"SAPIENZA" Università degli Studi di Roma
Facoltà di Ingegneria
Dipartimento di Ingegneria Civile, Edile e Ambientale

Dottorato di Ricerca in Ingegneria Idraulica
XXIII Ciclo

Tesi di Dottorato

Regionalization of flow duration curves in flow regimes of intermittent streams

Coordinatore: Prof. Ing. Antonio CENEDESE

Tutor: Prof. Ing. Lucio UBERTINI

Co-Tutor: Prof. Ing. Francesco NAPOLITANO

Dottorando: Ing. Maura Rianna

Roma, Dicembre 2010

Abstract

Low flow analysis is a very important and widely studied topic in hydrology and water resource design and management. The knowledge of low flows periods and of the amount of available water is in fact fundamental in a wide range of scenarios. Conventionally low flow analysis is used for the design and operation of public water supply schemes, but an other frequent application is the evaluation of dilution of waste water that arrives from industrial or domestic source into a river.

The number of summary information derivable from daily time series that describes low flow regimes of a river is quite large. The big amount of methods is based mainly on the types of available data and the required output.

One of the largely used tool to evaluate low flow and the river regime is the flow duration curve (FDC). This is a cumulative frequency curve that shows the percentage of time that a given discharge is equalled or exceeded during a fixed period. The choice of this tool was done because it is one of the most informative methods of displaying the flow characteristics of a stream throughout the range of discharge, without regard to the sequence of occurrence. These curves are derived in gauged river data, but are essential also in rivers where the systems for collecting and managing water information are inadequate.

The aim of this work is then to assess the reliability of existing regionalisation models of flow duration curves, to create a new model to calculate flow duration curves in rivers characterized by an intermittent regime and to apply a procedure of regionalisation starting from definition of homogeneous regions in the target area till arriving to the transfer of the hydrological information in ungauged basins.

In particular the work shows a new method for calculating flow duration curves and annual flow duration curves in intermittent basins. This methods joins the stochastic index theory (Castellarin 2004) that model the relationship between flow duration curves and annual flow duration curves of daily discharges to the Bayes's theorem that can be used to determine the probability of occurrence of a non-zero event, given that a zero event is already occurred.

Sommario

Lo studio delle portate minime è un argomento fondamentale e ampiamente studiato in diversi campi dell'idrologia e della gestione delle risorse idriche. La conoscenza dei periodi e della frequenza de i minimi di portata e della quantità d'acqua a disposizione in quei periodi è infatti di fondamentale importanza in un vasto campo di situazioni. L'analisi delle portate minime è normalmente utilizzata per la progettazione degli schemi di fornitura dell'acqua ed è frequente la sua applicazione per valutare la diluizione delle acque di scarico che provengono da fonti domestiche o industriali verso i corsi d'acqua.

Il numero di informazioni riassuntive derivabili dalle serie temporali giornaliere per descrivere il regime delle portate minime è elevato. Il tipo di informazione che si può ricavare dipende dal tipo di dati a disposizione e dal risultato richiesto.

Uno degli strumenti maggiormente utilizzati per valutare il regime delle portate minime di un corso d'acqua è la curva di durata delle portate (FDC). Questa rappresenta la curva di frequenza cumulata che mostra la percentuale di tempo in cui una determinata portata è uguagliata o superata. La scelta di questo strumento per rappresentare le portate minime è stata effettuata in quanto questo può essere considerato come uno degli strumenti più informativi per rappresentare le caratteristiche di un corso d'acqua. Queste curve si derivano con facilità in bacini strumentati ma risultano essenziali anche in quei corsi d'acqua in cui il sistema di raccolta e gestione delle informazioni idrologiche è carente.

L'obiettivo di questo lavoro è quello di valutare l'affidabilità dei diversi metodi di regionalizzazione delle curve di durata delle portate e in particolare di creare un nuovo modello che permetta di calcolare le curve di durata in bacini caratterizzati da regime intermittente. Inoltre si intende applicare una procedura di regionalizzazione che parta dalla definizione di aree omogenee nell'area di studio fino ad arrivare al trasferimento dell'informazione idrologica in bacini non strumentati.

In particolare il lavoro mostra una nuova metodologia per calcolare le curve di durata basate sull'intero periodo di registrazione (FDC) e quelle annuali (AFDC) in bacini caratterizzate da regime intermittente. Il metodo unisce la teoria dello Stochastic Index, che permette di modellare la relazione tra FDC e AFDC al teorema di Bayes che può essere utilizzato per determinare la probabilità di accadimento di un evento non zero, quando un evento pari a zero si è già verificato.

Table of contents

List of figures

List of tables

1

Introduction

The thought of river brings back the image of floods and high water levels. Besides a big amount of money of governments is used to protect environment and people from these events. Conversely the life of many communities is strictly connected with the availability of water for their water supply. In this context the low flow periods are critical for the water resource management.

The knowledge of low flows periods and of the amount of available water is fundamental for different purposes. There is in fact a wide range of scenarios where low flow information is needed. Traditionally one of the most common use is the design and operation of public water supply schemes. The low flow period length and the amount of available water is required to evaluate the probability that abstraction does not meet the anticipated water demand.

An other frequent application of low flows is the evaluation of dilution of waste water that arrives from industrial or domestic source into a river. Usually an official authorization is obligatory to release a pollutant and the discharge is avoided in rivers where there is no flow for period longer than a threshold.

Moreover ecosystems are more exposed to risks during low flow periods because of the decrease of oxygen dissolved, deterioration of water quality, the reduction in availability of habitats. At the end rivers are natural resources for recreational activities, and low levels and low velocities of the water can reduce the appeal of these natural environments.

As a matter of fact different and complex natural processes at a catchment scale cause low flows in a river. This can be seen as a conceptual model, compounded by interlinked reservoirs, each of them having recharge, storage and discharge phenomena. The recharge of the system is caused mainly by precipitations (solid or liquid), while the storage and the discharge depend on the geological and physiographic characteristics of the water basin. During dry seasons the most important processes are the releases of water from storage. Different aspects, natural and anthropogenic influence the low flow regime of the river. The natural factors that affect gain and losses, include hydraulic characteristics and extent of the aquifers, the types and infiltration characteristics of soils, the rate, frequency and amount of recharge, the evapo-transpiration rates from the basin, topography, climate and type and distribution of vegetation. Moreover anthropogenic impacts, as groundwater abstractions, change of the use of the soil and of the vegetation, construction of dams have to be considered in the low flow analysis.

Different measures and indexes can be used to analyse and estimate low flow regimes. One of the most informative methods of displaying the entire range of river discharge from low flows to flood events, is the flow duration curve. It is a relationship between any given discharge value and the percentage of time that this discharge is equalled or exceeded. This can be seen as the relationship between magnitude and frequency of streamflow discharges. These curves are derived in gauged river data, but surely this information for water management purposes is essential also in rivers where the systems for collecting and managing water information are inadequate.

Different solutions exist to answer this problem, and can be synthesized in two groups. The first one is based on statistical models. The flow estimation in ungauged sites is obtained by simple regression models, that permit to connect flow statisticals or parameters to known characteristics of the basins. The second group is represented by rainfall-runoff models that permit to simulate the evolution of the time series of river flows within a catchment. This method is useful to evaluate the interaction between the physical components of the water basin.

The aim of this work is to assess the reliability of existing regionalisation models of flow duration curves, to create a new model to calculate flow duration curves in rivers characterized by an intermittent regime and to apply a procedure of regionalisation starting from definition of homogeneous regions in the target area till arriving to the transfer of the hydrological information in ungauged basins.

1.1 Thesis outline

Chapter 2 presents a review of low flow hydrology, with analysis of low flows generating mechanism and analysis of low flows measures. In particular an explanation of flow duration curves is given, with particular attention on methods of construction of these curves and explanation of existing methods to regionalise flow duration curves.

Chapter 3 shows the studied area, that is the catchments belonging to Lazio region and a preliminary analysis on the discharge data. These include correlation analysis between discharge quantiles and geomorphologicalal characteristics, till arriving to definition of homogeneous regions, necessary for the regionalisation procedure.

Chapter 4 makes a review of existing methods of regionalisation of flow duration curves. Three main procedures can be founded by statistical, analytical and graphical methods. These are applied in the studied area following the literature, but adapting them at the case study. An evaluation of the reliability of these methods is done in the studied area using different indexes and a jack-knife validation procedure.

Chapter 5 shows a new method for calculating flow duration curves and annual flow duration curves in intermittent basins. This methods joins the stochastic index theory (Castellarin 2004) that model the relationship between flow duration curves and annual flow duration curves of daily discharges to the Bayes's Theorem that can be used to determine the probability of occurrence of a non-zero event, given that a zero event is already occurred.

Chapter 6 presents a regionalization procedure with a parsimonious number of parameters to estimate parameters in ungauged basins. The regionalization technique is based on a multiple regression analysis through the model parameters and geomorphologicalal characteristics.

Finally, in Chapter 7 a summary of the conclusions is presented, and future directions of research are discussed.

2

Low flow analysis and measurements

2.1. Introduction

The study of low flow cannot abstract from the definition of what low flow is. World Meteorological Organization (1974) defines low flows as the “flow of water in a stream during prolonged dry weather”. However, low flow is a seasonal phenomenon and differs from the drought phenomenon which is a natural and more general event and can be characterized by more than low flow streams (Smakhtin, 2001).

The low flow analysis is in fact connected with different topics such as the amount of water that is present in a river during the dry season of the year or the length of the time between flood events or the frequency of this period. The choice of the topic depends on the use of the stakeholder groups' needs. In fact, the knowledge of low flows is important in various engineering scenarios such as the development and design of water supply schemes, the waste load allocation, the design of dams or reservoirs, and the definition of the amount and quality of water for irrigation, domestic and recreation use.

Low flows originate from ground flow water, melting glaciers or surface water coming from lakes. The period of low flows is generally the same for the same region and occurs each year. According to Smakhtin (2001) low flows study is made up of two components: a temporal one that concerns the magnitude, the variability of flows and the length of these events, and a spatial component that is related to the regional distribution of low flows characteristics. This element in particular is dealt with the capability to obtain these properties in catchments where there is a lack of measured data.

To understand and study these elements of low flows, it is necessary to know the natural and anthropic causes of this phenomenon. In fact, several components influence low flows such as climate, geology, soil and topography but also abstractions and regulation of low flow domain through dams.

Literature on low flow ranges on different subjects. Various papers are related with a review of low flow study and analysis of different engineering applications (Riggs (1972), McMahon (1976) and Beran and Gustard (1977), Smakhtin (2001)). These papers approach the causes of low flows, the several techniques used for low flow analysis and define where it is possible to apply this analysis. Besides, they analyze the different characteristics that are computed by discharge data in order to obtain information on low flows used in different countries. Other important papers on this topic are those of Searcy (1959), Hall (1968), Riggs (1976), Vasak (1977), Kurdov (1977), McMahon and Mein (1986), Gustard (1989), Ponce and Lindquist (1990), Amusja et al., (1991), Heicher (1993), Demuth (1994), Tallaksen (1995) and Vogel and Fennessey (1995).

2.2. Low flow information

Low flow analysis is used in wide range of engineering applications often regulated by national and international water laws. In Europe, the Water Framework Directive (adopted in October 2000) establishes a framework for the surface water and groundwater resources. The Directive establishes government of catchments by river basin authorities and not by political boundaries. Moreover, the Directive prescribes that long term river basin plans, that are obtained using low–flow analysis, are defined for an integrated water resource management.

One of the most important and common uses of low–flow analysis is the design and operation of public water–supply schemes. In this field it is important to understand whether the water amount present in a river is sufficient for abstractions for different use like water treatment plants or reservoir storage facilities. The analysis of low flows is similar if abstractions or irrigation scheme are planned, although agricultural demand has a higher annual and inter-annual variability.

Water abstraction is also necessary for hydro-power use. This field requires an analysis of the complete range of flows, but low flow analysis is essential to decide how much water must bypass a hydro–plant to maintain downstream river ecology, and how much is available for hydropower production in dry season. For all these applications the forecast of flows is necessary to evaluate if restrictions on water use can minimize the risk of lack of water in the future.

Low flow analysis is often used also to estimate dilution of domestic or industrial discharge released into a river. In particular, legal authorizations are needed for discharge waste water. The frequency distribution of downstream water quality of a river is obtained through water quality models that use as input data rate and quality of the discharges. The complement of the cumulative distribution function or river flow duration curve is commonly used for develop the analysis.

Moreover, ecosystems are most vulnerable during the dry season because the reduction of water creates a decrease of dissolved oxygen, the fragmentation of the habitat and the deterioration of water quality. Flow duration curves or percentiles obtained from flow duration curves and low flow indexes as mean annual minima for a given duration are often used in this field.

In addition, rivers are also used for sport and recreation activities. In this case the artificial support of water levels and velocities can be important to maintain the attractiveness of a river.

2.3. Estimation, risk and forecasting

As seen above, low flow analysis is needed in different fields and for each situation different information is required. Simple indexes can be calculated, as recession constant, the mean value of the discharge time series and the proportion of baseflow.

Besides it is possible to calculate more informative characteristics such as cumulative distribution functions of daily flows or to use extreme values techniques. The first one represents the relationship between the discharge and the percentage of time that it is exceeded. The second one is used to calculate the non–exceedance probability of annual minima. The biggest difference between the two techniques is that flow duration curves use all the recorded time series and hence evaluate the percentage of time the entire time series exceeded. The extreme value technique is applied to annual minima data and permits to estimate the non–exceedance probability in years or the return period when the values are below a given value.

These estimates do not permit the forecasting of when low flow events will happen. Different methods exist to make long or low term estimation that allows calculating both the magnitude than the time of the low flow event. Obviously when forecasting time increases, the uncertainty of the prediction increases too and for very long lead times the long term statistical mean gives the best predictions.

2.4. Temporal and spatial scales of estimation

Different time scales can be used to perform low flow analysis but it is typically carried out using all available data with the minimum time scales. Sometimes it is necessary to use more appropriate time scales as weeks, months or seasons. For example, in designing and irrigation schemes, the analysis should focus on the season of the year when the abstractions for irrigation will take place. The same can be said for other use and then other period of the year. It could be then appropriate to calculate low flow indexes of different durations as 7-days, 10-days, 30-days and 90 days durations for specific design projects.

Besides, low flow problems occur over a wide range of space scales. These range from reaches of the dimension of 100 m within basin of 10 km^2, to basins of 1 million km^2 of area. Usually these kinds of basins in developed countries have a big number of gauging stations and long time series (often biggest than 50 years). Instead in developing countries or also where there is a good net of gauging stations, it often happens that continuous observational data is of poor quality and often there is a lack of gauging stations in some reaches or in little basins. In these cases it is not possible to estimate low flows characteristics directly from data but models could be used to assess low flow analysis.

Furthermore, low flow analysis can be done on the scale of river reach, catchment, national and international countries or globally. In this case the time and spatial scale of the problem will be a primary issue concerning the approach and the necessary data for the study.

2.5. The low flow cube

Different procedures and different kind of data can be applied during water management and supply scheme planning. The choice is done depending on the nature of the required output and the risk associated with the design decision. For example, the construction of a large reservoir will imply a high risk with the design decision. This kind of project needs gauging stations and the analysis of observed river flows. These data are necessary for hydrological design, typically the storage characteristics and spillway capacity. Oppositely, a small scale abstraction licence will not have high risk associated with it, it often requires in ungauged stations and the design will be based on low flow statisticals.

This complex scenario can be simplified using the "Design Scenario Cube" (Unesco, 1997), (Figure 2.1). The three dimensions of the cube represent the three design requirements and are defined in this way:

- the location of the design problem. This dimension establishes if there is a gauged station or not nearby the studied area;
- the operational requirement of the hydrological design and the financial capacity that determines if simple statistical or a long time series are required;
- the characteristics of the data that can be natural or necessitate to be naturalized. The catchment water use dimension discriminates between natural or artificially influenced flows.

The cube makes possible to create eight different possible combinations of these dimensions. Surely from this simple scheme it is possible to arrive to more complex situations.

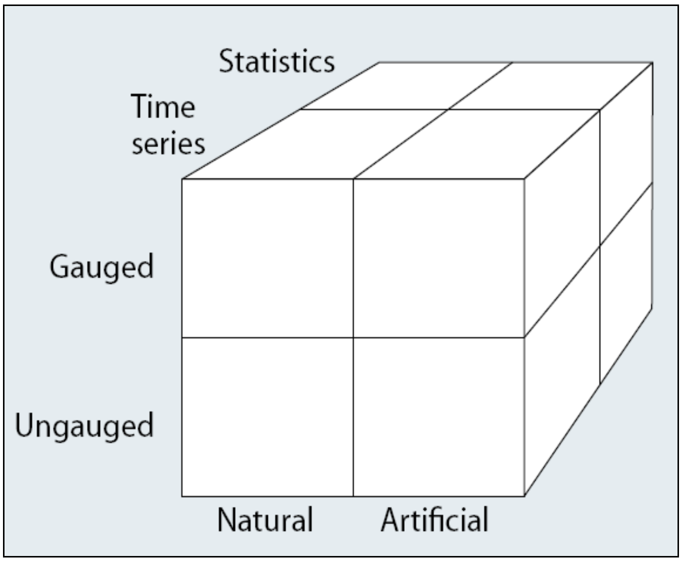

Figure 2.1 – Low flow cube (UNESCO, 1997)

2.6. Natural processes and regimes

The flow in a river is the result of complex processes that happen at a catchment scale. It schematizes the catchment as a conceptual model, constituted by interlinked reservoirs, where recharge, storage and discharge processes take place. While recharge of the catchment mainly depends on precipitation, storage and discharge are functions of physiographic characteristics. Low flows occur therefore after periods of no rain or when precipitation falls as snow. Besides this decreases the water stored in the soils and also the outflow to the river. The time of depletion depends on hydrological processes that operate near the channel as the storage properties within the catchment. Catchment soil and geology influence the capacity of the catchment's precipitation absorption and release as low flow.

Figure 2.2 shows the conceptual model of a catchment that have as main input precipitation, which then recharges different basin storages that are soil water, groundwater, snow, glaciers and lakes. The outflow from these storages contributes to river flows.

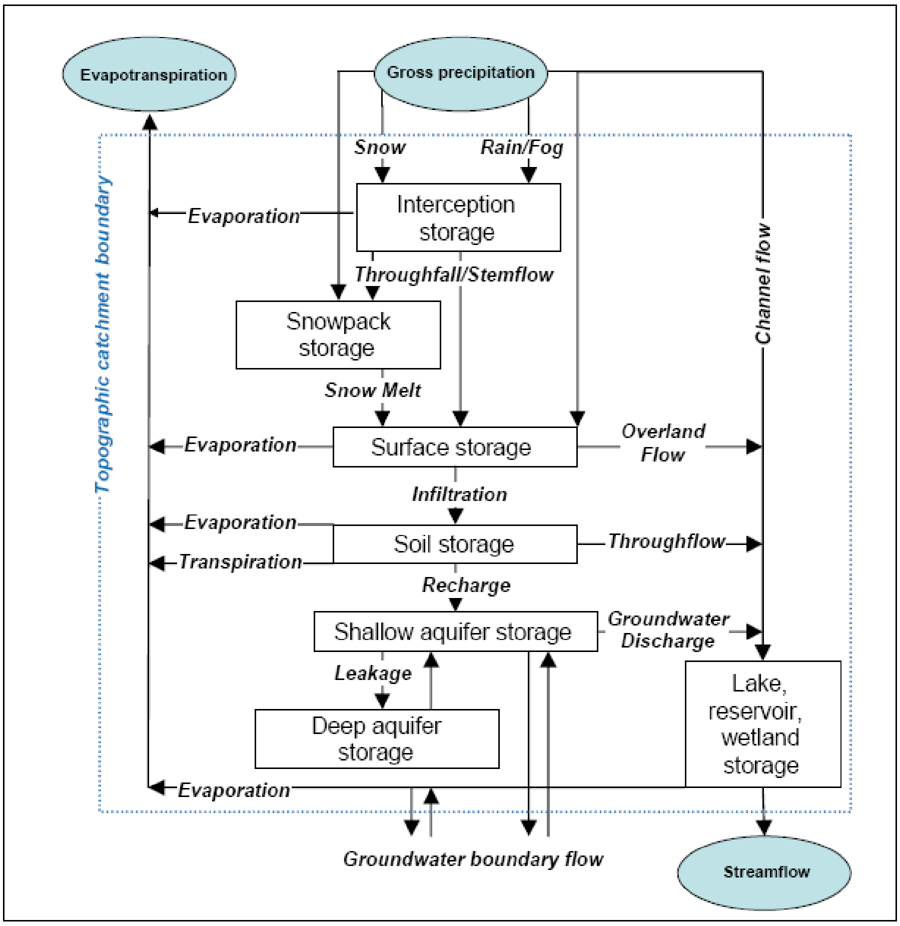

Figure 2.2 – Schematic conceptual model of a river basin, that shows the different catchment processes (WMO, 2008).

The next paragraphs will show different natural and anthropic factors that influence the spatial and temporal distribution of low flows.

2.6.1. Climate processes

Catchment input is precipitation or snowmelt, so low flow can decrease due to different causes, as:

- extended dry periods, when potential evaporation is higher than precipitation;
- extended periods of low temperature, when precipitation falls as snow.

Warm periods and dry weather periods are associated with high pressure systems, when high temperatures, high radiation input, low humidity and wind increases evaporation and transpiration rates. Conversely, snowfall and snow storage happen during periods of low temperatures, associated with cold, polar air masses and decreasing temperatures. In this situation the absence of snowmelt, creates an accumulation of solid water and then low flows (Bowles and Riley, 1976; Gerard, 1981; Collins, 1982; Fountain and Tangborn, 1985; Gurnell, 1993; Hopkinson and Young, 1998).

Winter low flows or summer low flows in mid and high latitude climates occur both annually. In low latitude climates different dry season and then distinct low flow periods can occur during the year.

In arid and semiarid climates, the combination of low precipitation and higher evaporation causes minimal river networks and ephemeral rivers. These kinds of basins are characterized by prolonged periods of zero flows and episodic high flows, often in the form of flash floods.

Climate processes have then a big influence on low flows, due to the influence on the magnitude and the variation of temperature, the potential evaporation and the precipitation that modifies the input water in the catchment. Climate maps furnish good information to understand low flows conditions of an area. However, climate varies spatially, particularly in mountain regions where there are strong altitude dependent temperature and precipitation gradients and temporally, at inter-annual, decadal and centennial time scales.

2.6.2. Catchment processes and storage

While climatic processes influence the increasing or decreasing of water in the catchment, in terms of deficit in one season or surplus in another the soil, vegetation and geology characteristics affect how these surplus and deficit propagate to the stream-flow. The most significant properties are soil moisture and groundwater storage, aquifer properties and hydraulic resistance between aquifer and river.

In particular, in impermeable urban surface, in sloping non–vegetated land, in compacted soils or exposed rocks it is common that water is stored in micro-depressions, and after having filled these, water goes to the stream-flow. Overland flows depend on precipitation intensity and whether it exceeds infiltration rates of the soils. As the soil is saturated, water may flow vertically towards the aquifer to recharge groundwater storage or move laterally through permeable soil layers toward the stream. Available soil moisture capacity is really important for low flows because furnishes a support to high annual transpiration. By means of macro-pores, cracks on the rocks and pipes, water can recharge aquifer or move toward the stream without recharging the soil layer. In semiarid or arid aquifers recharge occurs through river beds of ephemeral rivers and originates from high precipitation on mountains regions.

Water arriving from the soil to the aquifers increases the groundwater level. The groundwater discharge to the stream occurs where stream channels intersect the main phreatic surface in a draining aquifer. For low flows to be sustainable: (i) the draining aquifer must be recharged seasonally with adequate amounts of moisture; (ii) the water table must be shallow enough to be intersected by the stream; and (iii) the aquifer's size and hydraulic properties must be sufficient to maintain flows throughout the dry season (Smakhtin, 2001). A different example of groundwater re-emergence can occur where relatively slow moving groundwater drainage in fracture zones above the main water table has a significant lateral component which intersects the ground surface in the vicinity of channels (springs) (Smakhtin, 2001).

Geology obviously can at the reverse cause losses of water from the basin. These processes can be synthesized in (i) groundwater recharge from streamflow where the phreatic surface lies below the channel; (ii) bed losses, where unconsolidated alluvial material underlies the river channel; and (iii) losses to relatively dry soils forming the banks of streams (Smakhtin, 2001).

Physiographic characteristics as soil type and geology are therefore the key components for low flows. Many studies reported the presence of a strong correlation between geological category and flow rates or low flow discharges (e.g. Armbruster, 1976; Smith, 1981; Musiake et al., 1984; Bingham, 1986; Aucott et al., 1987; Rogers and Armbruster, 1990, Nikic & Radoja (2009)). Different studies have found that catchment characterized by unconsolidated sedimentary rocks have normally low yields on low flow periods. On the contrary, metamorphosed sedimentary and igneous rocks showed higher flow values compared to their basin size. Furthermore, literature has shown that stratified drift have a key role in the generation of low flows (Cervione et al., (1982), Barnes (1986), Morrisey et al., (1988), Ries (1994) and Risle (1994)). White (1977) showed that karst, limestone and dolomite rocks have decreasing effects on low flows. Felton and Currens (1994) studied the relationships between high and low flow and karst springs, and provided some bibliography on hydrological processes in karst formations.

Lakes and reservoirs hydraulically connected with rivers have usually a strong impact on low flows. The adequate water level in a lake should be maintained during the dry season to allow lateral outflow into a stream. Exhaustive studies that deepen the effects of lakes on low flows are not wide enough but the importance of lakes for low flows in some regions may be derived from the inclusion of lake related parameters in prediction models for low flows (Gerasimenko, 1972; Vladimirov, 1976; Kuusisto, 1987; FREND, 1989; Sakovich, 1990).

2.6.3. Human influences on low flows

The anthropic influence is really important for low flows and can cause water increase or decrease in the river.

The more important human impact on low flows is abstraction within sub–surface drainage area. This decreases the level of phreatic surfaces and therefore potential re–emergence for groundwater in stream channels. Localised reductions in the level of the water table may affect either hydraulic gradients or the length of channel that intersects the phreatic surface. The effects of groundwater pumping near the head of a perennial river may result in groundwater table depletion through interception of recharge water and induced recharge of the aquifer from the river itself (Smakhtin ,(2001)). Various studies concern groundwater abstractions on low flow (Owen ,(1991), Gustard et al., (1992), Bickerton et al., (1993), Clausen et al., (1994) and Fendekova and Nemethy (1994)).

Another important human impact of the flow is the change in forestation. This affects mainly evapotranspiration losses from the soil, affecting gain and losses to alluvial storage. In addition, the presence of vegetation increases interception losses and disturbance of the soil structure. Different authors as Banks, (1961), Rowe, (1963); Schulze, (1979), Swank et al., (1988); Smith and Bosch, (1989); Keppeler and Ziemer, (1990), Meier et al., (1997), Wilgen et al., (1997) studied this problem. Land use changes affect strongly low flows since it changes infiltration or evaporation characteristics and involves the groundwater recharge (Simmons and Reynolds, 1982; Warner, 1984; Ferguson and Suckling, 1990).

All these actions indirectly affect low flows, but there are processes that directly act on low flows, removing or adding water in the stream. A synthesis of these processes is given here.

Water abstractions for industrial, agricultural and domestic use. These processes decrease the amount of water of the river and affect mainly the dry season and the frequency of these periods (Kottegoda and Natale, 1994; Wilber et al., 1996; Eheart and Tornil, 1999). It is important to remember that irrigation returns water to river channels and irrigation return flows can be a big part of a stream's water balance (Blodgett et al., 1992). According to some sources (Hall and Du Plessis, 1984; McKenzie and Roth, 1994) up to 40% of water initially used for irrigation returns into the stream.

Discharges of water from industrial and domestic sources. These can worsen the quality of water downstream. Pirt and Simpson (1983) and Pirt (1989) illustrated how multiple abstractions and effluent discharges may affect the dry season flow.

Construction of dams and river flow regulation regime. This regulation can increase or decrease low flow discharge levels depending on the operational management of the reservoir. This can be seen as the most important impact on a river's low flow regime (Biggs, 1982; Muzik, 1986; Gustard and Cole, 1987; Harboe, 1988; Sherrard and Erskine, 1991; Bonacci et al., 1992; Walker and Thoms, 1993; Finlayson et al., 1994).

Different human impacts can affect low flow regimes and many rivers of the world with perennial regime have became intermittent whilst many rivers have been created artificially. A complete review of the relationship between hydrological processes, low flows and anthropogenic impacts is given in Tallaksen and Van Lanen (2004) whilst Dingman, (2002) gives a complete representation of basin processes.

2.7. Low flows in different hydrological regimes

The regime represents the seasonality of river flows and characterizes the duration and timing of low flows. The regime can be evaluated by using different years of stream-flow data.

This paragraph will show the peculiar attributes of different regimes and how these are modified by catchment processes.

2.7.1. Rain dominated regimes

2.7.1.1. Climates with no distinct dry season

The main characteristic of regimes in tropical or temperate climates is that usually they are perennial. The seasonal distribution of rainfall will determine whether there is an evident dry period. Commonly in temperate climates the strong period of low flow is the end of a warm season when there is a lack of precipitation for a period of different weeks.

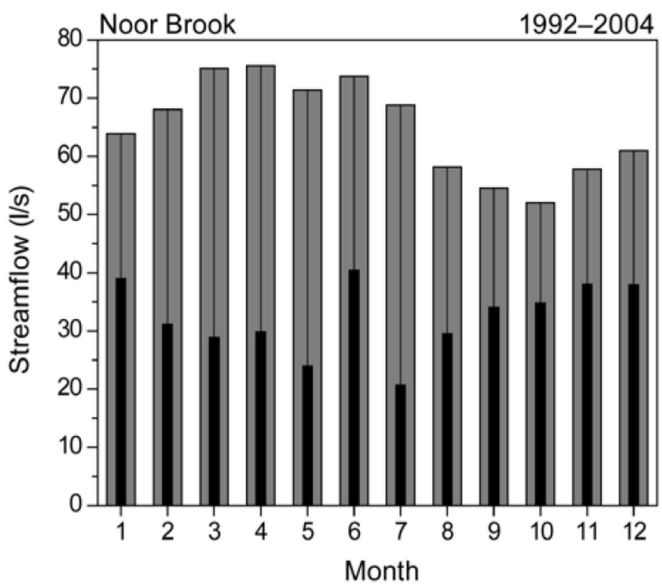

Figure 2.3 – Hydrological regime of the Noor Brook catchment (WMO, 2008).

The Figure 2.3 shows hydrological regime for a catchment in the Netherlands. The slowly responding Noor Brook drains the south eastern part of a chalk plateau and has a thick, unsaturated zone and a multilayered aquifer system with substantial deep storage. Due to this large store, the hydrograph of the Noor does not show a strong low flow season (WMO, 2008).

2.7.1.2. Climates with dry season (tropical and temperate)

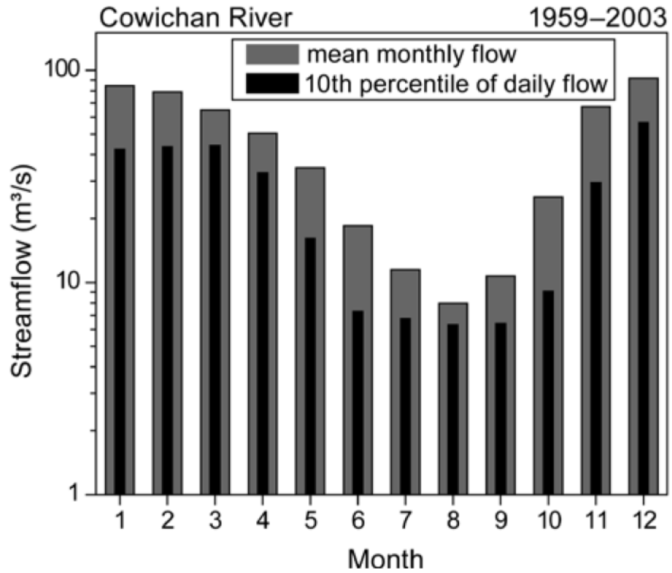

Figure 2.4 – Hydrological regime of Vancouver Island catchment (WMO, 2008)

Rivers belonging to climates characterized by dry season present high seasonality and the stream-flow diminishes with decreasing rainfall. In these kinds of climates high transpirations create very low flows, unless an aquifer or lake storage is present. Small catchments are often characterized by absence of flow for long periods.

Dry seasons are usually very similar from year to year and inter–annual variations of low flows depend mainly on the groundwater storage at the starting point of the dry season. Fig 2.4 shows an example of the catchment regime of Vancouver Island. The basin is characterized by a Mediterranean climate with a distinct dry season in the summer period from June to September. The y axe is represented in logarithmic scale in order to include the big range of discharges.

2.7.1.3. Dry climates

Streamflows in these climates are often intermittent or ephemeral so the river is dry during most part of the year.

2.7.2. Snow dominated regimes

This kind of regime is typical in high latitudes and also in mountain areas. Snow dominated regimes differ by rain dominated regimes because they do not follow the annual precipitation cycle but the water is stored as snow. Low flows occur usually in winter season, while stream-flows augment during spring season when temperature increases and snowmelt phenomenon happens. A second dry season can verify during the warm season depending on the amount of rain and snowmelt contribution. In high mountains, very low temperatures can cause soil freezing and thus very low base flows.

In regions with low mountains, snowmelt lasts few days or weeks. On the contrary, in high mountains snowmelt can go on for long periods and gradually move from lower altitudes to highest. When snowmelt ends, the stream flow reduces strongly.

Figure 2.5 shows an example of a snow dominated catchment from the mountainous area of British Columbia in Canada. The high flow period occurs during the summer from April to August. The catchment shows a single low flow period during the winter characterized by the highest low

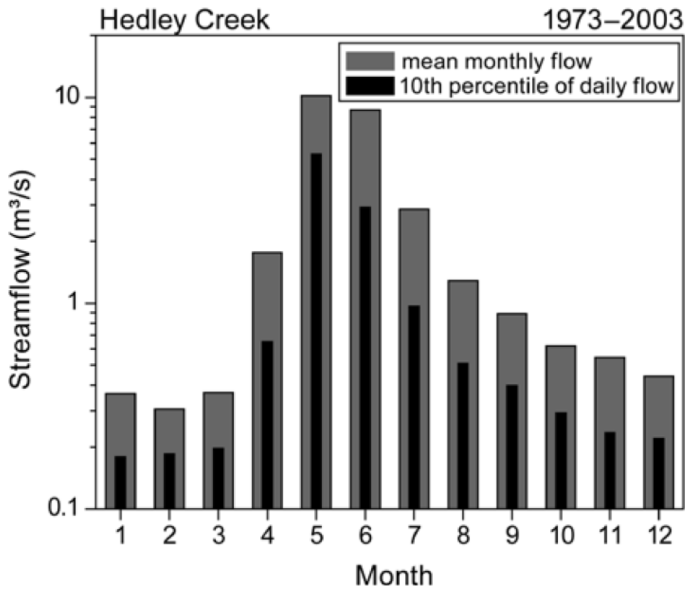

Figure 2.5 – Snow dominated area from the mountainous area of British Columbia in Canada (WMO, 2008)

flow.

2.7.3. Glacial regimes

Like snow regimes, glacial regimes are represented by winter low flows with the difference that snowmelt starts later and lasts longer in the warm season. The main difference between these two kinds of regimes occurs after snow-melting. With increasing temperatures glacier melt increases after the snow has melted from the glacier's surface, while in a non–glaciated catchment, the streamflow recedes to a secondary low flow season (WMO, 2008). The amount of water during the dry season depends on the percentage of glacial coverage in the catchment as it possible to observe in the Blue River basin in Figure 2.6 covered for only 6% by glaciers. The regime in this case is similar to a snow dominated regime with a secondary low flow summer season.

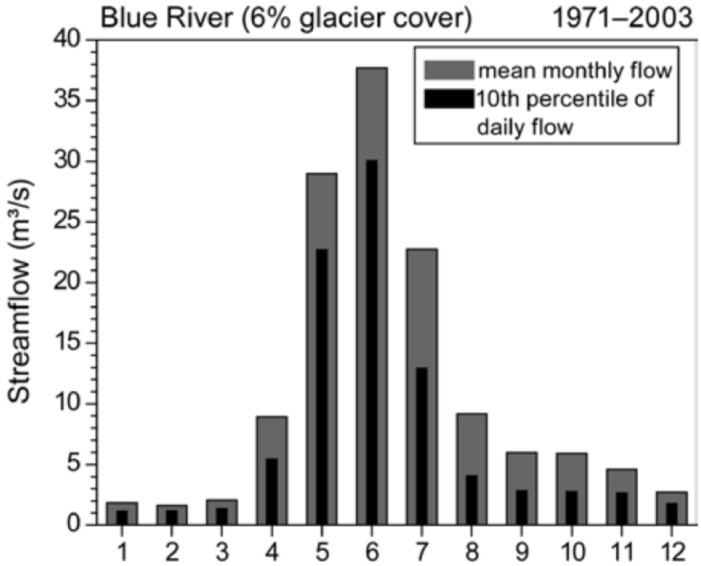

Figure 2.6 – Hydrological regime of Blu River Basin (WMO, 2008).

