Chapter 2

Foundations of the Study

2.1 Introduction

The role and importance of mathematical models in the planning field has been addressed by many authors (e.g. Forrester, 1969, Steiss 1974, Karlqvist 1978, Allen 1997). By simulating a part of "reality" or a *system*[1] using a set of rules and algorithms that mimic the behaviour and relationships of the observed data, a modeller may gain expertise, get a deep understanding of the underlining processes and their mutual interactions, forecast future trends, and estimate likely outcomes of plausible scenarios.

However, data availability must be "carefully considered" (Wilby 1997) before any modelling attempt is carried out. This implies that a *model*[2] should have variables that can be obtained or derived either from existing data bases or by direct surveying. This remark is of great importance when a model is projected to become a planning tool; in other words, it must avoid variables that cannot be estimated because there is a lack of technical capabilities, or their acquisition is too costly or, even worse, it is too complex or even impossible to acquire them. If these guidelines are not observed, a model, perhaps interesting from a theoretical point of view, would be unpractical and most probably misleading in the realm of planning. Due to the reasons stated above, it is worthy at the present stage of this research to have an overview of the relevant information available for the chosen Study Area.

---

[1] In general a system is "an orderly complex of elements and patterns constituting a functioning whole" (McArthur, 1992). More formally (based on Casti, 1984), let A be a set of abstract states (i.e. which occur in both space and time) of a natural system N and P be some range of potentially observed values. Then an observable of N is a mapping $x : \mathrm{A} \rightarrow \mathrm{P}$. "An observation, by contrast, is the concrete realization of an observable". Usually the potentially observed values are a subset of a n dimensional euclidean space, i.e. $\mathrm{P} \subset \mathbb{R}^n$. Then, "a natural system N comprises the sets A and P, together with a collection of observables $\{x_i\}$. In other words, a system N is defined by what we see, and the usual labels for systems, such as a 'national economy', a 'vibrating string', a 'chemical plant', and so forth, are interpreted only as conventions or labels introduced to account for situations where most observers are equipped to 'see' with the same set of observables, using the same sets A and P. Thus, by changing the sets or the observables, the same system N may be seen quite differently by different observers."

[2] A formal system M (i.e. a man-made construct composed of axioms, symbols, equations, and rules of logical inference) is called a model of N only if 1) it establishes a faithful correspondence between its elements and the observables $\{x_i\}$ and linkages of N; and 2) it can be used for describing N to a prescribed degree of accuracy (Casti, 1984). Moreover, if a model includes a set of rules describing the transitions from one state in time to another then it can be used for making specific predictions of N.

2.2 General Description of the Study Area

The Study Area is located to the south and southeast of Stuttgart, Germany. It comprises the upper catchment of the Neckar River upstream of the Plochingen gauging station (N3 530 930 and E5 396 740 m in Gauss-Krueger coordinates) covering an area of about 4002 km^2. The Neckar is a right-bank tributary of the Rhine; it is 367 km long and flows 40% of its course in direction north and northeast within the Study Area. The Upper Neckar Catchment (see Figure 2.1) is bounded by the north-western edge of the Swabian Jura on the right bank side of the Neckar and by the Black Forest on its left bank. Its elevation ranges from 240 to 1014 m.a.s.l. (from DEM LFU) and has a mean elevation of 546 m.a.s.l.. Slopes are in general mild; 90% of its area has slopes varying from 0° to 15°, although some areas in the Swabian Jura or in the Black Forest may have values as high as 50°.

The main geological formations in the Upper Neckar Catchment are originated in the Triassic and Jurassic periods, both corresponding to the Mesozoic Era. The main formations are composed of altered *keuper*, claystone-jura, claystone-keuper, limestone-jura, loess, sandstone and shelly limestone (*Muschelkalk*). Conversely, the river bed of the Neckar and its tributaries are relatively young compared with the previous formations. They are mainly composed of Quaternary sediments originated mainly from the erosion of outcrops of aforementioned rock types (Geyer and Gwinner 1991).

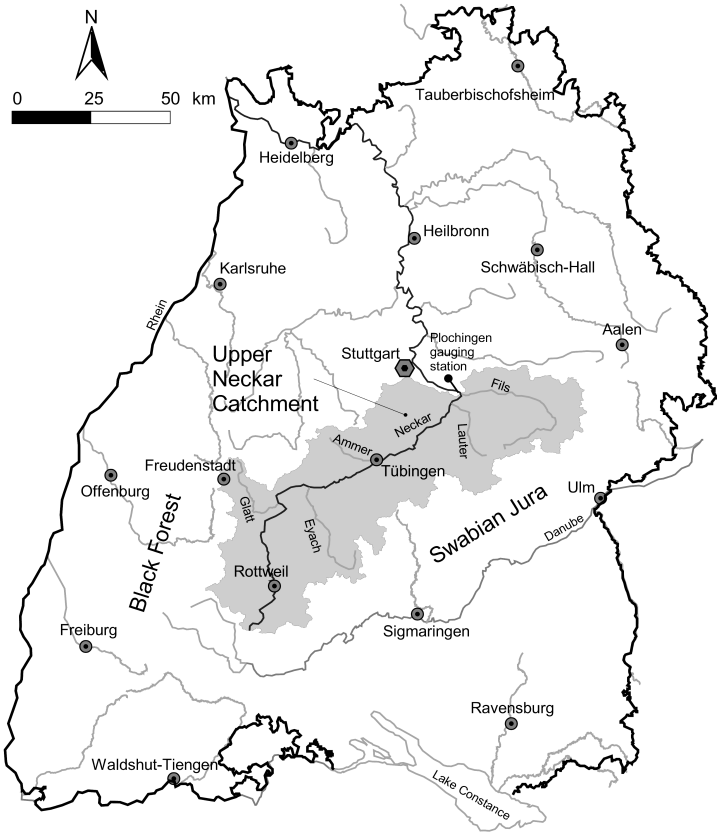

Figure 2.1 Map showing the location of the Upper Neckar Catchment within the State of Baden-Württemberg, Germany.

The climate of the Study Area can be classified as *Cf* according to Köppen's notation (1918), i.e. moist mid-latitude climates with mild winters. This climatic type is characterized by having warm-to-hot summers with generally mild winters, and it is wet all seasons. The coldest and hottest months in

the Study Area are January and July respectively. The daily mean air temperature in the former is about -0.8°C whereas in the latter is about 17°C according to the daily mean temperature readings available for the period 1961 to 1990 (DWD). Although the climate of the area is moderate, a maximum annual range of about 47.4°C has been observed in the past decades.

The annual variation of precipitation in the Study Area exhibits a multimodal distribution. Precipitation-events may arise the whole year round, the rainiest month being June and the driest one October, whose monthly means are 126 mm and 64 mm respectively (according to daily readings from 1961 to 1995, DWD). The mean annual precipitation observed during this period is 908 mm.

With regard to land use, the Study Area has endured rapid land use transitions from cropland or grassland to built-up area or industrial usages from 1960 to 1993. Among the principal driving forces behind these land use changes are the following:

1. The high level of affluence of the Stuttgart Region originated by the steady technological and industrial development during the last decades. According to recent information, the Greater Stuttgart Region is among the top ten richest regions of the European Union (GDP expressed in purchasing power standard, Eurostat 1996), and having a relatively low unemployment rate of 5.75% (SLA, 1996).

