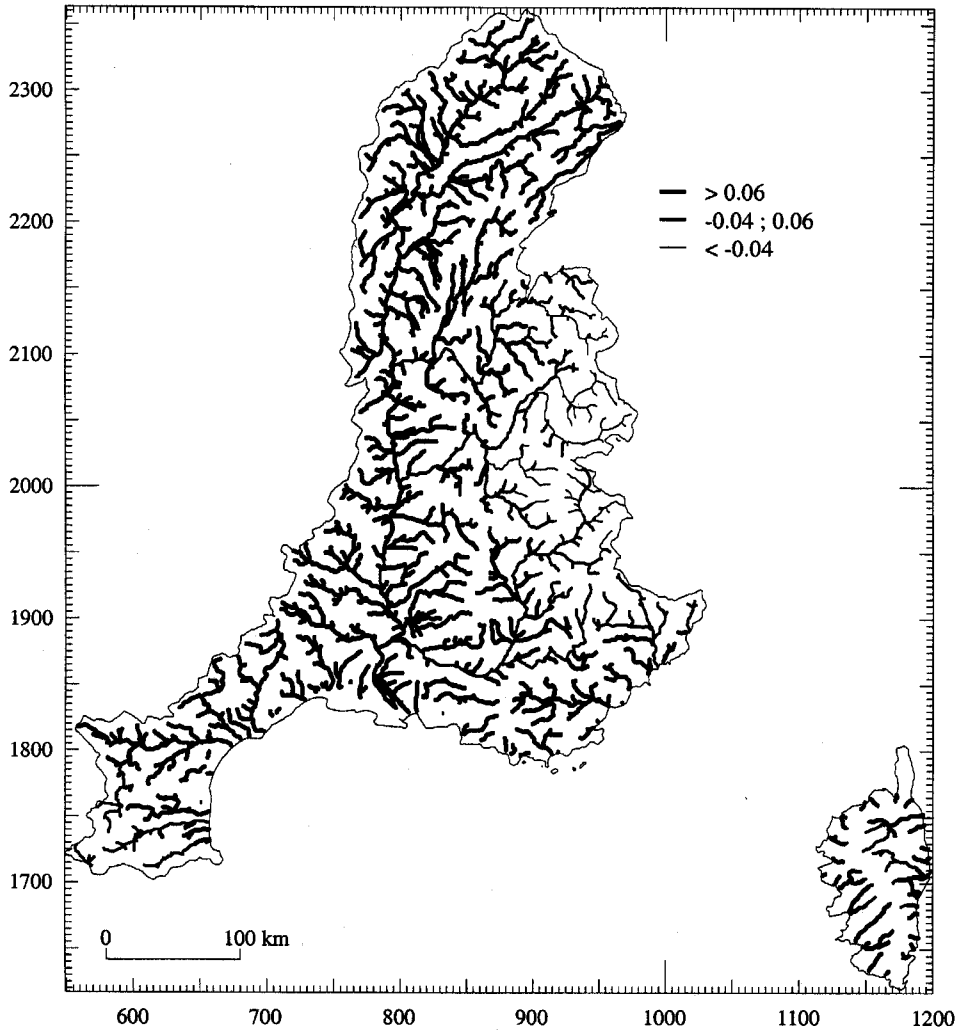

Fig. 8. *Map of the first spatial component S_1.*

Table 3. Equations for interpolating spatial components (weight coefficients).

Spatial components	Empirical formula	Determination coefficient (%)
S_1	$S_1^* = 0.08$ if Z15 < 1400 m $S_1^* = 0.08\text{-}0.277\ (Z15\text{-}1400)/3000$ if Z15 ≥ 1400 m	95
S_2	$S_2^* = -0.613\ (Z15/5000) + 0.445$ if Z15 ≥ 2400 m $S_2^* = 0.447\ (Z15/5000) - 0.067$ if Z15 < 2400 m	82
S_3	$S_3^* = 1.456\ (Z05/5000)^2 - 2.716\ (Z05/5000)^3 + 0.092\ (I/10)\text{-}0.033\ (DMF/100) + 0.017$	67
S_4	$S_4^* = -1.204\ (Z30/5000) - 2.023\ (Z30/5000)^2 + 0.014\ (DMB/100) - 0.031\ (DS/100) + 0.163$	69
S_5	$S_5^* = 0.607\ (Z20/5000)^2 + 0.707\ (DMB/100)^2 - 0.081\ (DMF/100) + 0.518\ (YG/1000) + 0.659\ (XG/1000) + 0.152\ (PCL/100) - 1.632$	50