In many regions of the word, water obtained from snowmelt is an important source for water resources, in particular during dry seasons.

2.8. Low flow measures

The number of summary information derivable from daily time series that describes low flow regimes of a river is quite wide. The large number of methods is based mainly on the types of available data and the required output. Furthermore, it depends also on the not clear definition of low flow event, expressed as annual minima, as a threshold discharge or can be indicated by the time that the discharge is lower than a fixed value. Other reasons for diversification are the different methods of expressing the frequency. It is possible in fact, just as in extreme value analysis, consider it as the proportion of time that a discharge is exceeded (flow duration curves) or as the proportion of years that a given low flow occurs. Other methods can be obtained given the different durations or averaging periods that are required in many applications.

In this work the term "low flow measures" will be used to identify specific values and methods as well as graphics developed to analyze graphically the low flow regime of a river.

The domain of low flow measures can be defined arbitrarily. Following Smakhtin (2001), it is possible to specify an upper and a lower bound. Different values can be chosen for the upper boundaries, more or less conservative. The less cautious value is the mean annual runoff (MAR), which is the mean value of the available annual flow totals. Dividing this value by the number of seconds in a year it is possible to obtain the long term mean daily discharge, defined as mean daily flow (MDF). The other possible boundary is the Median Flow (MF) obtained ranking the time series. This value is more conservative because streamflow time series are often positively skewed and MF is often smaller than MAR.

The lowest value for the low flow domain can be identified in the absolute minimum value of the daily discharge (AMF).The information content of this value depends on the length of the time series and the presence of a lowest measuring threshold value for the gauging system. The lowest bound for low flow hydrology is surely zero that occurs on intermittent and ephemeral rivers.

2.8.1. Base flow and base flow index

Base flow is an important component of hydrograph obtained from groundwater storage or other sources. To recognize baseflow in the hydrograph, it is possible to use separation techniques that permit its division in a quick component and a delayed component. Through most of the dry season streamflow is composed essentially of baseflow, while in the wet season the hydrograph is the sum of the direct response of the catchment to rainfall and the water stored component.

Baseflow separation techniques try to estimate the surface runoff component of a flood and concentrate on baseflow separation starting from flood events (event based methods). Other types of methods try to generate baseflow hydrographs from long time series using some kind of digital filters that permits to obtain from the daily streamflow a quick flow and a base flow hydrograph. The most important methods can be found on Nathan and McMahon, 1990a, Chapman, 1991, Sittner et al., 1969; Birtles, 1978; Boughton, 1988; Smakhtin and Hughes, 1993, Pettyjohn and Henning ,1979, Sloto and Crouse 1996, Sokolov 1974, Potter and Rice 1987.

A review on baseflow separation methods can be found on Hall (1971).

The aim of baseflow separation techniques is to obtain objective quantitative indexes related to the long term baseflow response of a catchment. Such indexes include mean annual baseflow volume and long-term average daily baseflow discharge. However, the main low flow index obtained from base flow is Baseflow Index (BFI). This measure was introduced first by Lvovich (1972) and then developed by the Institute of Hydrology (1980) to estimate the influence of geology on low flows. The index is calculated as the ratio between the volume of the baseflow and the volume of total streamflow. The index is adimensional and range from 0 to 1. The highest values of BFI are obtained in catchments able to sustain river flow during dry seasons and namely where there is a high groundwater contribution. Lowest values of BFI, on the contrary are typical of ephemeral streams. Different authors found characteristic values of BFI for a number of rivers in certain regions (FREND, 1989; Smakhtin and Watkins, 1997).

Baseflow index is usually highly correlated with hydrological properties of soil, geology and other storage descriptors, such as lake percentage.

Knowledge of baseflow regimes is important for the development of catchment management strategies (especially for drought conditions) establishing relationships between aquatic organisms and their environment, estimation of small and medium water supplies, water quality and salinity management, calculating water budgets of lakes, etc, (Smakhtin, 2001).

2.8.2. Recession analysis

The shape of the recession curve, that is to say the falling limb of the hydrograph; reveals the gradual depletion of water stored in a catchment during periods of little or no precipitation. The recession curve reflects how different catchment storages and processes control the river outflow (WMO, 2008).

The main governing factor of the stream recession rate is the catchment geology (e.g. transmissivity, storativity of the aquifers) and the distance from stream channels to basin boundaries (Smakhtin, 2001). Rivers where groundwater has a high influence have a slow recession rate, whereas a fast rate is representative of rivers draining impermeable catchments with limited storage.

Overland, interflow and groundwater flow have different characteristic recession rates, also if the range of these rates can overlap. In low flow context the baseflow recession rate is the more important component. Baseflow recession can be expressed by various recession equations, (e.g. Toebes and Strang, 1964; FREND, 1989; Clausen, 1992; Wittenberg, 1994; Griffiths and Clausen, 1997) but the most common is the exponential one identified by a recession constant which is the parameter of the model.

There are two main estimation methods used to calculate this index and they mainly consist in the computation of the slope of a master correlation curve (MRC) obtained between the envelope of different recession curves, or in performing a separate calculation of parameters for each separate recession segments (IRS).

The MRC method tries to overcome the variability in individual segments by constructing a mean recession curve. Different methods exist to calculate a master recession curve such as the correlation method (Knisel, 1963; Hall, 1968; Brutsaert and Neiber, 1977; Beran and Gustard, 1977; Institute of Hydrology, 1980; FREND, 1989) and the matching strip method (Nathan and McMahon, 1990a; Sugiyama, 1996). Several authors highlighted the lack of objectivity of these methods and the fact that they do not give a satisfactory representation of the recession curves. For these reasons authors like Jones and McGilchrist, 1978, Petras, 1986, Bako and Hunt (1988) tried to construct alternative and more objective recession curves.

IRS procedure, in reverse, consists in calculating a recession constant for each individual recession segment. This method takes clearly into account the variability of recession curves. Sample statisticals such as mean and variance can be calculated on this set of recession constant.

Martin (1973) proposed another method, the half-flow period, that is calculated as the time in days required to cut by half the baseflow. This value can range from 7-21 days for quick baseflows and 120-150 days for slowly receding streams. Nathan and McMahon (1990a) showed that this method is more physically based than recession constant method.

Recession analysis is used for a wide range of engineering applications like water resource assessment, planning and management. Usually this technique is used for irrigation, water supply, hydroelectric power plants and waste dilution, estimating groundwater resources in a catchment, and hydrograph analysis (Anderson and Burt, 1980; Collins, 1982; Bako and Owoade, 1988; Radczuk and Szarska, 1989; Curran, 1990; Griffiths and Clausen, 1997; Moore, 1997). Boughton and Freebairn, (1985) and Korkmas, (1990) used recession curves for estimating the available drainage storage.

Two important uses of recession constant are those of Kelman (1980) that through these values fitted a stochastic daily streamflow model and Kottegoda et al., (2000) that applied recession characteristics for the generation of continuous daily streamflow time series.

A complete recent review of streamflow recession analysis and its various applications has been treated by Tallaksen (1995).

2.8.3. Low-flow frequency analysis

Most statistical methods are concerned with extreme values, in particular with minima values in low flow analysis. The low flow frequency analysis obtains the proportion of years when a flow is exceeded (return period or recurrence interval).

The Low Flow Frequency Curve (LFFC) is constructed on the basis of historical records of annual flow minima (daily or monthly minimum discharges of flow volume).

Usually different theoretical distribution are used to forecast beyond the limits of observed probabilities and to improve low flow estimates because the available observed flow records are not sufficient. In this case it is not possible to know the true parent distribution for low flows but it is important to associate reasonable distribution and estimate its parameters. The fitting procedure consists of different graphical and statistical methods that are used to decide the best theoretical distribution function. The distributions usually applied in literature are different forms of Weibull, Gumbel, Pearson Type III and log-normal distributions. Literature largely studies what are best distributions to fitting minima data in some regions and using different averaging intervals (Matalas, 1963; Jozeph, 1970; Prakash, 1981; Beran and Rodier, 1985; Loganathan et al., 1985; McMahon and Mein, 1986; Singh, 1987; Waylen and Woo, 1987; Khan and Mawdsley, 1988; Sefe, 1988; Leppajarvi, 1989; Polarski, 1989; FREND, 1989; Nathan and McMahon, 1990b; Russell, 1992; Loaiciga et al., 1992; Durrans, 1996; Lawal and Watt, 1996a; Lawal and Watt, 1996b; Bulu and Onoz, 1997).

Many of these studies are interested in evaluating the lower limit of the distribution.

An universally defined distribution to fit low flow data is not yet identified although many studies consider this problem (Vogel and Kroll (1989), Pearson (1995) and Vogel and Wilson (1996)). Various studies suggest the use of non–parametric methods for low flow frequency analysis in order to avoid the specification of a parent distribution (Tasker (1987), Loaiciga and Marino (1988) and Lall (1995)).

One of the main issues of low flow frequency analysis is the presence of observed streamflow time series of zero flow data. Zero data are typical of arid climates and very cold regions where the streams can be frozen in winter. Zero values can be present in a record also because the gauging station has a streamflow limit and the river level is below this threshold (censored data). The presence of zero flows has an important effect and it is not possible to ignore it in the statistical analysis of low flow series. Distribution fitted to series with zero flows will result in a positive probability of negative streamflows unless the distribution is explicitly constrained to have a lower bound of zero. Such results are physically meaningless (Smakhtin, 2001). In addition, the flexibility of distribution is reduced constraining it to zero lower bound. Haan (1977) suggested a procedure based on conditional probability adjustment. This procedure adds a new parameter that represents the probability that an observation is zero, while continuous distribution is fitted on non–zero data. The result is adjusted with the new parameter. Gordon et al., 1992; Stedinger et al., 1993). Bulu (1997) has used the theorem of total probability in low flow frequency analysis. Durrans et al. (1999), studied zero data caused by zero flow or censored data and suggested to treat them as censored data in uncertainty case. In arid and semiarid regions, rivers are often intermittent or ephemeral and the annual minimum value is often zero so in these cases, it is not possible to apply a low flow frequency analysis.

It is possible to obtain many information and indexes by LFFC. The more important indexes estimated by LFFC are:

- slope of LFFC;
- break in the curve near the modal value;
- 7-day 10 year low flow (7Q10) and 7 day 2 year low flow (7Q2);
- Dry Weather Flow.

The slope of LFFC can be estimated as the difference between two flow values, normalized by catchment area, from low and high probability domain. The steepness of the slope indicates the variability of low flow regime.

The LFFC can present a break in the curve near the modal value. Velz and Gannon, (1953) identified this as the point where a change in drought characteristics occurs. Flows characterized by higher frequencies are those that have a tendency to normal conditions and can not be considered as drought flows. Though not a general feature of the LFFC, this may be interpreted as a condition at which a river starts getting water exclusively from a deep subsurface storage (Smakhtin, 2001).

The 7-day 10 year low flow ($7Q_{10}$) and 7 day 2 year low flow ($7Q_2$) are the most used indexes in the USA. These values represent the lowest average flows that occur for a consecutive 7-day period at the recurrence interval of 10 and 2 years.

The dry weather flow is the annual series of minimum 7 day average flow (Hindley, 1973, Pirt and Simpson, 1983; Gustard et al., 1992). It is used mainly in the UK for abstraction licensing. The 7 days average permits to eliminate the problem of day by day variability of the river flow and also measurement errors.

Literature widely studied these indexes and made use of them in different fields like drought studies, design of water supply systems, classification of streams for potential waste dilution, regulation of waste disposals to streams, supply of certain in-stream discharges, etc. (Chiang and Johnson, 1976; Refsgaard and Hansen, 1976; Male and Ogawa, 1982; Aron and Emmanuel, 1982; Biswas and Bell, 1984; Riggs, 1985; Paulson and Sanders, 1987; Cumming Cockburn Ltd, 1990a, Lung et al., (1990))

Low flow frequency analysis belongs to extreme events frequency analysis but the study of it is limited compared to flood frequency analysis.

2.8.4. Continuous streamflow events and streamflow deficit

Information about the length of period below a selected threshold value is not provided by LFFC, neither by flow duration curves. Furthermore, these measures lack of information concerning the streamflow deficit that can be obtained trough a continuous time series. In fact,the construction of FDC excluded this information. Two time series that have a very different low flow sequence can have a similar shape of FDC. There are various existing approaches to overcome this problem. For example Velz and Gannon (1960), recommended to evaluate the durations of the longest periods which are necessary to obtain a defined small percentage (1–10%) of the MAR. These indexes are similar to those obtained by FDC but without loosing information on the time sequencing of discharges. It is necessary calculate these intervals for every year, rank these values and plot them in different ways.

Another approach is based on the theory of runs (Yevjevich, 1967) called threshold level method. This approach is used for yielding estimates of the frequency of low flow periods and is used in designing reservoirs where reservoir releases are made to support downstream abstractions. The run in this context is the number of days when a discharge is below a defined threshold flow. The method, called also threshold level method or “method of crossing theory”, is applied to a global dataset in Fleig et al. (2006). Figure 2.7 shows definition of timing, duration and volumes of deficit below a threshold discharge.

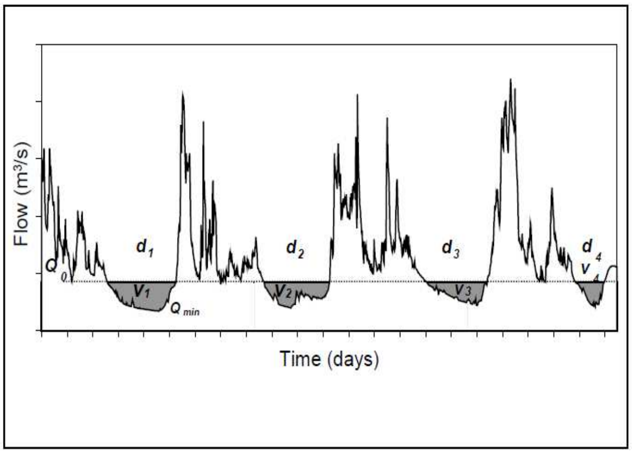

Figure 2.7 – Representation of threshold values and timing, duration and volumes of deficit.

The threshold value considers the purpose of the study, the type of flow regime and the available data. For perennial flow regimes, usually it is used as limit value of the discharge with 70-90% exceedence on FDC. Tate and Freeman (2000) instead suggested using the discharge with 20% exceedence on FDC as threshold in ephemeral rivers where there is a flow only after significant rainfall event. An alternative is to use a variable threshold. Defined the truncation level Q_0, it is possible to define also these deficit characteristics:

- The duration, which is the period of time where the flow is below the threshold level. This period is defined as drought duration, low flow spell, or run length (d_i in the Fig 2.7);
- The volume or severity, that is he cumulative water deficit (vi in the Figure 2.7);

- The intensity, also defined as drought magnitude, that is the ratio between deficit volume and deficit duration;
- The minimum flow of each deficit event (Q_{min} in the Fig 2.7).

Using the series of deficit characteristics it is possible to derive other indexes as the average deficit duration or average deficit volume. Additionally, it is possible to rank these data, assign a probability or return period trough plotting position and plot them against an assigned return period fixed the threshold value. It is then possible to associate a probability distribution to these points.

An example of application of this theory can be found in Zelenhasic and Salvai (1987). They used daily stream-flow time series to obtain all components of drought deficit, duration, start date, number of continuous events in a given time interval, the largest stream-flow deficit and the largest duration in a given time interval. The Institute of Hydrology (1980), using different names, handled the same problem.

Smakhtin and Watkins (1997) suggested to analyze the durations (or deficits) and number of spells below a threshold flow and to display the results in the form of a histogram and/or a cumulative frequency curve. These methods want to give a quick impression of how responsive the river is on the basis of the available record.

Literature largely studied run analysis to identify and characterize annual hydrological droughts (Weisman, 1978; Dracup et al., 1980; Sen, 1980a,b; Kosinsky, 1984; Tlalka and Tlalka, 1987; Chang and Stenson, 1990; Wijayaratne and Golub, 1991; Clausen and Pearson, 1995; Moye and Kapadia, 1995; Burn and DeWit, 1996; El-Jabi et al., 1997; Cigˇizogˇlu and Bayazit, 1998).

The use of this approach is required in those fields where continuous available river discharge is needed for domestic water supply, irrigation, power generation, dilution of industrial pollutants, recreational planning, fish migration etc. Moreover, it is used to assess the storage capacity of a reservoir (Pegram et al., (1980), McMahon and Mein (1986), Klemes (1987) and Midgley et al., (1994)).

It is also possible to study in ephemeral rivers the frequency of durations of continuous zero-flow periods using common statistical procedures. This problem is a specific case of run analysis which indicates the likelihood of extended periods of no–flow or drought (e.g. Armentrout and Wilson, 1987).

2.8.5. Flow duration curves

The flow duration curve (FDC) is a cumulative frequency curve that shows the percentage of time that a given discharge is equalled or exceeded during a fixed period. It is one of the most informative methods of displaying the flow characteristics of a stream throughout the range of discharge, without regard to the sequence of occurrence. It is possible to observe this curve as the relationship between magnitude and frequency (Smakhtin, 2001). This graph in the usual construction does not have a probability meaning because discharge is correlated to successive time intervals and discharge characteristics are dependent on the season; hence the probability that discharge on a particular day exceeds a specified value, depends on the discharge on preceding days and of the time of the year (Mosley and McKerchar, 1993, Ganora et al. (2010)). Despite this, the FDC is often seen as the complement of the cumulative distribution function of the considered streamflow at a site. FDC construction is attributed to Clemens Herschel and is dated back to 1880 whilst the general use of the curve is dated 1915, and its theory is discussed in Foster (1934). Searcy, (1959) summarized the properties and application of period of record FDCs and this topic is yet widely studied as evidenced from a review article [Smakhtin, 2001] as well as in papers on the regionalization of FDCs [Croker et al., 2003], uncertainty analysis of FDCs [Yu et al., 2002], the development of a stochastic model for FDCs [Cigizoglu and Bayazit, 2000; Sugiyama et al., 2003], and use of FDCs for watershed management [Good and Jacobs, 2001].

2.8.5.1. Construction of flow duration curves

Two main methods exist to construct flow duration curves and they are the following:

- The total period method (FDC);
- The calendar year method (AFDC).

A FDC is constructed by ranking the data in a decreasing order of magnitude, assigning flow values to class intervals and calculating the number of occurrence in each class. Afterwards it is possible to calculate cumulated class frequency and express them as a percentage of the total number of the time steps in the record period. At last, lower boundaries of every discharge class are plotted against the percentage point, (Searcy, 1959)

Another method to calculate these curves is through the definition of FDC as the complement of the cumulative distribution function based on the complete period of record. Vogel and Fennessey in 1994 introduced a variety of non parametric approaches for calculating the FDC. The procedure for the determination of the curve is simple:

- rank the observed discharges data in ascending order;
- plot each observation Q_p (p-th quantile of daily streamflows) versus its duration D_i or exceedence probability p.

The excedeence probability p is expressed as:

$$p = 1 - P\{Q_p \leq q\} \qquad (2.1)$$

It is possible to calculate p using the Weibull plotting position after which it is possible to define the duration D_i as:

$$D_i = 100(p) = 100 \cdot \left(1 - \frac{i}{n+1}\right) \text{, for } i = 1,2,3,\ldots, n \qquad (2.2)$$

where n is the length of the sample.

The calendar year method or annual based flow duration curve is firstly defined by Barrows, (1943), and Saville and Watson, (1933) and reintroduced by LeBoutillier and Waylen (1993) and Vogel and Fennessey (1994).

The method of calculation of the AFDC is the following:

- for each year or hydrological year a FDC is calculated using the procedure of the period of record FDC;
- for each exceedence probability p using all annual FDC it is possible to calculate the measures of central tendency and namely mean and median (by summarizing the interannual variability of the different annual FDC).

Starting by this construction Vogel and Fennessey (1994) introduced also confidence intervals for quantile of flow duration curves. This is done because when one focuses on a particular quantile or percentile of an FDC, the FDC does not, by itself, expose the uncertainty associated with a particular quantile estimate (Vogel and Fennessey, 1994). The construction of non–parametric confidence intervals can be easily done considering the n values of discharges for each order as a random sample from which one can estimate $100(1 - p)$%.

The AFDC and the FDC are complementary rather than competitive concepts (Castellarin, 2004). FDCs shows the entire range of observed river discharges and, Fennessey (1994), Hughes and Smakhtin (1996), and Smakhtin et al. (1997) showed that they can be effectively used for filling gaps and for extending daily streamflow series, and, when a regional FDC model is available, to generate streamflow series at ungauged river basins. This kind of curve presents, however, a negative implication: for example, Searcy (1959, p. 21) warns that, "To say that a flow-duration curve based on a 15-year record represents the distribution of the yearly flow is incorrect." In fact the definition of quantile Q_p, calculated using the procedure of the period-of-record FDC, is the discharge that was exceeded p percent of the time over the entire period of record on which it is based. If the period of record is long enough to obtain the limiting distribution of streamflow or if the particular period of record corresponds to a particular design life, this interpretation may be quite useful. This differs from flood and low-flow frequency analysis, where the interpretation of the frequency curve does not depend on the period of record, (Vogel and Fennessey, 1994). Obviously the number of years used for the frequency analysis leads to less sampling errors in quantile estimates, and therefore to more precise frequency curves.

At opposite AFDC have been shown to be quite useful for making probabilistic statements about wet, typical and dry years, for computing confidence intervals associated with the AFDC representing the typical hydrological condition and for assigning return periods to individual AFDCs [Vogel and Fennessey, 1994]. Besides, the median annual FDC is not influenced by the occurrence of extreme low-flow periods or extreme floods over the period of record, yet it still captures the frequency and magnitude of daily streamflow in a typical year. Since their introduction, a number of investigators have found AFDCs to be quite useful to solve a wide range of problems (Claps and Fiorentino, (1997); Smakhtin and Toulouse, (1998); Good and Jacobs, (2001); Sugiyama et al.,(2003)).

Different stochastic models are created to relate FDC and AFDC. The stochastic models are formed by a deterministic component that must reproduce the seasonality associated with daily flow series and by a stochastic component that must reproduce both the persistence and frequency distribution of the daily flow series.

LeBoutillier and Waylen (1993) introduced for first a five parameter stochastic model of daily streamflows which relates the FDC to the AFDC. The stochastic model developed by LeBoutillier and Waylen (1993) can reproduce the AFDC for a typical year though their model significantly underestimates the variability of observed AFDCs around the central AFDC. Later Castellarin et al. (2004) developed a mathematical model based on the index flood method that relates the period of record flow duration curve (FDC) to the mean and variance of the annual flow duration curve. This approach permits to construct confidence intervals associated with AFDCs at ungauged sites, (2) assign return periods to individual AFDCs (3), develop regional models of flow duration curves, (4) generate daily streamflow series at ungauged sites, and (5) develop a generalized stochastic model of daily streamflow, (Castellarin et al. (2004)).

Other probabilistic and parametric representations of FDC have been suggested by Quimpo et al., (1983), Mimikou and Kaemaki (1985), Fennessey and Vogel (1990) and LeBoutillier and Waylen (1993).

2.8.5.2. Time Units

FDC can be constructed using different time resolutions: daily, weekly, monthly and annual. The most informative time resolution is the daily one. The construction of the FDC can be done also using other time resolution, more appropriate to the studied problem by using the moving average method and then calculating these curves using m-day or m-months average time series. The choice of the time unit, such as the day, the week or the month depends on the use of the graph. The details of the variations in flows are obscured if the time unit is long (Searcy, 1959). Using monthly or annual data could be unsatisfactory for many streams because does not permit to evaluate the variability of the flow.

The effects of varying time unit are different depending on the different river regime. If the flow of the river is regular day by day, the change of the time, from daily to weekly to monthly time unit will not change significantly the curves. On the other hand, on "flashy" rivers characterized by sudden flows that last only few hours, the change of time unit will have a really large influence.

More details on the use of different time units can be found on Searcy, 1959.

2.8.5.3. Standardization

To improve the readability of the curve, discharge is often plotted on a logarithmic scale and sometimes the frequency axis is represented on a normal probability scale. When the logarithm of daily mean flow is normally distributed, the plot will follow a straight line.

This procedure facilitates the readability of the graph and permits to compare different curves.

Standardization through division by the average flow is used to compare the regime of different basins. The shape of FDC is connected with the physiographic and climatic characteristics of the basin, and therefore to the catchment response (WMO, 2008). A time series with low variability is reflected in a flat curve typical of permeable catchments, or with a strong regularity caused by the presence of storage (Figure 2.8).

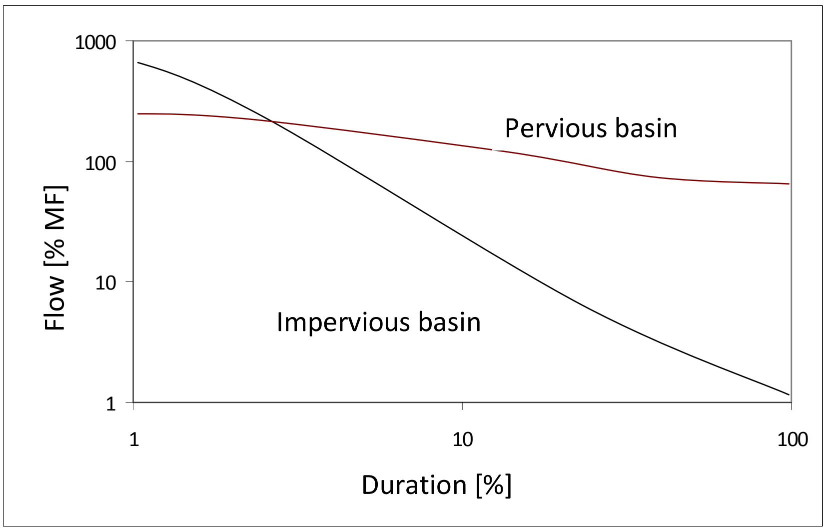

Figure 2.8 - Different curves for different permeability of the soil (WMO, 2008).

The FDCs characterized by steep curve are typical of impermeable catchments with high variability and sudden flows (Fig 2.8).

Fig 2.9 shows the different shapes of curves from various climatic regions. River Honokohau from tropic regions with warm and humid climate, where there is not a distinct dry season, is perennial and characterized by a large variability. The Arroyo Seco at Soledad, California, positioned in a temperate dry summer climate is characterized by high inter–annual precipitation variability and the flow becomes zero.

River Lambourne in the United Kingdom is located in a temperate maritime climate without a distinct dry season. The FDC shows little variability in the streamflow and the presence of a summer baseflow that influences the outflow.

Nigardselv River in Norway has a cold climate with no distinct dry season. The catchment has 75% coverage and in the early summer season snowmelt increases the summer flow. The FDC is steep reflecting high-flow variability with low flow in the winter and higher flow in the summer.

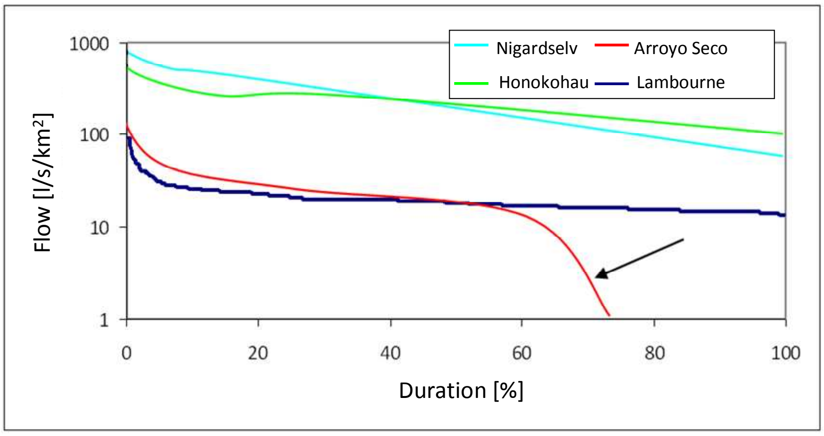

Figure 2.9 - Different FDC from different climatic regions (WMO, 2008)

The shape and general interpretation of any FDC has been directly or indirectly illustrated by Searcy (1959), Vogel and Fennessey (1994), Hughes and Smakhtin (1996), Mngodo (1997) and Smakhtin et al. (1997).

2.8.5.4. Low flow indexes obtained by FDC

Various indexes can be estimated by FDC. Particular importance has the low flow section of the FDC which is the part of the curve below the Mean Flow or Q_{50}. This part of the curve gives information about the storage capacity of the river and the contribution of the baseflow to the streamflow. The slope of the curve, steep or flat, indicates the baseflow contribution and the permeability of the basin.

Other indexes can be extracted from FDC. In particular Q_{95}, the 95 percentile flow, or rather the flow that is exceeded 95% of the period of record is a key index of low flow. The percentile used as low flow index depends on the type of the river being studied. However, the flows within the range of 70-99% time exceedance are usually widely used as design low flow. Some used example indices are: one- or n-day discharges exceeded 75, 90, 95% of the time called $Q_{75(7)}$, $Q_{75(10)}$, $Q_{90(1)}$, $Q_{95(1)}$, $Q_{95(10)}$.

Other indexes can be estimated to evaluate streamflow variability such as the ratio $Q_{20/Q50}$ and the ratio $Q_{50/Q90}$, which in particular represent the low flow variability. The percentage of time that a stream is at zero flow conditions, indicates the intermittency characteristics of a river (Gorgens and Hughes, 1982; Smakhtin et al., 1995).

2.8.5.5. Application of FDC

FDCs are widely used in hydrology and engineering. In particular, FDC is a key tool for the sustainable management of water resources. The management of abstractions requires estimating FDC from gauged data or from ungauged sites through statistical or physical models in order to obtain:

- the natural regime;
- the historical regime that includes the impact of abstraction and discharges returned to the river;
- a target regime that maintains the ecology of the river at an acceptable level (WMO, 2008).

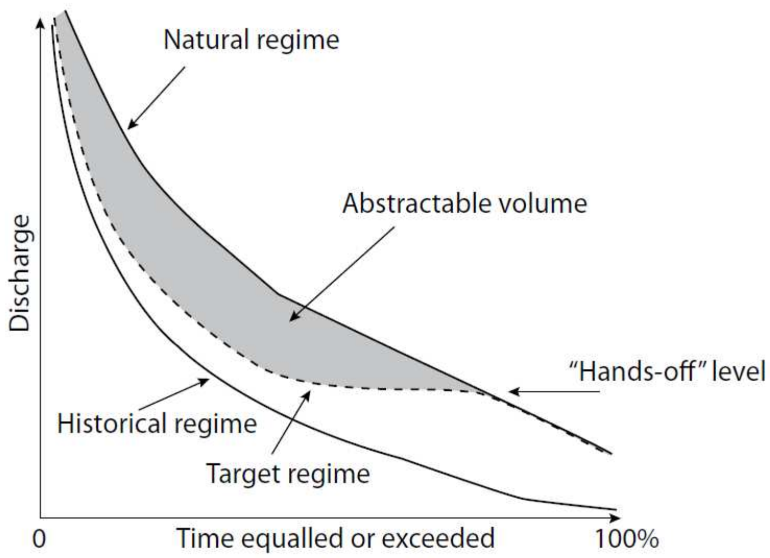

Figure 2.10 – River abstraction concepts obtained by an FDC (WMO, 2008).

From FDC estimated by a model for ungauged sites or calculated from the data for the gauged one, it is possible to evaluate the abstractable volume and the hands–off flow. The abstractable volume represents the amount of water that can be withdrawn from a river without creating problems to the river's ecosystem and adverse impact on the downstream uses. The hands–off flow is the discharge after which the abstraction of the discharge has to cease. Fig 2.10 shows the river abstraction concepts based on FDC. Pirt and Simpson, (1983); Gustard et al., (1992); DWAF, (1995); Mhango and Joy, (1998) studied the application of FDC on this field.

Another important use of FDC is the analysis for public water supply and agriculture. FDCs are used mainly for the preliminary design of simple abstraction schemes in public water supply projects whilst the primary application of water for agricultural use is to supply water for irrigation. The planning's key element is the assessment of water availability for the definition of potential areas that can be irrigated by the supply river. The assessment is carried out by considering the critical period of the year. In this case a seasonal FDC can be used to compare potential demand to available water.

FDC is also used to estimate the dilution of domestic or industrial discharges intended for a river. Normally a legal authorization is necessary to discharge waste water in a river. Water quality models are used to evaluate the quality of the receiving stream and to determine the frequency distribution of the downstream water quality. The Q_{95} obtained from the FDC is the most used flow variable to evaluate downstream water quality and to determine a constraint on discharge consent.

Other important use of FDC is the evaluation of available water for hydropower. In fact, small–scale hydropower schemes use water from the river without artificial storage. The conventional method to evaluate the availability of water in a river is the FDC. The design must evaluate the fluctuating power demands, the protection of downstream abstractors' interest and the ecosystem's health (WMO, 2008).

Ecosystem protection is one of the many uses of FDC. The low flow period create a critical situation for the rivers and their habitat due to the reduction of the river's depth, the increasing temperature and the reduction in the dissolved oxygen. Percentiles from FDC can be used to estimate a minimum value of the river discharge in a way that if the discharge falls under this value, abstraction should be reduced or ceased.

Literature has widely referred to the several applications and studies on FDC. Vogel and Fennessey (1994) showed the several applications of the FDC in the hydrological practice. Searcy (1959) was the first to summarize the FDC applications including catchment geology on low flow, hydropower and stream water quality uses. Warnik (1994) showed the use of FDC for hydropower feasibility studies for river run operations. Male and Hogawa (1984) suggested the use of FDC in the project phase of waste water treatment plants when a trade off is necessary between water quality requirements and costs. Alaouze (1989, 1991) estimated the optimal release schedule from reservoirs by using a procedure based on FDC. Pitman (1993) and Mallory and McKenzie (1993) used FDCs in flow diversions designs. Estes and Osborn (1986) and Gordon et al., (1992) illustrated the use of FDC for the assessment of river habitats in estimation of instream flow requirements. Hughes and Smakhtin (1996) showed a non–linear spatial interpolation approach based on FDC to extend the observed daily time series and to generate flow time series at ungauged sites (Smakhtin et al., 1997), and to restore the natural streamflow sequence at ungauged sites (Smakhtin, 1998). Gustard and Wesselink (1993), Lanen et al. (1997) and Smakhtin et al. (1998a)) used a FDC as a tool for rainfall–runoff model calibration and/or for the comparison of flow-time series simulated for different scenarios of development. Hughes et al., (1997) proposed a model which is based on FDCs and is designed to convert the original tabulated values of estimated ecological instream flow requirements for each calendar month into a time series of daily reservoir releases. Good and Jacobs, (2001) published works concerning the use of FDCs for watershed management.

3

Case study and preliminary analysis

3.1. Introduction

In hydrology, the analysis of geomorphologicalal and climatic characteristics with the greatest influence on the surface runoff and consequently on the hydrological response of a basin, has always been a very important subject. The existence and definition of this relationship is interesting both in theoretical and practical applications. On the one hand a framework for the role of geomorphologicalal and climatic properties on the formation of the water flow (either for extreme, maximal, minimal discharges or regular conditions) is necessary. On the other hand, these relations may be very useful for the evaluation of the water flows in areas where the discharge measurements are either insufficient or non-existent.

Following this concepts the aim of this note is to show the phisiographical characteristics of studied area, to underline the hydrometric and geomorphologicalal data used and evaluate the extent of the correlation between these data. At last a paragraph that contains an homogeneity analysis of the discharge data to identify hydrometric data characterized by the same hydrological distribution is shown. This analysis and results obtained will be the starting point for the following chapters.

3.2. Central Italy study area

The basins of the target area include the Tiber basin as well as its sub-basins, the basins located in northern Lazio and those located north and south of the Tiber River. These areas show considerable lithological variability that affects the large geomorphologicalal structures.

3.2.1. Geological characteristics

The geology of the area is connected to volcanic activity, since one-third of the region is covered with this kind of geological substrate. The volcanic area has developed in a NW-SE direction, while the north-western part of the target area involves a metamorphic substrate dating back to the Paleozoic period.

