POLITECNICO DI TORINO

Facoltà di Ingegneria
Corso di Laurea Magistrale in Ingegneria Civile

Tesi di Laurea Magistrale

Analysis of Event Runoff Coefficients for Flood Estimation in the UK

Relatori:

Prof. Pierluigi Claps

Prof. Ross Woods

Candidato:

Cristina Floris

Dicembre 2017

To my father

I would like to express my gratitude to my main supervisor Prof. Pierluigi Claps from Polytechnic of Turin, who allowed me to realize my Final Thesis Project at the University of Bristol and gave me all the support necessary for this formative experience.

Furthermore, I especially would like to thank Dr. Ross Woods, Reader in the Civil Engineering Department at the University of Bristol. He introduced me to my first research experience, being always helpful and supportive, and he gave me the opportunity to increase my knowledge in depth.

I would like to thank all the PhD students from the Water Management and Research Group at the University of Bristol for sharing this experience with me and for the incredible support, even in tough moments. In particular, a special mention goes to Wouter Knoben, always helpful, patient and available to listen to thoughts and doubts about the project.

The deepest gratitude goes to my mother. I owe her everything, not only for the support she gave me during my educational path, but mostly for the person I have become. She always pushed me to improve myself and encouraged me to leave the nest, flying first to Turin, then to Bristol, even if the situation was hard for both of us. She represents a model for me.

I dedicate this thesis to my father. I would not have become the person I am and I would not have reached this goal without all his precious teachings. I do believe that he would be proud of me.

Last but not least I would like to thank all my loved ones, my family and my friends, who supported me during all these years of study. If I managed to achieve this important result is also thanks to them.

Cristina Floris

Summary

List of Figures

List of Tables

1. Introduction

Floods are one of the natural and most destructive hazards in our earth. They occur when water overflows or inundates land that is normally dry. Usually it happens in periods of excessive rain in which the rivers or streams overflow their banks, or dams and levees are overwhelmed by the flow or breached. Also the timing of occurrence can be different: some floods may take days to develop, permitting people to evacuate, other kinds, called flash floods, are instead so fast that water covers everything in its path creating huge damages to property and foremost loss of people lives.

1.1. Flood estimation and its importance

The social and economic impacts that a destructive flood may create are so huge that it is necessary to deal with these hazards and study the assessment of flood occurrence. It is hence important to introduce the related concepts of flood risk assessment and flood frequency estimation. The first one focuses on the assessment of the probability of a flood occurrence and the magnitude of its consequences. The second one bases on the estimation of the peak river flow considering a certain return period, or rather the average interval between floods of a certain magnitude at a site.

Apart from the generic hazards mentioned before, flood estimation is fundamental for the hydrologic design that is the key and starting point of any kind of water resources planning and management. Usually the different problems the hydrological design focuses on are divided in two different categories.

The first one is the *water control* that includes drainage, flood control and some environmental aspects as pollution abatement, insect, sediment and salinity control. On the other hand there is the *water use* and management related to domestic and industrial water supply, hydropower generation, fish and wildlife improvements. The hydrological study focuses on different aspects considering the variety of problems. For example in the first category are projects such as levees, or river hydraulics installation whose design bases on extreme events of short duration, as the peak of discharge during a flood or the minimum flow over a dry period. While for projects of reservoirs or dams it is necessary to have knowledge of the whole flow hydrograph over a period of minimum a year, considering the different purposes of the work.

1.2. Statement of the problem

The United Kingdom presents a relatively extensive flood data, and a kind of climate that can be considered benign in relation with other countries. In spite of this, there are substantial flood damages in the UK, as for example in floods occurred in 2007, 2013/14 and 2015. However, the strong irregularity

in time in which natural phenomena and floods occur makes difficult to do a proper flood frequency estimation except at sites with a long flood records available.

It is possible to underline that there is a seasonal effect on the occurrence of floods. Usually in the UK, except in few cases, floods occur in the winter season, the period in which the wetness of the soil tends to be high and the soil moisture reaches the condition of saturation, so that the water gathers on the surface of the soil. In summer instead, soils tend to be dry so that precipitation is absorbed, suppressing the formation of floods. This result has been confirmed also by the study of Zavatteri (2012) who highlighted that for the 74% catchments the analyzed soil moisture deficit influenced the flood production. The alternation of wetter and drier periods leads in the long period to have years richer or poorer in terms of flood occurrence, and huge problems in a reliable flood estimation. In contrast, for the USA, Berghuijs et al. (2015) observed that other factors (e.g., snowmelt and soil moisture) other than just precipitation control the magnitude of floods across the majority of Unites States, noticing a large disparity between the dates of maximum precipitation events and the dates of flooding.

Another aspect that likely interferes with problems on flood estimation, not only in UK, but in all Europe is the climate change. Analysis of rainfall estimated by climate models highlighted changes in the extreme precipitations all over Europe, due to the increasing of greenhouse gas conditions [Giorgi et al,. 2001]; different studies report that climate change is bringing wetter winters in Northern Europe and drier summers in southern Europe [Hulme et al., 2002]; a result from the Regional Climate Model (RCMs) by Giorgi et al., (2001) underlines an increase in the frequency of precipitation events exceeding 30 mm day^{-1}. Bloschl et al. (2017), analyzed a large dataset of flood observations across the Europe to assess the shifts in time of river floods during the past five decades, with the result of total shifts that goes from -65 to +45 days in 50 years. Considering that floods are the result of the seasonal interaction of rainfalls, snow processes and soil moisture, they observed four distinguished regions with distinct drivers: the north-eastern Europe showed earlier snowmelt; the North Sea, later winter storms; the western Europe along the Atlantic coasts, instead, earlier soil moisture maximum; parts of Mediterranean coast, stronger Atlantic influence in winter [Bloschl, 2017]. All these aspects mentioned above, together with modelling errors, probably influence the uncertainty in flooding estimation studies, hence, it is important to take into account all of them with the purpose of decreasing uncertainty and, in turn, the destructive effects flooding causes on our earth.

1.3. UK flood estimation methods

The usual methods in United Kingdom for floods estimation are reported in the Flood Estimation Handbook. This publication is the result of a 5-years (1994-1999) research programme at the Institute of Hydrology in Wallingford, Oxfordshire and it has been written to develop and implement some

procedures for rainfall and flood frequency estimation in the UK. The FEH (Flood Estimation Handbook) method based on the Flood Studies Report (FSR) published by the Natural Environment Research Council (NERC, 1975). It divided United Kingdom in 11 geographical regions, taking into account Ireland, and for each one of them it calculated a flood frequency curve that linked the flood magnitude (m^3/s) to exceedance probability by using the product of a regional dimensionless growth factor and an estimate of the index flood, intended as the mean annual maximum flood, obtained from observations or from regression models on catchments characteristics. The Institute of Hydrology added to this work some new features as the introduction of the method of L-moments and the region of influence approach creating the proper FEH methodology currently used.

Within the 5 published volumes , the FEH actually proposes for the flood estimation two distinguished method categories: statistical analysis of flood peak data and rainfall-runoff methods.

1.3.1. Flood index method

The statistical approach, based on the FSR methodology, focuses on the construction of the flood frequency curve Q_T as the product of the index flood QMED and the growth curve x_T. The index flood QMED for gauged catchments can be estimated as the median of annual maxima (AMAX) if the length of data records is larger than 13 years, from peaks-over-threshold with data between 2 and 13 years length, and for catchments with less than 2 years of data or ungauged catchments, using regression models on catchments descriptor or data transfer from a pooling group of similar catchments. The value of QMED obtained with the best estimation may be adjusted taking into account the climatic variation. The growth curve, instead, depends on the return period and its estimation varies whether the catchment is gauged or not. In the first case the assessment is done by using data records, and usually it is better that the record length exceeds two times the return period of primary interest. Otherwise it is necessary to pool data from groups of catchments. Once that QMED and the growth curve are estimated it is possible to derive the flood frequency curve, adjusting for catchments urbanization and finally constructing a design hydrograph.

1.3.2. FSR/FEH rainfall-runoff method

The FSR runoff rainfall method bases instead on a unit hydrograph or a losses model by considering three parameters: the unit hydrograph time to peak TP, which controls the temporal characteristics of the runoff response to rainfall; the standard percentage runoff SPR that influences the volumetric characteristics of the runoff response; the baseflow BF that represents the river flow before a flood event. If possible, the model parameters are derived from observed rainfall and runoff records, otherwise they may be estimated from physical and climatic descriptors of the catchment. This method has actually been described in its simple form. Currently a modified version called revitalized FSR/FEH rainfall-runoff method is used, based on the Revitalised Flood Hydrograph (ReFH) rainfall-runoff

model. It adds to the previous model some features such as the information of soil moisture content or the interaction between direct runoff and baseflow, that lead to a more realistic representation of flood hydrology. A simple scheme of the revitalized FSR/FEH rainfall-runoff method is reported in Figure 1. It is possible to notice the connections between the three components together with the input variables and model parameters. When a flood event is simulated, the loss model is used to estimate the fraction of the total rainfall volume turned into direct runoff. The direct runoff is then routed to the catchment outlet using the unit hydrograph convolution in the routing model and, finally, the baseflow is added to the direct runoff to obtain total runoff [Kjeldsen, 2007].

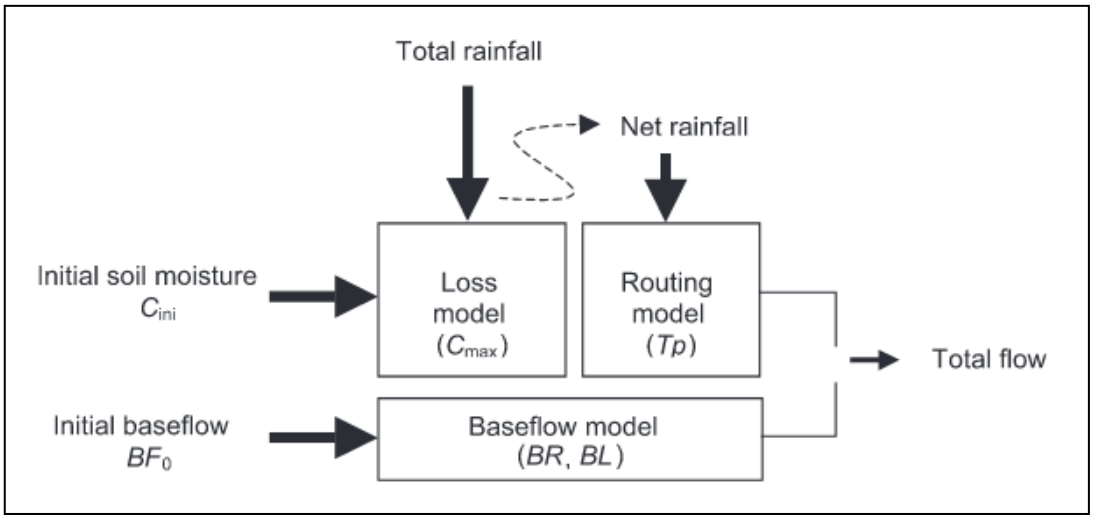

Figure 1: Schematic representation of the ReFH model

1.3.3. Grid to Grid model

Another rainfall-runoff model, not mentioned in the FEH work but proposed by Bell et others (2007), is the Grid to Grid method (G2G), useful for different purposes such as floods estimation or problems related to water resources. This method, still in development, was born with the intent to include the environmental change on hydrologically sensitive systems. It is a high-resolution flow routing and spatially-distributed runoff production model that is able to translate climate model estimates of current or future scenarios of rainfall and potential evaporation (PE) into estimated river flow at daily or sub-daily time-step [Bell et al., 2007]. The name" Grid to Grid Model" is due to the configuration of the model based on grids of different coverage and resolution, allowing themselves to be coupled directly with climate models, land-surface schemes and oceans models. In particular the study among UK based on a 1 km grid resolution. A schematic of this model is presented in Figure 2. The routing model is based on the simple discrete approximation of the kinematic wave equation with lateral inflow, for land and river flows both. The runoff production model, instead, is based on the catchment-based CEH grid

Model by Bell and Moore (1998) that assesses the direct runoff considering the storage capacity of each grid squares, even though maintaining a water balance for each one.

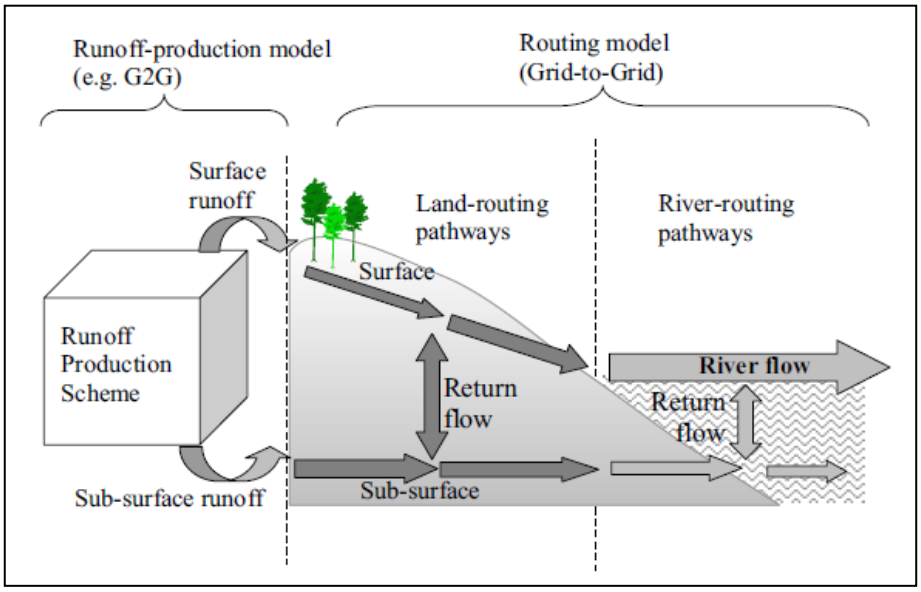

Figure 2: Schematic of the Grid-to-Grid model structure.

1.4. Uncertainty of flood estimation in UK

Among these approaches, in this work the attention is given to the statistical one, to its uncertainty related to ungauged catchments and in particular to the QMED value, which contains the largest error in its calculation. This has been done taking into account the work done by Kjeldsen (2014).

The first step to be done is to define how the uncertainty is measured. It is common to use the variance or the square root of the variance (standard deviation) to quantify the uncertainty of a random variable such as design flood estimates [Kjeldsen, 2014]. When a random variable, called x, is assumed normally distributed, the standard deviation s_x is often used to provide the 68% or 95% confidence intervals as:

$$x \pm s_x \qquad 68\%\ confidence\ interval$$

$$x \pm 2s_x \qquad 95\%\ confidence\ interval$$

Often in hydrology it is common to use the log-transformed variable normally distributed, instead of the random variable itself. Hence, the random variable $x = \ln q$ can be transformed using the exponential function $q = \exp(x)$. It is possible to define the confidence intervals by applying the exponential function as showed in the following:

$$[e^x \cdot e^{-s_x}\ ;\ e^x \cdot e^{s_x}] = \left[\frac{q}{fse}; q \cdot fse\right] \quad 68\%\ confidence\ interval \qquad \text{[Eq. 1]}$$

$$[e^x \cdot e^{-2s_x}\ ;\ e^x \cdot e^{2s_x}] = \left[\frac{q}{fse^2}; q \cdot fse^2\right] \quad 95\%\ confidence\ interval \qquad \text{[Eq. 2]}$$

The fse variable presented in the second term of the equation is the factorial standard error, defined using the exponential function to the estimate of the standard deviation of the log-transformed variable as:

$$fse = e^{s_x} \qquad \text{[Eq. 3]}$$

It represents the amplitude of the confidence interval related to the QMED value in flood estimation with FEH procedure.

The second step bases on the studying of uncertainty referring to the assessed QMED. In ungauged catchments the calculation of this value can be obtained in two different ways. The first one, reported in Equation 4, occurs when QMED is estimated only from catchments descriptors[1].

$$\ln QMED = 2.1170 + 0.8510 \cdot \ln[AREA] + 1.8734\left[\frac{1000}{SAAR}\right] + 3.4451 \cdot \ln[FARL] - 3.080\ BFIHOST^2 \qquad \text{[Eq. 4]}$$

The second one, when the assessment by catchments descriptor is adjusted by using data transfer from a nearby gauged donor site (Equation 5), where: s and g stand for gauged and ungauged sites, cds is the catchement descriptor estimate at the gauged and ungauged site; obs is the observed value at the gauged site; adj is the adjusted value at the subject site; α is a function of geographical distance.

$$QMED_{s,adj} = QMED_{s,cds}\left(\frac{QMED_{g,obs}}{QMED_{g,cds}}\right)^{\alpha} \qquad \text{[Eq. 5]}$$

This latter procedure therefore bases on the use of local data to compensate for the inability of the model represented in Equation 4 to contemplate complex catchment hydrology by considering only simplistic and lumped catchments descriptors.

[1] AREA is the catchment extension in km^2; SAAR is the average annual rainfall in the standard period (1961-1990), in mm; FARL is the Flood Attenuation by reservoirs and Lakes index that represents the degree of flood attenuation attributable to reservoirs and lakes in the catchment above a gauging station; BFIHOST is the baseflow index derived using 29-class Hydrology of Soil Types (HOST).

In the work made by Kjeldsen (2014) the uncertainty of QMED has been studied for return period of 2, 5, 30 and 100 years and considering median annual maxima peak flow data (AMAX) of 715 rural catchments.

The factorial standard error has been calculated in this case using the log-transformed residuals of the T-years estimate:

$$e = \ln \widehat{Q_T} - \ln Q_T \qquad \text{[Eq. 6]}$$

In which Q_T is the design peak flow value with a T-year return period obtained by fitting a Generalized Logistic (GLO) distribution directly to the at-sited data, while $\widehat{Q_T}$ is the corresponding estimate using the FEH procedures, as if the site was ungauged, using the two mentioned possible ways for ungauged catchments.

The results for factorial standard errors for different return period is reported in Table 1.

Return period	Fse (regression only)	Fse (regression + donor)
2	1.47	1.42
5	1.48	1.43
30	1.52	1.47
100	1.54	1.50

Table 1: Factorial standard errors (fse) for prediction at ungauged sites with and without using donor transfer.

The first clear result that it is possible to underline is that the prediction of uncertainty using donor adjustments is more reliable in comparison with estimation done considering only catchments descriptors method.

Furthermore, it is clear that uncertainty increases with higher return period. Taking into account the return period of 100 year flood, the 95% confidence interval obtained using donor sites will be of the order of minus 55% and plus 125% derived from the fse labeled. This means that if we have an estimated discharge value of 30 m^3/s, with a confidence level of 95%, the value can be between 13.3 and 65.5 m^3/s, revealing a considerable uncertainty and practical problem on any kind of hydrological design [G. Giani, 2017].

1.5. Uncertainty of flood estimation in other countries.

1.5.1. Italy

The problem of flood estimation uncertainty does not affect the United Kingdom only. Italy for example, although it lacks of a common uniform methodology for whole regions of the country,

presents a similar flood index method for the regions of its north-west part, Piedmont, Aosta Valley, western Lombardy and western Apennine Emilia. The Flood index method is, also in Italy, one of the most used for regional frequency analysis. It was proposed for the first time in 1960 by Dalrymple and bases on the estimation of the design peak flow for a certain measure section through the product between a flood index and a growth curve. There are different methods for the evaluation of the two quantities mentioned, whose uncertainty in the estimation is given considering the standard deviation. The flood index may be assessed by using historical observation for gauged basins, otherwise for ungauged catchments it is possible to adopt multi-regression models, or the rational method. The growth curve instead, is usually obtained by adapting a statistical distribution to the data of a pooling set of basins having similar characteristics and estimating its parameters, in order to have a different growth curves for each region. The final result of the approach, as in UK, permits to build the flood frequency curve. It is interesting trying to make a comparison between flood index method and its uncertainty in two different European countries, in particular giving the attention to the QMED value. With this purpose, a study by Laio, Claps et al., (2011) is proposed, which uses L-moments and their dimensionless ratios as statistical variables to be transferred to the ungauged sites avoiding the limitations of choosing a priori distribution to describe the sample data, and the difficulty to assess a reliable and stable configuration of regions having the same characteristics. In particular, they chooses L-moment of order one (the mean), the coefficient of L-variation (L_{CV}) and the L-skewness (L_{CA}) that allows to reconstruct the flood frequency curve, but also to recreate a flood index framework in which the mean is the scale factor (QMED) and the other two L-moments may be intended as descriptors of the dimensionless growth curve.

The results of the regression model used to estimate the QMED are interesting and show that the suitable model is the lnQMED, depending on the catchment area, on the mean annual precipitation, on a permeability index and on the coefficient of the Intensity-Duration curve of the average of maximum rainfall. In term of uncertainty of the estimation, the value of standard deviation reported for lnQMED is 0.340. It means that in term of the Factorial Standard Error proposed by Kjeldsen (2014), the uncertainty is given by a factor of 1.4, slightly lower than for UK results, whose fse is equal to 1.47. Even if this study is conducted only on 70 catchments, it may be considered as an important step, to be extended to a wider database, and maybe implemented also for UK.

1.5.2. USA

It is important also to expand the view about flood estimation and its uncertainty in the world. The attention has been focused on United States of America, in particular to the Western, Eastern and Central states of California, Pennsylvania and Kansas. The approaches used for estimating the frequency of flood peak discharge and flood hydrographs in ungauged sites may be divided in two groups: methods based on statistical regression analysis of data from gauged stations and methods based on

rainfall characteristics that use algorithms to convert rainfall excess to flood runoff. The interest in this report is given to statistical methods, that have been developed by U.S. Geological Survey (USGS) with the Federal Highway Administration and the Federal management Agency for all the states.

In California, a regional regression analysis has been used to develop a set of equations for estimating magnitude and frequency of floods in ungauges sites. These relations have been obtained for each of the six regions in which California has been divided and permit to relate the annual exceedance probability flow (inverse of the return period) calculated from peak flow records to basin characteristics of the associated drainage catchments. The discharge for a selected return period is given by the product between a regression coefficient with basin characteristics (drainage area, mean annual precipitation, mean basin elevation) raised different regression coefficients. Regression equations, being statistical models, must be interpreted taking into account model errors that represent the differences between predicted and observed values of discharge. Uncertainty in estimation is given considering the standard deviation of prediction for each region. For a return period of 2 years the largest standard deviation occurs in the Central Coast, with a value of 0.477 and a corresponding factorial standard error of 1.61. The lowest in the North Coast with a standard deviation of 0.23 and a corresponding factorial standard error of 1.26. For a return period of 100 years instead, the largest standard deviation occurs in the Central Coast with a value of 0.24 and a factorial standard error of 1.27, the lowest standard deviation occurs in the South Coast with a value of 0.17 and a respectively factorial standard error of 1.18.

In Pennsylvania regression analysis for flood estimation bases on a multiple linear regression model from MathSoft. The country has been divided in two different regions due to the fact that flooding in the Northwestern part (Region A) appeared not to be related to flooding in the rest of the State (Region B). The regression equations proposed for different return period, in this case, depend on the drainage area and on the percentages of forest, urban cover development, presence of reservoir. The uncertainty, being the equation expressed in a logarithmic form, is represented by a residual standard error in Log units. The lowest return period considered in Pennsylvania analysis is ten years old. In this case Region A appears to have the largest residual standard error (0.18) that corresponds, in terms of factorial standard error, to a factor of 1.51. For a return period of 100 years instead, Region A has a residual standard error of 0.23, that corresponds to a factorial standard error of 1.7.

In Kansas, regression model for predicting magnitudes and frequency of peak streamflow for a certain return period were developed using USGS computer program Weighted-Multiple-Linear regression. The model is expressed in a logarithmic form, relating the dependent variable (peak streamflow for a selected annual exceedance probability) to basin characteristics by regression model coefficients, estimated through the software. In this State the uncertainty of the model is given in term of the average variance of prediction (AVP), for different return period and for both hydrologic regions Kansas is

divided to. For a return period of 2 years the largest AVP occurs in Region 2 (0.114), and it corresponds to a factorial standard error of 2.17. In the same region, for a return period of 100 years, AVP is the largest (0.067), with a factorial standard error of 1.81.

Analyzing three of the United states it is possible to notice that, if California appears to have more reliable flood estimation than UK, Kansas and Pennsylvania result to have the same in certain cases, in others worse grade of uncertainty. This comparison, however, refers only to regression analysis of statistical models, hence it is not possible to consider these results completely satisfying and absolute.

1.6. Statement of this work

The large uncertainty observed in the existing flood estimation methods by Kjeldsen et al. (2014) pushes us to investigate around new procedures in order to find a potential way to reduce the uncertainty in the UK.

Mainly the study turns around the concept of runoff coefficient which is defined as the portion of rainfall that becomes direct runoff during an event [R. Merz, G.Bloschl, L. Parajka, 2006]. Runoff coefficients are a key variable for design and engineering practice in general, leading to represent runoff generation in catchments. They are fundamental for describing basin response, giving the possibility to underline changes from event to event or from season to season, and are also really useful for comparison between catchments, to highlight how different landscapes, soils and morphology can "transform" rainfall into event-based runoff [T. Blume et al., 2007]. Furthermore, they can be used in event-based derived flood frequency models that permit to understand the flood frequency control in a climate or hydrological regime by analyzing rainfall frequencies. In hydrology research, the analysis of the controls of runoff coefficients is still an active research topic. Different studies in fact analyzes runoff coefficients in relation to different problems. Cerdan et al., (2004) analyzed 345 rainfall-runoff events in three catchments of different sizes in France to study scale effects in the runoff generation process finding a decrease of runoff coefficients as area increases. Naef (1993) analyzed the largest ten floods of 100 Swiss catchments stating that runoff coefficients should be treated as random variables, considering the complexity of catchments conditions. Mertz et al., (2006) estimated about 50000 event runoff coefficients in Austrian catchments, finding patterns that they considered due to climate variability in Austria.

This thesis, instead, focuses on the analysis of runoff coefficients to have a clear vision about runoff processes all around the United Kingdom. Increasing the knowledge about these processes may lead to a better calibration of the methodologies used for flood estimation, and reduced uncertainty in flood estimation. The precise aim of this work is the spatio-temporal analysis of runoff coefficients by using a data set of UK basins. The spatial and temporal patterns of runoff coefficients have been studied taking

into account rainfall data and catchments information about soil moisture, geomorphology and catchments dimensions. The main questions addressed in this research are: (i) how do time-averaged runoff coefficients vary across the UK?; (ii) How do within-year runoff coefficients vary across the UK? (iii) How do within-month runoff coefficients vary in the UK, in term of standard deviation?; (iv) How reliable can UK runoff coefficients be predicted for ungauged catchments?

2. Data

United Kingdom has got a widespread system of gauging stations across all the country. The National River Flow Archive (NFRA) is the primary archive that provides hydrometric data for the UK. It gives the access to daily, monthly and flood peak river flow data from over the 1500 gauging stations of the country. All this data are provided to the NRFA by "The Measuring Authorities" that have the responsibility for collection and processing hydrometric data. In particular, the principle Measuring Authorities are the Environment Agency (EA) in England, Natural Resources Wales (NRW) in Wales, the Scottish Environment Protection Agency (SEPA) in Scotland and the Rivers Agency (RA) in Northern Ireland. Also a variety of public and private bodies and research organizations contributes with additional flow measurements, such as the Centre of Ecology and Hydrology (CEH), which hosts a number of data centers and data resources on behalf of the UK research community.

This work is based on a dataset of rainfall and flow data related to 76 rural benchmark catchments widespread in all the United Kingdom, with the exception of the Northern Ireland that has been excluded from this analysis. In particular, the dataset consists in hourly discharge and rainfall data, on a temporal period of ten years, from 1999 until 2008.

The data used in this work have been provided by Gemma Coxon, Postdoctoral Research Associate at University of Bristol (School of Geographical Science), who requested them directly from the Environment Agency (EA) and Centre of Ecology and Hydrology (CEH). CEH website gives a free access only to daily data for about 1500 gauging stations, while the analysis of this work needed to be conducted on an hourly temporal scale. Furthermore, it has been necessary the use of catchments descriptors, related to the soil moisture, geomorphology and dimension of the catchments considered. Their values have been obtained for the analyzed basins from CEH website.

Some words need to be spent to talk about the catchments studied in this work. These catchments in fact, are called benchmark catchments because they are usually sited in rural parts of the country, free from human disturbances, river engineering and water abstractions. Thus, they may be used for understanding the climate-driven changes in river flow and long-term hydrological variability. Gauging stations placed in these basins belong to the UK Benchmark Network (UKBN), whose information are reported in the "User Guide: A network of near natural catchments" from the Centre of Ecology and Hydrology. It is an archive that manages 146 benchmark catchments characterized by natural flow regimes, good and consistent data quality, long record lengths, representative UK hydrology. Considering that the satisfaction of all these requirements is difficult, each benchmark catchment has been assigned a score based on their benchmark suitability at different flow regimes.

2.1. Rainfall data

Hourly rainfall data used in this work have been obtained with a two step process. The first step, based on the calculation of daily rainfall data, has been carried on by Keller et al. (2015) within a project of the Centre of Ecology and Hydrology (CEH-GEAR). The second step, instead, focused on the passage from daily rainfall data, to hourly data by using quality-controlled hourly rain gauge data from over 1300 observation stations across the country.

The Gridded Estimates of Areal Rainfall (CEH-GEAR) dataset was developed in order to provide a reliable 1 km gridded estimates of daily and monthly rainfall for Great Britain and Northern Ireland from 1890 to 2012. The dataset used for evaluation of daily rainfall data was derived from the Met Office historical weather observations for the UK, which includes a national database of daily and monthly rain gauged observations. However, it is necessary to be cautious on using CEH-GEAR dataset for estimation before 1961 because of the poor gauge density that characterized part of UK before that year. Considering that distribution of rain gauges is not uniform across the UK, the first step required has been a normalization of total rainfall rain gauges before interpolation procedures. It has been done by using the Average Annual Precipitation AAR, whose values were derived for a 10 km grid using monthly data from approximately 13100 rain gauges, then, in turn, they have been contoured and gridded at a 1 km resolution. The resolution of 1 km was chosen because it aligns with AAR grids, and because it fits with gauges density around all the UK. The interpolation methodology chosen for CEH-GEAR to derive daily and monthly 1 km grids has been the natural neighbour method, as it produces smooth rainfall surfaces without the boundary discontinuities. This method basically is a development of the Thiessen approach for whom a Thiessen polygon is a polygon within which no other operational gauge is closer. Reconstructing polygons for each grid point, two layers of polygons overlapped are obtained. Each rain gauge that has part of its original Thiessen polygon overlapped by the Thiessen polygon for the grid point is included in the rainfall interpolation at the grid point, and the weight associated with rain gauge is proportional to the area of overlap [Keller et al., 2015].

The second step consisted on the construction of a 1 km gridded hourly rainfall dataset for the UK by disaggregating the daily Gridded Estimates of Areal Rainfall (CEH-GEAR) dataset using comprehensively quality-controlled hourly rain gauge data from three different sources across the country: hourly rainfall accumulations from UK locations with records of minimum 10 years taken from the Met Office Integrated data Archive (MIDAS); hourly rainfall accumulations from the Scottish Environmental Protection Agency (SEPA) and tipping bucket rain gauge data (TBR) from the Environment Agency (EA). After the latter source was purified by problems of high-frequency tipping, the whole dataset has been subjected to quality controls (QC) tests, divided in two categories: tests in which values were automatically treated as suspect, and tests where values were treated as suspect if

additional evidence was available either from TBR data or QC procedures. Furthermore, it has been necessary to identify dry spells in the rain gauges to distinguish them from malfunctions in gauges. To guarantee an overall quality of the data, was also necessary to validate the QC procedure against historic extreme rainfall events. For each hourly rain gauge the accumulated daily total was compared with the corresponding grid square in the UKCP09 5 km gridded daily rainfall dataset. Basically, the quality of the hourly dataset was checked by accumulating the hourly values to a daily total for comparison with a quality controlled gridded daily rainfall product, retaining only values that showed a good agreement.

2.2. Discharge data

Measurements of river flows are the primary information required to study the hydrological cycle. Hydrometry thus focuses mainly on their evaluation. After 1960 a network of over 1500 gauging stations permits to create a large database of river flow measurements, managed, as already said, by the National River Flow Archive together with other Measuring Authorities.

UK rivers are typically short, shallow and subject to artificial disturbance. This fact makes the research of accuracy in relation to the fitness of purpose the data are derived for, quite challenging.

Different methods are employed to obtain the discharge river values. Usually river flows are measured indirectly using a stage-discharge relation between river level and river flow. This relation is simply achieved when a gauging station is provided by a weir with known hydraulic characteristics, otherwise it requires at first, the calculation of stream velocity, then to combine this value with the cross-sectional area of the river to have the measurement of flow. This procedure is applied throughout the flow range, leading to stable stage-discharge relation. However, even if gauging stations tend to be placed in hydraulically controlled conditions, sometimes boundary conditions such as changes in bed profiles following a flood or the aquatic plant growth occur, with the necessity of updating these relations.

Measurements of flow rivers are done by using diverse kind of instrumentation. In some gauging stations instruments actuated by floats in a stilling well measures and records in time gauging stations stage, while solid state loggers are used to record the water level. The largest number of gauging stations, instead, transmit river levels directly to a processing centre, usually by telephone line.

The stage-discharge relation is not required for ultrasonic gauging stations. In this case, in fact, flows are computed on-site where the times are measured for acoustic pulses to traverse a river section along an oblique path in both directions. The mean river velocity is related to the difference in the two timings and the flow is then assessed using the river's cross-sectional area. The accuracy expected with such methods may be compromised by high suspended sediment concentrations or heavy weed growth, which create an obstacle to the acoustic signal.

In the case of electromagnetic gauging stations, flow data may be computed on-site. The technique requires the measurement of the electromotive force (emf) induced in flowing water as it cuts a vertical magnetic field generated by means of a large coil buried beneath the river bed or constructed above it. This emf is sensed by electrodes at each side of the river and is directly proportional to the average velocity in the cross-section.

Even though the different methodologies used to record and compute flow river data, all the values of flow are usually recorder for sub-daily intervals, and are reported in cubic meters for second. All the information around discharge data in the UK have been taken from the Centre of ecology and Hydrology (CEH) website.

2.3. Catchment descriptors

As already said, United Kingdom presents a large number of gauging stations widespread in all the country, that permit to have flow and rainfall data easily available. However, in presence of sites that lack of data, it is necessary the knowledge of physical and climatological properties of catchments in order to transfer information such flood peak data of gauged basins (pooling-group), to ungauged ones with similar hydrological characteristics. Using these descriptors, together with key variables such the flood index, gives the possibility to realize flood estimation at ungauged sites. Indeed, an estimation based on catchments descriptors is more affected of uncertainty than estimation made by using peak flow data, nevertheless, it is useful to have a provisional assessment in flood design schemes.

It is furthermore important to explain how these descriptors are calculated. Nowadays different organizations have access to digital terrain models, that represent the geodetic surface of the earth over a regular grid. In particular, the Institute of Hydrology Terrain Model (IHDTM) by Morris and Flavin (1990), refers to digitalized river information taken from a 1:50000 OS maps that permit to give river valleys a correct position. This model is based on the steepest route to neigh bouring grid nodes and it includes a 50 m x 50 m grid of drainage path directions, that allows to derive catchments boundaries automatically. Therefore, using digital terrain models provides not just an automation in deriving catchments information, but also detailed values of the interested quantities.

Within the list of almost 30 descriptors computed in the Flood Estimation Handbook, only a few of them results interesting for this work. They have been obtained from the Centre of Ecology and Hydrology (CEH) website and their meaning will be described in the following Paragraphs. The information used next are taken from Flood Estimation Handbook (FEH) and from the Centre of Ecology and Hydrology (CEH) website.

2.3.1. Drainage Area (AREA)

The drainage area is the surface catchment area draining to a certain gauging station, projected onto a horizontal plane and defined in square kilometers. Information about catchments area has been obtained for more than the 99% of UK catchments from the IHDTM. Some difficulties in the evaluation occurred in particular for catchments with subdued relief, with necessity to supply information about drainage directions by measuring agencies. In other few cases, instead, anomalous values of runoff may affect the river flows, due to the fact that catchment boundary does not follow the topographical border.

2.3.2. Mean drainage path slope (DPSBAR)

The Mean Drainage Path Slope is a landform descriptor that represents an index of overall catchment steepness developed for the Flood Estimation Handbook. The IHDTM defines an outflow direction based on the steepest route at each grid node. Considering the distance between the two nodes and the difference in altitude, it has been possible to calculate the inter-nodal slope for all nodal pairs within a catchment and so on the DPSBAR for each catchment. The index is expressed in meters per kilometer with values larger than 300 in mountainous terrain and smaller than 25 in the flattest parts of the country. From Figure 3 and Figure 4 it is possible to notice the UK terrain distribution: it is mostly dominated by catchments with moderate slopes, with presence, in Scottish Highlands, of a small set of very steep catchments.

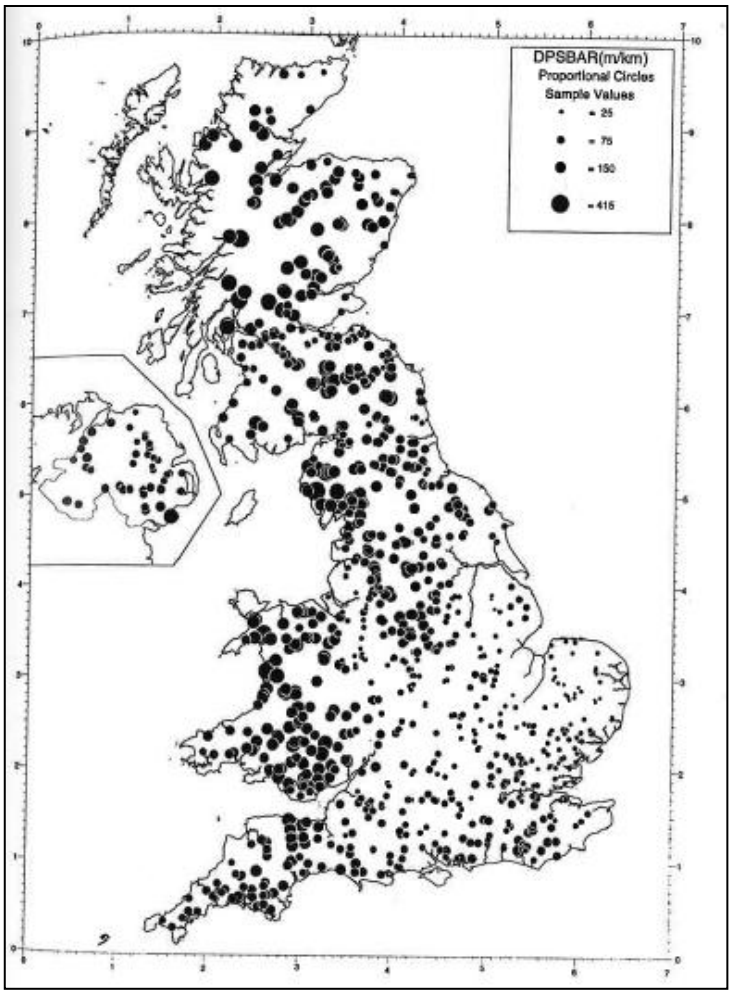

Figure 3: DPSBAR values for 943 gauged catchments.

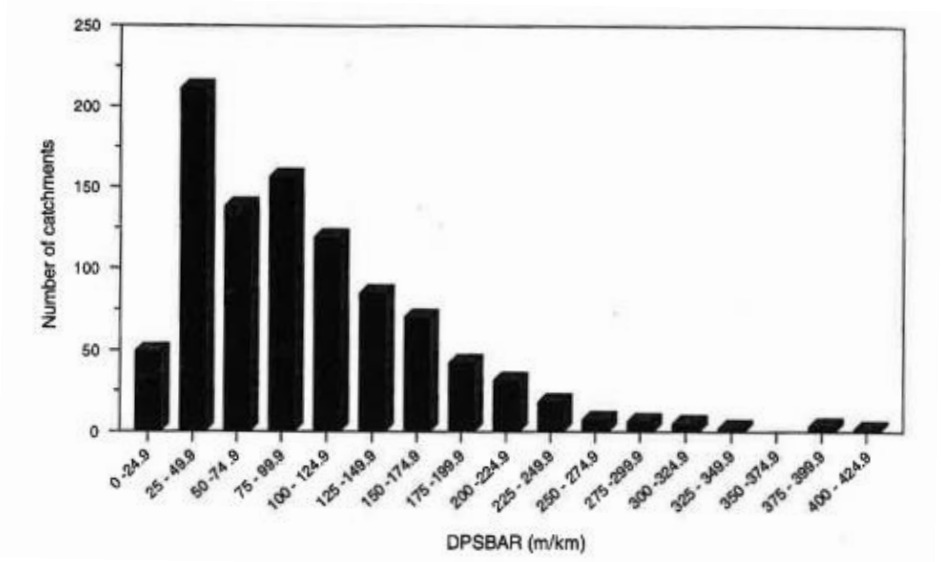

Figure 4: Numerical distribution of DPSBAR values.

