CHAPTER 9

Perspectives

This chapter deals with questions that are currently widespread in frequency analysis. More specifically, we will tackle multivariate frequency analysis using copulas and non-stationarity. We have also included a less recent field concerning regional frequency analysis (GREHYS, 1996) which is still under study. First, however, it would be worth taking a look at Bayesian statistics, although in frequency analysis, this approach remains somewhat marginal (Section 9.1).

9.1 BAYESIAN FREQUENCY ANALYSIS

9.1.1 The Principles of Bayesian Analysis

In the Bayesian approach, the values that the parameters $\theta = (\theta_1, \theta_2, \cdots, \theta_p)$ defining the model can take are uncertain rather than fixed, the opposite of the situation in classical statistics (also known as *frequentist statistics*). In Bayesian analysis, the uncertainty concerning the parameters is represented by a probability distribution called *prior* or *a priori* which is denoted $p(\theta)$. This distribution is established based on the information available *a priori* that does not result from a series of observations $x = (x_1, \ldots, x_n)$, but instead comes from other sources that can be either subjective (expert's or manager's knowledge…) or objective (previous statistical analyses). This means that the Bayesian approach has the advantage of formally incorporating into the analysis the knowledge that is available regarding the parameters of interest.

Once the available observations and the *a priori* distributions have been specified, Bayes' theorem is applied to combine the *a priori* information about the parameters with the information contained in the data, using the likelihood $p(x|\theta)$. Bayes' theorem is expressed as:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta}$$

Thus Bayes' theorem operates as an "information processor" to update the *a priori* knowledge in light of the observations (Bernier *et al.*, 2000).

The diagram of Figure 9.1 illustrates the basic steps in Bayesian reasoning that lead to an *a posteriori* inference.

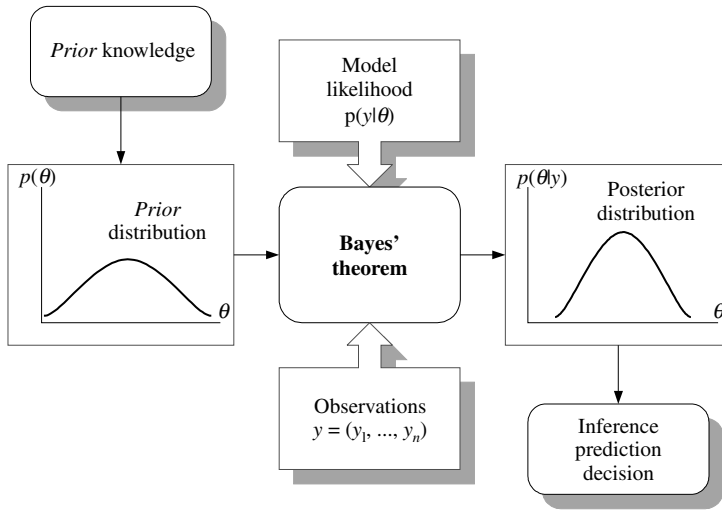

Fig. 9.1 Diagram illustrating the basic steps in Bayesian analysis (taken from Perreault, 2000).

The theoretical principles of Bayesian statistics are described in detail in the reference books by Berger (1985), Bernardo and Smith (1994) and Parent and Bernier (2007). For details on more practical aspects of applying the Bayesian approach, the reader is referred to Gelman *et al.* (2004) and Bernier *et al.* (2000).

9.1.2 The Advantages of the Bayesian Approach in Frequency Analysis

Several of the advantages of the Bayesian method in frequency analysis listed below are borrowed from Perreault (2000, 2003).

- Bayesian inference makes it possible to compute the joint probability (the *a posteriori distribution*) of all the unknown parameters, which in turn allows us to reach several conclusions simultaneously. We can not only obtain estimations of the parameters or functions of these quantities, such as percentiles (which are the values of interest in frequency analysis), but we can also evaluate their accuracy directly. It should be highlighted

that unlike most classical methods of frequency analysis, Bayesian inference is not based on any asymptotic assumption. As Katz *et al*. (2002) pointed out, the standard error of parameters estimated using the maximum likelihood method can sometimes not be applied in practice because of the asymptotic assumptions on which this frequentist approach is based.

- Hydrologists often have unquantified information about the phenomenon of interest before developing a project. The Bayesian approach allows them to formally include their own expertise in their considerations by choosing *a priori distributions*. The possibility to combine this information with the data is even more important because in hydrology, the samples are usually of limited size.
- Managers who are required to make decisions based on the results of a frequency analysis find several advantages in applying the Bayesian method. The idea of updating the *a priori* knowledge of the experts based on information available in the data can be very interesting and reassuring to them. In addition, they are sensitive to the socio-economic aspects of managing water resources. As a consequence, they think that statistical analyses and the assessment of the socio-economic consequences of their decisions should not be handled separately (Bernier *et al*., 2000). Now Bayesian analysis makes it possible to properly integrate the models in a whole decision-making process. This only requires that the analysis includes a function that quantifies the loss incurred for the different options. This is called a "complete" Bayesian analysis, the objective being to choose the optimal solution among all the possibilities.
- Historical data are often available for the estimation of events such as floods (see Section 3.2.5). This information is often of poorer quality than the data from more recent flow measurements. The Bayesian approach offers a framework that is well suited to incorporate these less reliable data.
- As we have already indicated, the result of a Bayesian analysis, corresponding to an *a posteriori* distribution, provides a more complete inference than the corresponding analysis based on maximum likelihood. More precisely, since the objective of an analysis of extreme values is usually to estimate the probability that future events reach a certain extreme level, it seems natural to express this quantity by means of a predictive distribution (Coles, 2001, Section 9.1.2).

9.1.3 Applications of Bayesian Inference in Frequency Analysis

The absence, in most practical cases, of explicit analytical solutions for *a posteriori* distributions, combined with the complexity of the resulting numeric computations, meant that the Bayesian method fell into disfavor. But recent advances in the field of numerical computation, which apply Markov chain Monte Carlo methods (MCMC) to implement Bayesian methods, have overcome this drawback (Robert and Casella, 2004). This has led to a widespread use of the Bayesian approach in various areas of statistical application (Berger, 1999). Surprisingly, the application of Bayesian methods within the water sciences community remains quite limited, contrary to other fields (medicine, finance, economics and biology, just to name a few). However, as we have already pointed out, the Bayesian perspective is conceptually closer to the concerns of the manager in water sciences than the so-called "classical" approach.

In hydrology, the analysis of flood probabilities (the estimation of percentiles) from a Bayesian perspective has not been the subject of much attention. Some of the first work in this regard includes Kuczera (1982, 1999) and Fortin *et al.* (1998). Coles and Tawn (1996) applied the Bayesian technique of elicitation of expert opinion[26] to integrate the knowledge of the hydrologist into the frequency analysis of maximum precipitations. This makes it possible to obtain a credibility interval (equivalent to the frequentist confidence interval) for the quantile corresponding to a 100-year return period. Parent and Bernier (2003a) developed a Bayesian estimation of a peak over a threshold model (POT) by incorporating an informative semi-conjugate *a priori* distribution. This approach makes it possible to avoid using MCMC algorithms, which can be difficult to implement. Katz *et al.* (2002) cite other applications of the Bayesian approach to frequency analysis.

9.2 MULTIVARIATE FREQUENCY ANALYSIS USING COPULAS

9.2.1 Using Copulas in Hydrology

In many areas of statistical applications, including hydrology, the analysis of multivariate events is of particular interest. For example, dam builders

[26] In Bayesian inference the term "elicitation" refers to the action of helping an expert to formalize his knowledge in order to define the *prior distribution*. The steps in an elicitation procedure include: definition of the quantities to elicit, the so-called elicitation itself, which is done through discussion and questions with the expert, the translation of this knowledge into the *a priori* distributions, and finally the verification by the expert of the concordance between these distributions and his own experience.

need to design structures as a function of river discharge, and this discharge increases considerably during the larger flows of springtime. Flows are usually described in terms of three main characteristics: peak flow, volume and duration. Because these three variables are correlated, three univariate frequency analyses cannot provide a full assessment of the probability of occurrence of the event in question. In addition, a univariate frequency analysis can overestimate the severity of a particular event (Yue and Rasmussen, 2002) and as a consequence, raise the construction costs of the structure.

Another application of multivariate frequency analysis in hydrology is to combine the risk downstream from the confluence of several rivers or from a cascade of several sub-watersheds. For several applications, the peak discharge is the result of the combination of discharges in several intermediate watersheds. Thus it is important to take into account any dependence between discharges.

The mathematical theory of univariate models for extreme events has been well established. Based on this theory, it is possible to develop multidimensional models for extreme values, which are the boundary distributions of the joint marginal distribution of the maxima.

Hydrologists apply several classical distributions for the multivariate analysis of extreme events, but such an approach is rarely efficient, partly because the available multivariate models are not very well suited to the representation of extreme values. For a long time, the normal model dominated statistical studies of multivariate distributions. This model is attractive because the conditional and marginal distributions are also normal. However this model is limited and thus it was necessary to find alternatives to the approach using a normal distribution. An abundant statistical literature exists dealing with multivariate distributions.

However, many multivariate distributions are direct extensions of univariate distributions (the bivariate Pareto distribution, the bivariate gamma distribution, etc.). These distributions have the following drawbacks:

- Each marginal distribution belongs to the same family.
- Extensions beyond the bivariate case are not clear.
- The marginal distribution parameters are also used to model the dependence between the random variables.

In hydrology, the most commonly applied multivariate distributions are the multivariate normal, bivariate exponential (Favre *et al*., 2002), bivariate gamma (Yue *et al*., 2001), and bivariate extreme value distributions (Adamson *et al*., 1999). In the case of the multivariate normal distribution, the measure of dependence is summarized in the correlation matrix. In

most cases, using a multivariate normal distribution is not appropriate for modelling maximum discharges as the marginal distributions are asymmetric and heavy-tailed. In addition, the dependence structure is usually different than the Gaussian structure described with Pearson's correlation coefficient. It should also be noted that in the case of more complex marginal distributions, such as mixtures of distributions that are widely used in the practice of modelling heterogeneous phenomena, it is not possible to apply standard multivariate distributions.

One construction of multivariate distributions that does not suffer from the drawbacks already mentioned is the concept of copulas. A copula is extremely useful for implementing efficient and realistic simulation algorithms for joint distributions. Copulas are able to model the dependence structure independently of the marginal distributions. It is then possible to build multidimensional distributions with different margins, the copula mathematically formalizing the dependence structure. The crucial step in the modeling process depends on the choice of the copula function that is best suited to the data and on the estimation of its parameters. Copulas have been widely used in the financial field as a way of determining the "Value at Risk" (VaR) (for example, Embrechts *et al*., 2002, 2003; Bouyé *et al*., 2000). Other areas of application include the analysis of survival data (Bagdonavicius *et al*., 1999) and the actuarial sciences (Frees and Valdez, 1998). However in the field of hydrology, application of copulas still remains marginal.

Hydrologists are concerned with determining statistical quantities such as joint probabilities, conditional probabilities and joint return periods. Once a copula family has been selected, its parameters estimated and the goodness-of-fit tested, these quantities can be computed very simply and directly from the copula expression. Yue and Rasmussen (2002) and Salvadori and De Michele (2004) showed how to calculate these probabilities in the case of a bivariate distribution. Generalization in the multivariate case is straightforward.

The theory of copulas is available in general reference books such as Joe (1997) and Nelsen (2006). Genest and Favre (2007) introduced the required steps for multivariate modeling using copulas in a didactic way, with an emphasis on hydrological applications. Consequently, here we will provide only a quick summary of the theory of copulas.