2. The excellent transportation infrastructure available in the region has contributed to improve the accessibility from anywhere in the countryside to all urban centres and vice versa. This fact together with the high income per capita of the region's inhabitants has modified the commuting behaviour of a large part of the population. This behavioural change is one of the reasons for the steady growth of car ownership in the region since 1974, at an average rate of 2.5% per year, whose absolute value rose up to 0.529 car/inh in 1997 (SLA) (see Figure 2.2).

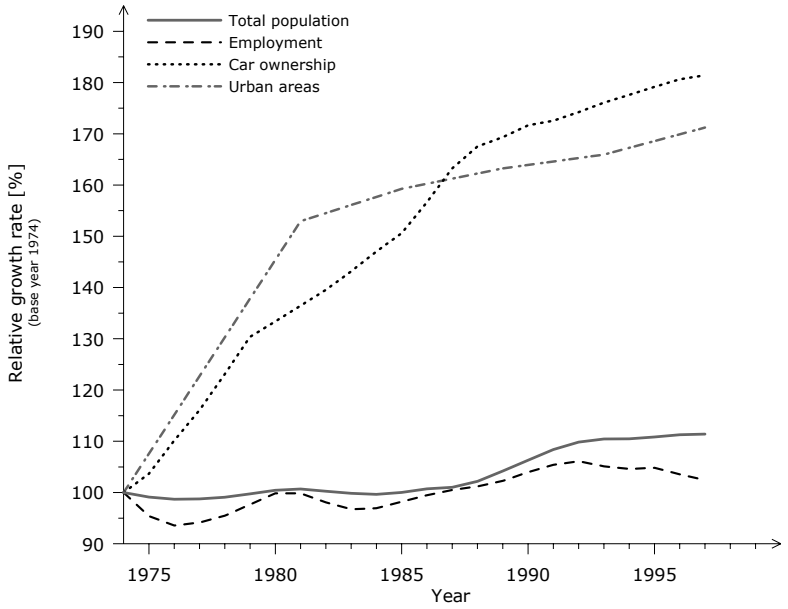

Figure 2.2 Relative growth ratio of population, car ownership, employment, and share of urban space (i.e. residential areas, commerce, manufacturing, and transportation) expressed in percentage. The base year for all indicators is 1974 (SLA).

As a direct consequence of that, a large part of the population would prefer to commute a long distance from home to work daily rather than a shorter one in order to have a dwelling with enough floor space located in a green, quiet, and peaceful environment where to live instead of a flat in a congested, noisy and crowded urban district. A further consequence of that is the rapid growth of floor space per capita in the past decades. This indicator increased in average from less than 15 m²/inh in 1950 to more than 38 m²/inh in 1997 (BBR 2000). In general, the demand for urban space, which includes transportation, commerce, and manufacturing, has grown 71% from its value in 1974, whilst the population has increased merely 11% in the same period.

3. Another driving force behind this phenomenon is in a lesser scale the population growth, especially its immigration component (from 1987 onwards, see Figure 2.2). As mentioned above, Stuttgart City as well as the counties located within the Study Area (i.e. Rems-Murr, Göppingen, Esslingen, Reutingen, Böblingen, Tüblingen, Zollernalbkreis, Freudenstadt, Rottweil, Schwarzwald-Baar-Kreis, and Tuttlingen) are ranked among those with the highest gross income per capita in Germany, namely, more than € 25 000/inh. in 1997 (Bundesanstalt für Arbeit in BBR 2000). This fact plus the very good living conditions in the region can be considered as proximate sources explaining the relatively high attractiveness of this region to immigrants.

The aforementioned land use changes of the Study Area have triggered a rapid land cover change as can be seen in Figure 2.3.

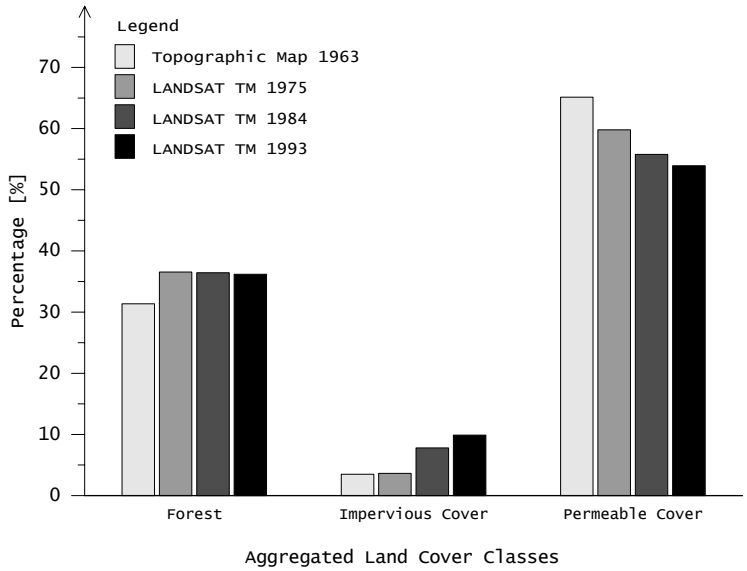

Figure 2.3 Land cover changes observed in the Upper Neckar Basin from 1960 to 1993. (Sources: for 1960: topographic maps of the area at scale 1 : 25 000 from LVA; for 1975, 1984 and 1993: LANDSAT TM scenes).

2.3 Conceptualisation of the Runoff Process at the Mesoscale Level

The hydrologic cycle within a drainage basin has been defined as a "*sequential, dynamic system in which water is the major throughput"* (Chow, 1964). This system is dynamic because it comprises several intertwined spatial phenomena, or processes, that are changing constantly over time. It is

sequential because there are inputs, an output, and a working fluid (i.e. water) called throughput passing through the system.

Based on the previous definition, and since one of the objectives of this study is to determine statistically significant variables representing the main processes involved in the system at mesoscale level, as well as their evolution over time; the available information has been divided, for analytical purposes, into two major categories, namely:

1. Inputs or explanatory variables, and
2. Output or explained variable.

In general, the complexity of the water cycle is due not only to the intertwined linkages of the physical, biological, and geological processes involved, but also due to the different temporal and spatial scales at which these processes interact. In addition to that, the water cycle can be disturbed from its normal behaviour by exogenous processes such as land cover changes, which are mainly caused by anthropogenic activities taking place within a basin. The spatial domain at which the water cycle spans varies from 10^{-8} m (at a molecular scale) to 10^{7} m (at a planetary scale). Nevertheless, the present study will only analyse the system at mesoscale catchments, i.e. those whose length ranges from 10^{2} to less than 10^{5} m, or whose area is greater than a few hectare but less than 5000 km^2 (Dooge, 1988).

Because of these reasons and others related with posterior analyses, the available input data has been further subdivided into three main subcategories, namely:

- **Physiographical factors**. This category comprises all those variables that can be regarded as constant or quasi-static since the period needed to appreciate a significant change has an order of magnitude greater than 10^{5} years. These factors comprise basin and channel characteristics mentioned by Chow, 1964. Among them are the following: 1) geological formations that constitute the basin's underground; 2) the basin's soil layers and their specific soils types; and 3) geometric factors of the drainage basin such as slope, aspect, shape, size, elevation, and drainage density. The symbol G symbolizes this subcategory.
- **Land cover types**. These variables have in general a slow changing rate over time (excluding some local exceptions, land use and land cover seldom change more than 5% per year, Robinson et al. in Meyer and Turner, 1998). The order of magnitude of a time interval necessary to perceive a significant change in their values varies from place to place, but in general, it would be ranging from 10^{0} to 10^{1} years. These variables stand for the observable consequences of anthropogenic activities happening within a basin. Land cover variables and their associated land uses are denoted by the symbol U.
- **Climatic or meteorological factors**. These variables are characterized by huge changes in their order of magnitude in very short periods of time. The period in which a significant change can be expected ranges from 10^{1} to even less than 10^{-4} years (Kleeberg and Cemus, 1992). In general, their behaviour exhibits some periodicity combined with partly chaotic and stochastic processes. This category comprises the following variables: precipitation, evaporation, solar radiation,

temperature, circulation patterns (closely related with relative humidity and wind velocity, among others). These variables are represented by the symbol M .