The Volcanic period in the region starts from the end of Pliocene Epoch and creates an acid volcanic activity. Successively four volcanic districts has developed with high amount of potassium. The four districts defined by this geological activity are:

- Vulsino District (starting from Toscana region and characterized by the presence of a big volcano depression, now took up by Bolsena Lake);
- Vicano District (it is placed on the south of Bolsena Lake, it was a central structure, characterized by a typical shape (stratovolcano), with the highest part of the mountain occupied by the caldera);
- Sabatino District (it is possible encounter this large and recent volcanic area (since 600000 to 40000 years ago) going to the south east part of the Lazio Region. The activity has a typical explosive character with the centre going from Sacrofano town to Monte Razzano, Monte Sant'Angelo and Baccano)
- Colli Albani District (this recent district (since 60000 to 20000 years ago) is connected with the Apennines mountain range. This volcanic area has the huge extension (1500 km2), and biggest lava volumes (290 km^3).

The other important geological domain is the Apennine dorsal positioned in the south-east of the region. This macro system is mainly composed of carbonate sediments dating back to the Mesozoic period and deposited in different sedimentation environments. Most recent sediments in the area are quaternary deposits that make up coastal plains and river valleys.

Looking for the others geological domains it is possible to underline the presence of a metamorphic substrate dated back to Paleozoic Period and positioned in the northern part of the studied area level with Fiora Basin.

From a tectonic point of view the orogenetic action that create the actual structure of the Apenninic dorsal is dated in the Neogenic period and is developed in different phases.

Most recent sediments in the area are quaternary deposits that constitute coastal plains and river valleys. Particular importance have the Agro Pontino that is constituted by eolic sediments that represent ancient dune sand.

3.2.2. Geomorphology

The geomorphological characteristics of the region are closely connected to the geological domains. In fact, the big geomorphological domains coincide with recognized geological structures: the big volcanic districts, the Apennine dorsal, the coastal plains and the remaining Tiber valley. In these big geomorphological structures, it is possible to identify uniformity and therefore distinctive morphotypes.

- Alkaline-potassic volcanic districts characterized by central activity (Vico and Colli Albani). It is possible to recognize big central structures, characterized by the typical volcanic shape, with little slope and wide calderas in the upper parts;
- Alkaline-potassic volcanic districts characterized by areal activity (Vulsini and Sabatini districts) are more flat than previous and characterized by the presence of many emission centres morphologically easily detectable.

The hydrographic net of the volcanic districts is easily identifiable for the centrifugal drainage pattern, for the steep rock face and the steps caused by the alternation of the different lythologies.

Besides the area studied contains different types of karst structures.

3.2.3. Hydrographic structure

The hydrographic structure is controlled by the Tiber River system in the northern part of the region and by the Liri-Garigliano river systems in the south of Tiber. The Tiber River basin covers about 17.200 km^2 and represents the main watercourse of the area.

The Tiber River has an Apennine trend in its initial reach and flows with a torrential regime. Along its right bank, the river collects water from different volcanic districts. From its left bank, it receives water from the Apennine carbonate structures. These contributions stabilize the regime. At the level of confluence with Farfa River the Tiber river change direction and assume a trend transversal to the previous part. In this reach in the right bank it receives the water from the Sabatini structures, while in left bank it receives water from Aniene River that drains Prenestini Mountains and Colli Albani. The Tiber River shows a large difference in hydrographic structure between the basins belonging to the right and left banks. This difference is due to the different ways in which the volcanic systems move, compared to those of carbonate structures characterized by a lower drainage density.

The Liri-Garigliano basin make up about 4.900 km^2 and, with the Tiber River, supply the area with 80% of the total runoff. Within these river basins, the permeability characteristics and morpho-topographic structures, mainly represented by carbonate platform deposits, determine the highly effective infiltration and consequently slow development of the hydrographic network and low overland flow.

The river basins in northern Lazio have been formed on geological formations with low permeability and a hydrological regime characterized by high overland flow from autumn to winter, when their discharges are 3-4 times higher than those in summer.

The karst system is also particularly developed both in the mountains in the north-east area and in the Apennine dorsal, where there are mostly extended karst shapes of large dimension. Moreover, this area has a highly variable climate due to two major bio-climate regions, temperate and Mediterranean, and the relative transitional regions.

The significant differences in the study area are highlighted in the map of the digital elevation model in Figure 3.1, where it is possible to recognize main basins belonging to the region.

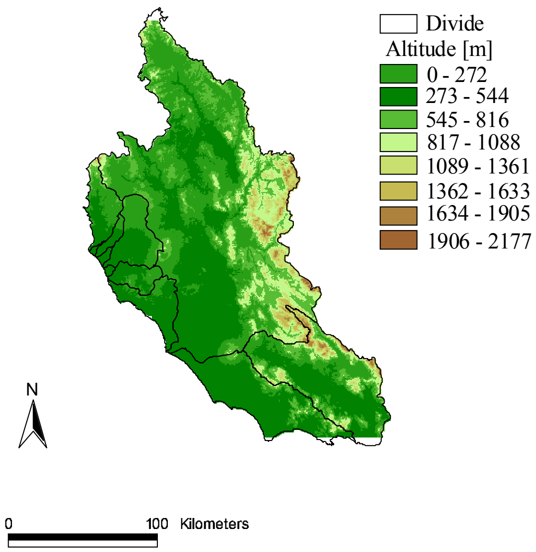

Fig 3.1 - Digital Elevation Model of the studied area. The main basins in the area are divided with the black lines.

The hydrological data used for the analysis came from the Ufficio Idrografico e Mareografico of Lazio Region. At least 6 years' worth of daily recorded discharges from 26 stations in the study area were used.

Quantiles Q_{95}, Q_{50} and Q_1, were estimated. In particular, minimum discharge Q_{95} is widely used in Europe and was chosen because of its importance for many applications relating to water management, as in Gustard et al. (1992), Smakhtin (2001) and Laaha and Bloshl (2007).

Each station selected for this study was considered as a separate basin. Thus for basins that do not have upstream stations, the quantiles were calculated directly from the discharge data, by adapting the data of each station to the best possible distribution and then calculating the relative quantiles.

The nested basins were, however, divided into sub-basins separated by several measurement stations; the quantiles were calculated as the difference between the quantiles of the upstream and downstream stations. This last estimation is more robust than the quantile calculated from the difference between the hydrographs, but it requires the mean maximum and minimum flow rates to be synchronized in different stations. Furthermore, this method reduces the spatial dependence of discharge data (Laaha and Bloshl, 2007). It is necessary to bear in mind that, in the absence of isochrones, error may also be high. All quantiles Q_{95}, Q_{50} and Q_{1} have also been standardized with respect to the area of the sub-basin to obtain the specific quantiles q_{95}, q_{50} and q_{1} respectively $[m^3\ s^{-1}\ km^{-2}]$. Figure 3.2 shows the geographical representation of quantiles q_{1}, q_{50} and q_{95}, where it is evident that the largest values of specific quantiles are in the south-east of the area that coincides with the Apennine dorsal.

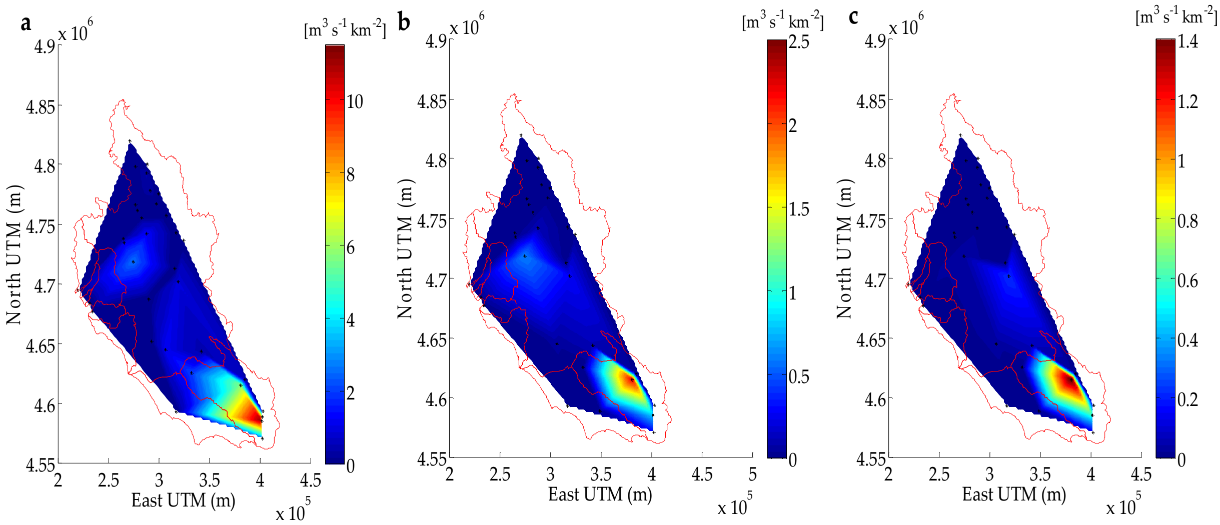

Fig 3.2 - Geographical representation of specific quantiles q_{1} (a), q_{50} (b) and q_{95} (c). The biggest values of quantiles are located in the south-east (Apenninic area), and the dimension of the region of interest decreases from higher (q_{1}) to lower quantiles (q_{95}). The black points in the figures represent the gauge stations in the study region.

In addition, the climatic characteristics were calculated by means of rainfall data from 118 rain gauges over the period 1985-2009. The average annual rainfall over the basin was calculated using the Thiessen method. Finally, the coefficient of variation of annual precipitation was calculated and then used to represent the temporal variability of rainfall in the area. Figure 3.3 represents the mean annual precipitation on each sub-basin and the coefficient of variation of annual precipitation. Again, in this case the orography influences the amount of rainfall over the basin, while the coefficient of variation is lower in the Apennines and higher in the coastal areas. The layout of the digital terrain model, precipitation, variation coefficient and quantiles of the discharge (in particular the maximum) highlight how strongly the orography influences the hydrological behaviour of the region.

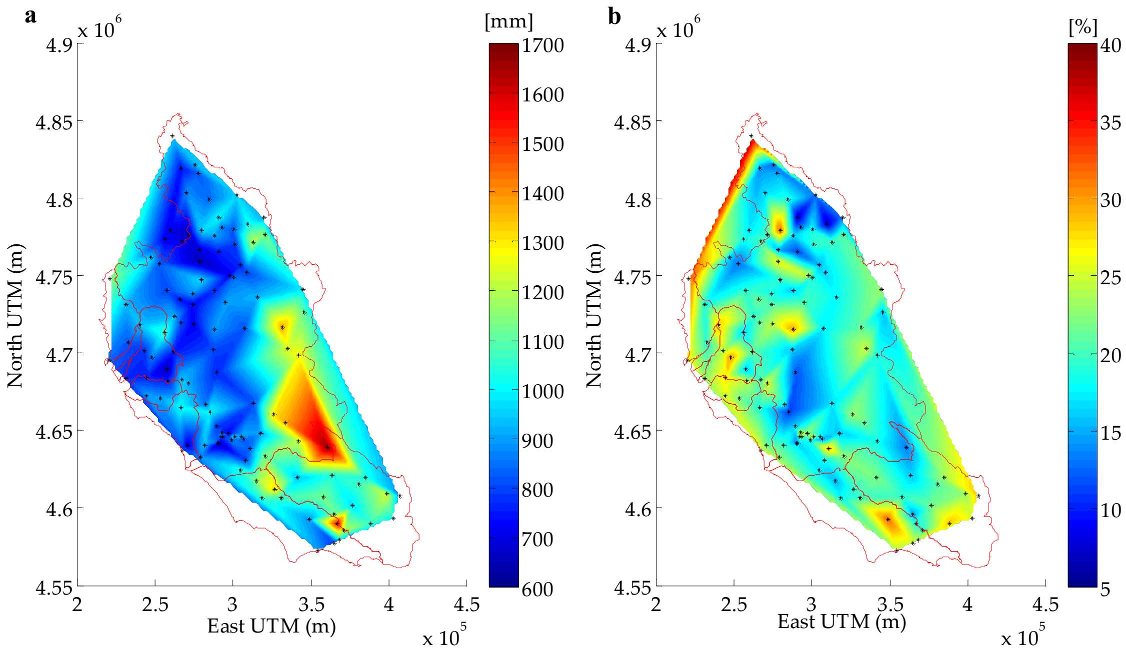

Figure 3.3 - Geographical representation of the mean annual precipitation MAP (a) and of the coefficient of variation CV of annual precipitation (b). The black points represent the locations of rain gauges in the region.

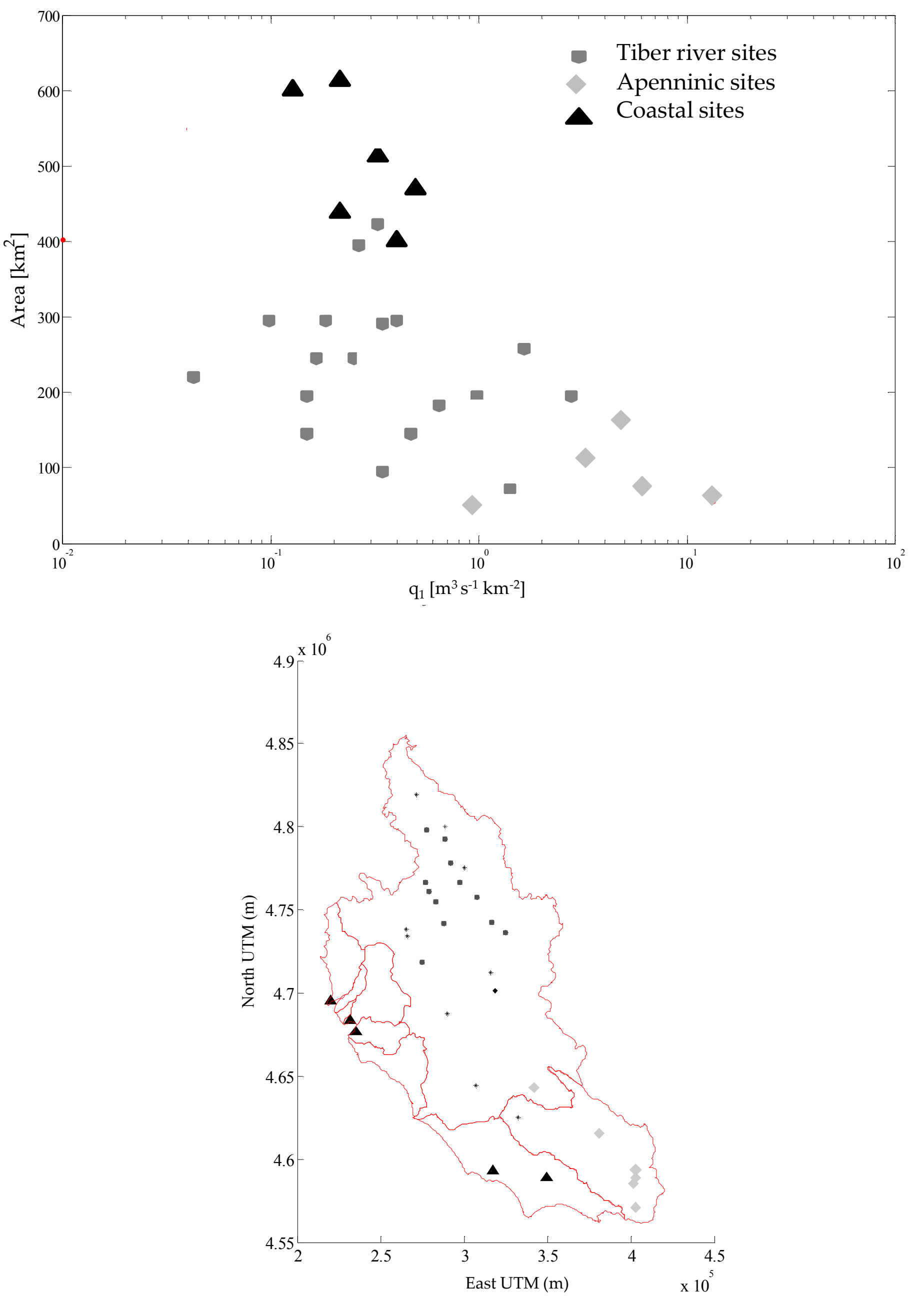

Figure 3.4 - The scatter plot of basin area against specific quantile of discharge q_1 used as qualitative analysis for preliminary identification of structures in the dataset (a) (semi-logarithmic representation). It is possible to recognize three different regions. In the figure below (b), points of the scatter plot are also identified geographically. The geographical representation points represent the flow level stations.

Figure 3.4 shows the basin area against the specific quantile of discharge q_1. This scatter plot was used as preliminary qualitative analysis to individuate the presence of structures in the data. Using the scatter plot, it was possible to identify the presence of three data sets: Coastal, Apennines and Tiber basin stations were identified. The same type of behaviour is present in the scatter plot of quantiles q_{50} and q_{95}.

These groups are formed according to the climatic conditions of the area. In particular, it is possible to note how the stations characterized by greater discharges are those in the Apennine region, characterized by higher precipitation depths.

3.3. Application and results

3.3.1. Selection of geomorphologicalal variables and correlations analysis

Correlation analysis allows for the evaluation of variables that influence quantiles in the target area.

The choice of characteristics for the basin that were considered for this study was guided by the analysis of the interactions among the flow regime, climate and physical characteristics.

The geomorphologicalal characteristics were calculated for each basin, using GIS data to obtain the area and maximum height of the basin. The Corine Land Cover (CLC) map, taken from the CORINE program in Lazio and Umbria, was used to evaluate soil use. The hydrogeological map of the area was used to define the percentages of different lithological structures for each basin.

Table 3.1 shows the variables as well as the symbols used. Table 3.2 shows the minimum, average and maximum values of some geomorphoclimatic indexes. The characteristics considered are sub-basin area A (km^2); the maximum, average and minimum elevation of the basin (H_{max}, H_{mean}, H_{min}) in meters above sea level (m a.s.l.); the value $\Delta H = H_{mean} - H_{min}$ in meters; the percentage of impervious substrate in the basins (F_D); and Mean Annual Precipitation MAP (mm) calculated for each basin. The values in Table 3.2 demonstrate the high heterogeneity and complexity of the study region.

Table 3.3 shows the correlation matrix estimated for all the dependent variables with data from the study region. Table 3.4 shows p-values at the 0.05% level when testing the hypothesis of no correlation; small p-values indicate significant correlation. Table 3.3 shows that the variables most highly correlated with specific quantiles are area, elevation of the basin and lithological characteristics. The correlation of specific quantiles with the area is particularly important. It can be noted that the coefficient diminishes from minimum to maximum discharges. One hypothesis is that the maximum discharges are influenced not only by the basin area but also by the average rainfall, which causes high flow.

The average annual rainfall influences the maximum discharges, which increase as expected when rainfall increases. The coefficient of variation appears to not greatly affect the discharges. The lower discharges are strongly linked to the lithological characteristics of the soil (Nikic and Radonja, 2009).

In particular, the percentage of carbonate and impervious substrate and thus the permeability characteristics of the basins are strongly linked to minimum discharges. This is probably because higher minimum discharges occur in basins with groundwater flow even in summer. This would also explain the negative correlation with the percentage of impermeable substratum F_D. Altitude, on the other hand, mainly affects maximum discharges. In particular, it was noted that the stations with the highest discharges are those in the Apennines, which are characterised by higher precipitation depths. This can be explained by the phenomenon of orographic rainfall.

Furthermore, Tables 3.3 and 3.4 show large p-values and thus low correlations, probably caused by the wide variability of the study area.

Table 3.1 - Predicting variables and annotation.

A	Area	km^2
H_{mean}	Mean Altitude	m
MAP	Mean annual precipitation	mm
CV	Coefficient of Variation of precipitation	%
F_A	Calcareous substrate	%
F_D	Impervious substrate	%
G_A	Agricultural areas	%
G_B	Forested areas	%

Table 3.2 - Minimum, average and maximum value of the geomorphologicalal and climatic indexes for the basins in the region.

	A (km^2)	H_{min} (m)	H_{max} (m)	H_{mean} (m)	ΔH (m)	F_D (%)	MAP (mm)
Minimum	31.08	2.0	389.0	32.3	31.3	0.00	650.00
Average	400.05	106.4	1414.6	479.7	396.5	6.53	1066.15
Maximum	981.23	368.0	2200.0	2031.0	2012.0	65.26	1350.00

Table 3.3 - Correlation Matrix Between Discharge Quantiles and geomorphologicalal Characteristics.

	q_1	q_{50}	q_{95}	A	H_{mean}	MAP	CV	F_A	F_D	G_A	G_B
q_1											
q_{50}	0.75										
q_{95}	-0.34	-0.23									
A	-0.45	-0.48	-0.49								
H_{mean}	0.42	0.33	0.31	-0.08							
MAP	0.34	0.01	0.05	0	0.13						
CV	-0.09	-0.13	0.02	0.19	0.33	0.12					
F_A	0.58	0.08	0.44	0.11	-0.17	-0.29	0.04				

F_D	0.20	0.13	-0.44	0.32	-0.11	-0.62	-0.19	0.42			
G_A	0.21	0.54	0.16	0.22	0.02	0.41	0.52	-0.04	0.22		
G_B	-0.41	-0.26	0.12	-0.32	-0.29	0.12	-0.23	-0.42	0.16	0.02	

Table 3.4. p-value Matrix to Test the Hypothesis of no Correlation (critical value equal to 0.05%).

	q_1	q_{50}	q_{95}	A	H_{mean}	MAP	CV	F_A	F_D	G_A	G_B
q_1											
q_{50}	0.02										
q_{95}	0.25	0.43									
A	0.12	0.08	0.09								
H_{mean}	0.13	0.25	0.26	0.79							
MAP	0.19	0.98	0.94	0.99	0.66						
CV	0.79	0.66	0.98	0.51	0.25	0.67					
F_A	0.05	0.79	0.11	0.72	0.54	0.32	0.95				
F_D	0.46	0.66	0.11	0.26	0.72	0.01	0.51	0.13			
G_A	0.45	0.05	0.56	0.45	0.98	0.13	0.06	0.95	0.43		
G_B	0.13	0.37	0.67	0.25	0.32	0.67	0.39	0.12	0.56	0.98	

3.3.2. Cluster analysis trough homogeneity test

Hierarchical cluster analysis is used to identify relatively homogeneous groups of variables based on selected characteristics, using Ward's algorithm (Ward, 1963), the most commonly used agglomerative clustering technique in the regionalisation context. This technique produces spherical clusters that are all approximately the same size. The aim is to link various objects into clusters using a measure of similarity or distance. This approach begins with *n* groups, each of which contains one case. Two of the cases are then combined to obtain a single cluster. At the next stage, either a third case is added to the cluster or two other cases are merged into a new cluster. Ward's method unites the two groups in order to minimise the error sum of squares.

It is also necessary to test whether the data observed at different sites in a homogeneous region arise from a common regional distribution. If the test fails, the association with the region is reconsidered and the procedure is repeated until the region can be considered homogeneous. The two homogeneity tests used were developed by Hosking and Wallis (1997) and estimate the degree of heterogeneity of a group of sites in order to evaluate whether they can be considered homogeneous. The tests are based on the L-moment ratios (L_{CV}, L-skewness and L-kurtosis) defined by Hosking and Wallis (1997).

The first heterogeneity measure is calculated as:

$$H_1 = \frac{(V - \mu_V)}{\sigma_V} \tag{3.1}$$

The H_1 measure is based on the sample variance of L-moment ratio L_{CV}, which Hosking and Wallis (1997) define as the most significant parameter to individuate homogeneous regions and here is identified as V.

The parameter V in the H_1 formulation can be calculated as:

$$V = \frac{\sum n_i \left(L_{cv}^i - \overline{L}_{cv}\right)^2}{\sum n_i} \tag{3.2}$$

where n_i is the number of observations in station i. L_{cv}^i and $\overline{L}_{cv}$ are the L_{CV} of station i and the mean regional L_{CV}.

The mean μ_V and standard deviation σ_V of the chosen dispersion measure are estimated using this procedure: the mean regional L-moment ratios are used to evaluate the parameters of a kappa distribution. This allows for the calculation of the repeated simulation of a homogeneous region in which the recorded lengths of its sites are the same as those of the observed data. In this case 500 homogeneous regions were generated. Mean μ_V and standard deviation σ_V are then obtained from these simulations.

The region can be assumed to be homogeneous if the H_1 is sufficiently small. Hosking and Wallis (1997) suggest that the region may be assumed to be "acceptably homogeneous" if $H_1 < 1$, "possibly homogenous" if $1 < H_1 < 2$ and "definitively heterogeneous" if $H_1 > 2$.

The H_1 only measures heterogeneity in the dispersion of the samples, since it is based solely on the differences between the sample L_{CV} in the region. Hosking and Wallis (1988) also give an alternative heterogeneity measurement, which we call H_2 and is obtained using the same procedure as that of H_1 measurement but is based on L_{CV} and *L-skewness* at the same time. H_2 has similar acceptability limits as the H_1 statistical. Hosking and Wallis (1997) judge H_2 to be inferior to H_1, stating that it rarely yields values larger than 2 even for highly heterogeneous regions.

Cluster analysis was then applied, using explanatory variables with the higher correlation coefficient. First of all the basins' area, altitude and geographical coordinates were used to cluster sites. In this way three regions were obtained, which coincided with the Apenninic, coastal and Tiber River stations as already identified from the scatter plot. The Hosking and Wallis (1997) homogeneity tests were applied for these regions to the Q_1 values, the annual mean values and Q_{95} values. The results of these tests are depicted in Table 3.5, where it can be seen that the more constraining H_1 statistical values are often bigger than the threshold value that identifies a heterogeneous region. A very high heterogeneity was detected in the Tiber River region in particular. To solve this problem, the percentage of substrate (volcanic or carbonatic) was added to the other variables used to cluster sites and a different configuration of the regions was hypothesized. After this procedure the Tiber River basins were divided into two regions: the left bank and the right bank of the river. The other two regions coincide with those initially identified.

Figure 3.5 shows the four regions identified with the cluster analysis, while Table 3.6 shows the results of the two tests for the Tiber River basins after the division. The values for the H_1 statistical for the two Tiber regions are lower than in the first configuration; the H_2 statistical is less than 2 in all cases. The obtained results are influenced by intersite correlation due to the nested structure of the region, although Hosking and Wallis (1988) showed that regional heterogeneity affects the accuracy of regional flood frequency quantiles more significantly than does intersite correlation. Hence, because of the great heterogeneity of the region, the tests seem to have been passed.

Table 3.5 - Results of the homogeneity tests by Hosking and Wallis (1997) for annual maxima Q_1, mean values Q_{50} and annual minima Q_{95} for the three regions initially identified.

	H_1	H_2
Tevere Q_1	4.008	0.938
Tevere Q_{50}	12.538	1.103
Tevere Q_{95}	1.798	1.815
Coastal Q_1	0.043	0.536
Coastal Q_{50}	3.127	0.407
Coastal Q_{95}	2.060	0.649
Appenninic Q_1	0.325	0.236
Appenninic Q_{50}	3.003	0.395
Appenninic Q_{95}	3.313	1.178

Table 3.6 - Results of the Hosking and Wallis (1997) Homogeneity Tests for Annual Maxima Q_1, Mean Values Q_{50} and Annual Minima Q_{95} for the Two Tiber River Regions.

	H_1	H_2
Tiber Carbonatic Q_1	1.246	0.582
Tiber Carbonatic Q_{50}	2.632	1.239
Tiber Carbonatic Q_{95}	1.842	0.812
Tiber Volcanic Q_1	0.124	0.423
Tiber Volcanic Q_{50}	0.764	0.512
Tiber Volcanic Q_{95}	1.314	0.981

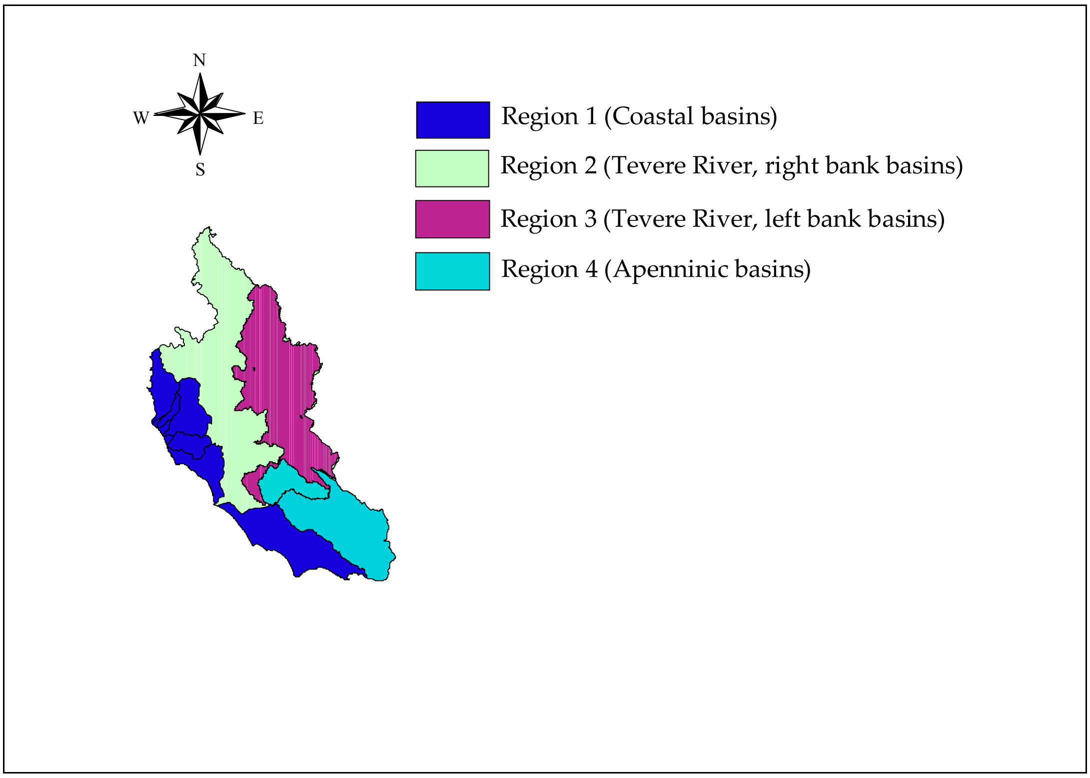

Fig 3.5 - Homogeneous regions recognized in the studied area. Four regions are detected according with the geomorphologicalal conditions of the area.

3.4. Aposelemis basin

Here are here presented also some characteristics about Aposelemis river basin in the Crete Island. The study area includes two watersheds, located in the north eastern part of Crete Island in Greece. The basins are the Lassithy Plateau, and the Aposelemis torrent basin. The first one is a closed mountainous basin. It has a average altitude of 1124 m and a total area of 128 km^2. The surface water is discharged through a sinkhole system created in limestones at the southwest end of the watershed, where the stream flow at a runoff gauging station (Tsakiris et al. 2006).

Conversely Aposelemis basin has a total area of 122 km^2 and an average altitude of 464 m. The outlet of this basin is positioned in the Cretan sea. The main stream of Aposelemis begins from Kastamonitsa village, in the same area of a group of karstik springs. It is known a connection between these springs and the Plateau sinkholes, because of the increase of the discharge from the springs after the increase of the water entered in the sinkholes.

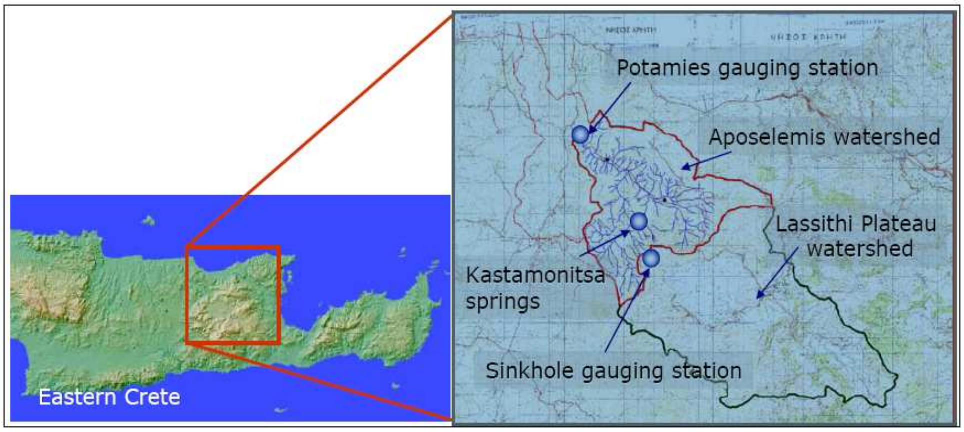

Fig 3.6 Aposelemis river basin (Tsakiris et al. 2006).

4

Comparison between flow duration curves regionalization methods

4.1 Methods of low flow estimation in ungauged sites

Low flow measures and calculation methods seen in the Chapter 2 can be easily calculated when flow time series are available and an adequate information can be provided.

Data are in fact necessary for different engineering problems as physical and biochemical process studies, hydrological modelling, forecasting, water resources planning and operation.

However many management and engineering decisions are required in catchments where measured data are not available. Hydrological networks have declined in many regions of the world and the system for collecting water information is not always existent or inadequate. This situation surely need to be solved, and many regionalization methods exist to do this. Different possible approaches to classify these methods exist and surely all arbitrary (WMO, (2008), Smakhtin (2001), Castellarin et al. (2004), furthermore because sometimes the different approaches are the sum of different techniques.

In this thesis we will consider a classification approach, that mainly considers regionalization of flow duration curves methods. A rough classification in two categories can be based on this:

- Parametric methods;
- Graphical methods. This method comprises techniques that use standardized graphical representation of FDC's with a regional validity.

The first category of techniques uses a parametric representation of flow duration curves; in these methods the calculation in ungauged sites is obtained through different methods (multiple regression models, kriging, etc.). This approach can be further divided in two categories, the statistical and analytical methods. The first one is based on the assumption that the flow duration curve is the complement of the cumulative distribution function, while the second one uses an analytical expression fitted to the empirical values, but does not use any hypothesis built on the probability theory.

The second category comprises techniques that use standardized graphical representation of FDC's with a regional validity.

In this chapter following Castellarin et al. (2004) three different regionalization methods are applied in the study area, due to evaluate the best performing regional model in this region. Moreover it is employed a jack knife technique to evaluate the uncertainty of the regionalization methods and to evaluate the reliability of these methods in ungauged basins.

4.2 Parametric methods

4.2.1. Statistical methods

Flow duration curves are often considered as deterministic objects, that can reproduce the regime feature of a river in a given section. Conversely the statistical models have a probabilistic approach to FDC and are characterized by the view of these curves as the complement of the cumulative frequency distribution. Although Ganora et al. (2009) showed that theoretically FDC could not be considered as a probability curve since discharge is correlated between successive time intervals and discharge characteristics are dependent on the season, these methods have known a big distribution.

The common steps for the application of this approach are these:

- A suitable frequency distribution is chosen as the parent distribution for the data;
- The parameters of the region are estimated through the observed data at the gauging stations;
- The evaluation of FDC obtained in ungauged basins is done through different methods. In particular the multiple regression model is widely used to connect region characteristics to curve parameters, in order to estimate their values in ungauged basins.

Different papers are involved in this topic. Usually statistical procedures to regionalize FDC use stochastic models to represent them and besides many frequency distributions are fitted to the FDC. The most used parent distribution adopted to fit the data is the log-normal frequency distribution. The first author that proposed it for the low part of the curve (the interval $0.50 \le p \le 0.99$) was Beard (1943), followed by Fennessey and Vogel (1990). The form of the excedeence probability is:

$$P(Z > z_p) = 1 - (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{z_p} \exp\left(-\frac{1}{2}x^2\right) dx \tag{4.1}$$

z_p in this expression is the p-th percentile of the normally distributed random variable Z characterized by zero mean and unit variance:

$$z_p = \lfloor \ln(t_p) - \mu \rfloor / \sigma \quad (4.2)$$

t_p is the p-th percentile of the lognormally distributed random variable with parameters μ and σ. The values of the parameters of the log-normal distribution for the different gauging stations was estimated minimizing the sum of the squared residuals between the empirical curve and the stochastic model in the range $0.50 \leq p \leq 0.99$. Then they used a multiple variate analysis to estimate the two parameters in ungauged basins.

Conversely Le Boutillier and Waylen (1993) studied a method to regionalize the annual based AFDC. While Searcy (1959) has described it as less accurate than the complete record curve, Le Boutillier and Waylen (1993) has recognized it as the only FDC which represents the interannual behaviour. They suggested a mixed five parameters log-normal distribution as the parent distribution in order to describe many streamflow regimes and the effect of different climatic and geological factors as seasonality, basin area and storage. Cluster analysis was then applied to identify groups of similar basins.