9.2.2 Definition and Properties of Copulas

Definition

A copula in dimension p is a multivariate distribution C, with uniform marginal distributions on the interval [0,1] ($U[0,1]$) such that

- $C : [0,1]^p \rightarrow [0,1]$;
- C is bounded and p-increasing
- The marginal distributions C_i of C are such that: $C_i(u) = C(1,\ldots,1,u,1,\ldots,1) = u$ whatever $u \in [0, 1]$.

It can be deduced from the definition that if $F_1,\ldots,F_p$ are univariate distributions, then $C(F_1(x_1),\ldots,F_p(x_p))$ is a multivariate distribution with marginal distributions $F_1,\ldots,F_p$, since $F_i(x_i)$, $i = 1,\ldots,p$ is distributed according to a uniform distribution. Copulas are very useful for the construction and simulation of multivariate distributions. In simplified form, a copula can be viewed as a transfer function that makes it possible to link the marginal distributions and the multivariate joint distribution.

Figure 9.2 illustrates a copula in the form of a transfer function.

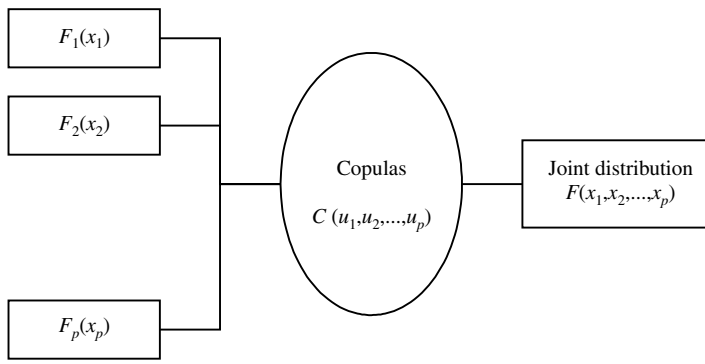

Fig. 9.2 Illustration of copulas as a transfer function.

The following theorem (Sklar, 1959) highlights the high potential of copulas for the construction of multivariate distributions.

Sklar's Theorem

Let F denote a probability distribution in dimension p with continuous marginal distributions $F_1,\ldots,F_p$; then F has the following unique representation in the form of a copula:

$$F(x_1,\ldots,x_p) = C\ (F_1(x_1),\ldots,F_p(x_p)) \tag{9.1}$$

It can be deduced from Sklar's theorem that it is possible to separate the marginal distributions and the multivariate dependence structure.

Schweizer and Wolff (1981) showed that copulas are able to take into account the whole dependence between two random variables X_1 and X_2 in the following way: let g_1 and g_2 be two strictly increasing functions on the domain of definition of X_1 and X_2. Thus the variables obtained by

the transformation, namely $g(X_1)$ and $g(X_2)$ have the same copula as X_1 and X_2. This means, as stated by Frees and Valdez (1998), that the way in which X_1 and X_2 are linked is included in the copula, independently of the measurement scale of each variable. Schweizer and Wolff (1981) also proved that two classical nonparametric measures of correlation can be simply expressed by means of copulas – Kendall's tau and Spearman's rho.

9.2.3 Types of Copulas

Archimedean Copulas

There are several families of copulas applied for modeling a set of variables. The Archimedean class of copulas, popularized by Genest and MacKay (1986), is the family most widely applied in practice, due to its interesting analytical properties and the ease with which it can be simulated.

A copula C in dimensions p is said to be Archimedean if there exists a strictly decreasing continuous function φ: $[0, 1] \rightarrow [0, \infty]$ satisfying the condition $\varphi(1) = 0$ so that

$$C(u_1,\cdots,u_p) = \varphi^{[-1]}\{\varphi(u_1)+\ldots+\varphi(u_p)\}$$

where φ^{-1} represents the inverse of the generator φ.

The most common examples of Archimedean copulas are the Gumbel, Frank and Clayton families. The Clayton model is frequently applied in life cycle analysis, under the name gamma frailty model; see Oakes (1982) for example.

Table 9.1 summarizes the main Archimedean copulas and their generators in the bivariate case. The generator determines the Archimedean copula in a unique way (they can only differ by a constant).

Table 9.1 Main bivariate Archimedian copulas and their generators.

Family	Generator $\varphi(t)$	Parameter α	Bivariate copula $C_\varphi(u_1, u_2)$
1. **Independence**	$-\ln t$	–	$u_1 u_2$
2. **Clayton** (1978) Cook and Johnson (1981) Oakes (1982)	$t^{-\alpha} - 1$	$\alpha > 0$	$(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha}$
3. **Gumbel** (1960) Hougaard (1986)	$(-\ln t)^{\alpha}$	$\alpha \geq 1$	$\exp\left[-\left\{(-\ln u_1)^{\alpha} + (-\ln u_2)^{\alpha}\right\}^{1/\alpha}\right]$
4. **Frank** (1979)	$\ln\left\{\dfrac{\exp(\alpha t) - 1}{\exp(\alpha) - 1}\right\}$	$\alpha \neq 0$	$\dfrac{1}{\alpha}\ln\left[1 + \dfrac{\{\exp(\alpha u_1) - 1\}\{\exp(\alpha u_2) - 1\}}{\exp(\alpha) - 1}\right]$

The family of Archimedean copulas also includes the Galambos copula (1975), the Ali-Mikhail-Haq copula (1978), the Cuadras-Augé copula (1981), the Hüsler-Reiss copula (1989) and the Genest-Ghoudi copula (1994).

Meta-elliptical Copulas

The elliptical class of copulas holds considerable practical interest, because they can be easily applied in dimension p (p>2). In addition, this class constitutes a generalization of the classical multivariate normal distribution. More precisely, a distribution is said to be elliptical if it can be represented in the form of

$$Z_{p\times 1} = \mu_{p\times 1} + RA_{p\times p}\mathbf{u}_{p\times 1} \sim EC_p(\mu, \Sigma, g),$$

where μ is a location vector, R is a positive random variable, $\mathbf{A}$ is a $p \times p$ matrix such as $\mathbf{AA}^{\mathrm{T}} = \Sigma$, $\mathbf{u}$ is a uniformly distributed vector on the sphere of dimension p and g is a scale function.

In the particular case where $g(t) \propto e^{-t/2}$, we find the classical multivariate normal distribution. Other generators correspond to multivariate Student, Cauchy and Pearson type II distributions, among others (see e.g. Fang *et al*., 2002 or Abdous *et al*., 2005).

The marginals of elliptical distributions are entirely determined by the generator g and, moreover, are all identical. However, by transforming the variables individually to make the marginals uniform, we obtain a copula known as meta-elliptical, into which we can now insert the desired margins.

In other terms, it is possible to build a vector $X_{p\times 1} = (X_1,\cdots,X_p)$ with marginals $F_1,\ldots,F_p$ from an elliptical vector $(Z_1,\ldots,Z_p)$ by using

$$X_i = F_i^{-1}\{Q_g(Z_i)\},\ 1 \le i \le p$$

where Q_g is the common marginal of the variables $Z_1,\ldots,Z_p$ and F_i^{-1} denotes the inverse of the distribution function of the marginal F_i of X_i. This construction, based on a meta-elliptical copula, is written as

$$X_{p\times 1} \sim ME_p(\mu, \Sigma, g; F_1,\cdots,F_p).$$

Fang *et al*. (2002) and Abdous *et al*. (2005) discuss various properties of meta-elliptical copulas. Among these, the most important is the fact that Kendall's tau (τ) between two components of an elliptical vector is independent of the function g and is linked to the correlation r between the two variables by the formula $r = \sin(\tau\pi/2)$.

Genest *et al*. (2007) developed a trivariate frequency analysis of flood peak, volume and duration by applying meta-elliptical copulas.

9.2.4 Estimation of Copula Parameters

Without loss of generality, let us suppose that the dimension $p = 2$ and that a parametric family of copulas (C_θ) has been chosen to model the dependence between two random variables X and Y. Given a sample $(X_1,Y_1),\ldots,(X_n,Y_n)$ from (X,Y), how should the parameter θ be estimated?

Several methods exist for the estimation. The most direct approaches involve inverting Kendall's τ or Spearman's ρ. These two methods can only be applied if there is a direct relation between these nonparametric measures of dependence and the parameter of the copula. In addition, they can only be used if the parameter is a real number (and therefore unidimensional). Below we present an efficient approach that is based solely on ranks. This approach follows Genest's school of thought (discussed for example in Genest and Favre, 2007) which advocates ranks as the best summary of the joint dependence of random pairs.

Maximum Pseudolikelihood Method

In classical statistics, the maximum likelihood method is a well-known alternative to the method of moments and is usually more efficient, especially when θ is multidimensional. In the current case, an adaptation of this method is required, as we would like the inference concerning the dependence parameters to be based exclusively on ranks. Such an adaptation was described briefly by Oakes (1994) and was later formalized and studied by Genest *et al.* (1995) and by Shih and Louis (1995).

The maximum pseudolikelihood method assumes that C_θ is absolutely continuous with density c_θ and involves the maximization of the log-likelihood based on ranks in the following form:

$$l(\theta) = \sum_{i=1}^{n} \log\left\{c_\theta\left(\frac{R_i}{n+1}, \frac{S_i}{n+1}\right)\right\},$$

where R_i and S_i are the respective ranks of X_i and Y_i, $i = 1,\ldots,n$.

The above equation corresponds exactly to the expression obtained when the unknown marginal distributions F and G in the classical log-likelihood method

$$l(\theta) = \sum_{i=1}^{n} \log[c_\theta\{F(X_i),G(Y_i)\}]$$

are replaced by rescaled versions of their empirical counterparts, i.e.:

$$F_n(x) = \frac{1}{n+1}\sum_{i=1}^{n} 1(X_i \le x) \text{ and } G_n(y) = \frac{1}{n+1}\sum_{i=1}^{n} 1(Y_i \le y),$$

where 1 denotes the indicator function equal to one when the inequality holds and zero otherwise.

In addition, it is immediate that $F_n(X_i) = R_i/(n + 1)$ and $G_n(Y_i) = S_i/(n + 1)$ for all $i \in \{i,\ldots,n\}$.

9.2.5 Goodness-of-fit Tests

Recently, goodness-of-fit tests for copulas have been developed. But although the field is in constant development, only a few papers have actually been published in the literature on this topic. Goodness-of-fit tests for copulas fall into three general categories:

- goodness-of-fit tests based on the probability integral transformation, for example Breymann *et al*. (2003);
- goodness-of-fit tests based on the kernel estimation of the copula density, for example Fermanian (2005);
- goodness-of-fit tests based on the empirical process of copulas, for example Genest *et al*. (2006) and Genest and Rémillard (2007).

We prefer the third type of tests. The first implies conditioning on successive components of the random vector and has the drawback of depending on the order in which this conditioning is done. The second category of tests depends on various arbitrary choices, such as the kernel, the window size and the weight function, which make their application cumbersome.

9.2.6 Application of Copulas for Bivariate Modeling of Peak Flow and Volume[27]

A flow is composed of three main characteristics: the peak, the duration and the volume. Univariate frequency analyses of these quantities result in the over-estimation or under-estimation of risk (De Michele *et al*., 2005), which can have disastrous consequences. The following example concerns the bivariate analysis of peak flows (maximum annual daily discharge) and the corresponding volumes for the Harricana River. The watershed, with a area of 3,680 km^2, is located in the northwest region of the province of Quebec, Canada. The data considered for this application include the annual maximum discharge X (in m^3/s) and the corresponding volume Y (in hm^3) for 85 consecutive years starting in 1915 and ending in 1999. The peak flow follows a Gumbel distribution with mean 189 [m^3/s] and standard error 51.5 [m^3/s]. The volume is adequately modeled by a gamma distribution with mean 1043.88 [hm^3] and standard error 234.93 [hm^3].

