PROBABILISTIC FLOW DURATION CURVES FOR USE IN ENVIRONMENTAL PLANNING AND MANAGEMENT

Pierluigi CLAPS and Mauro FIORENTINO

Dipartimento di Ingegneria e Fisica dell'Ambiente
Università della Basilicata
Via della Tecnica, 3, 85100 Potenza, Italy

Abstract. When dealing with the planning and management of water uses in a river, the knowledge of the probabilistic structure of minima of daily or of multiple-day flows is often required. As a tool for straightforward determination of different levels of flow minima, flow duration curves (FDCs) are particularly suited for planning purposes. In this paper, FDCs referred to annual samples are interpolated with lognormal curves and their probabilistic structure is obtained through the statistical analysis of the two lognormal parameters. Distribution of these parameters is shown to be normal, so that the discharge with a given duration related to a return period T can be easily evaluated. To build FDCs in ungauged basins, relations between the moments of the parameters and catchment characteristics have been investigated, with reference to the data available in the Basilicata region (Italy). For both parameters, most of the variance of the first moment can be explained by the Base Flow Index (BFI), which can be estimated from geology. The second moment can be derived considering that the coefficient of variation is constant over the whole region. Since the curves are considered in dimensionless form, estimation of the mean annual runoff is finally needed to obtain the dimensional probabilistic FDCs in ungauged sites.

1. Introduction

Availability of analysis and simulation tools for the flow regime, and the low flows in particular, is fundamental when dealing with the planning and management of water uses in a river. In most cases, particularly to ensure respect of the law regulations concerning water quality, evaluations of low flow indices with respect to a given return period are sufficient. For instance, one of the main concerns in environmental management is related to the compatibility of the remaining instream flow with the life of the habitat. Most of the industrialized countries have promulgated regulations concerning the evaluation and the respect of the minimum instream flow for habitat preservation. The evaluation of this characteristic discharge should be obtained with the help of ecologists and biologists, but some rules have been proposed for its selection within procedures for the determination of quality indexes (*e.g.* [1], [2]). In practice, this rules require the knowledge of the probabilistic structure of minima of daily or (more often) weekly flows.

Statistical analysis of low flows does not present particular problems if a minimum amount of data is available. However, evaluation of low flow indices are often needed in ungauged sites, and then the analysis becomes quite difficult.

Literature on the statistical modeling of flow minima presents abundance of models, (see *e.g.* [8]), even because several definitions apply to the low flow concepts. The great number of existing models and of definitions of low flows are due to the interaction between the nature of the process and the objectives of the study. *Gustard* [9] classified low flow measures based on (a) different definitions of a low flow event (*e.g.* discharge or volume below a threshold, run length), (b) different expressions of frequencies (*e.g.* proportion of time, as for flow duration curves, or proportion of years, as for flow frequency curves), (c) different aggregation intervals (*i.e.* average daily or weekly or monthly flows).

Identification of the correct low flow measure for the problem at hand is certainly a nontrivial task, since the use of quite different class of models can be required as a consequence of the choice. Moreover, the issues arising in the environmental planning and management are sometimes very different from one another and this introduces the need for flexible modeling tools, which can give a viewpoint of the low flow process, with respect to the related return period, in its evolution over time.

This paper presents an attempt to fill a gap in this area, through the introduction of the return-period viewpoint in the flow duration curves (FDCs), which already describe river flows according to the proportion of time of their availability. Advantages in using flow duration curves derive by the nature itself of the tool, because of the possibility of investigating the proportion of time in a year in which a discharge is exceeded.

This can be particularly useful, for instance, in cases in which there are difficulties in the modulation of the wastewater treatment or industrial effluents and the emphasis is on the duration of periods with ‘high’ concentration of pollutants. Another, more significant, case in which the indication of a single minimum instream flow value is clearly inadequate occurs when concurring uses of the water resource operate on the same water body. These situations originated revised criteria, such as the Instream Flow Incremental Methodology (see [15]), for the determination of scenarios of water resource uses compatible with the physical habitat.