Claps and Fiorentino (1997) used a three parameter log-normal distribution to represent one year of data of daily discharges. The adopted relation is:

$$\log(q_A - q_0) = \alpha + \beta z \quad (4.3)$$

q_A in this relation represents the dimensionless daily discharge $q_A = Q/E[Q]$, where $E[Q]$ is the average daily discharge in the year considered. q_0 is the discharge lower bound and represents an adjustment parameter, while z represents the normal reduced variate and corresponds to the frequency of the discharge q. The authors show that is not frequent to need the q_0 lower bound in Mediterranean climates, then the model adopted need a calibration of only two parameters. They show that the distribution of parameters obtained for different years for a station is normal and then propose regional models to estimate the parameters of the normal distributions for ungauged sites.

Other authors have used different distributions to represent FDC. In particular Singh et al. (2001) proposed a normal frequency distribution to represent the 10-day standardized and normalized streamflow series observed in the Himalayan area. The standardization and normalization of the data is performed by dividing the observations by the mean annual flow, and through a power transformation of the data. The regionalization is obtained relating the model parameters to the basin area.

Iacobellis (2008) used a beta distribution to fit the data.

Castellarin et al. (2007) regionalized a stochastic index flow model of flow duration curves. This model has the value to integrate the period of record flow duration curves with the annual based flow duration curves. They chose a kappa distribution to represent the frequency regime of daily streamflows in their study area and used a physically sensible multiregression relationship to estimate parameters from geomorphoclimatic characteristics.

4.2.2. Analytical approaches

Many papers proposed to adapt an analytical function to empirical flow duration curves in order to regionalize it. Singh (1971) suggested the use of a model according to which the standardized discharge, that correspond to a certain percentage of time D is a simple power function of the drainage area, while the constant and the exponent of the relation are regionally varying functions of D. Quimpo et al. (1983) have fitted an exponential and power relation on daily flow duration curves of Philippines. The expression of the two parameters relationships is:

$$q(D) = Q_A \cdot \exp(-cD) \tag{4.4}$$

$$q(D) = Q_A \cdot D^{-c} \tag{4.5}$$

In these relations *q(D)* is the discharge per unit area of the basin associated with the duration D, Q_A and c are positive constants.

Quimpo et al. (1983) identified a relationship between the basin area and the parameter Q_A in the way to estimate the FDC in ungauged basins. Besides they provided a contour map of *c* for the Philippines area.

Mimikou and Kaemaki (1985) tested different analytical equations (polynomial and logarithmic) to represent the monthly FDC's observed in northern Greece. At the end they proposed to use a cubic polynomial model to represent monthly FDC. Moreover in order to regionalize the model they proposed four regression equations that relate model parameters to climatic and geomorphological characteristics.

Franchini and Suppo (1996) suggested a parametric model to represent the lower part of daily flow duration curves ($D \geq 0.30$):

$$Q(D) = c + a \cdot (1 - D)^b \tag{4.6}$$

Where a, b and c are the positive model parameters. Parameters are estimated forcing the model to pass through three points (D_i ,Q (D_i)) with i=1,2,3. The estimation of parameters of $Q(D_i)$ equations in ungauged basins is done through regression models.

A different procedure is proposed by Yu et al. (2002). In fact they suggested to estimate daily discharges Q_p (corresponding to different percentiles p=10, 20,…) and not the curve parameters in ungauged basins through geomorphological indexes. These values of Q_p can then be used to construct flow duration curves in ungauged sites. Besides they evaluate the reliability of the proposed approach and estimated confidence intervals for the FDC through a bootstrap cross-validation.

4.2.3. Graphical approaches

The graphical approach is based on these steps:

- First of all it is necessary standardize the flow duration curve for each basin for all gauged river basins, through division for a mean or median discharge;
- After that through average of the standardized FDC in different sites the graphical regional flow duration curve is estimated;
- Finally the flow duration curve for a particular basin can be estimated multiplying the mean or median discharge (flow index) estimated in the gauged stations to the dimensionless FDC.
- The estimation on ungauged basins can be done relating the flood index to geomorphoclimatic characteristic of the area through regression models.

The approach in this form was defined by Smakhtin et al. (1997), but different papers are involved in the definition of graphical methods to estimate regional flow duration curves. In particular it is important to cite Furnes (1959), Gustard et al. (1992), Studley (2001).

Besides one of the last published approaches for regionalize FDC that use a graphical methods is that of Ganora et al. (2009). They model the FDC as non parametric object and to regionalize it they consider the dissimilarity between curves. A linear norm is in fact used to obtain a dissimilarity matrix between the curves. This matrix is related to an other one composed by the differences between all possible values of each descriptors within the set of basins. A cluster analysis is then used to identify homogeneous regions. Every region is characterized by a single dimensionless flow duration curve.

4.3 Validation analysis and indexes of reliability

A validation analysis is then necessary to evaluate accuracy of the regional estimates. This procedure allows one to simulate the presence of ungauged basins and is assessed on the following procedure (Castellarin et al. 2004):

- Consider to have a *N* gauging stations belonging to the study region;
- One of the stations of the homogeneous region is removed from the sample, in the way to have *N-1* components of the sample;
- A regression analysis is carried out with the *N-1* sample and a new equation between streamflow data and geomorphologicalal and climatic characteristics is defined;

- New parameters with the new equation are then calculated for the removed station and are used to calculate the flow duration curve;
- The procedure must be applied for all the stations.

The new N flow duration curves (that are here called jack knife FDC) are then compared with the empirical ones. This kind of procedure permits to evaluate the capacity of the models to work in ungauged basins condition and then to assess their robustness and reliability.

The flow duration curves calculated using parameters obtained through jack knife validation are then compared with empirical flow duration curves in order to assess the accuracy of the model.

To evaluate the performance of the model, the following indicator was considered (Castellarin et al., 2004):

$$\varepsilon_{s,j} = \frac{\hat{Q}_{s,j} - Q_{s,j}}{Q_{s,j}} \cdot 100 \tag{4.7}$$

where $Q_{s,j}$ and $\hat{Q}_{s,j}$ indicate the daily streamflow, empirical and estimated through regionalisation, associated with duration j for station s. From these values it is possible to obtain the mean relative error $\bar{\varepsilon}_s$ and its standard deviation $\sigma_{\varepsilon,s}$ for a station as:

$$\bar{\varepsilon}_s = \frac{1}{N'} \cdot \sum_{j=1}^{N'} \varepsilon_{s,j} \tag{4.8}$$

$$\sigma_{\varepsilon,s} = \sqrt{\frac{1}{N'} \cdot \sum_{j=1}^{N'} \left(\varepsilon_{s,j} - \bar{\varepsilon}_s\right)^2} \tag{4.9}$$

where *N'* represents the number of durations that are considered to calculate mean and standard deviation.

In addition, the average of *N* values of Equation 4.8, $\bar{\varepsilon}$, and of Equation 4.9, σ_{ε}, with *N* corresponding to the number of sites, gives an indication of the performance of the model.

It is also possible to graphically represent the mean and median of the distribution of the *N* relative errors $\varepsilon_{i,j}$ and the 100(*p*/2) % and 100[1-(*p*/2)] % percentiles, by identifying the interval about the median containing the 100(1-*p*) % of the *N* relative errors, against durations *j*, to evaluate the uncertainty of the regional FDC for all durations.

The mean relative error for a given duration *j* can be calculated as:

$$\bar{\varepsilon}_j = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_{j,i} \tag{4.10}$$

Another performance index E_s that can be used is the Nash-Sutcliffe efficiency method, calculated for each station as:

$$E_s = 1 - \frac{\sum_{j=1}^{N'} (\hat{Q}_{s,j} - Q_{s,j})^2}{\sum_{j=1}^{N'} (Q_{s,j} - \overline{Q}_s)^2} \tag{4.11}$$

where $\hat{Q}_{s,j}$ is the estimated value for each duration j and site s, $Q_{s,j}$ is the empirical value for each duration j and site s and $\overline{Q}_s$ is the mean value. The value of this index can range between 1 and $-\infty$.

The value of this index can range between 1 and $-\infty$.

The values of E_s are used to calculate three indexes of the effectiveness of the model:

- P_1 defines the percentage of cases over N stations in which $E_s > 0.95$;
- P_2 defines the percentage of cases over N stations in which $0.50 < E_s < 0.95$;
- P_3 defines the percentage of cases over N stations in which $E_s < 0.5$.

4.4 Implementation of the regional models

This study applies three different categories of regional models, to evaluate their reliability in the study area. Only one of the different proposed procedures was chosen for each category, and was applied in the Lazio study area.

It was chosen to apply both FDC and AFDC regionalisation methods. The FDC regionalization was adopted for the analytical and graphical method, while the AFDC regionalization was chosen for the statistical approach. The main characteristics of the different methods were maintained, but they were adapted to the case study.

Each method used a different composition of the homogeneous regions. In fact the best one was evaluated for each method to have the less relative errors and biggest efficiencies. Fig. 4.1 shows the study area and the used sites.

The first step of the work has regarded the application of the regionalization method for each approach and then it is assessed the evaluation of their reliability. Next paragraph will explain the details of the applied methods.

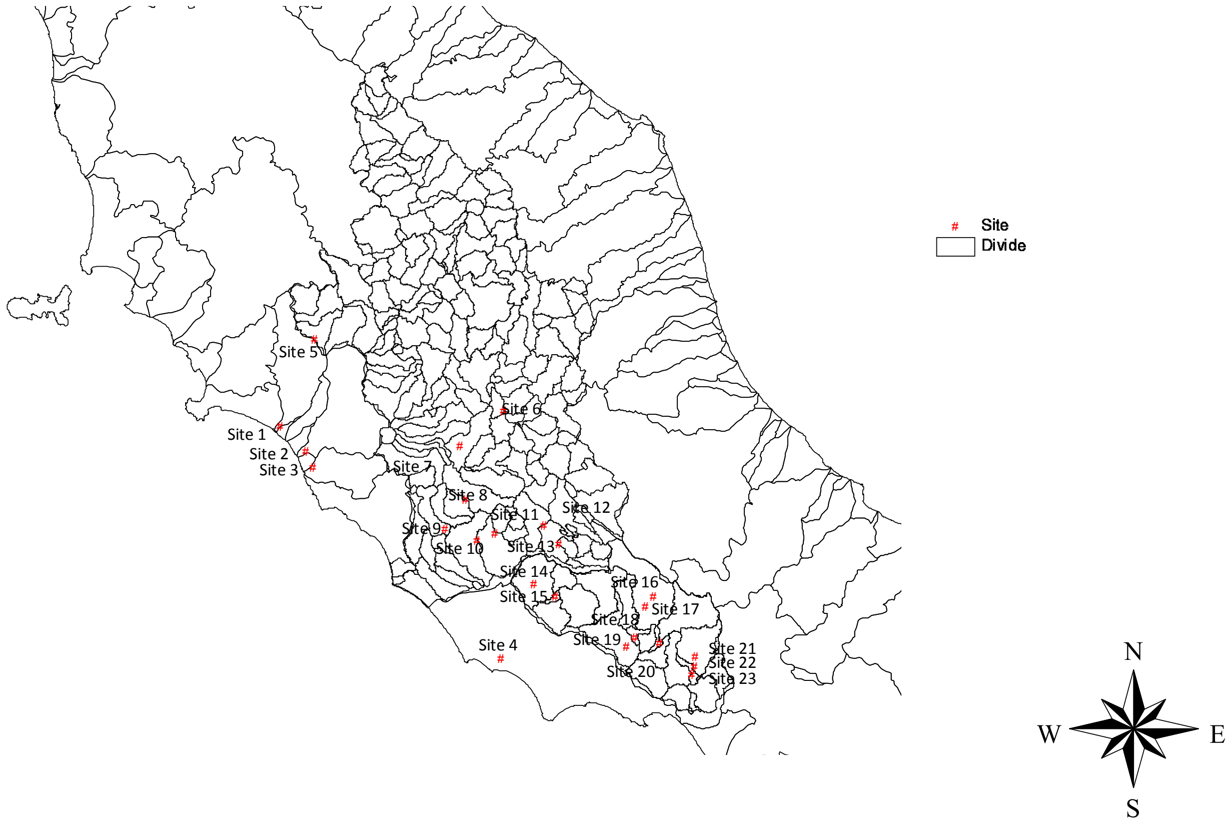

Fig.4.1 - Study area and gauge sites.

4.4.1. Statistical approach

The method developed by Leboutillier and Wailen (1993) was chosen as starting point to apply a statistical regionalisation approach.

The first step was to calculate the empirical annual flow duration curve for each station as defined by Fennessey and Vogel (1994). Following their procedure:

If we suppose to have *n* years of observations, for each year a FDC can be calculated using a non parametric approach (Vogel and Fennessey, 1994) as seen in 2.8.5 paragraph.

Each curve will differ from the other and from the period of record FDC, because of the expected interannual hydrological variability. In particular some sections of the curves will be above and some below the period of record FDC. To summarize this interannual variability can be calculated a measure of central tendency, as the mean or the median of the annual FDC for each exceedance probability *p* or Duration D_i.

For each station of the study area a frequency distribution was fitted and the same distribution was chosen for all stations in the homogeneous region. The best parent distribution was selected through the Kolmogorov-Smirnov fit test, on the basis of the distance value, always less than 0.10. The GEV max distribution with parameters evaluated with the L-moment method was identified as the best distribution for all stations. In this analysis all stations was joined together to obtain one homogeneous region, because the best results in the regionalization phase was found in this way.

Figure 4.2 shows an example of the fitting of the best distribution for a station in the region, while Table 4.1 displays the parameters of the distribution for the different stations.

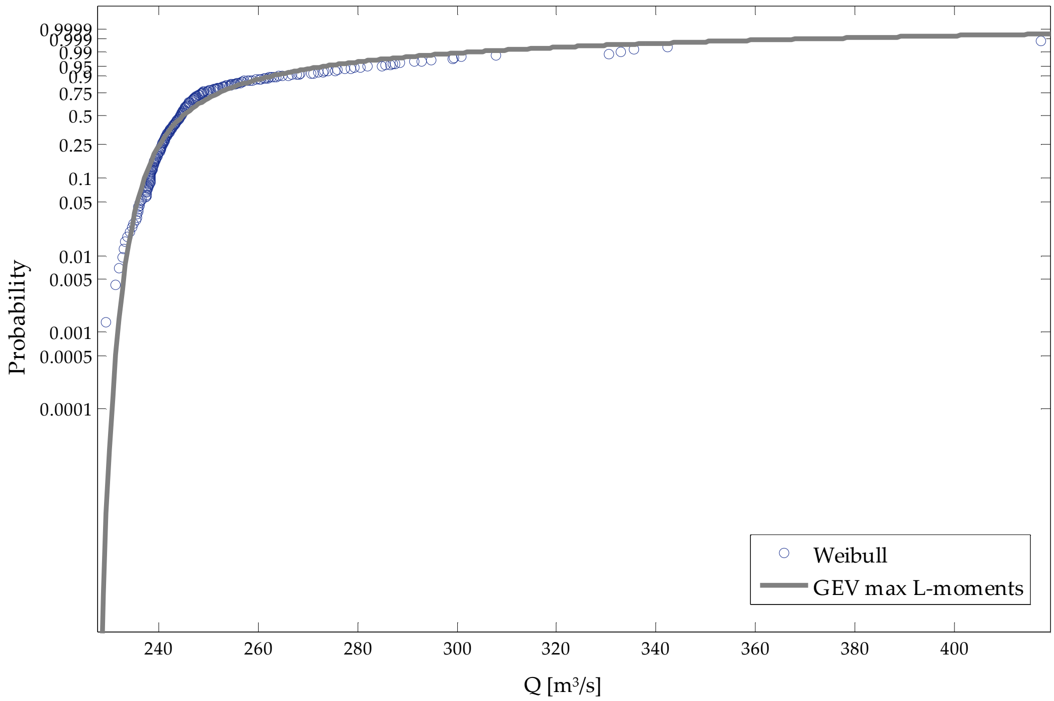

Figure 4.2 – Fitting of the Gev max distribution with parameters obtained with L-moments method

Table 4.1 – Parameters for the study area for the GEV max distribution.

	κ	λ	ξ
Site 18	0.74	0.71	1.69
Site 14	0.63	0.01	1.36
Site 19	0.38	23.41	0.6
Site 17	-0.12	0.81	16.28
Site 11	0.36	2.91	5.77
Site 3	0.7	0.46	1.06
Site 1	0.34	3.21	1.03
Site 8	0.13	40.84	7.21
Site 12	0.3	6.75	3.09
Site 10	0.4	0.1	4.33
Site 20	0.58	0.36	0.49
Site 22	0.42	10.57	4.95
Site 21	0.23	3.63	9.01
Site 23	0.47	23.7	1.89
Site 16	0.53	1.92	0.72
Site 7	0.43	5.65	242.8
Site 13	0.3	1.45	3.33

Site 2	0.3	1.76	1.99
Site 6	0.24	5.89	5.17
Site 9	0.36	16.83	17.15
Site 15	0.46	0.29	0.77
Site 4	0.61	0.48	1.18

The regionalization of the model is obtained through multiple regression analysis, using a stepwise procedure. For this type of statistical regression model, the order of entry of the predictor variables is based on statistical criteria. Variables that correlate more strongly with the dependent variable will be given entry priority.

The stepwise procedure is used to define regionalisation models to calculate the parameters in ungauged basins. Three kind of models (linear, exponential and logarithmic) are evaluated:

$$\hat{\vartheta} = A_0 + A_1\omega_1 + A_2\omega_2 + A_n\omega_n + \vartheta; \quad (4.12)$$

$$\hat{\vartheta} = A_0 \cdot \omega_1^{A_1} \cdot \omega_2^{A_2} \cdot \omega_n^{A_n} + \vartheta'; \quad (4.13)$$

$$\hat{\vartheta} = A_0 + \ln(A_1 \cdot \omega_1) + \ln(A_2 \cdot \omega_2) + \ln(A_n\omega_n) + \vartheta''. \quad (4.14)$$

In these equations $\hat{\vartheta}$ is the perfect estimated parameter value, A_i, i=1,2,..n are the coefficients of the model, ω_i are the explanatory variables and ϑ is the residual of the models.

Scale parameter λ and shape parameter κ and location parameter ξ are estimated by equations obtained by multiple regression. Two different kind of models linear and exponential was evaluated as the best for relating model parameters and geomorphoclimatic characteristics:

$$\xi = a + b \cdot A + c \cdot MAP + d \cdot F_d + e \cdot G_a + f \cdot G_b + t \cdot H_{\min} + \theta'; \quad (4.15)$$

$$k = m + n \cdot H_{\min} + o \cdot F_a + p \cdot A + r \cdot F_d + \theta''; \quad (4.16)$$

$$\lambda = g \cdot (A^h) \cdot (F_d^{\,i}) \cdot (H_{\min}^{\,l}) + \theta'; \quad (4.17)$$

Symbology for the different geomorphoclimatic characteristics are synthesized in the Table 3.1.

Table 4.2 shows the coefficients values and the R^2 for each relationship.

Table 4.2 –Values of the coefficients for the regression models and R^2 value.

ξ	Coefficients
a	32.63
b	0.01
c	0.01
d	-0.1
e	-0.31
f	-0.29
t	0.01
R^2	0.41

κ	Coefficients
m	0.41
n	0.00
o	-0.01
p	0.02
r	0.00
R^2	0.32

λ	Coefficients
g	10.25
k	0.22
i	-0.82
l	0.27
R^2	0.84

Moreover the procedure of jack knife validation was applied in order to assess the reliability of the model. Figure 4.3, 4.4, 4.5 shows the scatter plots of the parameters obtained through general regression against those obtained through jack knife validation.

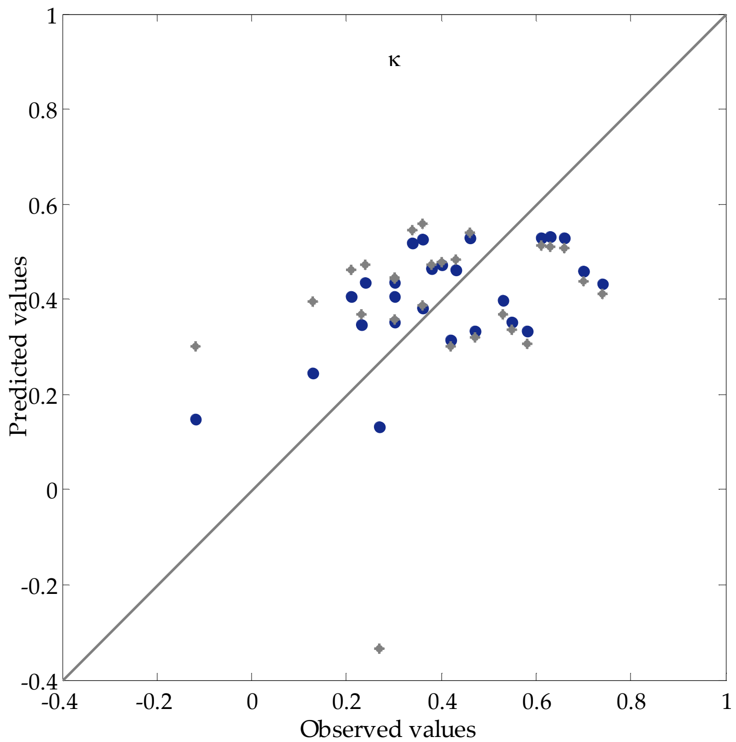

Figure 4.3 – Scatter plot of the *k* estimated parameters versus those obtained through regression analysis and jack knife validation.

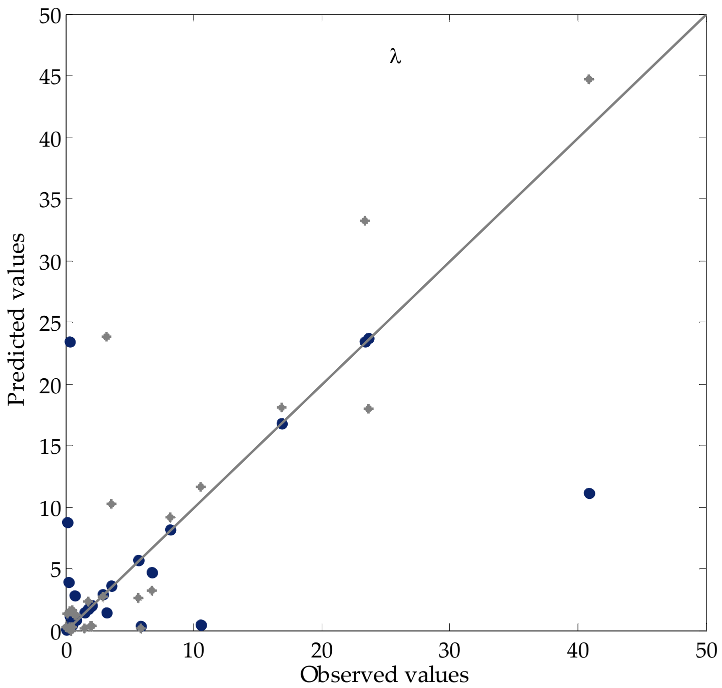

Figure 4.4 – Scatter plot of the λ estimated parameters versus those obtained through regression analysis and jack knife validation.

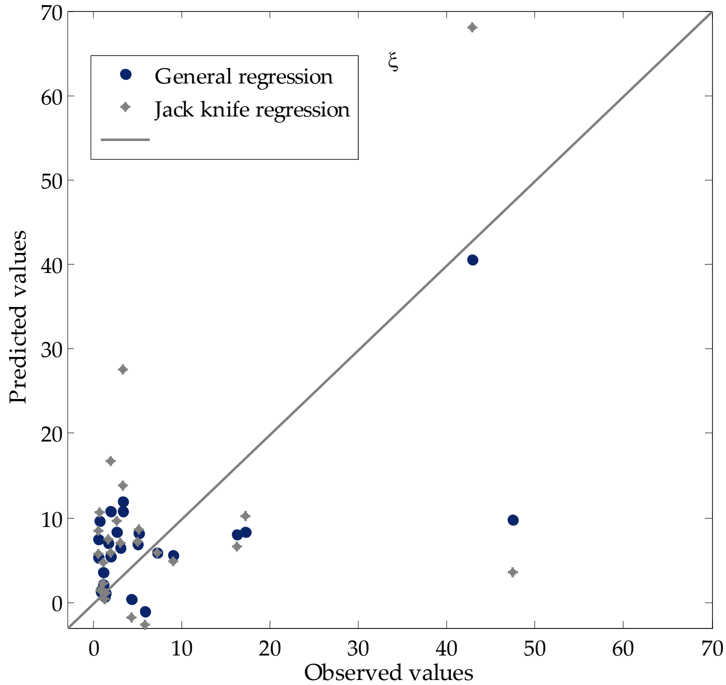

Figure 4.5 – Scatter plot of the ξ estimated parameters versus those obtained through regression analysis and jack knife validation.

The parameters estimated through jack knife validation was applied to calculate new flow duration curves. Finally the indexes of efficiency and reliability was calculated to evaluate the robustness of the results and to compare the different regionalisation methods.

Figure 4.6 shows the indexes of effectiveness for the statistical approach, that shows that there is a low percentage of P_1 results, and similar results for P_2 and P_3 results.

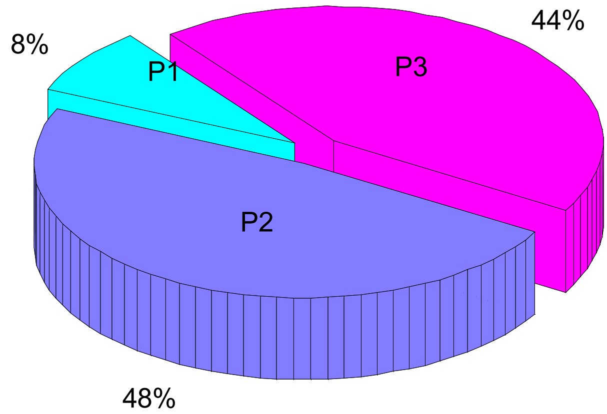

Figure 4.6 – Representation through pie graph of the indexes of effectiveness for the statistical approach.

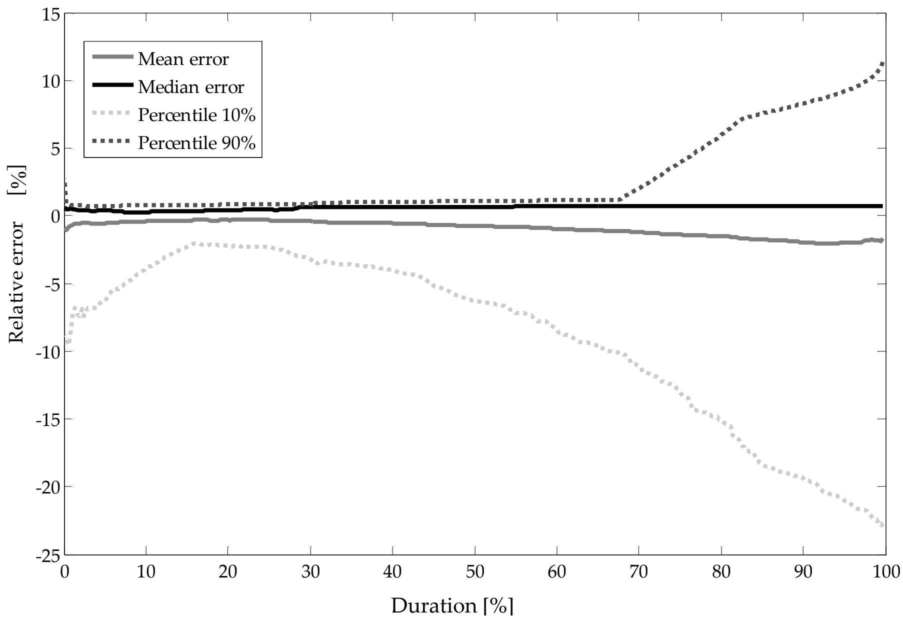

Figure 4. 7 –Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the statistical approach.

Besides Figure 4.7 represents the relative error graph for different durations and shows that the error increases for highest durations. Conversely the mean and median relative error are almost constant and stay inside a range of – 3% - 1%.

Furthermore Table 4.3 shows results for the indexes $\bar{\varepsilon}$, and σ_{ε}.

Table 4. 3 – Results for the average $\bar{\varepsilon}$, and σ_{ε}.

$\bar{\varepsilon}$	0.94 %
σ_{ε},	4.8

4.4.2 Parametric method

In this case the method proposed by Mimikou and Kaemaki (1985) was applied. For each station an empirical flow duration curve was calculated using the non parametric methodology (Vogel and Fennessey, 1994). So an analytical curve was fitted on the empirical curves. Two kind of analytical functions was identified for the study area on the basis of the highest coefficient of determination R^2. It was identified a rational function for the area of the Liri-Garigliano basin, while the exponential model was identified for the remaining stations.

The expression of the two models is:

$$Q = \frac{(a + bD)}{(1 + cD + dD^2)} \tag{4.18}$$

$$Q = f - g \cdot \exp(-hD)^i \tag{4.19}$$

Table 4.4 and 4.5 resume parameters values for the different stations.

Table 4.4 – From up to down parameters for the study area for the rational function.

	a	b	c	d
Site 15	3.32	0.00	-0.05	0.01
Site 11	128.61	39.14	1.63	0.01
Site 3	196.53	-1.91	4.35	-0.03
Site 12	304.60	202.49	3.88	0.11
Site 10	57.20	22.82	28.79	0.39
Site 22	617.16	159.63	1.95	0.02
Site 21	315.32	412.74	9.11	0.06
Site 23	1333.85	383.26	2.21	0.12
Site 13	70.77	56.71	4.72	0.13
Site 2	263.58	613.02	22.26	2.53
Site 6	413.49	21.73	-0.14	0.05
Site 9	1454.81	991.93	2.99	0.01

Table 4.5 – From up to down parameters for the study area for the exponential model.

	f	g	h	i
Site 4	63.26	63.89	0.41	-0.76
Site 19	702.08	701.47	0.04	-1.16

Site 14	271.90	272.34	0.09	-0.82
Site 1	941.57	948.54	0.06	-0.42
Site 8	1031.16	2278.14	1.69	-0.10
Site 7	879.08	654.22	0.23	-0.57
Site 17	31.27	58.13	1.46	-0.06

Besides two examples of the fitting of these curves to empirical data is shown in Fig 4.8 and 4.9.

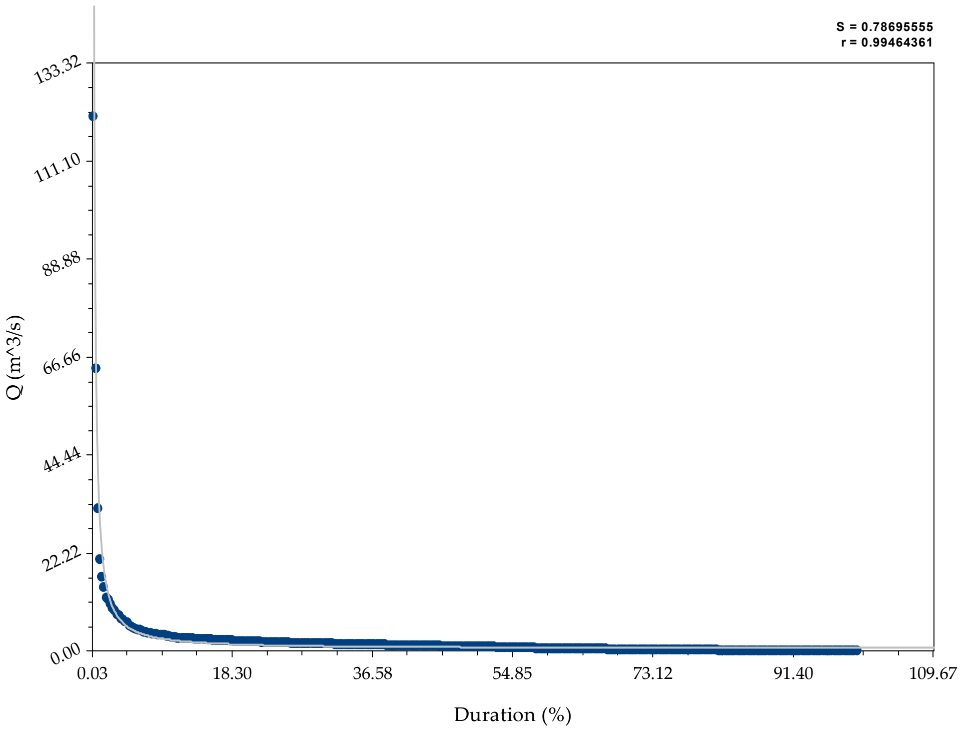

Figure 4.8 – Fitting of the exponential function to the observed data.

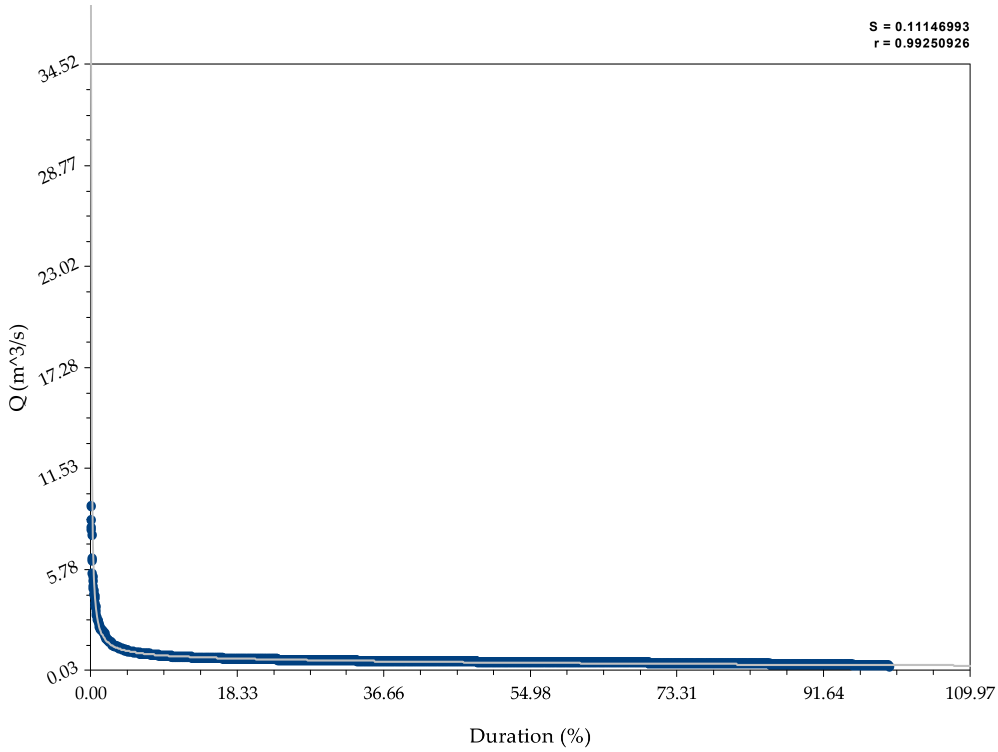

Figure 4.9 – Fitting of the rational function to the observed data.

At this point a regionalisation procedure based on multiple regression was applied as for the statistical method and equations between parameters and basin characteristics was developed. Also in this case different equations (linear, exponential and logarithmic) was tested on the empirical data and these was the chosen relationships for the rational expression:

$$a = e \cdot (A^f) \cdot (F_c^g) \cdot (F_d^h) + \theta; \tag{4.20}$$

$$b = i \cdot \left(A^l\right) \cdot (H_{\min}{}^m) \cdot (F_d)^n + \theta'; \tag{4.21}$$

$$c = o \cdot (H_{\max}{}^p) \cdot (F_a{}^q) \cdot (H_{mean}{}^r) \cdot \left(MAP^s\right) + \theta''; \tag{4.22}$$

$$d = t \cdot \left(A^u\right) \cdot \left(F_d{}^v\right) \cdot (G_a{}^w) \cdot (G_b{}^z) + \theta'''; \tag{4.23}$$

The expressions obtained for the exponential model was:

$$f = l_1 + m_1 \cdot A + n_1 \cdot H_{\min} + o_1 \cdot H_{mean} + p_1 \cdot G_a + q_1 MAP + r_1 \cdot F_a + \boldsymbol{\theta}''''; \tag{4.24}$$

$$g = l_2 + m_2 \cdot A + n_2 \cdot H_{\min} + o_2 \cdot G_a + p_2 \cdot G_b + q_2 MAP + \theta'''''; \tag{4.25}$$

$$h = l_3 + m_3 \cdot H_{mean} + n_3 \cdot F_d + o_3 \cdot A + p_3 \cdot F_a + \theta''''''; \tag{4.26}$$

$$i = l_4 + m_4 \cdot A + n_4 \cdot H_{mean} + o_4 \cdot F_a + p_4 \cdot F_d + \theta'''''''; \tag{4.27}$$

At this point the same procedure applied for the statistical method was developed here. Results are shown in the scatter plot (Figure 4.10, 4.11, 4.12, 4.13) and in the relative error graph (Figure 4.15). Moreover Figure 4.14 and reliability indexes in Table 4.7 resume results for the method. It is possible to highlight that results are very good in this case.