[27] A major portion of this section is taken from Genest and Favre (2007).

The scatter plot of ranks shown in Figure 9.3 suggests a positive association between peak flow and volume as measured by Spearman's rho (which represents the correlation of ranks). Computation of the two standard nonparametric measures of dependence gives us $\rho_n = 0.696$ for Spearman's rho and $\tau_n = 0.522$ for Kendall's tau.

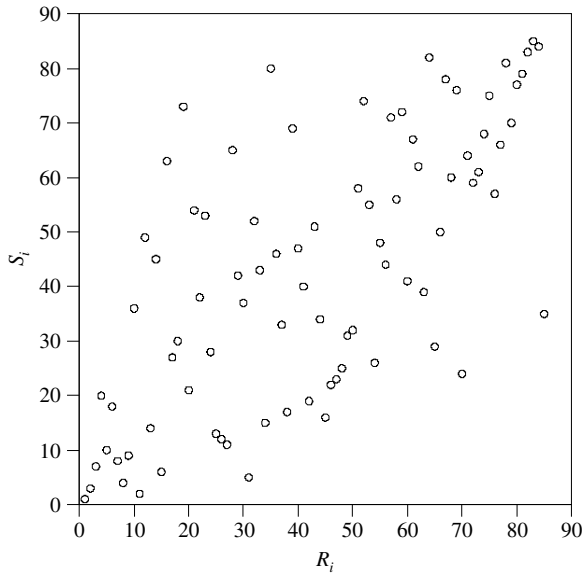

Fig 9.3 Scatter plot of the ranks of peak flow (R_i) and corresponding volume (S_i).

In order to model the dependence between the annual peak and the volume, some 20 families of copulas were considered. These could be classified into four broad categories:

1. Archimedean copulas with one, two or three parameters, including the traditional Ali-Mikhail-Haq (Ali *et al*., 1978), Clayton (1978), Frank (Nelsen, 1986; Genest, 1987) and Gumbel-Hougaard (Gumbel, 1960) families. The families of Kimeldorf and Sampson (1975), the class of Joe (1993) and the BB1-BB3, BB6-BB7 described in the book by Joe (1997, p. 150-153) were also considered as potential models.
2. Extreme values copulas, including (besides the Gumbel-Hougaard copula mentioned above) Joe's BB5 family and the classes of copulas introduced by Galambos (1975), Hüsler and Reiss (1989) and Tawn (1988).
3. Meta-elliptical copulas, more specifically the normal, Student and Cauchy copulas.

4. Various other families of copulas, such as those of Farlie-Gumbel-Morgenstern and Plackett (1965).

The Ali-Mikhail-Haq and Farlie-Gumbel-Morgenstern families could be eliminated off hand because the degree of dependence they span were insufficient to account for the association that we can observe in the data set. To discriminate between the other models we used a graphical tool described in detail Genest and Favre (2007). This graphic is constructed as follows: given a family of copulas (C_θ), an estimation θ_n of its parameters is obtained using the maximum pseudolikelihood method. Then 10,000 pairs of points (U_i, V_i) were generated from C_{θ_n}. The resulting pairs were transformed back into their original units by applying the inverse of the marginal distributions identified beforehand (as mentioned earlier, the domain of definition of a bivariate copula is the unit square, [0,1]×[0,1]). Figure 9.4 shows a scatterplot of these pairs with the observations superimposed for the five best families as well as the traditional bivariate normal model.

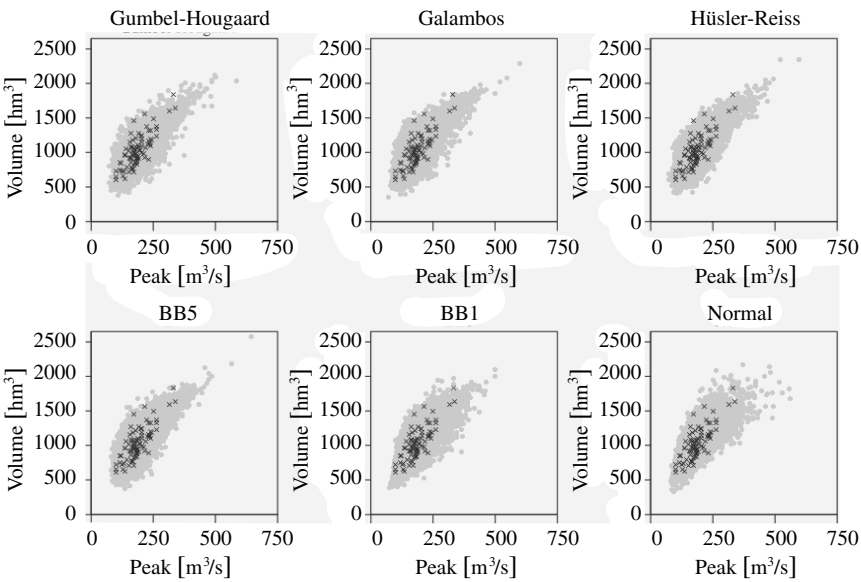

Fig. 9.4 Simulations of 10,000 peak flows and corresponding volumes using various families of copulas (Gumbel-Hougaard, Galambos, Hüsler-Reiss, BB5, BB1 and normal) and marginal distributions (Gumbel for the peak and gamma for the volume). The observations are superimposed, and indicated with the symbol *x*.

Keeping in mind the predictive ability that the selected model should possess, the normal copula was rejected due to the obvious lack of fit in the upper part of the distribution. Table 9.2 shows the definition of the five chosen copula models, as well as the domain of definition of their parameters. Note that four of these families of copulas belong to the extreme values class (Gumbel-Hougaard, Galambos, Hüsler-Reiss and BB5). The fifth (BB1) belongs to the two-parameter Archimedean class of copulas.

Table 9.2 Mathematical expressions of the five chosen families of copulas, with their parameter space

Copula	$C_\theta(u, v)$	Parameter(s)
Gumbel-Hougaard	$\exp\{-(\tilde{u}^\theta + \tilde{v}^\theta)^{1/\theta}\}$	$\theta \geq 1$
Galambos	$uv\exp\{(\tilde{u}^{-\theta} + \tilde{v}^{-\theta})^{-1/\theta}\}$	$\theta \geq 0$
Hüsler-Reiss	$\exp\left[-\tilde{u}\Phi\left\{\frac{1}{\theta} + \frac{\theta}{2}\log\left(\frac{\tilde{u}}{\tilde{v}}\right)\right\} - \tilde{v}\Phi\left\{\frac{1}{\theta} + \frac{\theta}{2}\log\left(\frac{\tilde{v}}{\tilde{u}}\right)\right\}\right]$	$\theta \geq 0$
BB1	$[1 + \{(u^{-\theta_1} - 1)^{\theta_2} + (v^{-\theta_1} - 1)^{\theta_2}\}^{1/\theta_2}]^{-1/\theta_1}$	$\theta_1 > 0,\ \theta_2 \geq 1$
BB5	$\exp\left[-\{\tilde{u}^{\theta_1} + \tilde{v}^{\theta_1} - (\tilde{u}^{-\theta_1\theta_2} + \tilde{v}^{-\theta_1\theta_2})^{-1/\theta_2}\}^{1/\theta_1}\right]$	$\theta_1 \geq 1,\ \theta_2 > 0$

Note: with $\tilde{u} = -\log(u)$, $\tilde{v} = -\log(v)$ and Φ standing for the cumulative distribution function of the standard normal.

Table 9.3 contains the parameters estimation obtained with the maximum pseudo-likelihood method as well as the corresponding 95% confidence interval (Genest *et al.*, 1995) for each of the five chosen models.

Table 9.3 Parameters estimated using the-maximum pseudolikelihood method and the corresponding 95% confidence interval.

Copula	Estimated parameter(s)	95% confidence interval
Gumbel-Hougaard	$\hat{\theta}_n = 2.161$	IC = [1.867, 2.455]
Galambos	$\hat{\theta}_n = 1.464$	IC = [1.162, 1.766]
Hüsler-Reiss	$\hat{\theta}_n = 2.027$	IC = [1.778, 2.275]
BB1	$\hat{\theta}_{1n} = 0.418,\ \hat{\theta}_{2n} = 1.835$	IC = [0.022, 0.815] × [1.419, 2.251]
BB5	$\hat{\theta}_{1n} = 1.034,\ \hat{\theta}_{2n} = 1.244$	IC = [1.000, 1.498] × [0.774, 1.294]

Goodness-of-fit tests based on the empirical process of copulas were then applied in order to differentiate between the five families of copulas (see Genest and Favre (2007), p. 363 and Tables 12 and 13). These tests rely on the parametric bootstrap method and show that at the 5% level,

none of the models still under consideration can be rejected. If the objective is to choose only one of the families, the one with the largest p-value in the goodness-of-fit test could be chosen. If we consider the test using the Cramér-von Mises process, the BB5 copula has the largest p-value ($p = 0.6362$, see Table 13 in Genest and Favre (2007)).

As an example, let us consider the following return period computation

$$T'(x, y) = \frac{1}{\Pr(X > x, Y > y)}$$

for various values of x and y corresponding to fixed univariate return periods. Now we have

$$\begin{aligned} \Pr(X_1 > x_1, X_2 > x_2) &= 1 - F_1(x_1) - F_2(x_2) + F(x_1, x_2) \\ &= 1 - F_1(x_1) - F_2(x_2) + C_\theta(u, v) \qquad (9.2) \\ \text{with } u &= F_1(x_1) \text{ and } v = F_2(x_2) \end{aligned}$$

Figure 9.5 illustrates the computation of this bivariate probability. Basically, we have: the area in (4) = 1 – area in (1) – area in (2) + area in (3) (equality translated as a probability in the first line of equation 9.2).

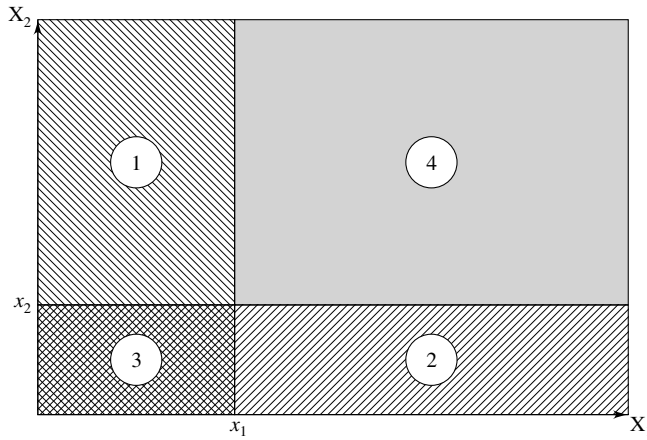

Fig. 9.5 Illustration of the computation of the bivariate probability $\Pr(X_1 > x_1, X_2 > x_2)$.

We denote $x_{[T]}$ (respectively $y_{[T]}$), the quantile corresponding to a peak flow (and respectively to a volume) of return period T. We applied equation 9.2 with the BB5 family of copulas and obtained the following bivariate return periods:

$$T'(x_{[10]}, y_{[10]}) = 16.16 \text{ [years]}, \; T'(x_{[50]}, y_{[50]}) = 84.00 \text{ [years] and}$$
$$T'(x_{[100]}, y_{[100]}) = 168.83 \text{ [years]}$$

If the peak flow and volume had been considered as independent, the return periods would have been $10^2 = 100$ [years], $50^2 = 2{,}500$ [years] and $100^2 = 10{,}000$ [years] respectively. By assuming such independence, the risk would have been highly underestimated.