In the Figure 2.4, the main subcategories in which the input data has been subdivided as well as a simplified representation of the system's evolution along the time axis has been depicted. In this schematic diagram, the output of the system, or simply runoff, has been denoted by the symbol Q .

2.4 Spatial Units

Any integrated water management plan or related studies should be accomplished within a spatial unit called watershed (Singh, 1995). A watershed (originally from the German word *Wasserscheide*) is an "area with natural hydrological boundaries draining to a particular watercourse or water body". This spatial unit comprises the surface and ground waters, soils, fauna, and flora in the drainage basin, as well as humans and their anthropogenic activities (Reimond, 1998). In this study, 46 watersheds of various sizes are analysed, each one corresponds to the drainage basin at a geographic point i where a runoff gauging station is located (see Figures 2.4 and 2.5). These locations define the shape and size of a spatial unit. These points have been chosen because only there the outflow of a given watershed is measured continuously by State Agencies (e.g. LfU, DWD).

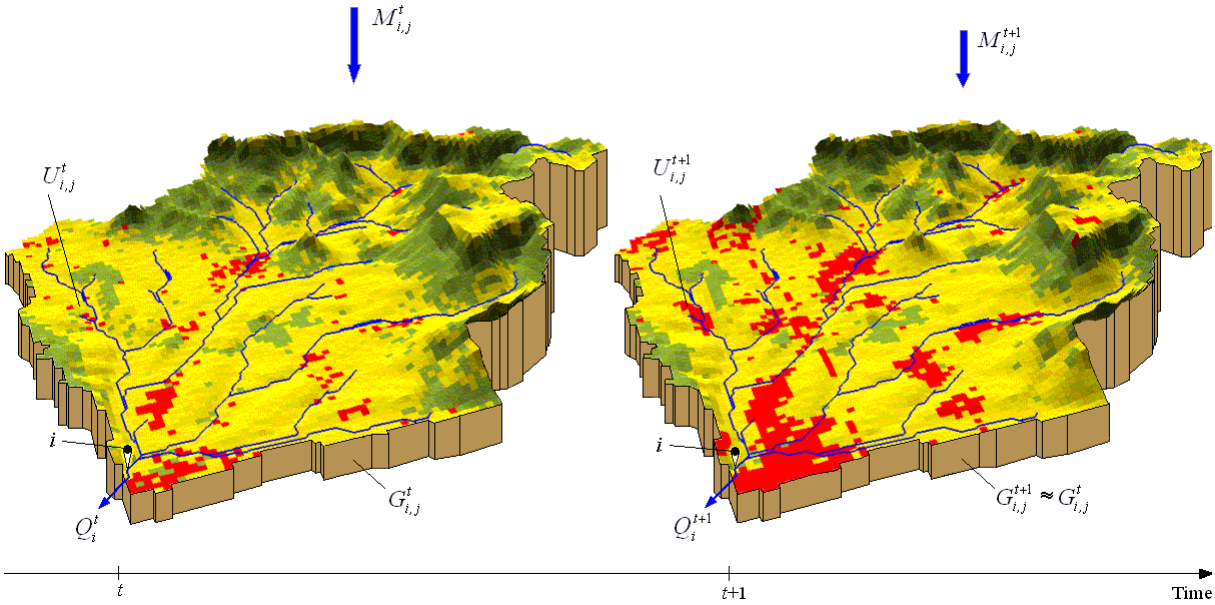

Figure 2.4 Schematic representation of the evolution of the system within a spatial unit i along the time axis t . The main inputs into the system are represented by the letters M (climatic factors), G (physiographical factors) and U (anthropogenic influences in the form of the land use/cover). The output is denoted by the letter Q (e.g. the total specific annual discharge in [mm]).

Figure 2.5 depicts the location of the stations within the Study Area and the delineation of corresponding subcatchments. Since the impacts of land cover change are cumulative (Dickert and Tuttle 1985), the delineation of the spatial units should be consistent with this phenomenon. Thus, a watershed draining to a point i should cover the watersheds of all those stations located upstream from it (e.g. the watershed Nº 4 comprises the sub-catchments Nº 4 and Nº 30, see Figure 2.5). In the

Appendix 1, a correspondence table shows how the sub-catchments depicted in Figure 2.5 have been aggregated to shape each spatial unit.

The delineation of a watershed is based on the algorithm proposed by Jenson and Domingue (1988). This algorithm requires as input data a DEM free of "sink holes". This means that the DEM has to be corrected in advance to ensure that a given drainage basin must have a single pour point, which normally corresponds to that one with the lowest elevation. Using this "corrected" DEM the delineation algorithm calculates the flow direction and the flow accumulation for each cell of the DEM. Based on this additional information, a watershed of any given point within the DEM can be obtained. The boundaries for the 46 watersheds were estimated by means of this procedure.

In order to validate the polygons that enclose the 46 watersheds, their areas were estimated and then compared with the official drainage area for the corresponding gauging station (hydrological yearbooks LfU). The results are very satisfactory as can be seen in Figure 2.6.

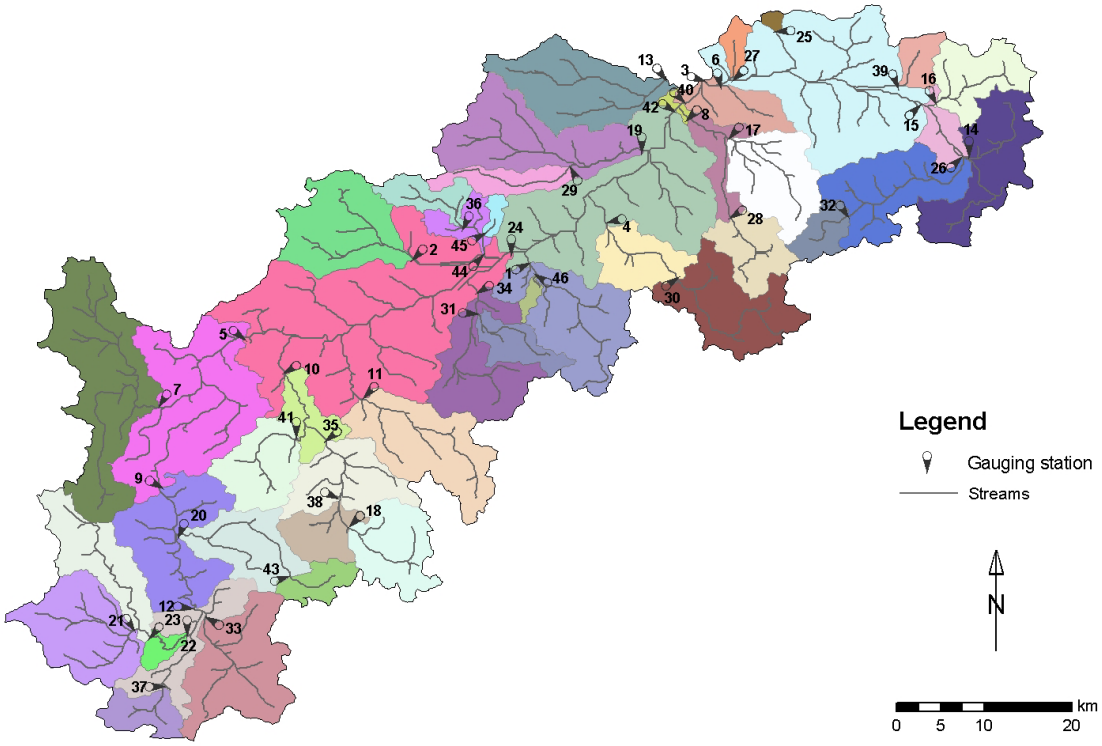

Figure 2.5 Subcatchments and stream network delineated from a corrected DEM (30×30 m, LfU).