A major issue to face is when there is no data available in the river section of interest. Techniques for estimating the low flow regime in ungauged sites (see [8] for a recent review) necessarily require that a certain low flow variable is selected. This is one of the main reasons for putting emphasis on some ‘typical’ low flow indices, such as Q(95) (95% percentile from the FDC) and on $Q_{7,T}$ (minimum 7-day average discharge with a return period of T years) with T usually assuming the values 2 and 10.

An impressive amount of work, done sometimes with reference to a whole country, is available in literature as concerns regionalization of such indices. However, where the single ‘characteristic’ minimum is inadequate as a support to decisions, these efforts can result insufficient. Our suggestion in this field, which is presented in the second part of the paper, goes in the direction of regionalizing the whole flow duration curve, which contain a greater amount of information. This approach can somewhat increase the overall uncertainty, but gives answers to a wider range of problems.

2. Probabilistic approach to Flow Duration Curves

Flow duration curves are usually defined [10] as a function of discharge indicating the percentage of time that the discharge is exceeded. This definition does not refer to a particular time span of data nor to the time of aggregation of discharge, even though the average daily discharge is usually considered. However, other definitions for the FDC can be given. For instance, in the Italian literature, a more strict reference to a time span of a year and a daily discretization is usually assumed (*e.g.* [6], [7]), so that the actual duration, in days of the year, is used as the index of exceedance.

In all of the literature references to the duration curves they are treated as deterministic objects, reproducing the features of the flow regimes of a river in a given section. If the percentage is considered as a fraction of the year, the most common way to intend the duration curve is to link each discharge to the expected fraction of the year in which the discharge is exceeded.

In this paper, in view of the institution of a probabilistic approach to FDCs, the frequency viewpoint is adopted, and the generic FDC is referred to one year of data of average daily discharge.

In the number of analytic representation proposed for the duration curves (*e.g.* [10],[7]) the lognormal representation appears as the most reasonable one, even because it corresponds to the simplest probability distribution that can be used for representation of daily flows (see *e.g.* [18]). The three-parameter lognormal distribution can be adopted in general, according to the relation

$$log(q - q_0) = \alpha + \beta z \qquad (1)$$

in which q represents the dimensionless daily discharge, *i.e.* $q=Q/E[Q]$ where $E[Q]$ is the average discharge in the year considered, q_0 is an adjustment parameter with the meaning of a discharge lower bound, and z is the normal reduced variate representing intrinsically the frequency of the discharge q. In this analysis, however, a two-parameter lognormal representation is adopted (implying $q_0 = 0$), even because in Mediterranean climates it is quite difficult to find a discharge lower bound.

One of the problems which often appear, with reference to the analytical representation of FDCs, is the presence of zero-flows. In this case, the analytical description is limited to the non-zero values and an additional parameter, that is the frequency of zeros, is to be taken under consideration. This problem equally affects the analysis based on the construction of probabilistic models of flow minima, and its handling is the subject of papers usually found in the field of the hydrology of semi-arid regions**.**

Comparing (1) with the definition of Normal reduced variate $z = [log(q) - \mu] / \sigma$, it is evident that α and β are estimates of the first two moments of the sample, that is:

$$\alpha = \mu[log(q)]; \quad \beta = -\sigma[log(q)], \tag{2}$$

and, consequently, their computation is immediate. An example of determination of interpolating lognormal curve is shown in Figure (1).