This example clearly highlights the importance of taking into account the dependence between these two quantities, and in a more general sense, the importance of applying a multivariate approach in cases such as the one illustrated.

Favre *et al.* (2004) developed another application for copulas for the bivariate frequency analysis of peak flow and volume. In this context, Poulin *et al.* (2007) highlighted the importance of taking into account the dependence between the largest values (modeled in copula theory by the concept of tail dependence), mainly if extrapolation is required.

9.3 FREQUENCY ANALYSIS IN A NON-STATIONARITY CONTEXT

Classical frequency analysis is based on the assumption of underlying independent and identically distributed (i.i.d.) random variables. The second of these assumptions is not valid in a non-stationarity context, induced by climate changes for example. Basically, strict-sense stationarity means that the distribution remains constant over time. From a practical point of view, statisticians are used to relying on second-order stationarity, which implies that the first two moments (mean and variance) do not vary over time. The frequency analysis of a non-stationary series calls for a different understanding than the conventional approach involving stationarity. In fact, in the context of climate change, the distribution parameters and the distribution itself are likely to be modified. As a consequence, the exceedance probability used to estimate the return period also varies over time.

Several recent methods make it possible to take into account, at least partly, non-stationarity in the context of a frequency analysis.

The simplest method, which is not very recent (Cave and Pearson, 1914) involves first removing the trend, for example, by differentiating the series one or several times and then choosing the distribution and estimating its parameters. Below we give an overview of some recent approaches, taken mostly from a recent literature review (Khaliq *et al.*, 2006).

9.3.1 Method of Moments that Change Over Time

Strupczewski and Kaczmarek (2001) as well as Strupczewski *et al.* (2001 a, b) incorporate a linear or parabolic trend in the first two moments

of the distribution. To apply this method, the distribution parameters are expressed as a function of the mean and the variance. The time-dependent parameters are estimated using the maximum likelihood method or the weighted least-squares approach. The validity of this method was proven by using non-negative asymmetric distributions with the two first moments expressed analytically as a function of the parameters. Six distributions were under consideration: the normal distribution, the 2-parameter lognormal, 3-parameter lognormal, gamma, Pearson type III and Gumbel distributions. For example, when considering a Gumbel distribution and a linear trend for both mean and variance, which is to say $\mu(t) = \mu_0 + \mu_1 t$ and $\sigma(t) = \sigma_0 + \sigma_1 t$, the Gumbel distribution parameters can be expressed as

$$\beta = \sqrt{6}\,\frac{\sigma_0 + \sigma_1 t}{\pi},\ \alpha = \mu_0 + \mu_1 t - 0.5772\beta$$

This means that the baseline distribution parameters are replaced with the following trend parameters (μ_0, μ_1, σ_0, σ_1). It is obvious that this approach generates a family of distributions, meaning that each time (t) has a distribution of its own.

9.3.2 Non-Stationary Pooled Frequency Analysis

Cunderlik and Burn (2003) propose a second-order non-stationary approach for flood frequency analysis by assuming non-stationarity in the first two moments, or in other words of the mean and variance of a time series. The quantile function at the local scale is splitted into a local non-stationary part that includes location and scale parameters and a regional stationary part. With this method the quantile function is expressed as $Q(F,t) = \omega(t)q(F)$ where $Q(F,t)$ is the quantile with probability F at time t, $\omega(t) = \mu(t)\sigma(t)$ is a local time-dependent component and $q(F)$ is a regional component independent of time. The parameters of the part dependent on time are estimated by splitting the time series into a trend portion and a time-dependent random variable representing the residual. This last variable models irregular fluctuations around the trend. The validity of this method is proven by assuming a linear trend for $\mu(t)$ and $\sigma(t)$.

Covariables Method

The statistical modeling of extremes is usually handled by one of the three main approaches: generalized extreme values distribution (GEV), peak over a threshold (POT, for example Reiss and Thomas, 2001) or point processes (for example Coles, 2001).

The idea underlying the covariables approach is to integrate changes that have taken place in the past directly into the techniques of

frequency analysis in order to extrapolate in the future. It is possible that the statistical characteristics of extreme values vary as a function of low frequency climate indices such as El Niño-Southern Oscillation (ENSO). Modifications in the extreme values and their links with ENSO have been modeled by incorporating covariables into the distribution parameters (McNeil and Saladin, 2000; Coles, 2001; Katz *et al*., 2002; Sankarasubramanian and Lall, 2003). Given a vector of covariables $V=v$, the conditional distribution of the maximum is assumed to be GEV with parameters dependent on v. For example, by considering the time as an explanatory variable, we can assume that the location parameter, the logarithm of the scale parameter and the shape parameter of the GEV distribution are linear functions of time. However, the shape parameter is often considered to be independent of time because the other two parameters are usually more important (McNeil and Saladin, 2000; Zhang *et al*., 2004) and this parameter is also difficult to estimate even in the classical case of stationarity. Covariables can be inserted in the same manner for peak over a threshold models (McNeil and Saladin, 2000, Katz *et al*., 2002) and in approaches based on point processes.

Approaches Based on Local Likelihood

Trend analysis is usually applied to determine an appropriate trend relation for frequency analysis in a context of non-stationarity. If *a priori* information about the structure of change exists, this approach can be better than defining a pre-specified parametric relation. In this context, the semi-parametric approach based on local likelihood (Davison and Ramesh, 2000; Ramesh and Davison, 2002) can be very useful for an exploratory analysis of extreme values. In these semi-parametric approaches, the classical models for reproducing the trends in annual maximum or POT series are estimated by applying a local calibration, while the parameters of these models are estimated separately for each time by weighting the data in the appropriate manner. This method leads to an estimation of parameters that is time-dependent and produces local estimations of extreme quantiles.

Estimation of Conditional Quantiles

In this type of approach, the conditional distribution of quantiles is studied in the presence of covariables or predictors, which means that the distribution of the parameters or of the moments is conditioned by the state of the covariables.

The best known method is based on quantile regression. This parametric approach was developed by Koenker and Basset (1978). It relies on the estimation of conditional quantiles by minimizing the weighted sum of asymmetric deviations by attributing different weights

to the positive and negative residuals applying optimization techniques. Let p be the conditional quantile defined by the following regression: $Y_p(t) = \psi_p(V_t) + \varepsilon_p(t)$ where $\psi_p(.)$ is a linear or nonlinear function linking the p conditional quantile and the climate indices and $\varepsilon_p(t)$ is a white noise with zero mean and variance σ_p^2. The noise process can be either homoscedastic (with constant variance), or heteroscedastic (the variance depends on the explanatory variable). Let $V = (V_{1t}, V_{2t}, \ldots, V_{mt})$ be the vector of m covariables, corresponding for example to climate indices. The function ψ is obtained by resolving the following minimization problem:

$$\min_{\psi(V_t)} \sum_{t=1}^{n} R_p\left[Y_p(t) - \psi_p(V_t)\right] \text{ where } R_p(u) = \lfloor |\mu| + (2p-1)u \rfloor/2$$

If the regression function is assumed to be linear then for $p = 0.5$, the regression is defined by means of $\psi_p(V_t) = a_p(V_t)$ where a_p is a $m \times 1$ regression coefficient vector for the p-quantile. Koenker and d'Orey (1987) developed an algorithm for estimating a_p by using linear programming.

Another approach based on local likelihood makes it possible to insert covariables into the local estimation with the aim of estimating the conditional quantiles. For local likelihood, the weights are chosen by applying, for example, a kernel method (Sankarsubramanium and Lall, 2003).

Despite the existence of all these methods, a fundamental problem remains to be solved. The return period is defined as the mean of the recurrence interval, measured on a very large number of occurrences. Now, in a context of non-stationarity, the mean changes over time and as a consequence, the concept of return period makes no sense. As a result, serious thought is required to find a concept that can replace return period in a non-stationarity context.

9.4 REGIONAL FREQUENCY ANALYSIS

A regional analysis of hydrometeorogical variables can answer to two main needs:

- We need to be able to use these parameters on a spatial basis. We are thinking here especially of hydrometeorological data concerning rain or temperature, but also of relation parameters for the regionalization of discharges.
- We need robust local estimations, a characteristic that ideally can be improved by simultaneously taking into account the data from several stations in a single region.

We can catch a glimpse of the possibility of solving problems in areas where measurements are insufficient or nonexistent, and also the possibility of improving local estimations. The interactions between these approaches are straightforward and are illustrated in Fig. 9.6.

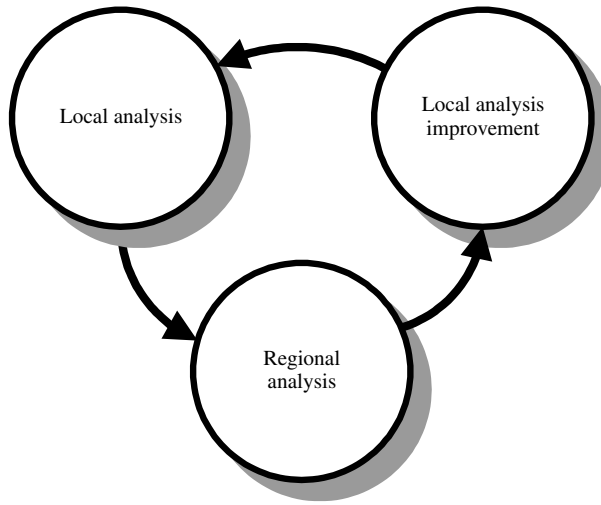

Fig. 9.6 Concept of regional analysis.

It is impossible to solve this problem in this section because the literature – and attempts at finding a solution – is too extensive. Also, research on this topic is still in fast development (Ouarda *et al.*, 2006; Castellarin *et al.*, 2007; Neykov *et al.*, 2007; Ribatet *et al.*, 2007; Viglione *et al.*, 2007). All we can accomplish here is to introduce a concept that is indispensable for hydrologists. Moreover it is important to be aware that in these techniques of regionalization, frequency analysis in the strictest sense plays a relatively modest role in comparison to all the other aspects of hydrology – rainfall-runoff modeling, for example.

We have chosen to divide the techniques of regional analysis into two main categories (GREHYS, 1996): the approach using the *anomaly method* and the one using *regional frequency analysis*. In both cases the basic hypothesis to be considered is that the phenomenon under study shows a certain degree of *homogeneity* over the region of interest. This hypothesis can be assessed either at the level of a coherent spatial structure of the variable of interest (residual method), or of the homogeneity of behavior allowing the introduction of auxiliary explanatory variables (for example, a definition of the flood index by a series of parameters), or even a combination of these two explanatory approaches.

9.4.1 Anomaly Method

The anomaly method (Peck and Brown, 1962) is based on the idea that the variable of interest is correlated to variables that are easy to define for each point in the region to be mapped (altitude, distance from a coast or a mountain chain, slope orientation, etc.).

We make the assumption that we have a sample of the variables of interest for N stations. Three phases are then involved:

1. Construction of an explanatory model

A multiple regression model for the response variable z can be constructed as follows:

$$\hat{z} = a_0 + a_1 w_1 + \cdots + a_p w_p \tag{9.3}$$

where $w_1,\ldots,w_p$ denote the p explanatory variables and $a_0,\ldots,a_p$ are the p+1 parameters estimated by least-squares using data from the stations $i = 1,\ldots,N$. For each station i the model implementation results in an anomaly (or residual):

$$r_i = \hat{z}_i - z_i \tag{9.4}$$

which models the difference between the value z_i observed at the station i and the value $\hat{z}_i$ estimated by the model.

After this first step we obtain a series of N residuals r_i (see for example Fig. 9.7) for which the variance is much lower than the variance of the basics field z_i.