2.5 Physiographical Factors

The physiographical factors considered in this study have been derived from different sources, mainly: 1) DEM available for the Study Area with a spatial resolution of 30×30 m (from LfU); 2) a digitized soil map of the State of Baden-Württemberg at the scale 1 : 200 000 (LfU-IER-ILPÖ); 3) a digitized geological map of the State of Baden-Württemberg at the scale 1 : 600 000 (LfU-IER-ILPÖ). A detailed explanation of the meaning, relevance, and calculation of these factors is given below.

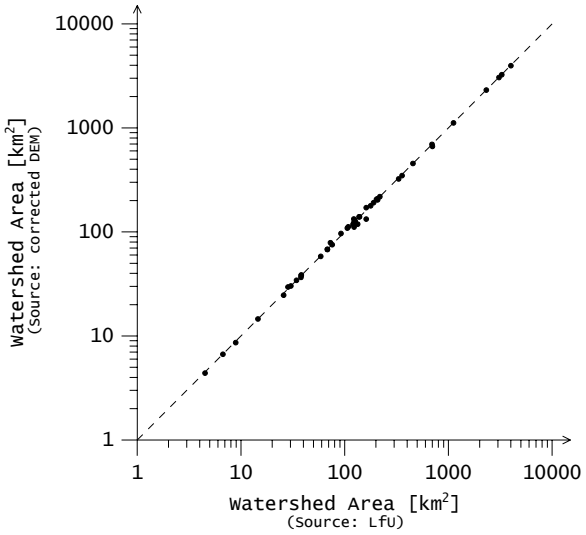

Figure 2.6 Comparison between the area of a given watershed estimated from a corrected DEM and its official drainage area (from LfU).

2.5.1 Basin's Area [km²]

According to the several studies conducted by Leopold and Miller (1956) and Hack (1957) among others, the area of a basin is positively correlated with its average discharge and inversely correlated with peak discharges (Dalrymple in Chow 1964). Because of this empiric evidence, the basin's area has been considered as a probable explanatory variable. Its estimation is as follows.

Let x_{i1} be the area of the spatial unit Ω_i in [km²]. A spatial unit Ω_i may be composed of one or several subcatchments as indicated in Appendix 1, hence its area is calculated as follows:

$$x_{i1} = A_i = \sum_{k \in \Omega_i} a_k \tag{2.1}$$

where

a_k is the area of a cell k (i.e. $30 \times 30\,\mathrm{m}^2 = 0.9 \times 10^{-3}\,\mathrm{km}^2$) that is fully contained within the spatial unit Ω_i.

2.5.2 Mean Slope [°]

Let x_{i2} be the arithmetic mean of the slope (i.e. gradient) in degrees of all cells k located within a spatial unit Ω_i. In general, the slope at a given point k is defined as the plane parallel to the topographic surface that is represented by an array of elevation points contained in the DEM ($\mathbf{Z}$). This plane can be mathematically represented as a vector $\vec{\mathbf{S}}_k$ that indicates the maximum rate of change in altitude. As a vector, it has two components, namely: 1) its magnitude called gradient or simply "slope", which is equal to the first derivative of the elevation z_k with respect to $\vec{\mathbf{S}}_k$; and, 2) its direction with respect to the North, or simply aspect ($\boldsymbol{\psi}$) (Burrough 1986).

In order to calculate the slope of a cell k, a plane is fitted by least squares to the elevation of its eight neighbours and itself. Then the gradient is estimated by the average maximum algorithm proposed by Burrough (1986). Finally, the mean slope within a spatial unit is

$$x_{i2} = \frac{1}{N_i} \sum_{k \in \Omega_i} \left\| \vec{\mathbf{S}}_k \right\|, \tag{2.2}$$

where

N_i is the number of cells within the spatial unit Ω_i

$$\left\| \vec{\mathbf{S}}_k \right\| = s_k = \tan^{-1}\left(\sqrt{\left(\frac{dz_k}{dx}\right)^2 + \left(\frac{dz_k}{dy}\right)^2} \right) \frac{180}{\pi} \ [°], \tag{2.3}$$

z_k is the elevation of cell k contained in the DEM [m], and

x, y the Cartesian directions of the reference system of the DEM [m].

The figure 2.7 shows the slope map s_k derived for the Study Area.

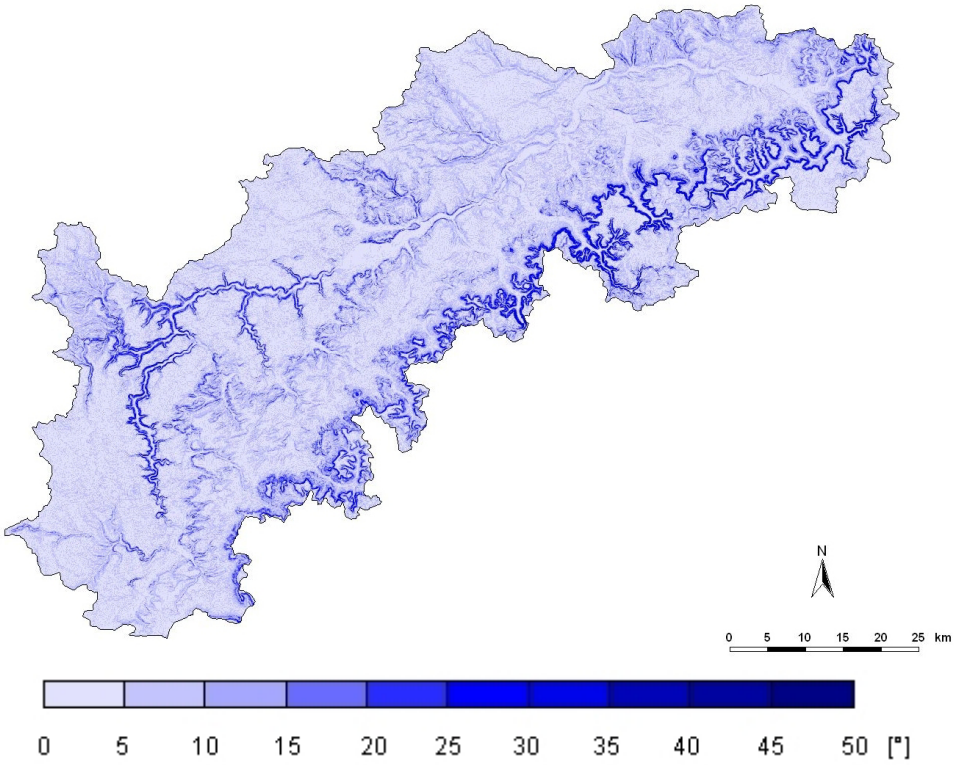

Figure 2.7 Slope map in [°] for the Study Area derived from a DEM (LfU).