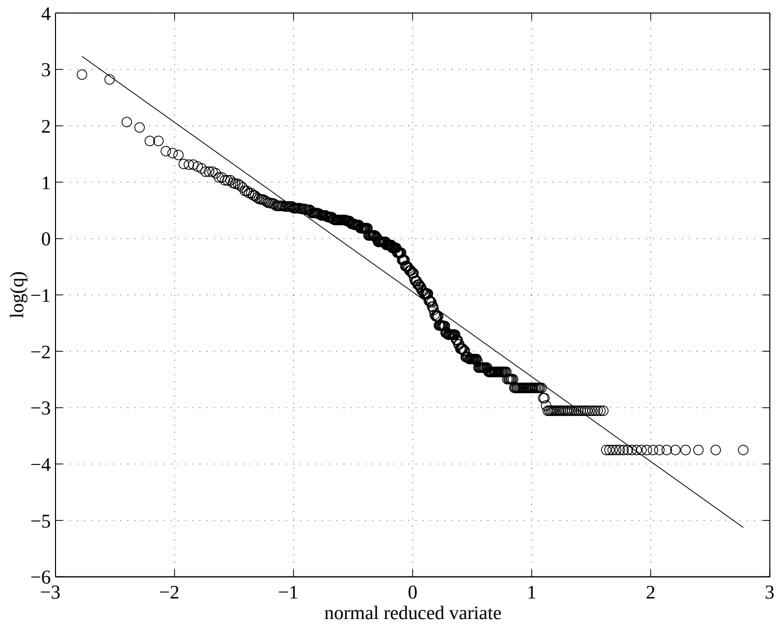

Figure 1. Station 2.1: flow duration curve and lognormal interpolating curve for the year of minimum annual runoff

The above representation is referred to one year of data. Therefore, parameters α and β are uniquely referred to the year at hand. Considering that relation (1) varies from one year to another, a general probabilistic representation can be adopted in the form

$$X_{F,\phi} = \alpha_\phi + \beta_\phi \; z(F) \tag{3}$$

with $X=log(q)$ and F corresponding to the "duration" of q. F is the frequency, expressed as $F=n/(365+1)$, where n is the actual duration of q in terms of number of days in which q is exceeded. ϕ is the probability of non-exceedance of the value X_F among years.

Once the probabilistic flow duration curve (3) is defined, its actual determination relies of the distribution of parameters α and β. Analysis of this issue, presented below, attains the local variability of FDC and, more in general, the structure of the daily riverflow process.

3. Structure of Probabilistic Flow Duration Curves

To investigate the probabilistic structure of FDCs, we analyzed daily runoff data from the 14 gauging stations with at least 5 years of records available in the Basilicata region in Italy. Characteristics of the data set are presented in Table 1, which also reports the values of the Base Flow Index (BFI) defined later in the paper.

Lognormal flow duration curves for the above stations were computed on data arranged on the hydrologic year (starting on Oct. 1). The sets of parameters α and β estimated from all of the series were then analyzed in order to assess their probabilistic structure.

This analysis showed clearly that the distribution of α and β is normal (see Figure 2) in all of the stations examined. This result is theoretically justified, since they are respectively the estimate of the mean and of the standard deviation of a normal variable. The former is normal-distributed while the latter follows a χ^2 distribution with $n-1$ degrees of freedom, where n is the sample length. However, the χ^2 distribution closely approximates the normal when n is large, as in this case where n=365.

The structure of the probabilistic flow duration curve (3) is then easily determined when data is available, by derivation of $\alpha_\phi = \mu_\alpha + u_\phi \, \sigma_\alpha$ and $\beta_\phi = \mu_\beta + u_\phi \, \sigma_\beta$, with u_ϕ as the standard normal variate.

It is worth adding that if statistics related to aggregated data are needed (as, for instance, the minimum average weekly flow for a given return period) the same procedure applies equally, even though the FDCs are to be recomputed with reference to aggregated data. In alternative, one can develop rules analogous to these reported by the *Institute of Hydrology* [10] p. 29, concerning the link between the discharge $Q_D(95)$, averaged over *D* days, and $Q_{10}(95)$.

TABLE 1. Main features of the data series analyzed.

Code	Station	Series Length (*years*)	Basin area (km^2)	Mean basin elevation (*m.a.s.l.*)	BFI	Mean annual runoff (*mm*)
1.0	BRADANO at Tavole Palatine	22	2743	407	0.29	84
1.1	BRADANO at San Giuliano	21	1631	440	0.252	127
1.2	BRADANO at Ponte Colonna	40	459	560	0.346	136
1.3	SAGLIOCCIA at Acqua Fredda	7	15.6	410	0.504	39
2.0	BASENTO at Menzena	28	1405	664	0.342	274
2.1	BASENTO at Gallipoli	40	848	893	0.382	338
2.2	BASENTO at San Vito	6	157	940	0.496	331
2.3	BASENTO at Pignola	43	57.6	1015	0.451	568
3.1	SAURO at Acinello	5	222	780	0.256	216
3.2	AGRI at Tarangelo	34	507	870	0.59	623
3.3	AGRI at Grumento	17	278	886	0.587	760
3.4	AGRI at Le Tempe	41	174	933	0.538	803
4.0	SINNI at Valsinni	33	1142	752	0.506	592
4.1	SINNI at Pizzutello	48	233	932	0.503	1029