2.3.3. Proportion of time catchment soils are wet (PROPWET)

PROPWET is a wetness index, developed for Flood Estimation Handbook, that represents the proportion of time soils are wet. It is therefore a measure of time that catchment soils are defined as wet or dry. The definition of PROPWET comes in fact after the concept of Soil Moisture Deficit (SMD). When soils are at the optimal condition of field capacity (SMD=0), a catchment is most likely to produce floods, with large values of SMD instead, dry soils tend to inhibit flood formation. The values of Soil Moisture Deficit are though theoretical and are used to represent large and often heterogeneous parts of the soil. For this reason it has been necessary to introduce a method to define the soil moisture content instead of choosing a random threshold. The peak over thresholds (POT) data were used to divide daily SMD data in days where POT flood occurred, and days in which they did not. Plotting them as percentage of each subset, it is clear from Figure 5 that the threshold between the occurrence of flood or not is the value of SMD equal to 5.7 mm. Hence 6 mm is considered the value above whom the catchment is defined dry.

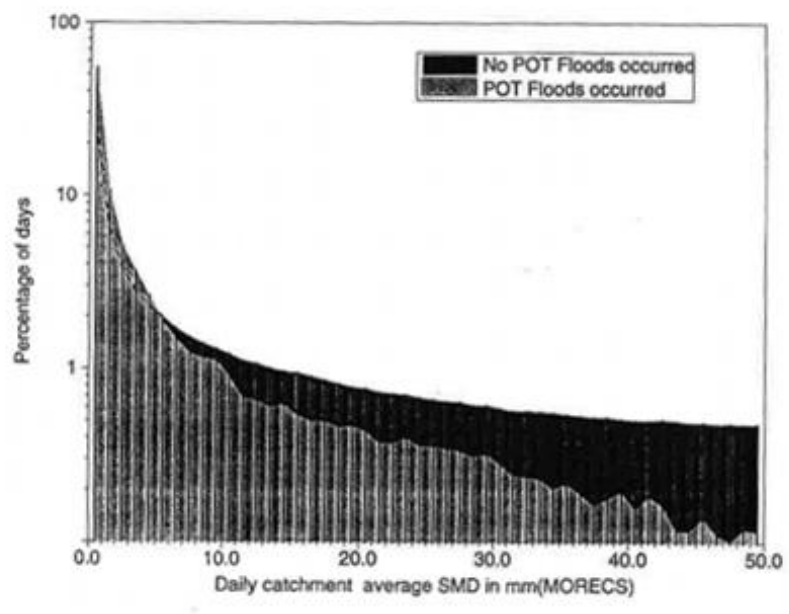

Figure 5:Distribution of daily SMD for floods and not floods days

The range of values PROPWET assumes is represented by a percentage: the wettest catchments have values bigger than 80%, while the driest parts of the country have values smaller than 20%. The highest values of PROPWET are found in northern Scotland. At the Cassley at Duchally, for example, a PROPWET of 0.84 means that the SMD threshold was exceeded for 84% of the time during the period 1961-90. As it is possible to underline from Figure 6, catchments soils are typically wetter for a large proportion of time in the north and north-west part of the country.

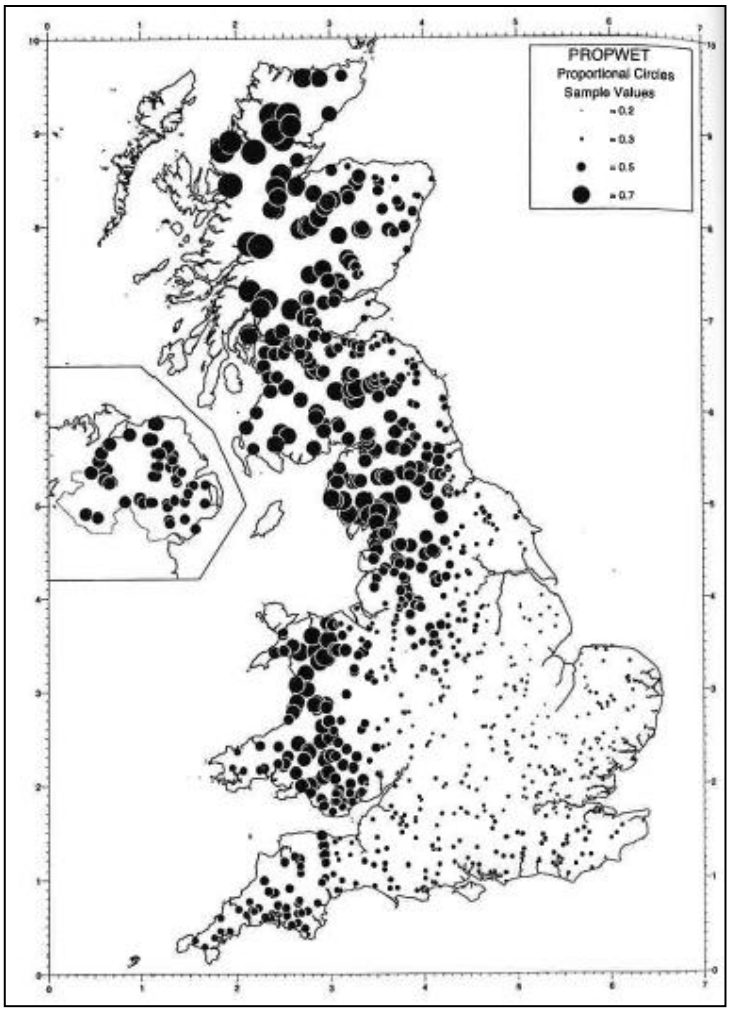

Figure 6: PROPWET values for 943 gauged catchments.

2.3.4. Baseflow index from the Hydrology Of Soil Types Classification (BFIHOST)

Another important descriptor used in the analysis of this work is the baseflow index BFIHOST. This index is a measure of catchments responsiveness derived using the 29-class Hydrology Of Soil Types (HOST) classification. In its simple definition it is the long-term average proportion of flow that occurs as baseflow, in this case defined taking into account the different soils properties. The HOST dataset is available as a 1 km grid and it records, for each grid square, the percentage associated with each HOST class present. Using IHDTM boundaries for each gauged catchment, it has been possible to index the soil characteristics of each catchment, which, related to runoff response, permitted the estimation of the BFIHOST.

The HOST classification has been an important step for accuracy in all concerns hydrological analysis. The previous classification of soils presented in the FSR, used only five class for all the country. Boorman et al., (1995) calculated the baseflow index (BFI) and the Standard Percentage Runoff (SPR) by considering 29 HOST class. The HOST classification is based on conceptual models that reproduce processes within the soil or the substrate. Basically they have three different physical settings:

- A soil on a permeable substrate in which there is a deep aquifer or groundwater (i.e. at >2 m depth). Classes: 1,2,3,4,5,6,13,14,15.
- a soil on permeable substrate in which there is normally a shallow water table runoff (i.e. at <2 m depth). Classes: 7 to 12.
- a soil (or soil and substrate) which contains an impermeable or semi-permeable layer within 1 m of the surface. Classes: 16 to 29.

In the following Figures is reported the HOST classification represented in the UK map and the soil characteristics of the 29 classes.

Class number	Class description
1	Soft sandstone, weakly consolidated sand
2	Weathered/fissured intrusive/metamorphic rock
3	Chalk, chalk rubble
4	Soft Magnesian, brashy or Oolitic limestone and ironstone
5	Hard fissured limestone
6	Hard coherent rocks
7	Hard but deeply shattered rocks
8	Soft shales with subordinate mudstones and siltstones
9	Very soft reddish blocky mudstones (marls)
10	Very soft massive clays
11	Very soft bedded loams, clays and sands
12	Very soft bedded loam/clay/sand with subordinate sandstone
13	Hard (fissured) sandstones
14	Earthy peat
15	River alluvium
16	Marine alluvium
17	Lake marl or tufa
18	Colluvium
19	Blown sand
20	Coverloam
21	Glaciolacustrine clays and silts
22	Till, compact head
23	Clay with flints or plateau drift
24	Gravel
25	Loamy drift
26	Chalky drift
27	Disturbed ground
34	Sand
35	Cryogenic
36	Scree
43	Eroded Blanket Peat
44	Raw Peat
50	Unsurveyed
51	Lake
52	Sea

Figure 7: Soil-geology classes used within the HOST project.

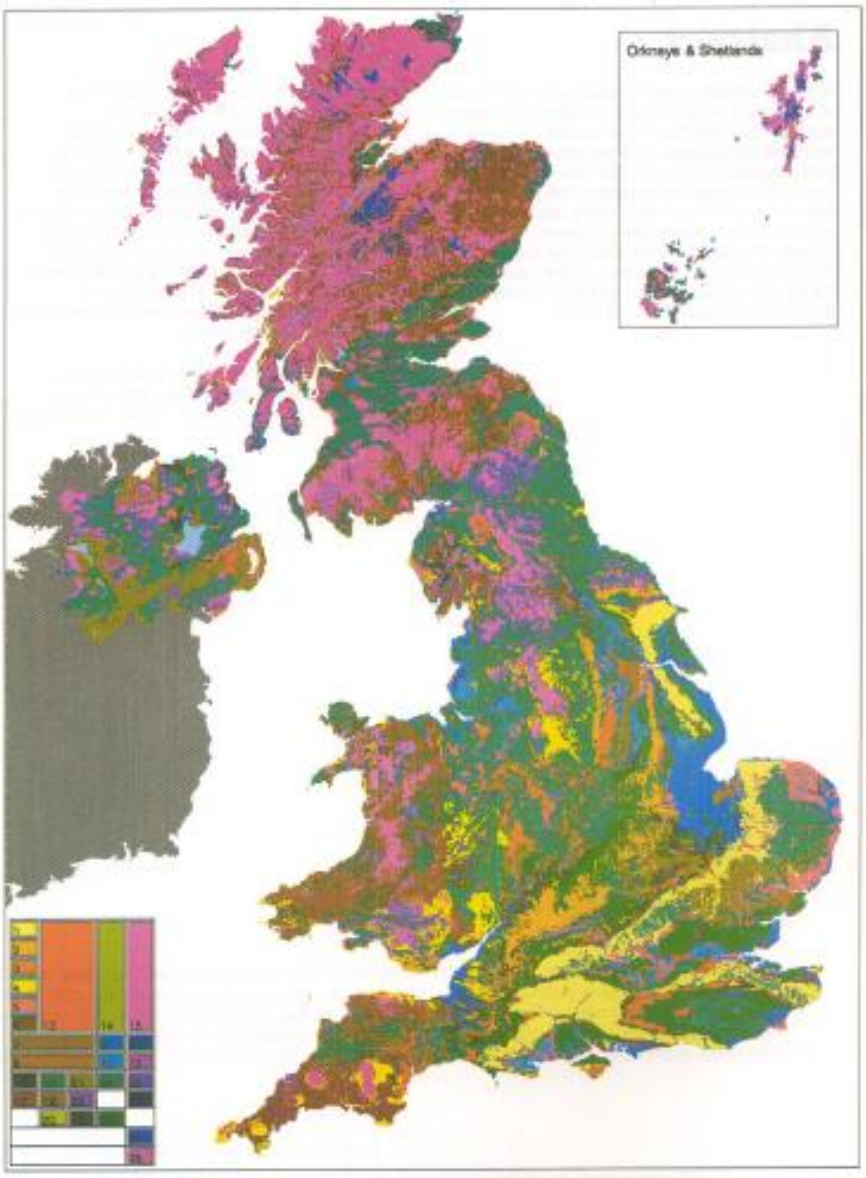

Figure 8: Dominant HOST class for each 1 km grid square.

2.3.5. Standard-period Average Annual Rainfall (SAAR)

This index represents the average annual rainfall for the standard period 1961-90 in Great Britain and Northern Ireland. It is provided by the Met Office on a 1 km grid, with average rainfall held to a resolution of 1 mm. This descriptor has not been used in this work. Instead of this index, the Mean Annual Precipitation (MAP) calculated on the temporal series or rainfall data analyzed (1999-2008) has been used. It is important, though, to mention this descriptor because it is used in flood estimation by statistical methods for ungauged catchments.

2.3.6. Flood Attenuation by Reservoirs and Lakes (FARL) index

Another important descriptor to consider is the index FARL, that takes into account the flood attenuation realized by reservoirs and lakes. This kind of attenuation tends to be more relevant when storage elements are directly linked to the streamflow. However, some effects occurs also in isolated reservoirs or lakes, considering that the amount of rainfall that falls on them will be lost from flood generation process. The catchment descriptor FARL is defined as the product of the individual local index values α, obtained evaluating the local effect that a reservoir or a lake has on the catchment outlet. It is close to the unity when there is a little influence in flood attenuation, when the importance of attenuation increases, the value of the index decreases. Figure 9 represents a spatial distribution of FARL index values. The range of values is about 0.557 and 1. The largest indices are found in the north Wales, Cumbria, Northumberland, Scotland and Northern Ireland . The lowest values, instead, are mainly in upland parts of the country. In this work it has been decided not to use this descriptors, considering that the data set analyzed refers only to rural benchmark catchments, not largely affected by reservoirs and lake flood attenuation.

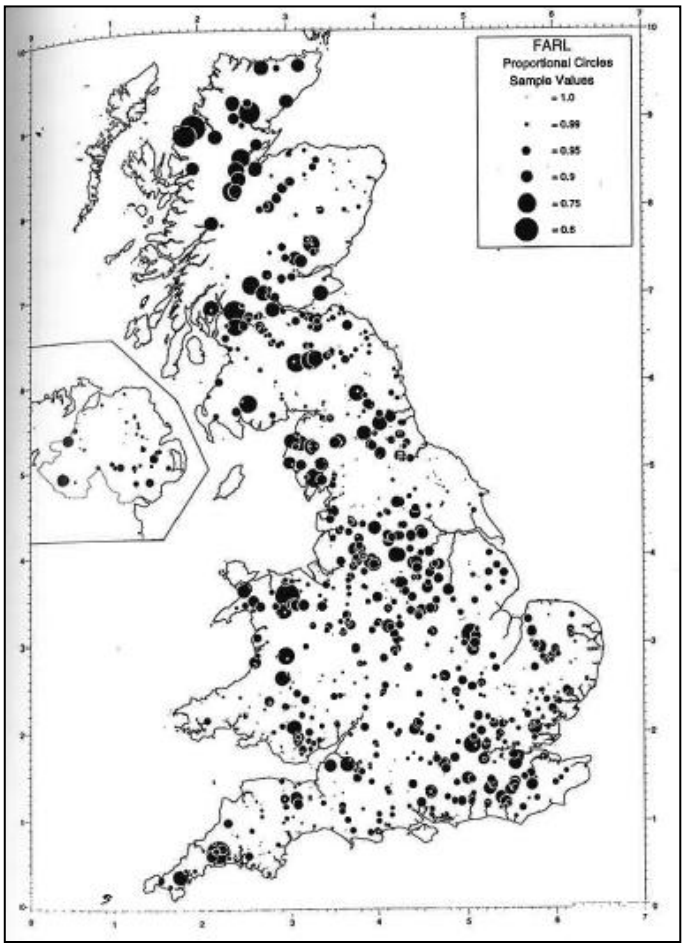

Figure 9: FARL value for 943 catchments.

3. Methods

In this chapter all methods used to calculate and to analyze runoff coefficients will be described. The work turns around the spatial and temporal variability of runoff coefficients in relation to controls considered varying in space and time respectively. Once runoff coefficients are characterized by spatial and temporal pattern, it is possible to summarize these information through different parameters and to use them for an estimation of mean runoff coefficients in ungauged catchments. The work thus can be divided in two diverse parts, whose main steps are shortly described in the following section, before a precise development for each one of them.

The first part of the work bases on the analysis of 76 benchmark catchments with the purpose of investigating on their characteristics. The analysis started with baseflow hygrographs separation. Different filters and methodologies are applied and analyzed due to choose the one that fits properly on catchments hydrographs in term of reliable quantity of baseflow volume.
The next step is the separation of different flood events, analyzing temporal distribution of rainfall data besides the hydrograph separation. Then the calculation of important characteristics is done for each event, such as the evaluation of runoff coefficients, volume of precipitation, the peak flow, volume of quick flow, etc., that have been ordered in a database containing more or less 22.000 events related to the analyzed catchments. Finally the analysis of runoff coefficients for each catchment of the database is done. After a cautious choice of the events to consider in the work, it was possible to realize a spatial and temporal analysis of them, referring to the by R. Merz et al. (2006) and Merz and Bloschl (2009).
The second part of the work bases on the objective to use the results obtained for the 76 benchmark catchments analyzed, to attempt prediction of runoff coefficients for ungauged catchments. It was done considering within the dataset, 51 catchments as the gauged ones, and the other 25 as if they were ungauged, due to make a comparison between the estimated and actually calculated parameters. Estimation of ungauged parameters has been done by realizing some regression trees with catchments descriptors as input. Finally, errors due to this procedure have been analyzed.

This work has been conducted by analyzing hourly rainfall and discharge data referred to time series of ten years (1999-2008) for each one of the 76 catchments. Their area ranges from 4.4 km^2 to 1500 km^2, all of them are retained suitable for the analysis because, being benchmark catchments, no anthropogenic impacts occur.

3.1. Baseflow separation

One of the most important steps for hydrological analysis is the hydrograph separation. It is a key point for the understanding of the hydrological cycle, whose knowledge is hard to be completely clear. Hydrological cycle is usually simplified in a system, in turn subdivided in three subsystems related one to each other: the *atmospheric water system,* that is based on the processes of precipitation, evaporation, interception and transpiration; the *surface water system,* that is based on the processes of overland flow, surface runoff, subsurface and groundwater outflow, runoff to streams and oceans; the *subsurface water system,* that is based on the processes of infiltration, groundwater recharge, subsurface flow and groundwater flow.

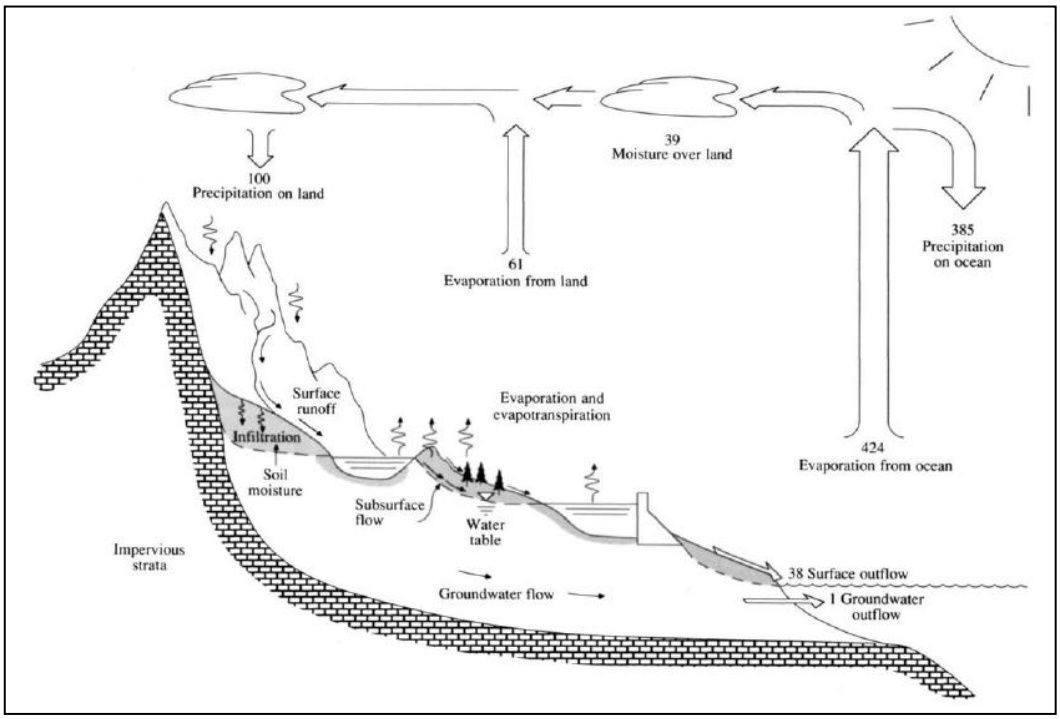

Figure 10: Schematic of the hydrological cycle ["Applied hydrology" by Chow et al., Figure 1.1.1].

The main interest in this study is given to the surface water system. Indeed, it is relevant to understand the runoff process and hence the main components that characterize the hydrograph of a certain cross section of a catchment. The streamflow in fact is made by different flow elements that reach a cross section with increasing delay: the direct surface runoff, that flows directly overland; interflow or subsurface flow, which flows in the shallow depths of the surface; the groundwater baseflow, characterized by water that slowly percolates until the groundwater storage and feed the stream from the deeper layers of surface.

This theoretical classification goes against some directly observed mechanism of runoff and furthermore, it is not possible to distinguish all the different components of the hydrograph mentioned above. For this reason streamflow is considered given by a two component-process, characterized by

baseflow and quickflow. In next section the approaches used for hydrograph separation will be described.

3.1.1. Baseflow separation methods.

Hydrograph separation methods are basically divided in two categories: tracer-based and non tracer-based methods.

The first methods use chemical and physical means for dividing baseflow from streamflow, such as environmental isotopes and chemical tracers. They permit to have accurate information about the separation of groundwater from surface water and are a reliable reference to make a comparison with non tracer methods. However, they require a lot of time to collect data, realize analyses and to obtain results, so they would be too onerous to be used for a large number of catchments.

Non tracer-based methods are divided in different techniques. Graphical methods as the straight line, fixed-base, variable slope methods, were spread in the early twentieth century [Linsley et al., 1988]. Even if they are really intuitive in their application, they are suitable only for single peaks events, they lack physical bases in interpolating the base-flow, having too much uncertainty in the decision of events starting and finishing points.

Filtering methods, instead, are typically designated for long term, daily time scale, data records. Among them, the smooth minima base-flow separation method of the United Kingdom Institute of Hydrology stands out.

Nowadays the major category of non tracer-based methods is represented by digital filters. These methods consider total stream-flow as being composed of both quickflow and baseflow and apply signal processing techniques to a streamflow time series in order to remove the high-frequency quickflow signal to obtain the low-frequency baseflow signal [Li et al., 2013].

Digital filters can be distinguished in one-parameter filter, two-parameter filter or three-parameter filter considering the number of parameters required for its application. In Table 2 the most common digital filters presented in literature are reported.

One-parameter Filters
Lyne and Hollick Filter (Lyne and Hollick, 1979)
Chapman and Maxwell Filter (Chapman and Maxwell, 1996)
Conceptual method (Mugo and Sharma, 1999, Nathan and McMahon, 1990)
Two-parameter Filters
Boughton Algorithm (Boughton, 1993)
Recursive Digital Filter (Eckhardt, 2005)

Three-parameter Filters
IHACRES (Jakeman and Hornberger , 1993)

Table 2: Digital filters.

Usually they bases on a parameter defined as the "recession constant" and on the baseflow index, that allows to summarize the results of baseflow separation. The first one can be evaluated through a recession analysis, studying the recession limb of the hydrograph (see section 3.1.1.5. for more details). The recession analysis bases on the study of the streamflow hydrograph recession curve, in order to understand the relation between groundwater and surface water during low flow periods [Tallaksen, 1995; Hall, 1968; Smakhtin, 2001; Thomas, 2015]. The second one, defined as the long term ratio of baseflow to total streamflow, can be obtained from in situ measurements or estimated by statistical methods based on catchments properties. These methods are simple to be automated, they provide reasonable results, but, on the other hand, it is difficult to decide whether or not one is better than another.

Among all the categories of methods presented above, only digital filters and the UKIH method will be used in this work. Among the digital filters, it has been decided to use only one parameter digital filters, whose recession constant has been estimated with recession analysis. Although the wide arrays of values proposed in literature for the different parameters of the filters, the variability of each catchment characteristics is too significant to adopt them uniformly. Digital filters with more than one parameters were avoided by following the purpose of simplicity. The using of complex models in fact is not always synonymous of better estimations. The UKIH method instead, has been tested because it is the main method used in UK for baseflow separation, as suggested in the Low Flow Studies report by the Institute of Hydrology (1980). All the analysis have been computed by using the software Matlab.

3.1.1.1. Lyne and Hollick Filter

Lyne and Hollick Filter is based on signal analysis and depends on the recession constant k. It separates the total streamflow (Q) in quickflow (Q_q) and baseflow (Q_b). The first step is the evaluation of quickflow, the second one is the computation of baseflow as a simple difference between total streamflow and quickflow.

$$Q_{q(i)} = k \cdot Q_{q(i-1)} + \frac{1+k}{2} \cdot (Q_i - Q_{i-1}) \quad for \quad Q_q \geq 0 \qquad \text{[Eq. 7]}$$

$$Q_{b(i)} = Q_i - Q_{q(i)} \qquad \text{[Eq. 8]}$$

The initial condition defined for the LH method sets the total streamflow equal to the baseflow so that for example $Q_{b(i)} = Q_1$. Taking into account a range of recession constants that goes from 0 to 1, it is clear that for $k = 0$, in the rising limbs of the hydrograph the baseflow would be obtained dividing in half the sum of the streamflow at the current steps and at the previous one, in the decreasing limbs the baseflow would be equal to the streamflow. Besides, if $k = 1$, when $Q_i > Q_1$ the baseflow obtained would be equal to the streamflow, when $Q_i < Q_1$ the same condition is applied due to the fact that baseflow could not be bigger than the total streamflow.

3.1.1.2. Chapman and Maxwell

Chapman and Maxwell proposed a digital filter that assumes the baseflow as a weighted average of the direct runoff (Q_d) and the baseflow at the previous step (Q_b), eliminating the uncertainty of the previous method which considers the baseflow constant in absence of quickflow.

$$Q_{b,i} = k \cdot Q_{b,i-1} + (1-k) \cdot Q_{d,i} \qquad \text{[Eq. 9]}$$

By the substitution $Q_d = Q - Q_b$ Chapman and Maxwell method can be reformulated as in the following formula, in which baseflow depends on the total streamflow (Q), on baseflow at the previous step, on the recession constant (k). The constraint to be applied is: $Q_b \leq Q_i$.

$$Q_{b,i} = \frac{k}{2-k} \cdot Q_{b,i-1} + \frac{(1-k)}{(2-k)} \cdot Q_i \qquad \text{[Eq. 10]}$$

3.1.1.3. Conceptual method

The conceptual method was proposed for the first time from Nathan and McMahon in 1990 and later used by Mugo and Sharma to evaluate the runoff components in forest catchments from Kenya. It can be represented schematically in Figure 11:

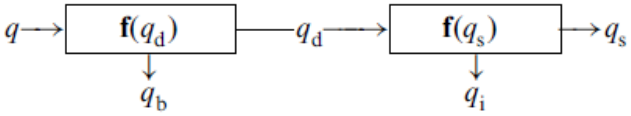

Figure 11: Scheme of conceptual method.

The total streamflow (q) passes through the runoff filter $\boldsymbol{f}(q_d)$, that divides the direct runoff (q_d) from the baseflow (q_b). The direct runoff that has been filtered passes through a surface runoff filter $\boldsymbol{f}(q_s)$,

that divides it in interflow (q_i) and surface runoff(q_s). The method can be represented by the following equations.

$$q_{d,k} = \beta_d \cdot (q_{d,k-1}) + \frac{(1+\beta_d)}{2} \cdot (q_k - q_{k-1}) \qquad \text{[Eq. 11]}$$

$$q_{s,k} = \beta_s \cdot (q_{s,k-1}) + \frac{(1+\beta_s)}{2} \cdot (q_{d,k} - q_{d,k-1}) \qquad \text{[Eq. 12]}$$

$$q_{b,k} = q_k - q_{d,k} \qquad \text{[Eq. 13]}$$

$$q_{i,k} = q_{d,k} - q_{s,k} \qquad \text{[Eq. 14]}$$

Where β_d is the direct runoff filter parameter, β_s is the surface runoff filter parameter, q_k is the streamflow at the time step k, q_{bk} is the baseflow at the time step k, q_{dk} is the filtered runoff response at the time step k, q_{ik} is the interflow at the time step k, q_{sk} is the filtered runoff response at the time step k. Taking into account Equation (7), it is possible to reformulate an equation (15) as the following one, in which β_d has been substituted with recession constant k and $q_{b,k} \leq q_k$.

$$q_{b,k} = \frac{(1-k)}{2} \cdot (q_k + q_{k-1}) + k \cdot q_{bi-1} \qquad \text{[Eq. 15]}$$

Reformulating Lyne and Hollick method, expressing quickflow as the difference between total streamflow and baseflow, the expression found is the same one presented by this method. For this reason in next paragraph only Lyne and Hollick method will be mentioned.

3.1.1.4. UKIH method

This method has been developed from the United Kingdom Institute of Hydrology in 1992. It consists in dividing the daily flow data in non-overlapping consecutive groups of five days and find the sequence of minima that then characterize the baseflow data. The precise steps mentioned in the report of IH are written in the following.

- Divide the mean daily flow data into non-overlapping blocks of five days and calculate the minima for each of those blocks, and let them be called $Q_1, Q_2, Q_3, \ldots \ldots Q_n$.
- Consider in turn (Q_1, Q_2, Q_3), (Q_2, Q_3, Q_4), (Q_{i-1}, Q_i, Q_{i+1}), etc. In each case, if $0.9 \cdot central\ value < outer\ values$, then the central value is an ordinate for the baseflow line. Continue this procedure until all the data have been analyzed to provide a derived set of baseflow ordinates $QB_1, QB_2, \ldots QB_n$ which will have different time periods between them.
- By linear interpolation between each value, estimate each daily value of $QB_1, \ldots QB_n$
- If $QB_i > Q_i$ then set $QB_i = Q_i$.

All the methods described in this paragraph have been applied considering hourly data of flows.

3.1.1.5. Recession analysis

As already said in the previous paragraph, in this study of a wide variety of UK catchments, it has been decided not to consider a fixed value of recession constant from literature, but to evaluate it by using a recession analysis. Studies about recession analysis started with Boussinesq in 1887 and Maillet in 1905. Recession analysis is mostly related to groundwater flow and to separation of flow components. Except from that, it is very useful in the field of water management: for example for water supply, calibration of Rainfall-Runoff models, hydroelectric power-plants [Tallaksen 1994]. In particular, recession analysis is based on the study of decreasing limbs of the hydrograph, defined as recession curves, that contain a large number of information about aquifer characteristics and storage properties.

One of the first quantitative expressions of recession curve was given in 1887 by Boussinesq, who developed a non linear differential equation governing unsteady flow from an unconfined aquifer to a stream, based on idealized conditions such as the absence of evapo-transpiration, leakage and recharge. Boussinesq himself linearized the equation, referring on the Depuit system that assumes the vertical flow components negligible, and null the effect of capillarity above the water table. In Dupuit-Boussinesq equation, $Q(t)$ is discharge at time t (m^3/s), Q_0 is discharge at start of recession (m^3/s), k is the recession constant (L/s).

$$Q(t) = Q_0 \cdot \exp(-kt) \qquad \text{[Eq. 16]}$$

Usually flow series contain different recession segments due to frequently rainfalls that occur in catchments exposed to humid climate. For this reason it is necessary to gather the information from each recession segment to evaluate the recession coefficient characterizing the catchment. One method consists in the construction of a recession master curve, that solves this problem by trying to find a mean recession curve, from whom it is possible to evaluate the coefficient required. The second possible

method consists in the evaluation of a constant for each recession element and to calculate the average value to obtain a mean coefficient characterizing the basin.

In this study it was decided to summarize recession behavior by using the method of Brusaert and Nieber (1997), that bases on a correlation method in which the dependence on the interval of time t is removed by plotting the series of flow $Q(t)$ with the rate of change in flow for a differential dt.

$$-\frac{dQ}{dt} = c \cdot Q^d \qquad \text{[Eq. 17]}$$

This equation can be also expressed in a simpler linear form assuming $d = 1$, which in term of differences can be written as:

$$(Q_{t-1} - Q_t)/\Delta t = -c \cdot (Q_{t-1} + Q_t)/2 \qquad \text{[Eq. 18]}$$

It is possible to evaluate the recession the recession constant inverting the previous formula as in the following:

$$k = \exp(-c) \qquad \text{[Eq. 19]}$$

Another factor that requires to be taken into account, is the evaluation of points that define the recession curves in the hydrograph. For sure, they belong to the decreasing limb of hydrograph, but it is important to define which is the starting point for each of them and how long recession segments should be.

The starting point of recession segments can be defined as a constant value or as a variable one, it is possible to define more than one value, considering the number of catchments analyzed. It can be individuated also as the discharge at a particular time step after a rainfall period or after a peak and it implicates that each recession segment will have a different discharge starting value.

There are many possibilities also in the evaluation of the recession length: it could be constant or variable. Usually a recession curve has a minimum length of 4 or 10 days length [Tallaksen, 1994], but it depends on the climate characterizing each catchment.

For the study of 76 basins spread in United Kingdom, the same rules for defining the recession curves have been applied. No length constraints have been considered, but recession segments have been

defined considering only the part of hydrograph that belong to the decreasing limb, and characterized by a period of 2 days without rain. Furthermore, only values of discharge smaller than the median value of flow for each catchment have been taken into account for the evaluation, to avoid including the recession immediately after flood events.

3.2. Estimation of event runoff coefficients

In this section the procedure followed to obtain runoff coefficients will be explained.

3.2.1. Event selection

An unavoidable step required in order to calculate the event runoff coefficients is the rainfall runoff event selection. Within the 10 year series of data it was necessary to identify the rainfall runoff events in order to gather important information for each catchment such as the volume of quickflow, the volume of baseflow and the volume of rainfall associated to each event.

As a result from the hydrograph separation methods selected, hydrographs with clear baseflow and quickflow underlined have been obtained. Hence, for each catchment, the rainfall runoff events have been selected, by using a simple procedure implemented in Matlab that matched the event separation as it would have been done manually. The first step was the identification of the starting and ending point of each event. The starting point of each event was estimated searching through the data for time steps in which baseflow and total streamflow had the same value, but followed by time steps in which total streamflow values were bigger than baseflow values. The ending point, instead, has been estimated searching for time steps in which baseflow and total streamflow had the same value, but in the previous time step characterized by a total streamflow bigger than the baseflow, and in the following time step by the same values of baseflow and total streamflow. In Table 3 a simple schematic of the conditions applied for event selection is reported.

Time step	i-1	i	i+1
Start of events	Streamflow ≥ Baseflow	Streamflow = Baseflow	Streamflow > Baseflow
End of events	Streamflow > Baseflow	Streamflow = Baseflow	Streamflow ≥ Baseflow

Table 3:Event selection condition applied.

Each rainfall runoff event is characterized by: a volume of total streamflow that is represented by the area underlying the streamflow curve within the time interval between the starting and the ending time steps; a volume of baseflow represented by the area underlying the baseflow curve within the same time interval; the volume of quickflow, obtained as the difference between the previous quantities.
The rainfall volume associated with each event has been considered referring to a larger time interval. It was necessary, in fact, taking into account a range of 24 h before the event [Giani, 2017] in which the

rainfall increased the soil moisture, easing the forming of a runoff event. This assertion may be true in the case in which the catchment soil is far from the field capacity and/or it is permeable. Otherwise, in impermeable catchments, or when the soil is nearly saturated, rainfall tends to be transformed immediately in overland flow, without giving birth to infiltration processes.

3.2.2. Estimation of runoff coefficients

Runoff coefficients then were estimated as the ratio of event quickflow volume and event rainfall volume. Even if small, with the procedure described before, all events were considered in the analysis, and so their respectively runoff coefficients. However, it is necessary to do some considerations.

Evaluation of runoff coefficients has been straightforward thanks to the procedure that has been used. In fact each event is clearly separated from another, so that total flow is nearly equal to the baseflow between events. Instead, if the direct runoff at the end of an event was significantly larger than zero, there would have been and underestimation of the runoff coefficients.

The procedure used to select events based uniquely on the baseflow hydrograph separation obtained by previous analysis, without relying on rules or thresholds to consider only a certain type of events. Besides, hydrograph separation methods are automated procedures based on general criteria that assess the pattern of baseflow without contemplating the hydrological behavior of each catchment. Even if different methods are tested, none of them will assess the exact pattern of baseflow. For this reason, it is possible that events identified with the event separation method are not actual events. Cases, for example, in which direct runoff is larger than zero, but the volume of rain is null for all event duration have been excluded from the analysis whether they were due to baseflow separation imprecision or rainfall errors, or to rare cases in which other sources create direct runoff. In terms of runoff coefficients, all values that tend to infinite have been excluded.

For the same reason, due to the inaccuracy of baseflow separation methods, it was necessary to apply another threshold, excluding runoff coefficients bigger than the unity. It is in fact impossible that direct runoff are larger than rainfall volume. As Figure 10 simply shows, precipitation volume that falls on the earth surface may follow different paths, nevertheless all the amount of water involved in them should be at least the same of volume precipitation.

Furthermore, it has been decided to focus the attention only on events of a certain magnitude, in term of rainfall volume involved in each one of them. This work in fact, has the objective to investigate around runoff events that may generate flooding: events whose rainfall volume was less the 5 mm were not considered relevant with this purpose.

3.2.3. Database

All the information about catchments and their rainfall-runoff events have been gathered in a unique database. Each event was identified with a specific ID in order to avoid confusion between events of different catchments. Then, for each one of them, relevant information were summarized in Table 4. By using a simple Matlab structure containing all this information is thus possible to investigate about all events characteristics.

Database information	
ID event	[-]
Name of the catchment	[-]
Peak flow of the event	[mm/h]
Volume of streamflow	[mm]
Volume of baseflow	[mm]
Volume of quickflow	[mm]
Volume of rainfall	[mm]
Peak rainfall value	[mm/h]
Starting date of the event	[year-month-day-hour]
Ending date of the event	[year-month-day-hour]
Duration of the event	[h]
Catchment area	[km^2]
MAP	[mm]
PROPWET	[-]
DPSBAR	[m/km]
BFIHOST	[-]
Latitude	[°]
Longitude	[°]

Table 4: Database information.

3.3. Spatio-temporal variability of runoff coefficients

The main purpose of this thesis is the analysis of spatial and temporal variability of runoff coefficients. The knowledge of this information leads to investigate about runoff processes developments across the

country, which characteristics influence catchments runoff response, considering different kind of soils and climate, how runoff coefficients vary in time, considering both annual and monthly variability.

Two studies conducted in Austria have been taken as a cue for this work: "Spatio-temporal variability of event runoff coefficients" by Merz, Bloschl, Parajka (2006) and “A regional analysis of event runoff coefficients with respect to climate and catchments characteristics in Austria” by Merz and Bloschl (2009). The former is the starting point of spatio-temporal analysis of runoff coefficients for Austrian catchments. Merz et al. (2006) developed a methodology to estimate around 50000 event runoff coefficients from 337 Austrian catchments by using automated procedures of baseflow hydrograph separation and event selection. They managed to find a regional pattern in runoff coefficients distribution, but without relating explicitly catchment characteristics to runoff response. The latter, instead, based on the runoff coefficients dataset created in the previous work, focused on the objective to relate spatial and temporal distribution to catchment characteristics and climate. Reasons of time did not permit to work with a large dataset and to realize deep statistical analysis as done in these papers. However, with means available, spatial and temporal variability of runoff coefficients have been investigated, as precisely described in the following sections.

3.3.1. Spatial Variability

The goal of the analysis of spatial variability of runoff coefficients is to investigate, how these coefficients, representative of runoff generation in catchments, vary within the United Kingdom. Multiple factors are responsible of different runoff response in the country.

First of all there is a variety of soil types, classified in 29 groups by the HOST study classification, that implies diverse processes followed by the rainfall inside and outside the soil.
Also the land of use is a relevant factor. Catchments characterized by a manmade environment differs from benchmark catchments, in which usually urban extension is less than the 10% of the whole area. Presence of human extension is often synonymous of impermeable surfaces, that obstacle rainfall infiltration, immediately transformed in direct flow. Rainfall that falls in catchments decreases the concentration time to get to a cross section, with higher peaks of flow in hydrographs, and related problems to the neighboring elements . An example may be represented by rivers directly surrounded by buildings that in flooding do not have the proper space to contain the increasing flow, which is channeled between banks and it has a higher peak flow as a result.
Another relevant factor is the climate. United Kingdom is in fact well known for the variability of its weather within days, seasons, years and places. The presence of Atlantic Ocean with its warm waters from the West, the influence of mainland Europe, together with coastlines and islands are key elements for this variability.

In order to analyze the influence of these effects on runoff coefficients values it has been necessary to consider different catchment descriptors.
The primary hypothesis evaluated to consider the effect of soil types and land use on runoff coefficients, was the calculation of the SCS curve number [US-SCS, 1972; Dingman, 1994]. The SCS method provides an index of the mean catchment response in the range that goes from 0 to 100 by classifying the soil type, the land use and antecedent rainfall. Multiple reasons suggested to discard this method. First of all, the difficulty to associate for each catchment the type of soil among the 29 classes to the four classes proposed by the method. Secondly, from the land use percentages obtained for each catchment from the CEH website, was difficult to understand in which category of soil coverage each catchment belongs, considering though, that all the 76 catchments are benchmark, hence without large variability in the land use, except from some mountainous catchments in the North of UK.

For this reason the decision that has been taken was to analyze, among catchments, the variability of some descriptors representative of soil condition and morphology. Runoff coefficients have been therefore related to a descriptor that takes into account the proportion of time in which the soil is wet (PROPWET); to a descriptor that contemplates the morphology of the soil with a average steepness of lands (DPSBAR); to a descriptor representative of the responsiveness of catchments, by giving the proportion of flow that comes as baseflow in relation to the 29 soil type classes (BFIHOST).
The climate effect, instead, has been considered by using the mean annual precipitation (MAP), that represents the hydro-climatic situation of a basin, and a surrogate measure of the antecedent soil moisture content of the soil.