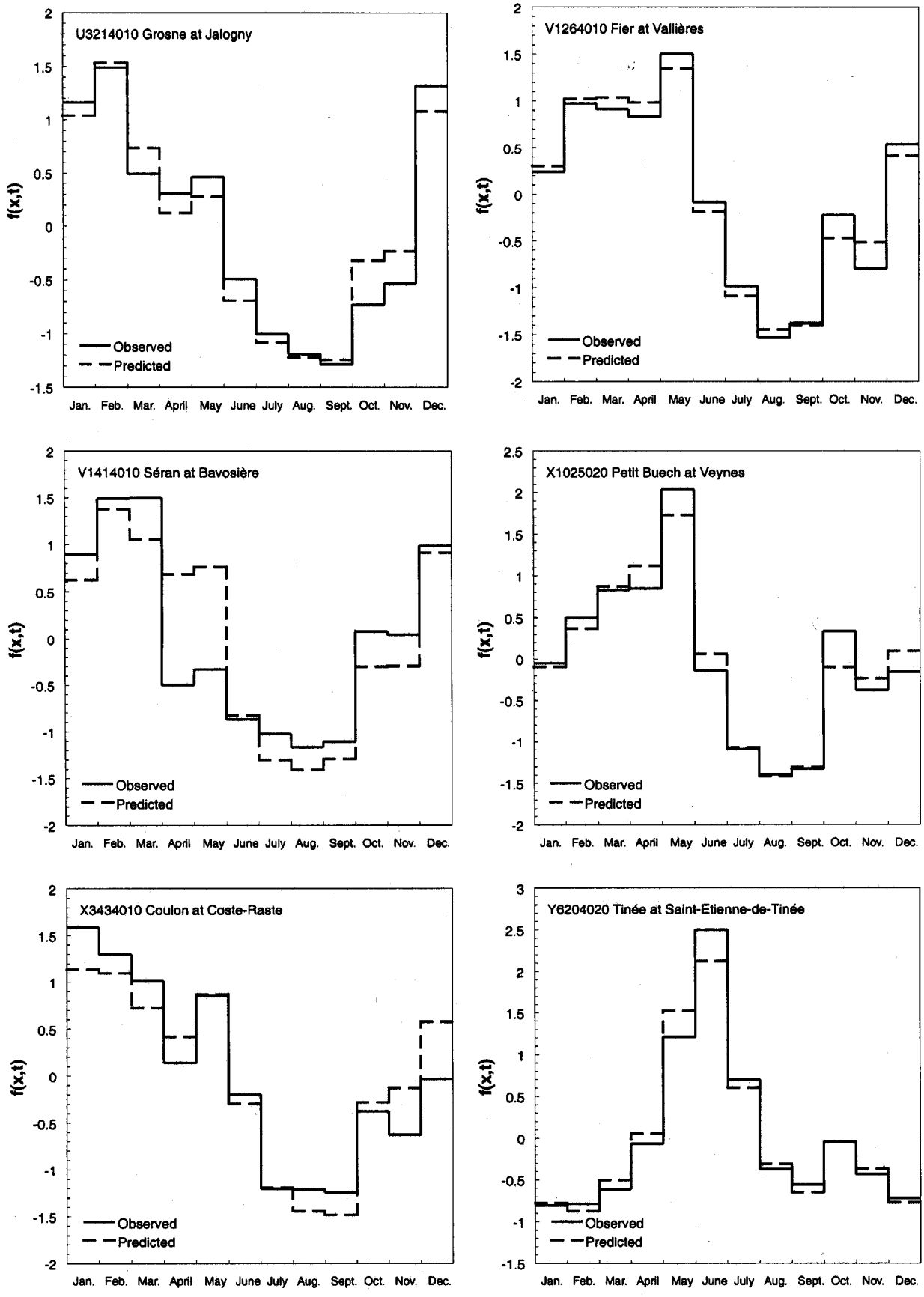

Fig. 9. *Observed and calculated standardised regimes for six of the validation series.*

Table 4. Nash-Sutcliffe criterion for the predicted standardised monthly discharge of the validation stations