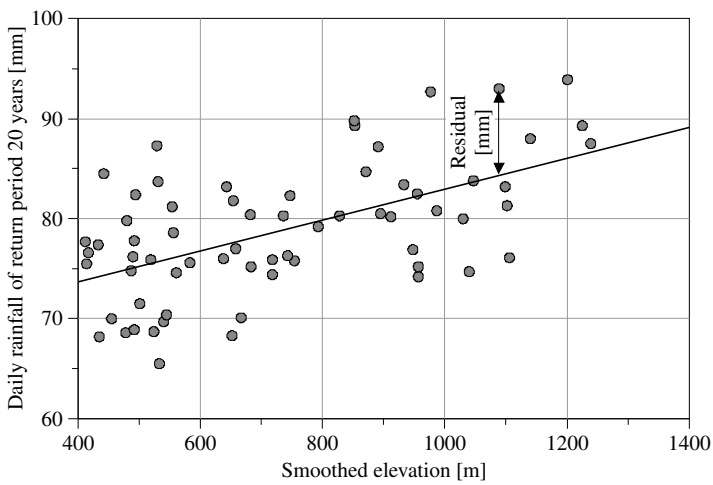

Fig. 9.7 Relation between daily rainfall corresponding to a 20-year return period and smoothed altitude for a window of 6 × 6 kilometers, based on Meylan (1986).

The several variants of the method differ mainly with regards to the choice of the explanatory variables.

2. Spatial interpolation of residuals

In this step, we must make the assumption that the residual field shows a coherent *spatial structure*, an assumption that can be verified by studying the *variogram* of residuals. It is then possible to regionalize the residuals on a regular grid, by applying the *kriging* method, for example.

3. Reconstruction of the variable of interest

For each point of the grid where it is possible to determine the p explanatory variables of the regression model, we can easily reconstruct the estimation of the variable of interest z as follows:

$$\hat{z}(x,y) = a_0 + a_1 w_1(x,y) + \cdots + a_p w_p(x,y) + r(x,y) \tag{9.5}$$

where $r(x, y)$ represents the value of the "spatial" residual at point (x,y).

Applications

The various applications of the anomaly method, the majority of which are applied for rainfall studies, can basically be distinguished by the choice of the explanatory variables.

De Montmollin *et al*. (1979) used smoothed altimetric information for a 500 meters square to model annual rainfall modules for the Mentue watershed.

Jordan and Meylan (1986 a, b) used smoothed altitude on a 6 km × 6 km window to study maximum daily rainfall in western Switzerland. In addition, a technique to take into account uncertainties in the data was applied by using a *structural variogram* (Meylan, 1986). Figure 9.7 shows the relation between daily rainfall corresponding to a 20-year return period and smoothed altitude, while Fig. 9.8 illustrates the result of regionalization.

Bénichou and Le Breton (1987) applied a principal components analysis of the landscape to select the explanatory variables for statistical rainfall fields. This approach made it possible to define the types of landscape that are most relevant for understanding rainfall fields: entrenchment effect, general west-east slope, general north-south slope, north-south pass, etc.

9.4.2 Regional Frequency Analysis

In the absence of a more highly developed approach, which does not yet exist, the *flood index* method is commonly used in practice to estimate flood discharge.

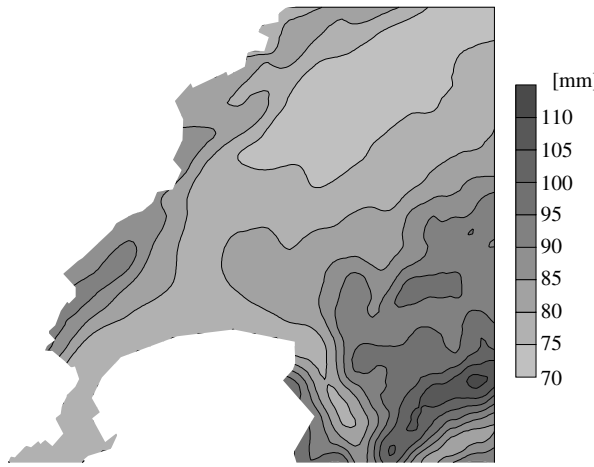

Fig. 9.8 Map of daily rainfalls [mm] corresponding to a 20-year return period for western Switzerland, based on Meylan (1986).

The flood index method was created for the purpose of regionalizing discharges (Dalrymple, 1960). It is usually used for large watersheds.

The basic hypothesis is that the region is homogenous, which is to say that the frequency distributions $Q_i(F)$ of maximum annual floods at each of the N stations i are identical, with only one different scale factor, which is called the *flood index*. Thus we can write:

$$Q_i(F) = \mu_i q(F),\ i = 1, \cdots, N \tag{9.6}$$

where μ_i, the mean maximum annual flood, is the flood index representing a scale factor,

$q(F)$ is the standardized regional frequency distribution.

It should be noted that $q(F)$ is also called the *region curve* (NERC, 1975) or *regional growth curve* (NERC, 1983) or also the *growing factor*. The general procedure introduced by Dalrymple (1960) remains one of the most popular for a regional approach to runoff (GREHYS, 1996).

The regional model has at least two parameters:

- The *flood index* μ_i, whose value varies from one measurement station to another and which must be regionalized if the goal is to apply the model for points where measurements are not available;
- the *growth curve* $q(F)$, identical for each point in each of the *homogeneous regions*.

The Hosking and Wallis Procedure

Hosking and Wallis (1993) proposed the following method for regional frequency analysis.

The mean of the flood discharges at station i expressed in equation (9.6) is given by $\hat{\mu}_i = \overline{Q}_i$, which is to say, the mean of Q_{ij}: annual flood discharges for the year j at station i. Other estimators can be considered, such as the median or the trimmed mean.

The following normed data

$$q_{ij} = \frac{Q_{ij}}{\hat{\mu}_i},\ j = 1, \cdots, n_i;\ i = 1, \cdots, N \tag{9.7}$$

therefore serve as the basis for the estimation of $q(F)$. In equation (9.7), n_i denotes the total number of years of observations.

This method is based on the hypothesis that $q(F)$ is known and that only the p parameters $\theta_1,\ldots,\theta_p$ need to be estimated.

The procedure consists of estimating, separately for each station i, the parameters $\hat{\theta}_k^{(i)}$ which are the "local" parameters. Next, these parameters are combined as follows to give the regional estimation:

$$\hat{\theta}_k^{(R)} = \frac{\sum_{i=1}^{N} n_i \hat{\theta}_k^{(i)}}{\sum_{i=1}^{N} n_i} \tag{9.8}$$

Equation (9.8) is a weighted mean, where the estimation for station i has a weight proportional to n_i (the sample size). These estimations lead to $\hat{q}(F)$, the regional growth curve.

Hosking and Wallis (1993) also proposed several scores that allow to check the data (using a measure of the *discordance* D_i), identify the homogeneous regions (with a measure of *heterogeneity* H) and choose the regional distribution to use (by way of a measure of *fit* Z).

Example of Application

Niggli (2004) studied the discharges of 123 rivers in western Switzerland and after dividing the area into three "homogeneous" regions and removing watersheds with an area less than 10 km^2, obtained the growth curves shown in Fig. 9.9.

To regionalize the flood indices, Niggli (2004) proposed the following equation for two of the regions, the *Plateau* and the *Pre-Alps*:

$$\mu = 0.0053A^{0.78}ALT^{0.68}\exp(0.58DRAIN - 0.025FOREST + 0.023IMP) \tag{9.9}$$

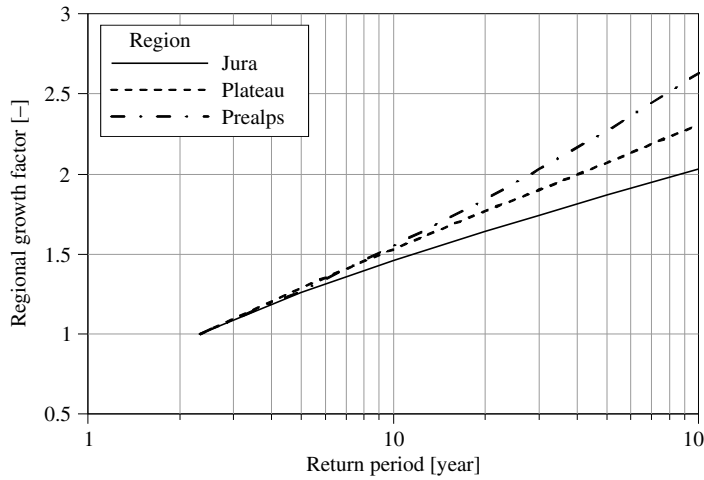

Fig. 9.9 Growth curve $q(T)$ for three "homogeneous" regions, namely the Jura, the Plateau and the Pre-Alps (Switzerland). Based on Niggli (2004).

where μ is the flood index under study,
A is the area of the watershed [km^2],
ALT is the mean altitude of the watershed [m],
DRAIN is the drainage density [km^{-1}] defined by L/A where L is the length of the hydrographic network,
FOREST is the proportion of forested area [%],
IMP is the proportion of impermeable area [%].

Thus Niggli (2004) applied a multiple linear regression with five explanatory variables selected on the basis of a "stepwise" regression procedure.

9.4.3 Other Approaches

This paragraph gives as an example an approach that is difficult to attribute to either the *anomaly method* or to *flood index* techniques.

Often, a technique for estimating flood discharge in small watersheds (generally those with an area less than 10 km^2 (OFEG, 2003)) is advocated. This kind of technique falls in between the anomaly methods and the flood indices methods. Anomaly methods focus essentially on the regionalization of residuals. The flood index methods look at a homogeneous region to construct a "global" frequency model. The intermediate method discussed here is based on a conceptual hydrological model, with some of the parameters being *explanatory variables* (slope, runoff coefficient, IDF) and others being calibration parameters validated

for the entire region or that can be regionalized (see the parameters α and β below).

This means that the *Rational Formula,* for example, is written as:

$$Q(T) = uC_r i(T,d)A \tag{9.10}$$

where $Q(T)$ is the flood discharge corresponding to return period T [m^3/s],

C_r is the runoff coefficient of the watershed [-],

$i(T,d)$ is the rainfall intensity with return period T and duration d [mm/h],

d is the duration of the rain event involved, usually considered as equal to the concentration time t_c of the watershed [min],

A is the area of the watershed [km^2],

and $u = 0.278$ is a constant used to change the measurement units.

The intensity of rainfall with return period T and duration d can be described using a Montana formula (see equation (8.14)) rewritten as:

$$i(T, d) = \varepsilon_1(T)d^{\varepsilon_2(T)} \tag{9.11}$$

where $\varepsilon_1(T)$ and $\varepsilon_2(T)$ are the parameters of the Montana relation for the given return period T.

The most delicate aspect of this technique consists of formulating a relation for the concentration time t_c (which determines the duration d of the rainfall under study). Niggli (2004) proposed the following equation (9.12), which summarizes the basics of the different forms of this equation found in the literature:

$$t_c = \beta\left(\frac{A}{p}\right)^{\alpha} \tag{9.12}$$

where p is the mean slope of the watershed [%], and

α and β are two parameters to be estimated.

Based on equations (9.11) and (9.12), Niggli (2004) expresses the equation for the Rational Method (equation(9.10)) as follows:

$$Q(T) = uC_r\beta^{\varepsilon_2(T)}p^{-\alpha\varepsilon_2(T)}\varepsilon_1(T)A^{1+\alpha\varepsilon_2(T)} \tag{9.13}$$

This model has two calibration parameters, α and β. Niggli (2004) and several other authors note that the parameter α can often be considered as a constant and propose setting it equal to 0.5. Parameter β serves to characterize the watershed and can be regionalized or used to define *homogeneous regions*. Therefore and applying this principle, Niggli (2004), for example, has divided western Switzerland into three homogeneous regions within which β can be considered as a constant: *Jura* ($\beta = 0.87$); *Plateau* ($\beta = 0.42$); *Pre-Alps* ($\beta = 0.87$). This approach, as Niggli noted

(2004), is only a preliminary and not very satisfying attempt, and needs to be reworked to be fully operational. In particular, as the parameter β obtained is the only calibration parameter, it expresses the total explained variance and it reaches a value that has no plausible physical reality.