Expression (3) refers to the dimensionless expression of the probabilistic FDC. Given a probability ϕ of non-exceedance, the related discharge $Q_{F,\phi}$ of duration $F\cdot(365+1)$ comes from

$$Q_{F,\phi} = exp(\alpha_\phi + \beta_\phi \; z(F)) \; E(Q) \tag{4}$$

However, since in defining q the mean $E[Q]$ was used as the average discharge in the year considered, it is important to exclude that the average annual discharge is correlated to parameters α and β. In that case, the global average could not be used independently of the probability of non-exceedance considered.

Nevertheless, the analyses made on our samples completely excluded the presence of correlation between the annual mean discharge and the FDC parameters. Therefore, the global average can confidently be used in (4) to build dimensional curves.

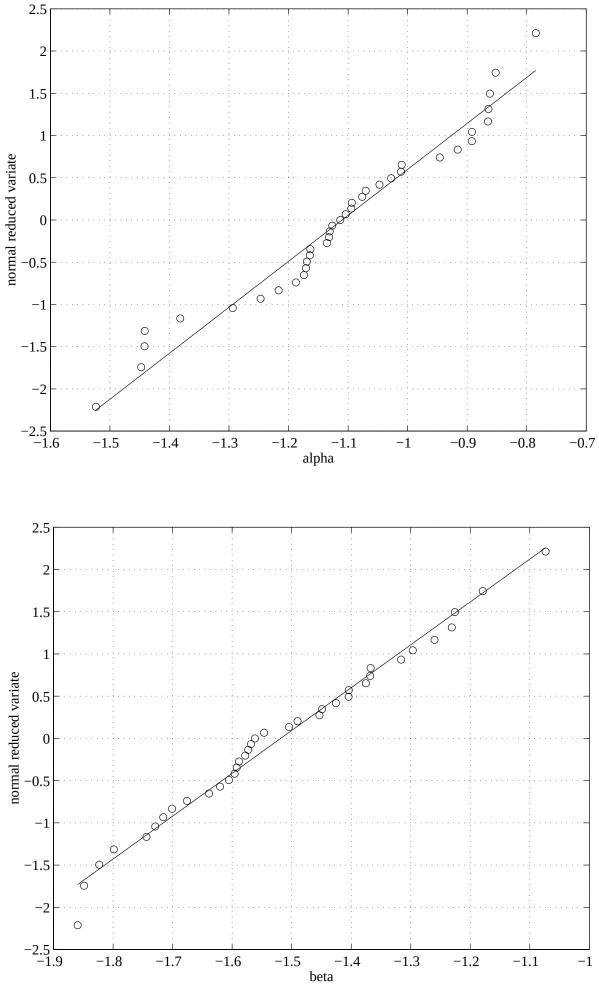

Figure 2. Station 2.1: Fitting of the normal distribution to parameters α and β.

Based on how flow data are arranged, α and β for a given station are correlated, because, for a lognormal variable, it is:

$$\alpha = \mu[log(q)] = log(\mu[q]) - 1/2\ \sigma^2[log(q)] = \\ = -1/2\ \sigma^2[log(q)] = -1/2\ \beta^2 \qquad (5)$$

given that $\mu[q]$=1. However, when α and β are estimated on the real world data, this theoretical link will appear only as a cross-correlation, because of the approximations in the probabilistic model and in the estimation methods. In any case, this circumstance must be accounted for in the regional analysis, where the dependence of parameters must be preserved.

On the data series available, α and β were computed as estimates of $\mu[log(q)]$ and $-\sigma[log(q)]$, respectively, obtained with the criterion of the overall best fitting. In other words, α and β were derived as the least-squares parameters of the straight line (1) with q_0=0. This method seems more coherent with the definition of duration curves, which does not underline, in principle, a probability distribution of flow data.

In our data set, the estimates of α and β obtained in each station, from each year of data, are clearly correlated, but, owing also to the estimation method adopted, the correlation coefficients obtained were not particularly high.