All these quantities have been related to runoff coefficients for all the catchments, with the purpose to find a regional pattern able to divide UK in regions depending on these characteristics. Furthermore, each catchment has been analyzed singularly investigating for each event on the relation between runoff coefficients, quickflow and precipitation volumes.

3.3.2. Temporal variability

The analysis of temporal variability of runoff coefficients allows to investigate on how runoff processes develop within time, in particular within seasons. It is known that the different regions of UK have variable climates. Places in the South and in the East tend to be drier, warmer, sunnier and less windy, than places in the West and in the North. Spring and summer usually are characterized by better weather conditions, synonymous of drier soils and higher temperatures, than autumn and winter, usually colder and wetter. The purpose is hence the knowledge of a temporal pattern of runoff coefficients, that may lead to characterize whole UK with a temporal pattern. This interest is directly related to the main goal of the research this work bases on. In fact, assessing a temporal pattern leads to prediction of mean

runoff coefficients versus time and their uncertainty for ungauged catchments. Standard deviation has been chosen to describe the within-month random variability of runoff coefficients.

For this reason, by using the data of ten years records, it has been possible to study the trend of mean runoff coefficients versus months. It has been done taking into account each singular catchment at first, then the whole set. Another important step was based on studying the correlation between mean runoff coefficients and their standard deviation within months and years. This kind of relation leads in fact to understand the variability of runoff coefficients values in relation to their magnitude. Some other analysis have been done relating runoff coefficients to events duration.

To summarize the information about the patterns obtained, fitting curves have been used, characterizing with their parameters the temporal variability investigated. The pattern between mean runoff coefficients versus months has been fitted by a cosine shape curve represented by two parameters: an amplitude defined as $\bar{\delta}$, and a mean value of runoff coefficient defined as $\overline{rc}$, which permit to summarise monthly variability for each catchment.
The relation between mean runoff coefficients and their standard deviation within months has been fitted with straight linear functions characterized by their slope and their intercept for each catchment. The knowledge of these quantities leads to associate uncertainty variability (standard deviation) to a determined value of mean runoff coefficient.

3.4. Regression trees

The results found through spatial and temporal variability of runoff coefficients analysis have finally been summarized in order to do prediction of these values and their standard deviation for ungauged catchments.

Spatial variability analysis leads to classify regional patterns considering soil properties, topography and average annual precipitation. The data set of 76 catchments is not large enough to create distinguished regions in which considering determined ranges of runoff coefficients. More rainfall, discharge data and consequently runoff events should be considered for creating a reliable dataset of runoff coefficient values within each region. The analyses conducted, though, are a starting point that allows to individuate different regional patterns.

Temporal variability analysis instead, allows to predict mean runoff coefficients behaviour in time and the related magnitude of uncertainty (standard deviation). By using the four parameters mentioned above ($\bar{\delta}$, $\overline{rc}$, intercept and slope of the relation mean-sd of runoff coefficients) calculated for a group of gauged catchments, it is possible to do some predictions for ungauged catchment. Considering to know for ungauged basins only some descriptors of the soil, through $\bar{\delta}$ and $\overline{rc}$ estimated for similar catchments it is possible to reconstruct the temporal pattern of runoff coefficients within months. In the

same way, the knowledge of the relation between mean and standard deviation of runoff coefficients for gauged catchments permit to recreate the variability of runoff coefficients around its average within months for ungauged ones.

In order to do the prediction about spatial and temporal variability of runoff coefficients, the dataset of 76 catchments was divided in two groups, 51 of them have been used as "gauged" catchments, the other 25 as "ungauged" catchments. This subdivision was done with the objective to compare predicted values of parameters with their actual calculated values. The choice of catchments to be considered "gauged" or "ungauged" was done ordering all the data set in function of the descriptor more influencing the response of prediction, and creating two groups that had the same average value of this descriptor. In this case PROPWET, was selected as the descriptor that permit to obtain more reliable results.

After this selection was done, regression trees were used to realize the prediction of the four parameters mentioned above. Regression trees, or decision trees are an instrument useful to predict response to data. They requires some input variables (predictors) and a response variable. Starting from the root of the tree, that represents the beginning of the tree, input variables are checked and classified in relation to their values. This classification has as result a binary tree with branches that subdivide the input variables and whose leaf nodes represent the response variable values.
A simple example is given in Figure 12. In this case, the input predictor are called x_1 and x_2. The first decision is whether x_1 is smaller than 0.5. If yes, following the left branch, the tree classifies the data as type zero. Otherwise, following the right branch, the value of the second predictor x_2 must be checked, and the procedure continues as at the start of the decisional process.

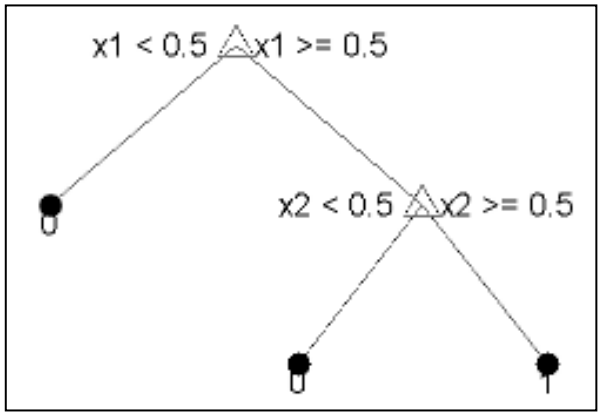

Figure 12: Schematic of a simple regression tree.

The regression trees were used to predict the four parameters mentioned above: $\bar{\delta}$ and $\overline{rc}$, that leads to recreate the monthly pattern of mean runoff coefficients; intercept and slope of the relation Mean-Sdev of runoff coefficients, to recreate this relation for ungauged catchments.

Different predictors were used to do this analysis, for each one of the parameter to be estimated. At first only one descriptor (PROPWET) was used as input, then others, such as DPSBAR, BFIHOST, AREA were added as input variables to improve the quality of the estimation. The root mean square error (RMSE) was used to have a measure of the error due to these predictions. Its formulation is reported in Equation 20: where $\hat{x}$ represents the estimated variable, x represents the observed value, n, is the number of elements of the sample considered.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x} - x)^2}{n}} \qquad \text{[Eq. 20]}$$

4. Results

In this chapter all the results obtained from the analyses conducted will be reported. The starting point is the choice of the baseflow separation method to be used. Then, all the results about spatial and temporal variability of runoff coefficients will be reported, through plots, box-plots and maps. Some of them will regard the 76 catchments dataset, others, instead, singular basins. For this reason three representative catchments have been selected in order to explain the results obtained.

4.1. Baseflow separation

The proper separation method the work bases on was chosen considering how baseflow fitted the discharge pattern in the hydrograph. This was done for each catchment analyzing hydrographs with different baseflow separation methods implemented, and trying to highlight also the geographical position in which one method was more suitable than others. By using the data available, all methods were compared and the one that fitted hydrographs in the largest number of catchments was chosen.

4.1.1. Examples of baseflow pattern in different UK regions

The central part of Scotland, in particular the lands of Aberdeenshire, Tay, Tweed, Forth, is quite rainy and characterized by gauged catchments with different sizes, from 31 km^2 to 1500 km^2. Nonetheless, the pattern of baseflow is similar for each basin, and quite well represented by the UKIH method. It is the most balanced between the others: Chapman and Maxwell (CM) baseflow results quite low, while Lyne and Hollick's, instead, presents too many high peaks, even if the general trend is similar to the UKIH. In Figure 13 is shown a particular of catchment 94001, placed in the North of Scotland. It is possible to notice that CM baseflow is too low: a single event should be defined as the interval within the interception points between streamflow and baseflow. Hence it is not credible that all the hydrograph represented in Figure 13 belongs to a single event. On the other hand, LH method overestimates the baseflow, considering its high peaks.

Catchments placed in Yorkshire lacks of flow data, so that the best method to assess a possible baseflow is given by UKIH.

Flow data series change in the lands of East Anglia and in Thames catchments in which the values of discharge are relatively low in relation to flows observed in Scotland and Yorkshire. It is possible to say that the largest contribution to the total streamflow is given by the baseflow component. Also in this case the UKIH method seems the more reliable, while CM method is represented by a too low baseflow. It is difficult to decide whether UKIH method is better than LH because their trend is quite similar, except for some peaks in which LH probably overestimates the baseflow. In Figure 14 is shown a

particular of the total streamflow and different baseflow methods for a catchment placed in a East Anglia basin.

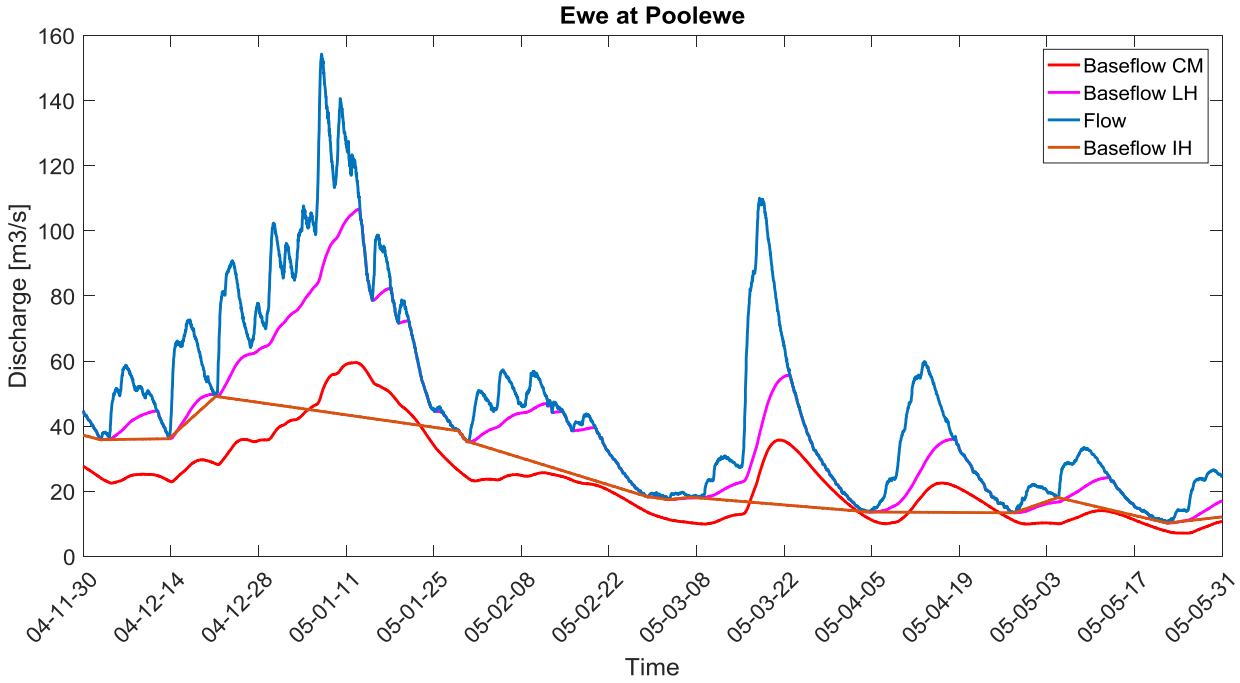

Figure 13: Zoom of streamflow and baseflow methods for catchment 94001, Ewe at Poolewe, Scotland.

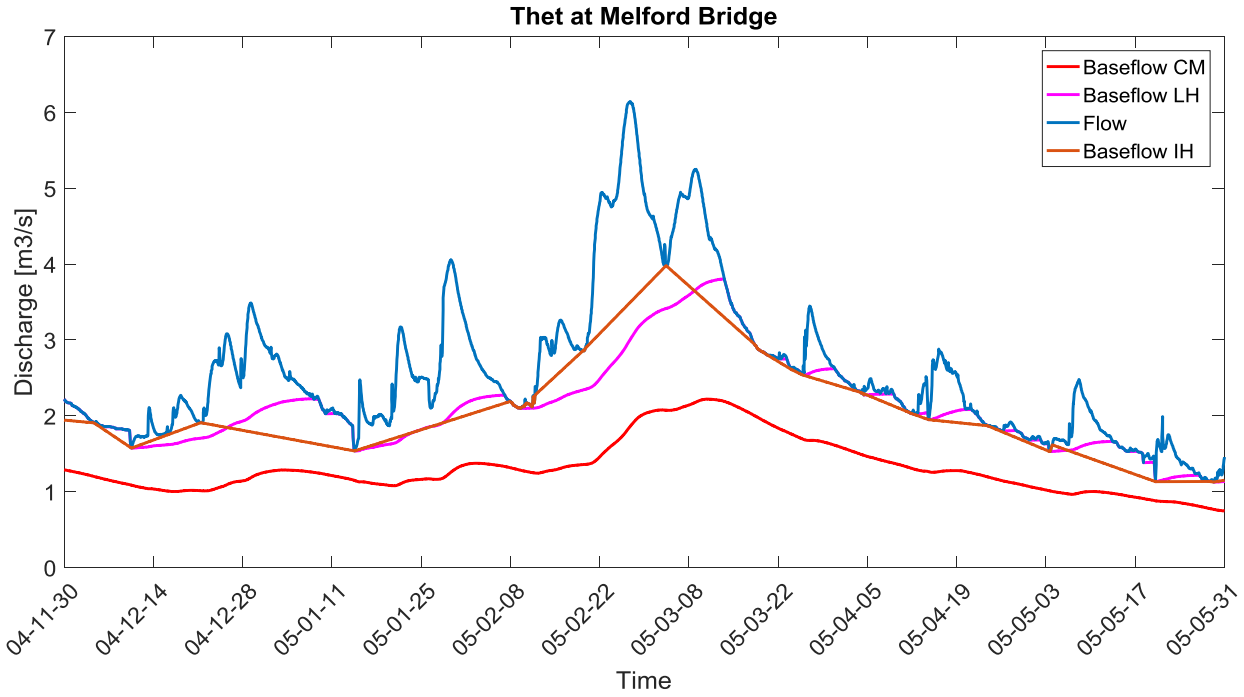

Figure 14: Zoom of streamflow and baseflow methods for catchment 33019, Thet at Melford Bridge.

The UKIH baseflow separation method appears to be the most suitable also for catchments placed in the North, South of Scotland and in Wessex. The only exception is given by basins of Cumbria and Lancashire where the Chapman and Maxwell method seems to give the most reliable trend of baseflow. Other catchments placed in Wales and a few in South Scotland are also well represented by the Chapman and Maxwell method, as it is shown in Figure 15. It is clear that between the three methods, CM is the most balanced, it does not overestimate the baseflow as LH method, and, on the other hand, in some peaks it does not underestimate it as the UKIH method. However, events separation considering CM or UKIH is quite the same.

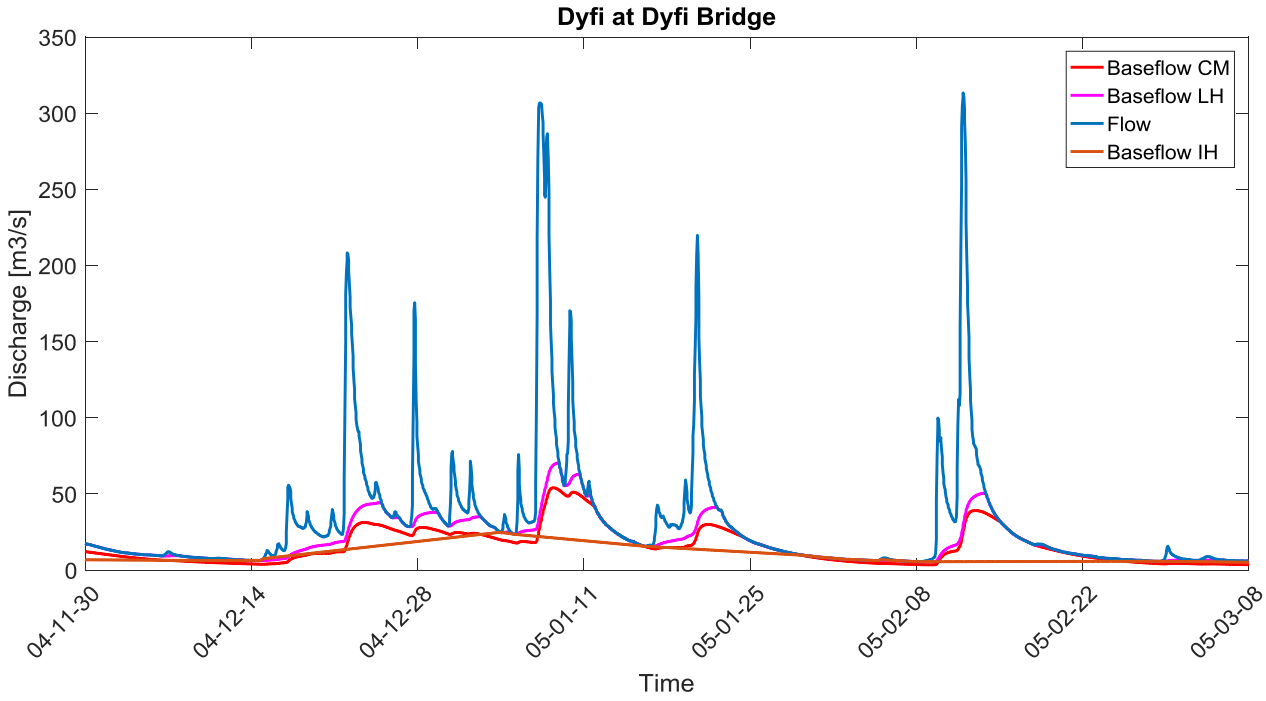

Figure 15: Zoom of streamflow and baseflow methods for catchment 64001, Dyfi at Dyfi Bridge.

4.1.2. Baseflow separation method chosen

The choice of a separation baseflow method is quite difficult because there is no certainty about its reliability. Digital filters bases on a constant that has been derived from catchments hydrograph, taking into account, in a small part, the rainfall-runoff behavior of the catchment. However, the noise that affects discharge data is too high to consider these values reliable. In Figure 16, for a generic catchment placed in Yorkshire, the time rate change in discharge $-dQ/dt$ versus the discharge Q is reported. It is clear that recession constants obtained from recession curves affected by such a noise are not considerable reliable. Rupp and Selker (2006) proposed a method to decrease the noise that affects the analysis which focuses on the time increment Δt. They suggest to change the length of Δt, considering the precision of data and the degree of noise. The time increment should be taken large enough to detect the signal being sought, but no so large to overwhelm it [Rupp et al., 2006].

Nevertheless, this correction method appeared to be onerous to be applied considering the short time available to develop this work. Hence, for different reasons the UKIH method has been chosen to separate baseflow from total streamflow.

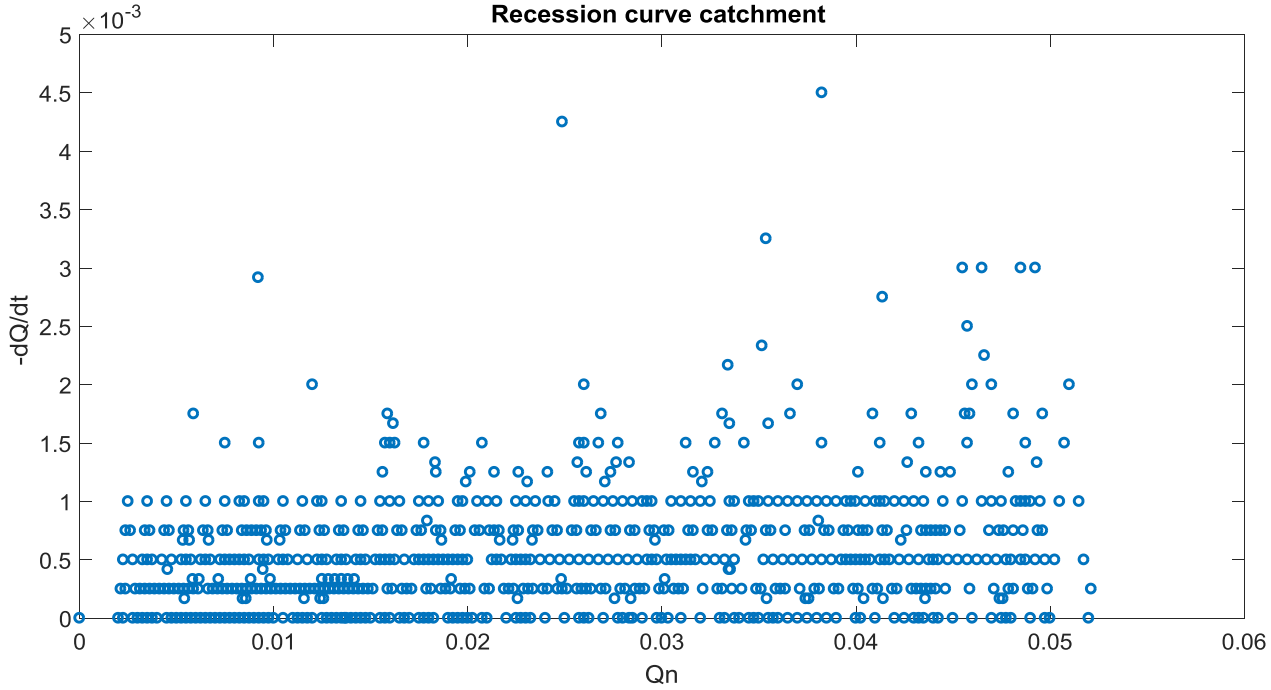

Figure 16: Effect of noise in recession analysis.

First of all, it fits better than the other methods for the majority of the catchments. Even though in some cases it underestimates the baseflow compared to other methods, it is considered a safety approach, because the amount of rainfall that becomes quickflow is considered larger. Furthermore, its uncertainty is not increased by the uncertainty of parameters required by the methods, as it happens with digital filters.

4.2. Three representative catchments

Three representative catchments of runoff regime among UK have been chosen in order to investigate around different ranges of runoff coefficients. In Figure 17 a map showing the location of the three catchments is reported. They belong to different UK regions: Thet at Melford Bridge belongs to East Anglia; Dyfi at Dyfi Bridge belongs to eastern Wales; Nevis at Claggan instead belongs to the North of Scotland. In the next three sections they will be analyzed in term of runoff coefficients ranges in relation to quickflow volume, precipitation volumes, runoff event duration and number of events with certain characteristics.

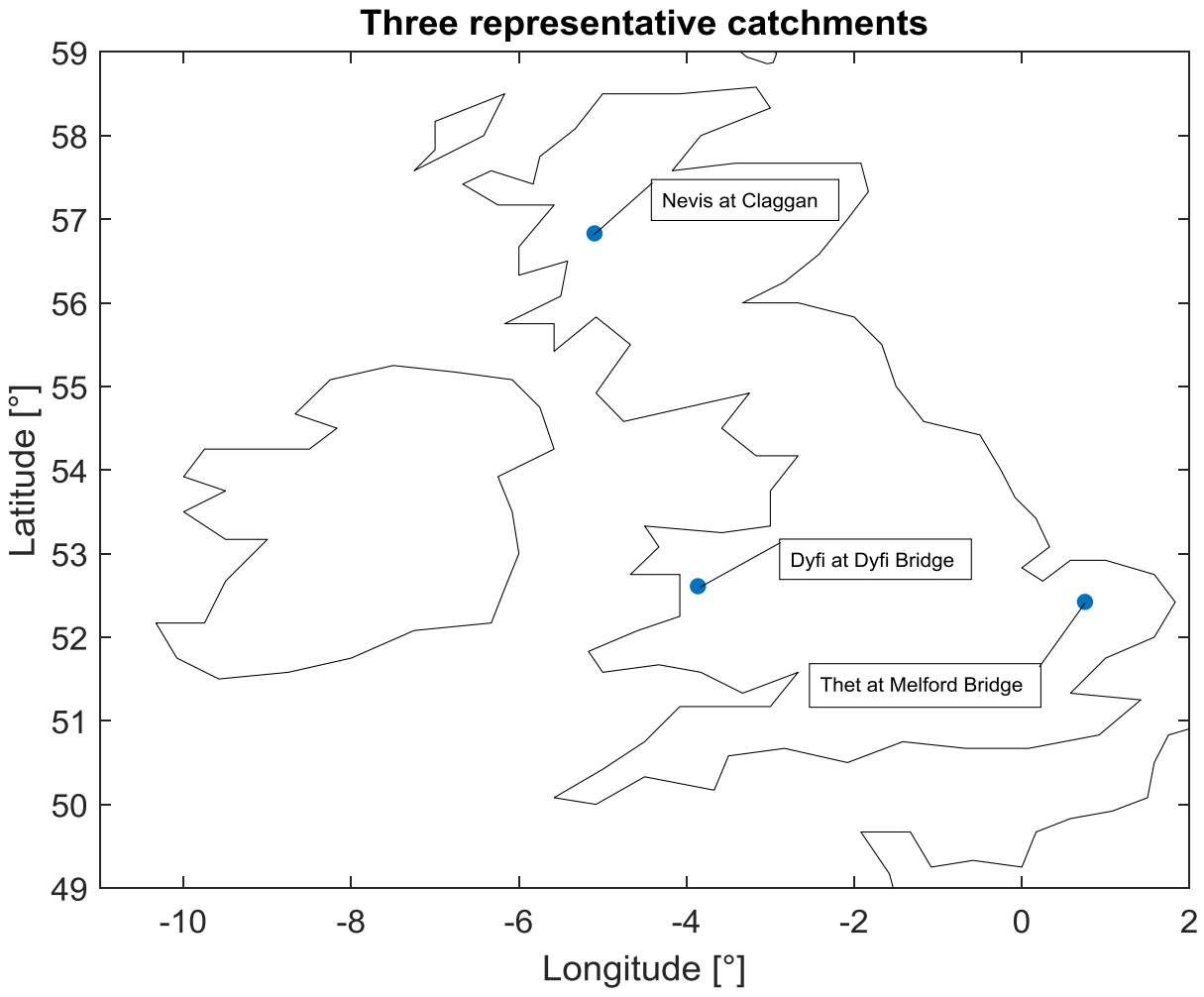

Figure 17: Placement of three representative catchments.

4.2.1. Thet at Melford Bridge

Thet at Melford Bridge is a catchment placed in the East Anglia, with an area of 316 km^2. It is predominantly a Chalk catchment largely overlain by Boulder Clay. The land use is primarily arable land, with some grassland and forestry.

The mean runoff coefficient for this catchment is quite small, equal to 0.0538. The number of events that characterizes this catchment is 290. The low values of runoff coefficients observed, are due to the morphology of the soil. Chalk soils in fact are soft, porous sedimentary carbonate rocks, that have the capacity to retain large volumes of stored water. For this reason rainfall tends to infiltrate as it falls on the ground, with the result of small amount of runoff on the surface.

It is interesting to show some results about correlation between runoff coefficients, volume of precipitation, quick flow and event duration.

In Figure 18 runoff coefficients values are presented in relation to different classes of precipitation. It is clear that larger volumes of precipitation leads to bigger values of runoff coefficients. In this case, even if this pattern is clear, median runoff values remain overall small. The largest number of events belong to the class with precipitation volumes smaller than 50 mm (67 events within 3mm-10mm range, 185 events within 10mm-50mm range). Only few events belong to the last class, showing that events with precipitation bigger than 100 mm do not occur often.

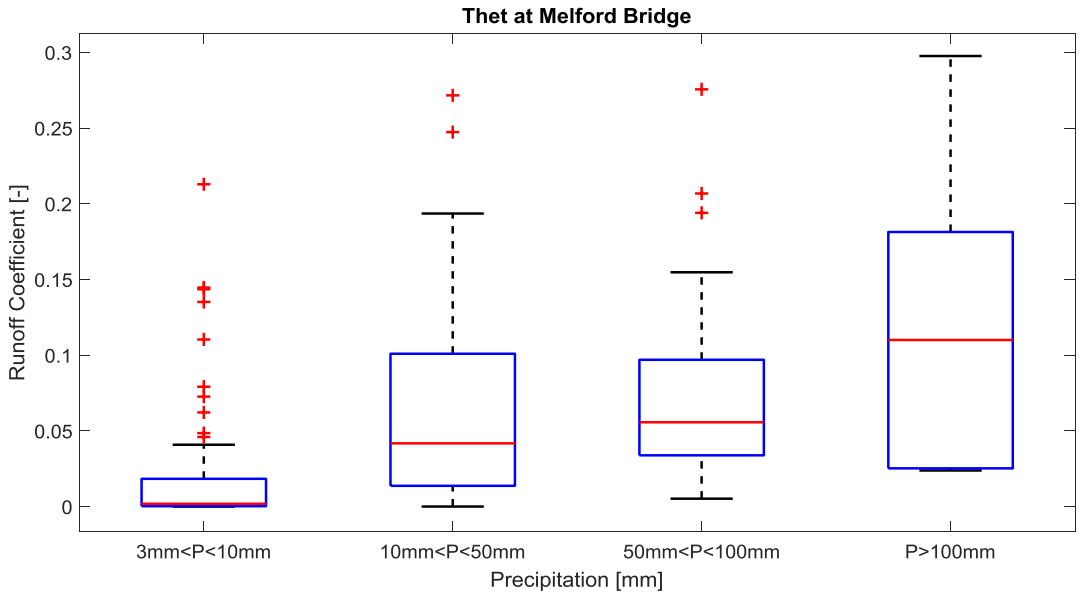

Figure 18: Runoff coefficients versus volume precipitation.

In Figure 19 the relation between runoff coefficients and volume of quickflow is reported. In this case the increasing pattern that is attended is more evident. The more the quickflow is present in a channel, the larger is the runoff coefficient. As attended, considering the low values of runoff coefficients, also quickflow tends to be small with a volume that reaches maximum 50 mm. The largest number of events, though, belong to the first two classes: 226 events have a quickflow smaller than 3 mm; 58 events have a quickflow volume between 3 and 10 mm.

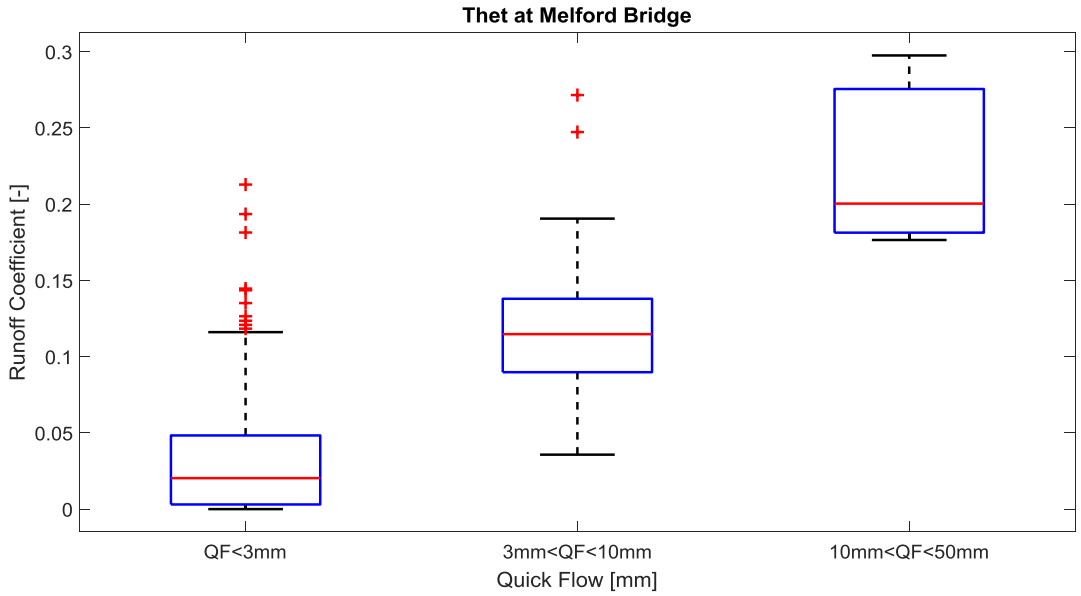

Figure 19: Runoff coefficients versus quickflow volume.

In Figure 20 the relation between runoff coefficients values and event duration is presented. Also in this case the pattern between runoff coefficients and event duration is increasing as the number of events that belongs to each class. The largest number of events has duration between 144 and 288 h (97 events) and larger than 288 h (103 events). The overall mean value of runoff coefficients, as already said is lower than 0.1.

Figure 21 shows the linear relation between precipitation and quickflow. As precipitation increases, also quickflow increases. The largest number of events though are related to small values of precipitation and quickflow, as underlined by the previous box plot reported above.

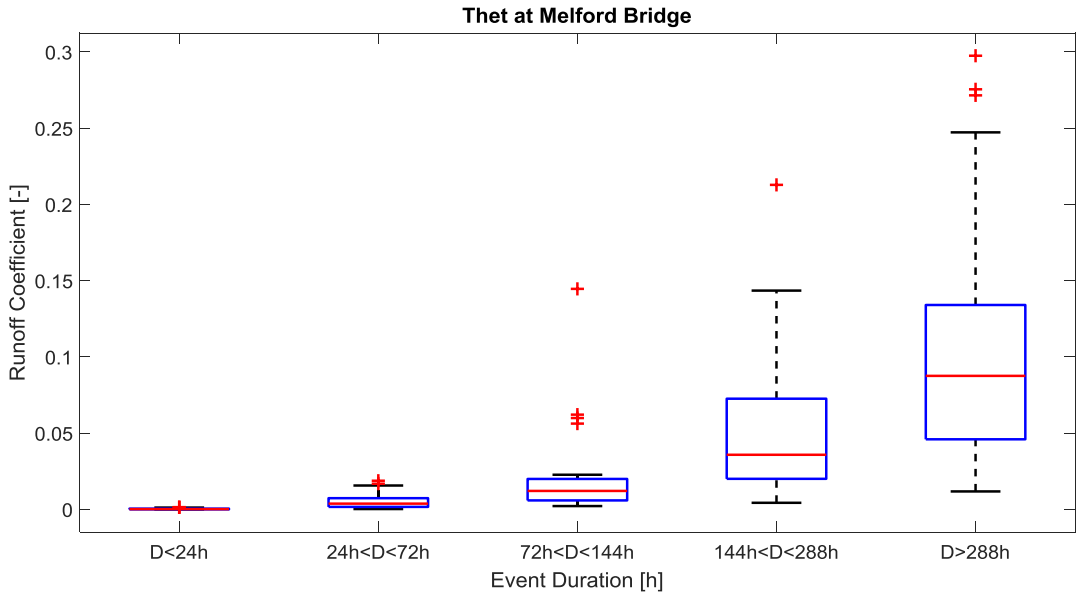

Figure 20: Runoff coefficients versus event duration.

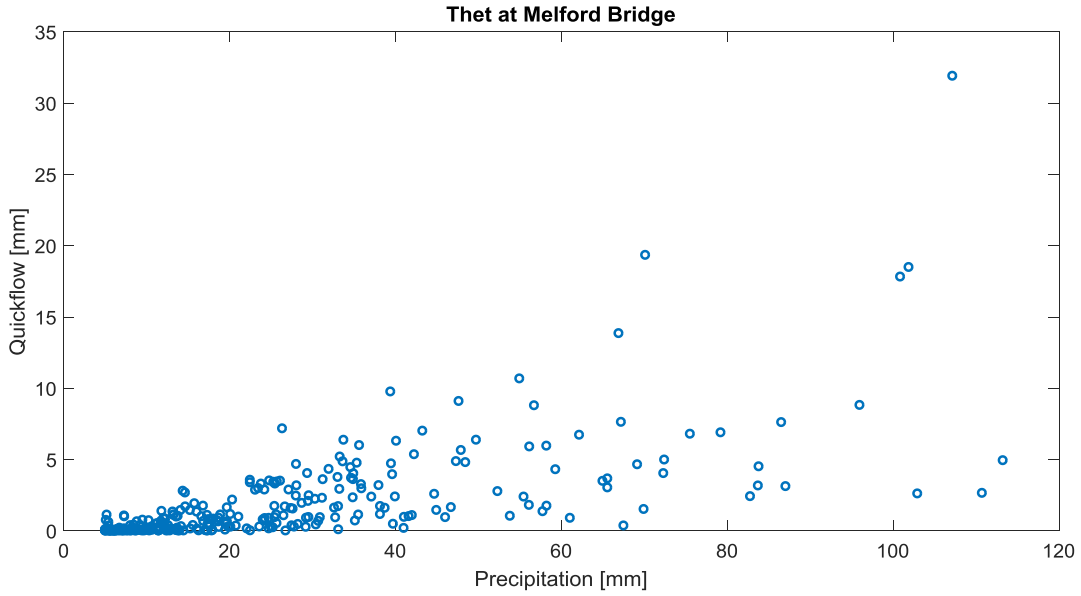

Figure 21: Quickflow versus precipitation.

Figure 22 reports runoff coefficients versus precipitation. In this case, though, dots representative of each event are colored in relation to the month they belong to. The trend is increasing, as already shown by the respective box-plot, but no particular pattern was found in relation to the month of each event.

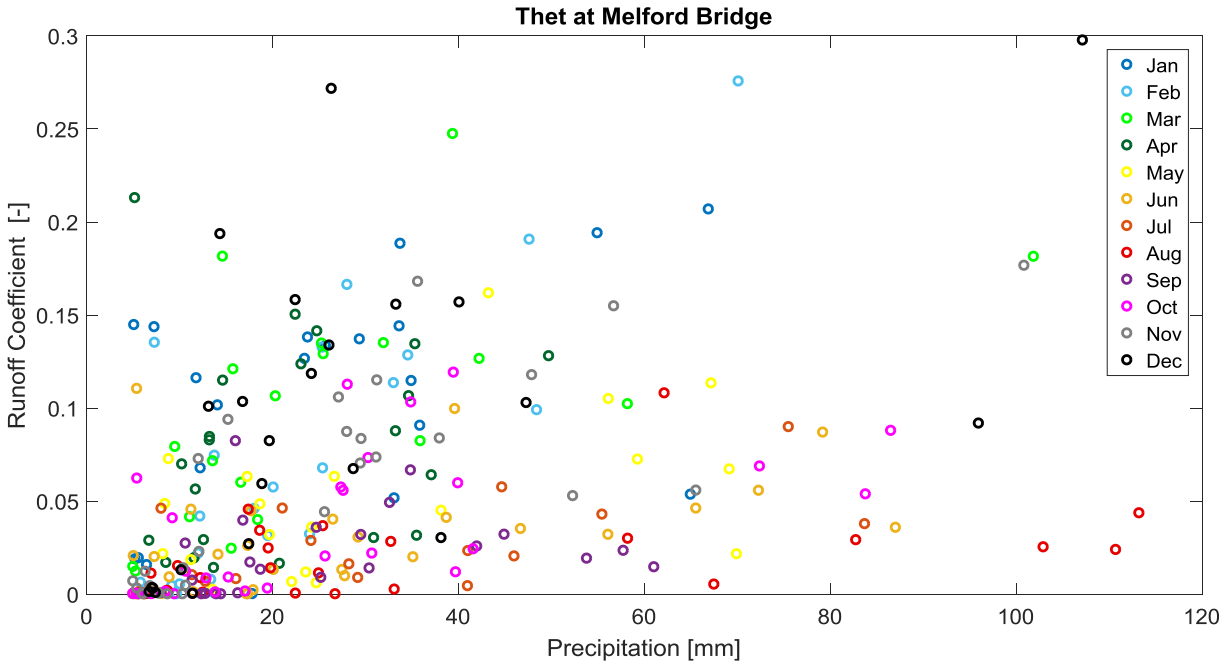

Figure 22: Runoff coefficients versus precipitation. Different colours of dots represent the months of the year.

4.2.2. Dyfi at Dyfi Bridge

Dyfi at Dyfi Bridge is a catchment placed in Wales with an area of 471.3 km^2. It is mainly characterized by impermeable Silurian formations, minor Boulder Clay and alluvium deposits. The catchment is mostly grassland or moorland, with extensive areas of forest.
Impermeable soil is the main cause of larger runoff coefficients. As precipitation falls on the ground, it tends to become runoff, skipping or reducing the infiltration process that in permeable or unsaturated soils occurs. The mean value in this case is 0.2873, evaluated considering the 289 events studied in this catchment.

In Figure 23 runoff coefficients values are presented in relation to different ranges of precipitation. The expected increasing pattern between precipitation and runoff coefficient values is respected, with a median runoff coefficient of 0.6 for precipitation bigger than 100 mm. In this case the number of events that belongs to each class in term of volume precipitation is quite balanced: 115 events are characterized by rain volume between 10 mm and 50 mm, 62 between 50 mm and 100 mm and 80 larger than 100 mm. Only 32 events are characterized by a precipitation smaller than 10 mm. It is possible to affirm, hence, that rainfall in this catchment is quite high.