2.5.3 Median of the Slope [°]

Non-linear problems generally have variables whose empiric probability density functions, or (PDFs), exhibit multimodal and skewed distributions. In such cases *the sample average is not a resistant*[3] *characterization of the centre of the data set* (Wilks, 1995). Instead of the arithmetic mean, other resistant measures of location such as the median or the trimmed means can be used. In order to estimate the median of the slope (s) in a spatial unit Ω_i the following definitions are needed. Let be $\{s_1, s_2, s_3, \ldots, s_k, \ldots, s_{N_i}\}_i$ the set of data containing the slope for each cell within a spatial unit Ω_i.

[3] A resistant method "*is not unduly influenced by small number of outliers or 'will data'* " (Wilks, 1995).

The same data set, but sorted in ascending order, can be denoted using parenthetical sub indices as $\{s_{(1)}, s_{(2)}, s_{(3)}, \ldots, s_{(k)}, \ldots, s_{(N_i)}\}_i$; where $s_{(1)}$ and $s_{(N_i)}$ represent the minimum and maximum slope values respectively. Based on this ranked set, the median can be defined as follows

$$x_{i3} = \begin{cases} s_{([N_i+1]/2)} & \text{if} \quad N_i \quad \text{is} \quad \text{odd} \\ \dfrac{s_{(N_i/2)} + s_{([N_i/2+1])}}{2} & \text{if} \quad N_i \quad \text{is} \quad \text{even} \end{cases} \tag{2.4}$$

2.5.4 Trimmed Mean of the Slope [°]

As mentioned before, another measure to obtain a robust and resistant measure of location with low sensitivity to outliers and extreme values is the trimmed mean. In the present study this indicator may play a significant role describing the water budget of a given basin because extreme slope values may come from both very steep slopes on mountainous regions and very flat areas such as water bodies. Both types of locations are not likely to have had a land cover change in the past, and hence their influence in the overall water budget of the given basin may be insignificant. In general, this variable can be calculated by

$$x_{i4} = \frac{1}{N_i - 2\rho} \sum_{k\,=\rho+1}^{N_i-\rho} s_{(k)} \qquad k \in \Omega_i, \tag{2.5}$$

where

ρ an integer equal to the rounded value of the product σN_i, and

σ the proportion of observations excluded from each tail of the data set $\{s_{(1)}, s_{(2)}, s_{(3)}, \ldots, s_{(k)}, \ldots, s_{(N_i)}\}_i$. In this case $\sigma = 0.15$, i.e. the 30 % extreme observations are excluded as a whole.

Another trimmed mean, i.e. x_{i5}, can be calculated using (2.5) but with $\sigma = 0.3$.

2.5.5 Mean Stream Slope [°]

Let x_{i6} be the arithmetic mean of the slope of all cells k belonging to the stream network $\mathcal{N}_i$ located within the spatial unit Ω_i, thus

$$x_{i6} = \frac{1}{N_{\mathcal{N}_i}} \sum_{k \,\in\, \mathcal{N}_i} s_k \qquad \mathcal{N}_i \subset \Omega_i, \tag{2.6}$$

where

$N_{\mathcal{N}_i}$ is the number of cells contained within the spatial unit Ω_i that belong to the stream network $\mathcal{N}_i$.

The stream network has been delineated based on the corrected DEM mentioned above. The delineation algorithm requires two additional grids, namely: the flow direction and the flow accumulation grid. The flow direction, which has also been derived from the DEM, defines for each cell k of the DEM the direction(s) in which the surface water may flow freely over the land surface. Using this additional information it is possible to calculate how many cells (or areas) are pouring into

a given cell k. This procedure is repeated for all cells within the Study Area and the result stored in the so-called flow accumulation grid Υ (Jenson and Domingue 1988). In order to have a stream with permanent flow of water over the year a minimum number of contributing cells is required. Such threshold is represented here with τ (Tarboton et al. 1993). Based on these definitions the stream network $\mathcal{N}_i$ can be defined as follows

$$\mathcal{N}_i = \left\{ \ (k) \ \middle| \ k \in \Omega_i \ \wedge \ \Upsilon_k \geq \tau \ \right\}. \tag{2.7}$$

The threshold $\tau = 300\,[\text{cells}]$ has been used in the present study following the recommendations of Tarboton et al. (1993), which corresponds to a drainage area of 27 ha. The stream network derived under these conditions is shown in Figure 2.8.

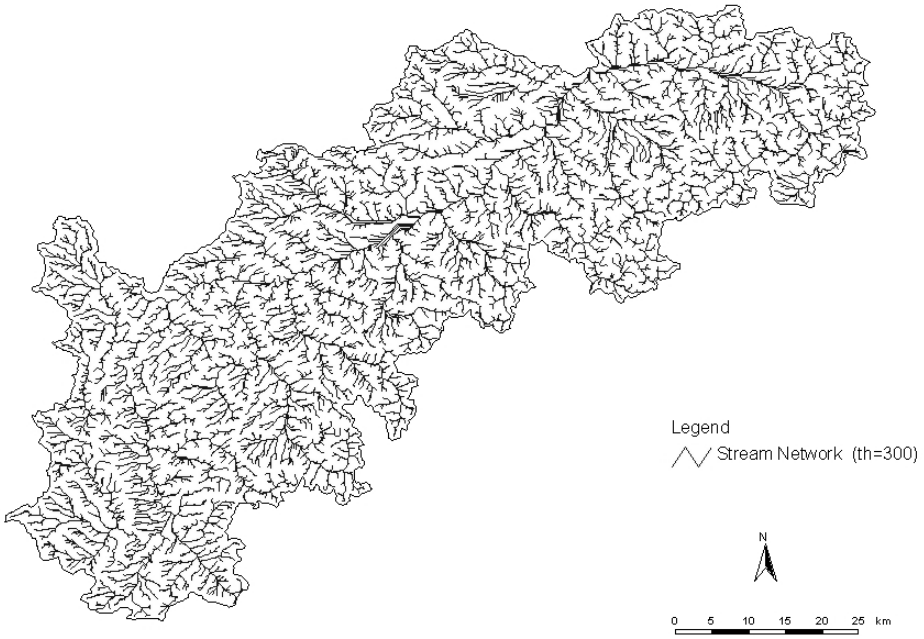

Figure 2.8 Drainage network of the Study Area derived from a DEM (30×30 m LfU) and a threshold value of 300 cells.

2.5.6 Mean Slope of the Areas Located at the Floodplains and Riparian Land of the Stream Network [°]

From previous research studies, it is known that floodplains and riparian land play a very important role in the hydrological cycle of a basin because they serve as potential reservoirs in case of extreme events; this is the reason why, apart from their high biodiversity, they are considered as very sensitive areas (Dalrymple in Chow 1964), and are often protected by law. Unfortunately, this is not always the case as the historical evidence points out (Meijerink and Mannaerts in Schultz 2000). In fact, many land cover and land use changes happened in these zones in the Study Area during the last 40 years as is depicted in Figure 2.9.