Station No	X (m)	Y (m)	Area (km^2)	Nash-Sutcliffe criterion
U3214010	777750	2158620	332	0.95
V0334010	927100	2156200	495	0.89
V1015810	867500	2133630	183	0.95
V1264010	878030	2106490	1350	0.98
V1414010	858880	2103420	158	0.70
V2814020	835980	2128100	317	0.95
V5045020	742007	1937047	507	0.85
X1025020	877420	1953610	318	0.95
X3434010	853450	1877200	333	0.90
Y1435410	623790	1806890	108	0.92
Y6204020	967257	1927492	167	0.97

timing of high and low flows. For example, Pardé (1955) considered three elementary regimes, each one reflecting one dominating genetic source of flow illustrated in Fig. 10:

- glacial regimes dominated by glacial meltwaters with maximum discharge in summer (e.g. the Arveyron at Pont du Bouchet);
- nival regimes characterised by high flows in late spring due to snowmelt (e.g. the Têt at Joncet);
- rainfall-fed regimes with high flows in winter and early spring and low flows in summer due to evaporation (e.g. the Breuchin at La Proiselière).

The glacier/snowmelt-fed regimes (i.e. glacial and nival)

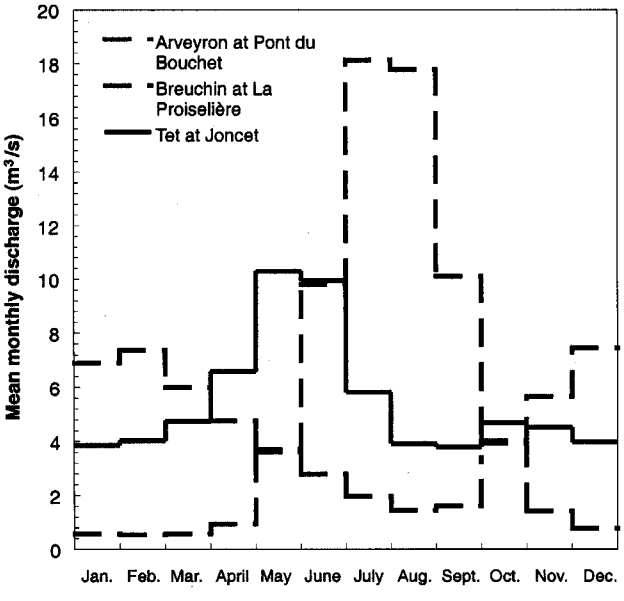

Fig. 10. *Example of flow patterns observed in the study area (the mean monthly discharges are calculated over the 1970–1994 period).*

occur in mountainous areas whereas rainfall-fed patterns concern basins at lower altitudes. Flow regime patterns of the latter are rather complex, i.e. they represent a combination of the three elementary regime types.

Other flow regime classification schemes (Gottschalk, 1985; Haines *et al.*, 1988) are based on some statistical similarity criterion to group long term mean monthly values and then assigning a regime type (from an existent classification scheme) to the groups obtained. Krasovskaia (1997) suggested a formal entropy-based scheme for grouping river flow regimes which considers genetic origin of flow, similarity between the series in the group and stability of seasonal flow.

The length of the series with the common observation period used in the present study is too short to perform reliable flow regime identification according to any scheme. Hence, a well-known classification scheme by Pardé (1955) is used as a check on the consistency of the results, i.e. to compare the interpolated mean monthly variation pattern with those from this scheme. Following Gottschalk (1985), the Euclidian distance is used:

$$d(q_{ref}, \mathbf{x}) = \sqrt{\sum_{t=1}^{12} (f(\mathbf{x}, t) - q_{ref}(t))^2} \qquad (11)$$

where $q_{ref}(t), t = 1,\ldots,12$ is a standardised model of flow regime, as a measure of the similarity to a set of specific standardised regime models. The mean monthly pattern interpolated is then assigned to the type of the closest standardised model of flow regime according to this distance. Since no information about the genetic sources of flow is considered, a cautious selection of the possible regime types avoids attributing, for example, series with high water in mid-summer (glacial regime type) to the tropical regime type.