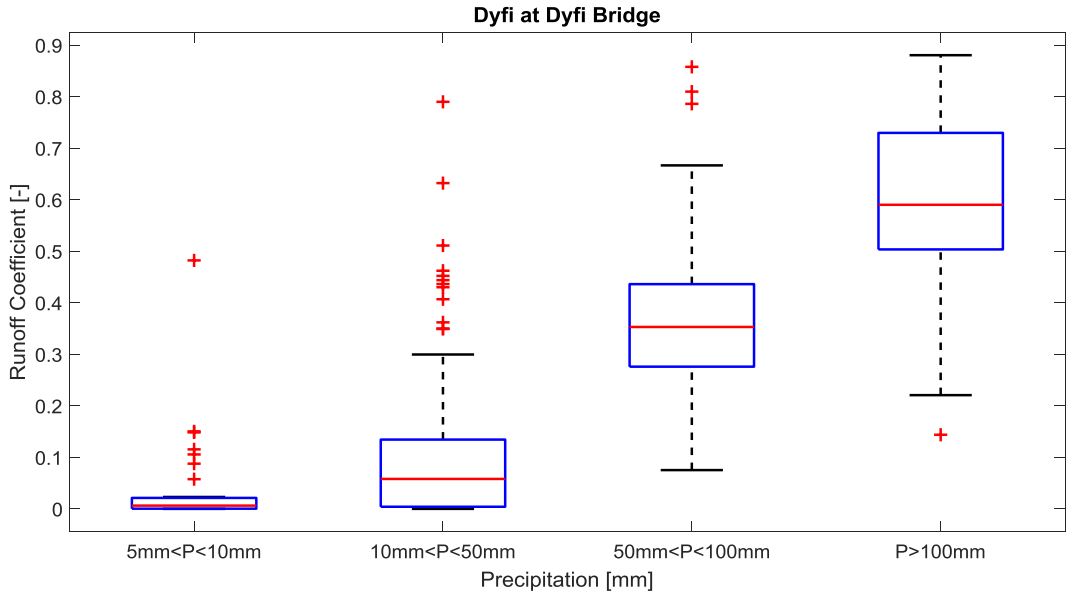

Figure 23: Runoff coefficients versus volume precipitation.

In Figure 24 the relation between runoff coefficients and volume of quickflow is reported. The attended increasing pattern between them is respected. The median runoff coefficient for the highest class of quickflow stands between 0.7 and 0.8. As the precipitation volume, also quickflow appears to be quite

high as runoff coefficients increase. Over 100 mm of quickflow are reached in this catchments, but they do not occur so frequently. In fact, looking at the number of events that belongs to each range in which quickflow has been divided, 108 events have a quickflow smaller than 3 mm; 30 events between 3 mm and 10 mm; 77 events between 10 mm and 50 mm; 36 events between 50 mm and 100 mm, 38 events over 100 mm. Therefore, only really high precipitation creates a quickflow with volume over 1 m, otherwise values remain mostly within 3 mm and 50 mm.

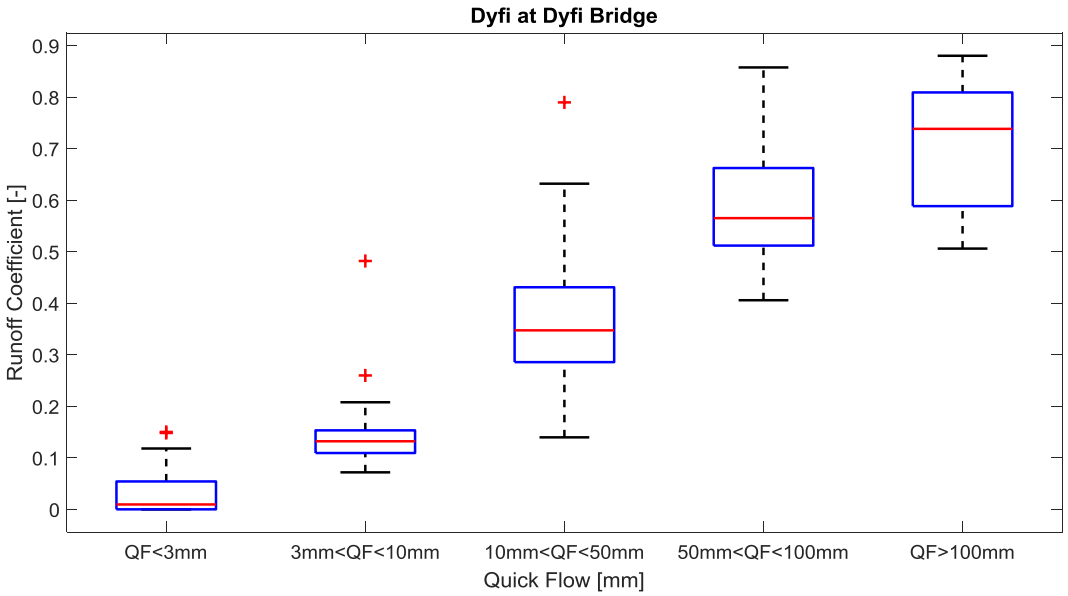

Figure 24: Runoff coefficients versus quickflow volume.

Figure 25 relates runoff coefficients values with events duration. The pattern is increasing as it is expected. It is possible to underline that short events are related to low runoff coefficients: for events shorter than 6 days the median runoff coefficients is lower than 0.1, for longer events it reaches values around 0.3 and 0.5. Considering the number of events present in each class of duration created, it is possible to notice that a little amount of them has a length smaller than 4 days (45 events lasted less than 1 day, 22 events lasted between 1 and 3 days, 24 events lasted between 3 and 6 days). The largest number of events, instead, usually lasts more than 6 days (73 events lasted between 6 and 12 days, 125 events lasted more than 12 days).

The relation between quickflow and precipitation is reported in Figure 26. It is clear that the pattern is linear, with a concentration of precipitation events between 10 and 50 mm. As the precipitation increases though, quickflow increases slower. This plot hence summarize the results obtained from box plots.

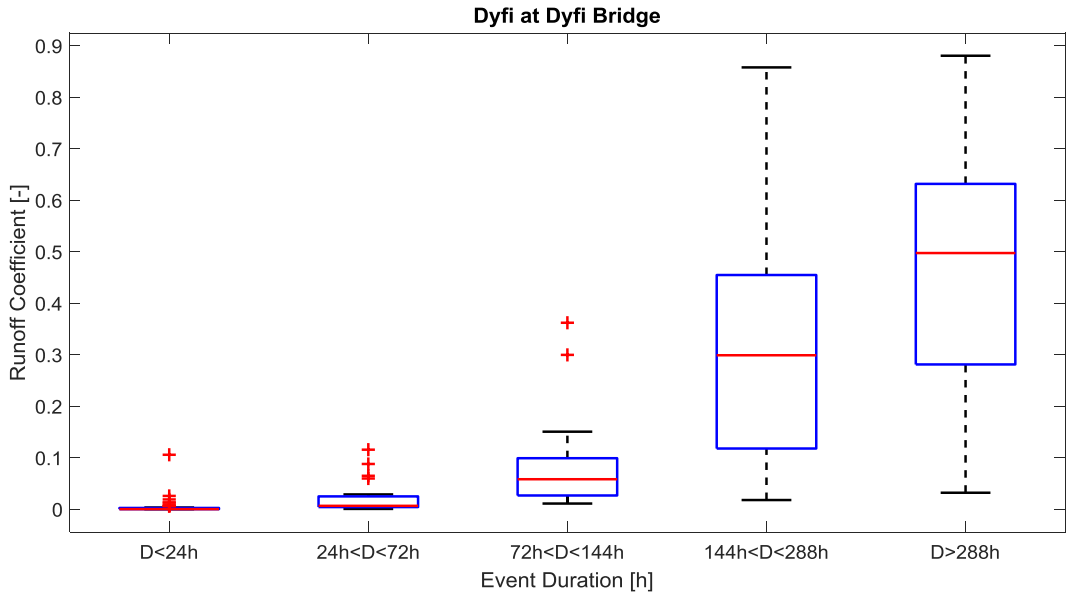

Figure 25: Runoff coefficients versus event duration.

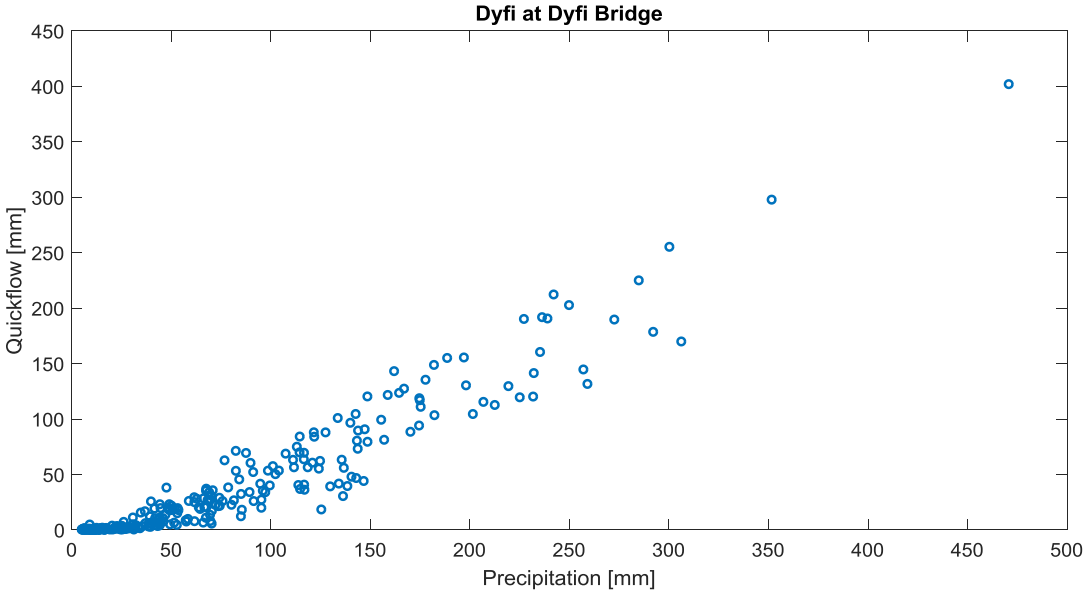

Figure 26: Quickflow versus precipitation.

Figure 27 reports runoff coefficients versus precipitation. In this case dots representative of each event are colored in relation to the month they belong to. The trend is increasing as already shown by the respective box-plot, but runoff coefficients values increase faster with precipitation than in the previous catchment. Furthermore, it is not possible, also in this case, to show a correlation within months and runoff coefficients pattern.

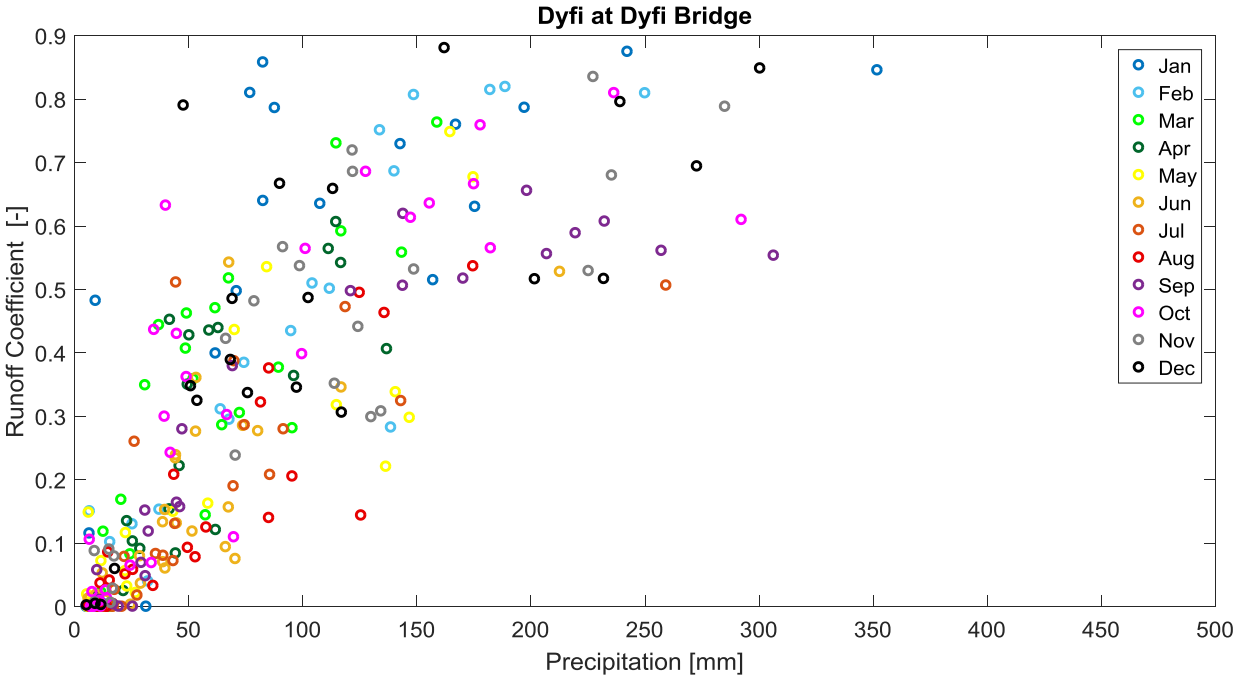

Figure 27: Runoff coefficients versus precipitation. Different colors of dots represent the months of the year.

4.2.3. Nevis at Claggan

Nevis at Claggan is a catchment placed in the North of Scotland, with an area of 69.2 km^2. It is a wet, steep-sided, high altitude catchment, draining southern slopes of Ben Nevis. It is characterized largely by impermeable bedrock with approximately 1/3 overlain by superficial deposits.

Also in this catchment, impermeable soil is the main cause of larger runoff coefficients. Infiltration processes are reduced or impeded, while rainfall produces runoff as it falls on the ground. The mean value of runoff coefficients between all runoff events is 0.4845, evaluated considering the 277 events studied in this catchment.

Runoff coefficients versus rainfall volumes are reported in Figure 28. The relation shows an increasing trend as it was expected. Runoff coefficients tend to be higher for events with precipitation larger than 50 mm showing median values around 0.6 and 0.7, while for smaller volumes median values are lower than 0.2. Taking into account the number of events present in each range of precipitation considered, it is possible to say that volumes larger than 100 mm occur more frequently. Only 28 events have precipitation between 5 mm and 10 mm, 74 events between 10 mm and 50 mm, 58 events between 50 mm and 100 mm, and 117 events have precipitation bigger than 100 mm.

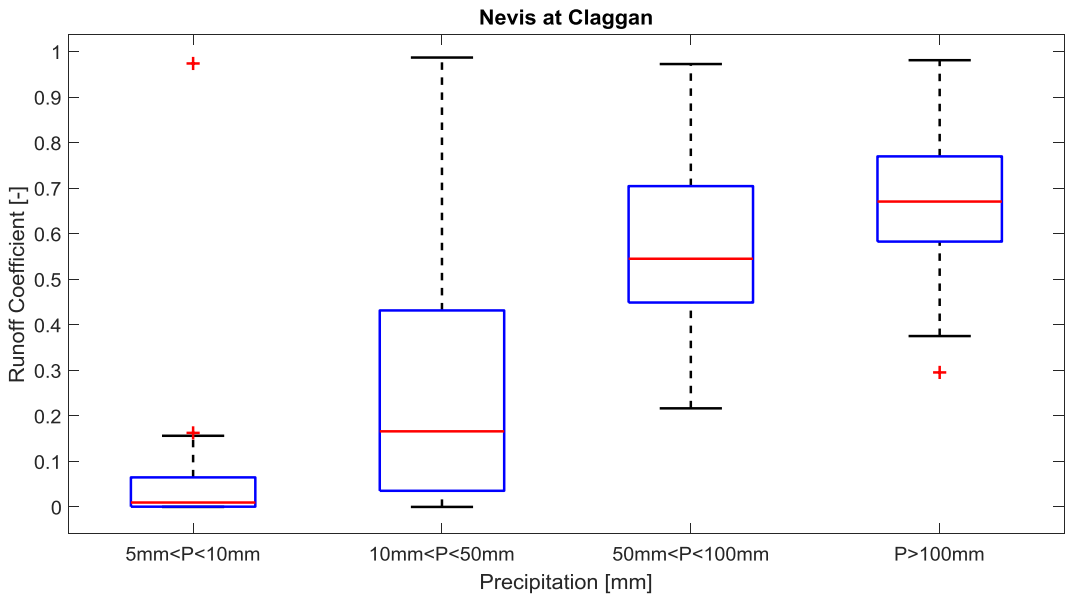

Figure 28: Runoff coefficients versus volume precipitation.

In Figure 29 the relation between runoff coefficients and quickflow shows an increasing linear pattern. As quickflow increases, runoff coefficients increases uniformly. The largest median runoff coefficient is around 0.75, and it occurs with direct flow larger than 100 mm. Taking into account the number of events that belong to the ranges created to study this relation, except for the class with quickflow between 3 mm and 10 mm, that contains only 13 events, the number of events in the remaining classes is balanced: 63 events with a quickflow smaller than 3 mm, 62 events with quickflow between10 mm and 50 mm, 64 events with a quickflow between 50 mm and 100 mm, and 75 events with quickflow larger than 100 mm.

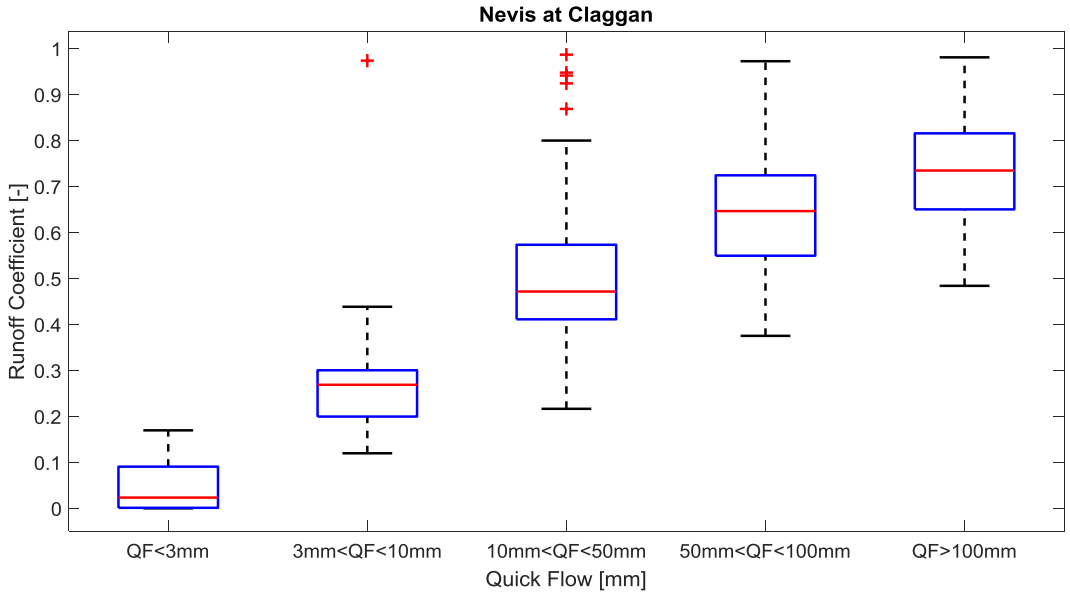

Figure 29: Runoff coefficients versus event quickflow volume.

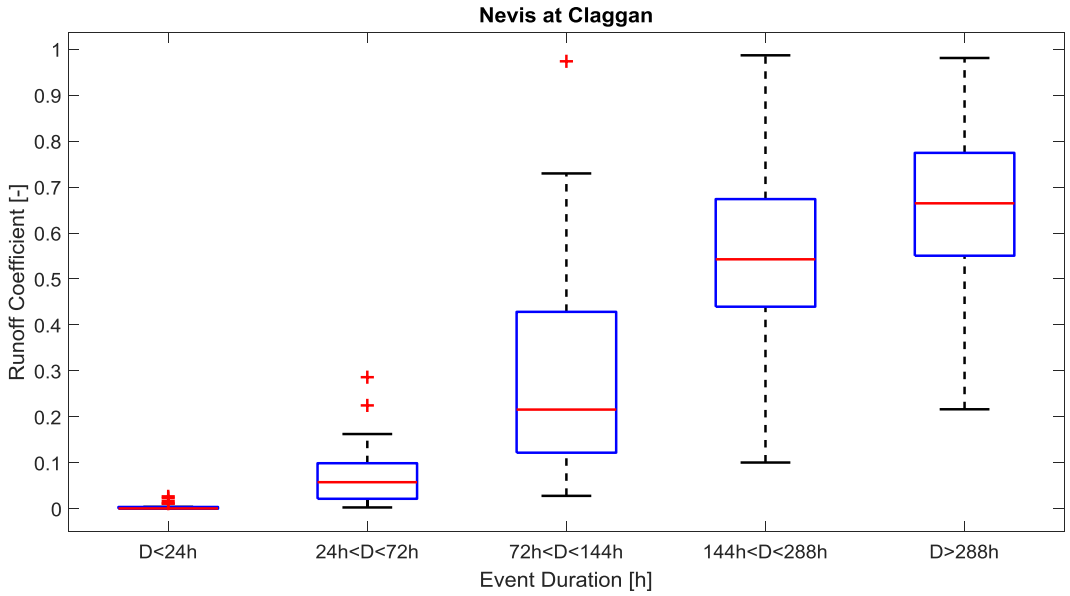

Figure 30: Runoff coefficients versus event duration.

In Figure 30 the relation of runoff coefficients with event duration is reported. The pattern is still linear, but a gap between median values of runoff coefficients occur considering events shorter and larger than 6 days. This gap is underlined also considering the number of events present in the different classes created to study event duration. 27 events are shorter than 1 day with a median RC value of 0.01, 24 events are between 2 and 3 days of duration with a median RC of 0.05, 28 events are between 3 and 6 days of duration with a median RC of 0.2. With larger events, the number of events for each class and their median RC increases: 65 events have duration between 6 and 12 days with a median around 0.55, 133 events have duration larger than 12 days with a median around 0.7.

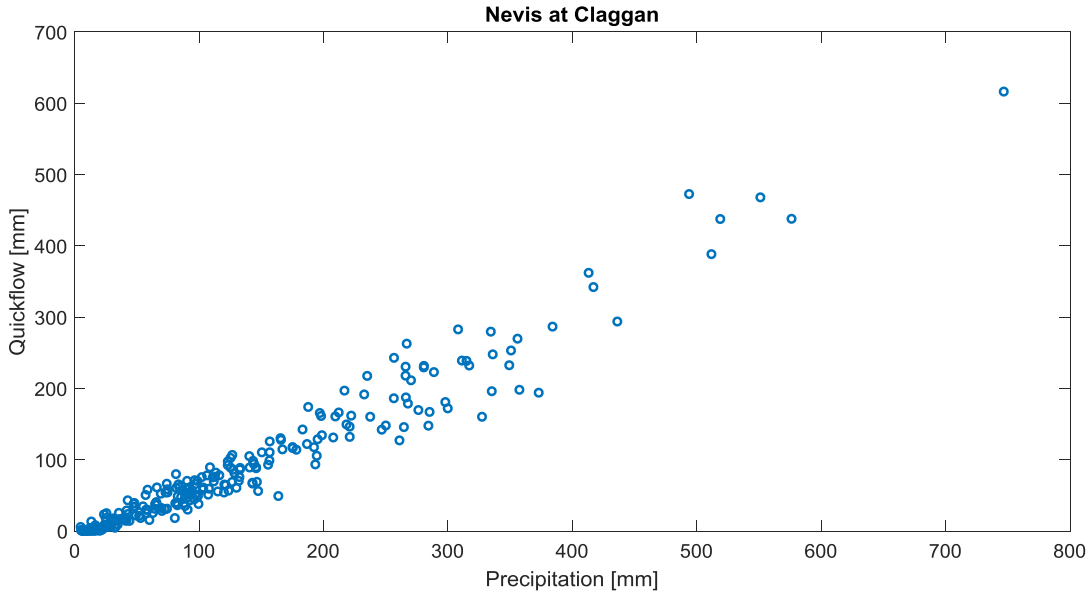

Figure 31: Quickflow versus precipitation.

Figure 31 shows precipitation versus quickflow. It summarizes the results obtained by the previous box plot. The trend is linear and increasing. This catchment is characterized by high precipitation with a maximum value of rainfall around 600 mm. Also in this case quickflow increases slower than precipitation, in fact if the half of events have precipitation bigger than 100 mm, only around 70 catchments have quickflow larger than 100 mm.

Finally in Figure 32 runoff coefficients versus precipitation are reported. It is clear, how impermeability influences the runoff response. In this catchment, more than in the catchments in Wales, runoff coefficients increase very fast at the increasing of precipitation.

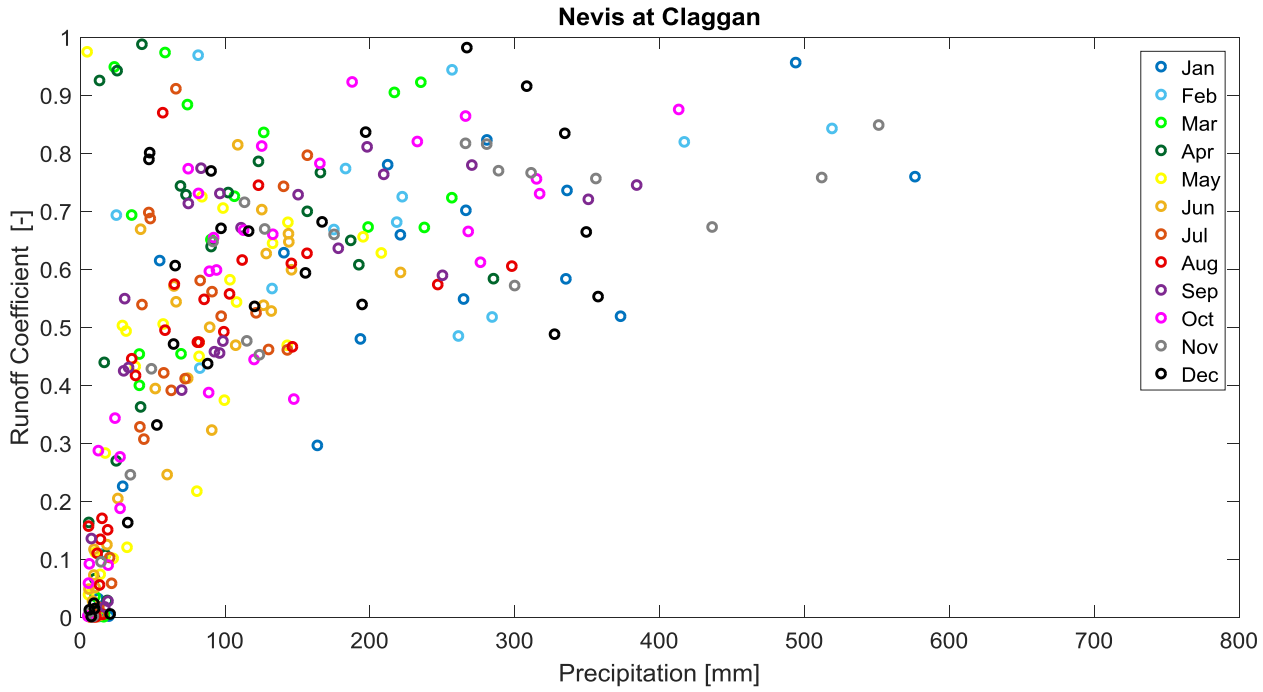

Figure 32: Runoff coefficients versus precipitation. Different colours of dots represent the months of the year.

4.3. Spatial variability

In this section results about spatial variability of runoff coefficients around the country will be reported. It is clear from the results obtained from each singular catchment analysis, that different ranges of runoff coefficients are due to hydrological causes. The three catchments analyzed belong to different regions with different characteristics in term of climate and soil type. The land use in not considered a relevant influencing factor in runoff coefficient results because all of the basins analyzed are benchmark, which means that they do not have significant areas affected by human changes such as urbanization or intensive farming. After a summary about the three catchments representative of runoff response in different regions, some consideration will be done considering all the dataset of 76 catchments.

4.3.1. Analysis of three catchments

Among the three catchments studied, Thet at Melford Bridge, placed in East Anglia, is characterized by the lowest runoff coefficients, with a maximum value that stands around 0.3. For each event the quickflow tends to be smaller than 50 mm, rainfall rarely exceeds 50 mm of volume and runoff events rarely exceed a duration of 12 days.

Dyfi at Dyfi Bridge is placed in the East part of Wales, and is characterized by intermediate values of runoff coefficients among the three basins considered. Its maximum value stands around the unity. For each event the quickflow tends to be smaller than 50 mm, most frequent values of precipitation are below 50 mm, but rainfall events tend to be longer than 12 days.
Nevis at Claggan is placed in the North of Scotland and it is characterized by the highest values of runoff coefficients. In this case the most frequent precipitation volumes exceed 100 mm, events duration tend to be mostly longer than 12 days. Quickflow among the dataset of events studied does not show predominant values among events: each class of volumes tends to have the same number of events.

Consideration about the different ranges of runoff coefficients are clearer looking at the cumulative distribution function of runoff coefficients showed in Figure 33. Thet at Melford Bridge has the 50% probability that runoff coefficients are lower or equal to 0.05; Dyfi at Dyfi Bridge has the 50% probability that runoff coefficients are lower or equal to 0.2; Nevis at Claggan has the 50% probability that runoff coefficients are lower or equal to 0.6.

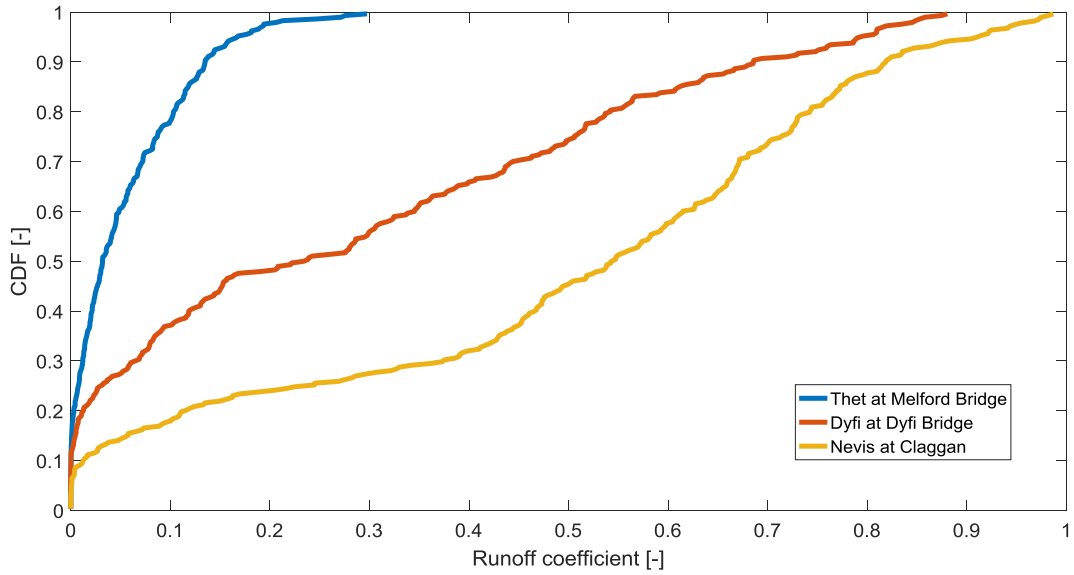

Figure 33: Cumulative distribution function of runoff coefficients for three representative catchments.

4.3.2. Analysis of 76 catchments

Spatial variability of runoff coefficients has been analyzed considering the descriptors whose meaning is explained in Chapter 2. Taking into account different ranges of soil properties and climate characteristics, values of runoff coefficient have been studied. The relation between runoff coefficients and these properties are being studied because they may provide a method for estimating runoff coefficients in ungauged basins of the UK.

In particular, the map represented in Figure 34 shows the spatial distribution of the mean value of runoff coefficient for each site studied. It is clear that in the south of UK runoff coefficients are the smallest. They tend to increase in the Midlands, Yorkshire and Wales, and they reach their highest values in the central-north part of Scotland. A qualitative analysis, hence, permits us to distinguish three regions: the south of UK with mean runoff coefficients smaller than 0.15; a second region made by the eastern Scotland, and central Wales with mean values between 0.15 and 0.35; a third region comprehensive of central and western Scotland with mean values larger than 0.35.

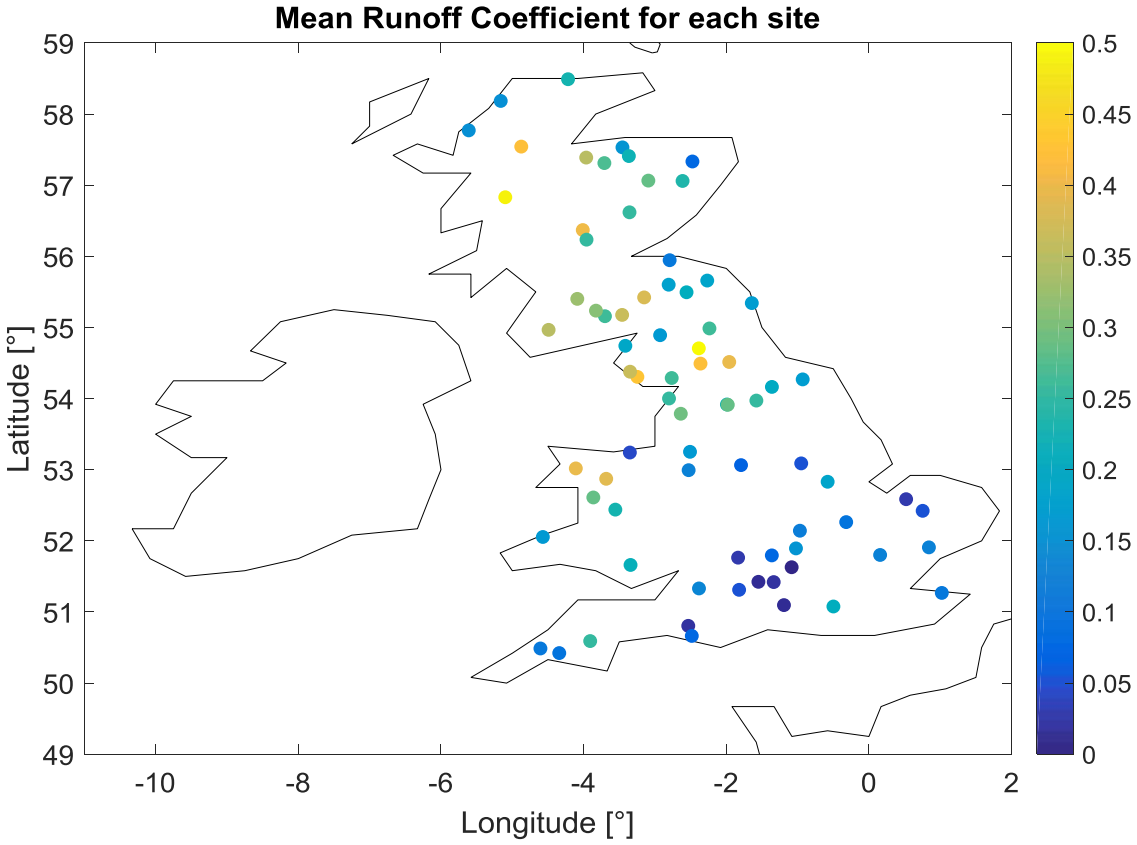

Figure 34: Mean runoff coefficient for each of the 76 sites analyzed.

4.3.2.1. PROPWET

PROPWET represents the wetness of the soil or rather the proportion of time in which the soil may be considered wet. Figure 35 shows the cumulative distribution functions of runoff coefficients, subdivided in relation to the values of PROPWET. The number of classes chosen to divide basins in relation to PROPWET, is the result of an iterative process done in order to maximize the distance between the cumulative distribution functions. Among UK it is possible to consider three different regions taking into account the wetness of the soil.

The region with a PROPWET smaller than 0.4 is placed in England and it comprehends the South East, London, East of England, East and West Midlands and the western part of the South West. As it is possible to notice from the CDF curve, there is the 80% of probability that runoff coefficients are lower than 0.2. The region with PROPWET between 0.4 and 0.6 regards mostly coastal places: it comprehends coasts in the South West and Wales, and eastern coasts of Yorkshire and the Humber, North East and Scotland. In this case from the CDF curve emerges a 80% probability to have values of runoff coefficients lower than 0.4.

The region with PROPWET larger than 0.6 regards western part of North West and Scotland, and the north-east part of Scotland. From the CDF curve is clear that for this region there is a 80% probability to have values of runoff coefficients smaller than 0.6.

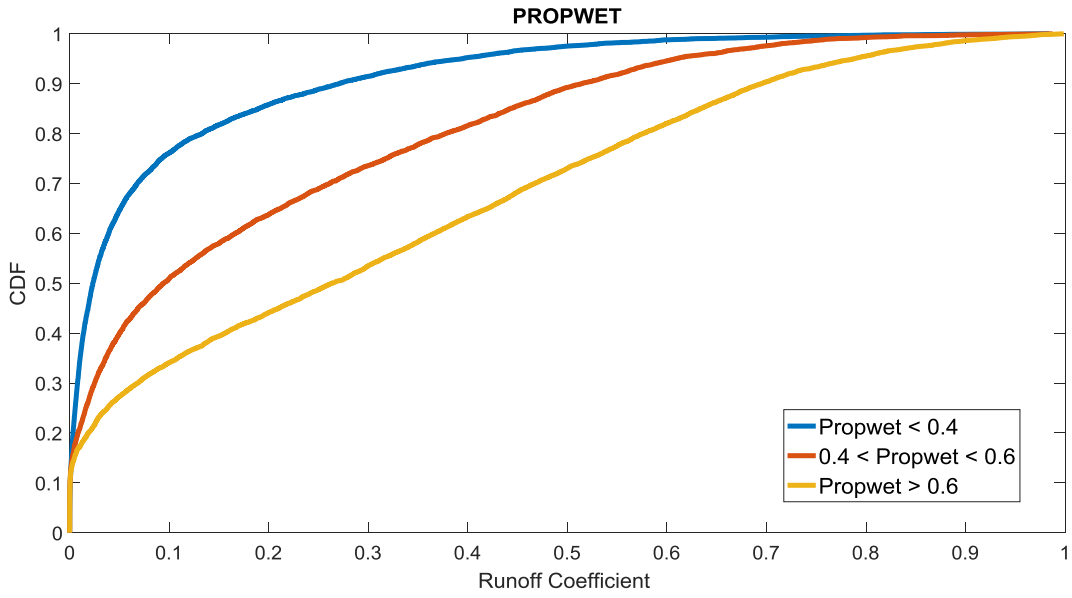

Figure 35: Distribution function of the event runoff coefficients of 76 catchments stratified by PROPWET.

Figure 36 shows a map in which each catchment is represented by a dot colored in relation to the PROPWET class it belongs to. Thet at Melford Bridge belongs to the first class of PROPWET, with a value of 0.31; Dyfi at Dyfi Bridge and Nevis at Claggan belong to the third class with values respectively of 0.66 and 0.81.

The results obtained are intuitive. Thet at Melford Bridge compared to the other two basins, has a smaller amount of precipitation. This facts reflects drier soils and so on, a smaller value of PROWET. This means that the Soil Moisture Deficit (SMD) was exceeded for the 31% of the time between 1960-90. The other two basins instead, characterized by higher precipitation, have soils wet for a larger

proportion of time: the SMD is exceeded for the 66% of the time for Dyfi at Dyfi Bridge, for the 81% of the time for Nevis at Claggan.

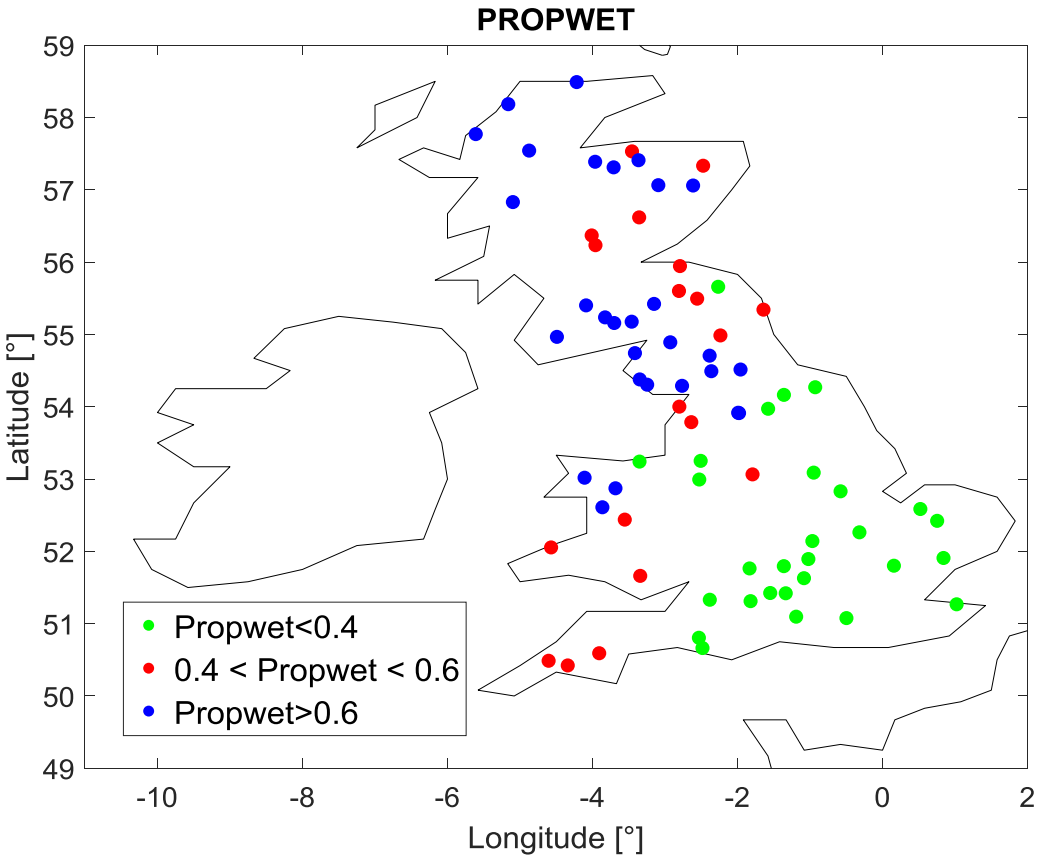

Figure 36: Map of PROPWET distribution for 76 catchments in UK.