9.4.4 Final Comment

Following this quick overview of regional analysis, we can draw the following conclusions:

- in the case of runoff, the approach will often be different depending on the size of the watersheds being studied.
- it can often be indispensable to combine the approach using *explanatory variables*, the method using the *hypothesis of regional homogeneity*, and finally, the approach based on *spatial continuity* (for example, by using a kriging technique).

It should be added that some authors have proposed Bayesian methods of estimation in order to combine the estimations from these different approaches (Niggli, 2004; Niggli and Musy, 2003).

Bibliography

Abdous B., Genest C. and Rémillard B., Dependence properties of meta-elliptical distributions, *In*: P. Duchesne and B. Rémillard (editors), *Statistical modeling and analysis for complex data problems*, GERAD 25th Anniversary Series, Springer, New York, pp. 1-15, 2005.

Adamson P.T., Metcalfe Q.V. and Parmentier B., Bivariate extreme value distributions: An application of the Gibbs sampler to the analysis of floods, *Water Resources Research*, 35(9): 2825-2832, 1999.

Ahmad M.I., Sinclair C.D. and Spurr B.D., Assessment of flood frequency models using empirical distribution function statistics, *Water Resources Research*, 24(8): 1323-1328, 1988.

Akaike H., Information theory as an extension of the maximum likelihood principle, *In*: B.N. Petrov and F. Csaki (editors), *2nd International symposium on information theory*, Budapest, Akademiai Kiado, pp. 267-281, 1973.

Ali M., Mikhail N. and Haq M., A class of bivariate distributions including the bivariate logistic, *Journal of Multivariate Analysis*, 8(3): 405-412, 1978.

Anderson R.L., *Distribution of the Serial Correlation Coefficient*, *The Annals of Mathematical Statistics*, Vol. 13: 1-13, 1942.

Bagdonavicius V., Malov S. and Nikulin M., Characterizations and parametric regression estimation in Archimedean copulas, *Journal of Applied Statistical Science*, 8(1): 137-154, 1999.

Baker V.R., Paleoflood hydrology and extraordinary flood events. *Journal of Hydrology*, 96(1-4): 79-99, 1987.

Beard L.R., Probability estimates based on small normal distribution samples, *Journal of Geophysical Research*, 65(7): 2143-2148, 1960.

Bénichou P. and Le Breton O., Prise en compte de la topograhie pour la cartograhie des champs pluviométriques statistiques, *La Météorologie*, 7ème série, 19: 23-34, 1987.

Ben-Zvi A. and Azmon B., Joint use of L-moment diagram and goodness-of-fit test: a case study of diverse series, *Journal of Hydrology*, 198(1-4): 245-259, 1997.

Berger J.O., *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York, 617 pp., 1985.

Berger J.O., *Bayesian analysis: A look at today and thoughts of tomorrow*, Technical Report 99-30, ISDS Discussion Paper, Duke University, 1999.

Bernardo J.M. and Smith A.F.M., *Bayesian Theory*. Wiley, New York, 573 pp., 1994.

Bernier J. and Veron R., Sur quelques difficultés rencontrées dans l'estimation d'un débit de crue de probabilité donnée, *Revue de Statistique Appliquée*, 12(1): 25-48, 1964.

Bernier J., *Sur les utilisations des L-Moments en hydrologie statistique*, Communication personnelle, 1993.

Bernier J., Parent E. and Boreux J.-J., *Statistique de l'Environnement. Traitement Bayésien de l'Incertitude.* Eds Tec&Doc (Lavoisier), Paris, 363 pp., 2000.

Bessis J., *La probabilité et l'évaluation des risques*, Masson, Paris, 132 pp., 1984.

Bobée B., The log-Pearson Type 3 distribution and its application to hydrology, *Water Resources Research*, 11(6): 681-689, 1975.

Bobée B. and Robitaille R., Correction of biais in the estimation of the coefficient of skewness, *Water Resources Research*, 33(12): 2769-2780, 1975.

Bobée B. and Degroseillers L., *Ajustement des distributions Pearson type 3, gamma, gamma généralisée et log-Pearson*, INRS-Eau, rapport de recherche 105, 63 pp., 1985.

Bobée B. and Ashkar F., *Sundry Averages Methods (SAM) for estimating parameters of the log-Pearson type 3 distribution*, INRS-Eau, rapport de recherche R-251, 28 pp., 1988.

Bobée B. and Ashkar F., *The Gamma family and derived distributions applied in hydrology*, Water Resources Publications, Littleton, 203 pp., 1991.

Bois P., *Contrôle des séries chronologiques par étude du cumul des résidus*, Colloque et séminaires ORSTOM, pp. 89-99, 1986.

Bouyé E., Durrleman V., Nikeghbali A., Riboulet G. and Roncalli T., *Copulas for finance - A reading guide and some applications*, rapport technique, Groupe de recherche opérationnelle, Crédit Lyonnais, Paris, 70 pp., 2000.

Bozdogan H., Model selection and Akaike's information criterion (AIC): The general theory and its analytical extentions, *Psychometrika*, 52(3): 345-370, 1987.

Breymann W., Dias A. and Embrechts P., Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance*, 3(1): 1-14, 2003.

Brunet-Moret Y., Statistiques de rangs, *Cahiers ORSTOM, Série hydrologie*, 10(2): 133-151, 1973.

Brunet-Moret Y., Homogénéisation des précipitations, *Cahiers ORSTOM, Série hydrologie*, 16(3-4): 147-170, 1979.

Burnham K.P. and Anderson D.R., *Model selection and multimodel inference: A practical information-theoretic approach*, 2nd edition, Springer, New York, 488 pp., 2002.

Castellarin A., Vogel R.M. and Matalas N.C., Multivariate probabilistic regional envelopes of extreme floods, *Journal of Hydrology*, 336(3-4): 376-390, 2007.

Cave B.M. and Pearson K., Numerical illustrations of the variate difference correlation method, *Biometrika*, 10(2-3): 340-355, 1914.

CEA, *Statistique appliquée à l'exploitation des mesures*, Tome 1, Masson, 296 pp., Paris, 1978.

Chow V.T., The log-probability law and its engineering applications, *Proceedings of the American Society of Civil Engineers*, 80(536): 1-25, 1954.

Chow V.T. (editor), *Handbook of applied hydrology*, McGraw-Hill, New York, 1468 pp., 1964.

Clayton D.G., A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, 65(1): 141-151, 1978.

Coles S.G. and Tawn J.A., A Bayesian analysis of extreme rainfall data, *Applied Statistics*, 45(4): 463-478, 1996.

Coles S., *An introduction to statistical modeling of extreme values*, Springer, London, 224 pp., 2001.

Cook D. and Johnson M., A family of distributions for modeling non-elliptically symmetric multivariate data, *Journal of the Royal Statistical Society: Series B*, 43(2): 210-218, 1981.

CTGREF (Centre Technique du Génie Rural, des Eaux et Forêts), *Utilisation de quelques tests en hydrologie et calcul d'intervalles de confiance*, Informations techniques, Cahier 31, 2, Groupement d'Antony, Division hydrologie - hydraulique fluviale, 4 pp., 1978.

Cuadras C.M. and Augé J., A continuous general multivariate distribution and its properties, *Communications in Statistics - Theory and Methods*, 10(4): 339-353, 1981.

Cunderlik J.M. and Burn D.H., Non-stationary pooled flood frequency analysis, *Journal of Hydrology*, 276(1-4): 210-223, 2003.

Cunnane C., Unbiased plotting positions - A review, *Journal of Hydrology*, 37(3-4): 205-222, 1978.

Cunnane C., *Statistical distributions for flood frequency analysis*, World Meteorological Organization, Operational hydrology report 33, WMO 718, 115 pp., Geneva, 1989.

Danielsson J. and De Vries C.G., Tail index and quantile estimation with very high frequency data, *Journal of Empirical Finance*, 4: 241-257, 1997.

Danielsson J., de Haan L., Peng L. and de Vries C.G., Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation, *Journal of Multivariate Analysis* 76: 226-248, 2001.

Dalrymple T., *Flood frequency analysis*, U.S. Geological Survey, Water supply paper 1543 A, 80 pp., 1960.

Dauphiné A., *Risques et catastrophes*, Armand Colin, Paris, 288 pp., 2001.

Davison A.C. and Smith R.L., Models for exceedances over high thresholds (with discussion), *Journal of the Royal Statistical Society* 52: 393-442, 1990.

Davison A.C. and Hinkley D.V., *Bootstrap methods and their application*, Cambridge series in statistical and probabilistic mathematics, 1, Cambridge University Press, Cambridge, 592 pp., 1997.

Davison A.C. and Ramesh N.I., Local likelihood smoothing of sample extremes, *Journal of the Royal Statistical Society: Series B*, 62(1): 191-208, 2000.

De Michele C., Salvadori G., Canossi M., Petaccia A. and Rosso R., Bivariate statistical approach to check adequacy of dam spillway, *Journal of Hydrologic Engineering*, 10(1): 50-57, 2005.

De Montmollin F., O*livier* R. and Zwahlen F., Utilisation d'une grille d'altitudes digitalisées pour la cartographie d'éléments du bilan hydrique, *Journal of Hydrology*, 44(3-4): 191-209, 1979.

De Souza P.B., Consuegra D. and Musy A., CODEAU: A database package for the assessment and analysis of hydrometeorological data, *In*: A. Verwey, A.W. Minns, V. Babović and Č. Maksimović (editors), *Hydroinformatics '94, Proceeding of the 1st International conference on hydroinformatics*, Balkema, Rotterdam, pp. 103-110, 1994.

El Adlouni S., Ouarda T.B.M.J. and Bobée B., Orthogonal Projection L-moment estimators for three-parameter distributions, *Advances and Applications in Statistics*, sous presse, 2007.

Embrechts P., Klüppelberg C. and Mikosch T., *Modelling extremal events for insurance and finance*, Springer, Berlin, 655 pp., 1997.

Embrechts P., McNeil A. and Straumann D., Correlation and dependence in risk management: Properties and pitfalls, *In*: M. Dempster (editor), *Risk management: Value at risk and beyond*, Cambridge University Press, Cambridge, pp. 176-223, 2002.

Embrechts P., Lindskog F. and McNeil A., Modelling dependence with copulas and applications to risk management, *In*: S. Rachev (editor), *Handbook of heavy tailed distributions in finance*, Elsevier, New York, chapter 8, pp. 329-384, 2003.

Fang H.B., Fang K.T. and Kotz S., The meta-elliptical distributions with given marginals, *Journal of Multivariate Analysis*, 82(1): 1-16, 2002.

Favre A.C., Musy A. and Morgenthaler S., Two-site modeling of rainfall based on the Neyman-Scott process, *Water Resources Research*, 38(12): 1307, doi:10.1029/2002WR001343, 2002.

Favre A.C., El Adlouni S., Perreault L., Thiémonge N. and Bobée B., Multivariate hydrological frequency analysis using copulas, *Water Resources Research*, 40: W01101, doi:10.1029/2003WR002456, 2004.

Fermanian J.D., Goodness-of-fit tests for copulas, *Journal of Multivariate Analysis*, 95(1): 119-152, 2005.