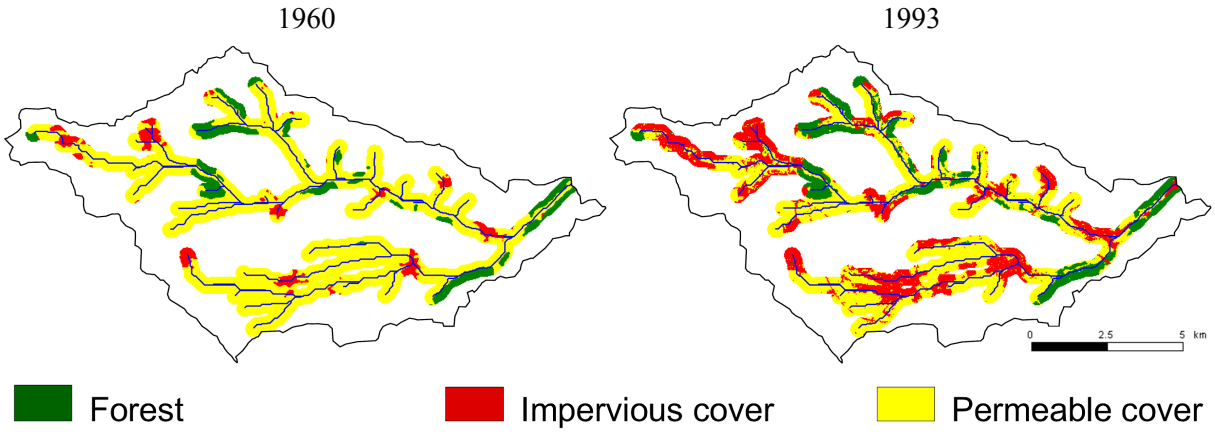

Figure 2.9 Sample of land cover and land use changes along floodplains and riparian zones along the main streams of river Körsch.

Based on these premises, it is reasonable to presuppose that if the land cover and the land use of those sensitive areas have changed in the past, the mean slope where these changes took place would have influenced the overall water budget of the basin; hence, they have been evaluated to estimate their expected influence.

Let x_{i7} be the arithmetic mean of the slope of all cells k located within the floodplains and riparian zones $\mathcal{B}_i$ belonging to the spatial unit Ω_i. Based on this definition this variable can be estimated as follows

$$x_{i7} = \frac{1}{N_{\mathcal{B}_i}} \sum_{k \in \mathcal{B}_i} s_k \qquad \mathcal{B}_i \subset \Omega_i \,, \tag{2.8}$$

where

$N_{\mathcal{B}_i}$ is the number of cells contained within the spatial unit Ω_i that belong to the riparian zones and floodplains $\mathcal{B}_i$ of the stream network $\mathcal{N}_i$.

The floodplains and riparian zones for each stream located within a spatial unit Ω_i have been defined using a simple concept, i.e. a cell k belongs to $\mathcal{B}_i$ only if its Euclidean distance from it to the closest cell belonging to the stream network $\mathcal{N}_i$ is less than or equal to the threshold δ [m], whose value in the present case is $\delta = 150$ m. The latter definition can be written formally as follows

$$\mathcal{B}_i = \left\{ (k) \,\middle|\, k \in \Omega_i \quad k^* \in \mathcal{N}_i \quad \wedge \quad \left\| \mathbf{r}_k - \mathbf{r}_{k^*} \right\| \leq \delta \right\}, \tag{2.9}$$

where

$\mathbf{r}_k$ denotes the position vector of the cell k in geographic coordinates, and

$\mathbf{r}_{k^*}$ represents the position vector of the cell k^* whose distance to cell k is smaller than or equal to δ.

2.5.7 Drainage Density [km^{-1}]

There is empiric evidence showing that the drainage density of a given basin x_{i8} correlates with many variables governing its water budget. This variable, originally introduced by Horton (1945) and complemented later by Melton (1958), is a function of climate, vegetation, rock and soil types, i.e. the

composition of the geological strata, rainfall intensity, infiltration capacity and relief are related with this variable. Strahler (in Chow 1964) also proposed that the drainage basin is a function of the runoff intensity (the volume rate of flow per unit of area of cross section), an erosion-proportionality factor, relief -which represents the potential energy of the system-, density of the fluid, and gravity. Carlston (1963, 1966) found that the drainage density has a negative correlation with the basin's base flow and a positive correlation with the mean annual flood.

The drainage density should consider all streams with permanent flow located within a given basin (Chow, 1964), as has been shown in Figure 2.8. This variable is estimated by

$$x_{i8} = \frac{\sum_{o}\sum_{j} \ell^{i}_{oj}}{A_i}, \tag{2.10}$$

where ℓ^{i}_{oj} is the length of the stream segment j belonging to order o of the stream network $\mathcal{N}_i$.

2.5.8 Shape Factor [-]

This morphological indicator may influence stream discharge characteristics. According to Chow (1964) "long narrow basins with high bifurcation ratios would be expected to have attenuated flood-discharge periods, whereas rotund basins of low bifurcation ratios would be expected to have sharply peaked flood discharges". Due to this reason it is considered as a potential explanatory variable.

The shape factor of a given basin i, x_{i9}, is defined as the ratio between the catchment length squared and its area. The length of a basin (L_i) is by definition the distance from the basin's pour point to the furthest point located on its perimeter (Chow, 1964). Hence it is calculated as follows

$$x_{i9} = \frac{L_i^2}{A_i}. \tag{2.11}$$

2.5.9 Fraction of North- and South-facing Slopes [-]

The proportion of south-facing slopes in the Northern Hemisphere is linked with the daily amount of received solar radiation, and this in turn is linked with the potential evapotranspiration (PET), and air temperature of those places (Hamon 1961, Frank and Lee 1966, Swift 1976, Leavesley 1983 and 1995). Due to these reasons, the fraction of both north- and south-facing slopes should be considered as a potential explanatory variable at mesoscale.

In order to estimate the percentage of north- and south-facing slopes, the direction of each vector $\vec{\mathbf{S}}_k$ with respect to the north axis has been stored in an array called the aspect grid $\boldsymbol{\psi}$, which is shown in Figure 2.10.

The aspect is given by

$$\psi_k = \tan^{-1}\left(-\frac{dz_k}{dy}\Big/\frac{dz_k}{dx}\right)\frac{180}{\pi}[^\circ]. \tag{2.12}$$

The reclassification of all cells ($\boldsymbol{\psi}^*$) into north- or south-facing slopes has been done as follows

$$\psi_k^* = \begin{cases} 1 & \text{if} \quad 0° \le \psi_k \le 45° \ \vee \ 315° \le \psi_k \le 360° \\ 2 & \text{if} \quad 135° \le \psi_k \le 225° \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

where ψ_k^* takes the value of 1 for slopes mainly toward the north, 2 for slopes mainly toward the south, and 0 otherwise. Then, the proportion of north-facing slopes x_{i10} within a spatial unit Ω_i have been calculated by

$$x_{i10} = \frac{\left|\left\{(k) \middle| \ k \in \Omega_i \ \wedge \ \psi_k^* = 1\right\}\right|}{N_i}. \tag{2.14}$$

In (2.14) the symbol $||$ denotes the cardinality of a given set, i.e. its total number of elements. Furthermore, the proportion of south-facing slopes is

$$x_{i11} = \frac{\left|\left\{(k) \middle| \ k \in \Omega_i \ \wedge \ \psi_k^* = 2\right\}\right|}{N_i}. \tag{2.15}$$