4.3.2.2. DPSBAR

DPSBAR descriptor was analyzed in the same way. Cumulative distribution functions of runoff coefficients were realized considering ranges of this descriptor, in order to maximize the distance between the CDM curves and therefore to distinguish catchments with different properties. Also in this case three classes were obtained.

The region with DPSBAR<60 comprehends the central and eastern part of England (London, South West, East of England) and a few catchments placed in Yorkshire and North West. From Figure 37 emerges that there is a probability of the 80% to have runoff coefficients smaller than 0.1 for basins placed in these regions.

The region with 60<DPSBAR<200 regards the rest of the country except for some catchments placed in the North of Scotland and in the Western coasts of Wales that have a DPSBAR>200. The former have a probability of the 80% to have runoff coefficients smaller than 0.4, the latter instead smaller than 0.7.

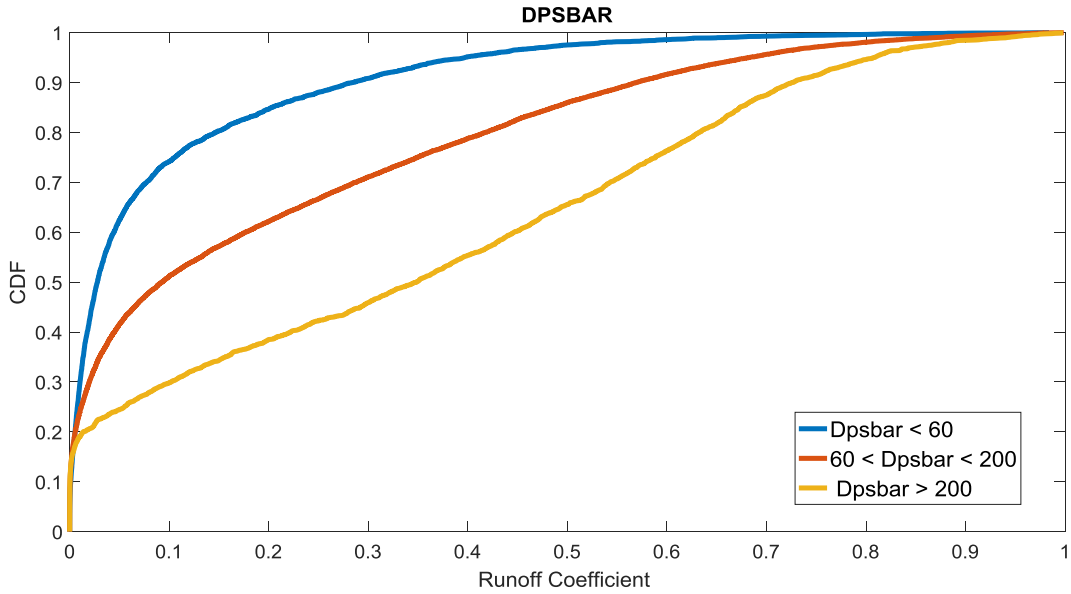

Figure 37: Distribution function of the event runoff coefficients of 76 catchments stratified by DPSBAR.

Figure 38 shows a map in which each catchment is represented by a dot colored in relation to the DPSBAR class it belongs to. Thet at Melford Bridge belongs to the first class of DPSBAR, with a value of 13.9; Dyfi at Dyfi Bridge and Nevis at Claggan belong to the third class with values respectively of 269.8 and 441.8.

The results obtained in term of runoff coefficients reflect what it is expected considering the morphology of the soil. Less steep catchments of the first region analyzed tend to slow down the flow, favoring the water to become delayed runoff, rather than direct runoff. This fact is combined to more permeable soils that facilitate the infiltration of water, decreasing the amount of flow that generates superficial runoff. The other two regions, instead, are characterized by steeper lands and impermeable soils, favoring the direct flow as the values of runoff coefficients show.

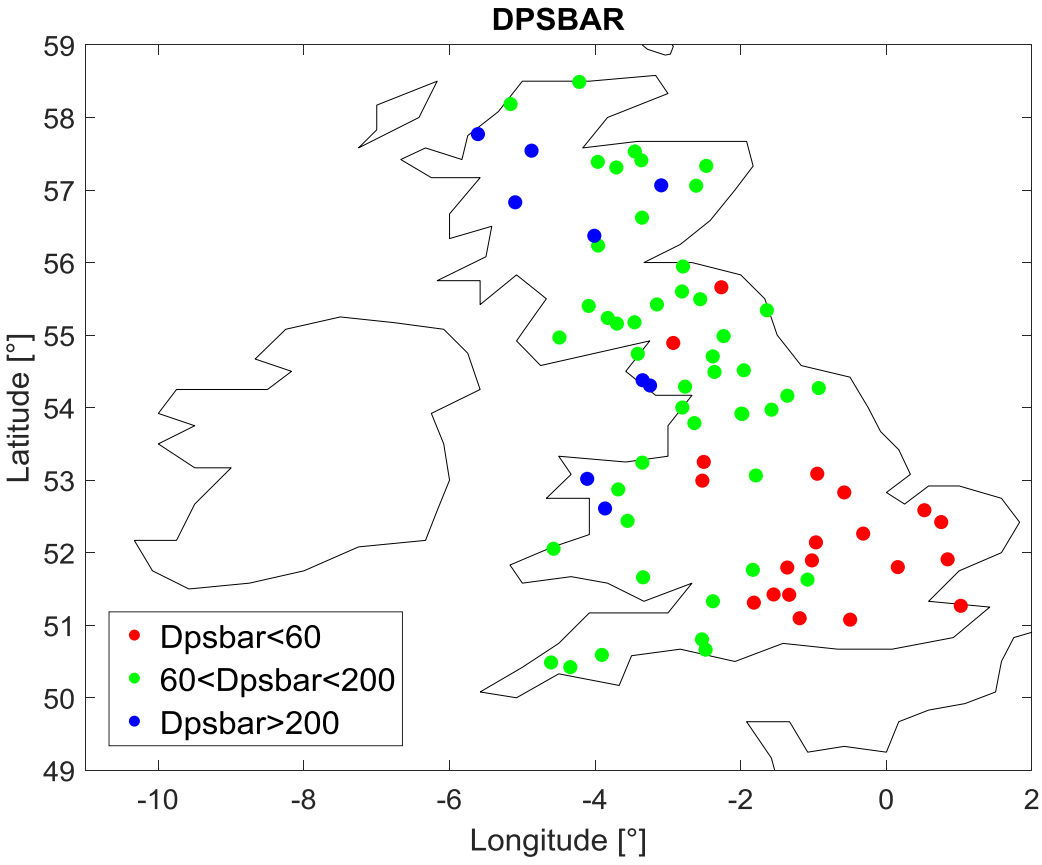

Figure 38: Map of DPSBAR distribution for 76 catchments in UK.

4.3.2.3. MAP

The mean annual precipitation descriptor was studied in the same way of the previous descriptors. Figure 39 shows the cumulative distribution functions obtained. The ranges of MAP that better distinguish UK regions are three.

The region with MAP smaller than 800 mm is mainly characterized by catchments placed in the East of England, London and South East. For these regions there is a probability of the 80% that runoff coefficients are smaller than 0.1.

The region with 800 mm <MAP<2000 mm comprehends all the rest of the country except for catchments in the West of Scotland, and in the coasts of North West and Wales that have MAP>2000 mm. There is respectively a probability of the 80% to have runoff coefficients smaller than 0.4 and 0.6.

In Figure 40 it is possible to notice in which class of MAP the three catchments belong. Thet at Melford Bridge belongs to the first class with a MAP or 713 mm; Dyfi at Dyfi Bridge and Nevis at Claggan belong to the third class with MAP values respectively of 2028 mm and 3275 mm.

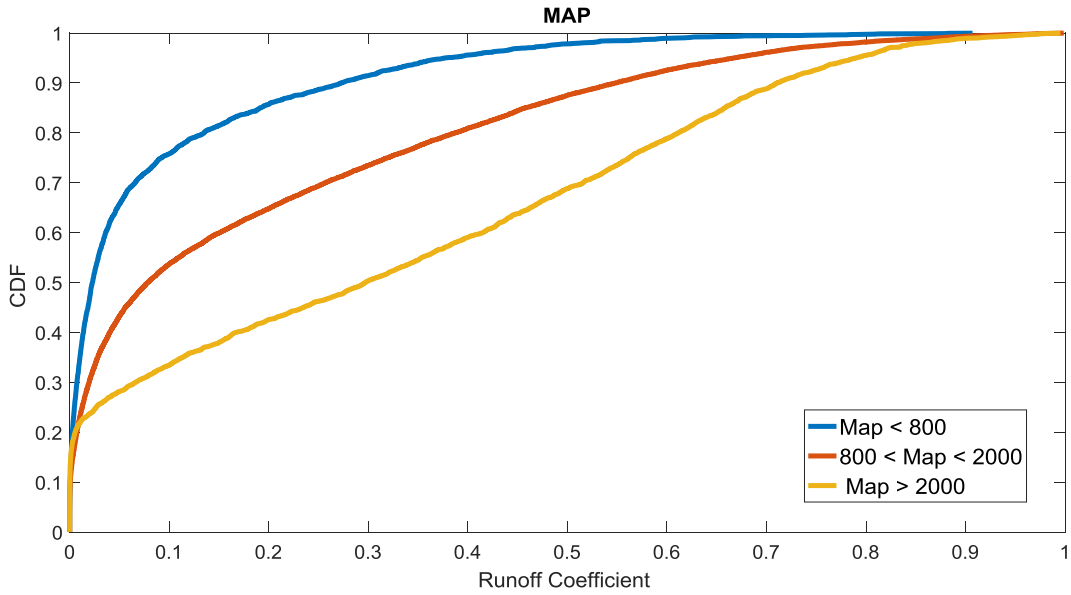

Figure 39: Distribution function of the event runoff coefficients of 76 catchments stratified by MAP.

The spatial distribution of mean annual precipitation agrees with results obtained for the other basins characteristics. More precipitation within a year implies more water on the soil and more direct flow. If soils are permeable they reach faster the saturation condition favouring the direct runoff, if they are impermeable certainly runoff will be increased.

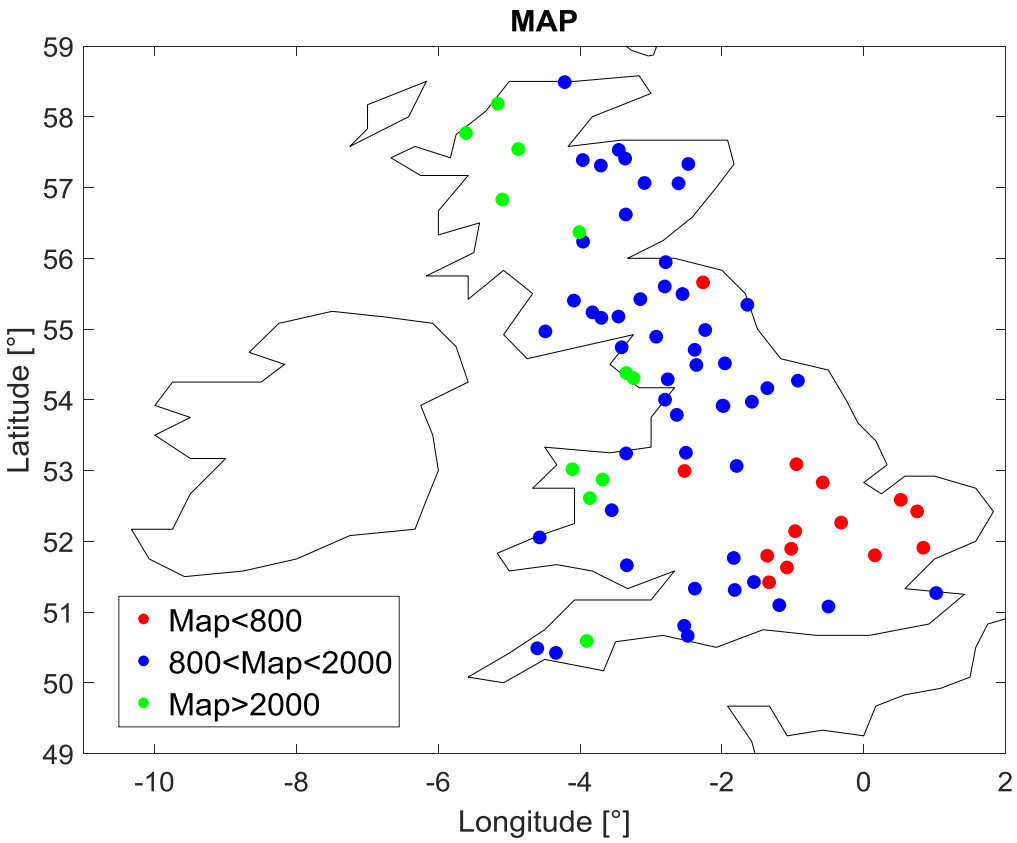

Figure 40: Map of MAP distribution for 76 catchments in UK.

4.3.2.4. BFIHOST

The analysis of BFIHOST among the 76 catchments reveals that UK can be divided in two big regions taking into account runoff coefficients values and BFIHOST descriptor of each basin.

Region with BFIHOST larger than 0.5 are placed mostly in the south of UK. Only few catchments of this category belong to Scotland and South West. Scotland and Wales are instead mostly characterized by BFIHOST smaller than 0.5. Looking at the cumulative distribution functions reported in Figure 41, there is a probability of the 80% to have runoff coefficients smaller than 0.5 for the former, smaller than 0.1 for the latter.

BFIHOST represents, as already said, the part of runoff that comes as baseflow taking into account the soil types. The kind of soil that characterizes the catchments is hence relevant for the amount of baseflow that grows inside the ground.
Thet at Malford Bridge, for example, has got a BFIHOST equal to 0.707. Being the soil made of porous Chalks, rainfall simply infiltrates in the ground creating a groundwater flow that, with a certain delay, will become runoff. Even if in this region rainfall has lower volumes, and smaller runoff, its main part is done by baseflow.
Dyfi at Dyfi Bridge and Nevis at Claggan have respectively a BFIHOST of 0.478 and 0.428. The values are similar, and reflect the impermeable soils of both the regions.

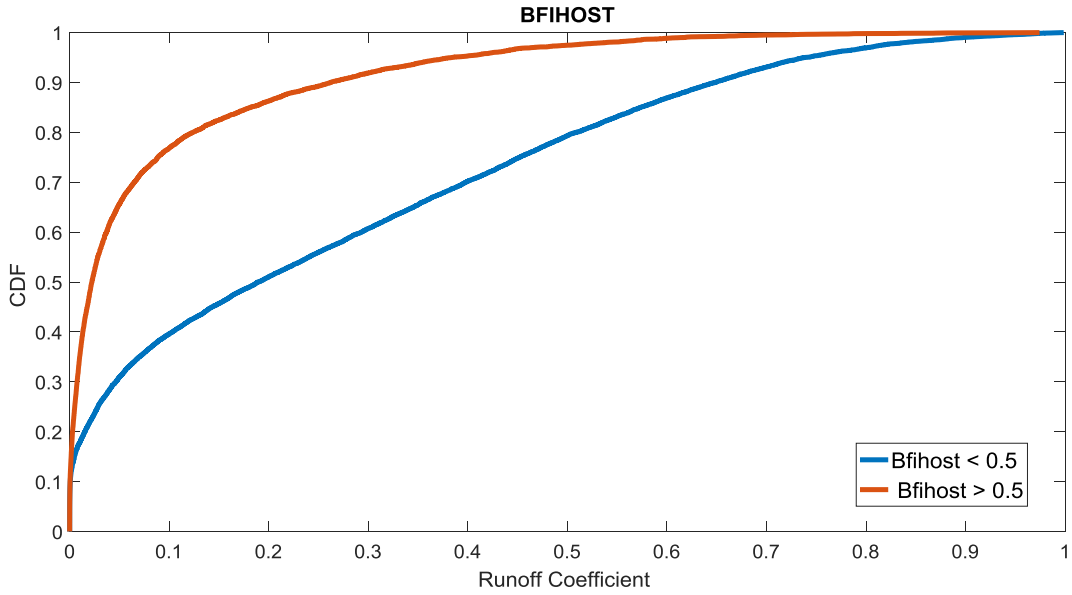

Figure 41: Distribution function of the event runoff coefficients of 76 catchments stratified by BFIHOST.

The two regions individuated are shown in the map presented in Figure 42.

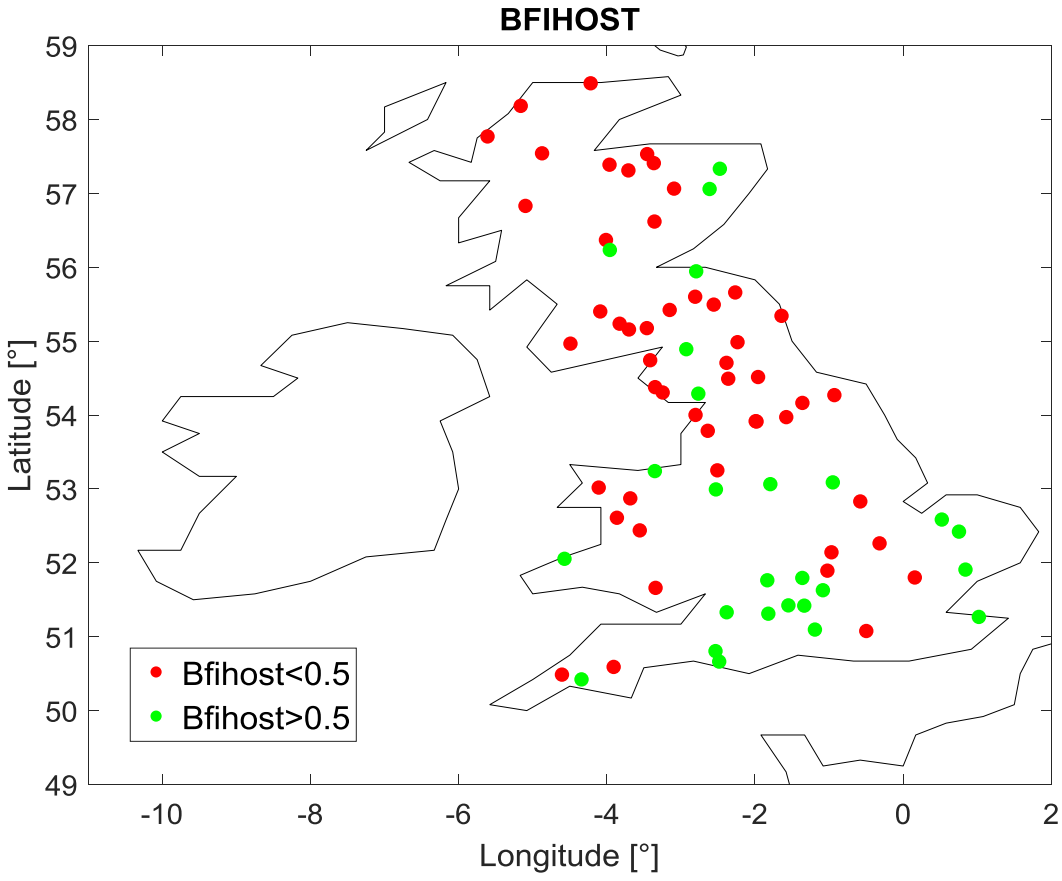

Figure 42: Map of BFIHOST distribution for 76 catchments in UK.

4.4. Temporal variability

In this section temporal variability of runoff coefficients will be discussed, firstly for the three catchments chosen as representative in this study, secondly for the whole dataset.

4.4.1. Analysis of three catchments

As already said in Chapter 3, temporal variability of runoff coefficients has been studied considering runoff coefficients versus months and their standard deviation related to their mean values.

In particular, this relation has been analyzed considering the mean value of runoff coefficients for each month and for each year of the data set. In the first case, the mean runoff coefficient for each month was calculated as the mean value of coefficients that belong to the same month of the ten years the dataset was formed. For all the events occurred in the month of January within the ten years of data considered, for example, the mean of the respective runoff coefficients was calculated and plotted against its standard deviation. Hence, the 12 dots fitted in this plot represent mean monthly runoff coefficients for each month. Secondly, the same was done calculating the mean value of runoff coefficients from events that happened in the same year. In this case, the study regarded a dataset of 10 years, so ten are the fitted dots characterizing the annual mean-standard deviation relation, plotted in the following figures for three catchments.

Therefore, the runoff coefficient trend was fitted with a cosine shape curve, in order to characterize the pattern with the amplitude δ and mean value of the cosine curve $\overline{rc}$. It was assumed, though, that the cosine function always has its maximum in the month of January. Hence, there is no necessity to define an offset parameter.

$$y = cos\left(2 \cdot \pi \cdot \frac{time - 0.5}{12}\right) \qquad \text{[Eq. 21]}$$

4.4.1.1. Thet at Melford Bridge

In Figure 43 the monthly temporal pattern of runoff coefficients is reported. The overall median value of runoff coefficients stands around 0.1 for all the winter and spring months, while it decreases below 0.05 in summer and autumn. In particular, the maximum runoff coefficient is reached in February and December with values around 0.3. Even if it is not quite pronounced, an inverse bell pattern characterize the temporal variability of runoff coefficients.

Figure 44 shows the relation between mean monthly runoff coefficients and their standard deviation. This plot underlines a linear increasing straight function between them. The more is the mean value of

runoff coefficient, the more is its standard deviation. It means that more uncertainty affects larger runoff coefficients. From the slope (0.6492) and intercept (0.0123) of the fitting curve it is possible to recreate this important relation.

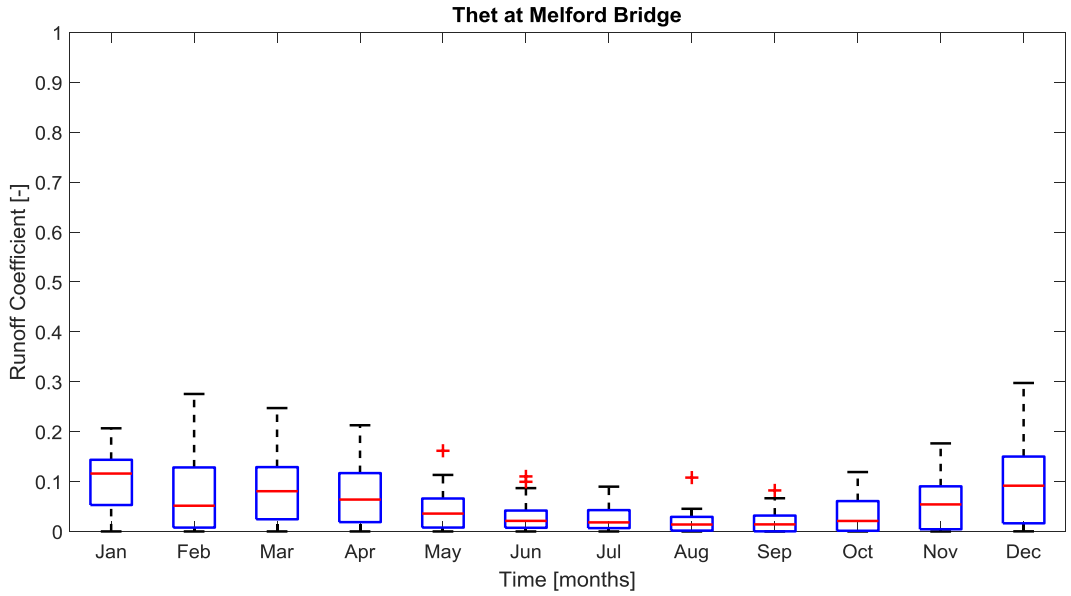

Figure 43: Runoff coefficients versus month for the catchment Thet at Melford Bridge.

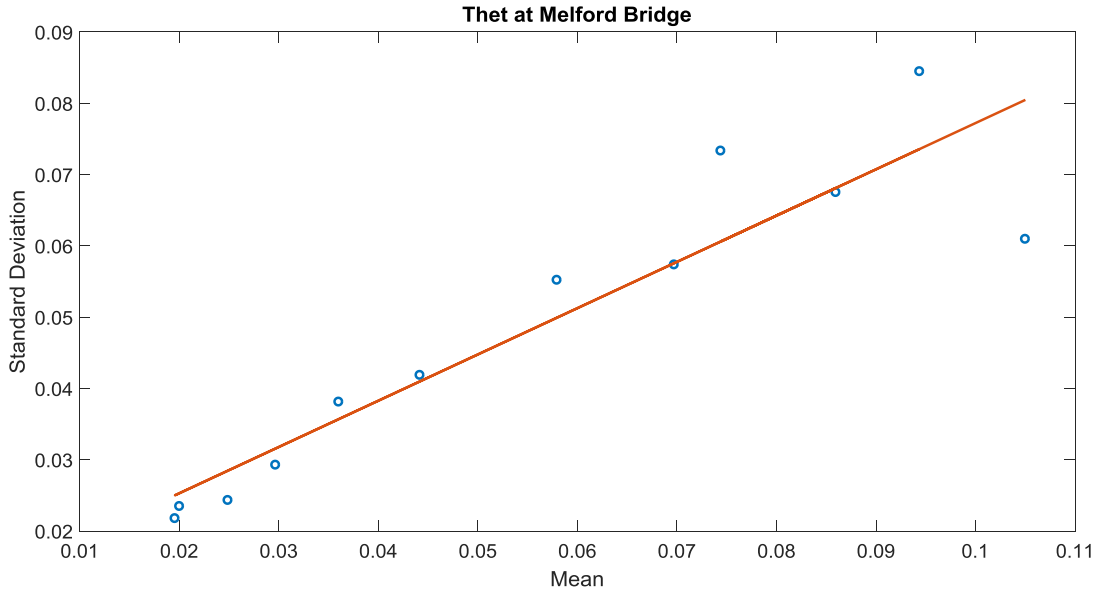

Figure 44: Mean monthly runoff coefficients versus standard deviation for the catchment Thet at Melford Bridge.

The mean value of runoff coefficients for each year from the dataset (1999-2008) is plotted against its standard deviation in Figure 45. The pattern is still linear and increasing. Also considering annual mean values of runoff coefficients, higher values assume higher standard deviation. With a slope of 0.6530 and an intercept of 0.0211, the fitting curve permit to recreate this pattern.

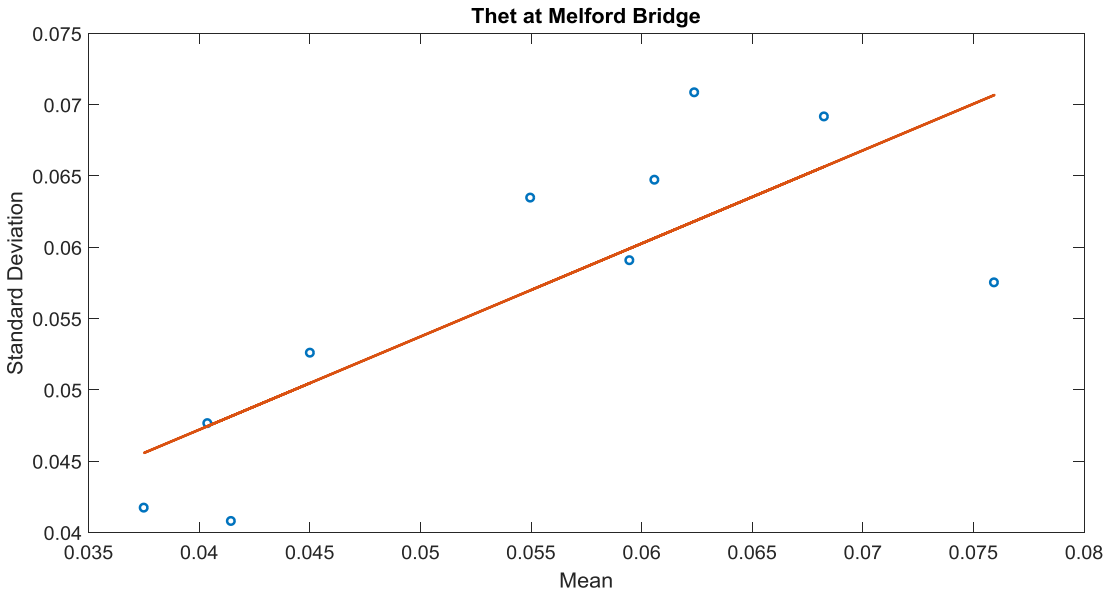

Figure 45: Mean annual runoff coefficients versus standard deviation for the catchment Thet at Melford Bridge.

Finally, monthly mean values are plotted with a cosine shape fitting curve (Figure 46). In this catchment there is a deviation between the fitting line and the points, even though the bell inverse shape of the temporal trend is respected. The amplitude of the fitting curve and its mean value, permit to recreate this temporal pattern for catchments similar to this one. For this catchment, the amplitude δ of the cosine curve evaluated is equal to 0.0328 while the mean value $\overline{rc}$ is 0.0552, according with the mean value calculated from runoff coefficients values. Actually the mean value obtained from then fitting curve and the one calculated (0.0538) differs with an error of 0.001, due to the approximation of the values. A countercheck was done with Excel to prove this fact.

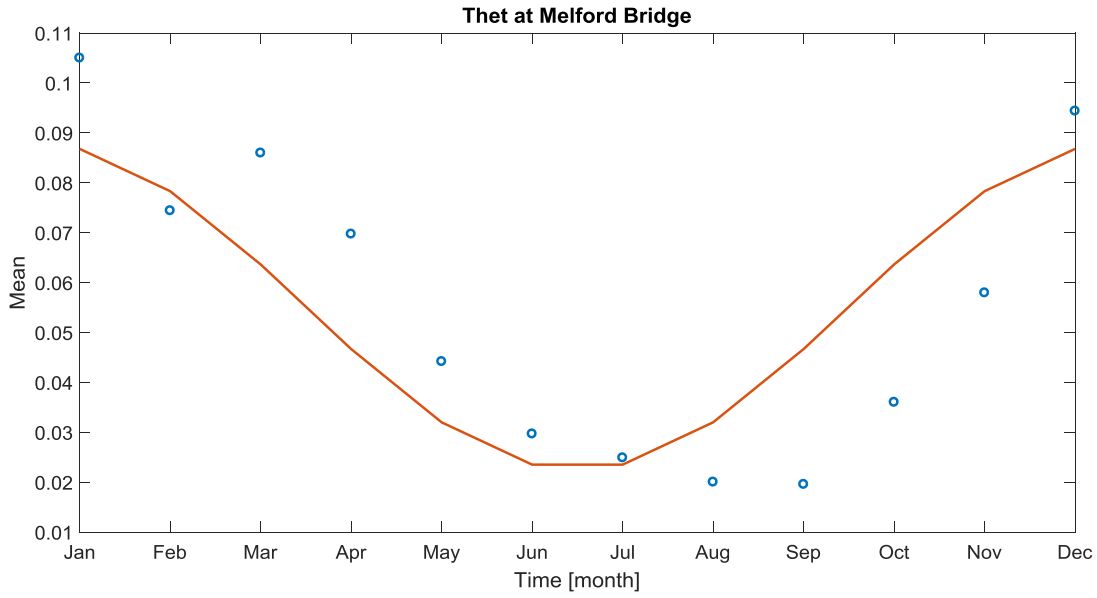

Figure 46: Fitting of the temporal pattern of runoff coefficients at Thet at Melford Bridge.

4.4.1.2. Dyfi at Dyfi Bridge

Runoff coefficients values versus time are shown in Figure 47. The trend has a reversed bell shape, more pronounced than the previous catchment analyzed. The highest runoff coefficients are observed in the months of January and December with respectively a median value of 0.63 and 0.49 and a maximum value of 0.89 for both of them. Summer appears to be the period with less runoff, with a median value of 0.1 and a maximum of 0.49 in August.

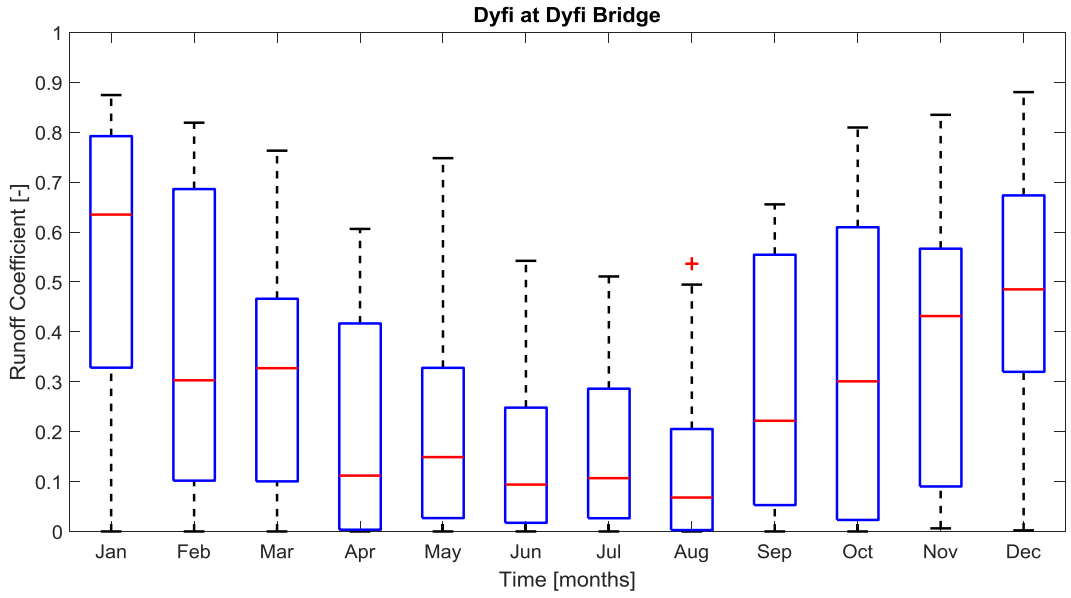

Figure 47: Runoff coefficients versus month for the catchment Dyfi at Dyfi Bridge.

Figure 48 shows the mean monthly value of runoff coefficients against its standard deviation for each month. The fitting curve of this relation is linear, increasing and straight with a slope of 0.4009 and an intercept of 0.1210.

The same relation made with mean values within the ten years of data is reported in Figure 49. The slope of the curve in this case is 0.3192, while the intercept is 0.1764. From this plot instead, it is clear that the linear straight pattern realized does not fit well the dots plotted.

It is possible to observe also, that in both cases the deviation around each mean value increases as the mean value of runoff coefficient increases.

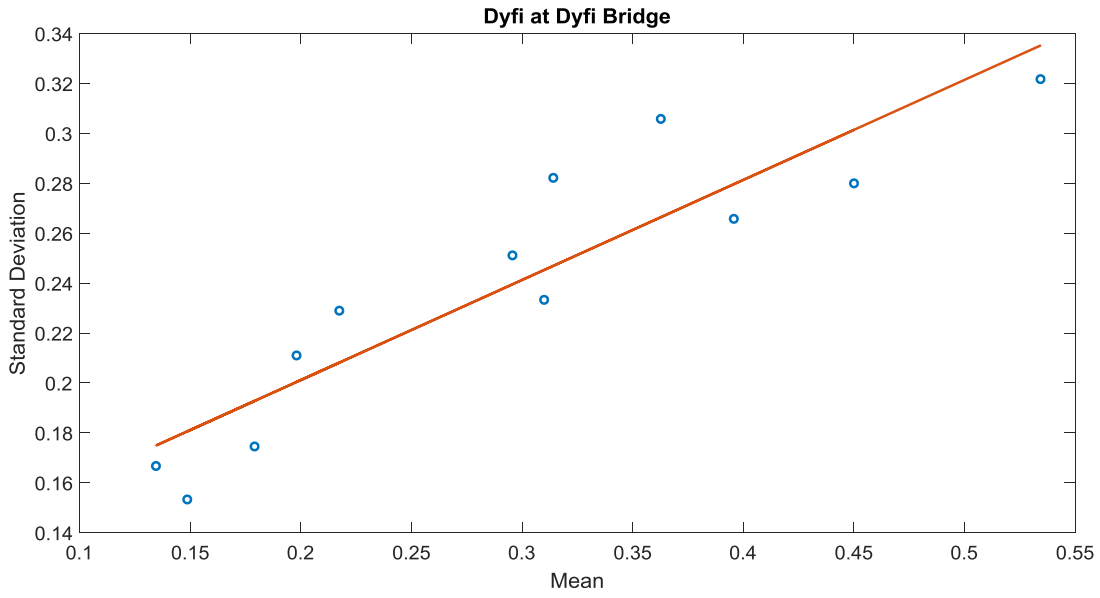

Figure 48: Mean monthly runoff coefficients versus standard deviation for the catchment Dyfi at Dyfi Bridge.

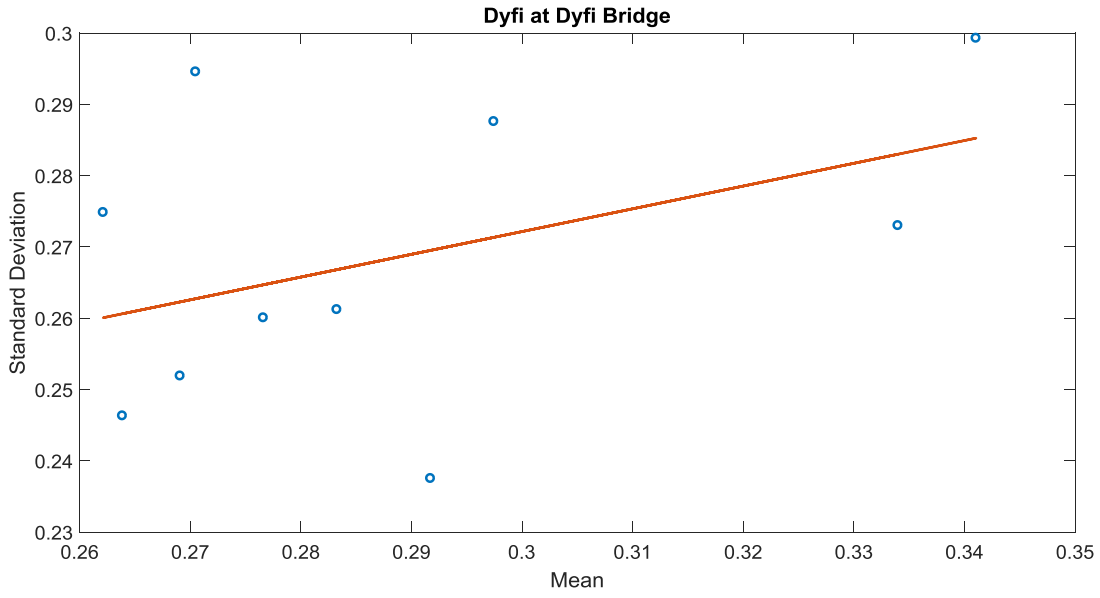

Figure 49: Mean annual runoff coefficients versus standard deviation for the catchment Dyfi at Dyfi Bridge.

Finally, monthly mean values are plotted with a cosine shape fitting curve (Figure 50). In this catchment the deviation between the fitting line and the points is less stronger than in the other catchment, so that the bell inverse shape of the temporal trend is respected quite well. For this catchment, the amplitude of the cosine curve evaluated is equal to 0.1593 while the mean value is 0.2952, according with the mean value calculated from runoff coefficients values. Also in this case the calculated average runoff coefficient and the mean value obtained from the fitting curve differ one to each other with an error of 0.008.

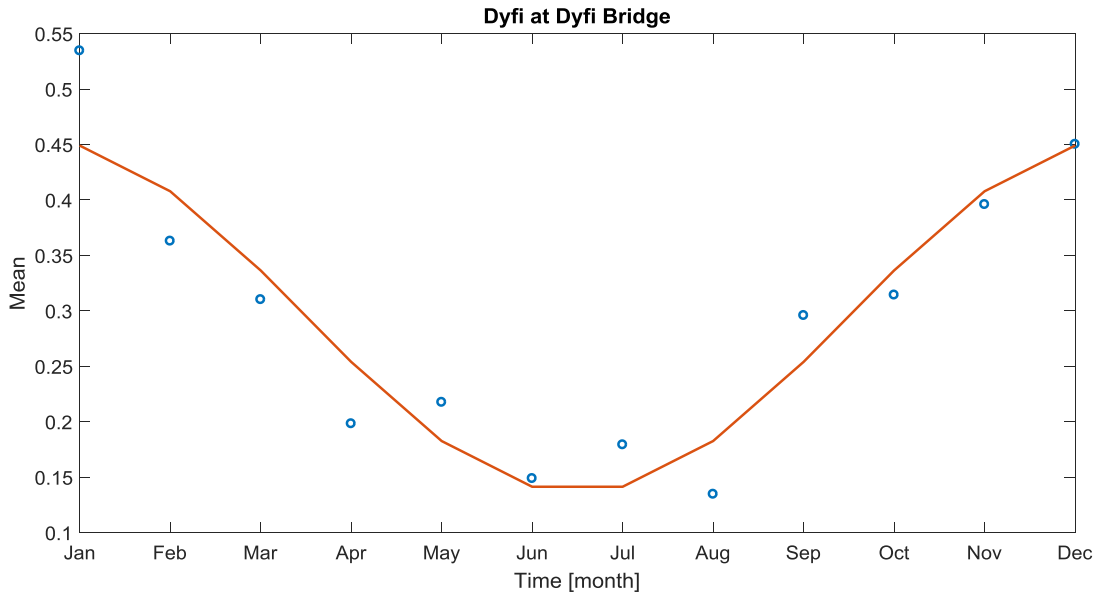

Figure 50: Fitting of the temporal pattern of runoff coefficient at Dyfi at Dyfi Bridge.