The patterns of the interpolated mean monthly values were compared with a number of the flow regime patterns from the classical flow regime classification by Pardé (1955), where a detailed description of different types can be found. Figure 11 illustrates the *a priori* selection of regime types for standardised values. The Glacial (A and A′ in Fig. 11) and Nival (B and B′ in Fig. 11) regime types have a common genetic source *viz.* glacier/snowmelt. Both have high water in mid and late summer and spring-early summer, respectively, and low water in winter and differ from each other in the time of occurrence and length of the high water periods. The Pluvial regime type (C and C′ in Fig. 11) is rain-fed, with winter high water and low water in summer. Three other regimes (transition nival, nivo-pluvial and pluvio-nival) are transitional between nival and pluvial patterns.

The similarity criterion (Euclidian distance) described earlier has been used to choose the flow regime pattern closest to the mean monthly pattern interpolated. A smoothing of the preliminary map was performed: channels

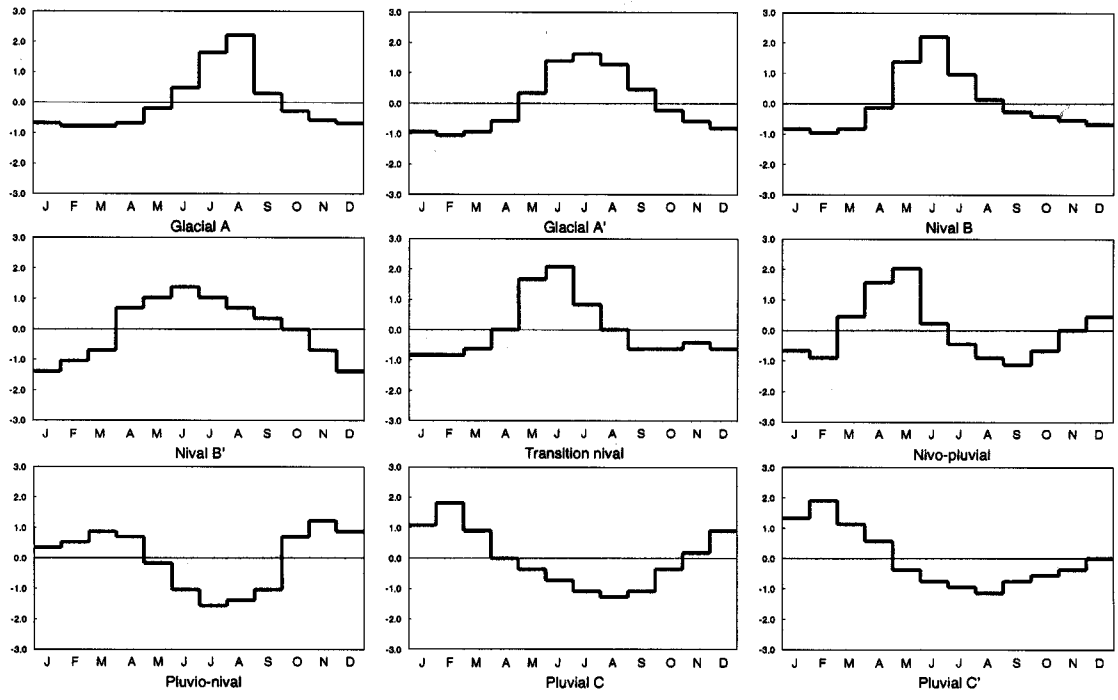

Fig. 11. *Standardised regime models (Pardé, 1955).*

less than 10 km long with the same mean monthly patterns upstream and downstream were attributed the regime type of the neighbouring longer channels.

The resulting map (Fig. 12) demonstrates a sound picture of flow regimes attributed to the interpolated mean monthly flow patterns. The glacier/snowmelt-fed regimes were found in the Alps and the Pyrenees in contrast to the northern part of the study area, where rainfall and evaporation processes govern seasonal flow patterns. The natural hierarchy imposed by glacier/snowmelt processes is respected along the main rivers and demonstrates glacial regimes in headwaters and pluvial ones in alluvial plains (i.e. from A′, B, transition mountainous regimes to C or C′).

The pluvial regime C dominates the Saône River basin, whereas C′ (found in areas with more abundant winter rainfall) is more like a Mediterranean regime. It is noticeable that no channel showed the Nival B′ flow pattern.

Besides the automated classification procedure finds its limits in the Dombes area to the North of Lyon. Some rivers are assigned to the pluvio-nival type. The name does not reflect the origin of the monthly variability: the flow patterns are influenced here by manipulation of water in artificial swamps.