Table 4.6 - Values of the coefficients for the regression models and R^2 value.

a	Coefficient
e	16.64
f	0.12
g	0.63
h	0.45
R^2	0.64

b	Coefficient
i	-2.36
l	0.54
m	-1.34
n	4.01
R^2	0.98

c	Coefficient

o	35321
p	-0.83
q	1.245
r	-0.55
s	-4.21
R^2	0.86

d	Coefficient
t	0.001
u	6.42
v	9.26
w	-3.97
z	-7.32
R^2	0.99

f	Coefficient
l_1	6279.07
m_1	0.36
n_1	21.86
o_1	-4.32
p_1	312.16
q_1	-47.79
r_1	-4.41
R^2	0.95

g	Coefficient
l_2	-11027.9
m_2	0.14
n_2	-0.40
o_2	129.64
p_2	126.96
q_2	-1.28
R^2	0.7

h	Coefficient
l_3	0.62
m_3	-0.01
n_3	0.07
o_3	0.001
p_3	-0.003
R^2	0.99

i	Coefficient
l_4	-2.00
m_4	0.001

n_4	0.001
o_4	-0.006
p_4	0.02
R^2	0.99

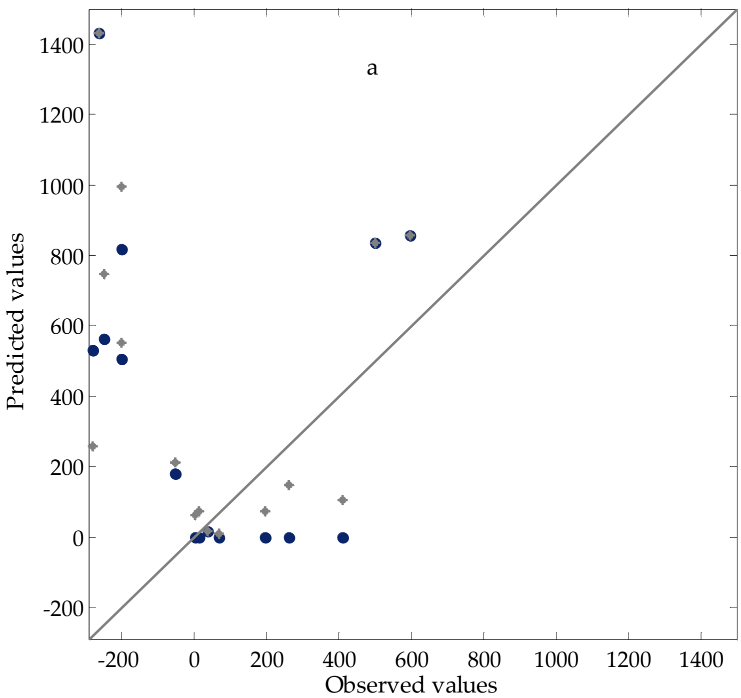

Figure 4.10 - Scatter plot of the *a* parameters estimated versus those obtained through regression analysis and jack knife validation.

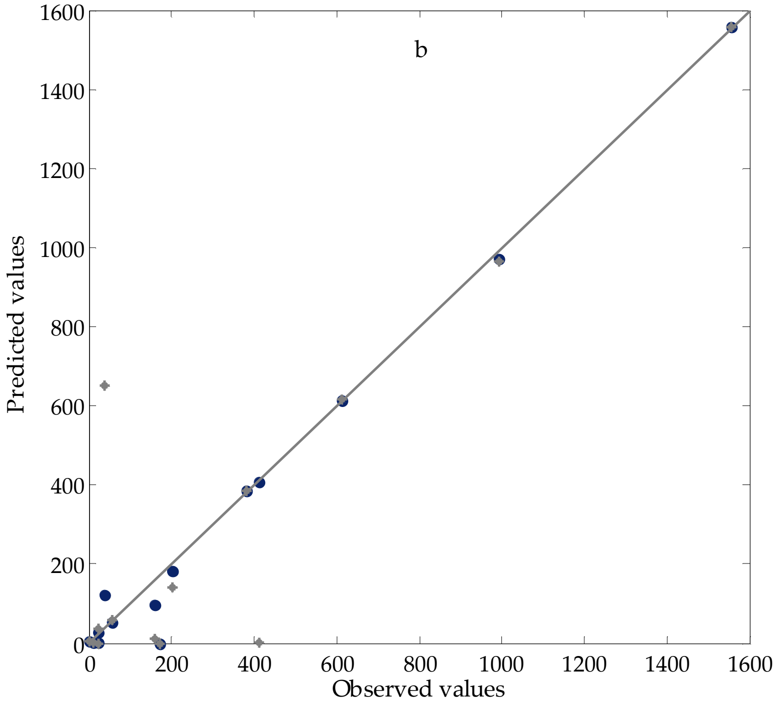

Figure 4.11 - Scatter plot of the *b* parameters estimated versus those obtained through regression analysis and jack knife validation.

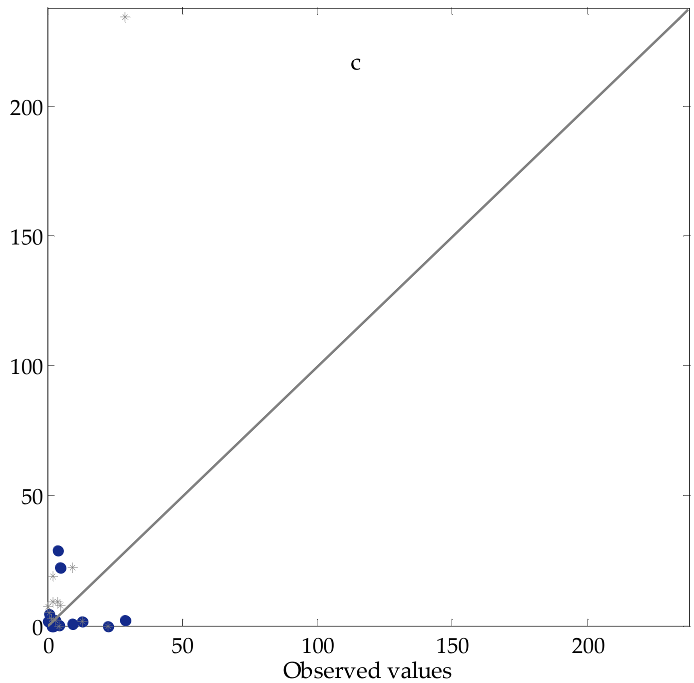

Figure 4.12 - Scatter plot of the *c* parameters estimated versus those obtained through regression analysis and jack knife validation.

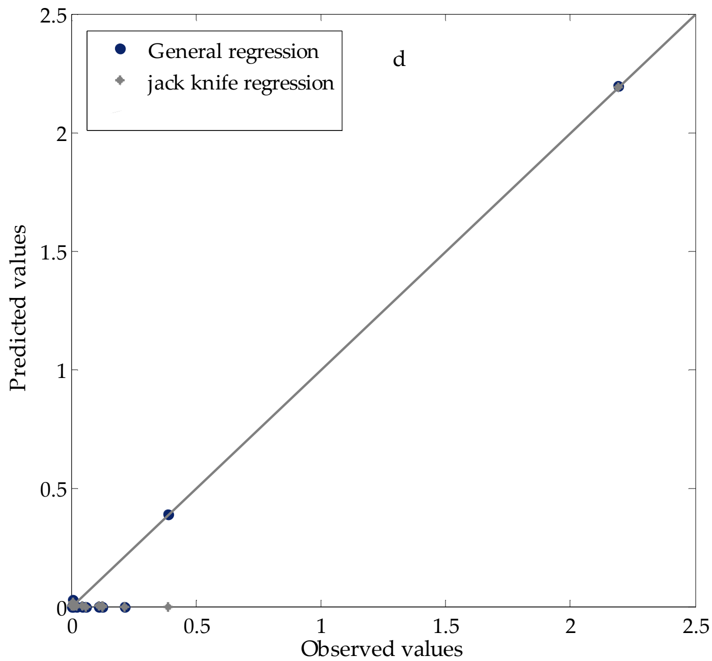

Figure 4.13 - Scatter plot of the *d* parameters estimated versus those obtained through regression analysis and jack knife validation.

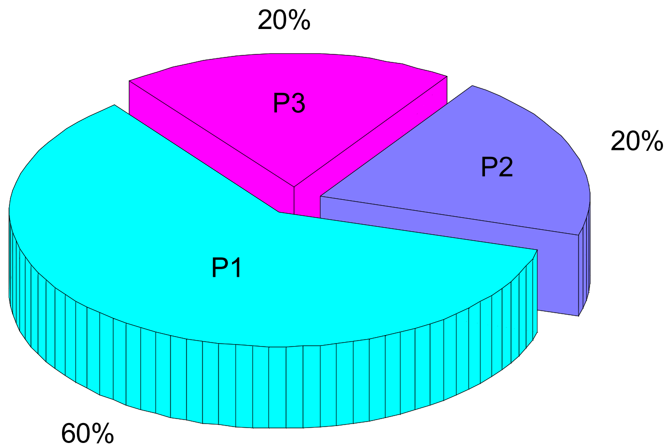

Figure 4.14 - Representation of the global results through pie graph of the indexes of effectiveness for the parametric approach.

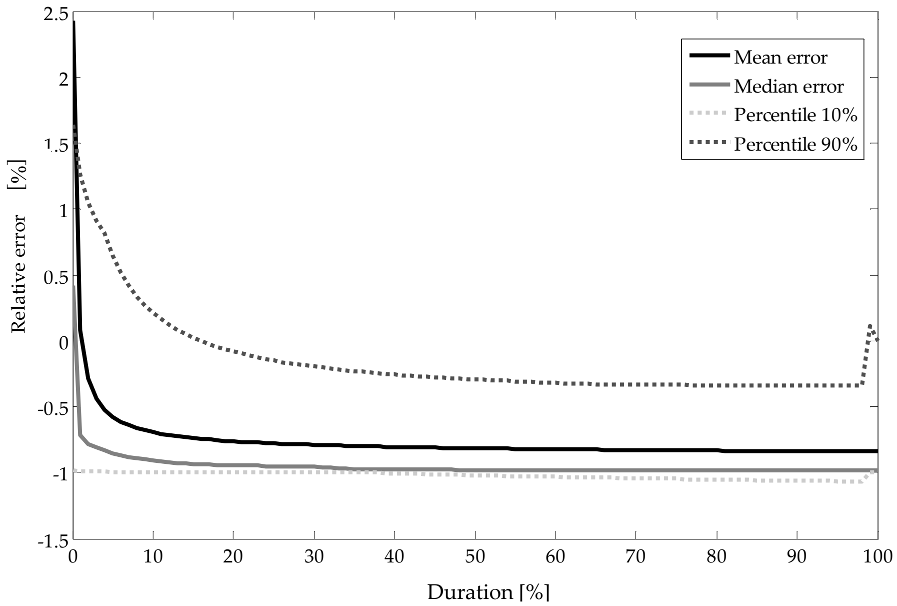

Figure 4.15 - Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the parametric approach.

Table 4.7 – Results for the average $\bar{\varepsilon}$, and σ_ε.

$\bar{\varepsilon}$	-0.7 %
σ_ε,	0.35 %

4.4.2. Graphic approach

The graphical approach proposed by Smakhtin et al. (1997) was developed. The first step was to standardize the FDC for all stations dividing for the flow index (usually the mean and median value). The curves was then separated depending on the homogeneous regions. For each set of FDC a mean curve was calculated in order to obtain the regional FDC. Finally the index values was related to the catchment characteristics using multiple regression method.

$$Q_{mean} = a + b \cdot \ln(A) + c \cdot \ln(H_{\max}) + d \cdot \ln(F_b) + e \cdot \ln(G_a) + \theta; \quad (4.28)$$

Table 4.8 - Values of the coefficients for the regression models and R^2 value.

Q_{mean}	Coefficients
a	163.834
b	11.66144
c	44.73149
d	38.70365
e	-156.793
R^2	0.658

Eq. 4.28 shows the shape of the relationship obtained for the flow index. The index value as expected is function of the basin contributing area, of the maximum altitude in the basin and of the soil characteristics of the basins. This equation confirms the results obtained in case study chapter.

The jack knife procedure was here applied in a different configuration. First of all a jack knife procedure was used to calculate a new regional curve. Then every station belonging to the sample was removed one by one and a new regional curve was estimated for each station. Then the same process was performed for the index value in order to estimate new equations by multiple regression. At this point multiplying the regional curve obtained for a certain station to the index value obtained from the jack knife equation is possible to obtain the jack knife FDC.

Table 4. 9 - Parameters for the study area for the GEV max distribution.

	q index
Site 15	0.6315
Site 4	1.83655
Site 18	4.17803
Site 14	0.03721
Site 19	46.2865
Site 17	13.3317
Site 11	20.0105
Site 3	2.56063
Site 1	7.87924
Site 8	323.883
Site 12	27.6027
Site 10	0.63779

Site 20	0.97543
Site 22	65.7298
Site 21	36.5342
Site 23	77.3456
Site 16	5.84802
Site 7	248.392
Site 13	6.25474
Site 2	6.59701
Site 6	35.6794
Site 9	309.365

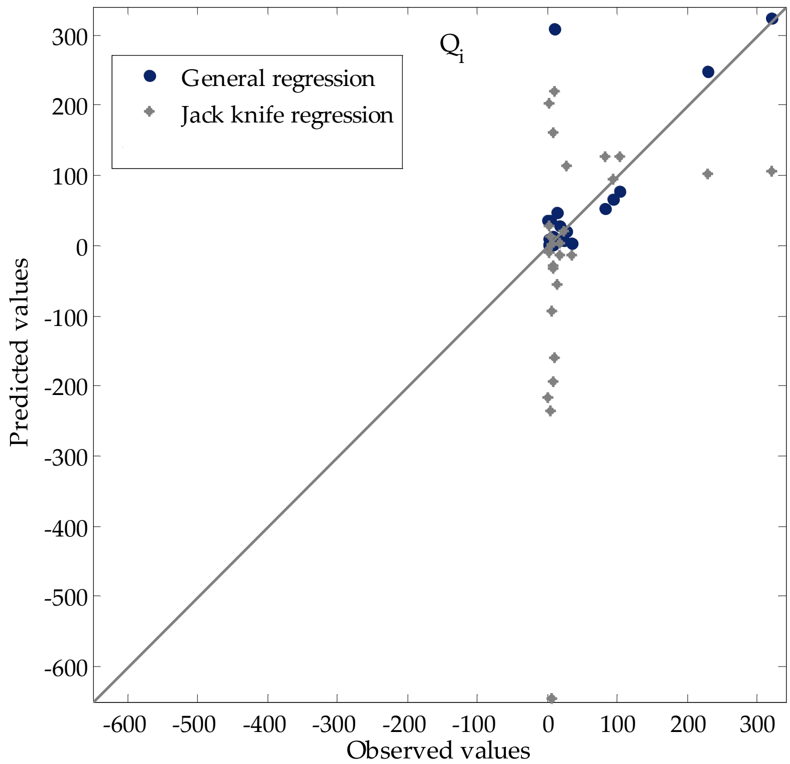

Figure 4.16 - Scatter plot of the Q_i parameters estimated versus those obtained through regression analysis and jack knife validation.

All the curves are then used to calculate relative errors (Figure 4.18) and reliability indexes shown in the Table 4.10 and in the Figure 4.17.

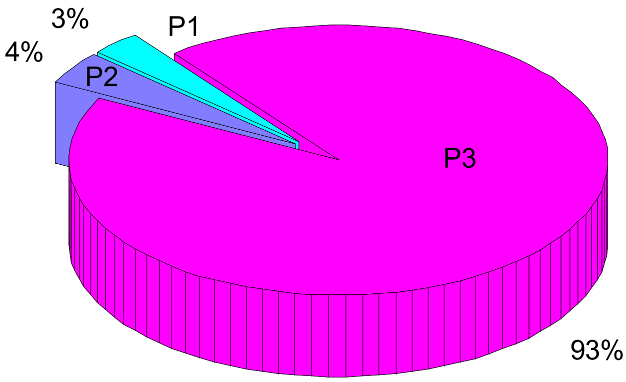

Figure 4.17 - Representation of the global results through pie graph of the indexes of effectiveness for the graphic approach.

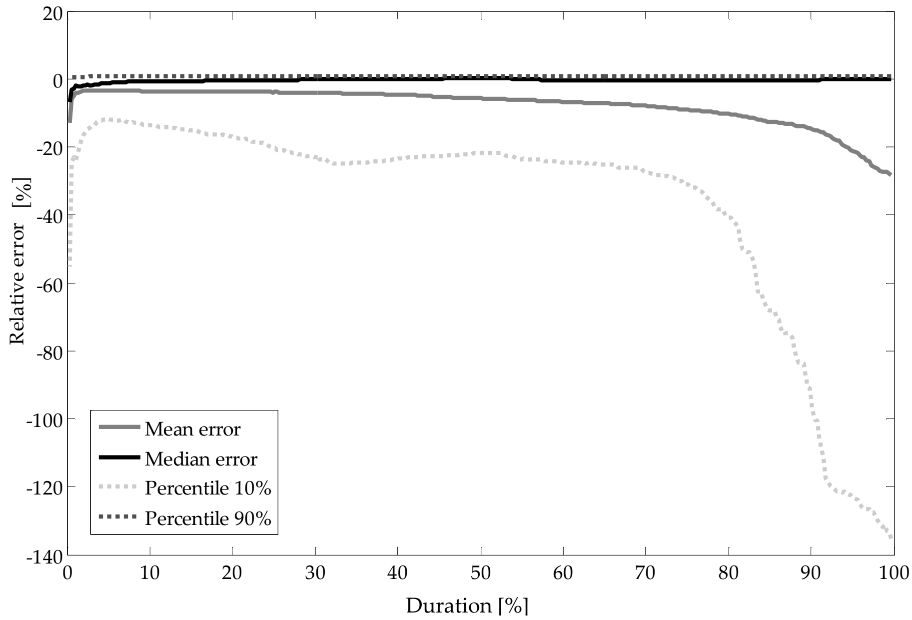

Figure 4.18 - Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the graphic approach.

Table 4.10 – Results for the average $\bar{\varepsilon}$, and σ_{ε}.

$\bar{\varepsilon}$	-7.7 %
σ_{ε},	6.6 %

4.5 Results

The analysis shows that the regionalization method that gives worst results is the graphic approach. This is probably caused by the use of an unique parameter, that makes difficult to represent the geographic variability of the curves in the study area that is particular heterogeneous (as shown in the Chapter 3). The other two methods provide similar results, also due to the comparable procedure. We want here compare these results with those obtained by Castellarin et al. (2004). They found that for ungauged sites results of the three models was comparable. This seems in contrast with our results that shows that the graphic method has the worst outcomes, with the 93 % of P_3 value and the 3 % of the P_1 against the 20 % of the P_1 value for the parametric approach and the 44 % for the statistical one. Moreover their results show that the graphic methodology has the least poor fit value P_3 in contrast with our results. These different results show also that the reliability of different methods in a study area have to be proved in the study area before the use.

5

Stochastic index methodology for the use with zero data

5.1. Introduction

An accurate representation of the river regime, as already seen, is essential in several engineering applications, such as the analysis of hydroelectric feasibility, reservoir and lake sedimentation, resources management and environmental planning. In this field a fundamental environmental problem is the discharge of wastewater on rivers. In fact, legislation regarding this matter is designed to put a stop to this process on rivers characterized by no flow for long periods.

The FDC can be used as a tool to represent the streamflow frequency regime correctly and can be applied to all of these hydrological applications. The FDC can be easily calculated as the complement of the cumulative distribution function (*cdf*) from gauged river data.

The existent methods do not allow for the calculation of flow duration curves on basins with an intermittent regime, where the above-mentioned environmental problems caused by waste water, are critical. These basins are characterized by flash floods and have no base flow, or their base flow is restricted to the wet periods of the year, or even zero flow in the other periods of the year.

5.2. Zero data in hydrological time series

Presence of zeros on discharge time series can be real or can take place when the discharge is underneath a threshold and the instruments cannot take any measurements, as in the case of censored data (Durrans et al., 1999). Different studies have focused on various methods to work with zero data on frequency analysis (e.g. Haan, 1977, Kilmartin and Peterson, 1972; Wang and Singh, 1995; Jennings and Benson, 1969), while others have concentrated more on the techniques to work with censored data (e.g., Tate and Freeman, 2000; Kroll and Stedinger, 1996).

The occurrence of zero events can be expressed through probability theory by substituting a non-zero probability mass with a zero value. This creates a discontinuity in the density function from which the hydrological series is obtained with discontinuity in the zero value. However this solution can create problems with the assumption of continuity made in frequency analysis. Jennings and

Benson (1969) highlighted the potential problems encountered when a continuous distribution is fitted with data with zero values.

The literature presents three methodologies for approaching the zero data, summarized in Haan (1977):

- The first methodology is to add a small constant to all observations, such as 1% of the mean magnitude, and fit a continuous distribution, such as the Log-Pearson type III distribution onto the data (Subcommittee on Hydrology, (1966)). This approach can move the discontinuity represented by the zero state, and does not solve the problem created by the discontinuity.
- The second method ignores the zero values and considers only the non-zero ones and then corrects the results for the entire period recorded. Haan (1977) and Wang and Singh (1995) showed, however, that this method is biased, since it ignores all zero values in the data.
- The third method is based on the theorem of total probability (Jennings and Benson, 1969).

W.G. Strupczewski et al. (2003) on the other hand, proposed a different method based on the hypothesis that the unit impulse response of a linearized kinematic diffusion (KD) model is a probability distribution suitable for frequency analysis of hydrologic samples with zero values.

Woo and Wu (1989) and Wang and Singh (1995) developed empirical three-parameter models for the frequency analysis of hydrologic data containing zero values starting from the theorem of total probability.

Occurrence of zero events can be expressed using the probability theory by placing a non zero probability mass on a zero value, i.e. $P(Y = 0) \neq 0$; where Y is the random variable, and P is the probability mass. Woo and Wu (1989) gave the definition of non exceedance probability of a hydrological random variable Y, that can take on zero values as:

$$P\{Y < y\} = P\{Y = 0\} + (1 - P\{Y = 0\}) \cdot P\{Y < y | Y > 0\} \tag{5.1}$$

Or in distribution form:

$$F(y) = P\{Y = 0\} + (1 - P\{Y = 0\}) \cdot F'(y | Y > 0) \tag{5.2}$$

Where F is the distribution function of (), while *F'(|)* represents the conditional distribution function. It is here explained how obtain these two equations.

It is proved that the distribution function of a random variable, is monotonic, whether continuous or not. The distribution function of this variable can be decomposed (Wu et al., 1979) in:

$$F(y) = F_c(Y) + P\{Y = 0\}; \qquad 0 < y < \infty \tag{5.3}$$

This follows the rules of monotonic functions in real variable function theory and physical basis. In this equation F and F_c symbolize the distribution function of the random variable *Y* and of the continuous part *Y*>0. *y* represents a particular value of *Y*.

But the density function of this random variable, calculated through derivation from the previous equation, is discontinuous with discontinuity at the zero value, and has the form:

$$f(y) = f_c(y) + \Pr\{Y = 0\}\delta(y) \qquad 0 < y < \infty \tag{5.4}$$

In this equation $f_c(y)$ is the continuous function of y>0, δ(y) is the Dirac delta function.

This equation goes against the assumption of continuity made in conventional frequency analysis and does not meet the condition to be a probability density function:

$$\int_{-\infty}^{+\infty} f_c(t)dt = 1 - P\{Y = 0\} < 1 \tag{5.5}$$

Moreover when y=0,

$$\frac{dF(y)}{dy} \to \infty \quad , \tag{5.6}$$

That confirms that *F(y)* has a discontinuity at the zero value (*y*=0).

In addition it is possible to write:

$$F_c(Y) = P\{Y < y, Y > 0\} \tag{5.7}$$

And then

$$F_c(Y) = P\{Y > 0\}P\{Y < y | Y > 0\} = (1 - P\{Y = 0\}) \cdot F'(y | Y > 0) \tag{5.8}$$

If Equation 5.8 is put on Equation 5.3 it is possible to find Equation 5.1 and 5.2.

Literature is mainly involved on definition of methods to work with zero data on frequency analysis but is lack of methods that permit to calculate FDC and AFDC in presence of zero data. Croker et al. (2003) illustrate a regional model, based on the theory of total probability, for predicting FDCs in a region containing ephemeral catchments, but no model was found that calculate with the same parameters period of record FDC and AFDC in ephemeral basins. Rianna et al. (in press) showed a method based on the stochastic index model identified by Castellarin et al. (2004) to evaluate FDC in intermittent basins and showed a technique to regionalize the model.

It is known that the period of record FDC is incomplete because with it, one can only make steady state probabilistic statements about streamflow exceedances (Castellarin et al. 2004). But in low-flow frequency analysis, is sometimes necessary to make probabilistic statements about a given calendar or water year. AFDCs have been shown to be quite useful for making probabilistic statements about wet, typical and dry years, for computing confidence intervals associated with the AFDC representing the typical hydrologic condition and for assigning return periods to individual

AFDCs (LeBoutillier and Waylen (1993) and Vogel and Fennessey (1994) Castellarin et al. (2004). It was identified only one paper by Croker et al. (2003) that illustrate a regional model, based on the theory of total probability, for predicting FDCs in a region containing ephemeral catchments, but no model was found that calculate with the same parameters period of record FDC and annual based FDC in ephemeral basins.

Then the main objective of this chapter is to construct a model to represent daily FDC and annual AFDC that works also in basins belonging to dry climates characterized by intermittent and ephemeral regimes due to the important engineering problems existing in these kind of basins.

In order to reach this objective, the model combines the stochastic index methodology (Castellarin et al., 2004) with the theory of total probability. Stochastic index methodology permits to model the relationship between the FDC and the mean and variance of the AFDC and to maintain the variability of observed AFDCs without resorting to empirical approximations regarding the serial structure of the daily streamflows. Moreover it enables the calculation of conditional distribution $F(y|Y>0)$. The theory of total probability allows for the evaluation in percentage terms of how long the river is dry.

5.3. Methodology

5.3.1. Definition of period of record FDC

Flow duration curves bring together the discharge values and how often this discharge is equalled or exceeded. For a series of daily flows, the FDC can be seen as the complement of the cumulative distribution function of the daily streamflows based on the complete recording of flows.

A non parametric approach to construct the FDC can be used (Vogel and Fennessey, 1994) (as seen in 2.8.5 paragraph).

It is possible to build FDC with different time resolutions of the discharges but daily FDC offer the most detailed way of examining the duration characteristics of a river.

5.3.2. Stochastic index approach for period of record FDC

An approach to model daily streamflows and estimate FDC is based on the stochastic index flow method (Castellarin et al., 2004), that is similar to the flood index approach (Dalrymple, 1960).

The approach assumes that the daily streamflow X can be found by multiplying an index flow equal to the annual flow AF to a dimensionless daily streamflow X',

$$X = AF \cdot X' \qquad (5.9)$$

The climatic conditions and annual precipitation for a given basin influences *AF*. The probability density function, $f_{X'}$, of standardized flows is correlated with the geomorphological characteristics of the basin.

Using this formulation, it is possible to calculate FDC for the complete recording period of flows as the complement of the cumulative distribution function (cdf) of *X*, F_X given by:

$$F_x(x) = P\{X \le x\} = \int_{xl}^{x} f_x(u)du = P\{AF \cdot X' \le x\} = \int_{\Omega_{X'}} \int_{afl}^{x/z} f_{AF,X'}(vv,z)dvd \qquad (5.10)$$

Ω_Y = domain of a given random variable *Y*;

f_X = *pdf* of *X*;

$f_{AF,X'}$ = joint probability distribution of *AF* and *X'*;

x_l and af_l = lower bounds of $\Omega_{X'}$ and Ω_{AF}, respectively.

If it is assumed that *AF* and *X'* are independent, then $f_{AF,X'}$ equals the product of the two marginal distributions, and it is possible to write:

$$F_X(x) = \int_{\Omega_{X'}} f_{X'(z)} \int_{af_l}^{x/z} f_{AF}(v)dvdz = \int_{\Omega_{X'}} f_{X'}(z)F_{AF}(x/z)dz \qquad (5.11)$$

where

F_{AF} = *cdf* of *AF*

$f_{X'}$ = *pdf* of *X'*.

The FDC can be estimated by plotting the variable *X* against the duration, equal to 100(1-F_X) (Castellarin et al. 2004).

5.3.3. Stochastic index approach for period of record FDC in presence of zero data events

The problem of the presence of zero data can be solved using theorem of total probability which is used to determine the probability of occurrence of a non-zero event, given that a zero event has already occurred (Jennings and Benson, 1969).

The theorem is given by:

$$P(X > x) = P(X > x \mid X = 0)\, P(X = 0) + P(X > x \mid X \neq 0) P(X \neq 0) \qquad (5.12)$$

Given that $P(X > x \mid X = 0)$ is zero, as the river flow cannot be negative, Equation 5.13 reduces to:

$$P(X > x) = P(X > x \mid X \neq 0)\, P(X \neq 0) \qquad (5.13)$$

If one calls $P(X \neq 0) = p_{nz}$, it is possible to write Equation 5.14 as:

$$P(X > x) = P(X > x \mid X \neq 0)\, p_{nz} \qquad (5.14)$$

This relationship can be written in the form of the cumulative probability distributions:

$$1 - P(X \le x) = [1 - P(X \le x \mid X \neq 0)]p_{nz} \qquad (5.15)$$

Thus, it is possible to obtain $P\ (X \leq x)$:

$P\ (X \leq x) = p_{dry} + p_{nz} P\ (X \leq x|X \neq 0)$ (5.16)

p_{nz} = the percentage of time that the river is flowing (i.e. $P\ (X \neq 0)$). p_{nz} can be estimated using the plotting position formulation;

p_{dry} = the percentage of time the river is dry, equal to 1- p_{nz}.

Therefore the conditional distribution $P\ (X \leq x|X \neq 0)$ can be calculated using the stochastic index model:

$$P(X \leq x \mid X \neq 0) = \int_{\Omega Xnz'} f_{Xnz'}(z) \cdot F_{AFnz}(x/z)dz \qquad (5.17)$$

$f_{Xnz'}$ = probability density function of non zero X' values;

F_{AFnz} = cumulative distribution function of non zero AF values.

The calculation of the conditional distribution $F_{AFnz} = P\ (AF \leq af\ |AF \neq 0)$ and $f_{X'nz} = P\ (X' \leq x' \mid X' \neq 0)$ with positive values of the series, carried out using a fitting procedure. The empirical frequency distribution conditioned by $AF > 0$ and $X' > 0$ can be calculated on non zero values using a modified Weibull plotting position (Wang and Singh, 1995). In fact it is possible to consider a situation whereby the observed ordered-time series has size n ($y_1, \ldots. y_k, 0, \ldots, 0$) in ascending order of magnitude, where $y_1, \ldots y_k$, are all positive while the others n-k are zero values. To calculate the Weibull plotting position it is not possible to use all the n values, and the formulation 100(1-m/n+1) , for m=1,…,n, but it is necessary to use the formulation with only the k positive values:

D_i =100 (1-i/k+1), for i=1,…k (5.18)

The general formulation of the stochastic index flow model for the use with zero values is obtained putting Equation (5.17) into Equation (5.16):

$$\text{P}\ (X \leq x) = p_{dry} + p_{nz} \cdot \int_{\Omega Xnz'} f_{Xnz'}(z) \cdot F_{AFnz}(x / z)dz \qquad (5.19)$$

5.3.4. Annual based flow duration curves

The construction of annual based AFDC can be done following Vogel and Fennessey (1994) as seen in the 2.8.5 paragraph.

As above seen FDC for different years will differ from the other and from the period of record FDC, because of the expected interannual hydrological variability. In particular some sections of the curves will be above and some below the period of record FDC. To summarize this interannual variability a measure of central tendency can be calculated, as the mean or the median of the annual FDC. Then the mean or median value of the n discharges values will be estimated for each exceedance probability p or Duration D_i. It was demonstrated by Vogel and Fennessey (1994) that the annual based FDC are comparable to period of record FDC except for durations above 0.8 (low flows). In fact in this case the period of record FDC is always lower than the mean or median AFDC. The differences between the two curves are caused by the higher sensitivity to hydrological extremes of the period of record FDC, due to the particular period chosen for the analysis. Furthrmore AFDCs can be applied to estimate flood flow indexes, as well as low flow and water quality indexes, which are usually calculated from the probabilistic structure of daily or weekly mean flows (Claps and Fiorentino, 1997). In fact, the recurrence interval T of AFDCs can be easily expressed as a function of the nonexceedance probability, or equivalently, the percentile p of the AFDCs (Castellarin et al. 2004). The frequency analysis of the streamflow regime for dry years, is equal to the recurrence interval T in years calculated as 100/*p*; T equals 100/(100-*p*) otherwise. As a result, the 2-year AFDC is the 50-percentile of AFDCs, while the 10-year AFDC is the 10-percentile of AFDCs if the interest is in droughts, or the 90-percentile of AFDCs if the frequency analysis focuses on the streamflow regime of wet year.

5.3.5. Stochastic index approach for annual flow duration curves (AFDC)

The daily streamflow for a given year can be called X_j, with j=1,…,n, for n=365. The construction of the AFDC can be done, arranging in ascending order the daily value to obtain:

$$X_{(1)} \leq X_{(2)} \leq ..... \leq X_{(n)} \tag{5.20}$$

$X_{(r)}$, with rank r=1,2,…,n represents the r-th order statistical of the n random variables X_j (Balakrishnan and Rao, 1998).

Equation 5.9 can be written again, using the order statistical, and make possible to obtain the X(r) value in terms of:

$$X_{(r)} = AF \cdot X'_{(r)} \tag{5.21}$$

We assume that the unordered daily streamflows are independent and daily distributed, with cdf equal to $F_{x'}$. This hypotesis is not completely valid, because is known the high degree of serial correlation of daily discharges. However the assumption is maintained because the dependence structure of daily discharge does not influence the AFDC and FDC as long as the interannual variability of annual flows is preserved.

Equation 5.21 makes possible to calculate the mean AFDC, through estimation of expected value of $X_{(r)}$, $E[X_{(r)}]$ for r=1,2,..,n:

$$E[X_{(r)}] = E[AF]E[X'_{(r)}], \text{ for } r=1,2,\ldots,n \tag{5.22}$$

This calculation is possible due to the assumed independence of AF and $X'_{(r)}$.

$E[AF]$ and $E[X'_{(r)}]$ can be estimated from the *pdf* of AF and $X'_{(r)}$. In particular $E[AF]$ can be easily calculated:

$$E[AF] = \int_{-\infty}^{+\infty} x \cdot f(x)dx \qquad (5.24)$$

x represents the random variable *AF*, while *f(x)* represents the *pdf* of *AF*.

Conversely the calculation of $E[X'_{(r)}]$ can be done using the theory of order statisticals under the iid hypotesis (Balakrishnan and Rao, 1998), provided the *pdf* of *X'(r):*

$$f_{(X'_{(r)})}(x) = r\binom{n}{r}[F_{X'}(x)]^{r-1}[1 - F_{X'}(x)]^{n-r}\frac{d}{dx}F_{X'}(x) \qquad (5.24)$$

Determined the expected values of *AF* and $X'_{(r)}$, the graph of the mean AFDC can be obtained plotting the n values of $E[X'_{(r)}]$ against the corresponding duration expressed in terms of the rank.

The Equation 5.21, given in the form of order statisticals is also useful to evaluate the interannual variability of the AFDCs. In fact the expected values of *AF* and $X'_{(r)}$ can be used to calculate the standard deviation of $X'_{(r)}$:

$$\sigma(r) = \sqrt{E[AF]^2 E[X'^2_{(r)}] - E^2[X_{(r)}]} \qquad (5.25)$$

The equation can be written under the assumption of independence of *AF* and $X'_{(r)}$.

5.3.6. Stochastic index approach for annual based FDC in presence of zero data events

In the same way the theorem of total probability can be simply applied to calculate annual based flow duration curves in presence of zero data.