4.4.1.3. Nevis at Claggan

Runoff coefficients values versus months are shown in Figure 51. The trend in this catchment does not reveal a proper reversed bell shape, there is a decreasing of the median values of runoff coefficients during summer months in comparison with autumn and winter, but their mean values appear quite high, remaining between the range that goes to 0.5 to 0.7. The highest runoff coefficients are observed from December to May, reaching in each month a maximum near the unity. The highest median runoff coefficient values are observed in February and March with a value of 0.7.

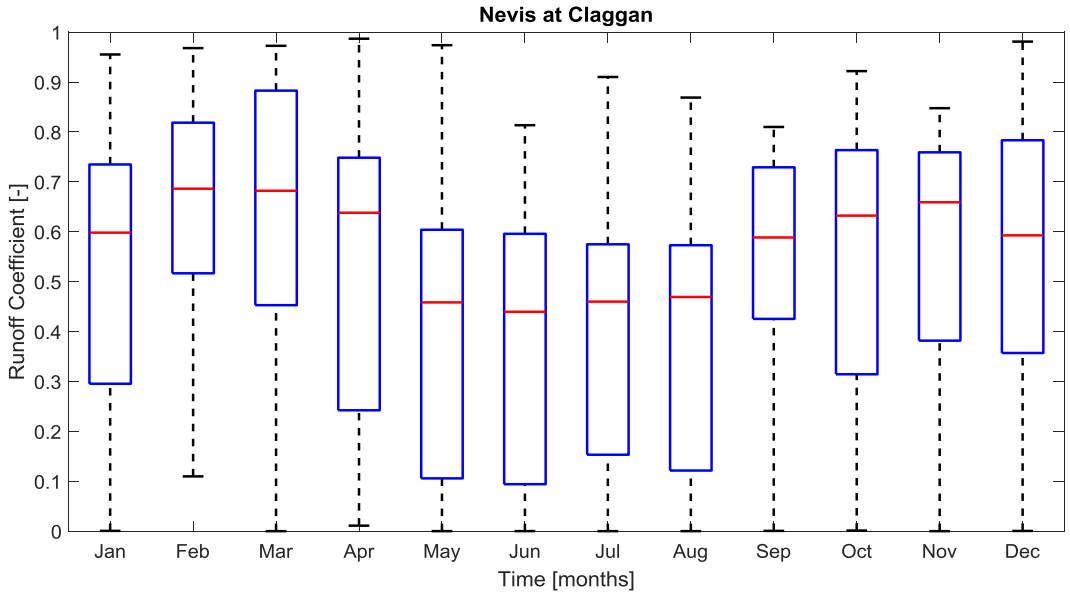

Figure 51: Runoff coefficients versus month for the catchment Nevis at Claggan.

Figure 52 shows the mean monthly value of runoff coefficients against its standard deviation for each month. The fitting curve of this relation is linear and straight with a slope of 0.0364 and an intercept of 0.2627. This fitting line tends to be almost horizontal. In this catchment in fact there is not a high variation of runoff coefficients during the year, and this fact is reflected by a deviation from the mean that stays constant within the year.

The same relation made with mean values within the ten years of data is reported in Figure 53. The slope of the curve in this case is -0.1411, while the intercept is 0.3547. The pattern has got a negative slope. This means that considering a set of ten years of data, the increasing of mean runoff coefficients probably brings to have less difference in their mean values between summer and winter periods. Having so, a median value almost constant, also the standard deviation tend to decrease as the differences between mean values decrease.

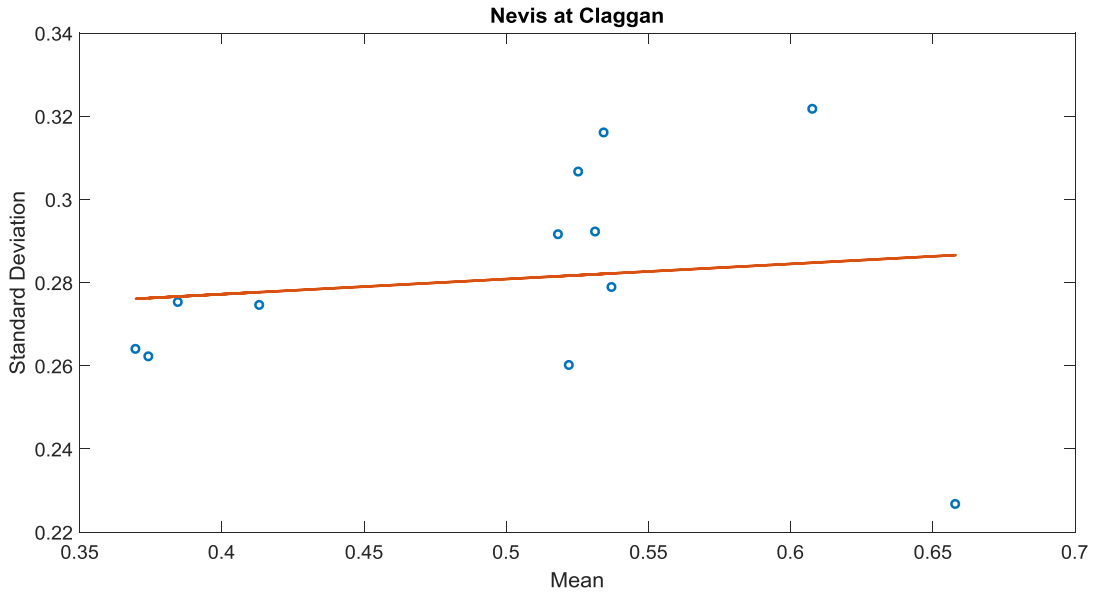

Figure 52: Mean monthly runoff coefficients versus standard deviation for the catchment Nevis at Claggan.

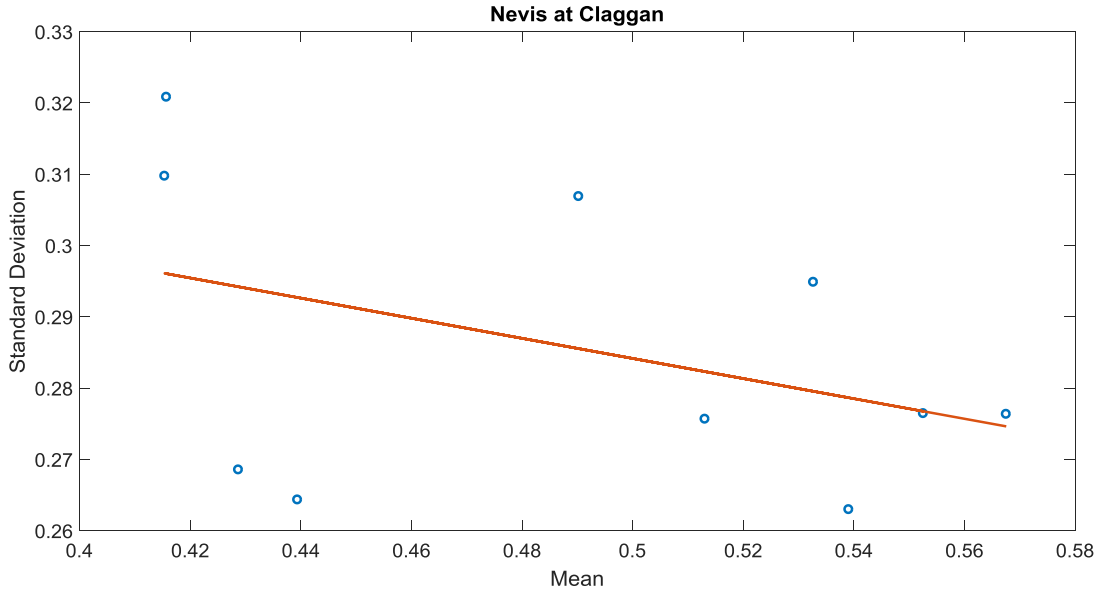

Figure 53: Mean annual runoff coefficients versus standard deviation for the catchment Nevis at Claggan.

Finally, monthly mean values are plotted in Figure 54 with a cosine shape fitting curve. In this catchment the deviation between the fitting line and the points is stronger than in the other two catchments, so that the bell inverse shape of the temporal trend is not respected well. For this catchment, the amplitude of the cosine curve evaluated is equal to 0.0963 while the mean value is 0.4980, according with the mean value calculated from runoff coefficients values. Also in this case the

calculated average runoff coefficient (0.4845) and the mean value obtained from the fitting differ one to each other with an error of 0.01.

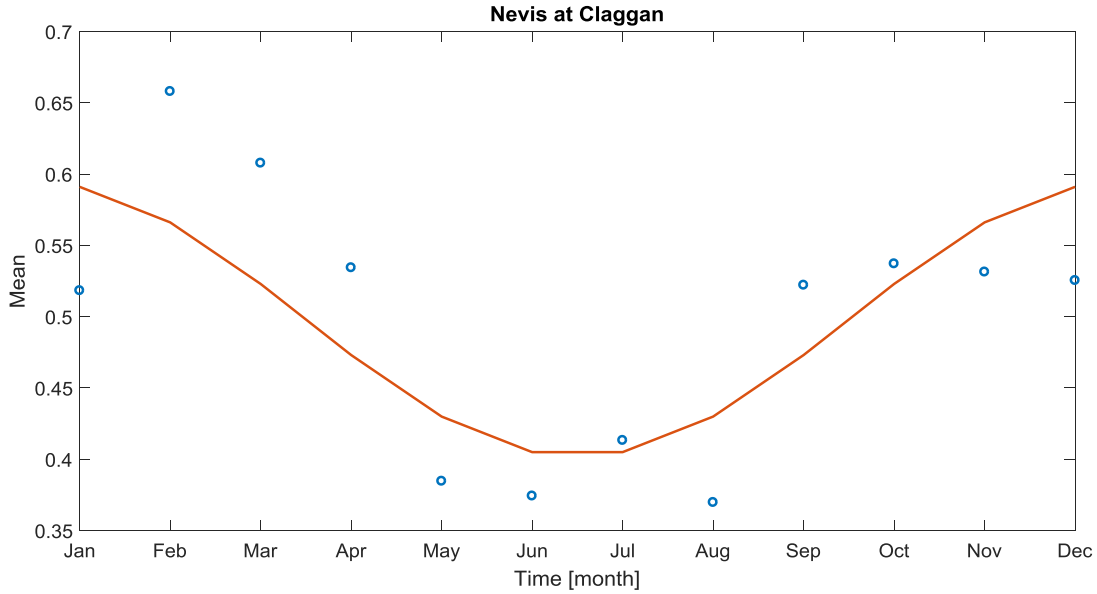

Figure 54: Fitting of the temporal pattern of runoff coefficient at Nevis at Claggan.

4.4.2. Analysis of 76 catchments

To be thorough the temporal variability of runoff coefficients has been reported also for the 76 catchments together (Figure 55). All runoff coefficients for each month of the year that goes from 1999 to 2008 were averaged. A reverse bell shape is observed also in this case with a mean runoff coefficient of 0.19. The amount of out-liers observed in summer months is explained with the fact that in summer it is more frequent the presence of short rainfall events with a high intensity. For this reason, even if short, some runoff events are related to high precipitation, and hence to high runoff coefficient. It goes against the presence of small runoff that usually occurs in summer, and this fact is reflected in the presence of values far from the average that are interpreted by the software Matlab as out-liers.

We observe in Figure 56 that the pattern between standard deviation and mean values of runoff coefficient is not linear anymore, but it follows a logarithmic shape. This means that the standard deviation increases faster at the increasing of runoff coefficients values.

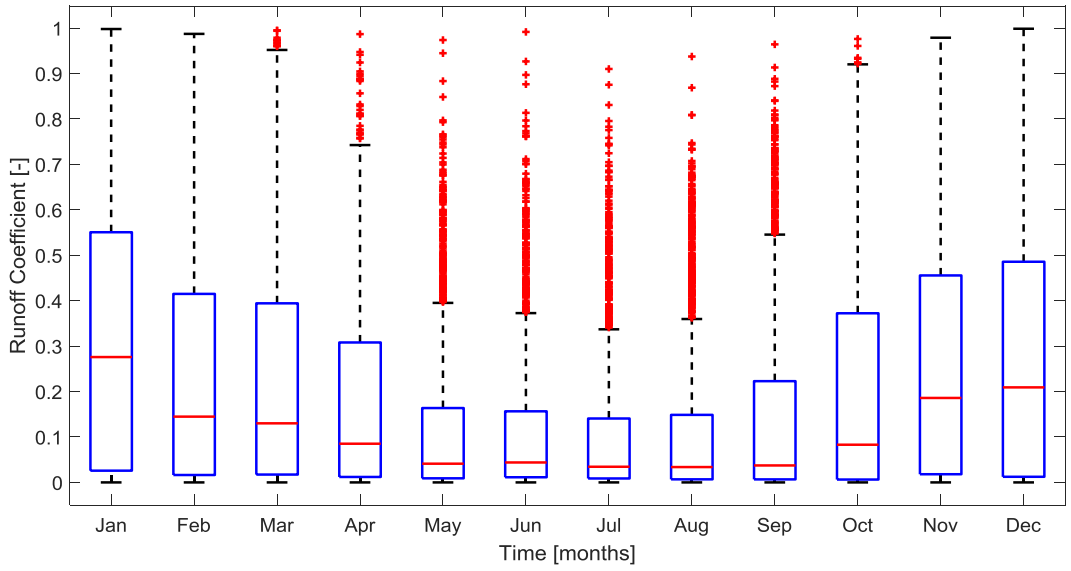

Figure 55: Runoff coefficients versus time (months) for all the 76 catchments analyzed.

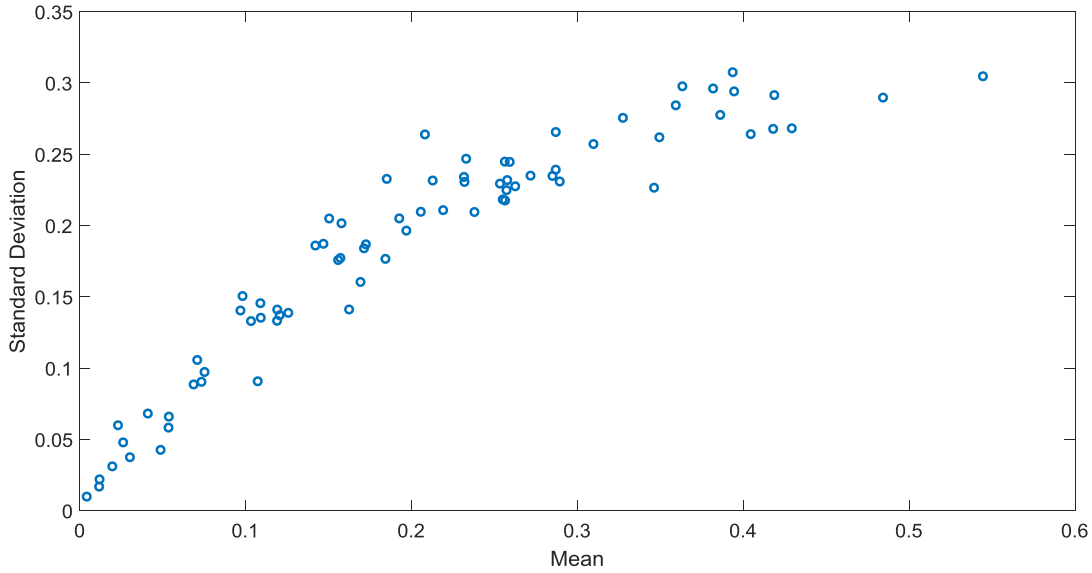

Figure 56: Mean runoff coefficient versus standard deviation for all the 76 catchments analyzed.

Another important result is reported for each catchment in Figure 57 to show the relation between mean value of runoff coefficients and the amplitude δ of their monthly temporal pattern. It is possible to notice that at the increasing of the mean runoff coefficients also the amplitude increases. The only exception is given by catchments that have a mean runoff coefficient larger than 0.4. Most of them in fact, belong to regions such as Scotland in which runoff coefficients tend to have a almost constant mean value in each month of the year. For Nevis at Claggan was shown that the reverse bell shape of the monthly variability was not pronounced as in other catchments. It is reflected in a lower amplitude in

the cosine fitting curve and hence in dots that have high mean runoff coefficients but amplitude lower than 0.1.

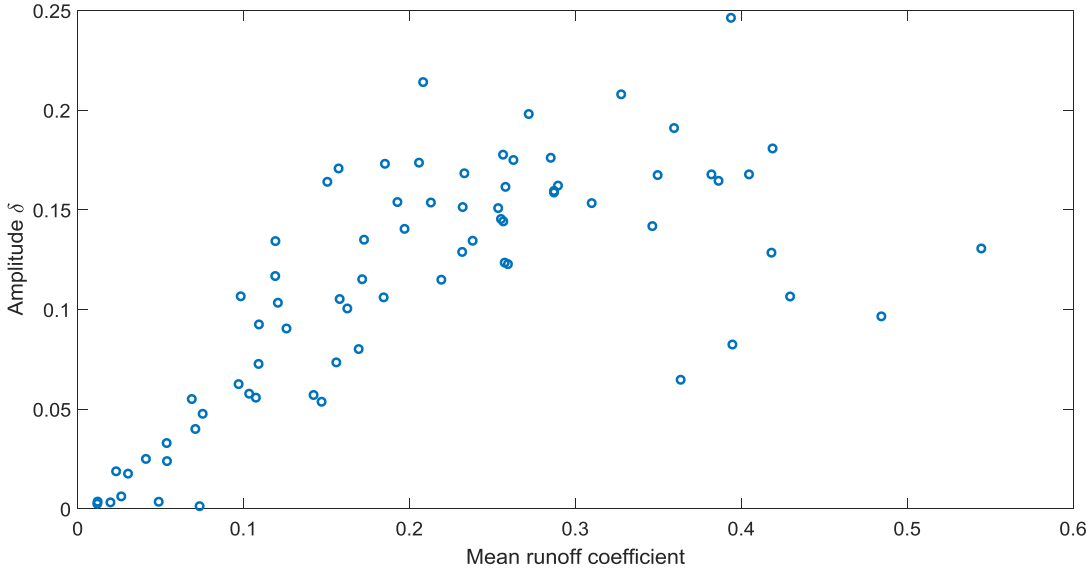

Figure 57: Mean runoff coefficient versus the amplitude of the temporal variation of runoff coefficients within month for all the 76 catchments.

Finally the scatter plot for the mean value of the cosine fitting curve $\overline{rc}$ (Figure 58) and the amplitude δ (Figure 59) are reported for all the 76 catchments. The same is done for the slope (Figure 60) and intercept (Figure 61) of the mean-standard deviation fitting curve. As a counter result $\overline{rc}$ values are the same plotted in Figure 34. The map with slope values shows higher results in the south of the country presumably due to a higher variation of runoff coefficients between warm and cold seasons. It was shown from singular catchments analysis that similar mean runoff coefficients within a year implied a constant standard deviation at the increasing of runoff coefficient values.

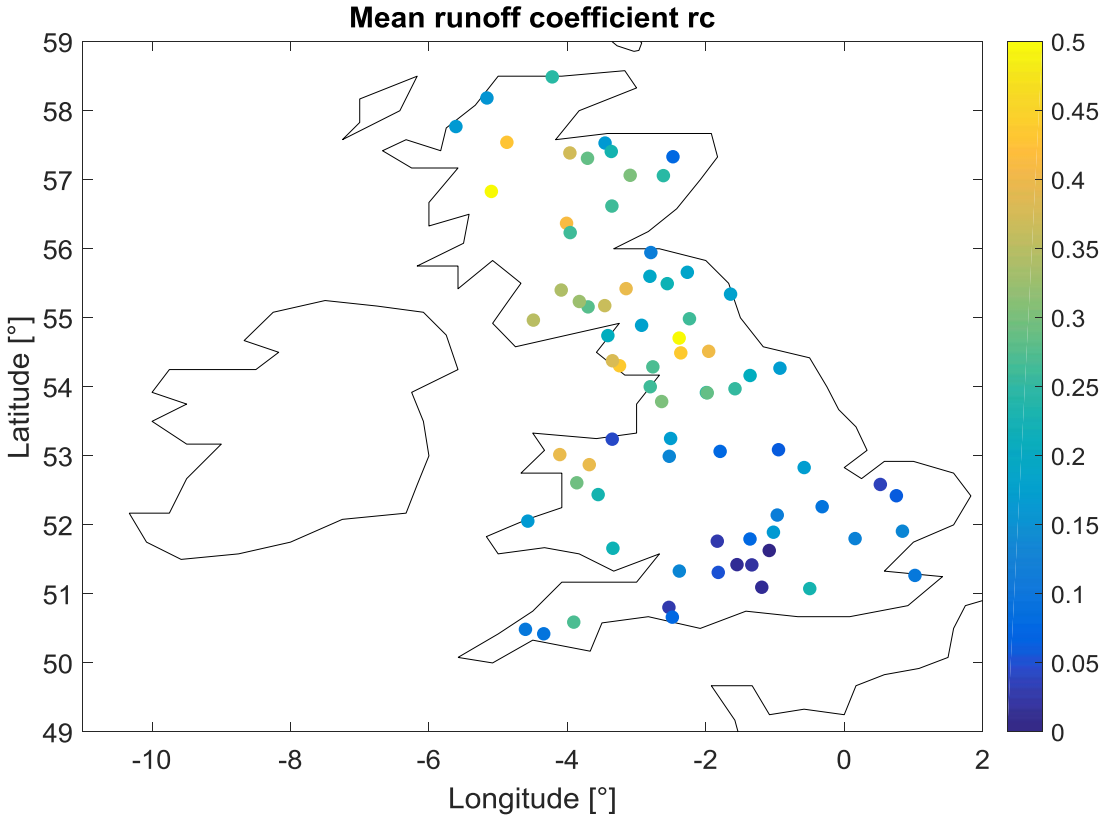

Figure 58: Mean runoff coefficient $\overline{rc}$.

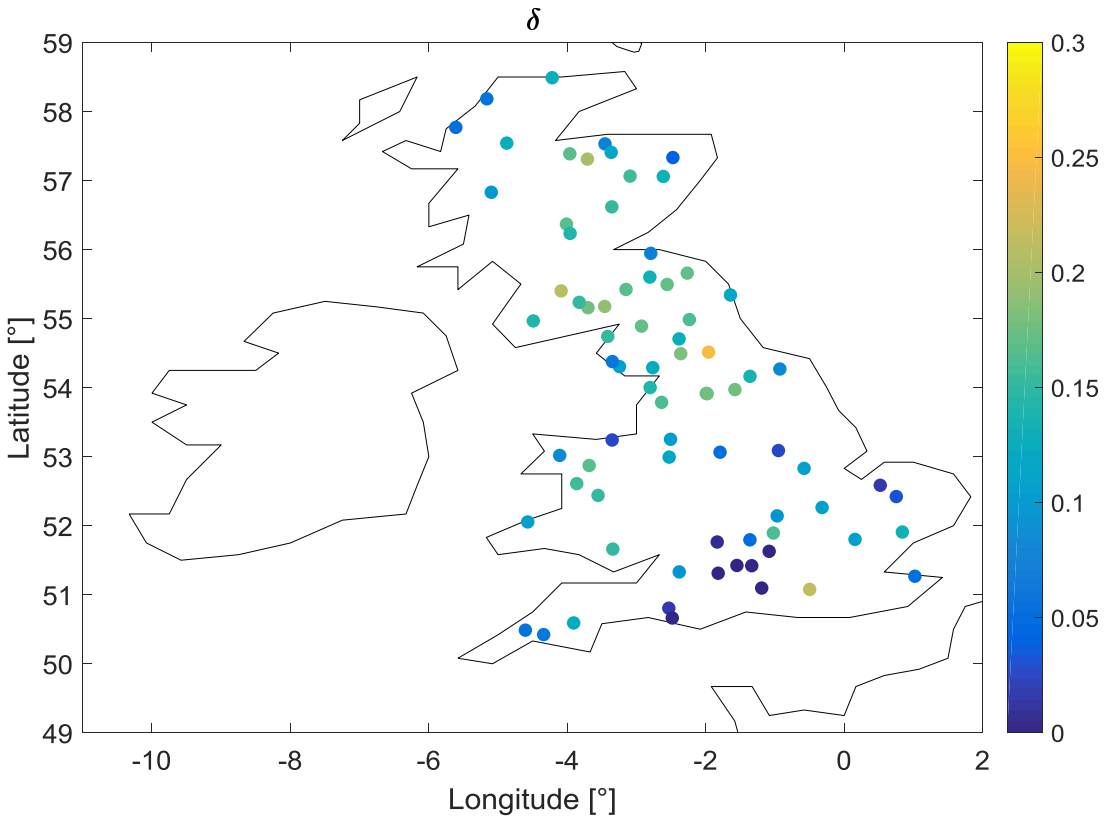

Figure 59: Amplitude δ

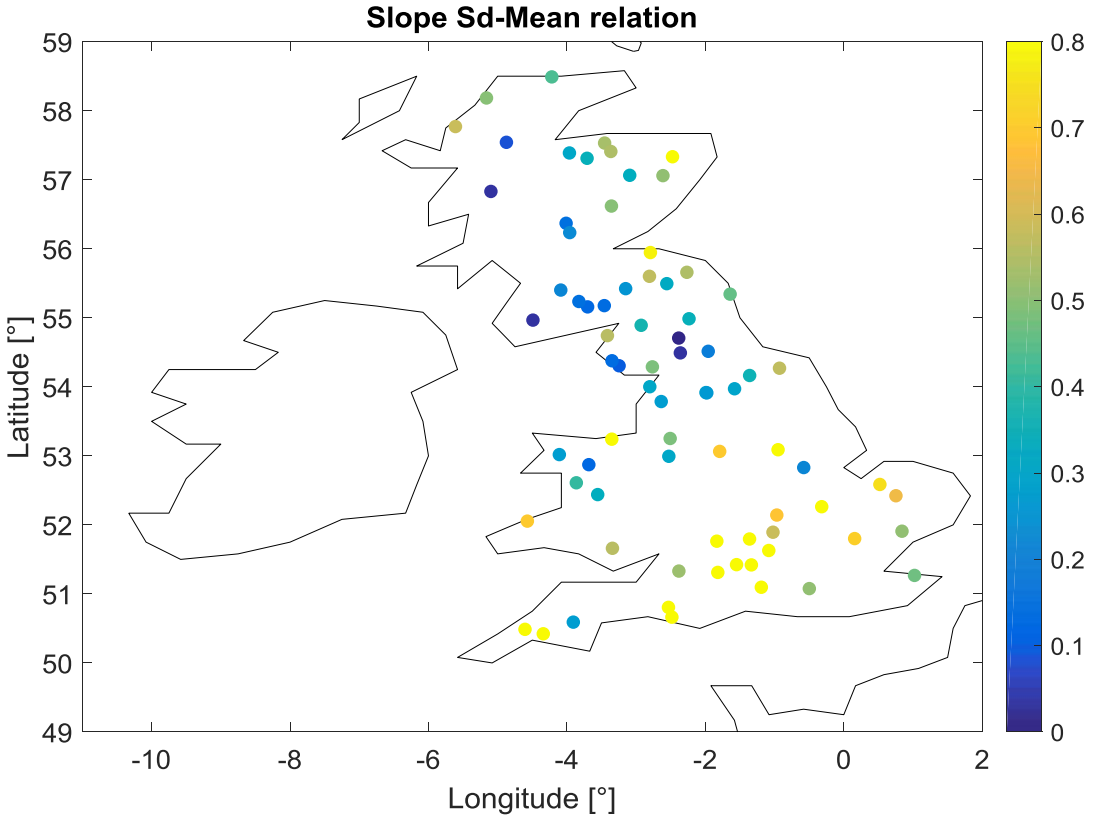

Figure 60: Slope of the fitting curve Mean-Standard deviation.

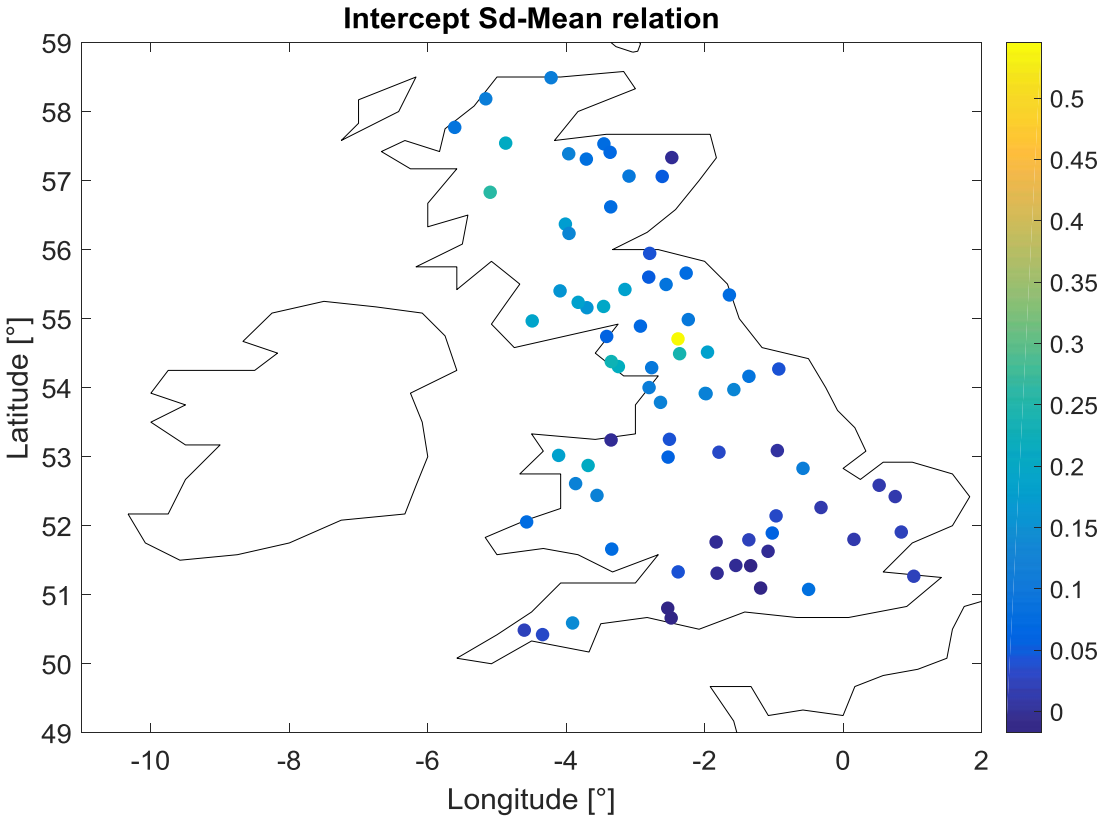

Figure 61: Intercept of the fitting curve Mean-Standard deviation.

4.5. Regression trees

As mentioned in Chapter 3, this work is divided in two parts. The first one with its results, regarded the analysis of spatial and temporal variability of runoff coefficients in relation to 76 catchments spread in all UK, in specific the results of three representative catchments have been reported above. The second part, instead, discussed in this paragraph, bases on the prediction of runoff coefficients for ungauged catchments, by using information such descriptors of soil properties and climate of gauged catchments with similar characteristics to the ungauged ones. The dataset was divided in two groups: the first one contains 51 catchments considered as "gauged", the second 25 catchments considered as "ungauged".

The set of 51 basins was used to create regression trees in order to be able to do predictions about the temporal pattern of runoff coefficients and their uncertainty in term of standard deviation. As already said above, the temporal pattern of runoff coefficient is known once the amplitude δ and the mean runoff coefficient $\overline{rc}$ from the cosine fitting curve are known. In the same way, the deviation of runoff coefficients from their mean value is known once the intercept and the slope of the mean-standard deviation relation are known.

For each of these four characteristics, regression trees were realized considering different input descriptors. At first PROPWET, secondly PROPWET and DPSBAR, then PROPWET, DPSBAR and BFIHOST, and finally PROPWET, DPSBAR, BFIHOST, AREA. The regression tree will be shown only for the fourth tentative because, taking into account more soil information, sharpener and more reliable results were obtained. It was decided, though, not to consider MAP as predicting descriptor because it contains redundant information: it is a surrogate control of the soil moisture content, better expressed in term of proportion of time catchments may be considered wet or dry by PROPWET. This descriptor resulted to be the most influencing one in the prediction. The set of catchments to be considered as gauged and ungauged chosen in order to have the same mean value of PROPWET gave better results in prediction (measured by RMSE).

The prediction has been done realizing first of all the regression trees for each of the four characteristics we are interested in ($\overline{rc}, \delta, slope\ and\ intercept$) by using the data of 51 catchments. For example, for the prediction of the mean runoff coefficient, the tree subdivides catchments descriptor in relation to their values with the result to have in its leaf nodes values of the runoff coefficients that take into account the diverse spatial variability for each catchment. Through a predict function of Matlab, it was possible to enter in the tree using catchment descriptors of ungauged catchments, that we consider to know, and hence do the prediction of the interested quantities. These ungauged catchments are actually known, so the calculation of the same parameters were compared to the estimated ones, in order to have an idea of the uncertainty of this prediction method.

The uncertainty in this procedure is evaluated in term of the root mean standard error (RMSE), that has been calculated for each regression tree studied. Tables with comparison about the percentage of error between calculated and predicted values will be reported for each quantity in Appendix A. These errors are simply estimated as the difference between predicted values from regression trees and calculated values from the dataset.

4.5.1. Prediction of the mean runoff coefficient

The regression tree used to do the prediction of the mean runoff coefficient is reported in the following picture (Figure 62). By using a prediction function in Matlab the software managed to calculate the mean runoff coefficient of 25 catchments considered as ungauged by entering inside the tree.

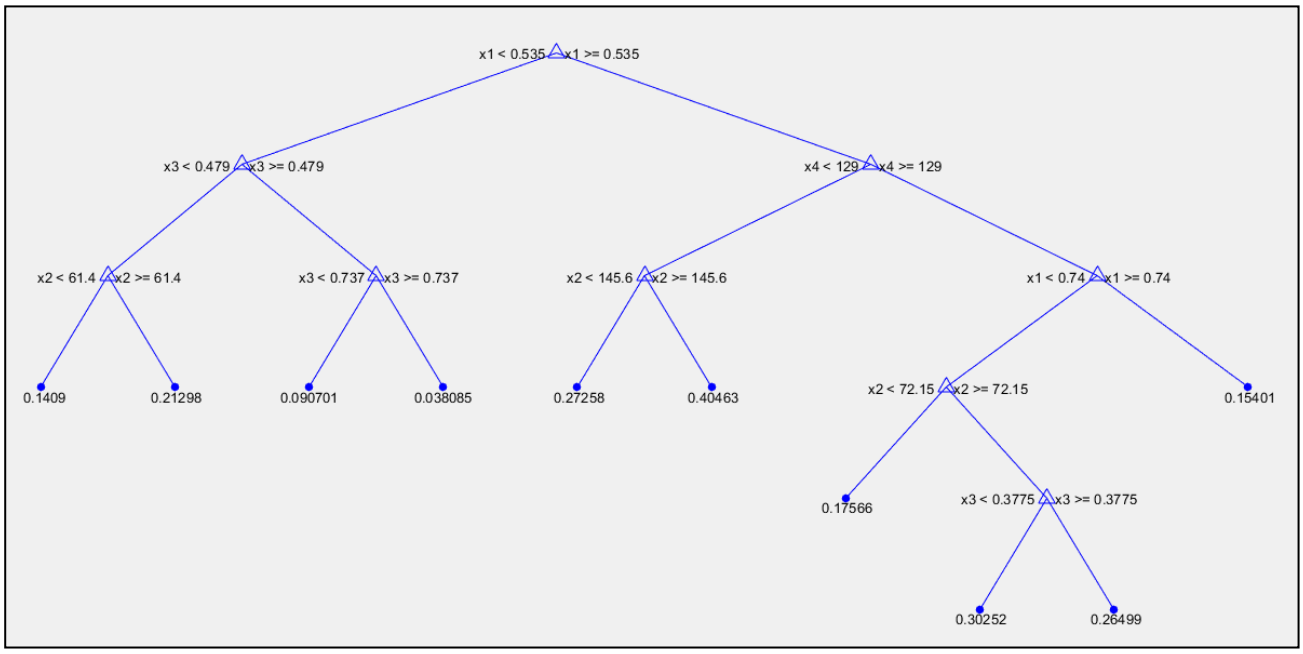

Figure 62: Regression tree for prediction of mean runoff coefficient RC using PROPWET(x_1), DPSBAR(x_2), BFIHOST(x_3), AREA(x_4) as input predictors.

Four attempts were done considering each time a bigger number of descriptors as input. It is possible to establish the error committed in the different cases. Analyzing Figure 63 from figures (a) to (d), it is clear that the dots plotted tend to assume a linear pattern as the number of predictor increases. Actually, doing a qualitative analysis, cases (c) and (d) are quite similar, but the RMSE is slightly smaller in case (d). For this reason it is possible to affirm that all the descriptors influence the final predicted values, reducing uncertainty in the estimation.

RMSE			
(a)	(b)	(c)	(d)
0.1247	0.1099	0.0997	0.0849

Table 5: RMSE values for regression three realized using as input predictors: (a) PROPWET; (b) PROPWET, DPSBAR; (c) PROPWET, DPSBAR, BFIHOST; (d) PROPWET, DPSBAR, BFIHOST, AREA.

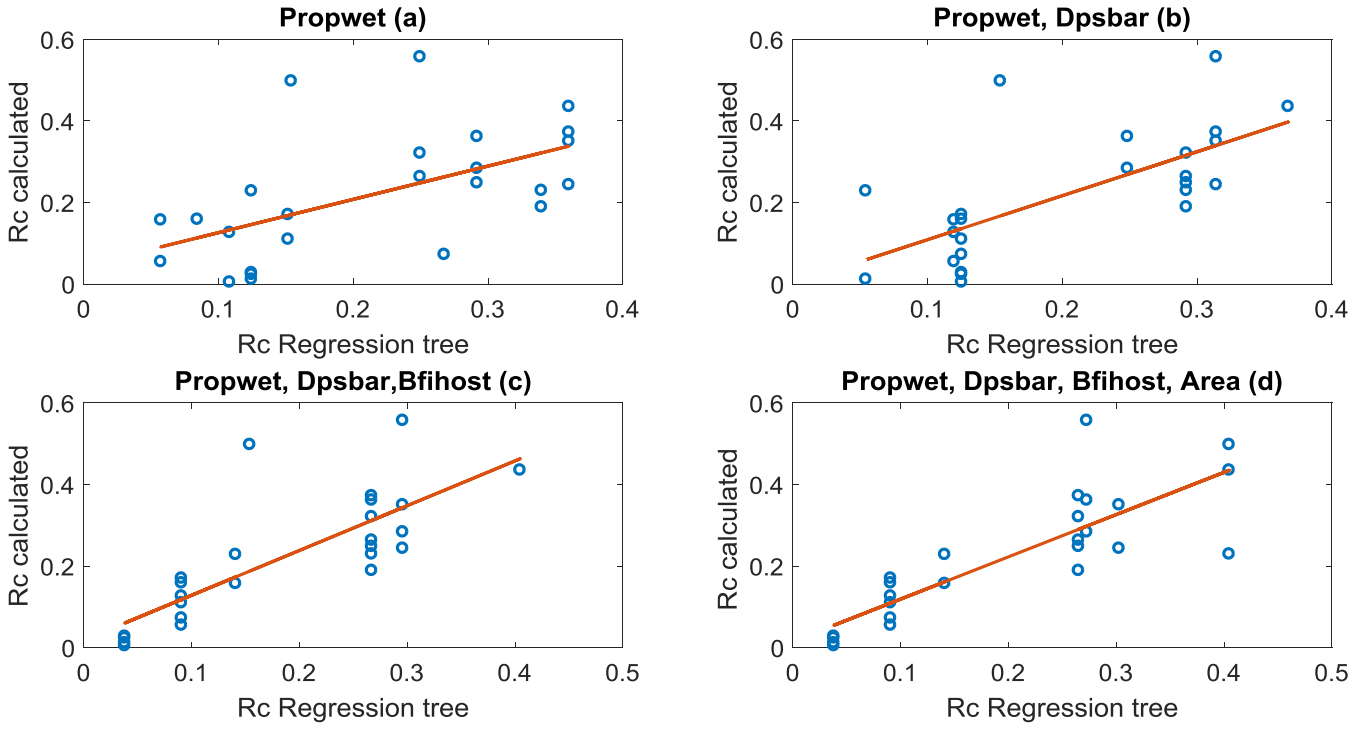

Figure 63: Relation between runoff coefficient values calculated and predicted from regression trees for 25 catchments of the dataset.

In term of errors between predicted values and computed ones, the results indicate that this method of prediction is not very reliable. Tables with errors of the prediction are reported in Appendix A (Table 10).

Only 14 catchments among 25 reveal a prediction in their mean runoff coefficient values that is less than the 30%, but sometimes the large percentage error is a consequence of the observed value being very close to zero. Considering the results obtained, (see Appendix A) for catchments 35, 39 and 47 probably the prediction with trees is not able enough to gather these basins information and relate them to reliable mean runoff coefficients. Catchment 4 has le lowest percentage error in the prediction equal to 0.5%.