4. Regional variability of flow duration curves

Evaluation of the dimensionless flow duration curve for a given return period in a ungauged site requires, as seen previously, the determination of the probability distribution of α and β. This is equivalent to the determination of their moments μ_α, σ_α, μ_β and σ_β.

Evaluation of moments of α and β without data is possible if they are in some way related to climatic and physical characteristics of the basins. Investigation of such relationships is the subject of the regional analysis described hereafter. Moments of α and β computed for the gauging stations examined in our case study, reported in Table 2, are the reference data for this regional analysis.

The dimensional FDC is then obtained through evaluation of the mean annual runoff. Possible methods for the estimation of the mean annual runoff will not be discussed here, as many approaches to this task are available in literature (see *e.g.* [10] p. 32, [4]).

4.1. Regional estimation of the mean of FDC parameters

Among the catchment characteristics used to regionalize low flow parameters, a peculiar position is occupied by the Base Flow Index (BFI), defined [12], [10] as a measure of the proportion of the river runoff that derives from groundwater sources. It is intended as a catchment characteristic and, being so, its computation is made on the entire available data set of daily runoff.

To serve as a catchment feature, the BFI requires to be computed with a method that ensures robust estimation. The method suggested by *Lvovitch* [12] has been repeatedly confirmed as providing reliable estimates. With this method the base flow curve is traced by connecting 'turning points' searched among a subset of the daily runoff time series. The subset is made up of the minima of N days non-overlapping periods.

TABLE 2. Moments of parameters α and β.

code	α			β		
	mean	std. dev.	CV	mean	std. dev.	CV
1.0	-1.420	0.271	-0.191	-1.707	0.223	-0.131
1.1	-2.144	0.908	-0.424	-2.185	0.485	-0.222
1.2	-1.800	0.480	-0.267	-2.118	0.372	-0.176
1.3	-0.871	0.433	-0.497	-1.406	0.407	-0.290
2.0	-1.249	0.319	-0.255	-1.713	0.288	-0.168
2.1	-1.110	0.184	-0.166	-1.519	0.197	-0.130
2.2	-0.555	0.216	-0.390	-1.052	0.220	-0.209
2.3	-0.621	0.177	-0.286	-1.092	0.159	-0.146
3.1	-1.264	0.278	-0.220	-1.801	0.231	-0.128
3.2	-0.337	0.117	-0.347	-0.730	0.112	-0.153
3.3	-0.298	0.114	-0.383	-0.671	0.149	-0.222
3.4	-0.299	0.098	-0.327	-0.694	0.129	-0.186
4.0	-0.655	0.107	-0.163	-1.098	0.111	-0.101
4.1	-0.712	0.175	-0.246	-1.121	0.160	-0.143

The selection of N, which is actually a smoothing parameter, affects the estimate of the BFI. Suggestions from the literature ([10], [3]) and a sensitivity analysis performed on our data [16] all converge on the 'best' value N=5.

In our case study, the BFI emerged clearly as the parameter which better describe the variability of μ_α and μ_β, confirming a similar result obtained by *Manciola and Casadei* [13]. The relation is particularly meaningful when applied to μ_β (Figure 3**)** also because if the stations in the Bradano basin are excluded (codes 1.#) the coefficient of determination jumps to 0.93. Such a separate consideration of the Bradano basin could be justified by some particular hydrogeologic features, currently under investigation. However, we assume that the result of Figure 3 is acceptable as it is.

In agreement with the correlation existing between α and β it was then assumed more correct to determine μ_α as a function of μ_β, rather than of the

BFI. The former relation was established on a regional basis (Figure 4**),** and is again characterized by a high degree of correlation between the variables.