It is in fact possible to write the expected value of a random variable *X* as:

$$E[X] = 0 \cdot (1 - p_{nz}) + E[X|X > 0] \cdot p_{nz} \qquad (5.26)$$

In terms of order statisticals is then possible to write:

$$E[X_{(r)}] = E[AF_{nz}] \cdot E[X'_{nz(r)}] \cdot p_{nz} \qquad (5.27)$$

It is important to show that the expected values of *AF* and *X'* are calculated through knowledge of the pdf of *AF* and $X'_{(r)} > 0$.

It is also possible to find to the expression of the interannual variability of the annual based flow duration curve. In fact for a random variable *X*:

$$E[X^2] = E[X^2|X > 0] \cdot p_{nz} + 0 \cdot (1 - p_{nz}) \qquad (5.28)$$

From this we obtain:

$$E[X^2] = E[AF^2_{nz}] \cdot E[X'^2_{nz(r)}] \cdot p_{nz} \qquad (5.29)$$

It is then possible to calculate the standard deviation $\sigma[X]_{(r)}$:

$$\sigma[X]_{(r)} = E[X^2] - \{E[X]\}^2 \qquad (5.30)$$

$$\sigma[X]_{(r)} = \sqrt{E[AF^2_{nz}]E[X'^2_{nz(r)}] \cdot p_{nz} - E^2[AF_{nz}]E^2[X'_{nz(r)}]p^2_{nz}} \qquad (5.31)$$

5.4. Results of stochastic index

A more thorough description is shown here for the procedure and results for two stations that have intermittent regime localized in the Region 4 (as defined in the Chapter 3), which corresponds to the Apennine region and for the Greek station. Seven years of data, from 2003 to 2010, is available for sites n.3 and n.5. Conversely sixteen years of data, from 1973 to 1989 are available for the station of Aposelemis in Greece.

Figure 5.1 shows the study area and the stations used in the analysis.

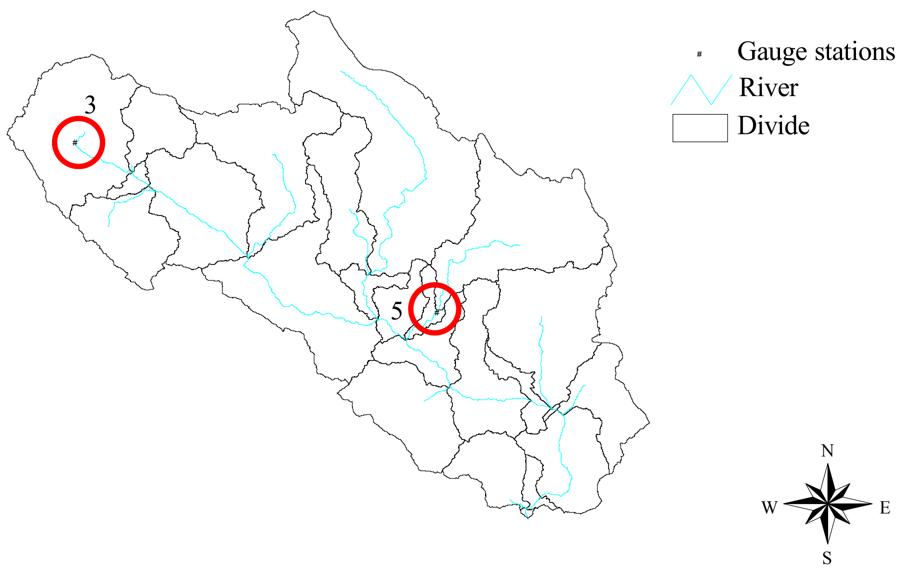

Figure 5.1. Study area, with representation of gauged sites.

Initially *AF* and *X'* were calculated for each time series, and a distribution was fitted on these new series to calculate the FDC using the stochastic index model.

Thus firstly the FDC is calculated using the stochastic index method, without using the theory of total probability. Then the new method is used to calculate FDC when zero data are present The probability of non zero flow p_{nz} and the complement p_{dry} is calculated for each sites and then zero data are separated from the time series. p_{nz} is 95% and 97% respectively for sites n. 3 and n. 5 respectively, while is 43% for Aposelemis station. These values are calculated using the standard plotting position formulation.

Subsequently, a fitting procedure is applied to non zero data. Empirical distribution for AF and X' is calculated using the formulation that considers only non zero data.

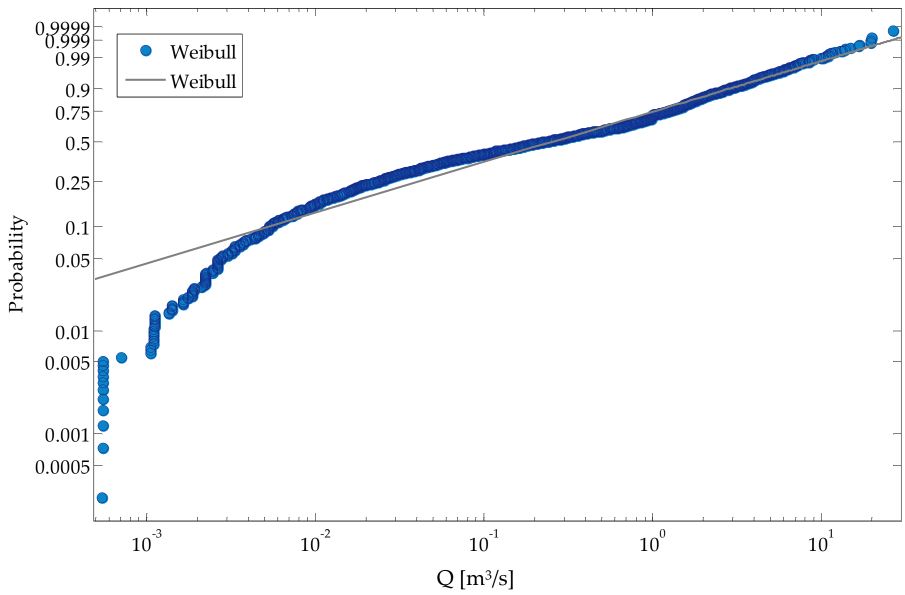

Figure 5.2 - Fitted Conditional Frequency Distribution $F_{X'NZ}=P\ (X' \leq x'|X' \neq 0)$ for sites n.5. The best distributions is the Weibull distribution.

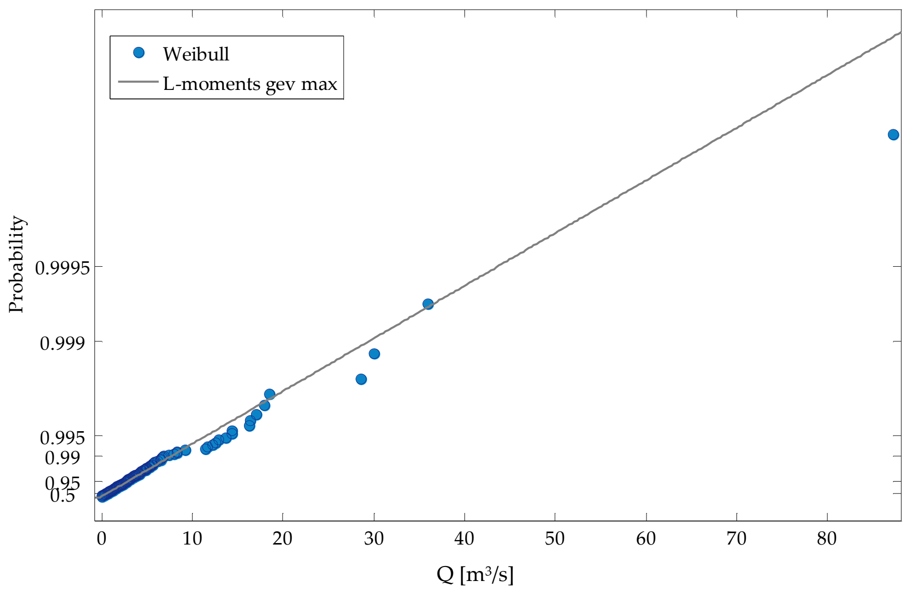

Figure 5.3 - Fitted Conditional Frequency Distribution $F_{X'NZ}=P\ (X' \leq x'|X' \neq 0)$ for n.3 and for Aposelemis station. The best distributions is GEV max for site n. 3.

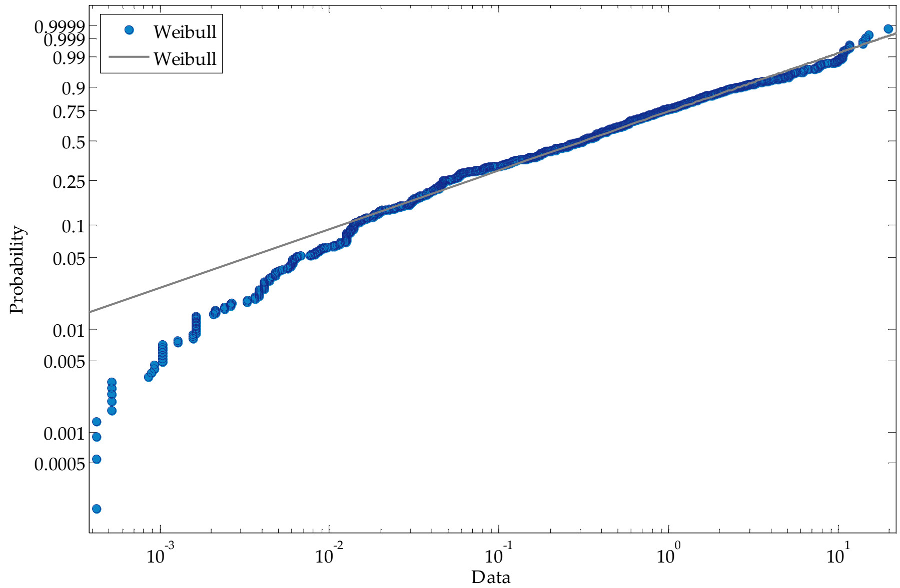

Figure 5.4 - Fitted Conditional Frequency Distribution $F_{X'NZ}=P\ (X' \leq x'|X'\neq 0)$ for Aposelemis station. The best distributions is Weibull distribution for Aposelemis station.

Figures 5.2, 5.3 and 5.4 represent the fitting of distributions to *X'* non zero data. The distributions that fit the values better are Weibull distribution for *X'* non zero data for station n. 5 and for Aposelemis and GEV Max with parameters obtained using L-moments for *X'* data without zero of station n. 3. Normal distribution is the best distribution for *AF* non zero data.

Calculation of the conditional distribution $F_{AFnz}= P\ (AF \leq af \mid AF \neq 0)$ and $f_{X'nz} = P\ (X' \leq x' \mid X' \neq 0)$ with positive values of the series, it is then performed.

At this point it becomes possible to calculate the probability of non zero data using the new formulation of the stochastic index model:

$$P(X \leq x) = p_{dry} + \cdot p_{nz} \cdot \int_{\Omega Xnz'} f_{Xnz'}(z) \cdot F_{AFnz}(x/z)dz \qquad (5.32)$$

Figure 5.5, 5.6 and 5.7 represent the FDC for the three stations. The results are represented in semi logarithmic scale. The broken black line represents the FDC calculated using the Weibull plotting position. The grey full line represents the results calculated using the more general method to take into account the presence of zero data and the bright grey broken line is the FDC calculated using the stochastic index method without zero data. It is possible to highlight that for the station n.3 there is the worse fitting. This is probably caused by the particular shape of the empirical curve, created by the large abstractions in the area.

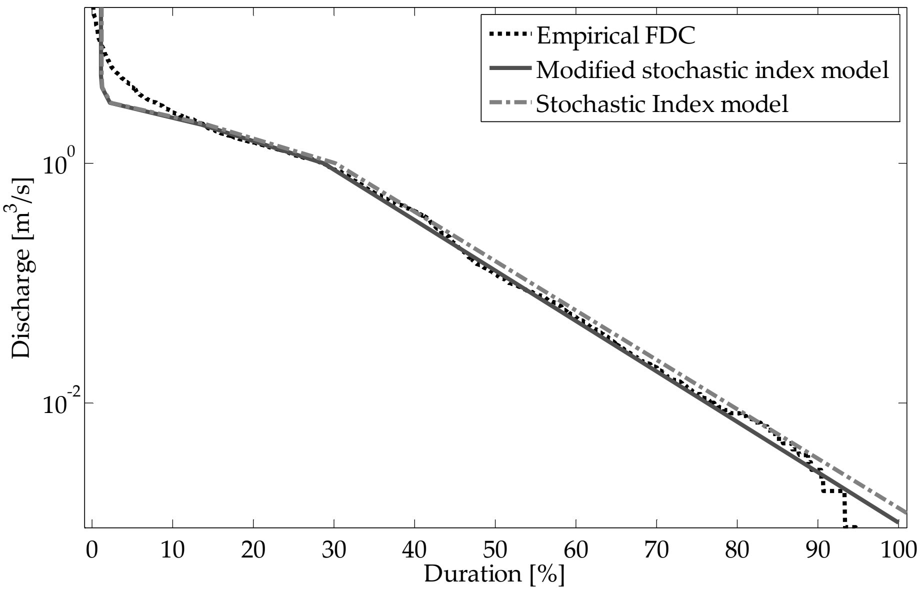

Figure 5.5 - Fitted Flow duration Curves (*P (X>x)*) for site n.5. Black broken line is the empirical FDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the FDC calculated using the standard stochastic index method.

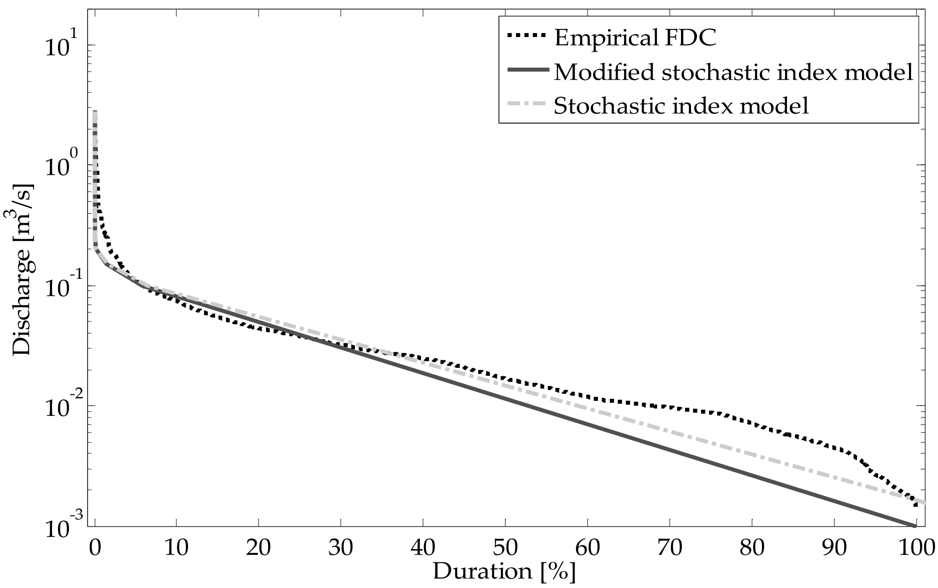

Figure 5.6. Fitted Flow duration Curves (*P (X>x))* for site n.3. Black broken line is the empirical FDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the FDC calculated using the standard stochastic index method.

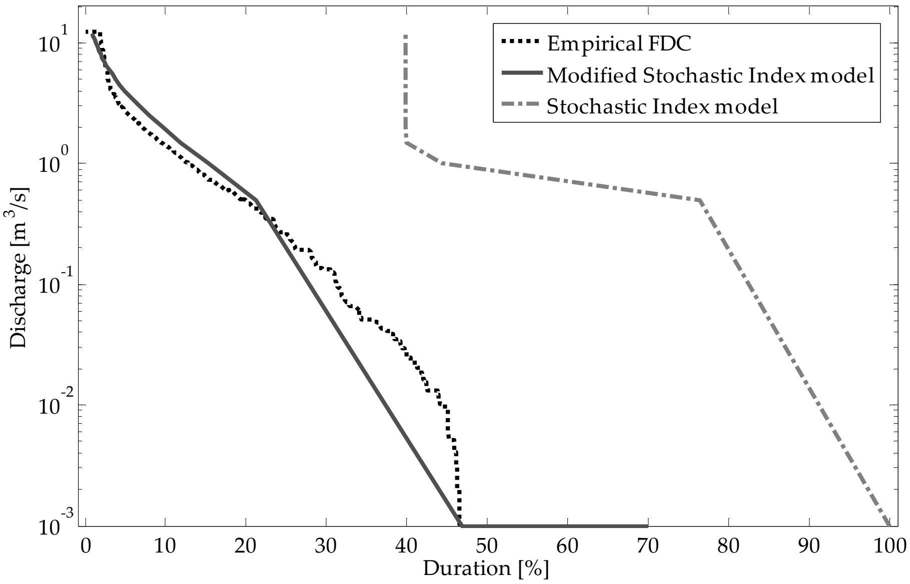

Figure 5.7. Fitted Flow duration Curves (*P (X>x)*) for Aposelemis station. Black broken line is the empirical FDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the FDC calculated using the standard stochastic index method.

Figure 5.8, 5.9 and 5.10 represent the AFDC calculated for station n. 3, n. 5 and Aposelemis. Results are represented with the same simbology than FDC curves. Is clear that there is not a big difference between curves calculated using the new methodology and the old one.

Figure 5.11, 5.12 and 5.13 represents standard deviation calculated for the three stations. Here results obtained with the technique of stochastic index modified are better than those estimated with the simplest approach.

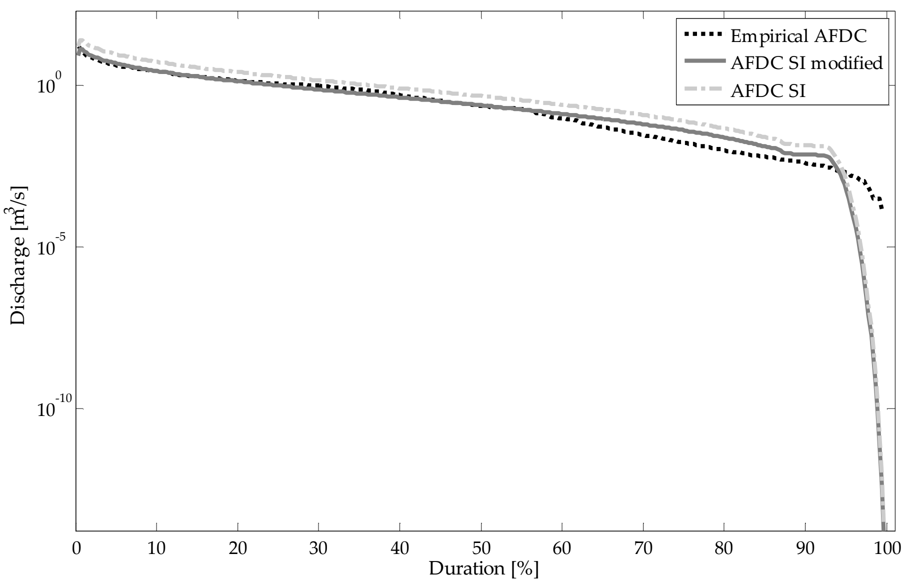

Figure 5.8 - Fitted AFDC (*P (X>x)*) for site n.5. Black broken line is the empirical AFDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the AFDC calculated using the standard stochastic index method.

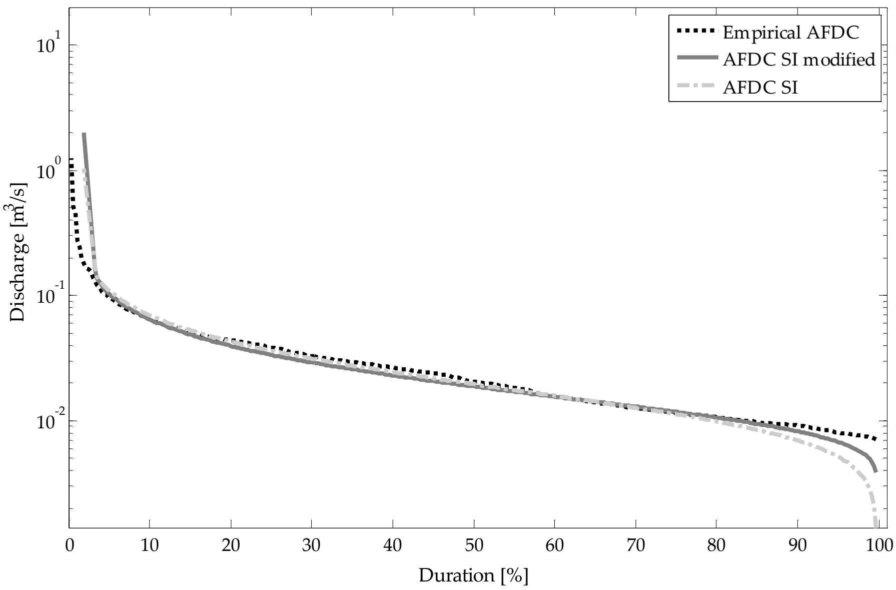

Figure 5.9 - Fitted AFDC (*P (X>x)*) from up to down for site n.3. Black broken line is the empirical AFDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the AFDC calculated using the standard stochastic index method.

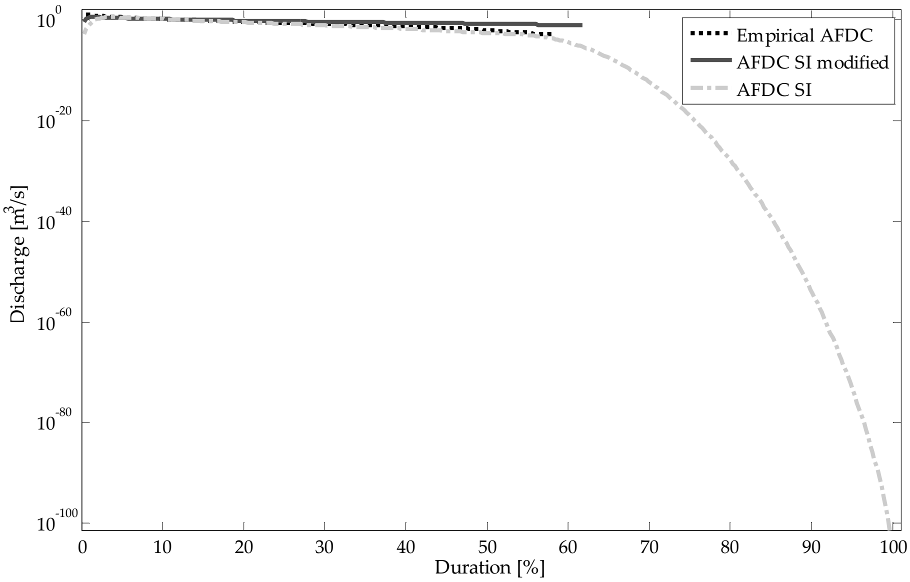

Figure 5.10 - Fitted AFDC (*P (X>x)*) for Aposelemis station. Black broken line is the empirical AFDC, obtained through Weibull plotting position, grey full line is calculated using the stochastic

index method modified to take into account the presence of zeros and bright grey broken line is the AFDC calculated using the standard stochastic index method.

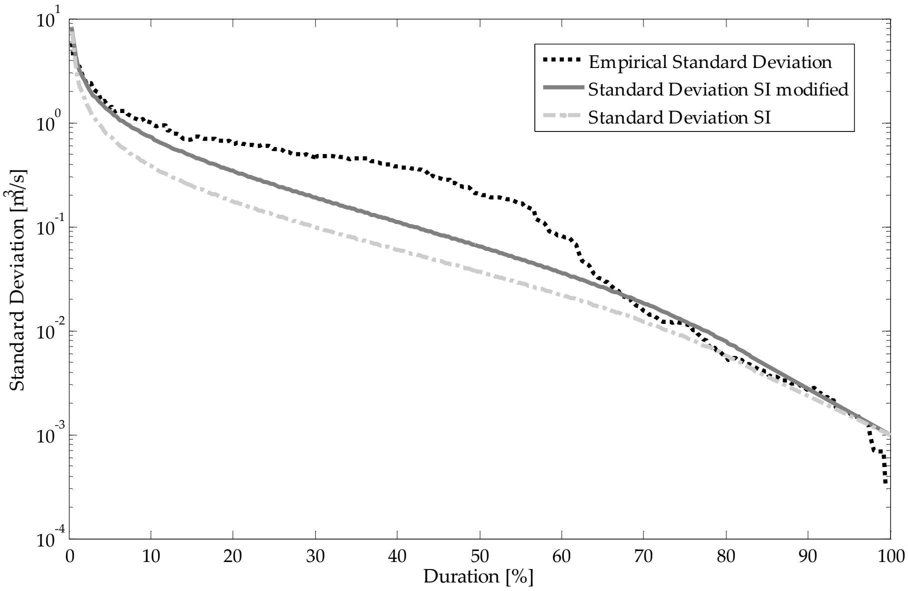

Figure 5.11 - Fitted standard deviation curves (P (X>x)) from up to down for site n.5. Black broken line is the empirical standard deviation, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the standard deviation calculated using the standard stochastic index method.

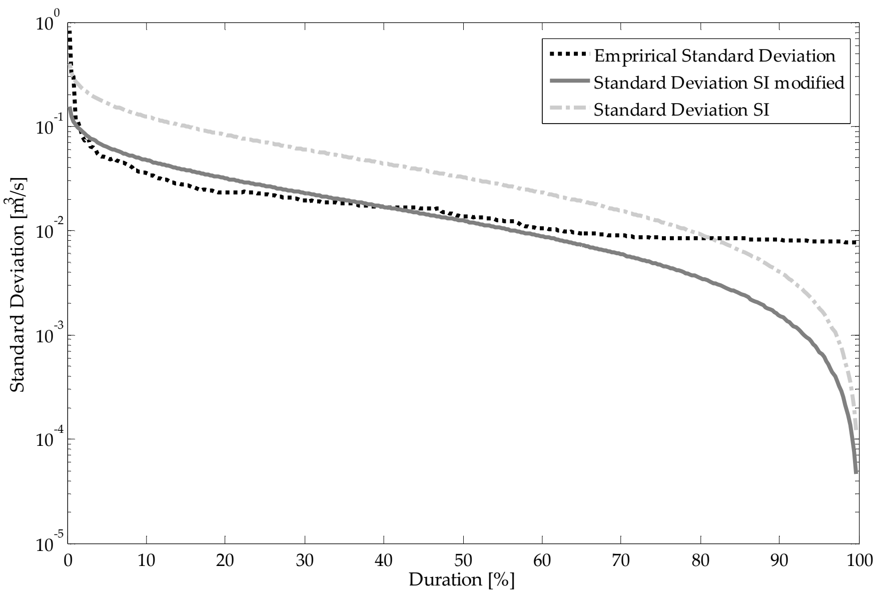

Figure 5.12 - Fitted standard deviation curves (P (X>x)) from up to down for site n.3. Black broken line is the empirical standard deviation, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the standard deviation calculated using the standard stochastic index method.

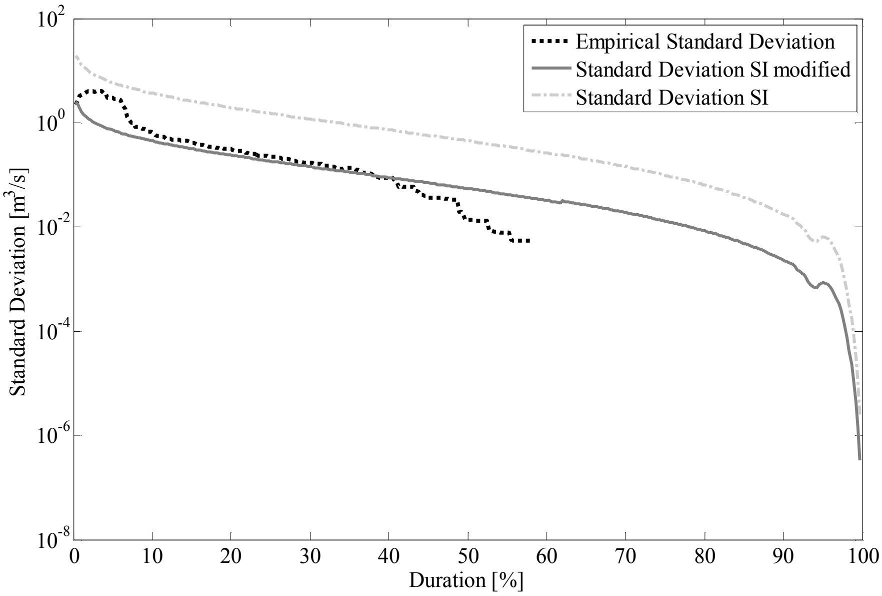

Figure 5.13 - Fitted standard deviation curves (P (X>x)) from up to down for Aposelemis site. Black broken line is the empirical standard deviation, obtained through Weibull plotting position, grey full line is calculated using the stochastic index method modified to take into account the presence of zeros and bright grey broken line is the standard deviation calculated using the standard stochastic index method.

In order to evaluate the accuracy of the modified method, the root mean square error (RMSE) and Nash-Sutcliffe (NS) model efficiency coefficients are calculated.

The formulation of RMSE is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q_o^t - Q_m^t)^2}{n}} \tag{5.33}$$

where Q_0^t is the observed discharge at the time t, and Q_m^t is modeled discharge at the time t.

The formulation of Nash-Sutcliffe efficiency coefficient is:

$$NS = 1 - \frac{\sum_{t=1}^{T}(Q_0^t - Q_m^t)^2}{\sum_{t=1}^{T}(Q_0^t - \overline{Q}_0)^2} \tag{5.34}$$

where $\overline{Q}_0$ is the mean value of the observed discharges.

Nash–Sutcliffe efficiencies can range from −∞ to 1. An efficiency of 1 (NS =1) represents a perfect match of the modelled discharges to the observed data. An efficiency value of 0 (NS = 0) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero (NS < 0) occurs when the observed mean is a better predictor than the model.

Results for RMSE and NS efficiency are shown in Table 5.1 for the stations with intermittent regime. These efficiency indexes are calculated for the FDC, AFDC and standard deviation using stochastic index in the standard formulation and using the modified formulation. It is possible to see that the results for the modified formulation are better than for the standard formulation. In particular, the worst results are those for site n. 1 (without naturalization). The table 4, shows the results of the accuracy indexes for the site n. 1 after the naturalization procedure. It is then defined that the modified method is influenced by the abstractions and by the change of the shape of the flow duration curve.

Table 5.1. RMSE and Nash-Sutcliffe efficiency on the FDC calculated for the three stations with intermittent regime and for the standard and modified formulation of Stochastic index model. The Nash-Sutcliffe efficiency is calculated on logarithms of data.

Site	RMSE SI [m^3/s]	RMSE SI modified [m^3/s]	N-S Efficiency SI	N-S Efficiency SI modified
5	6.10	5.20	0.94	0.95
3	5.70	7.00	0.96	0.94
Aposelemis	30.30	7.22	0.42	0.98

Table 5.2. RMSE and Nash-Sutcliffe efficiency on the AFDC calculated for the three stations with intermittent regime and for the standard and modified formulation of Stochastic index model. The Nash-Sutcliffe efficiency is calculated on logarithms of data.

Site	RMSE SI [m^3/s]	RMSE SI modified [m^3/s]	N-S Efficiency SI	N-S Efficiency SI modified
5	4.13	1.04	-2.88	0.75
3	3.53	1.89	0.93	0.97
Aposelemis	1.04	0.68	0.21	0.92

Table 5.3. RMSE and Nash-Sutcliffe efficiency on the standard deviation calculated for the three stations with intermittent regime and for the standard and modified formulation of Stochastic index model. The Nash-Sutcliffe efficiency is calculated on logarithms of data.

Site	RMSE SI [m^3/s]	RMSE SI modified [m^3/s]	N-S Efficiency SI	N-S Efficiency SI modified
5	0.75	0.58	0.43	0.68
3	0.46	0.44	-1.55	-1.36
Aposelemis	1.23	0.49	-1.12	0.98

The results of the application of the identified five parameters model in the study area and for the Aposelemis station (Greece) show that it can be used satisfactorily to represent FDC, AFDC and interannual variability on rivers with zero flow. Besides it is shown that the percentage of zero data highly influences the response of the method. In fact the application of the modified model produced very good approximations of these curves when there is a big percentage of zero data, while produces a low difference when there is a low percentage. Surely an application in a larger number of basins is necessary to evaluate better the performance of the technique.

6

Regionalization of stochastic index model

6.1. Introduction

A regionalization approach is necessary to permit the calculation of the stochastic index model modified also in ungauged basins. In fact as already seen (Chapter 4) many management and engineering decisions are required in catchments where measured data are not available.

This chapter presents a regional scale application of this model.

The analysis was carried out in Region 4, which corresponds to the Apennine region and comprises eight stations (Figure 6.1). The regionalisation approach was tested in this region involving basins with intermittent flow. Six out of the eight sub-basins are considered to have a permanent regime. For this reason, the stochastic index flow model was applied to these stations with the classical implementation. This means that for these stations the value of p_{nz} is 1 because there is always flow in the river.

The analysis is developed through a multiple regression model between the geomorphoclimatic characteristics of the basins and the parameters of the model. A cross validation analysis through jack knife procedure is assessed in order to evaluate the reliability of the proposed model.

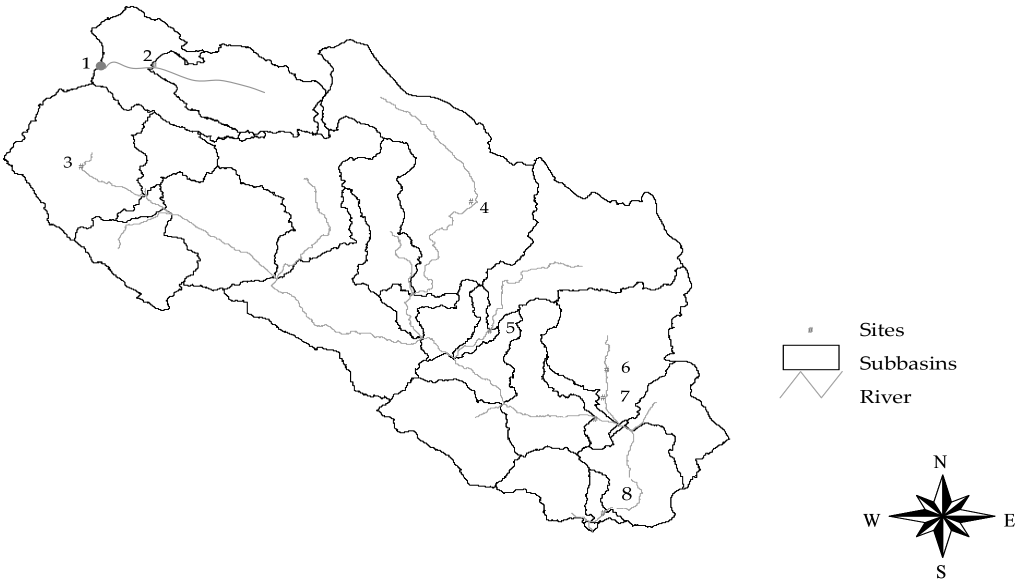

Figure 6.1 - Region 4 (Apenninic basins), with representation of gauged sites used to regionalize the stochastic index model.

6.1.1. Choice of the best parent distribution for AF and X’

To develop a regionalisation model, all the sites must have the same parent distribution; it is then necessary to choose distributions that closely fit *AF* and *X’* data for all stations. It is also important to choose the distributions with the fewest parameters in order to have a parsimonious model.

For *AF* values of the different sites, several distributions have a good fit. The best distribution for the data was the normal one. This result was expected because of the low skewness of the annual flow and the central limit theorem. Moreover, this distribution has only two parameters, the mean μ and standard deviation σ.