Figure 64 shows the error in term of the difference between the calculated values and the predicted ones. Catchment 47 and 14 in this case show the worst error in the map, marked by the yellow and violet dots respectively.

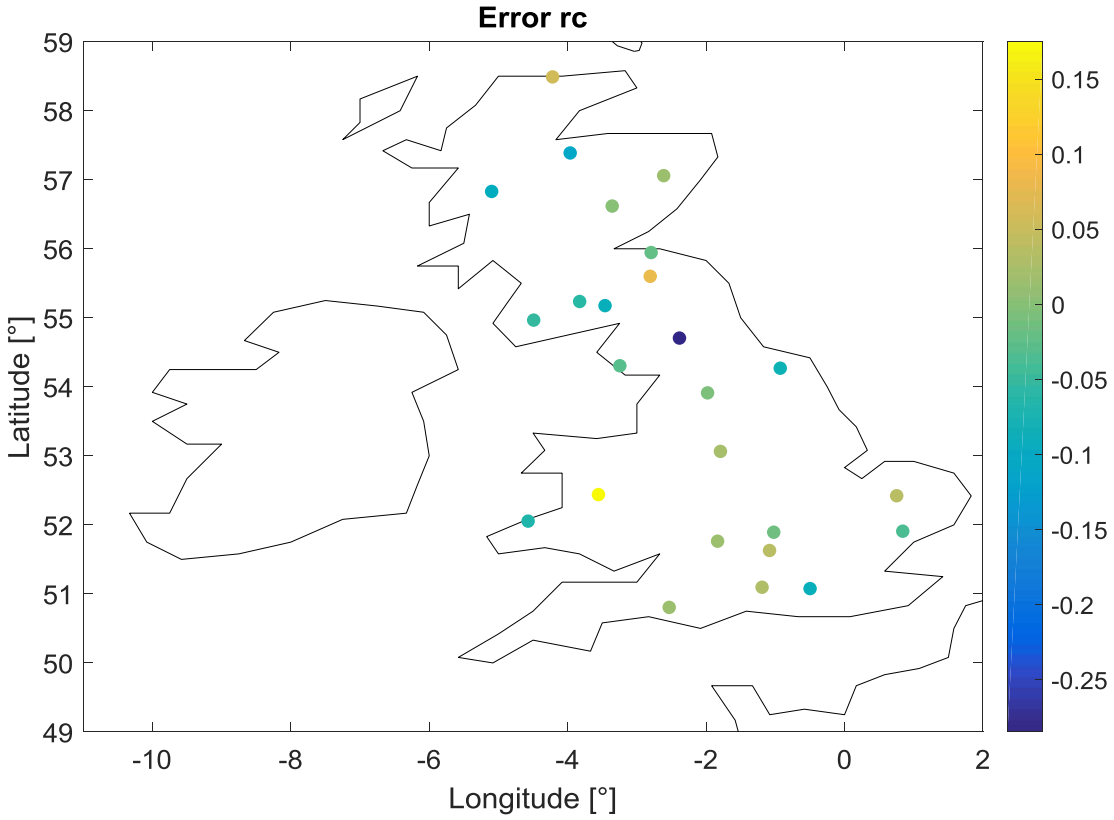

Figure 64: Error of the mean runoff coefficient for 25 "ungauged" catchments.

4.5.2. Prediction of δ

The regression tree used to do the prediction of the amplitude of the cosine fitting curve is reported in Figure 65. By using a prediction function in Matlab the software managed to calculate the values of δ for 25 catchments considered as ungauged.

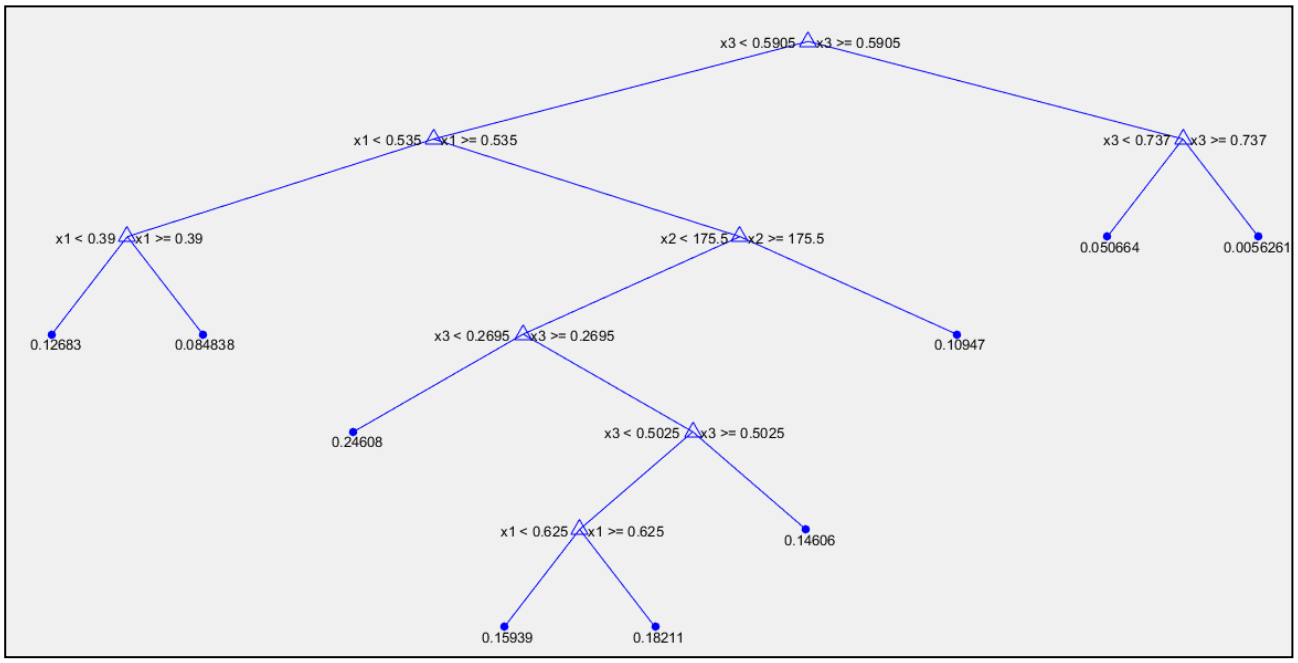

Figure 65: Regression tree for prediction of amplitude $\boldsymbol{\delta}$ *using PROPWET(x_1), DPSBAR(x_2), BFIHOST(x_3), AREA(x_4) as input predictors.*

Four attempts were done using a different number of descriptors as predictors. It is possible to establish the error committed considering at first one descriptor, and then one more at the time. Analyzing Figure 66 from figures (a) to (d), it is clear that the dots plotted tend to assume a more evident linear pattern as the number of predictor increases. Though, cases (c) and (d) show the same results. It means that in this case the information of the catchment Area was not relevant in the prediction.

RMSE			
(a)	(b)	(c)	(d)
0.0548	0.0512	0.0363	0.0363

Table 6: RMSE values for regression three realized using as input predictors: (a) PROPWET; (b) PROPWET, DPSBAR; (c) PROPWET, DPSBAR, BFIHOST; (d) PROPWET, DPSBAR, BFIHOST, AREA.

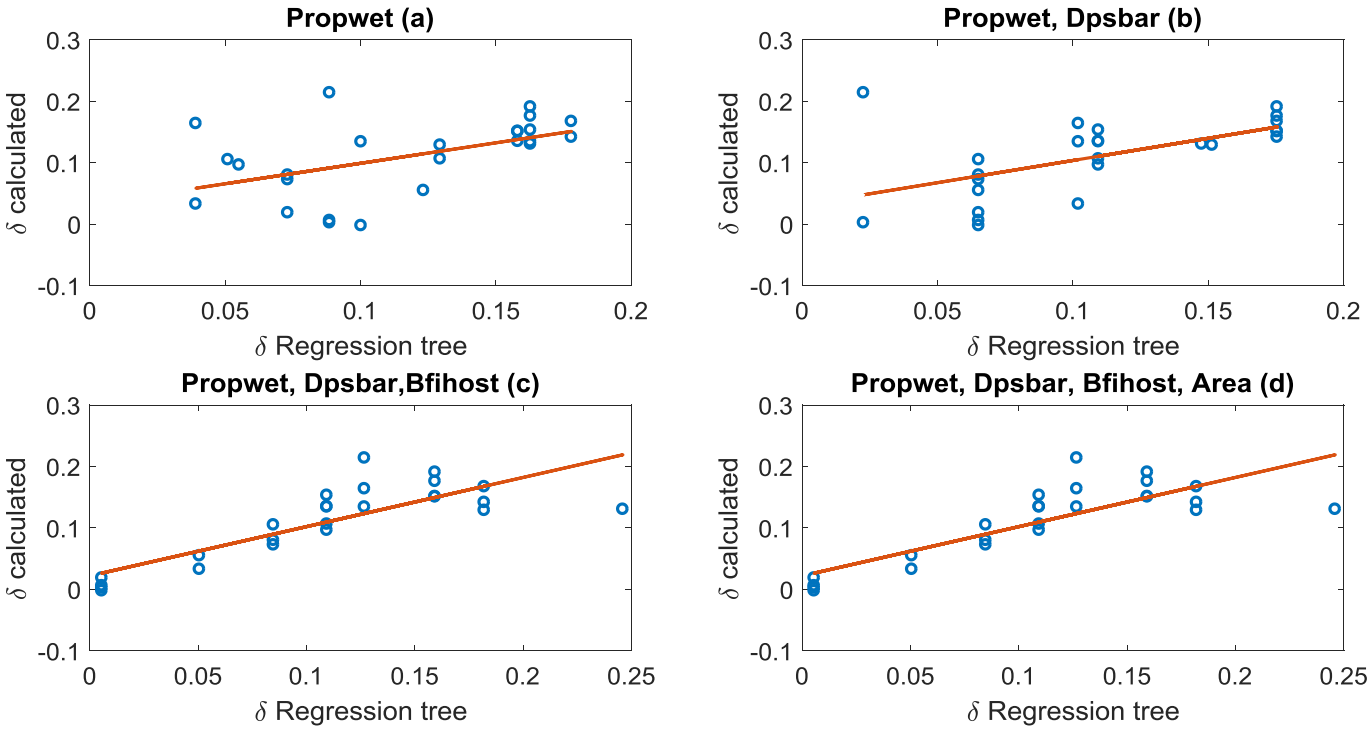

Figure 66: Relation between amplitude values calculated and predicted from regression trees for 25 catchments of the dataset.

The prediction of δ shows better result than the previous cases, even if also in this analysis it is possible to notice some catchments for whom probably this prediction procedure is not completely suitable. 18 catchments among 25 show a percentage error lower than 30%. Also in this case, catchment 35 and 39 and 14 show error over the 70%, due to the fact that the observed values are close to zero

In Appendix A (Table 11) the table with the values of the error in the prediction of the amplitude of the temporal pattern of runoff coefficients is shown. The prediction appears reliable for the majority of the catchments except for number 14 and 38 identified by the yellow and violet dots in Figure 67.

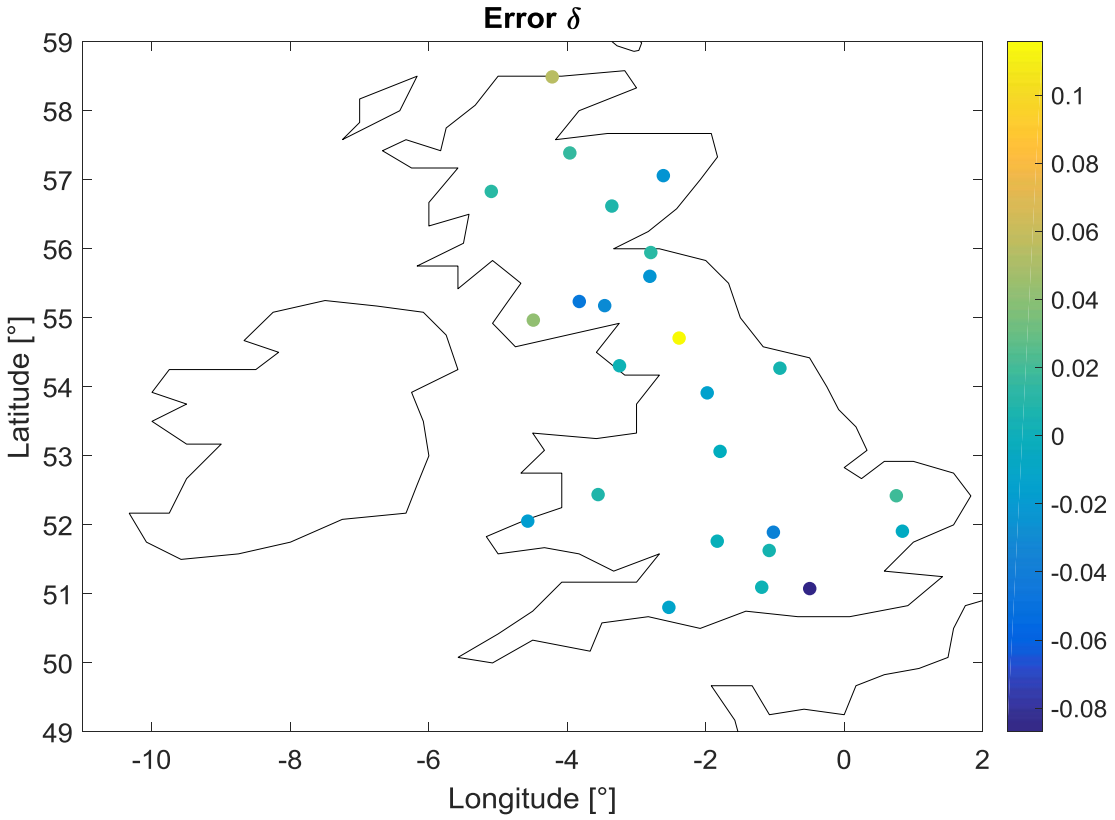

Figure 67: Error of the amplitude δ for 25 "ungauged" catchments.

4.5.3. Prediction of the slope of Mean-StDev relation of runoff coefficients

The regression tree used to do the prediction of the slope of the relation between mean and standard deviation of runoff coefficient is reported in Figure 68. By using a prediction function in Matlab the software managed to calculate the values of slope for 25 catchments considered as ungauged.

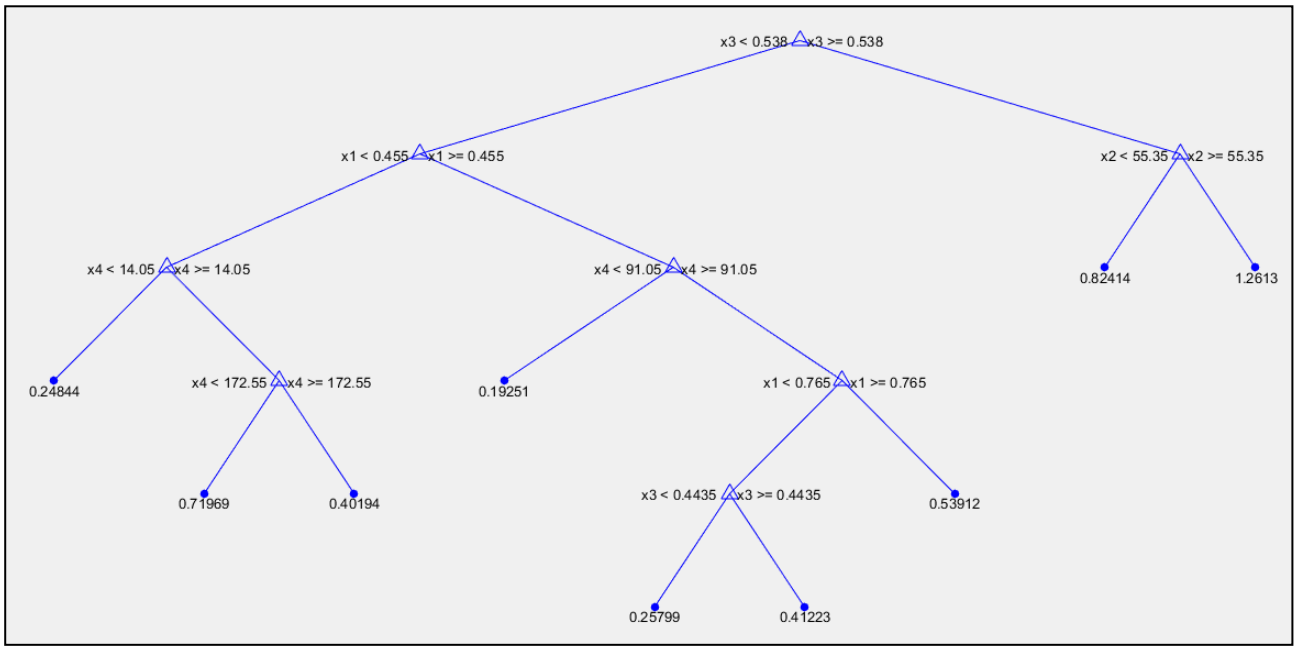

Figure 68:Regression tree for prediction of slope in Mean-StDev relation using PROPWET(x_1), DPSBAR(x_2), BFIHOST(x_3), AREA(x_4) as input predictors.

Four attempts were done with a different number of descriptors as predictors. It is possible to establish the error committed considering at first one descriptor, and then one more at the time. Analyzing Figure 69 from figures (a) to (d), it is clear that the dots plotted tend to assume a more evident linear pattern as the number of predictor increases. Actually cases (c) and (d) show similar results. The RMSE value though shows that also the basin area influences the result improving the prediction.

RMSE			
(a)	(b)	(c)	(d)
0.6186	0.5686	0.4545	0.4325

Table 7: RMSE values for regression three realized using as input predictors: (a) PROPWET; (b) PROPWET, DPSBAR; (c) PROPWET, DPSBAR, BFIHOST; (d) PROPWET, DPSBAR, BFIHOST, AREA.

Values of the RMSE show for this parameter are higher than for the other parameters. It is reflected in the magnitude of errors estimated as the difference between predicted and calculated values of slope.

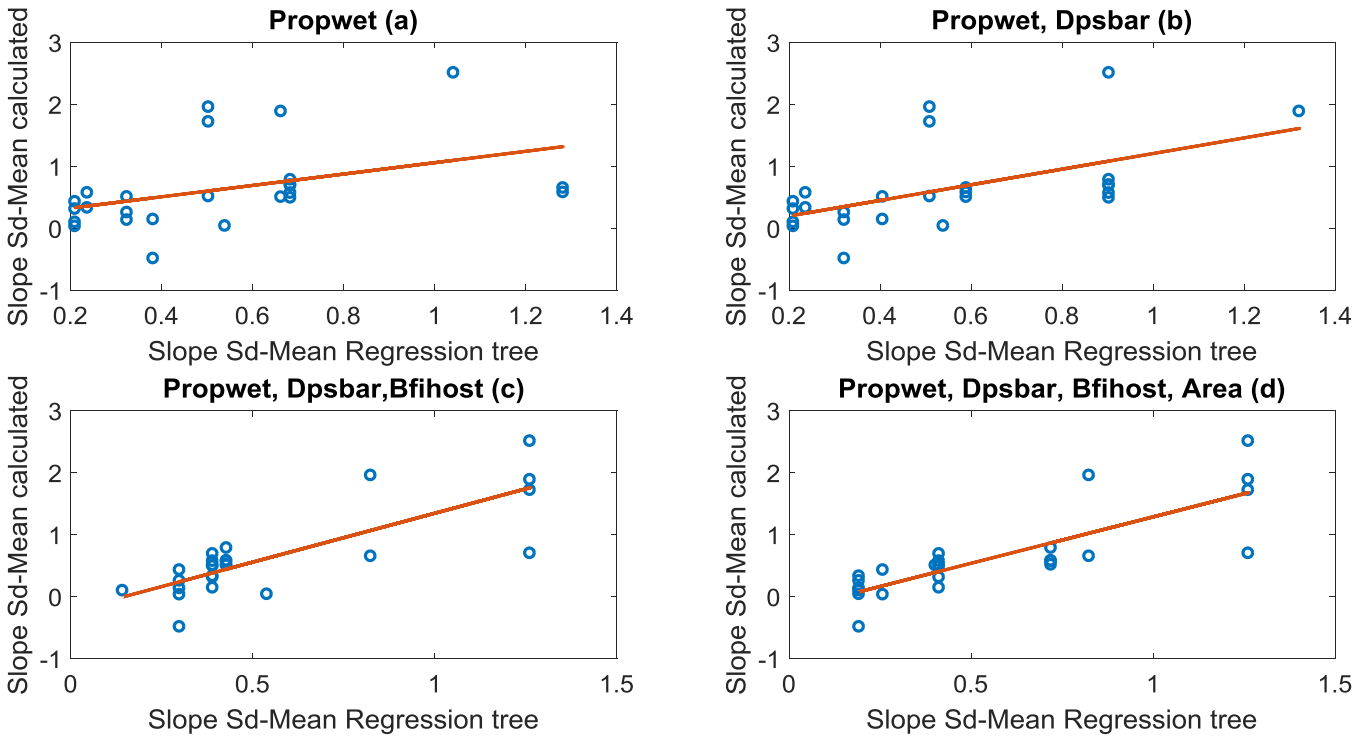

Figure 69: Relation between slope values in Mean-StDev relation calculated and predicted from regression trees for 25 catchments of the dataset.

The slope of the mean-standard deviation relation of runoff coefficient is definitely not assessable using regression trees. Errors larger than the 50% occur for several catchments. Only three catchments have a predicted slope affected by an error smaller than the 10%.If the temporal pattern of runoff coefficients is overall respected within a year due to the seasonality of rainfall directly related to runoff response, the same may not be stated for mean and standard deviation relation. Catchments 14, 69, 61,72,73 show completely unreliable results (Appendix A, Table 12).

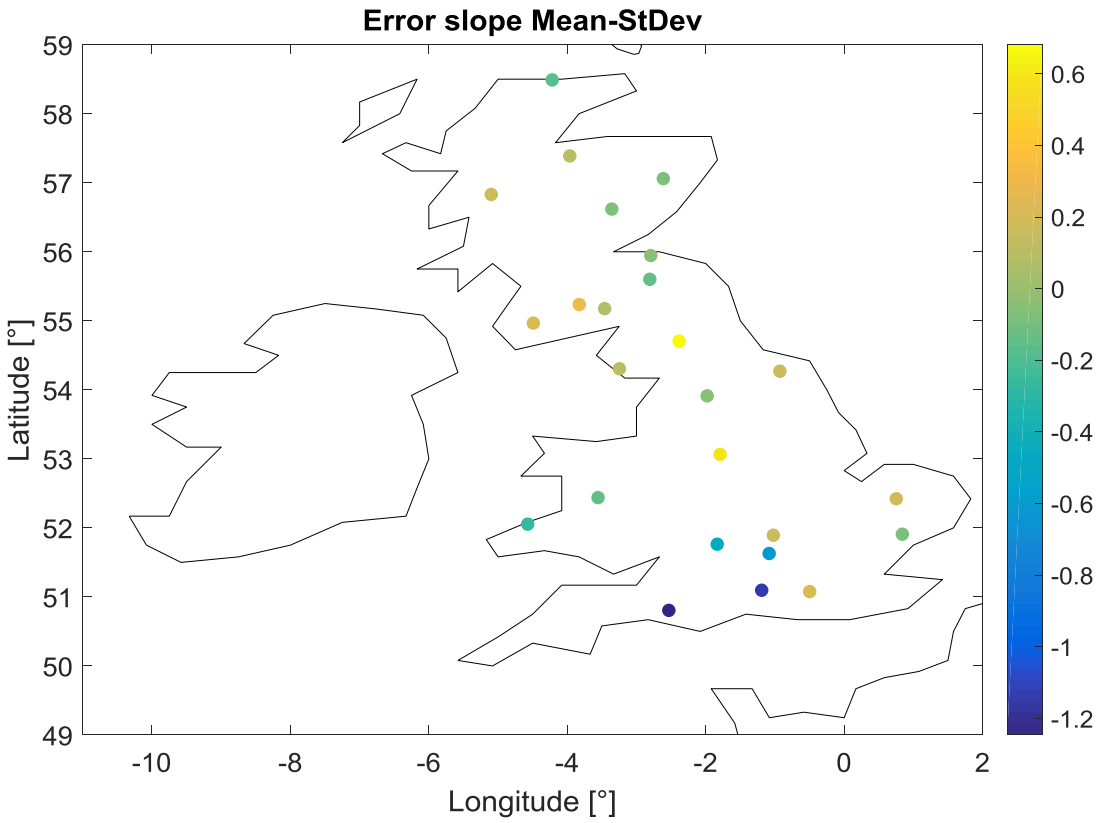

Figure 70: Error of the slope values in Mean-StDev relation for 25 "ungauged" catchments.

Figure 70 shows for each catchment the error for the parameter currently analyzed simply calculated as the difference between predicted and calculated values. The South of the country shows bigger values of errors in the prediction of the slope of the mean-StDev relation of runoff coefficients.

4.5.4. Prediction of the intercept in the Mean-StDev relation of runoff coefficients

The regression tree used to do the prediction of the intercept in the relation between mean and standard deviation of runoff coefficient is reported in Figure 71. By using a prediction function in Matlab the software managed to calculate the values of intercept for 25 catchments considered as ungauged.

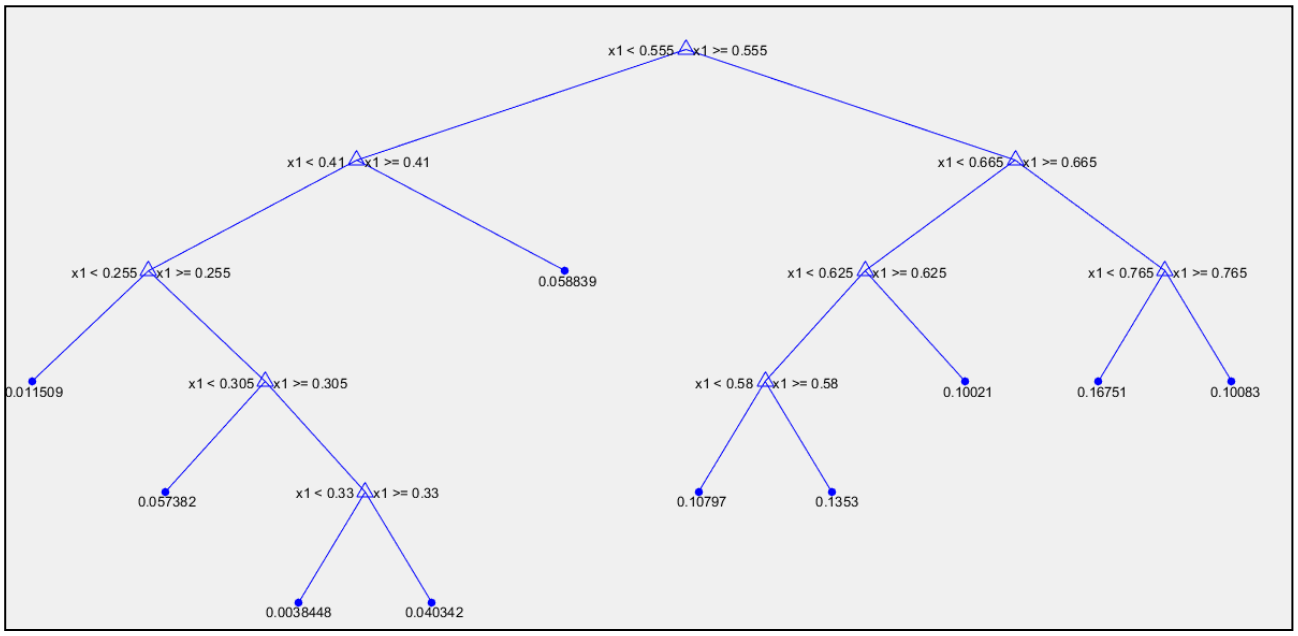

Figure 71: Regression tree for prediction of the intercept in Mean-StDev using PROPWET(x_1), DPSBAR(x_2), BFIHOST(x_3), AREA(x_4) as input predictors.

Four attempts were done with a different number of descriptors used as predictors. It is possible to establish the error committed considering at first one descriptor, and then one more at the time. Analyzing Figure 72 from figures (a) to (d), in this case the pattern the dots assume is linear and does not show relevant changes at the increasing of the descriptor used as input. Case (b) for example shows a slight increase of the RMSE considering the descriptor of land steepness. Considering the other two descriptor, instead, the prediction improves.

RMSE			
(a)	(b)	(c)	(d)
0.1053	0.1077	0.0932	0.0796

Table 8: RMSE values for regression three realized using as input predictors: (a) PROPWET; (b) PROPWET, DPSBAR; (c) PROPWET, DPSBAR, BFIHOST; (d) PROPWET, DPSBAR, BFIHOST, AREA.

Values of the RMSE in this case are quite slower that the slope ones, as in the previous two parameters ($\overline{rc}$, δ) studied.

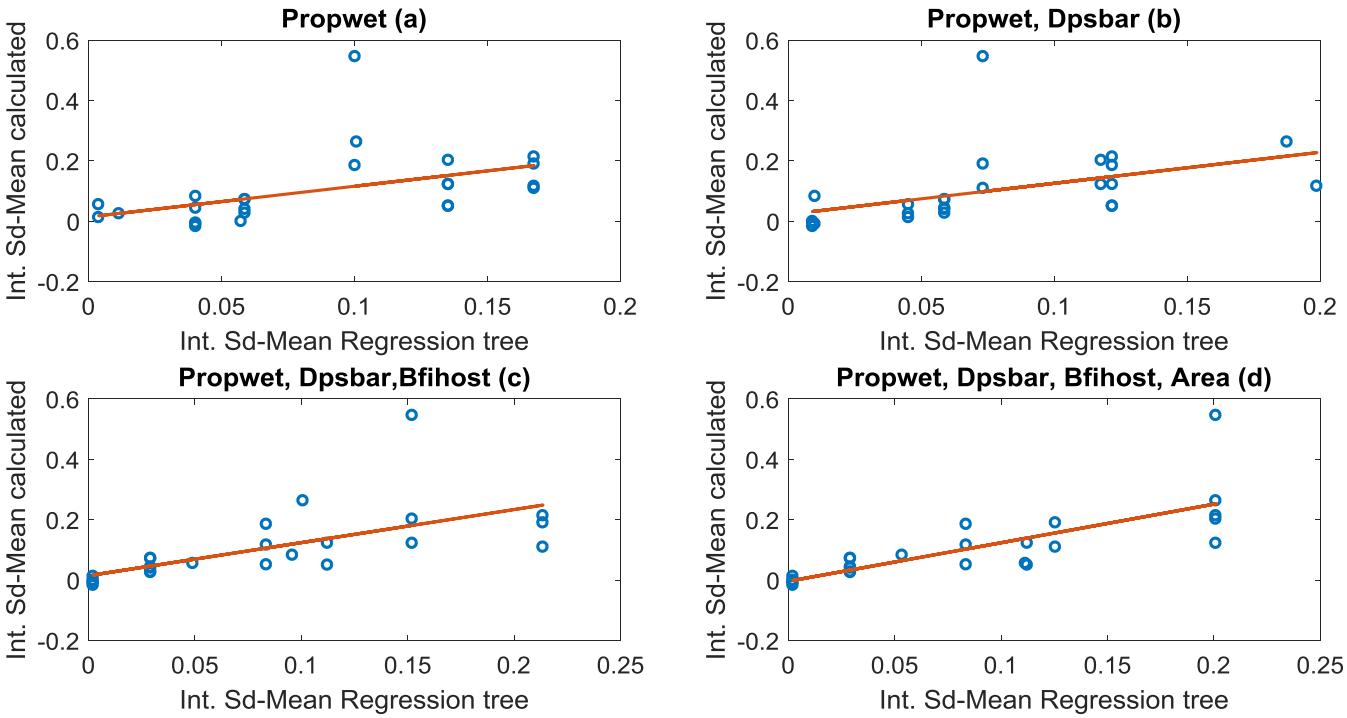

Figure 72: Relation between intercept values in Mean-StDev relation calculated and predicted from regression trees for 25 catchments of the dataset.

Giving a look to Table 13 in Appendix A, errors of prediction are lower than the 30% only for 9 catchments among 25. Hence it is not possible doing any prediction with similar results.

Map of errors related to the prediction of the intercept of the mean and standard deviation relation for runoff coefficients is shown in Figure 73. It is possible to notice that the prediction appears not to be reliable with this results.

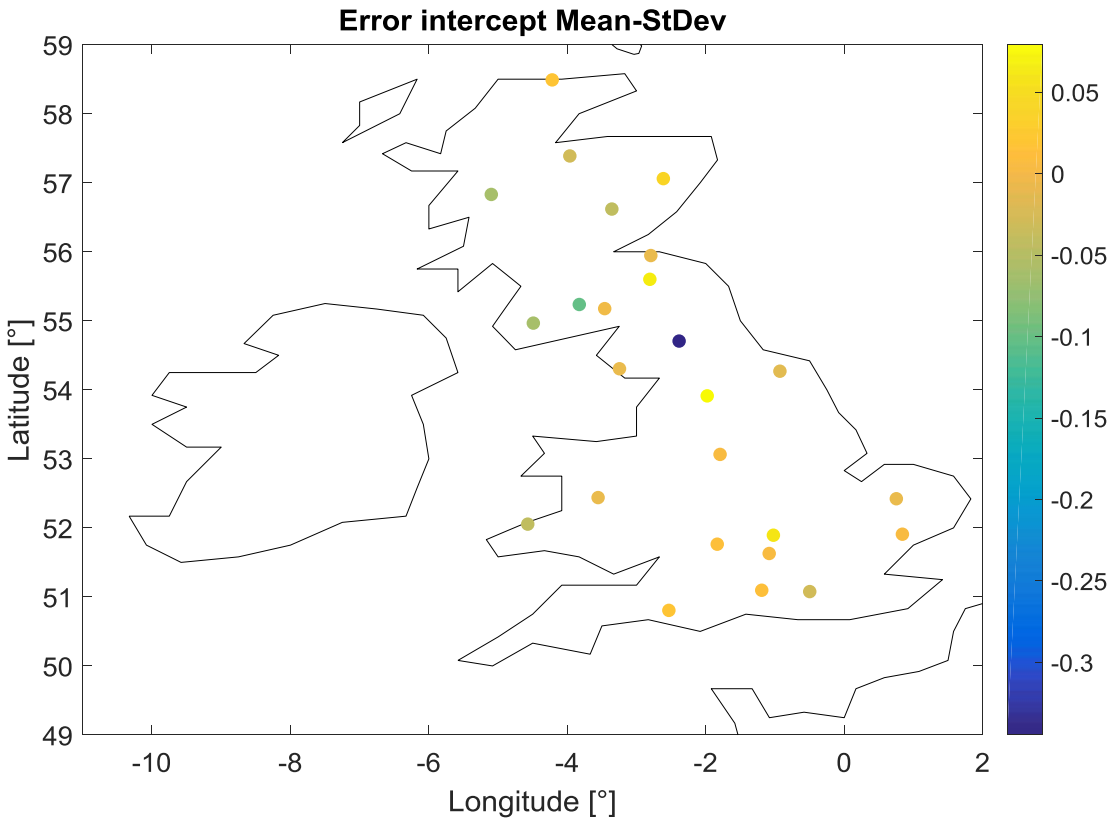

Figure 73: Error of the intercept values in Mean-StDev relation for 25 "ungauged" catchments.

5. Discussion

5.1. Summary of key results

The results obtained from the analysis of 76 catchments spread in the UK permit to do relevant assertions about spatial and temporal variability of runoff coefficients. The data set is not large enough to talk in absolute terms, considering also that only benchmark catchments have been studied, and Northern Ireland was not considered in the analysis. Taking into account each single event of each catchment, it was possible to do some overall considerations. Table 12 summarizes some results useful for them.

Precipitation [mm]	**P< 5**	**5 <P<10**	**10 <P<50**	**50 <P<100**	**P>100**
Thet at Melford Bridge	-	67	185	32	6
Dyfi at Dyfi Bridge	-	32	115	62	80
Nevis at Claggan	-	28	74	58	117
Quickflow [mm]	**QF<3**	**3 <QF<10**	**10 <QF<50**	**50 <QF<100**	**QF>100**
Thet at Melford Bridge	226	58	6	0	0
Dyfi at Dyfi Bridge	108	30	77	36	38
Nevis at Claggan	63	13	62	64	75
Duration [h]	**D<24**	**24 <D<72**	**72 <D<144**	**144 <D<288**	**D>288**
Thet at Melford Bridge	47	16	27	97	103
Dyfi at Dyfi Bridge	45	22	24	73	125
Nevis at Claggan	27	24	28	65	133

Table 9: Summary of events characteristics for the three representative catchments analyzed.

The driest part of UK is placed in the south of the country, mostly regarding the region of the South East, London, East of England and the western part of Wales. It is well distinguished from the other regions that tend to be wetter. Talking about them, it is difficult to make a geographical subdivision about the wettest part of the country. Some sites in the North of Wales and in Scotland (mainly in the eastern coasts and in the North) may be defined the wettest of the country.

The catchment chosen to represent the driest part of the country is Thet at Melford Bridge, whose 290 events within ten years of data were analyzed in term of precipitation, duration and volume. Taking this catchment as an example, and generalizing the result, it is possible to say that volumes of precipitation tend to be small in this part of the country, with a majority of events that involve less than 50 mm or rainfall volume. This fact is directly connected with the quickflow, that is lower than 3 mm for more than the half dataset of events considered. From the event duration analysis, instead, it is underlined that

events tend to be overall longer that six days. Hence, rainfall occur with high frequency and long duration, involving though small volumes of precipitation.

Some words need to be spent about the kind of soil present in the South East of the country. Being the soil mainly characterized by Chalks, the runoff response together with all the hydrological processes of basins are influenced by the presence of a high porosity. The rainfall tend to be absorbed by the soil as soon as it reaches the ground. Being its volume not so relevant, the saturation condition of the soil is rarely reached, so that the groundwater reservoir increases its volume and direct runoff does not reach high values of discharge. Hydrographs of this region in fact, considering in this case four different baseflow separation methods, show that the largest part of the discharge in Chalks catchments is given by the baseflow.

These statements are corroborated by considering the descriptors about soil properties and climate. PROPWET and MAP in particular show that this region is characterized by the smallest rainfall of the country with a mean annual rainfall lower than 800 mm, and a index of the proportion of time in which catchments can be considered wet lower than 0.4.
Besides, also the information about the altitude agrees with this result. This part of the country has the lowest DPSBAR index, with values below 60. It is clear that the less is the steepness of lands, the more the water is slowed down in its path.
Talking about the amount of baseflow related to each catchment, also the descriptor BFIHOST reflects the fact that baseflow is the component of total streamflow that contributes in the largest part. BFIHOST for this region is bigger than 0.5, and also the results from different baseflow filters tested for each catchment showed, although the different patterns, that baseflow dominated the amount of discharge.

The other two catchments analyzed belong both to regions with part of impermeable formations, that facilitate for definition the formation of runoff. In this case, in fact, the main important difference among them is principally due to the slightly different climate of the two regions. Dyfi at Dyfi Bridge is placed in the eastern part of Wales. In term of precipitation, taking into account a dataset of 289 events, half of the events have a rainfall volumes between 10 and 50 mm, the others, precipitation bigger than 50 mm. The quickflow is mostly less than 3 mm, but also higher values occur, probably in relation to the highest precipitation volume. Nevis at Claggan, instead, within its 277 events, is mostly characterized by rainfall events volumes larger than 100 mm, and quickflow values balanced in all the classes created to establish its magnitude. Both of the catchments though, are similar for the duration of events that tend to last more than 12 days.

In term of soil and climate descriptors though, there is not a comparison between them, considering that they belong to the same region distinguished in term of PROPWET, DPSBAR and BFIHOST. The first one is bigger than 0.8, as a synonymous of wet soil for the 80% of the time; the second is bigger than

200, taking into account that these catchments belong to mountainous lands, with a high index of steepness; BFIHOST instead is lower than 0.5, and it shows that baseflow volume does not contribute to the total streamflow as the most representative component, as it happened in the south of the country.

Therefore, all these statements lead to a better comprehension about runoff coefficients range of values. In the South of UK they are overall low. For Thet at Melford Bridge the mean value is 0.0538, and the maximum is above 0.3. The same may be said for the rest of benchmark catchments analyzed in these lands. Dyfi at Dyfi Bridge and Nevis at Claggan show higher values of mean runoff coefficients respectively equal to 0.2873 and 0.4845, and maximum values around the unity. To have an idea about the mean runoff coefficients for all the rest of the dataset, a table will be reported in Appendix A.

Talking about temporal variability of runoff coefficients, a cosine shape curve permitted to represent the monthly trend of runoff coefficients within a year.

For the driest region, it is clear that this pattern is less pronounced because runoff is rather small during the whole year. Even if summer months are characterized by less rainfall and hence a reduced direct runoff, the difference between mean runoff coefficients is not marked within 12 months. This fact is explained with the smallest value of amplitude δ (0.0328) among the three catchments considered.

Catchments spread in the rest of the country tend to have overall high values of runoff coefficients. Looking at the two representative basins chosen, it is possible to say that the precipitation regime is almost the same. The main important difference though, is in the variability of runoff coefficients within months. Dyfi at Dyfi Bridge shows in fact a defined reverse bell shape curve, in which runoff seasonality is well pronounced. It is possible to notice that runoff decreases in summer, while it shows its highest peaks in the winter season. The amplitude value is equal to 0.1593. Nevis at Claggan, instead, shows runoff coefficients that tend to be high during the whole year. There is a slightly decreasing of them during summer, but their mean value stays in a short range of values. The amplitude in this case is equal to 0.0963, higher than for Thet at Melford Bridge, but much more shorter than Dyfi at Dyfi Bridge.