The major rivers span a variety of climatic conditions and can propagate a persistent glacier/snowmelt signal from headwaters, whereas channel parts and tributaries at lower altitudes display pluvial regime (e.g. the Rhône River between Geneva and Lyon). Hence it is difficult to allocate one regime type to a whole river basin stretching over a number of altitude and physiographic zones, unless the river network itself is considered for this purpose.

Conclusions

Maps of river flow regime are a valuable source of information about the average seasonal behaviour of flow important for many applications. The problem of regionalisation of flow regimes has been focused on in this study. A computerised objective procedure for mapping standardised mean monthly runoff has been developed in the framework of a GIS. The technique suggested for interpolation of the standardised mean monthly runoff values is based on the EOF analysis, which permits identification of the temporal (amplitude functions) and spatial (weights) components of the input signal. The weights have then been interpolated applying regression dependencies with a number of basin characteristics (georegression). The residuals of the weights have then been interpolated using kriging techniques. The procedure described was applied successfully to interpolate standardised mean monthly flow patterns for rivers of south-eastern part of France, and a variety of flow regime patterns has been demonstrated.

The interpolation technique yields standardised monthly runoff patterns along the river network. This procedure seems to be sound when flow regime information is needed for areas extending over a number of geographical and altitude zones.

The interpolation scheme includes a formal approach to identify a reference flow regime type (from some chosen