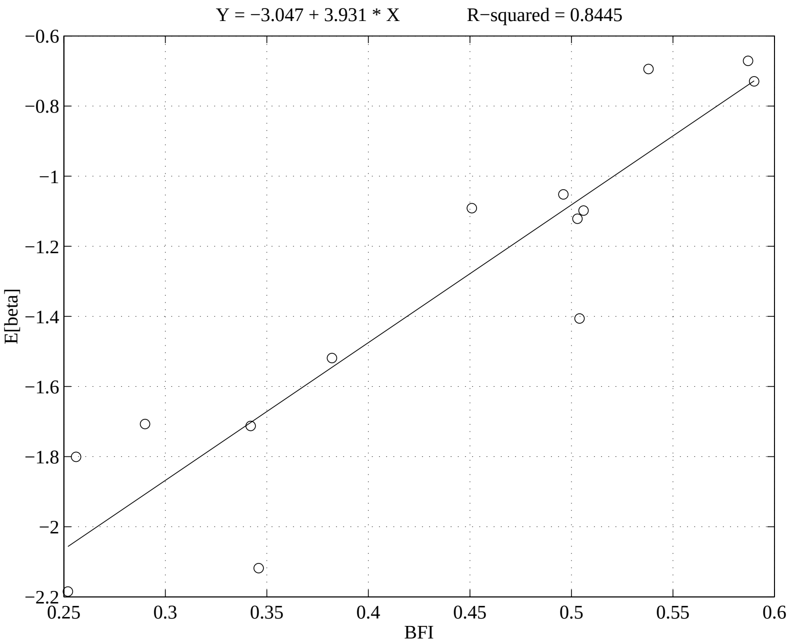

Figure 3. Relation between μ_β and BFI for all of the stations of Table 1.

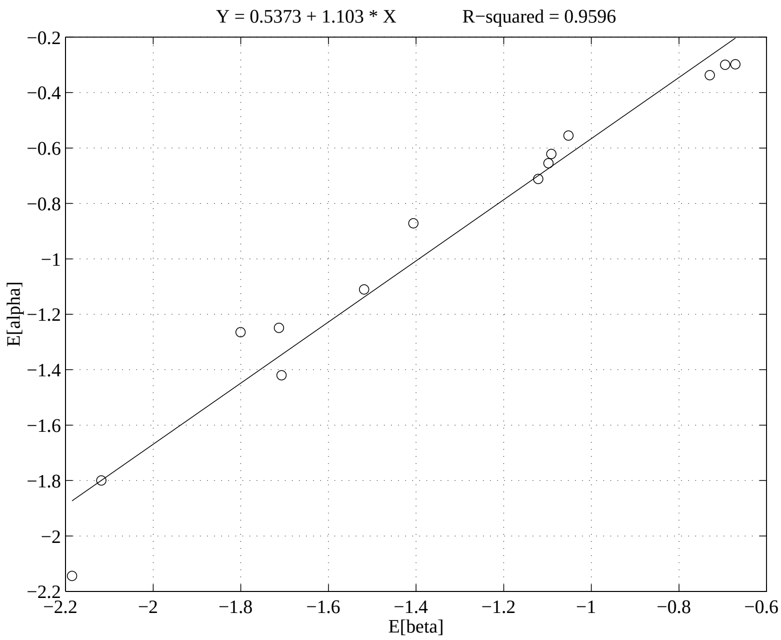

Figure 4. Relation between μ_β and μ_α for all of the stations of Table 1.

4.1.1. Determination of the BFI in ungauged basins

Low flow indices are strongly dependent by the hydrogeological characteristics of the basin. In some of the regionalization approaches [17], [5], a groundwater recession constant was used as a catchment parameter measuring the relative amount of groundwater runoff.

Even though these papers report procedures for the determination of the recession constant from hydrogeological maps, the literature available on the analogous procedures for determination of the BFI is certainly richer (see *e.g.* [14], [9]). *Gustard* [9] and *Demuth and Hagemann* [5] also provide wide literature surveys on methods and case-studies regarding relationships between low flow indices and geology.

Besides these methods it is always to consider the possibility of obtaining meaningful information from a very short amount of direct data. Indeed, in the *Low Flow Study report* [10] the main procedure suggested for obtaining the BFI of an ungauged basin was to collect at least one year of flow data, because it was reported that the BFI standard error was as low as 0.04. However, a corresponding analysis [16], relative to the data considered in this paper, showed that in all of the cases the standard error of the BFI was over 0.1 and often it was even greater than 0.2. Therefore, even if small data sets are available, information from geology seems, in general, necessary to obtain reliable estimation of the BFI.