Fortin V., Bernier J. and Bobée B., Rapport final du projet C3: *Détermination des crues de conception*, Chaire en hydrologie statistique, INRS-Eau, rapport R-532, 1998.

Frank M.J., On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$, *Aequationes Mathematicae*, 19(1): 194-226, 1979.

Frees E.W. and Valdez E.A., Understanding relationships using copulas, *North American Actuarial Journal*, 2(1): 1-25, 1998.

Galambos J., Order statistics of samples from multivariate distributions, *Journal of the American Statistical Association*, 70(3): 674-680, 1975.

Garrido M., *Modélisation des événements rares et estimation des quantiles extrêmes : méthodes de sélection de modèles pour les queues de distribution*, Thèse de doctorat, Université Joseph Fourier - Grenoble I, Grenoble, 231 pp., 2002.

Gelman A., Carlin J.B., Stern H.S and Rubin D.B., *Bayesian Data Analysis*, 2nd Edition, Chapman and Hall, Londres, 698 pp., 2004.

Genest C. and MacKay R., The joy of copulas: Bivariate distribution with uniform marginals, *The American Statistician*, 40: 280-283, 1986.

Genest C., Frank's family of bivariate distributions, *Biometrika*, 74(3): 549-555, 1987.

Genest C. and Rivest L.P., Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association*, 88(423): 1034-1043, 1993.

Genest C. and Ghoudi K., Une famille de lois bidimensionnelles insolite, *Comptes Rendus de l'Académie des Sciences de Paris*, 318(1): 351-354, 1994.

Genest C., Ghoudi K. and Rivest L.P., A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, 82(3): 543-552, 1995.

Genest C., Quessy J.F. and Rémillard B., Goodness-of-fit procedures for copula models based on the probability integral transformation, *Scandinavian Journal of Statistics*, 33(2): 337-366, 2006.

Genest C. and Favre A.C., Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, 12(4): 347-368, 2007.

Genest C. and Rémillard B., Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, *Annales de l'Institut Henri Poincaré*, 37, sous presse, 2007.

Genest C., Favre A.C., Béliveau J. and Jacques C., Meta-elliptical copulas and their use in frequency analysis of multivariate hydrological data, *Water Resources Research*, 43, W09401, doi:10.1029/2006WR005275, 2007.

Gibbons J.D. and Chakraborti S., *Nonparametric statistical inference*, 4th edition, Statistics: A Dekker series of textbooks and monographs, Marcel Dekker, New York, 680 pp., 2003.

Grandjean P., *Choix et élaboration des données de pluie pour l'assainissement rural et urbain dans le canton de Genève*, VSA, Journée « Niederschlag und Siedlungsentwäasserung », Zürich, 1988.

Greenwood J.A., Landwehr J.M., Matalas N.C. and Wallis J.R., Definition and relation to parameters of several distributions expressable in inverse form, *Water Resources Research*, 15(5): 1049-1054, 1979.

GREHYS (Groupe de Recherche en Hydrologie Statistique), Presentation andreview of some methods for regional flood frequency analysis, *Journal of Hydrology*, 186(1-4): 63-84, 1996.

Guillot P. and Duband D., La méthode du GRADEX pour le calcul de la probabilité des crues rares à partir des pluies, *In*: IAHS (editor), *Les crues et leur évaluation*, International Association of Hydrological Sciences, 84, pp. 560-569, 1967.

Guillot P., Les arguments de la méthode du gradex, base logique pour évaluer les crues extrêmes, *Symposium IDNDR*, Oklahoma, 1994.

Gumbel E.J., *Statistics of extremes*, Columbia University Press, New York, 375 pp., 1958.

Gumbel E.J., Bivariate exponential distributions, *Journal of the American Statistical Association*, 55(292): 698-707, 1960.

Gunasekara T.A.G. and Cunnane C., Expected probabilities of exceedance for non-normal flood distributions, *Journal of Hydrology*, 128(1-4): 101-113, 1991.

Harr M.E., *Lectures on reliability in civil engineering*, Swiss Federal Institute of Technology, Lausanne, 1986.

Harr M.E., Probabilistic estimates for multivariate analysis, *Appied Mathematical Modelling*, 13(5): 313-318, 1989.

Hiez G., L'homogénéité des données pluviométriques, *Cahiers ORSTOM, Série hydrologie*, 14(2): 129-172, 1977.

Hill, B.M., A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3(5): 1163-1174, 1975.

Hosking J.R.M., L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society: Series B*, 52(1): 105-124, 1990.

Hosking J.R.M. and Wallis J.R., Some statistics useful in regional frequency analysis, *Water Resources Research*, 29(2): 271-281, 1993.

Hörler A., *Die Intensitäten von Starkregen längerer Dauer für verschiedene ortschaften der Schweiz*, Gaz-Wasser-Abwasser, 57. Jahrgang, 12: 853-860, 1977.

Hougaard P., A class of multivariate failure time distributions, *Biometrika*, 73(3): 671-678, 1986.

Hüsler J. and Reiss R.D., Maxima of normal random vectors: between independence, and complete dependence, *Statistics and Probability Letters*, 7(4): 283-286, 1989.

Iman R.L. and Conover W.J., *A modern Approach to Statistics*, Wiley, New-York, 1983.

Joe H., Parametric families of multivariate distributions with given margins, *Journal of Multivariate Analysis*, 46(2): 262-282, 1993.

Joe H., *Multivariate models and dependence concepts*, Chapman and Hall, London, 399 pp., 1997.

Jordan J.P. and Meylan P., *Estimation spatiale des précipitations dans l'ouest de la Suisse par la méthode du krigeage*, IAS, 12: 157-162, 1986a.

Jordan J.P. and Meylan P., *Estimation spatiale des précipitations dans l'ouest de la Suisse par la méthode du krigeage*, IAS, 13: 187-189, 1986b.

Katz R.W., Parlange M.B. and Naveau P., Statistics of extremes in hydrology, *Advances in Water Resources*, 25(8-12): 1287-1304, 2002.

Kendall M.G., A new measure of rank correlation, *Biometrika*, 30(1-2): 81-93, 1938.

Kendall M.G., Stuart A., Ord K. and Arnold S., *Kendall's advanced theory of statistics*, Volume 2A: Classical inference and the linear models, 6th edition, Arnold, London, 448 pp., 1999.

Khaliq M.N., Ouarda T.B.M.J., Gachon P. and Bobée B., Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, *Journal of Hydrology*, 329(3-4): 534-552, 2006.

Kimeldorf G. and Sampson A.R., Uniform representations of bivariate distributions. *Communications in Statistics: Theory and Methods*, 4(7): 617-627, 1975.

Kite G.W., Confidence limits for design event, *Water Resources Research*, 11(1): 48-53, 1975.

Kite G.W., *Frequency and risk analysis in hydrology*, Water Resources Publications, Littelton, 224 pp., 1988.

Klemeš V., Dilletantism in hydrology: Transition or destinity?, *Water Resources Research*, 22(9): 177S-188S, 1986.

Klemeš V., Tall tales about tails of hydrological distributions. II, *Journal of Hydrologic Engineering*, 5(3): 232-239, 2000.

Koenker R. and Basset G., Regression quantiles, *Econometrica*, 46(1): 33-50, 1978.

Koenker R. and D'Orey V., Computing regression quantiles, *Applied Statistics*, 36(3): 383-393, 1987.

Koutsoyiannis D., Kozonis D. and Manetas A., a mathematical framework for studying rainfall intensity-duration-frequency relationships, *Journal of Hydrology*, 206: 118-135, 1998.

Kuczera G., Combining site-specific and regional information: an empirical Bayes approach, *Water Resources Research*, 18(2): 306-314, 1982.

Kuczera G., Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference, *Water Resources Research*, 34(6): 1551-1558, 1999.

Landwehr J.M. and Matalas N.C., Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resources Research*, 15(5): 1055-1064, 1979.

Lang M., *Communication personnelle*, CEMAGREF, Groupement de Lyon, 1996.

Langbein W.B., Annual floods and the partial-duration flood series, *Transactions of the American Geophysical Union*, 30(6): 879-881, 1949.

Larras J., *Prévision et prédétermination des crues et des étiages*, Eyrolles, Paris, 160 pp., 1972.

Lebart L., Morineau A. and Fenelon J.P., *Traitement des données statistiques*, 2ème édition, Dunod, Paris, 510 pp., 1982.

Lebel T. and Boyer J.F., *DIXLOI : Un ensemble de programmes Fortran 77 pour l'ajustement de lois statistiques et leur représentation graphique*, ORSTOM, Laboratoire d'hydrologie, Montpellier, 55 pp., 1989.

Linsley R.K., Flood estimates: How good are they?, *Water Resources Research*, 22(9): 159S-164S, 1986.

Mallows C.L., Some comments on C_p, *Technometrics*, 15(4): 661-675, 1973.

Masson J.M., *Méthode générale approchée pour calculer l'intervalle de confiance d'un quantile. Application à quelques lois de probabilité utilisées en hydrologie*, Note interne 6/1983, Laboratoire d'hydrologie et modélisation, Université Montpellier II, 41 pp., 1983.

Matheron G., *Estimer et choisir. Essai sur la pratique des probabilités*, Cahiers du Centre de morphologie mathématique de Fontainebleau, 7, Ecole Nationale Supérieure des Mines de Paris, ENSMP, Fontainebleau, 175 pp., 1978.

McNeil A.J. and Saladin T., Developing scenarios for future extreme losses using the POT method, *In*: P. Embrechts (editor), *Extremes and integrated risk management*, Risk Books, London, 397 pp., 2000.

Meylan P., Régionalisation de données entachées d'erreurs de mesure par krigeage, *Hydrologie Continentale*, 1(1): 25-34, 1986.

Meylan P. and Musy A., *Hydrologie fréquentielle*, HGA, Bucarest, 413 pp., 1999.

Meylan P., Grandjean P. and Thöni M., *Intensité des pluies de la région genevoise – Directive IDF 2001*, Etat de Genève, DIAE, DomEau, SECOE, 4 pp., 2005.

Meylan P., *A.F.P.C. – Analyse fréquentielle par comptage*, République et canton de Genève – Département de l'intérieur et de la mobilité – Direction générale de l'eau, AIC, Lausanne, février 2010.

Michel C., *Hydrologie appliquée aux petits bassins ruraux*, CEMAGREF, Division hydrologie, hydraulique fluviale et souterraine, Antony, 528 pp., 1989.

Miquel J., *Guide pratique d'estimation des probabilités de crues*, Eyrolles, Paris, 160 pp., 1984.

Morel-Seytoux H.J., Forecasting of flow-flood frequency analysis, *In*: H.W. Shen (editor), *Modeling of rivers*, chapter 3, Wiley, New York, 1979.

Morgenthaler S., *Introduction à la statistique*, Méthodes mathématiques pour l'ingénieur, 9, 2ème édition, Presses polytechniques et universitaires romandes, Lausanne, 328 pp., 2001.

Morlat G., Les lois de probabilités de Halphen, *Revue de statistique appliquée*, 4(3): 21-43, 1956.

Musy A. and Higy C., *Hydrology*: *A Science of Nature,* Science Publishers, Enfield (USA), 326 pp., 2010.

Naulet R., *Utilisation de l'information des crues historiques pour une meilleure prédétermination du risque d'inondation. Application au bassin de l'Ardèche à Vallon Pont-d'Arc et Saint Martind' Ardèche*. PhD thesis, Université Joseph Fourier - Grenoble 1 (France) et INRS-Eau Terre et Environnement (Québec), 2002.

Nelsen R.B., Properties of a one-parameter family of bivariate distributions with specified marginals, *Communications in Statistics: Theory and Methods*, 15(11): 3277-3285, 1986.