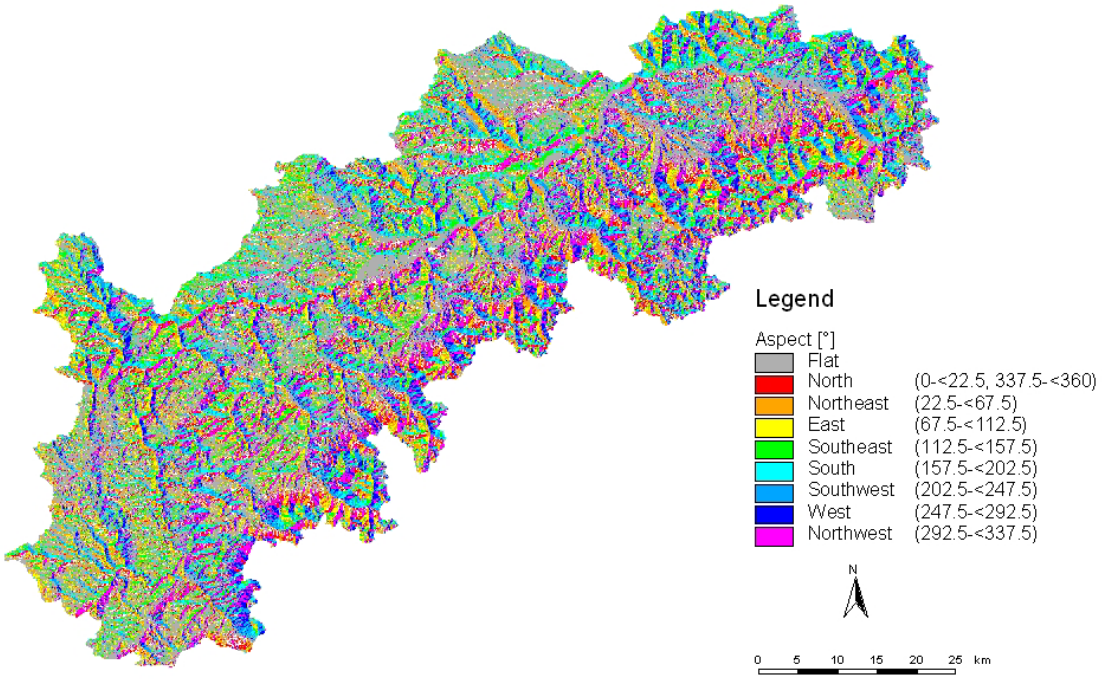

Figure 2.10 Aspects map in [°] for the Study Area derived from a DEM (30×30 m LfU).

2.5.10 Mean Basin Elevation and Difference between Maximum and Minimum Altitudes [m]

Both the mean elevation and the difference between maximum and minimum altitudes of a basin Ω_i are variables related with the potential energy of the system (Haggnett and Chorley 1969), as well as with the maximum air temperature and PET of the basin (Leavesley 1983). A number of validated models, for instance Hewlett and Hibbert (1967), Freeze (1974), Anderson and Burt (1978), and Beven and Kirkby (1979), have recognized the role of topography in determining areas of downslope movement of moisture, saturation-excess flow, and hence its influence in controlling throughflow

generation in a basin's hillslopes. Because of these reasons, the following indicators derived from a DEM shown in Figure 2.11 are considered potentially important explanatory variables.

Mean elevation has been estimated by

$$x_{i12} = \frac{1}{N_i} \sum_{k \in \Omega_i} z_k \,, \tag{2.16}$$

and the difference between maximum and minimum altitudes as

$$x_{i13} = \max(z_k) - \min(z_k) \quad k \in \Omega_i \,. \tag{2.17}$$

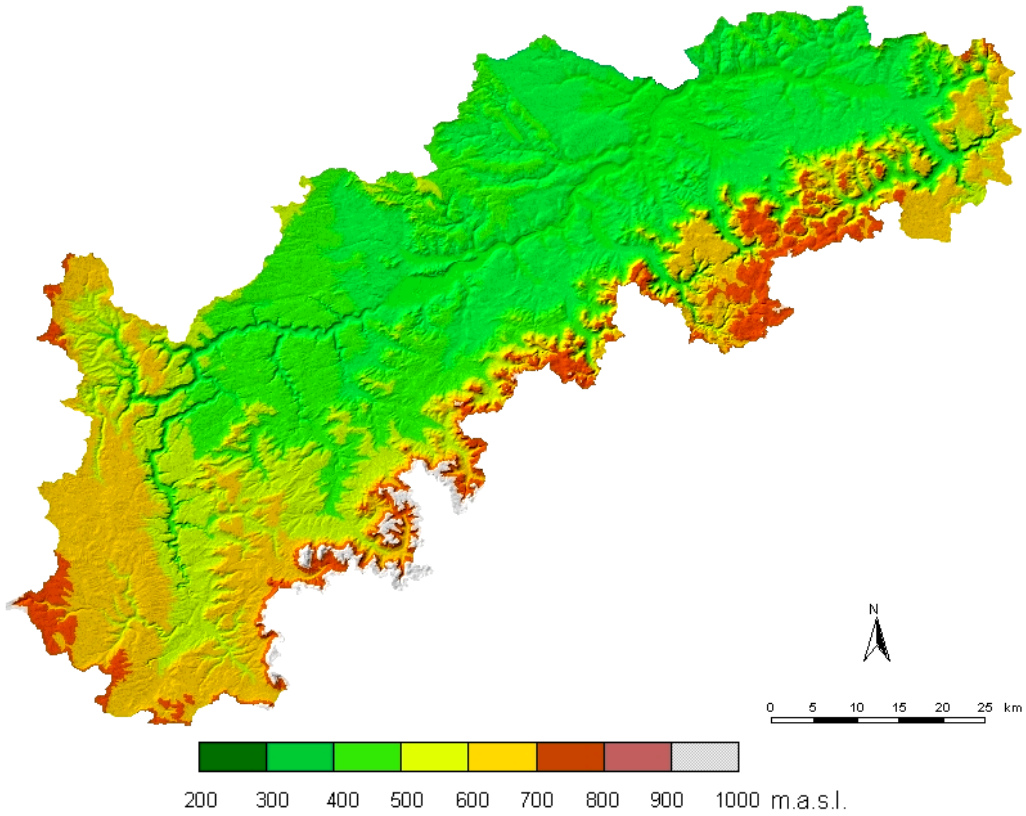

Figure 2.11 Topography of the Study Area represented by a DEM (LfU).

2.5.11 Fraction of Saturated Areas [-]

Many hydrological models (e.g. TOPMODEL or TOPLOG) assume that a wetness index based only on topography can be used to estimate the depth of the ground water table[4], which in turn, is a variable closely related with the runoff generation (Beven and Kirkby 1979, O'Loughlin 1981, Moore et al. 1991, Wolock 1993a). The most common and simple indicator, which is called the topographic index I_k, can be computed in the present case by

$$I_k = \ln\left(\Upsilon_k \frac{a_k}{c_k} \frac{1}{\tan s_k}\right). \tag{2.18}$$

[4] A water table by definition is the top of an unconfined aquifer, which correspond to a surface at atmospheric pressure defined by the level to which water will rise in a well (Chow, 1964).

Where a_k and c_k are the area and the size of the cell k (i.e. 30 m). In (2.18), the term $\Upsilon_k a_k / c_k$ represents the total upstream drainage area of the cell k per unit of contour, and $\tan s_k$, denotes the hydraulic gradient for saturated water flow (slope is expressed in radians). Υ_k is the flow accumulation grid as in Section 2.5.5. As a result of these calculations Figure 2.12, depicting a subcatchment of the Study Area, has been obtained.