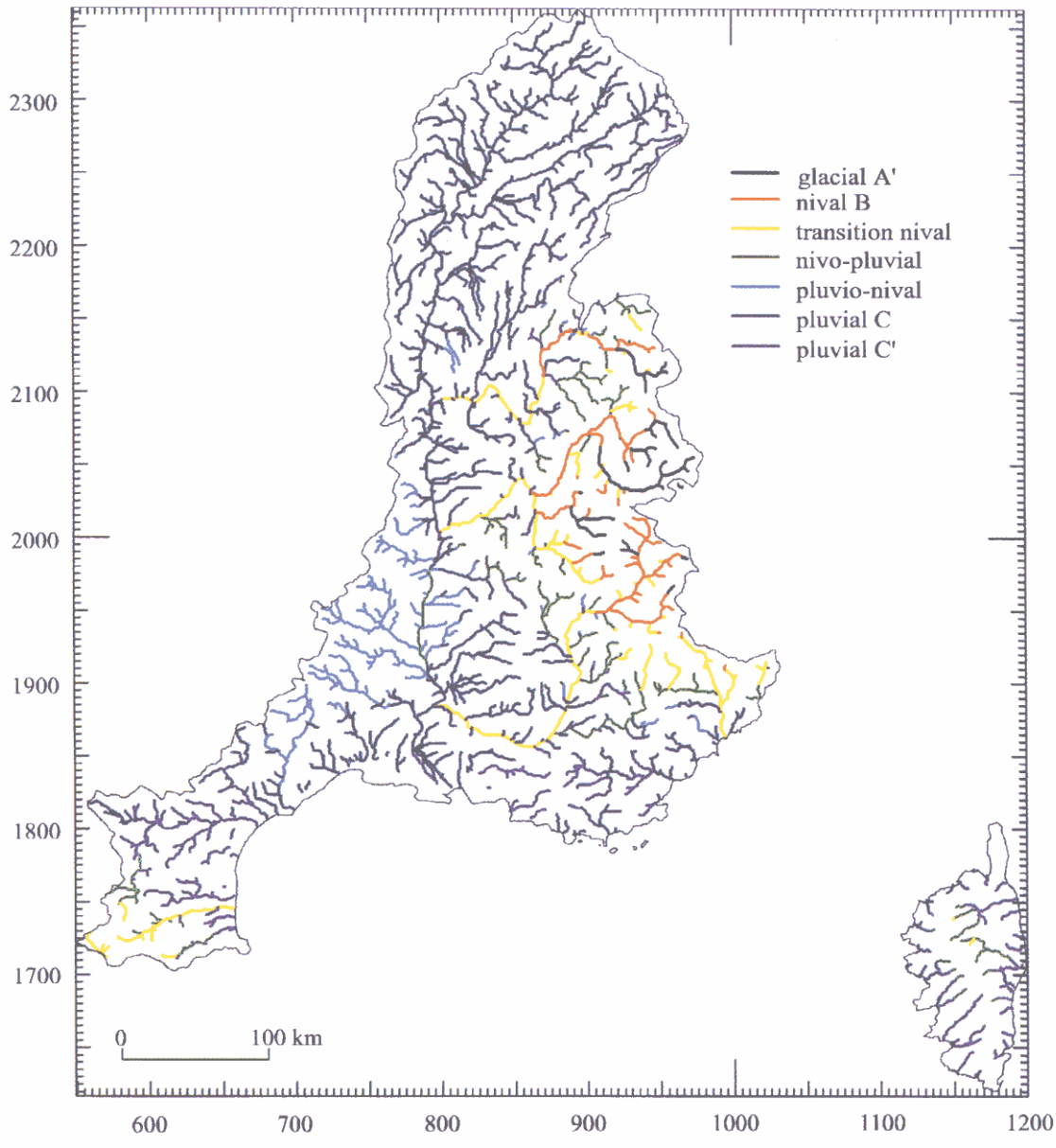

Fig. 12. *Flow regime map along the main river network of the south-eastern part of France including Corsica.*

classification scheme) of the interpolated mean monthly pattern selecting the closest reference type with the help of a similarity criterion (Euclidian distance).

A powerful combination of stochastic and deterministic methods used for computerised interpolation, formal identification of the regime type together with a sound GIS-supported and river net-based method for map presentation of the integrated output, make it a useful tool for regionalisation of river flow regimes.

In this approach, the regularity of the regime type from year to year is ignored. A further development would be to supplement the procedures described by taking the stability of river flow regimes into account and create a corresponding map as measure of reliability of the resultant information.

Acknowledgements

This work is a part of the French project GEWEX-Rhône of hydrological modelling and was supported by the Programme National de Recherche en Hydrologie (project "Regimes", n° 97-23). The authors wish to express their thanks to the French data base HYDRO and the General Engineering Branch (DTG) of Electricité de France for supplying the discharge data used in this study and to Catherine Guérin (CETP) for CORINE preprocessing.

References

Anonymous, 1996. *Le référentiel spatial de l'Eau en France* – Codification Hydrographique et BD CARTHAGE, 6 pp.

Arnell, N.W., Krasovskaia, I. and Gottschalk, L., 1993a. River flow regimes in Europe. In: *Flow Regimes from International Experimental and Network Data (FRIEND)*, vol. 1: Hydrological studies A. Gustard (ed), 112–121, Institute of Hydrology, Wallingford, Oxfordshire, UK.

Arnell, N.W., Oancea, V. and Oberlin, G., 1993b. *European river flow regimes.* Report to the European Environment Agency Task Force, Insitute of Hydrology, Wallingford, and Cemagref, Lyon, 24 pp.

Bartlein, P.J., 1982. Streamflow anomaly patterns in the USA and Southern Canada 1951–1970. *J. Hydrol.*, **57**, 49–63.

Bishop, G.D. and Church, M.R., 1992. Automated approaches for regional runoff mapping in the northeastern United States. *J. Hydrol.*, **138**, 361–383.

Braud, I. and Obled, C., 1991. On the use of Empirical Orthogonal Function (EOF) analysis in the simulation of random fields. *Stochastic Hydrol. Hydraul.*, **5**, 125–134.

Eastman, J.R., 1997. *Idrisi: User's guide. Version 2.0.* Clark University, Graduate School of Geography, Worcester, Massachusetts, USA.

Faulkner, D.S. and Prudhomme, C., 1998. Mapping an index of extreme rainfall across the UK. *Hydrol. Earth Syst. Sci.*, **2**, 183–194.

Gottschalk, L., 1985. Hydrological regionalization of Sweden. *Hydrol. Sci. J.*, **30**, 65–83.

Gottschalk, L., Jensen, J.L., Lunquist, D., Solantie, R. and Tollan, A., 1979. Hydrological regions in the Nordic countries, *Nordic Hydrol.*, **10**, 273–286.

Haines, A.T., Finlayson, B.L. and MacMahon, T.A., 1988. A global classification of river regimes. *Appl. Geogr.*, **8**, 255–272.

Hisdal, H. and Tveito, O.E., 1991. Generation of runoff series at ungauged locations using Empirical Orthogonal Functions in combination with kriging, *Stochastic Hydrol. Hydraul.*, **6**, 255–269.

Hisdal, H. and Tveito, O.E., 1993. Extension of runoff series by the use of empirical orthogonal functions. *Hydrol. Sci. J.*, **38**, 33–50.