Nelsen R.B., *An introduction to copulas*, Lecture notes in statistics, 139, 2nd edition, Springer, New York, 269 pp., 2006.

NERC (Natural Environment Research Council), *Flood studies report*, Volume I: Hydrological studies, NERC, London, 1975.

NERC (Natural Environment Research Council), *Review of regional growth curves*, Flood studies supplementary report 14, London, 6 pp., 1983.

Neykov N.M., Neytchev P.N., Van Gelder P.H.A.J.M. and Todorov V.K., Robust detection of discordant sites in regional frequency analysis, *Water Resources Research*, 43: W06417, doi:10.1029/2006WR005322, 2007.

Niggli M. and Musy A., A Bayesian combination method of different models for flood regionalisation, *In*: *Proceedings of the 6th Inter-regional conference on environment-water, Land and water use planning and management*, Albacete, 2003.

Niggli M., *Combinaison bayésienne des estimations régionales des crues : concept, développement et validation*, Thèse de doctorat 2895, EPFL, Lausanne, 229 pp., 2004.

Oakes D., A model for association in bivariate survival data, *Journal of the Royal Statistical Society: Series B*, 44(3): 414-422, 1982.

Oakes D., Multivariate survival distributions, *Journal of Nonparametric Statistics*, 3(3-4): 343-354, 1994.

O'Connell D.R.H., Osteena D.A., Levish D.R. and Klinger R.E., Bayesian flood frequency analysis with paleohydrologic bound data. *Water Resources Research*, 38(5), 1058, doi:10.1029/2000WR000028, 2002.

OFEFP (Office fédéral de l'environnement, des forêts et du paysage), *Manuel I de l'ordonnance sur les accidents majeurs OPAM*, OFEFP, Berne, 74 pp., 1991a.

OFEFP (Office fédéral de l'environnement, des forêts et du paysage), *Débits de crue dans les cours d'eau suisses*, Communications hydrologiques, 3 et 4, OFEFP, Berne, 1991b.

OFEG (Office fédéral des eaux et de la géologie), *Protection contre les crues des cours d'eau – Directives 2001*, Bienne, OFEG, 72 pp., 2001.

OFEG (Office fédéral des eaux et de la géologie), *Evaluation des crues dans les bassins versants de Suisse - Guide pratique*, Rapports de l'OFEG, Série Eaux, 4, OFEG, Berne, 114 pp., 2003.

Ouarda T., Rasmussen P.F., Bobée B. and Bernier J., Utilisation de l'information historique en analyse hydrologique fréquentielle, *Revue des sciences de l'eau*, 11(1): 41-49, 1998.

Ouarda T.B.M.J., Cunderlik J., St-Hilaire A., Barbet M., Bruneau P. and Bobée B., Data-based comparison of seasonality-based regional flood frequency methods, *Journal of Hydrology*, 330(1-2): 329-339, 2006.

Parent E. and Bernier J., Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. *Journal of Hydrology*, 283(1-4): 1-18, 2003.

Parent E. and Bernier J., Bayesian POT modeling for historical data. *Journal of Hydrology*, 274(1-4): 95-108, 2003.

Parent E. and Bernier J., *Le raisonnement bayésien. Modélisation et inférence*. Springer, Statistique et probabilités appliquées, Paris, 364 pp., 2007.

Payrastre O., *Faisabilité et utilité du recueil de données historiques pour l'étude des crues extrêmes de petits cours d'eau. Etude du cas de quatre bassins versants affluents de l'Aude*. Thèse. École nationale des ponts et chaussées (France), 2005.

Plackett R.L., A class of bivariate distributions, *Journal of the American Statistical Association*, 60(2): 516-522, 1965.

Peck E.L. and Brown M.J., An approach to the development of isohyetal maps for montainous areas, *Journal of Geophysical Research*, 67(2): 681-694, 1962.

Perreault L., *Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques*. Thèse de doctorat en statistique appliquée, ENGREF, Paris, 2000.

Perreault L., *Modélisation des débits sortants à Chute-des-Passes : application du mélange de distributions normales*, Rapport Institut de Recherche d'Hydro-Québec. IREQ-2002-127C, confidentiel, 41 pp, 2003.

Pickands III J., Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics*, 3(1): 119-131, 1975.

Poulin A., Huard D., Favre A.C. and Pugin S., Importance of tail dependence in bivariate frequency analysis, *Journal of Hydrologic Engineering*, 12(4): 394-403, 2007.

Proschan F., Confidence and tolerance intervals for the normal distribution, *Journal of the American Statistical Association*, 48(263): 550-564, 1953.

Pugachev V.S., *Théorie des probabilités et statistique mathématique*, MIR, Moscou, 471 pp., 1982.

Ramesh N.I. and Davison A.C., Local models for exploratory analysis of hydrological extremes, *Journal of Hydrology*, 256(1-2): 106-119, 2002.

Rao D.V., Estimating log Pearson parameters by mixed moments, *Journal of Hydraulic Engineering*, 109(8): 1118-1132, 1983.

Reiss R.D. and Thomas M., *Statistical analysis of extreme values*, 2nd edition, Birkhäuser, Bâle, 443 pp., 2001.

Restle E.M., El Adlouni S., Bobée B. and Ouarda T.B.M.J., *Le test GPD et son implémentation dans le logiciel HYFRAN PRO*, INRS-ETE, rapport de recherche R-737, 35 pp., 2004.

Ribatet M., Sauquet E., Grésillon J.M. and Ouarda T.B.M.J., A regional Bayesian POT model for flood frequency analysis, *Stochastic Environmental Research and Risk Assessment*, 21(4): 327-339, 2007.

Rissanen J., Modeling by shortest data description, *Automatica*, 14(5): 465-471, 1978.

Robert C.P. and Casella G., *Monte Carlo Statistical Methods*, 2nd Edition, Springer-Verlag, New York, 635 pp., 2004.

Rosenblueth E., Point estimates for probability moments, *Proceedings of the National Academy Sciences of the United States of America*, 72(10): 3812-3814, 1975.

Ross S.-M., *An Introduction to Probability Models*, 10th Edition, Academic Press, Elsevier, Burlington, USA, 784 pp., 2010.

Sachs L., *Applied statistics: A handbook of techniques*, 2nd edition, Springer, New York, 706 pp., 1984.

Salvadori G. and De Michele C., Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resources Research*, 40: W12511, doi:10.1029/2004WR003133, 2004.

Sankarasubramanian A. and Lall U., Flood quantiles in a changing climate: Seasonal forecasts and causal relations, *Water Resources Research*, 39(5): 1134, doi:10.1029/2002WR001593, 2003.

Saporta G., *Probabilités, analyse des données et statistique*, Technip, Paris, 493 pp., 1990.

Schwarz G., Estimating the dimension of a model, *The Annals of Statistics*, 6(2): 461-464, 1978.

Schweizer B. and Wolff E., On nonparametric measures of dependence for random variables, *The Annals of Statistics*, 9(4): 879-885, 1981.

Sherrod Ph.-H., *NLREG – Nonlinear regression analysis Program*, Brentwood, 2002.

Shih J.H. and Louis T.A., Inferences on the association parameter in copula models for bivariate survival data, *Biometrics*, 51(4): 1384-1399, 1995.

Siegel S., *Nonparametric statistics for the behavioral sciences*, 2nd edition, McGraw-Hill, New York, 399 pp., 1988.

Sklar A., Fonctions de répartition à *n* dimensions et leurs marges, *Publications de l'Institut de statistique de l'Université de Paris*, 8: 229-231, 1959.

Sneyers R., *Sur l'analyse statistique des séries d'observations*, Note technique 143, OMM, Genève, 192 pp., 1975.

Sokolov A.A., Rantz S.E. and Roche M.F., *Floodflow computation: Methods compiled from world experience*, Studies and reports in hydrology, 22, UNESCO Press, Paris, 294 pp., 1976.

Solaiman B., *Processus stochastiques pour l'ingénieur*, Collection technique et scientifique des télécommunications, Presses polytechniques et universitaires romandes, Lausanne, 241 pp., 2006.

Stedinger J.R., Vogel R.M. and Foufoula-Georgiou E., Frequency analysis of extreme events, chapter 18, *In*: D.R. Maidment (editor), *Handbook of Hydrology*, McGraw-Hill, New York, 1424 pp., 1993.

Stephens M.A., EDF statistics for goodness of fit and some comparisons, *Journal of the American Statistical Association*, 69(347): 730-737, 1974.

Strupczewski W.G. and Kaczmarek Z., Non-stationary approach to at-site flood frequency modelling II. Weighted least squares estimation, *Journal of Hydrology*, 248(1-4): 143-151, 2001.

Strupczewski W.G., Singh V.P. and Feluch W., Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation, *Journal of Hydrology*, 248(1-4): 123-142, 2001a.

Strupczewski W.G., Singh V.P. and Mitosek H.T., Non-stationary approach to at-site flood frequency modelling III. Flood analysis of Polish rivers, *Journal of Hydrology*, 248(1-4): 152-167, 2001b.

Sturges H., The choice of a class-interval. *Journal of the American Statistical Association*, 21: 65-66, 1926.

Takeuchi K., Annual maximun series and partial-duration series — Evaluation of Langbein's formula and Chow's discussion, *Journal of Hydrology*, 68(1-4): 275-284, 1984.

Tawn J.A., Bivariate extreme value theory: Models and estimation, *Biometrika*, 75(3): 397-415, 1988.

Thirriot C., Quoi de neuf depuis Noé ? *In*: *23èmes Journées de l'hydraulique*, Rapport général, Société hydrotechnique de France, Nîmes, 1994.

Ventsel H., *Théorie des probabilités*, MIR, Moscou, 563 pp., 1973.

Viglione A., Laio F. and Claps P., A comparison of homogeneity tests for regional frequency analysis, *Water Resources Research*, 43: W03428, doi:10.1029/2006WR005095, 2007.

Vogel R.M. and Fennessey N.M., L moment diagrams should replace product moment diagrams, *Water Resources Research*, 29(6): 1745-1752, 1993.

VSS (Association suisse des professionnels de la route et des transports), *Evacuation des eaux de chaussée - Intensité des pluies*, Norme SN 640 350, Zürich, 8 pp., 2000.

Wald A. and Wolfowitz J., An Exact Test for Randomnesss in the Non-Parametric Case based on serial Correlation, *The Annals of Mathematical Statistics*, Vol. 14: 378-388, 1943.

Wallis J.R., Catastrophes, computing, and containment: Living in our restless habitat, *Speculation in Science and Technology*, 11(4): 295-315, 1988.

Weiss L.L., Ratio of true to fixed interval maximum rainfall, *Journal of the Hydraulics Division ASCE*, 90(1): 77-82, 1964.

WMO (World Meteorological Organisation), Guide to hydrological practices Volume II : Analysis, forecasting and other applications, 168, OMM, Geneva, 304 pp., 1981.

WRC (U.S. Water Resources Council), *Guidelines for determining flood flow frequency*, Hydrology committee, WRC, Washington, 1967.

Yevjevitch V., *Probability and statistics in hydrology*, Water Resources Publications, Fort Collins, 302 pp., 1972.

Yue S., Ouarda T.B.M.J. and Bobée B., A review of bivariate gamma distribution for hydrological application, *Journal of Hydrology*, 246(1-4): 1-18, 2001.

Yue S. and Rasmussen P., Bivariate frequency analysis: Discussion of some useful concepts in hydrological application, *Hydrological Processes*, 16(14): 2881-2989, 2002.

Zhang X., Zwiers F.W. and Li G., Monte Carlo experiments on the detection of trends in extreme values, *Journal of Climate*, 17(10): 1945-1952, 2004.