4.2. Regional estimation of the variance of FDC parameters

The standard deviations of α and β do not show any particular dependence on the variables considered in the previous case, nor look particularly stable in the region considered. On the other hand, the coefficient of variation of both α and β appears to be quite constant among stations, as can be remarked from Figure 5, in which the coefficient of variation of β, computed for all of the stations, is compared to the 95% approximate confidence bands, whose computation was made according to [11], p.76.

Given the results shown in Figure 5 it can be concluded that the weighted average $cv_\beta = -0.159$ can be taken as the reference value for ungauged sites. The weights used were the number of observations, which are less than the numbers reported in Table 1 since the analysis was performed on data organized in hydrologic years.

The corresponding weighted average for α is $cv_\alpha = -0.274$. This rather high value does not give the possibility of drawing sufficiently reliable confidence bands, according to the restrictions associated with the approximation cited in [11]. However, the confidence on this 'regional' value is supported by the linear correlation existing between the coefficient of variations of α and β (which has R^2=0.872). From this linear relation, exactly the same average value arise for cv_α in correspondence of cv_β.

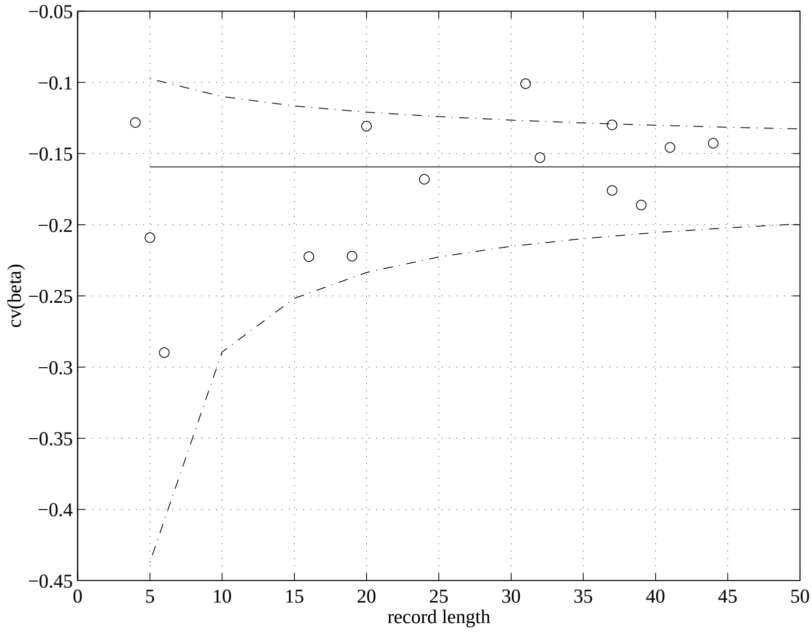

Figure 5. Variability of CV_β in the region investigated and 95% confidence bands. The weighted average of CV_β is -0.159.

5. Final remarks

In the field of the water quality management and planning some issues are emerging which require the availability of a comprehensive probabilistic description of river low flows. The methodological approach proposed in this paper goes in this direction, introducing a probabilistic description of the flow duration curves, which gives the possibility of a useful integration of the information presently provided by the commonly used low flow indices, such as Q(95) or $Q_{7,T}$.

Probabilistic flow duration curves can be built in gauging stations and can provide a different source of information besides that of the more traditional flow frequency curves. In addition, some criteria were presented here for regional estimation of their parameters, in view of the curve building for ungauged sites.

With reference to our case study, dealing with the river basins of the Basilicata region in southern Italy, the procedure proposed for construction of the probabilistic flow duration curves in a ungauged site can be summarized as follows:

– determination of the BFI of the basin
– evaluation of μ_β from the relation $\mu_\beta = 3.93 \cdot BFI - 3.05$

– evaluation of μ_α from the relation $\mu_\alpha = 1.103\ \mu_\beta + 0.537$
– determination of $\sigma_\alpha = -\ 0.274\ \mu_\alpha$ and $\sigma_\beta = -\ 0.159\ \mu_\beta$
– evaluation of the mean annual runoff

The first and last steps were not explicitly considered in this paper, as many methods are available in literature for the specific purposes.