The understanding of the different temporal pattern between Wales and Scotland, talking in general terms, results pretty difficult. First of all, there is a climate difference due to higher precipitation in Scotland than Wales, that involve more water in the soil. Even though the soil is in part formed by impermeable precipitation for both the representative catchments, there are also alluvial permeable deposits that involve infiltration processes. The classification of different 29 soil types with different runoff response should also been taken into account. Lots of processes, then, are not considered in this analysis such as the infiltration, the evaporation process and the interception from coverage. The differences in the amount of quickflow that belong to different classes, for the two last catchments, may

be due to these consideration or, for example, to the different canopy interception that can affect the forming of the runoff.

Another aspect to be considered is the evaluation of the uncertainty related to mean runoff coefficients. From the plot for the all dataset, it was observed for all the 76 catchments that the mean-standard deviation relation follows a linear increasing pattern, that was hence fitted with a straight fitting line. This result occurred considering the monthly runoff coefficient average and the annual. It is hence possible to note, that in general higher mean runoff coefficients are related to higher standard deviation.
This is not completely true in cases in which the mean runoff coefficient does not follow a seasonal variation (lower runoff in summer, higher runoff in winter). For Nevis at Claggan, in fact, the temporal monthly relation between standard deviation and mean runoff coefficient showed a linear pattern, that was almost constant. It may be interpreted saying that a low variation of mean runoff coefficients within a year, implies less deviation around the mean runoff coefficient values.
For other catchments instead, it is not possible to have reliable statements about the trend of the standard deviation. It was observed, for example, that within a year the majority of the mean monthly runoff coefficients had the same standard deviation, while for only two months the standard deviation was completely different. In this cases it is difficult to say that a linear increasing straight line represents well this relation.

The mean-standard deviation relation that takes into account the mean value of runoff coefficient for each year considered in the analysis, was realized to be thorough, but the result we are more interested in is the monthly standard deviation versus the mean, more efficient for investigating about the prediction of the mean runoff coefficient in each month of the year.

Talking about the second part of the analyses that bases on the use of regression trees, it is still not possible to make statements about these instruments for the prediction of monthly mean runoff coefficients. First of all because of the small number of catchment taken into account for the prediction. Secondly, because other investigation should be done in order to find the best combination of catchment characteristics that influences the prediction of runoff response by using regression trees.

From the result obtained, it is possible, though, to affirm that the regression trees worked better in the estimation of the mean runoff coefficient and cosine curve's amplitude. For both of these parameters, in fact, the percentage error in the estimation is less than the 30%. In particular for the amplitude, only 7 catchments among 25 showed errors bigger than the 30%.
The parameter predicted to reconstruct the mean-standard deviation relation instead, resulted to be completely unreliable. For sure the number of catchments analyzed is not enough, but a too large variability in the pattern between mean runoff coefficients and standard deviation was found, to say that

a straight line represents correctly this kind of relation. As a consequence, the prediction gave us too high errors.

5.2. Limitations

Some limitations affected the work since the first steps of the analyses. First of all, the data used in the computation referred to a period of ten years since 1999 to 2008, that can be considered the minimum temporal interval able to show reliable results in statistical calculation. The flow data had gaps for some catchments, so that the analysis implemented for them does not have complete discharge information and so, the whole amount of runoff events within the years. Also the quality of the data lacks of accuracy. The recession analysis done to estimate the recession constant required for the digital filters of baseflow separation showed some noise for small values of discharge, as it was shown in Figure 16.
All the baseflow separation methods tested reveal not to be enough good to estimate the baseflow, as it would be done by a visual inspection. The UKIH method used, overpasses the problem observed with digital filters, in too high or low volume of baseflow, but it is not possible to say with certainty that it represents the correct values of baseflow for each catchment.
Also the event selection can be criticized. It was implemented in order to consider events of every duration, with the duty, after the selection, to remove too short events in term of duration and in term of quickflow volume.
Even if snow affects only the most mountainous part of the country, no models were taken into account to estimate snowfalls and snow melting. The same can be said about processes of evapo-transpiration that were neglected from this analysis. No models of infiltration were used, but the surrogate concept of the wetness of the soil and soil moisture deficit were taken into account by PROPWET and MAP.
The results obtained from the prediction of the 4 quantities discussed above, demonstrate the main limitation of the work that is the small number of catchments used in the analysis. Hydrological processes involve a huge number of variables so that 51 gauged catchments are not enough to do a reliable estimation in term of spatial and temporal variability of runoff coefficients.

However, the analyses conducted on the 76 catchments show important results on the aspects that influence mostly runoff coefficients values. From the spatial variability analysis, it is clear that the climate and the wetness of the soil are the main controls on runoff coefficients. The geology was not deeply investigated, but the simplistic subdivision between Chalk soil and impermeable formations is enough to divide the UK in two main regions that have high and low values of runoff coefficients. Also the morphology of the soil summarized by the descriptor DPSBAR affected the spatial variability of runoff coefficients, while it was less influencing in term of prediction by using regression trees. It is difficult to understand whether the area of the catchments is relevant or not. The largest part of the work refers to volume of flow and rainfall in order to compare the results without considering differences in

term of basin extent. The prediction trees, instead, show that the catchments extent contributes in reducing the RMSE of the prediction. It is better to say that other investigation should be done to affirm that catchments' area is an influencing control.
The land of use was not investigated because no differences occurred in the catchments, being all of them benchmark, so not having any significant anthropogenic change. Talking about temporal variability, the climate and its seasonality are the most influencing factors in runoff coefficients values. More rainfall volumes and more flood event duration are causes of higher runoff coefficients, but more investigation should be done to establish if the larger precipitation volume that occurs in Scotland is the main cause of the less pronounced seasonal pattern observed in the Northern part of the country for runoff coefficients.

5.3. Comparison with similar studies

Spatial and temporal variability of event runoff coefficients was investigated also by other authors in order to understand the main influencing controls in their behavior. Merz et al. (2006), developed a study on 55000 flood events in 337 catchments placed in Austria and assessed that the spatial distribution of runoff coefficients is highly correlated with the mean annual precipitation (MAP), a little correlated with the soil type and land of use. The temporal distribution was instead represented by e Beta distribution whose parameters permit to distinguish six different Austrian regions. The climate variability in Austria was though considered the main cause of the regional patterns found, even if authors did not define the process controls on runoff coefficients (Merz et al., 2009).

The same authors in 2009 worked on the same topic considering a wider dataset of Austrian flood events (64000) from 459 catchments. The spatial variability was analyzed through a correlation analysis between runoff coefficients and catchments attributes, while the temporal distribution was studied comparing the deviation of event runoff coefficients from their mean, depending on event characteristics. They concluded that antecedent soil moisture conditions controls runoff coefficients to a higher degree than does event rainfall [Merz et al., 2009]. In their work, the land use, soil type and geology do not represent a major control on runoff coefficients. By contrast, the results of my study suggest that both event rainfall and soil properties have significant effect on runoff coefficients in the UK. The reasons for this difference are not understood and require further study.

Other authors obtained different results as Gottschalk & Weingartner (1998) that analyzed 17 Swiss watersheds and found that the differences in runoff coefficients could be partially explained by topographic and geomorphologic characteristics, such as altitude, slope, stream network density, and geology [Del Giudice et al., 2014]. Scherrer et al. (2007), analyzed a group of hill-slopes in Switzerland and stated, observing a high variability in the runoff coefficient values, that the parameters used to

represent features such as vegetation, slope, soil clay content, are nor usually directly linked to their response values.

Del Giudice et al., (2014), instead, investigated on the spatial variability of runoff coefficients for 50 watershed placed in Southern Pensinsular Italy with the purpose to improve runoff coefficients prediction for ungauged catchments, by using multivariate regression models taking into account morphological and annual climatic watershed characteristics both. The prediction of flood runoff coefficient was obtained for each gauged watershed by minimizing an objective function of error, in particular the square sum of errors SSE. The correlation between runoff coefficients, a geology-land use combination and climatic information were investigated. The second one gave the better result, and, by using a regression procedure (genetic algorithm GA), the calibration parameters that minimized the SSE were found. The estimated values of runoff coefficients were then compared to the observed values resulted from the application of the rational method for each gauged watershed. Therefore, also this work demonstrates that both soil and climate have a relevant influence on the prediction of event runoff coefficients. Other information such as the watershed area, the river length, the mean river slope and time of concentration, instead, did not show any kind of significant correlation.

A procedure similar to the one adopted in the present work was used by Loritz et al. (2015) for the prediction of mean runoff coefficients. Their investigation based on 60 meso-scale catchments in the Southwest Germany, taking into account both static and dynamic catchments characteristics. The former are physiographic properties, basin extent, catchment characteristics, the latter, instead, the rainfall amount, intensity and initial soil moisture. After they detected runoff events with an automatic method (aet), they managed to individuate which catchments were controlled mostly by storage or by intensity parameters by using random-Forest models (RF). Training two gradient boosted machines (gbm) for 20 storage and 20 intensity dominated catchments, and by using regression trees, they obtained prediction of runoff coefficients with a RMSE less than the 5%.

5.4. Implications for future research

The goal of predicting runoff coefficients for ungauged catchments push us to develop deeper investigations around the main steps that characterize this study, in order to enhance the accuracy of the results.

First of all, a spark for the future could be the studying of a new separation baseflow method. Only three simple methods were implemented in this work: the UKIH method and two one-parameter digital filters. All of them did not show completely satisfactory results in the evaluation of baseflow. Depending on the basins, they gave in response too high or too low baseflow volumes. Multi-parameter digital filters were not tested, because some parameters, except from the regression constant, required data not

available for the analyzed catchments. An hypothesis for the future could be the testing of these methods, with the purpose, though, to estimate all the parameters needed, focusing on which one is more influencing in the baseflow resulting pattern. Otherwise, more researches may be done in order to create a new baseflow method, able to give a reliable response of baseflow for a large dataset, without manual interventions. It is reasonable to think, considering the results obtained in the UK, that, apart from the recession limb of the hydrograph, necessary to visualize the period in which baseflow reveals to be the main component of streamflow, also the geological characteristics of the soil should be taken into account. In the UK in fact, the simple subdivision between Chalk soils, and impermeable formations, showed a large difference in the patterns of the hydrographs and hence in the baseflow volumes.

The consistent results obtained from spatial and temporal variability analyses of runoff coefficients allow to state that they may be predicted for ungauged catchments. Prediction of the temporal pattern within the year and about runoff coefficient uncertainty was done in this work by using recession trees from Matlab. The efficacy of these tools should be probably deepened more, investigating on the descriptors that may be more suitable to represent the catchment response in term of temporal variability. On the other hand, other researches should be done to achieve the knowledge of different regression methods able to catch the catchments characteristics that lead to a better reproduction of their temporal variability. As regards the uncertainty of the prediction, regression trees did not show enough reliable results to continue the research by using them, other methods should be investigated.

Also more analyses around an automated event selection could be a spark for the future. The method used in this work may be considered not fully satisfactory because it was based only on the baseflow separation method without taking into account the developing of peaks of flood or the characteristics time scale of the runoff dynamics. More research should be done to detect the interested events automatically, without requiring further threshold, as done in the current work.

5.5. This work inside the broader research project

It is important now to explain how the study and prediction of runoff coefficients for ungauged catchments is related to the decreasing of the uncertainty within the flood index evaluation approaches.

As already said in the introduction, runoff coefficients are defined as the portion of rainfall that contributes to river flow during a flood event. Therefore, the knowledge of the quickflow and rainfall volumes are necessary for their evaluation. Inverting the problem, the volume of quickflow related to a certain flood event, is known once rainfall volume and runoff coefficients are estimated. In particular, it is given by their product.

Every prediction of flood events, though, is related to a particular period of the year. For this reason the prediction of runoff coefficients has been studied in term of a distribution of them within the time. If we are interested on an event that occurs in March, for example, by the results of the regression, we may reconstruct the temporal pattern, and estimate the respective runoff coefficient. That value is not univocal single number, but it belongs to a range of values indicated by the standard deviation.

Once a range of runoff coefficients is defined for the prediction of a certain flood event within a determined period of the year, information of rainfall are required (e.g. [Stewart et al., 2013]), but for each month. The process of this evaluation represents the next step of this research. The idea is to associate a distribution to the pattern of rainfall versus the time, with the goal to have, as for runoff coefficients, a range of precipitation volumes related to a certain month of the year.

The quickflow volume obtained by this process should be incremented with the respective baseflow volume, whose estimation needs to be investigated with other methods, as already said in the previous section.

The result achieved from the consideration done above is, though, the volume of a flood event. The passage to have a proper value of discharge is expressed in Equation 22, where Q_s is the total streamflow, P is the precipitation, RC is the runoff coefficient, BF is the volume of baseflow and T is the duration of the event (comprising both rainfall event duration and catchment response time).

$$Q_s = \frac{QV + BF}{T} = \frac{P \cdot RC + BF}{T} \quad \frac{[mm]}{[h]} \qquad \text{[Eq. 22]}$$

The prediction of a flood event, comes with the determination of the flood size. For the same dataset used in this work, a flood analysis was carried on, showing that flood seasonality may be described by the seasonality of maximum effective rainfall events for more than the 74% of basins [Zavatteri, 2014]. This was the starting point for an investigation around rainfall-runoff mechanisms, by using bucket models an studying time scale of catchments in order to estimate the flood sizes [Giani, 2017]. All the steps mentioned are therefore fundamental for the prediction of the median peak flood. Once they will be developed with deeper analyses, it is reasonable to think that it will be possible to achieve the goal to obtain a flood index which overpasses the high uncertainty given by using the current statistical procedure.

6. Conclusion and future developments

The present work develops with the purpose to contribute around the problem of uncertainty in flood estimation. In particular, the main goal is related to statistical models used to predict the flood discharge for a certain return period by considering the product between a growth curve and a flood index. Indeed, the largest uncertainties are related to the flood index that, as Kjeldsen et al. (2014) observed, is affected by a factorial standard error of 1.5. The final objective is hence the evaluation of a flood index less affected by this uncertainty which is too large to make reliable predictions.

This thesis contributes in this project investigating around the concept of runoff coefficients. A dataset of 76 benchmark catchments was analyzed, considering hourly rainfall and discharge data for a period of ten years, from 1999 to 2008. The analysis of the catchments started with the application of baseflow separation methods, in order to estimate the baseflow volumes for each basin. The UKIH method was judged as the more reliable, because it gave a baseflow in compliance with the one that would have been traced manually.

In accordance with the baseflow separation method, the selection of runoff events was done establishing that each event was characterized by a starting and a ending time steps. Both of them represent the intersection points between the total streamflow and baseflow, but within whom, streamflow is bigger than baseflow. Rainfall volume of each event was though considered in relation to a wider temporal interval, that started 24 hours before the formal starting time step of each event.

The event selection permitted, hence, to create a database formed by 20000 events. Each event was characterized by a duration, rainfall volume, quickflow and baseflow volume, and a runoff coefficient. Runoff events with rainfall volumes lower than 5 mm were discarded from the analysis because the main interest was given to bigger events.

A spatial and temporal variability analysis was then conducted on the runoff coefficients of each event, in order to investigate the influencing controls on runoff coefficients values. The spatial variability was studied taking into account soil properties, geomorphology and climate properties through catchments descriptors such as PROPWET, DPSBAR, BFIHOST, MAP. The land use was not considered in this analysis. The temporal variability, instead, was analyzed studying runoff coefficients within months, and the relation between their deviation around their mean value.

It is possible to state that the soil properties, together with land steepness and climate have a relevant influence of runoff coefficients. Different classes of these descriptors permitted to divide UK in different regions characterized by different range of runoff coefficient. The temporal analysis revealed for each catchment a cosine shape pattern, with a larger or lower amplitude in relation to the

precipitation regime and geology of the soil. More investigation should be done to make statements around the relation between mean runoff coefficient and standard deviation, because the straight fitting line considered to represent this pattern is not considered satisfactory.

The prediction of mean runoff coefficients was then realized by using regression trees based on catchments descriptor as input data. The data set was divided in two groups considered similar in term of wetness of the soil: 51 catchments were considered as "gauged" and 25 as "ungauged". This was done with the objective to check if the calculated runoff coefficients for all the catchments agreed with the regression trees' prediction.

Regression trees prediction, though, seems to be more reliable for the estimation of the monthly temporal pattern of runoff coefficients. Both the estimation of the mean runoff coefficients and the monthly amplitude show percentage errors less than the 30% for the majority of the catchments considered. On the other hand, the prediction of the slope and intercept that permit to estimate the deviation of runoff coefficients around their mean value within the year, does not have good results, the percentage errors are too large for the majority of the catchments, so that this prediction cannot be considered enough reliable.
Overall, more investigation should be done to improve the prediction with these regression instruments.

In conclusion, this study is a relevant starting point in the knowledge of runoff coefficients and in their prediction for ungauged catchments in the UK. More investigation are required for increasing the reliability of the results. The consistency of spatial and temporal patterns of runoff coefficients obtained for a small dataset of 76 catchments, though, reveals a reasonable potential in the prediction of runoff coefficients for ungauged catchments.

7. Appendix A

7.1. Tables of predicting errors.

Mean runoff coefficient.

Catchment	$\widehat{RC}$ predicted	RC calculated	Error ($\widehat{RC}$- RC)	Error (%)
28	0.0907	0.1266	-0.0359	-28.3
35	0.0381	0.0050	0.0331	661.5
26	0.0907	0.0552	0.0355	64.5
30	0.1409	0.1573	-0.0164	-10.5
32	0.0381	0.0280	0.0101	36.0
39	0.0381	0.0120	0.0260	216.1
38	0.1409	0.2286	-0.0877	-38.4
41	0.0381	0.0239	0.0142	59.5
17	0.0907	0.1706	-0.0798	-46.8
7	0.0907	0.1101	-0.0194	-17.6
21	0.0907	0.0727	0.0180	24.7
49	0.0907	0.1590	-0.0682	-42.9
4	0.2650	0.2636	0.0014	0.5
8	0.2650	0.1894	0.0756	39.9
47	0.4046	0.2297	0.1749	76.2
2	0.2650	0.2483	0.0167	6.7
20	0.2726	0.2838	-0.0112	-4.0
66	0.2726	0.3619	-0.0893	-24.7
14	0.2726	0.5572	-0.2846	-51.1
69	0.2650	0.3211	-0.0561	-17.5
56	0.2650	0.3726	-0.1076	-28.9
72	0.3025	0.3503	-0.0478	-13.6
61	0.4046	0.4355	-0.0309	-7.1
76	0.3025	0.2437	0.0588	24.1
73	0.4046	0.4980	-0.0934	-18.8

Table 10: Comparison between mean runoff coefficient calculated and predicted.

Amplitude δ

Catchment	$\hat{\delta}$ predicted	δ calculated	Error ($\hat{\delta}$-δ)	Error (%)
28	0.1268	0.1341	-0.0072	-5.4
35	0.0056	-0.0022	0.0079	-351.2
26	0.0507	0.0328	0.0179	54.7
30	0.1268	0.1638	-0.0370	-22.6
32	0.0056	0.0060	-0.0004	-6.1
39	0.0056	0.0023	0.0033	143.4
38	0.1268	0.2138	-0.0870	-40.7
41	0.0056	0.0186	-0.0130	-69.7
17	0.0848	0.0799	0.0049	6.2
7	0.0848	0.0724	0.0124	17.1
21	0.0507	0.0548	-0.0042	-7.6
49	0.0848	0.1050	-0.0202	-19.2
4	0.1594	0.1506	0.0088	5.8
8	0.1095	0.1348	-0.0253	-18.8
47	0.1594	0.1511	0.0082	5.5
2	0.1095	0.1342	-0.0248	-18.5
20	0.1594	0.1759	-0.0165	-9.4
66	0.1594	0.1908	-0.0314	-16.5
14	0.2461	0.1304	0.1157	88.8
69	0.1095	0.1531	-0.0436	-28.5
56	0.1821	0.1672	0.0149	8.9
72	0.1821	0.1416	0.0405	28.6
61	0.1095	0.1063	0.0032	3.0
76	0.1821	0.1286	0.0535	41.6
73	0.1095	0.0963	0.0132	13.7

Table 11: Comparison between values of amplitude calculated and predicted.

Slope of the straight fitting curve of Mean-Sdev relation of runoff coefficients

Catchment	$\hat{s}$ predicted	s calculated	Error ($\hat{s}$-s)	Error (%)
28	0.4019	0.5023	-0.1003	-20.0
35	1.2613	1.8842	-0.6228	-33.1
26	0.8241	0.6492	0.1749	26.9
30	0.7197	0.5773	0.1424	24.7
32	1.2613	1.7160	-0.4547	-26.5
39	0.8241	1.9538	-1.1297	-57.8
38	0.7197	0.5117	0.2080	40.6
41	1.2613	2.5073	-1.2459	-49.7
17	0.7197	0.5671	0.1526	26.9
7	0.7197	0.7826	-0.0629	-8.0
21	1.2613	0.6974	0.5639	80.9
49	0.4122	0.6880	-0.2758	-40.1
4	0.4122	0.4905	-0.0783	-16.0
8	0.4122	0.5710	-0.1588	-27.8
47	0.1925	0.3284	-0.1359	-41.4
2	0.4122	0.5055	-0.0933	-18.4
20	0.1925	0.2534	-0.0609	-24.0
66	0.1925	0.1313	0.0612	46.6
14	0.1925	-0.4887	0.6812	-139.4
69	0.4122	0.1410	0.2712	192.4
56	0.4122	0.3093	0.1030	33.3
72	0.2580	0.0304	0.2276	749.2
61	0.1925	0.0951	0.0975	102.5
76	0.2580	0.4275	-0.1695	-39.6
73	0.1925	0.0364	0.1561	429.0

Table 12: Comparison between values of slope in the Mean-StDev calculated and predicted.

Intercept of the straight fitting curve of Mean-Sdev relation of runoff coefficients.

Catchment	$\hat{\imath}$ predicted	i calculated	Error ($\hat{\imath}$ - i)	Error (%)
28	0.0292	0.0250	0.0043	17.1
35	0.0022	-0.0005	0.0027	-530.7
26	0.0022	0.0123	-0.0101	-82.0
30	0.1114	0.0550	0.0564	102.6
32	0.0022	-0.0058	0.0080	-138.3
39	0.0022	-0.0092	0.0114	-124.2
38	0.0535	0.0824	-0.0289	-35.0
41	0.0022	-0.0172	0.0194	-112.9
17	0.0292	0.0433	-0.0141	-32.5
7	0.0292	0.0396	-0.0104	-26.2
21	0.0292	0.0268	0.0024	8.9
49	0.0292	0.0719	-0.0427	-59.3
4	0.0292	0.0714	-0.0422	-59.1
8	0.1123	0.0496	0.0627	126.5
47	0.1123	0.1221	-0.0098	-8.0
2	0.0836	0.0506	0.0330	65.2
20	0.2009	0.1218	0.0791	65.0
66	0.2009	0.2019	-0.0010	-0.5
14	0.2009	0.5454	-0.3445	-63.2
69	0.0836	0.1847	-0.1011	-54.7
56	0.0836	0.1160	-0.0324	-27.9
72	0.1255	0.1896	-0.0640	-33.8
61	0.2009	0.2131	-0.0122	-5.7
76	0.1255	0.1090	0.0165	15.1
73	0.2009	0.2627	-0.0617	-23.5

Table 13: Comparison between values of intercept in the Mean-StDev calculated and predicted.

7.2. Main relevant characteristics for the 76 basins dataset

Table 14 shows catchment descriptors, mean runoff coefficient and amplitude from the temporal pattern, slope and intercept from the straight fitting curve of Mean-Sdev relation of runoff coefficients.

MATLAB ID	Name of the Cathcment	Station Number	Area (km2)	MAP (mm)	BFIHOST	DPSBAR [m/km]	PROPWET	Mean RC	d	slope	intercept	N° event
1	Urie at Pitcaple	11004	198	959	0.562	88.4	0.53	0.071	0.040	1.318	-0.003	385
2	Dee at Woodend	12001	1370	1167	0.506	185.7	0.62	0.248	0.134	0.505	0.051	295
3	Dee at Polhollick	12003	690	1289	0.459	219.6	0.68	0.302	0.158	0.329	0.104	292
4	Ericht at Craighall	15025	432	1234	0.489	173	0.55	0.264	0.151	0.491	0.071	299
5	Ruchill Water at Cultybraggan	16003	99.5	2141	0.428	216.4	0.59	0.414	0.168	0.148	0.173	258
6	Allan Water at Kinbuck	18001	161	1565	0.507	92.8	0.59	0.265	0.145	0.217	0.132	296
7	Gifford Water at Lennoxlove	20007	64	820	0.527	112.5	0.43	0.110	0.072	0.783	0.040	341
8	Tweed at Boleside	21006	1500	1272	0.496	191.9	0.58	0.189	0.135	0.571	0.050	319
9	Leet Water at Coldstream	21023	113	729	0.388	34.5	0.3	0.184	0.173	0.543	0.077	243
10	Jed Water at Jedburgh	21024	139	962	0.436	111.6	0.57	0.223	0.173	0.318	0.096	283
11	Tima Water at Deephope	21026	31	1676	0.37	172.8	0.72	0.397	0.168	0.241	0.179	253
12	Coquet at Morwick	22001	569.8	937	0.393	110.1	0.44	0.173	0.115	0.462	0.073	276
13	South Tyne at Haydon Bridge	23004	751.1	1202	0.298	107.1	0.6	0.265	0.161	0.349	0.104	302
14	Trout Beck at Moor House	25003	11.4	1950	0.227	91.9	0.64	0.557	0.130	-0.489	0.545	223
15	Greta at Rutherford Bridge	25006	86.1	1178	0.241	68.3	0.62	0.399	0.246	0.177	0.183	248

16	Aire at Kildwick Bridge	27035	282.3	1252	0.385	99.6	0.62	0.242	0.168	0.284	0.144	291
17	Dove at Kirkby Mills	27042	59.2	1015	0.495	136	0.4	0.171	0.080	0.567	0.043	304
18	Crimple at Burn Bridge	27051	8.1	895	0.309	62.9	0.34	0.255	0.177	0.288	0.125	225
19	Swale at Crakehill	27071	1363	923	0.436	66.9	0.38	0.207	0.140	0.355	0.092	268
20	Eastburn Beck at Crosshills	27084	43.4	1141	0.315	114.2	0.62	0.284	0.176	0.253	0.122	233
21	Dove at Izaak Walton	28046	83	1180	0.651	141.3	0.46	0.073	0.055	0.697	0.027	398
22	Greet at Southwell	28072	46.2	711	0.623	44	0.27	0.056	0.024	0.865	0.009	171
23	West Glen at Easton Wood	31023	4.4	708	0.32	32.8	0.27	0.169	0.106	0.209	0.113	183
24	Kym at Meagre Farm	33012	137.5	637	0.309	26.1	0.24	0.096	0.106	0.930	0.013	127
25	Tove at Cappenham Bridge	33018	138.1	737	0.368	37.1	0.3	0.113	0.092	0.677	0.031	268
26	Thet at Melford Bridge	33019	316	713	0.707	13.9	0.31	0.055	0.033	0.649	0.012	290
27	Stringside at Whitebridge	33029	98.8	727	0.864	13.5	0.23	0.031	0.017	0.749	0.010	328
28	Colne at Lexden	37005	238.2	650	0.537	30.4	0.25	0.127	0.134	0.502	0.025	249
29	Pincey Brook at Sheering Hall	38026	54.6	648	0.388	24.1	0.31	0.119	0.103	0.708	0.011	196
30	Ray at Grendon Underwood	39017	18.8	690	0.238	28	0.32	0.157	0.164	0.577	0.055	219
31	Lambourn at Shaw	39019	234.1	797	0.839	59.2	0.32	0.020	0.003	2.030	-0.013	397
32	Coln at Bibury	39020	106.7	928	0.858	76.8	0.33	0.028	0.006	1.716	-0.006	455
33	Dun at Hungerford	39028	101.3	896	0.768	46.3	0.31	0.013	0.003	1.542	-0.002	450
34	Evenlode at Cassington Mill	39034	430	786	0.699	46.5	0.32	0.077	0.047	0.848	0.019	297

35	Ewelme Brook at Ewelme	39065	13.4	734	0.933	78.4	0.29	0.005	-0.002	1.884	-0.001	480
36	Great Stour at Horton	40011	345	814	0.706	50.9	0.34	0.109	0.056	0.474	0.025	286
37	Meig at Glenmeanie	4005	120.5	2305	0.389	289	0.76	0.422	0.128	0.087	0.209	274
38	Loxwood Stream at Drungewick	41025	91.6	909	0.321	56.4	0.35	0.229	0.214	0.512	0.082	243
39	Cheriton Stream at Sewards Bridge	42008	75.1	986	0.941	54.2	0.34	0.012	0.002	1.954	-0.009	480
40	East Avon at Upavon	43014	85.8	878	0.872	43.1	0.34	0.050	0.003	0.919	-0.006	353
41	Sydling Water at Sydling St Nicholas	44006	12.4	1213	0.879	128.9	0.38	0.024	0.019	2.507	-0.017	475
42	Wey at Broadwey	44009	7	964	0.783	117.7	0.38	0.076	0.001	1.263	-0.010	481
43	East Dart at Bellever	46005	21.5	2336	0.363	95	0.46	0.267	0.123	0.264	0.136	322
44	Tiddy at Tideford	47009	37.2	1405	0.591	121.2	0.48	0.098	0.062	0.929	0.035	337
45	Warleggan at Trengoffe	48004	25.3	1509	0.499	93.8	0.45	0.107	0.058	0.931	0.022	381
46	Wellow Brook at Wellow	53009	72.6	1092	0.644	68.5	0.37	0.131	0.090	0.523	0.043	200
47	Dulas at Rhos-y-pentref	54025	52.7	1370	0.439	161.4	0.59	0.230	0.151	0.328	0.122	131
48	Cynon at Abercynon	57004	106	1992	0.421	142.3	0.54	0.226	0.153	0.552	0.077	305
49	Teifi at Glanteifi	62001	893.6	1478	0.507	109.8	0.52	0.159	0.105	0.688	0.072	338
50	Dyfi at Dyfi Bridge	64001	471.3	2028	0.478	269.8	0.66	0.295	0.159	0.401	0.121	289
51	Glaslyn at Beddgelert	65001	68.6	3100	0.406	309.3	0.62	0.394	0.082	0.267	0.184	302
52	Wheeler at Bodfari	66004	62.9	949	0.696	108.6	0.38	0.042	0.025	1.506	-0.006	412
53	Dee at New Inn	67018	53.9	2295	0.312	149	0.71	0.397	0.164	0.120	0.205	271

54	Dane at Rudheath	68003	407.1	933	0.459	59.9	0.4	0.165	0.100	0.480	0.038	278
55	Weaver at Audlem	68005	207	797	0.502	27.4	0.34	0.126	0.117	0.311	0.062	254
56	Findhorn at Shenachie	7001	415.6	1337	0.451	138.9	0.68	0.373	0.167	0.309	0.116	277
57	Lossie at Torwinny	7006	20	1083	0.296	86.6	0.42	0.163	0.073	0.529	0.073	338
58	Ribble at Samlesbury	71001	1145	1454	0.37	95	0.57	0.300	0.162	0.264	0.120	276
59	Conder at Galgate	72014	28.5	1311	0.443	93.4	0.6	0.261	0.144	0.308	0.110	267
60	Kent at Sedgwick	73005	209	1872	0.514	153.5	0.71	0.267	0.123	0.483	0.097	305
61	Duddon at Duddon Hall	74001	85.7	2377	0.338	210.3	0.71	0.435	0.106	0.095	0.213	264
62	Esk at Cropple How	74007	70.2	2385	0.417	231.7	0.71	0.377	0.065	0.126	0.236	174
63	Ellen at Bullgill	75017	96	1180	0.488	78.4	0.62	0.204	0.154	0.558	0.053	298
64	Petteril at Harraby Green	76010	160	991	0.59	51.5	0.64	0.176	0.171	0.372	0.064	257
65	Eden at Kirkby Stephen	76014	69.4	1647	0.409	148.9	0.68	0.431	0.181	0.037	0.239	250
66	Kinnel Water at Redhall	78004	76.1	1600	0.431	98.7	0.62	0.362	0.191	0.131	0.202	255
67	Nith at Friars Carse	79002	799	1576	0.433	157.8	0.67	0.278	0.175	0.117	0.156	285
68	Nith at Hall Bridge	79003	155	1641	0.357	121.7	0.63	0.342	0.208	0.203	0.153	280
69	Scar Water at Capenoch	79004	142	1789	0.446	195.9	0.66	0.321	0.153	0.141	0.185	282
70	Avon at Delnashaugh	8004	542.8	1175	0.451	178.2	0.63	0.227	0.115	0.550	0.064	317
71	Dulnain at Balnaan Bridge	8009	272.2	1104	0.498	117.9	0.68	0.286	0.198	0.348	0.083	278
72	Cree at Newton Stewart	81002	368	1935	0.341	118.7	0.69	0.350	0.142	0.030	0.190	267
73	Nevis at Claggan	90003	69.2	3275	0.428	441.8	0.81	0.498	0.096	0.036	0.263	277

74	Ewe at Poolewe	94001	441.1	2480	0.365	222.1	0.83	0.156	0.053	0.581	0.093	370
75	Inver at Little Assynt	95001	137.5	2297	0.399	190.3	0.77	0.152	0.057	0.497	0.109	400
76	Naver at Apigill	96002	477	1427	0.338	111.7	0.73	0.244	0.129	0.427	0.109	323

Table 14: Main characteristics of the 76 catchments

8. References

Papers

Archer D.R., (1981). Seasonality of flooding and the assessment of seasonal flood risk. Proc. Instn. Civ. Engrs., Part 2, 1981, 70, December, 1023-1035.

Bayliss A.C. and Jones R.C. (1993). Peaks-over-threshold flood database: summary statistics and seasonality. Report No. 121. Institute of Hydrology. Wallingford.

Bell V. A., et al., (2007). Development of a high resolution grid-based river flow model for use with regional climate model output. Hydrol. Earth Syst. Sci., 11(1), 532-549, 2007.

Berghuijs W. et al., (2016). Dominant flood generating mechanisms across United States. Geophysical research letters, DOI: 10.1002/2016GL068070.

Blenkinsop S., Lewis E. et al., (2017). Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. Int. J. Climatol. 37:722-740(2017), DOI: 10.1002/joc.4735.

Bloschl G. et al., (2017).Changing climate shifts timing of European floods. Science 357, 588-590 (2017).

Blume T., Zehe E., Bronstert A., (2007). Rainfall-runoff response, event-based runoff coefficients and hydrograph separation, Hydrological Science Journal, 52:5, 843-862, DOI: 10.1623/hysj.52.5.843.

Boorman D. B., Hollis J. M. and Lilly A., (1995). Hydrology of Soil Types: a hydrologically-based classification of the soils of the United Kingdom. IH Report No. 126. Institute of Hydrology, Wallingford.

Chapman T., (1999). A comparison of algorithms for stream flow recession and baseflow separation. Hydrol. Process. 13, 701-714 (1999).

Chow V. T., Maidment D. R., Mays L. W., (1988). Applied hydrology. GB661.2.C43 1988 627 87-16860.

Claps P. and Laio F., (2008). Aggiornamento delle procedure di valutazione delle piene del Piemonte, con particolare riferimento ai bacini sottesi da invasi artificiali (Volume I).

Collischonn W. et al., (2013). Defining parameters for Eckhardt's digital baseflow filter. Hydrol. Process. 27, 2614-2622 (2013).

Del Giudice G. and Padulano R., (2014). Spatial prediction of the runoff coefficient in Southern Peninsular Italy for the index flood estimation. Hydrology Research 45.2 2014.

Flood estimation handbook (FEH), (1999). Centre for Ecology & Hydrology, Wallingford. Volumes I, III, IV and V.

Giani, G. et al. (2017). Improved flood estimation using hydrological process knowledge. Unpublished Master's thesis, Turin Polytechnic, Italy.

Gotvald A. J., Barth N. A., Veilleux A. G., Parret C., (2012). Methods for Determining Magnitude and Frequency of Floods in California, based on Data through Water year 2006. U.S. Geological Survey Scientific Investigations Report 2012–5113, 38 p., 1 pl.

Gustard A., Bullock A. and Dixon J. M., (1992). Low flow estimation in the United Kingdom, Report N°. 108 Institute of Hydrology.

Jennings, M. E., et al. (1994). Nationwide Summary of U.S. Geological Survey regional Regression Equations for Estimating Magnitude and Frequency of Floods for Ungauged Sites, 1993. U.S. GEOLOGICAL SURVEY. Water-Resources investigations Report 94-4002

Keller V. et al. (2006). Environment Agency R & D Project W6-101 Continuous Estimation of River Flows (CERF), Technical Report: Task 1.1: Estimation of Precipitation Inputs.

Keller V. et al., (2015). CEH-GEAR: 1 km resolution daily and monthly areal estimates for the UK for hydrological and other applications. Earth Syst. Sci. Data, 7, 143–155, 2015. DOI:10.5194/essd-7-143-2015.

Kjeldsen T.R. (2014). How reliable are design flood estimates in the UK?, Journal of Flood Risk Management, DOI: 10.1111/jfr3.12090.

Kjeldsen T. R., (2007). The revitalized FSR/FEH rainfall-runoff method. Flood estimation Handbook, Supplementary Report N. 1.

Laio F. et al. (2011). Spatially smooth regional estimation of the flood frequency curve (with uncertainty). Journal of Hydrology 408, 67-77.

Li L., Maier H. R., et al., (2013). Framework for assessing and improving the performance of recursive digital filters for baseflow estimation with application to the Lyne and Hollick filter. Environmental Modelling and Software 41 (2013) 163-175.

Loritz R. et al., (2015). Data mining methods for predicting event runoff coefficients in ungauged basins using static and dynamic catchment characteristics. Geophysical Research Abstracts. Vol. 17, EGU2015-11072, 2015.

Manfreda S., Giordano C., Iacobellis V., (2003). Stima dei deflussi di base mediante un filtro fisicamente basato. Giornata di studio: Metodi statistici e matematici per l'Analisi delle serie Idrologiche-Roma 2003.

Mei Y., Anagnostou E. N., (2015). A hydrograph separation method based on information from rainfall and runoff records. Journal of Hydrology 523 (2015) 636-649.

Merz R., Bloschl G., Parajka J., (2006). Spatio-temporal variability of event runoff coefficients. Journal of Hydrology (2006) 331, 591-604.

Merz R. and Bloschl G., (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. WATER RESOURCES RESEARCH, VOL. 45, W01405, DOI:10.1029/2008WR007163.

Moisello U. (1998). Idrologia tecnica. La Goliardica Pavese s.r.l., Pavia, Italy.

Mugo J. M. and Sharma T. C., (1999). Application of a conceptual method for separating runoff components in daily hydrographs in Kimakia forest catchments, Kenya. Hydrol. Process. 13, 2931-2939.

Nathan R. J. and McMahon T. A., (1990). Evaluation of Automated Techniques for base Flow and Recession Analyses. Water resources research, Vol. 26, n°. 7, pages 1465-1473, July 1990.

Osborn T. and Maraun D. (2008). Changing intensity of rainfall over Britain, Climatic Research Unit Information Sheet no. 15.

Tallaksen L. M., (1995). A review of baseflow recession analysis, Journal of Hydrology 165 (1995) 349-370.

Painter C. C., Heimann D. C., Lanning-Rush J. L., (2017). Methods for Estimating Annual Exceedance-Probability Streamflows for Streams in Kansas based on Data Through Water Year 2015. (ver. 1.1, September 2017): U.S. Geological Survey Scientific Investigations Report 2017–5063, 20 p., doi.org/10.3133/sir20175063.

Rupp D. E. and Selker J. S., (2005). Information, artifacts, and noise in $\frac{dQ}{dt} - Q$ recession analysis, Advances in Water Resources, February 2006. DOI: 10.1016/j.advwatres.2005.03.019.

Stewart M. K., (2015). Promising new baseflow separation and recession analysis methods applied to streamflow at Glendhu Catchment, New Zealand. Hydrol. Earth Syst. Sci., 19, 2587–2603. DOI:10.5194/hess-19-2587-2015.

Stuckey M. H. and Reed L. A., (2000). Techniques for estimating magnitude and frequency of peak flows for Pennsylvania streams. Water-Resources Investigations Report 00-4189.

Tan S. B. K., Lo E. Y., et al., (2009). Hydrograph Separation and Development of Empirical Relationships Using Single-Parameter Digital Filters. DOI:10.1061/(ASCE)1084-0699(2009)14:3(271).

Viglione A. et al., (2010). Quantifying space-time dynamics of flood events types. Journal of Hydrology 394 (2010) 213-229.

Wood E. F., Sivapalan M., Beven K., (1990). Similarity and scale in catchment storm response. Reviews of Geophysics, 28.

Zavatteri, M. et al. (2015). Flood Seasonality Analysis in Great Britain. Unpublished Master's thesis, Turin Polytechnic, Italy.

Zhang J., Zhang Y., Song J., Chang L., (2017). Evaluating relative merits of four baseflow separation methods in Eastern Australia, Journal of Hydrology 549 (2017) 252-263.

Website

http://www.metoffice.gov.uk/climate/uk/regional-climates

https://www.ceh.ac.uk/services/land-cover-map-2007

https://it.mathworks.com