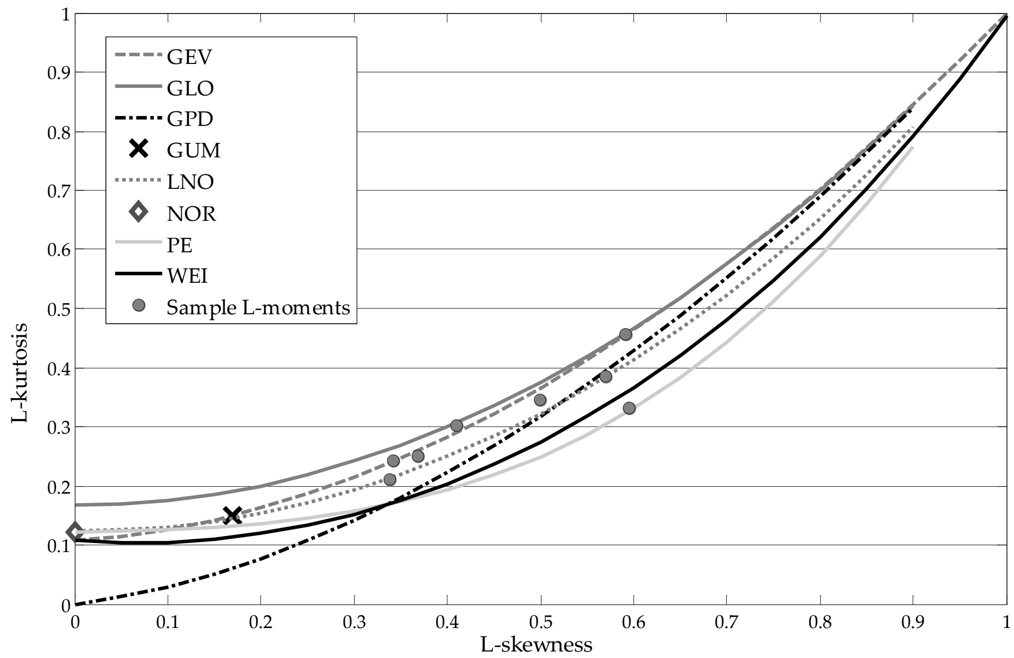

Figure 6.2 - L-moment ratio diagram of L-kurtosis vs. L-skewness, used to choose the parent distribution. Distributions in the diagram are: Normal (NOR), Gumbel (GUM), Generalized extreme values (GEV), Generalized logistic (GLO), Generalized Pareto (GPD), Generalized normal (LNO), Pearson type III (PE), Weibull (WEI).

To evaluate the best distribution for *X'* data, a method based on the L-moment ratio diagram that represents L-skewness versus L-kurtosis was used. Figure 6.2 demonstrates that the GEV distribution gives the best fit for almost all the stations. Observed parameters from the sites of the region are provided in Table 6.1.

A regression model was used to transfer model parameters to ungauged basins (see Section 6.1.2). Parameters of the normal distribution as well as scale parameter λ and shape parameter κ of the GEV distribution are estimated by equations dependent on geomorphologicalal characteristics. It is possible to evaluate location parameter ξ of the GEV distribution using L-moment formulation (Hosking and Wallis, 1997). λ_1,, the mean value of the standardized discharge sample *X',* is equal to unity. Location parameter ξ can thus be evaluated after the estimation of κ and λ parameters:

$$\lambda_1 = \xi + \lambda\{1 - \Gamma(1+\kappa)\}/\kappa \, . \qquad (6.1)$$

where $\Gamma(.)$ represents the gamma function.

The other parameter of the model is the p_{nz} value that evaluates the regime type of the basin and can range between 0 for basins with no flow in the period of measurements to 1 for rivers with permanent regimes.

6.1.2. Regression models

Stepwise regression analysis was performed for all stations in the region. For this type of statistical regression model, the order of entry of the predictor variables is based on an F-test.

A validation analysis, through jack-knife procedure, is then necessary to evaluate the accuracy of the regional estimates. This method allows a simulation of the presence of ungauged basins and is assessed using the procedure below (Castellarin et al., 2004).

One of the stations of the homogeneous region was removed from the sample and a new regression analysis was carried out without it. New parameters were then calculated for this station with the new equations; the results were used to calculate the FDC. The procedure was then applied to all the stations.

The stepwise procedure was then used to define regionalisation models to calculate the five parameters in ungauged basins. Three kinds of models (linear, exponential and logarithmic) were evaluated:

$$\hat{\vartheta} = A_0 + A_1\omega_1 + A_2\omega_2 + A_n\omega_n + \vartheta \,; \qquad (6.2)$$

$$\hat{\vartheta} = A_0 \cdot \omega_1^{\,A_1} \cdot \omega_2^{\,A_2} \cdot \omega_n^{\,A_n} + \vartheta' \,; \qquad (6.3)$$

$$\hat{\vartheta} = A_0 + \ln(A_1 \cdot \omega_1) + \ln(A_2 \cdot \omega_2) + \ln(A_n\omega_n) + \vartheta'' \,. \qquad (6.4)$$

where $\hat{\vartheta}$ is the perfect estimated parameter, A_i , i=1, 2, .., n are the coefficients of the model, ω_i are the explanatory variables and ϑ is the residual of the models.

The regression models identified for the five parameters of the modified stochastic index flow model are:

$$\mu = A_0 \cdot \left(A^{A_1}\right) \cdot \left(F_A^{\,A_2}\right); \qquad (6.5)$$

$$\sigma = A_3 \cdot \left(A^{A_4}\right); \qquad (6.7)$$

$$\kappa = A_5 \cdot \left(F_D^{\,A_6}\right) \cdot \left(H^{A_7}\right); \qquad (6.8)$$

$$\lambda = A_8 \cdot \left(A^{A_9}\right); \qquad (6.9)$$

$$p_{nz} = A_{10} \cdot \left(F_D^{\,A_{11}}\right) \cdot \left(A^{A_{12}}\right). \qquad (6.9)$$

The two parameters of the *AF* data depend on the area of basin A, while parameters of the *X'* data and p_{nz} also depend on the percentage of pervious substrate F_D and on the percentage of calcareous substrate F_A.

Figure 6.3 shows the scatter plots of parameters observed versus the parameters predicted, which were calculated from all sites using the jack-knife procedure. It can be seen that the parameters estimated using the jack-knife procedure are more scattered than those obtained with general regression; however, they are not so different as to imply that the models cannot be used.

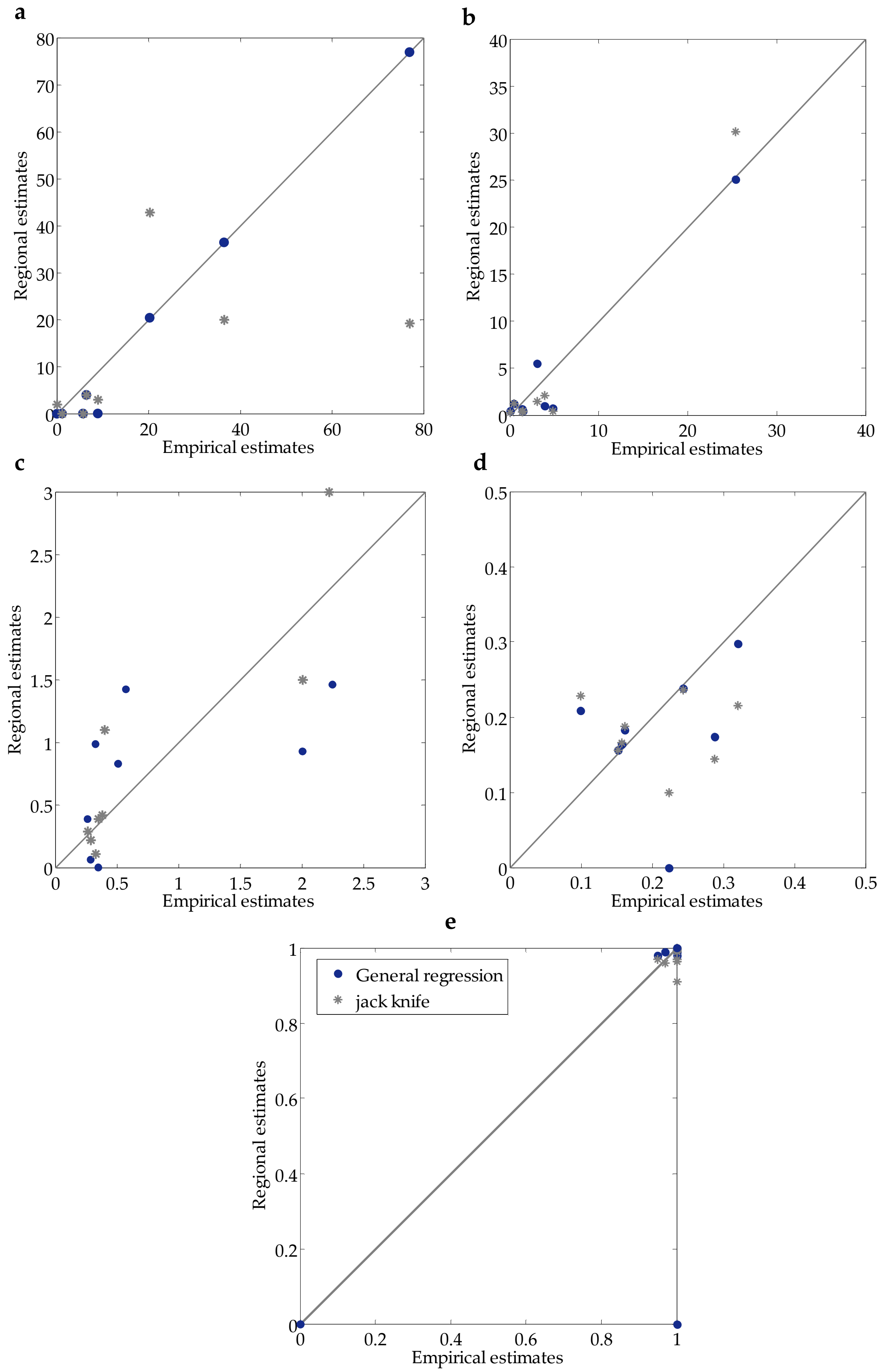

Figure 6.3 - Scatter plot of observed parameters vs. predicted parameters of the stochastic index flow model. Graphs (a) and (b) represent μ and σ parameters of the normal distribution. Graphs (c) and (d) represent κ and λ parameters of the GEV distribution, and Graph (e) is the scatter plot of the p_{nz} values (in probability form) .

6.1.3. Results of regionalisation

The flow duration curves calculated using parameters obtained through jack-knife validation were then compared with empirical FDC in order to evaluate the accuracy of the model.

To evaluate the performance of the model, the following indicator was considered (Castellarin et al., 2004):

$$\varepsilon_{s,j} = \frac{\hat{Q}_{s,j} - Q_{s,j}}{Q_{s,j}} \cdot 100 \qquad (6.10)$$

where $Q_{s,j}$ and $\hat{Q}_{s,j}$ indicate the daily streamflow, empirical and estimated through regionalisation, associated with duration *j* for station *s*. From these values it is possible to obtain the mean relative error $\bar{\varepsilon}_s$ and its standard deviation $\sigma_{\varepsilon,s}$ for a station as:

$$\bar{\varepsilon}_s = \frac{1}{N'} \cdot \sum_{j=1}^{N'} \varepsilon_{s,j} \qquad (6.11)$$

$$\sigma_{\varepsilon,s} = \sqrt{\frac{1}{N'} \cdot \sum_{j=1}^{N'} \left(\varepsilon_{s,j} - \bar{\varepsilon}_s\right)^2} \qquad (6.12)$$

where *N'* represents the number of durations that are considered to calculate mean and standard deviation.

In addition, the average of *N* values of Equation 6.11, $\bar{\varepsilon}$, and of Equation 6.12, σ_{ε}, with *N* corresponding to the number of sites, gives an indication of the performance of the model.

It is also possible to graphically represent the mean and median of the distribution of the *N* relative errors $\varepsilon_{i,j}$ and the 100(p/2) % and 100[1-(p/2)] % percentiles, by identifying the interval about the median containing the 100(1-p) % of the *N* relative errors, against durations *j*, to evaluate the uncertainty of the regional FDC for all durations.

The mean relative error for a given duration *j* can be calculated as:

$$\bar{\varepsilon}_j = \frac{1}{N} \sum_{h=1}^{N} \varepsilon_{h,j} \qquad (6.13)$$

Another performance index E_s that can be used is the Nash-Sutcliffe efficiency method, calculated for each station as:

$$E_s = 1 - \frac{\sum_{j=1}^{N'} (\hat{Q}_{s,j} - Q_{s,j})^2}{\sum_{j=1}^{N'} (Q_{s,j} - \overline{Q}_s)^2} \tag{6.13}$$

where $\hat{Q}_{s,j}$ is the estimated value for each duration j and site s, $Q_{s,j}$ is the empirical value for each duration j and site s and $\overline{Q}_s$ is the mean value. The value of this index can range between 1 and $-\infty$.

The E_s values are used to calculate three indexes of the effectiveness of the model:

- P_1 is defined as the percentage of cases over N stations in which $E_s > 0.95$;
- P_2 is defined as the percentage of cases over N stations in which $0.50 < E_s < 0.95$;
- P_3 is defined as the percentage of cases over N stations in which $E_s < 0.5$.

Figure 6.5 shows FDC results for two explanatory sites. Table 6.1 shows the observed parameters for the stations of the region 4. Tables 6.2, 6.3 and 6.4 show the results of the average relative error and the standard deviation of relative error for the region and the Nash-Sutcliffe efficiency index. Figure 6.6, 6.7 and 6.8 on the other hand, are a graphic representation of mean, median and 10% and 90% percentiles of relative errors for the regionalized FDC, AFDC and standard deviation..

Figure 6.4 and 6.5 show FDC results for two explanatory sites. Figure 6.6, 6.7 and 6.8 on the other hand, shows the results of the graphical representation of mean, median and 10% percentile of relative errors for the regionalized FDC, AFDC and standard deviation.

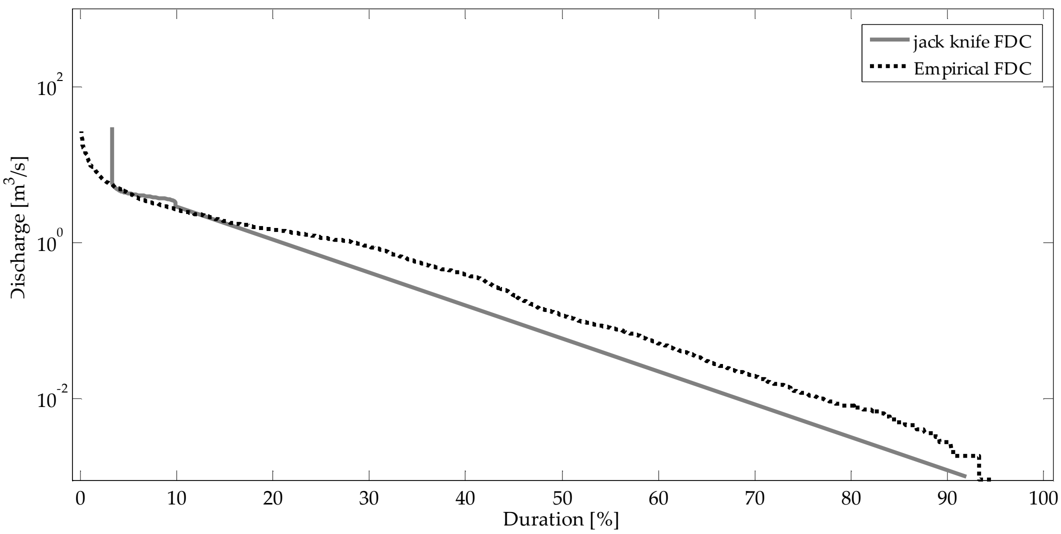

Figure 6.4 - Empirical and jack-knife flow duration curves for site n.5 (intermittent regime).

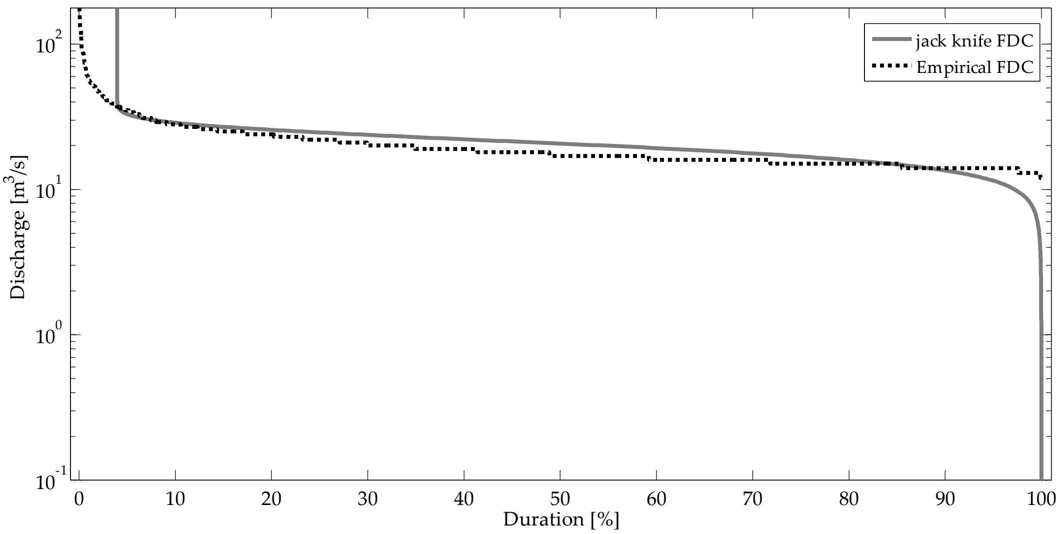

Figure 6.5 - Empirical and jack-knife flow duration curves for site n.1 (permanent regime).

The relative error graph for the regionalized FDC shows that the worst results are produced for smaller durations, but the error decreases and for durations greater than 20% is lower than 1%. The median value is less affected by the anomalous values and is always under 1% and very near to 0%. It is particularly important to note that the relative error calculated for the higher percentage of durations, which coincides with the lower part of the FDC, is very low. The efficiency results are quite good: in fact more than 50% of the sample is a very good fit and only 27% fits poorly.

The relative errors for the regionalized AFDC and standard deviation are quite good. Figures 6.9, 6.10 and 6.11 show the pie representation of the indexes of efficiency by Nash and Sutcliffe. The graphs summarizes the results obtained. The P_1 value, that represents the percentage of cases over the total in which E_s is higher than 0.95, is lowest for the regionalization of AFDC and standard deviation. In particular for the regionalization of the AFDC there is the totally absence of P_2 values, than of intermediate outcomes.

Although the regionalisation of the model and jack knife validation give not bad results. The implementation of the regionalisation model shows that the approach used based on the regression relationships, identified through stepwise regression relationships can be simply adopted in different geographical areas, changing the explicative variables. It is important highlight that the regionalisation approach was applied to a not extensive sample, due to the presence of only a region which comprises intermittent basins. Only a few stations were then used of which only two had zero flow. Hence, further studies are necessary to test the applicability in regions with more stations with zero flow. Besides the nested structure of the used region influences the results of the homogeneity tests and also results of regression models.

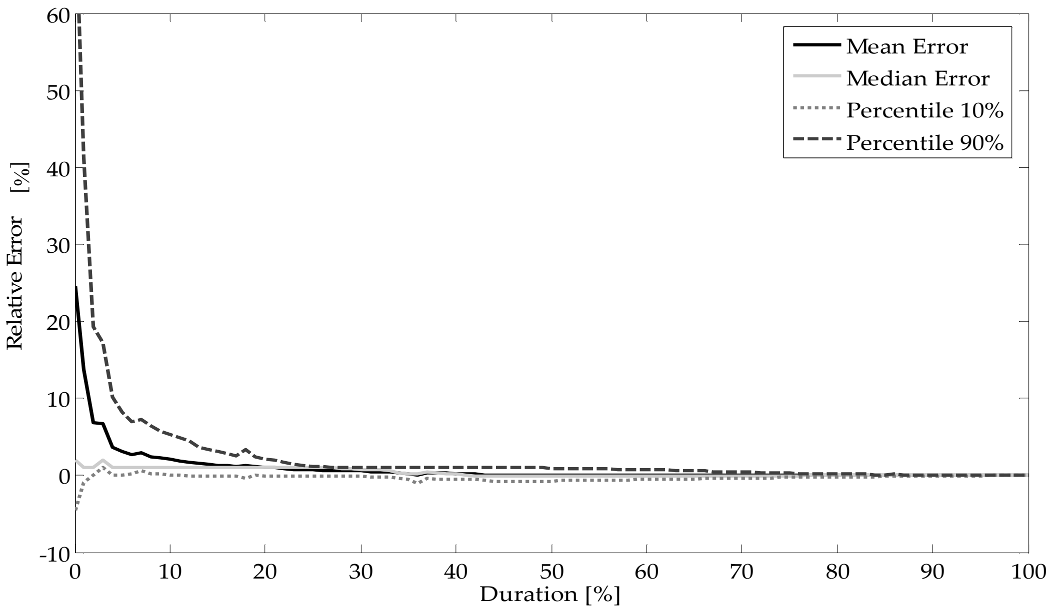

Figure 6.6 - Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the FDC.

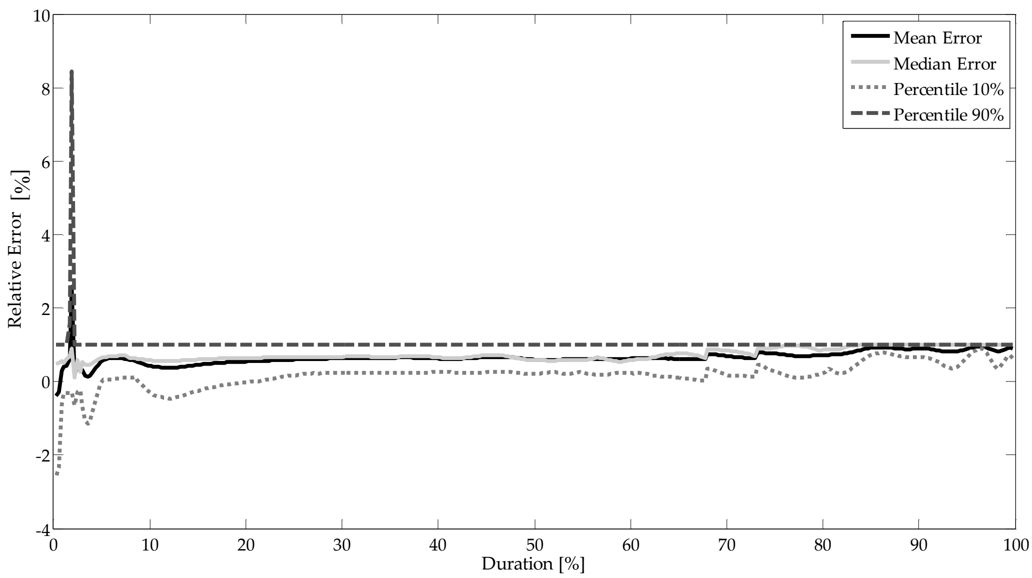

Figure 6.7 - Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the AFDC.

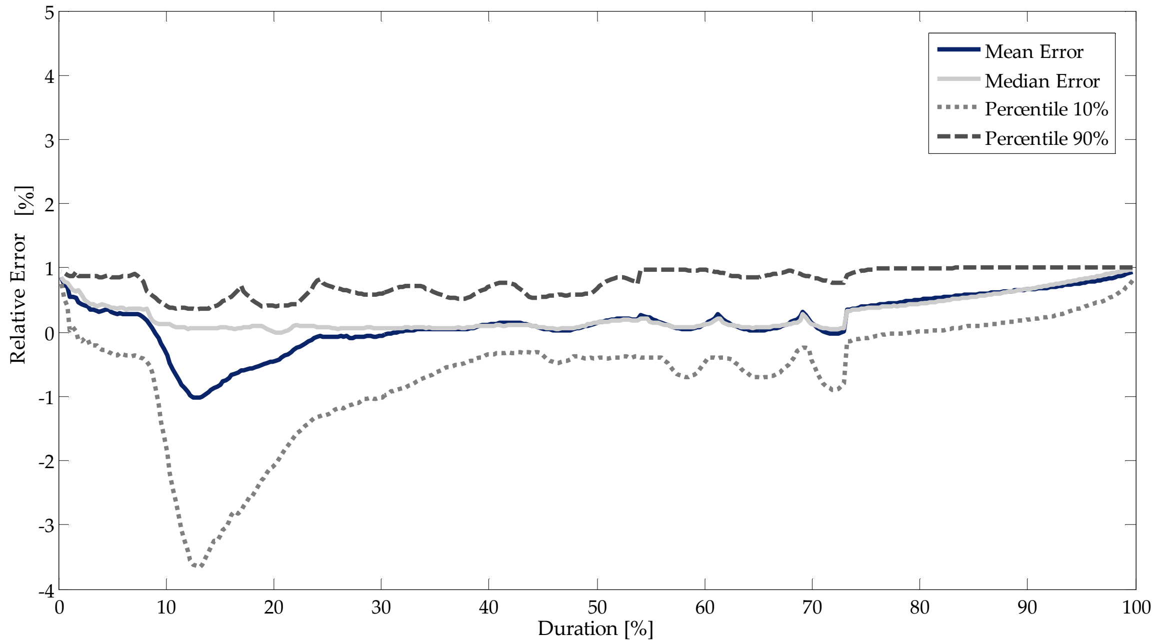

Figure 6.8 - Representation of mean (black full line), median (grey full line) and 10% percentiles (grey and black broken lines) of relative error for different durations for the standard deviation.

Table 6.1 - Observed parameters for the stations of the region 4, (Apenninic area) used for the regionalization approach. m and s are mean and standard deviation of the X' data, *k, l* and x are the shape parameter, scale parameter, and the location parameter of the GEV distribution and p_{nz} is the percentage of time that the river is flowing.

	m	s	k	l	x	p_{nz}
Site 1	20.26	3.03	0.33	0.16	0.83	1.00
Site 2	6.44	1.38	0.29	0.24	0.77	1.00
Site 3	0.04	0.02	0.57	0.32	0.40	0.95
Site 4	5.73	3.89	2.25	0.16	0.07	1.00
Site 5	1.12	0.49	2.01	0.15	0.08	0.97
Site 6	8.95	1.47	0.35	0.22	0.76	1.00
Site 7	36.47	4.88	0.26	0.10	0.91	1.00
Site 8	76.94	25.39	0.51	0.29	0.58	1.00

Table 6.2 - Indexes of reliability calculated on flow duration curves obtained from the regionalization model for the FDC.

$\bar{\varepsilon}$	0.96
$\sigma_{\varepsilon'}$	3.68

Table 6.3 - Indexes of reliability calculated on flow duration curves obtained from the regionalization model for the AFDC.

$\bar{\varepsilon}$	0.64
$\sigma_{\varepsilon'}$	0.39

Table 6.4 - Indexes of reliability calculated on flow duration curves obtained from the regionalization model for the sigma.

$\bar{\varepsilon}$	-0.03
σ_{ε},	0.61

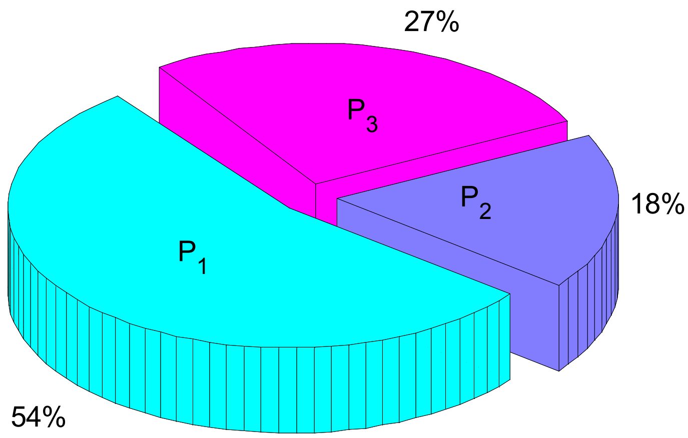

Figure 6.9 - Pie graph of the reliability results for FDC.

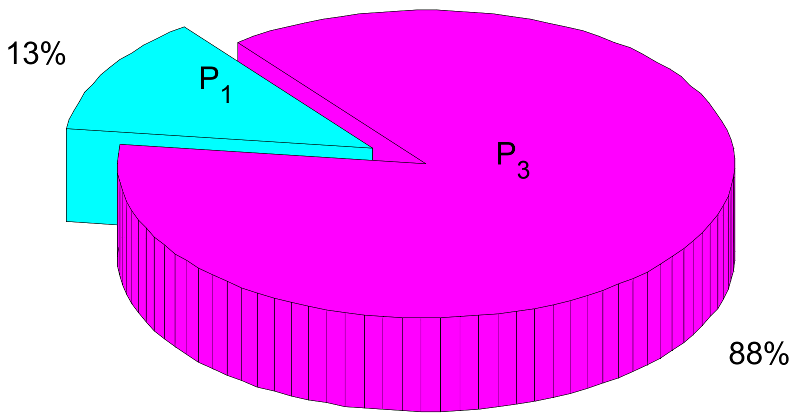

Figure 6.10 - Pie graph of the reliability results for AFDC.

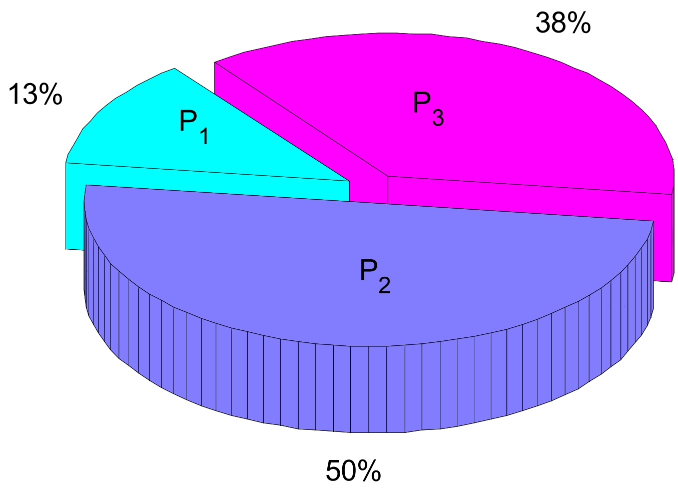

Figure 6.11 - Pie graph of the reliability results for the standard deviation.

7

Conclusions

Low flow analysis are a very important and widely studied topic in hydrology and water resource design and management. The knowledge of low flows periods and of the amount of available water is in fact fundamental in a wide range of scenarios. Conventionally low flow analysis is used for the design and operation of public water supply schemes, but an other frequent application is the evaluation of dilution of waste water that arrives from industrial or domestic source into a river.

Different and complex natural processes at a catchment scale can produce low flows in a river. In fact it is the result of complex processes that happen at a catchment scale. Different models can then be applied to study and schematize this phenomenon. In particular a conceptual model can be used to schematize the catchment. This is constituted by interlinked reservoirs, where recharge, storage and discharge processes take place. While recharge of the catchment mainly depends on precipitation, storage and discharge are functions of physiographic characteristics. Low flows occur therefore after periods of no rain or when precipitation falls as snow. Besides this decreases the water stored in the soils and also the outflow to the river. The time of depletion depends on hydrological processes that operate near the channel as the storage properties within the catchment. Catchment soil and geology influence the capacity of the catchment's precipitation absorption and release as low flow.

The number of summary information derivable from daily time series that describes low flow regimes of a river is quite large. The big number of methods is based mainly on the types of available data and the required output. Furthermore, it depends also on the not clear definition of low flow event, expressed as annual minima, as a threshold discharge or can be indicated by the time that the discharge is lower than a fixed value. Other reasons for diversification are the different methods of expressing the frequency. It is possible in fact, just as in extreme value analysis, consider it as the proportion of time that a discharge is exceeded (flow duration curves) or as the proportion of years that a given low flow occurs. Other methods can be obtained given the different durations or averaging periods that are required in many applications.

One of the largely used tool to evaluate low flow and the river regime is the flow duration curve (FDC). This is a cumulative frequency curve that shows the percentage of time that a given discharge is equalled or exceeded during a fixed period. The choice of this tool was done because it is one of the most informative methods of displaying the flow characteristics of a stream throughout the range of discharge, without regard to the sequence of occurrence.

The application of FDC described in this work and in the cited references, and the fact that a lot of papers were published on this topic starting from 1915 to now testify the big interest of the scientific community in this tool. Besides it appears clearly the usefulness and the wide typology of potential applications. The potential themes treated regarding FDC are the regionalization of FDCs [Croker et al., 2003], uncertainty analysis of FDCs [Yu et al., 2002], the development of a stochastic model for FDCs [Cigizoglu and Bayazit, 2000; Sugiyama et al., 2003], and its use in particular for watershed management [Good and Jacobs, 2001].

One of the most important categories regards surely the regionalization of these measures. In fact many management and engineering decisions are required in catchments where measured data are not available. Hydrological networks have declined in many regions of the world and the system for collecting water information is not always existent or inadequate. This situation surely need to be solved, and many methods exist to do this.

The first aim of this work was to show the phisiographical characteristics of studied area, to underline the hydrometric and geomorphologicalal data used and evaluate the extent of the correlation between these data. Besides the first step of regionalization was employed an homogeneity analysis of the discharge data to identify hydrometric data characterized by the same hydrological distribution is shown. Preliminary analysis showed the important role of horography, area and geology on the discharge. Secondly these variables permit to evaluate through cluster analysis the homogeneous regions in the study area. The reliability of these regions was assessed through standard homogeneity tests.

After that it was decided to review different existing regionalization methods and to evaluate their reliability in the study area. Three different procedures was developed and the evaluation of the different methods was done through a cross correlation analysis. It was highlighted that the method that gives the best results is the parametric one and in the second place the statistical approach. The worst results of the graphical methods was probably caused by the fact that this approach use an unique parameter. This can be surely good for the less number of parameter to estimate, but it makes more difficult to adapt the method to the data also because the big heterogeneity of the curves.

After that Chapter 5 treats of a method to construct a on site model to represent daily FDC and annual AFDC that works also in basins belonging to dry climates characterized by intermittent and ephemeral regimes due to the important engineering problems existing in these kind of basins.
In order to reach this objective, the model combines the stochastic index methodology (Castellarin et al., 2004) with the theory of total probability. Stochastic index methodology permits to model the relationship between the FDC and the mean and variance of the AFDC and to maintain the variability of observed AFDCs without resorting to empirical approximations regarding the serial structure of the daily streamflows. Moreover it enables the calculation of conditional distribution $F(y|Y>0)$. The theory of total probability allows for the evaluation in percentage terms of how long the river is dry.
The results of the application of the identified five parameters model in the study area and for the Aposelemis station (Greece) show that it can be used satisfactorily to represent FDC, AFDC and interannual variability on rivers with zero flow. In fact the application of the general model produced very good approximations of these curves. Surely an application in a larger number of basins is necessary to evaluate the performance of the technique.

A regionalization approach for this method is proposed in the chapter 6. The analysis is developed through a multiple regression model between the geomorphoclimatic characteristics of the basins and the parameters of the model. A cross validation analysis through jack knife procedure is assessed in order to evaluate the reliability of the proposed model. The proposed method of regionalization showed good results for the curves although the little number of stations used and the nested configuration of the catchment stations.

References

Alaouze, C.M., 1989. Reservoir releases to uses with different reliability requirements. Water Resour. Bull. 25 (6), 1163– 1168.

Alaouze, C.M., 1991. Transferable water entitlements which satisfy heterogeneous risk preferences. Aus. J. Agric. Econ. 35 (2), 197–208.

Amusja, A.Z., Ratner, N.S., Sokolov, B.L., 1991. Minimum riverflow: state of art and prospects for research. Trudy GGI (Trans. State Hydrol. Inst., Leningrad, USSR) 350, 3–28.

Anderson, M.G., Burt, T.P., 1980. Interpretation of recession flow. J. Hydrol. 46, 89–101.

Armbruster, J.T., 1976. An infiltration index useful in estimating low-flow characteristics of drainage basins. J. Res. USGS 4 (5), 533-538.

Armentrout, G.W., Wilson, J.F., 1987. Assessment of low flows in streams in northeastern Wyoming. USGS Water Resources Investigations Report, 85-4246, 30 pp.

Aron, G., Emmanuel, G.B., 1982. Evaluation of drought potentials and damages in the northeastern United States. Institute for Research on Land and Water Resources. Completion Report, 158 pp.

Aucott, W.R., Meadows, R.S., Patterson, G.G., 1987. Regional ground water discharge to large streams in the upper coastal plain of south Carolina and parts of north Carolina and Georgia. USGS Water Resources Investigations Report 86-4332, 28 pp.

Bako, M.D., Hunt, D.N., 1988. Derivation of recession constant using computer and numerical analysis. Hydrol. Sci. J. 33 (4), 357–367.

Bako, M.D., Owoade, A., 1988. Field application of a numerical method for the derivation of baseflow recession constant. Hydrol. Processes 2, 331–336.

Balakrishnan, N., Rao, C.R., 1998. Order statistics: An introduction, in Order Statistics: Theory and Methods, Handbk. Stat., vol. 16, edited by N. Balakrishnan and C. R. Rao, pp. 3 –24, Elsevier Sci., New York.

Banks, C.H., 1961. The hydrological effects of riparian and adjoining vegetation. For. S. Afr. 1, 31–46.

Barnes, C.R., 1986. Methods for estimating low-flow statistics for ungaged streams in the lower Hudson River basin, NY. USGS Water Resources Investigations Report 85-4070, 22 pp.