6. References

1. Binns N. A. (1982) *Habitat, quality index procedure manual*. Wyoming Game and Fish Department, Cheyenne, 209 pp.
2. Bovee K. D. and Milhous R. T. (1978) Hydraulic simulation in instream flow studies: Theory and techniques. Instream Flow Information Paper 5. U.S. Fish and Wildlife Service, 130 pp.
3. Casadei S. (1995) Analisi dell'indice di deflusso di base (BFI) in alcuni bacini dell'Italia Centrale, *Idrotecnica* n. 3 , 177-191.
4. Claps, P. and Silvagni G. (1996) Aspetti idrologici della gestione operativa delle risorse idriche in uno schema interconnesso, paper presented at the *III Simposio Italo-Brasiliano di Ingegneria Sanitaria-Ambientale*, Gramado (Brasile) 16-20 Maggio, 1996 (in italian).
5. Demuth S. and Hagemann I. (1995) Case study of regionalising base flow in SW Germany applying a hydrogeological index, in A. Gustard (ed.) *Flow Regimes from International Experimental and Network Data (FRIEND),* - UNESCO- Vol. I, Hydrological studies, pp. 86-98, Institute of Hydrology, Wallingford, Oxon, UK.
6. Evangelisti, G. (1965) *Impianti idroelettrici*, Patron Publ., Bologna, Italy (in Italian).
7. Franchini M. and Ferraresi M. (1988) Analisi regionale dei deflussi di magra tramite le curve di durata, *Proc. XXI Convegno Nazionale di Idraulica e Costruzioni Idrauliche*, L'Aquila, Sept. 1988, vol. I, 139-149 Maggioli Publ., Rimini, Italy (in Italian).
8. Gustard A. (editor) (1995) *Flow Regimes from International Experimental and Network Data (FRIEND)*, - UNESCO- Vol. I, Hydrological studies, Institute of Hydrology, Wallingford, Oxon, UK.
9. Gustard, A. (1995) Case study in regionalising low flows in the UK, in A. Gustard (ed.) *Flow Regimes from International Experimental and Network Data (FRIEND),* - UNESCO- Vol. I, Hydrological studies, pp. 68-85, Institute of Hydrology, Wallingford, Oxon, UK.
10. Institute of Hydrology, (1980) *Low Flow Studies*, Research Report, Institute of Hydrology, Wallingford, Oxon, UK.
11. Johnson N.L. and Kotz S. (1970) *Distributions in Statistics - Continuous Univariate Distributions*, Vol. I, J. Wiley, New York.
12. Lvovitch M.I. (1972) Hydrologic budget of continents and estimate of the balance of global fresh water resources. *Soviet Hydrology*, N. 4, 349-360.

13. Manciola P. and Casadei S. (1990) Flow duration curve estimation, in Vogt G. and H. Mickle (eds.): *Proc. XXI Annual Pittsburgh Conf. on Modeling and Simulation*,1467-1472, Univ.of Pittsburgh, Pittsburgh, MA.
14. Manciola P. and Casadei S. (1991) Low flow index and hydrogeological characteristics, *ASCE Proc. National Conf. on Hydraulic Engineering 1991*, 29 July - 2 August 1991, Nashville, TN, pp. 930-936.
15. Stalnaker C., Lamb B.L., Henriksen J., Bovee K. And Bartholow J. (1995) *The Instream Flow Incremental Methodology: a primer for IFIM*. U.S. Dept. of the Interior, National Biological Service, Biological Report 29, Washington D.C.
16. Urgesi R. (1993) *Analisi regionale delle curve di durata dei corsi d'acqua del versante ionico della Basilicata*, unpublished graduate thesis, Università della Basilicata, Potenza, Italy (in Italian).
17. Vogel R.M. and Kroll C.N. (1992) Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, *Water Resour. Res*. Vol. 28 , No. 9 , 2451-2458.
18. Yevjevich V. (1984) Structure of Daily Hydrologic Series, Water Resources Publication, Fort Collins, Colorado.

Acknowledgments

This work was supported by funds granted by the GNDCI of the Italian National Research Council (CNR) and by the NATO under the Linkage Grant "Assessment of Water Quality Monitoring Networks"