Holmström, I., 1963. On a method for parametric representation of the state of the atmosphere. *Tellus*, **15**, 127–149.

IFEN, 1996. *CORINE Land Cover: une base de données géographiques d'occupation du sol*, Agence Européenne pour l'environnement, 4 pp.

Korzun, V.I., 1978. *World water Balance and water resources of the Earth.* UNESCO Studies and Reports in Hydrology, 35, UNESCO, Paris.

Krasovskaia, I., 1997. Entropy-based grouping of river flow regimes. *J. Hydrol.*, **202**, 173–191.

Krasovskaia, I. and Gottschalk, L., 1992. Stability of river flow regimes, *Nordic Hydrol.*, **23**, 137–154.

Krasovskaia, I. and Gottschalk, L., 1995. Analysis of regional drought characteristics with empirical orthogonal functions. In: *New Uncertainty Concepts in Hydrology* Z. Kundzewicz (ed.) Cambridge University Press, Cambridge, UK.

Krasovskaia, I., Arnell, N.W. and Gottschalk, L., 1994. Flow regimes in northern and western Europe: development and application of procedures for classifying flow regimes. In: *FRIEND: Flow Regimes from International Experimental and Network Data*, vol. 221, P. Seuna, A. Gustard, N.W. Arnell, G.A. Cole (Eds), IASH, 185–193.

Krasovskaia, I., Gottschalk, L. and Kundzewicz, Z.W., 1999. Dimensionality of Scandinavian river flow regimes, *Hydrol. Sci. J.*, **44**, 705–723.

Leblois, E. and Sauquet, E., 1997. Scale effects in runoff mapping. In: *FRIEND: Flow Regimes from International Experimental and Network Data*, Projects H-5-5 (IHP IV) and 1.1 (IHP V), third report: 1994–1997. Cemagref Editions, Antony, pp. 109–118.

Leblois, E. and Sauquet, E., 1999. *Grid elevation models in hydrology – Part 1: Principles and literature review; Part 2: HydroDEM, User's manual.* Cemagref, Technical notes, Lyon, 80 pp.

Lvovich, M.I., 1973. *The World's Water.* Mir, Moscow, 213 pp.

Matheron, G., 1971. *The theory of regionalized variables and its applications.* Ecole des Mines, Fontainebleau, France, Cahier 5, 211 pp.

Meijerink, A.M.J., De Brouwer, H.A.M., Mannaerts, C.M. and Valenzuela, C.R., 1994. *Introduction to the use of geographic information systems for practical hydrology.* UNESCO-International Institute for Aerospace, the Netherlands, Publication Number 23, 224 pp.

Nash, J.E. and Sutcliffe, J.V., 1970. River, flow forecasting through conceptual models: 1. A discussion of principles. *J. Hydrol.*, **10**, 257–274.

Obled, Ch. and Creutin, J.D., 1986. Some developments in the use of empirical orthogonal functions for mapping meteorological fields. *J. Appl. Meteorol.*, **25**, 1189–1204.

Pardé, M., 1955. *Fleuves et rivières.* Collection Armand Colin, Paris, 224 pp.

Rao, A.R and Hsieh, C.H., 1991. Estimation of variables at ungaged locations by Empirical Orthogonal Functions. *J. Hydrol.*, **123**, 51–67.

Roche, M., 1963. *Hydrologie de surface.* Ed. Gauthier-Villars, Paris, 429 pp.

Rochelle, B.P., Church, M.R., Gebert, W.A., Graczyk, D.J. and Krug, W.R., 1988. Relationship between annual runoff and watershed area for the eastern United States. *Water Resour. Bull.*, **24**, 35–41.

Stidd, C.K., 1967. The use of eigenvectors for climatic estimates. *J. Appl. Meteorol.* **6**, 255–264.

Weingartner, R. and Aschwanden, H., 1992. Discharge Regime – the basis for the estimation of average flows. In *Hydrological Atlas of Switzerland*, Swiss National Hydrological and Geological Survey, Berne, Switzerland, pl. 5.2.

Zaikov, B.D., 1944. *Vnutrigodovoe raspredelenie rechnogo stoka na territorii Evropy* (Interannual distribution of runoff on the territory of Europe). In Russian. Gidrometizdat, Sverdlovsk-Moskva, 50 pp.