Barrows, H. K., 1943. Water Power Engineering: 3d edition, New York, Mc Graw-Hill Book Co., 791 p.

Beard, L. R., 1943. Statistical Analysis in Hydrology, Am. Soc. Civil Engineers Trans., v. 108, 1110-1160.

Beran, M. A. ,Gustard, A., 1977. A study into the low-flow characteristics of British rivers. J. Hydrol. 35, 147-157.

Beran, M.A., Rodier, J.A., 1985. Hydrological aspects of drought. UNESCO-WMO Studies and Reports in Hydrology, vol. 39, 149 pp.

Bickerton, M., Petts, G., Armitage, P., Castella, E., 1993. Assessing the ecological effects of groundwater abstraction on chalk streams: three examples from eastern England. Regul. Rivers Res. Manag. 8 (12), 121–134.

Biggs, B.J.F., 1982. The effects of low flow regulation on the biology and water quality of rivers with particular reference to an impoundment-regulated flow. In: River Low Flows: Conflicts and Water Use. New Zealand Ministry of Works and Development Water and Soil Miscellaneous Publication, vol. 47, pp. 28–42.

Bingham, R.H., 1986. Regionalization of winter low-flow characteristics of Tennessee streams. USGS Water-Resources Investigations Report 86-4007, 88 pp.

Birtles, A.B., 1978. Identification and separation of major base flow components from a stream hydrograph. Water Resour. Res. 14 (5), 791–803.

Biswas, H., Bell, B.A., 1984. A method for establishing site-specific stream design flows for wasteload allocations. J. Water Pollut. Control Fed. 56 (10), 1123–1130.

Blodgett, J.C., Walters, J.R., Borchers, J.W., 1992. Streamflow gains and losses and selected flow characteristics of Cottonwood creek, north-central California, 1982–1985. Water Resources Investigations Report 92-4009, 19 pp.

Bonacci, O., Tadic, Z., Trninic, D., 1992. Effects of dams and reservoirs on the hydrological characteristics of the lower Drava River. Regul. Rivers Res. Manag. 7 (4), 349–357.

Boughton, W.C., Freebairn, D.M., 1985. Hydrograph recession characteristics of some small agricultural catchments. Aust. J. Soil. Res. 23, 373–382.

Bowles, D.S., Riley, J.P., 1976. Low flow modeling in small steep watersheds. J. Hydraul. Div., ASCE 102 (HY9), 1225–1239.

Brutsaert, W., Neiber, J.L., 1977. Regionalised drought flow hydrographs from a mature glaciated plateau. Water. Resour. Res. 13 (3), 637–643.

Bulu, A., 1997. Statistical analysis of low flows with zerodischarges. FRIEND: Flow Regimes from International Experimentaland Network Data, Third Report, Cemagref, 167–170.

Bulu, A., Onoz, B., 1997. Frequency analysis of low flows by the PPCC tests in Turkey. FRIEND'97-Regional Hydrology: Concepts and Models for Sustainable Water Resources Management. IAHS Publication No. 246, pp. 133–140.

Burn, D.H., DeWit, W.J., 1996. Spatial characterization of drought events using synthetic hydrology. Can. J. Civil Engng 23 (6), 1231–1240.

Castellarin, A., Camorani, G., Brath, A., 2007. Predicting Annual and Long-Term Flow-Duration Curves in Ungauged Basins, Adv. in Water Resour., , 30/4, pp. 937 – 953.

Castellarin, A.,Vogel, R. M., and Brath, A., 2004. A stochastic index flow model of flow duration curves, Water Resour. Res., 40, W03104.

Cervione Jr., M.A., Melvin, R.L., Cyr, K.A., 1982. A method for estimating the 7-day, 10-year low flow of streams in Connecticut. Connecticut Water Resour. Bull. No. 34, 17 pp.

Chang, T.J., Stenson, J.R., 1990. Is it realistic to define a 100-year drought for water management? Water Resour. Bull. 26 (5), 823–829.

Chapman, T.G., 1991. Comment on Evaluation of automated techniques for base flow and recession analyses by R.J., McMahon, T.A. Water Resour. Res. 27 (7), 1783–1784.

Chiang, S.L., Johnson, F.W., 1976. Low flow criteria for diversions and impoundments. J. Water Resour. Plan. Manag. Div., ASCE 102 (WR2), 227–238.

Cigizoglu, H. K. and Bayazit, M., 2000. A generalized seasonal model for flow duration curve, Hydrol. Process., 14, 1053–1067.

Claps, P. and Fiorentino, M., 1997. Probabilistic flow duration curves for use in environmental planning and management, In: Harmancioglu NB, et al. (eds.) Integrated Approach to Environmental Data Management Systems, NATO ASI Series, Ser. 2 (31), 255–266, Kluwer Academy.

Clausen, B., 1992. Modelling streamflow recessions in two Danish streams. Nordic Hydrol. 23 (2), 73–88.

Clausen, B., Pearson, C.P., 1995. Regional frequency analysis of annual maximum streamflow drought. J. Hydrol. 173, 111–130.

Clausen, B., Young, A.R., Gustard, A., 1994. Modelling the impact of groundwater abstractions on low-river flow. FRIEND: Flow Regimes from International Experimental and Network Data, IAHS Publication No. 221, pp. 77–85.

Collins, D.N., 1982. Water storage in an alpine glacier. Hydrological Aspects of Alpine and High-Mountain Areas, IAHS Publication No. 138, pp. 113–122.

Croker K. M., Young M. D. Z. , Rees H. G., 2003. Flow duration curve estimation in ephemeral catchments in Portugal, Hydrol Sci J; 48(3):427–39.

Cumming Cockburn Ltd, 1990. Assessment of the biologically based low flow analysis technique. Ontario Ministry of the Environment, Canada, 10 pp.

Curran, J.C., 1990. Low flow estimation based on river recession rate. J. Inst. Water Environ. Manag. 4 (4), 350–355.

Dalrymple, T., 1960. Flood frequency analyses, U.S. Geological Survey, Water Supply Papers, 568 1543-A..

Demuth, S., 1994. Regionalisation of low flows using a multiple regression approach—a review. Proceedings of the XVIIth Conference of Danube countries, Budapesht, vol. 1, pp. 115–Conference of Danube countries, Budapesht, vol. 1, pp. 115–122

Dingman, S.L., 2002. Physical Hydrology. Second edition, Prentice Hall, Upper Saddle River, Nj, United States.

Dracup, J.A., Lee, K.S., Paulson, E.G., 1980. On the definition of droughts. Water Resour. Res. 16

(2), 297–302.

Durrans, S. R., Ouarda, T. B. M. J., Rasmussen, P. F., and Bobée, B., 1999. Treatment of zeroes in tail modelling of low flows, J. Hydrol. Eng., 4(1), 19–27.

Durrans, S.R., 1996. Low-flow analysis with a conditional Webull tail model. Water Resour. Res. 32 (6), 1749–1760.

Durrans, S.R., Quarda, T.B.M.J., Rasmussen, P.F., Bobee, B., 1999.Treatment of zeroes in tail modeling of low flows. J. Hydrol.Eng., 4 (1), 19–27.

DWAF: Department of Water Affairs and Forestry, 1995. Procedures to Assess Effluent Discharge Impacts, Pretoria, South Africa. 352pp.

Eheart, J.W., Tornil, D.W., 1999. Low-flow frequency exacerbation by irrigation withdrawals in the agricultural midwest under various climatic scenarios. Water Resour. Res. 35 (7), 2237–2246.

El-Jabi, N., Ashkar, F., Issa, M., 1997. Application stochastique du phe´nome`ne d'e´tiage. Can. J. Civil Engng 24 (2), 191–200.

Estes, C., Osborn, J.F., 1986. Review and analysis of methods for quantifying instream flow requirements. Water Resour. Bull. 22 (3), 389–398.

Felton, G.K., Currens, J.C., 1994. Peak flow rate and recessioncurve characteristics of a karst spring in the Inner Bluegrass, central Kentucky. J. Hydrol. 162, 99–118.

Fendekova, M., Nemethy, P., 1994. Low flows of Starohorsky Potok brook and their affecting factors. FRIEND: Flow regimes from International Experimental and Network Data, IAHS Publication No. 221, pp. 163–169.

Fennessey , N. M., and Vogel, R. M., 1990. Regional flow duration curves for ungauged sites in Massachusets, Journal of Water Resources Planning and Management, ASCE, Vol. 116 (4), pp. 530-549.

Fennessey NM., 1994. A hydro-climatological model of daily streamflows for the northeast United States. PhD dissertation, Tufts University, Medford, MA.

Ferguson, B.K., Suckling, P.W., 1990. Changing rainfall–runoff relationships in the urbanizing peachtree creek watershed, Atlanta, Georgia. Water Resour. Bull. 26 (2), 313–322.

Finlayson, B.L., Gippel, C.J., Brizga, S.O., 1994. The effects of reservoirs on downstream aquatic habitat. J. Aust. Water Works Assoc. 21 (4), 15–20.

Fleig, A., Tallaksen, L.M.,Hisdal H. and demuth, S.,2006. A global evaluation of streamflow drought characteristics, Hydrology and Hearth System Sciences, 10:535-552.

Foster, H. A., 1934. Duration curves. Trans., ASCE, 99, 1213-1267.

Fountain, A.G., Tangborn, W.V., 1985. Effect of glaciers on streamflow variations. Water Resour. Res. 21 (4), 579–586.

Franchini M, Suppo M., 1996. Regional analysis of flow duration curves for a limestone region. Water Resour. Manage.;10:199–218.

FREND: Flow Regimes From Experimental And Network Data, 1989. I: Hydrological Studies; II:

Hydrological Data, Wallingford, UK.

Furness LW., 1959. Kansas streamflow characteristics-part 1 - Flow duration. Kansas Water Resources Board technical report no. 1. 213 p.

Ganora, D., Claps, P., Laio, F., and Viglione, A., 2009. An approach to estimate nonparametric flow duration curves in ungauged basins, Water Resour. Res., 45, W10418.

Gerard, R., 1981. Regional analysis of low flows: a cold region example. Proceedings of the Fifth Canadian Conference, Canadian Society for Civil Engineers, vol. 1, pp.95-112.

Gerasimenko, 1972. Effects of lakes on minimum streamflow in West Siberia. Trudy GGI (Trans. State Hydrol. Inst., Leningrad, USSR) 188, 305–311.

Gorgens, A.H.M., Hughes, D.A., 1982. Synthesis of streamflow information relating to the semi-arid Karoo Biome of South Africa. S. Afr. J. Sci. 78, 58–68.

Good, J. C., Jacobs, J. M., 2001. Ecologically sustainable watershed management using annualized flow duration curves, paper presented at World Water and Environmental Resources Congress, Am. Soc. of Civ. Eng., Orlando, Fla., May.

Gordon, N.D., McMahon, T.A., Finlayson, B.L., 1992. Stream Hydrology (An Introduction for Ecologists). Wiley, NewYork, 526 pp.

Griffiths, G.A., Clausen, B., 1997. Streamflow recession in basins with multiple water storages. J. Hydrol. 190, 60–74.

Gurnell, A.M., 1993. How many reservoirs? An analysis of flow recession from a glacier basin. J. Glaciol. 39, 409–414.

Gustard A., A. Bullok and J.M. Dixon ,1992. Low flow estimation in the United Kingdom, Institute of Hydrology Report No. 108, Wallingford, United Kingdom.

Gustard, A., 1989. Compensation flows in the UK: a hydrological review. Regul. Rivers Res. Manag. 3 (1–4), 49–59.

Gustard, A., Cole, G.A., 1987. Towards a Rational Assessment of Residual Flows Below Reservoirs Regulated Streams: Advances in Ecology. Plenum Press, New York (pp. 267–275).

Gustard, A., Wesselink, A., 1993. Impact of land use change on water resources: Balquidder catchments. J. Hydrol. 145, 389– 401.

Hall, F.R., 1968. Base flow recessions — a review. Water Resour. Res. 4 (5), 973–983.

Hall, A.J., 1971. Baseflow recessions and baseflow hydrograph separation problem. Proceedings of the Hydrology Symposium, Institution of Engineers, Australia, Canberra, pp. 159–170.

Haan, C. T., 1977. Statistical methods in hydrology, Iowa State University Press, Ames, Iowa, 576.

Hall, G.C., Du Plessis, H.M., 1984. Studies of mineralization in Great Fish and Sundays rivers. Vol. 2: modeling river flow and salinity. Water Research Commission, Working Group for Mineralization. Pretoria, South Africa, 377 pp.

Harboe, R., 1988. Including daily constraints in a monthly reservoir operation model for low-flow. Adv. Water Res. 11 (2), 54–57.

Heicher, D.W., 1993. Instream flow needs: biological literature review. Susquehanna River Basin Commission No. 149, 37 pp.

Hindley, D.R., 1973. The definition of Dry Weather Flow in river flow measurement. J. Inst. Water Engng 27, 438–440.

Hopkinson, C., Young, G.J., 1998. The effect of glacier wastage on the flow of the Bow River at Banff, Alberta, 1951–1993. Hydrol. Processes 12 (10–11), 1745–1762.

Hosking, J. R. M., Wallis, J. R., 1997. Regional frequency analysis, Cambridge University Press, New York.

Hosking, J. R. M., Wallis, J. R., 1988. The effect of inter-site dependence on regional flood frequency analysis, Water Resour. Res., 24, 588–600.

Hughes, D. A., Smakhtin, V. Y., 1996. Daily flow time series patching or extension: A spatial interpolation approach based on flow duration curves, Hydrol. Sci. J., 41(6), 851– 871.

Hughes, D.A., O'Keeffe, J.H., Smakhtin, V.Y., King, J., 1997. Development of an operating rule model to simulate time series of reservoir releases for instream flow requirements. Water SA 23 (1), 21–30.

Institute of Hydrology, 1980. Low Flow Studies (1–4), Wallingford, UK.

Jennings, M. E. and Benson, M. A. 1969. Frequency curve for annual flood series with some zero 582 events or incomplete data, Water Resour. Res., 5(1), 276–280.

Jones, P.N., McGilchrist, C.A., 1978. Analysis of hydrological recession curves. J. Hydrol. 36, 365–374.

Jozeph, E.S., 1970. Probability distribution of annual droughts. J. Irrig. Drain. Div., ASCE 96 (4), 461–474.

Kelman, J., 1980. A stochastic model for daily streamflow. J. Hydrol. 47, 235–249.

Keppeler, E.T., Ziemer, R.R., 1990. Logging effects on streamflow: water yield and summer low flows at Caspar Creek in northwestern California. Water Resour. Res. 26 (7), 1669–1679.

Khan, L.R., Mawdsley, J.A., 1988. Reliable yield of unconfined aquifers. Hydrol. Sci. J. 33 (2), 151–171.

Kilmartin, R. F. , Peterson, J. R., 1972: Rainfall–runoff regression with logarithmic transforms 584 and zeros in the data, Water Resour. Res., 8(4), 1096–1099.

Klemes, V., 1987. One hundred years of applied storage reservoir theory. Water Resour. Manag. Ser., 1.

Knisel, W.G., 1963. Baseflow recession analysis for comparison of drainage basins and geology. J. Geophys. Res. 68 (12), 3649–3653.

Korkmas, N., 1990. The estimation of groundwater recharge from spring hydrographs. Hydrol. Sci. J. 35 (2/4), 209–217.

Kosinsky, V.de., 1984. Management of river basins during drought periods. Proceedings of the Fifth International Conference on Water Resources Planning and Management Water in the year 2000, Athens, Catania, pp. 57–69.

Kottegoda, N.T., Natale, L., Raiteri, E., 2000. Daily streamflow simulation using recession characteristics. J. Hydrol. Engng 5 (1), 17–24.

Kroll, C. N. , Stedinger, J. R., 1996. Estimation of moments and quantiles using censored data, 586 Water Resour. Res., 32(4), 1005–1012.

Kurdov, A.G., 1977. The problem of mimimum river flow calculation.Problems of river hydrology, Moscow University,Moscow, 1977. pp. 65–80.

Kuusisto, E., 1987. Winter and summer low flows in Finland. Aqua Fenn. 16 (2), 181–186.

Laaha, G. and Bloshl, G., 2007. A national low flow estimation procedure for Austria, Hydrological Sciences, Journal des Sciences Hydrologiques, 52 (4), 625–644.

Lall, U., 1995. Recent advances in non-parametric function estimation:hydrological applications. Rev. Geophys. (Suppl.), 1093–1102.

Lanen, H.A.J. van, Tallaksen, L.M., Kasparek, L., Querner, E.P., 1997. Hydrological drought analysis in Hupsel basin using different physically based models. In: FRIEND'97 — Regional Hydrology: Concepts and Models for Suatainable Water Resource Management, IAHS Publication No. 246, pp. 189–196.

Lawal, S.A., Watt, W.E., 1996a. Frequency analysis of low-flows using the Akaike information criterion. Can. J. Civil Engng 23, 1180–1189.

Lawal, S.A., Watt, W.E., 1996b. Non-zero lower limit in low flow frequency analysis. Water Resour. Bull. 32 (6), 1159–1166.

LeBoutillier, D. V. , Waylen, P. R., 1993: A stochastic model of flow duration curves, Water Resour. Res., 29(10), 3535–3541.

Leppajarvi, R., 1989. Frequency analysis of flood and low flow. FRIENDS in Hydrology, IAHS Publication No. 187, pp. 435–442.

Loaiciga, H.A., Marino, M.A., 1988. Fitting minima of flows via maximum likelihood. J. Water Res. Plan. Manag., ASCE 114(1), 78–90.

Loaiciga, H.A., Michaelsen, J., Hudak, P.F., 1992. Truncated distributions in hydrologic analysis. Water Resour. Bull. 28 (5), 853–863.

Loganathan, G.V., Kuo, C.Y., McCormic, T.C., 1985. Frequency analysis of low flows. Nordic Hydrology 16, 105–128.

Lung, W.S., Rapaport, R.A., Franco, A.C., 1990. Predicting concentrations of consumer product chemicals in estuaries. Environ. Toxicol. Chem. 9 (9), 1127–1136.

Lvovich, M.I., 1972. Hydrologic budget of continents and estimate of the balance of global fresh water resources. Sov. Hydrol. 4.

Male, J.W., Ogawa, H., 1982. Low flows of Massachusetts streams. Water Resource Research Center, Amherst Publication 125, 160 pp.

Male, J.W., Ogawa, H., 1984. Tradeoffs in water quality management. J. Water Resour. Plan. Manag., ASCE 110 (4), 434–444.

Mallory, S.J.L., McKenzie, R.S., 1993. Water resources modelling of flow diversions. Proceedings of the Sixth South African National Hydrology Symposium, Pietermarietzburg, South Africa, vol. 1, pp. 429–436.

Martin, G.N., 1973. Characterisation of simple exponential baseflow recessions. J. Hydrol. NZ 12 (1), 57–62.

Matalas, N.C., 1963. Probability distribution of low flows. USGS Professional Paper 434-A, USGS, Washington, DC.

McKenzie, R.S., Roth, C., 1994. Orange River Losses Study — Phase 1. BKS Report N P9762-02 Pretoria, South Africa.

McMahon, T.A., 1976. Low Flow analyses of streams: details ofcomputational procedures and annotated bibliography. University, Department of Civil Engineering, , Clayton (Australia),Monash Research Report No. 5/1976, 60 pp.

McMahon, T.A., Mein, R.G., 1986. River and reservoir yield. Water Resource Publication, CO, 368 pp.

Meier, K.B., Brodie, J.R., Schulze, R.E., Smithers, J.C., Mnguni, D., 1997. Modelling the impacts of riparian zone alien vegetationon catchment water resources using the ACRU model. Proceedings of the Eighth South African National Hydrology Symposium, Pretoria, South Africa, 13 pp.

Mhango, D.H.Z., Joy, D.M., 1998. Low flow characteristics and assessment of domestic water abstraction permits in Malawi.Proceedings Conference on Hydrology in a Changing Environment, Exeter, UK.

Midgley, D.C., Pitman, W.V., Middleton, B.J., 1994. Surface water resources of South Africa 1990. Water Research Commission Report No 298/5.1/94, Pretoria, South Africa.

Mimikou, M., Kaemaki, S. 1985. Regionalization of flow duration characteristics, J. Hydrol., 82, 77–91.

Mngodo, R.J., 1997. Flow duration characteristics of Southern African rivers. Acta Hydrotechnica 15/18 — Proceedings of oral presentations, FRIEND'97, pp. 49–63.

Moore, R.D., 1997. Storage-outflow modelling of streamflow recessions, with application to a shallow-soil forested catchment. J. Hydrol. 198, 261–270.

Morrisey, D.J., Randall, A.D., Williams, J.H., 1988. Upland runoff as a major source of recharge to stratified drift in the glaciated northeast. AWRA Symposium on Monitoring, Modeling, and Mediating Water Quality. May 1987, Syracuse, NY. AWRA Monograph Series No. 11, pp. 17–36.

Mosley, M. P., McKerchar, A. I., 1993. "Streamflow," Handbook of Hydrology. D. R. Maidment ed., McGraw-Hill, New York.

Moye, L.A., Kapadia, A.S., 1995. Prediction of drought length extreme order statistics using run theory. J. Hydrol. 169, 95–110.

Musiake, K., Takahasi, Y., Ando, Y., 1984. Statistical analysis on effects of basin geology on river flow regime in mountainous areas of Japan. Proc. Fourth Cong. Asian & Pacific Reg. Div. Int. Assoc. Hydraul. Res., Bangkok, APD-IAHR/Asian Institute Technology, vol. 2, pp. 1141–1150.

Muzik, I., 1986. The cause and effect of low flows on the Peace- Athabasca delta. Proceedings of the 16th Canadian Hydrology Symposium: Drought — the Impending Crisis? Regina, Saskatchewan, pp. 147–158.

Nathan, R.J., McMahon, T.A., 1990a. Evaluation of automated techniques for base flow and recession analyses. Water Resour. Res. 26, 1465–1473.

Nathan, R.J., McMahon, T.A., 1990b. Practical aspects of low-flow frequency analysis. Water Resour. Res. 26, 2135–2141.

Nikic, Z., Radonja, P., 2009: Modelling the influence of hydrogeological parameters on low flow in hilly and mountainous regions of Serbia, Hydrological Sciences, 54 (3).

Owen, M., 1991. Groundwater abstraction and river flows. J. Inst.Water Environ. Manag. 5 (6), 697–702.

Paulson, C.L., Sanders, T.G., 1987. Evaluation of design flow criteria for effluent discharge permits in Colorado. Colorado Water Resources Research Institute, Fort Collins, Completion Report No. 147, 312 pp.

Pearson, C.P., 1995. Regional frequency analysis of low flows in New Zealand rivers. J. Hydrol. NZ 30 (2), 53–64.

Pegram, G.G.S., Salas, J.D., Boes, D.C., Yevjevich, V., 1980. Stochastic properties of water storage. Hydrology Paper 100, Colorado State University Fort Collins, Colorado, 48 pp.

Petras, I., 1986. An approach to the mathematical expression of recession curves. Water SA 12 (3), 145–150.

Pettyjohn, W.A., Henning, R., 1979. Preliminary estimate of ground-water recharge rates, related streamflow and water quality in Ohio. Ohio State University Water Resources Center Project Completion Report No. 552, 323 pp.

Pirt, J., 1989. Estimation of river flows with particular reference to natural and artificial components. FRIENDS in Hydrology, IAHS Publication No. 187, pp. 361–371.

Pirt, J., Simpson, M., 1983. The Estimation of river flows. Severn Trent Water Authority, UK, 41 pp., append.

Pitman, W.V., 1993. Simulation of run-of-river schemes using monthly data. Proceedings of the Sixth South African Hydrology Symposium, Pietermarietzburg, South Africa, vol. 1, pp. 445–452.

Polarski, M., 1989. Fitting distributions to annual minimum flows of different durations. IAHS Publication No. 187, pp. 97–104.

Ponce, V.M., Lindquist, D.S., 1990. Management of baseflow augmentation: a review. Water Resour. Bull. 26 (2), 259–268.

Potter, K.W., Rice, J.M., 1987. Estimating baseflow volume and its uncertainty. Water Resour. Bull. 23 (2), 233–238.

Prakash, A., 1981. Statistical determination of design low flows. J. Hydrol. 51, 109–118.

Quimpo, R. G., Alejandrino, A. A., McNally, T. A., 1983. Regionalised flow duration curves for Philippines, J. Water Res. PL—ASCE, 109(4), 320–330.

Radczuk, L., Szarska, O., 1989. Use of the flow recession curve for the estimation of conditions of river supply by underground water. FRIENDS in Hydrology, IAHS Publication No. 187, pp. 67–74.

Refsgaard, J.C., Hansen, E., 1976. Economic value of low flow data. Nordic Hydrol. 7 (1), 57–72.

Rianna, M., Russo, F., Napolitano, F., Stochastic index model for intermittent 1 regimes: from preliminary analysis to regionalisation. Natural hazard and Heart System Sciences, In press.

Ries III, K.G., 1994. Development and application of generalizedleast-squares regression models to estimate low-flow. USGS Water-Resources Investigations Report 94-4155, 33 pp.

Riggs, H.C., 1972. Low flow investigations. Techniques of Water Resources Investigations of the USGS, Book 4, Analysis and Interpretation, Washington DC, 18 pp.

Riggs, H.C., 1976. Effects of man on low flows. Proceedings of theConference on Environment, Aspects Irrigation and Drainage,University of Ottawa, American Society of Civil Engineers, New York, pp. 306–314.

Riggs,H.C., 1985. StreamFlow Characteristics. Elsevier,Amsterdam.

Risle, J.C., 1994. Estimating the magnitude and frequency of low flows of streams in Massachusetts. USGS Water-Resources Investigations Report 94-4100, 34 pp.

Rogers, J.D., Armbruster, J.T., 1990. Low flows and hydrologic droughts. Surface Water Hydrology, Geological Society of America, Boulder, CO, pp. 121–130.

Rowe, P.B., 1963. Streamflow increases after removing woodlandriparian vegetation from a southern California watershed. For. J. 61, 365–369.

Russell, D.S.O., 1992. Estimating flows from limited data. Can. J. Civil Engng 19 (1), 51–58.

Sakovich, V.M., 1990. The representation of minimum runoff series in Karelia and North-Western region. Trans. Hydrolmeteorol. Center USSR 308, 129–136.

Saville, Thorndyke, Watson, J. D., 1933, An investigation of flow duration characteristics of North Carolina streams: Am. Geophys., Union Trans., p. 406-425.

Schulze, R.E., 1979. Hydrology and water resources of the Drakensberg. Natal Town and Regional Planning Reports, Durban, South Africa, vol. 42.

Searcy, J. C.: Manual of hydrology 2, Low flow techniques, flow duration curves, U.S. Geological Survey, Water Supply Paper, 1542–A., 1959.

Sefe, F.T.K., 1988. Some aspects of the hydrology of one-day annual minimum low flows of Malawian rivers. Hydrol. Process. 2, 75–91.

Sen, Z., 1980b. Critical drought analysis of periodic-stochastic processes. J. Hydrol. 46, 251–263.

Sherrard, J.J., Erskine, W.D., 1991. Complex response of a sand-bed stream to upstream impoundment. Regul. Rivers Res. Manag. 6, 53–70.

Simmons, D.L., Reynolds, R.J., 1982. Effects of urbanization on base flow of selected south-shore streams, Long Island New York. Water Resour. Bull. 18, 797–805.

Singh, R.D., Mishra, S.K., Chowdhary, H., 2001. Regional flow-duration models for large number of ungauged Himalayan catchments formplanning microhydro projects. J Hydrol Eng;6(4):310–6.

Singh, K.P., 1971. Model flow duration and stream flow variability. Water Resour. Res., AGU, 7(4) : 1031--1036.

Singh, V.P., 1987. Derivation of extreme value (EV) type III distribution for low flows using entropy. Hydrol. Sci. J. 32 (4), 521–533.

Sittner, W.T., Schauss, C.E., Monro, J.C., 1969. Continuous hydrograph synthesis with an API-type hydrologic model. Water Resour. Res 5 (5), 1007–1022.

Sloto, R.A., Crouse, M.Y., 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis. US Geologica Survey Water-Resources Investigations Report 96-4040, 46 pp.

Smakhtin, V. Y., 2001. Low-flow hydrology: a review. Journal of Hydrology. v240. 147-186.

Smakhtin, V. Y., Hughes, D. A., Creuse-Naudine, E., 1997. Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, Hydrol. Sci. J., 42(6), 919–936.

Smakhtin, V.Y., Hughes, D.A., 1993. The application of automated procedures for low-flow analysis of South African rivers. Proceedings of the Sixth South African National Hydrology Symposium, Pietermaritzburg, South Africa, vol. 1, pp. 329–336.

Smakhtin, V. Y., Toulouse, M., 1998. Relationships between lowflow characteristics of South African streams, Water South Afr., 24(2), 107– 112.

Smakhtin, V.Y., Watkins, D.A., Hughes, D.A., 1995. Preliminary analysis of low-flow characteristics of South African rivers. Water SA 21 (3), 201–210.

Smakhtin V.Y., Watkins, D.A., 1997. Low-flow estimation in South Africa. Water Research Commission Report No. 494/1/97, Pretoria, South Africa.

Smakhtin, V.Y., Watkins, D.A., Hughes, D.A., Sami, K., Smakhtina, O.Y., 1998a. Methods of catchment-wide assessment of daily low-flow regimes in South Africa. Water SA 24 (3), 173–185.

Smith, R.W., 1981. Rock type and minimum 7-day/10-year flow in Virginia streams. Virginia Water Resource Research Center, Virginia Polytechnology Institute and State University, Blacksburg, Bulletin, vol. 116, 43 pp.

Smith, R.E., Bosch, J.M., 1989. A description of the Westfalia catchment experiment to determine the effect on water yield of clearing the riparian zone and converting an indigenous forest to a Eucalyptus plantation. S. Afr. For. J. 151, 26–31.

Sokolov, B.L., 1974. Method for reducing single discharge measurements to the average discharge of the low-flow period. Sov. Hydrol. (Selected Papers) 4, 235–241.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. v(Ed.), Handbook of Hydrology. McGraw-Hill, New York(chap. 18).

Strupczewski, W. G., Weglarczyk, S., and Singh, V. P., 2003. Impulse response of the kinematic 604 diffusion model as a probability distribution of hydrologic samples with zero values, J. 605 Hydrol., 270, 328–35.

Studley, S.E., 2001. Estimated flow-duration curves for selected ungare sites in Kansas. US Geological Survey. Water-resources investigations report 01-4142.

Subcommittee on Hydrology, 1966. Methods of Flow Frequency Analysis, Inter-Agency Committee

607 on Water Resources, Notes on hydrological act, Bulletin 13, Washington D.C..

Sugiyama, H., 1996. Analysis and extraction of low flow recession characteristics. Water Resour. Bull. 32 (3), 491–497.

Sugiyama, H., Vudhivanich, V., Whitaker, A. C., Lorsirirat, K., 2003. Stochastic flow duration curves for evaluation of flow regimes of rivers, J. Am. Water Resour. Assoc., 39(1), 47–58.

Swank, W.T., Swift, J.R., Douglas, J.E., 1988. Streamflow changes associated with forest cutting, species conversion and natural disturbance. In: Swank, W.T., Crossley, D.A. (Eds.), Forest Hydrology and Ecology at Coweeta. Ecological Studies 66. Springer, New York, pp. 297–312.

Tallaksen, L.M., Van Lanen, H.A.J., 2004: Hydrological Drought - Processes and estimation methods for streamflow and groundwater. Development in Water Science, 48, Elsevier Science B. V., Amsterdam.

Tallaksen, L.M., 1995. A review of baseflow recession analysis. J.Hydrol. 165, 349–370.

Tasker,G.D., 1987. A comparison of methods for estimating low flow characteristics of streams.Water Resour. Bull. 23 (6), 1077–1083.

Tate, E.L., Freeman, S.N., 2000. Three modelling approaches for seasonal streamflow droughts in southern Africa: the use of censored data. Hydrol. Sci. J. 45 (1), 27–42.

Tlalka, A., Tlalka, J., 1987. Low discharge periods of rivers in the Upper Vistula river basin, Poland. Water Resour. Bull. 23 (2), 227–232.

Toebes, C., Strang, D.D., 1964. On recession curves. 1. Recession equations. J. Hydrol. NZ 3 (2), 2–15.

Tsakiris, G,Tigkas, D., Spiliotis, M., 2006. assessment of interconnection between two adjacent watersheds using deterministic and fuzzy approaches, European Water 15/16: 15-22.

United Nations Educational, Scientific and Cultural Organization (UNESCO), 1997. *Southern Africa - FRIEND*). Technical Documents in Hydrology No. 15, UNESCO, Paris.

Vasak, L., 1977. Low flow studies. A literature survey. Free University,Amsterdam, Netherlands.

Velz, C.J., Gannon, J.J., 1953. Low flow characteristics of streams.OH State Univ. Studies Engng Survey XXII (4), 138–157.

Velz, C.J., Gannon, J.J., 1960. Drought flow characteristics of Michigan streams. Water Resources Commission of Michigan, Lansing, Michigan.

Vladimirov, A.M., 1976. Stok rek v malovodnyj period goda (River flow during a dry weather period). Gidrometeoizdat, Leningrad, USSR.

Vogel, R.M., Fennessey, N.M., 1994. Flow duration curves: new interpretation and confidence intervals, J. Water Res. PL—ASCE, 120, 485-504.

Vogel, R.M., Fennessey, N.M., 1995. Flow duration curves. II. A review of applications in water resource planning. Water Resour. Bull. 31 (6), 1029–1039.

Vogel, R.M., Kroll, C.N., 1989. Low-flow frequency analysis using probability-plot correlation coefficients. J. Water Res. Plan. Manag. (ASCE) 115 (3), 338–357.

Vogel, R.M., Wilson, I., 1996. Probability distribution of annual maximum, mean, and minimum streamflows in the United States. J. Hydrol. Engng 1 (2), 69–76.

Walker, K.F., Thoms, M.C., 1993. Environmental Effects of Flow Regulation on the Lower River Murray, Australia. Regul. Rivers Res. Manag. 8 (1/2), 103–119.

Wang, C.Y., Singh, R. P., 1995. Frequency estimation for hydrological samples with zero data, 613 J. Water Res. PL—ASCE, 121(1), 98–108.

Ward J.H., 1963. Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc., 58(301), 236–244.

Warner, R.F., 1984. Man's impact on Australian drainage systems. Aust. Geographer 16, 133–141.

Warnick, C.C., 1984. Hydropower Engineering. Prentice-Hall, Englewood-Cliffs, NJ (pp. 57–73).

Waylen, P.R., Woo, M.K., 1987. Annual low flows generated by mixed processes. Hydrol. Sci. J. 32 (3), 371–383.

Weisman, R.N., 1978. Characterizing low flows using threshold levels. J. Hydraul. Div., ASCE 104 (IR2), 231–235.

White, E.L., 1977. Sustained flow in small Appalachian watersheds. J. Hydrol. 32, 71–86.

Wijayaratne, L.H., Golub, E., 1991. Multiyear drought simulation. Water Resour. Bull. 27 (3), 387–395.

Wilber, D.H., Tighe, R.E., O'Neil, L.J., 1996. Associations between changes in agriculture and hydrology in the Cache River basin, Arkansas, USA. Wetlands 16 (3), 366–378.

Wilgen, B.W., van Little, P.R., Chapman, R.A., Gorgens, A.H.M., Willems, T., Marais, C., 1997. The sustainable development water resources: History, financial costs and benefits of alien water resources: History, financial costs and benefits of alien plant control programmes. S Afr. J. Sci. 93, 404–411.

Wittenberg, H., 1994. Nonlinear analysis of flow recession curves. FRIEND — Flow Regimes from International Experimental and Network Data. IAHS Publication No. 221, pp. 61–67.
WMO: World Meteorological Organization, 1974. International Glossary of Hydrology, WMO, Geneva.

WMO: World Meteorological Organization, 2008. Manual on low flow estimation and prediction, WMO, Geneva.

Woo, M. K., Wu, K., 1989. Fitting annual floods with zero flows, Can. Water Resour. J., 14 (2), 617 10–16.

Wu, L.D., Bian, G. R., et al., 1979. Probability Theory, Education Press, Beijing,.

Yevjevich, V., 1967. An objective approach to definitions and investigations of continental hydrologic droughts. Colorado State University, Fort Collins, Hydrology Paper 23, 18 pp.

Yu, P. S., Yang, T. C., Wang, Y. C., 2002. Uncertainty analysis of regional flow duration curves, J. Water Res. PL—ASCE, 128(6), 424–430.

Zelenhasic, E., Salvai, A., 1987. A method of streamflow drought analysis. Water Resour. Res. 23 (1), 156–168.