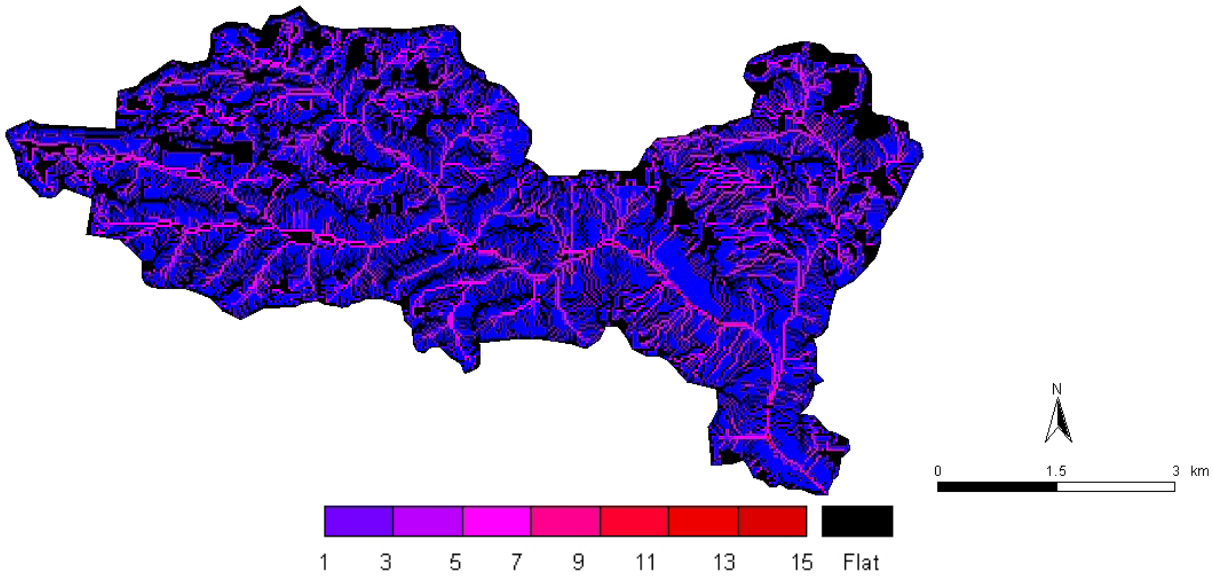

Figure 2.12 Topographic index derived according to the method proposed by Beven and Kirkby (1979) for the subcatchment No. 36 located in the Study Area.

Those areas of the catchment having higher values of the topographic index are prone to become saturated during a precipitation event, and hence, to contributing most to surface runoff generation (Beven and Kirkby 1979). Even though the topographic index has many shortcomings - for instance: 1) the amount of runoff generation depends not only on topographic characteristics, but also on the initial soil moisture and the soil type (Moore et al. 1991), and 2) the speed of the surface runoff would depend on the surface's roughness determined largely by the land cover - it would be interesting to investigate whether a relationship between the basin's discharge and the fraction of the basin's area prone to become saturated exists. Moreover, testing whether the land cover changes occurred on those areas with higher values of the topographic index would have influenced the discharge characteristics of the basin.

In order to compute the fraction of potential saturated areas (x_{i14}) within a given basin i only those cells having a topographic index greater than or equal to the upper quartile ($\mathrm{Q}_3 = \mathrm{P}_{75}$) of the empirical distribution function of the topographic index of a given spatial unit i would be taken into account. Hence, it is calculated by

$$x_{i14} = \frac{\sum_{k=\varsigma+1}^{N_i} I_{(k)}}{A_i} \qquad k \in \Omega_i, \tag{2.19}$$

where ς is equal to the rounding of $0.75 N_i$.

2.5.12 Mean Field Capacity [mm]

As mentioned in Section 1.2, a very important part of the water cycle takes place in the lithosphere. Hence, it would be reasonable to find out an observable that should be able to represent, at least partially, the important role of the soil matrix as a reservoir, and, at the same time, as a porous media where the subsurface flow would take place under the interaction of capillary and gravitational forces. Among many soil characteristics, the most commonly used are: field capacity, wilting point, and porosity.

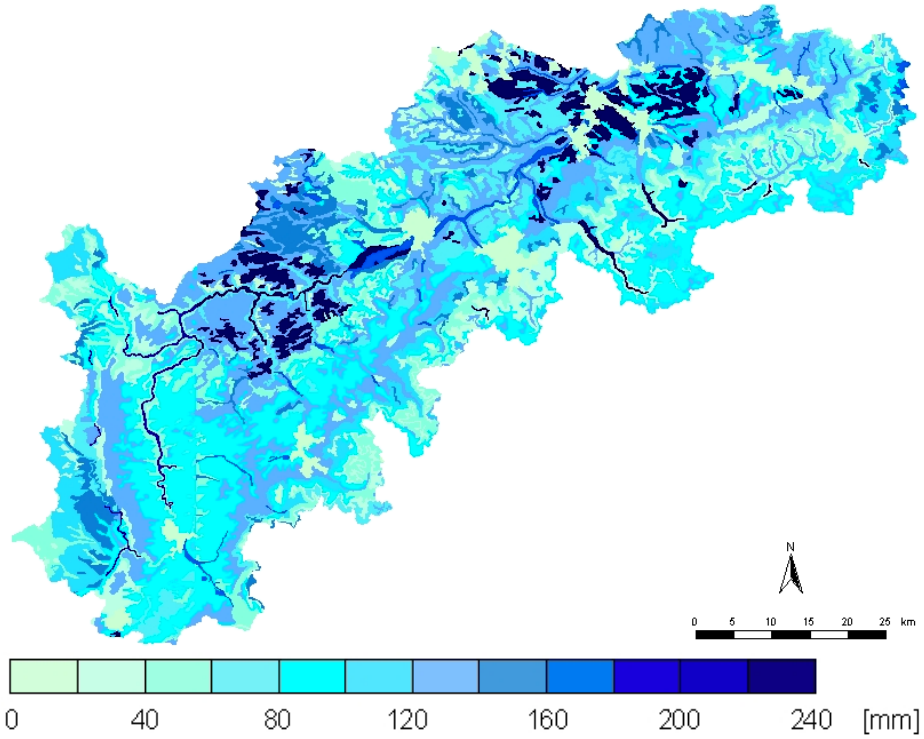

Figure 2.13 Main soil types and their associated field capacity in [mm] in the Study Area.

By definition, field capacity "is the amount of water held in a draining soil after gravity movement of water has largely ceased" (Ward and Robinson, 2000). Besides, it is intrinsically a soil type dependent characteristic. As a result of this soil water "constant", it can be stated, for instance, that catchments having mainly sandy soils would drain more rapidly than those having mostly clayish soils. Hence, the behaviours of such basins as to runoff generation would be different. Due to this reason, the average field capacity of the basin is considered as a potential explanatory variable, which is calculated as follows

$$x_{i15} = \frac{1}{N_i} \sum_{k \in \Omega_i} C_k \,, \tag{2.20}$$

where C_k is the observed field capacity in [mm] for a given cell k. The data used to estimate this indicator is shown in Figure 2.13 (LfU).

2.5.13 Fraction of Karstic Formations [-]

The kind of geological formations –along with their faults and interstices- underneath a given basin have a special relevance with regard to its water balance because they act as groundwater reservoirs

(i.e. aquifers), as well as pipelines for groundwater flow. The physical characteristics of a geological formation such as specific retention, specific yield, porosity and permeability would depend upon a number of factors, but in general, the materials that constitute them may be a determinant one (Ward and Robinson, 2000). This is the case when the underground is composed of limestone, a sedimentary rock full of fissures, sinkholes and caverns (i.e. karts) originated from groundwater flow erosion. Figure 2.14 shows the main geological formations of the Study Area (LfU).

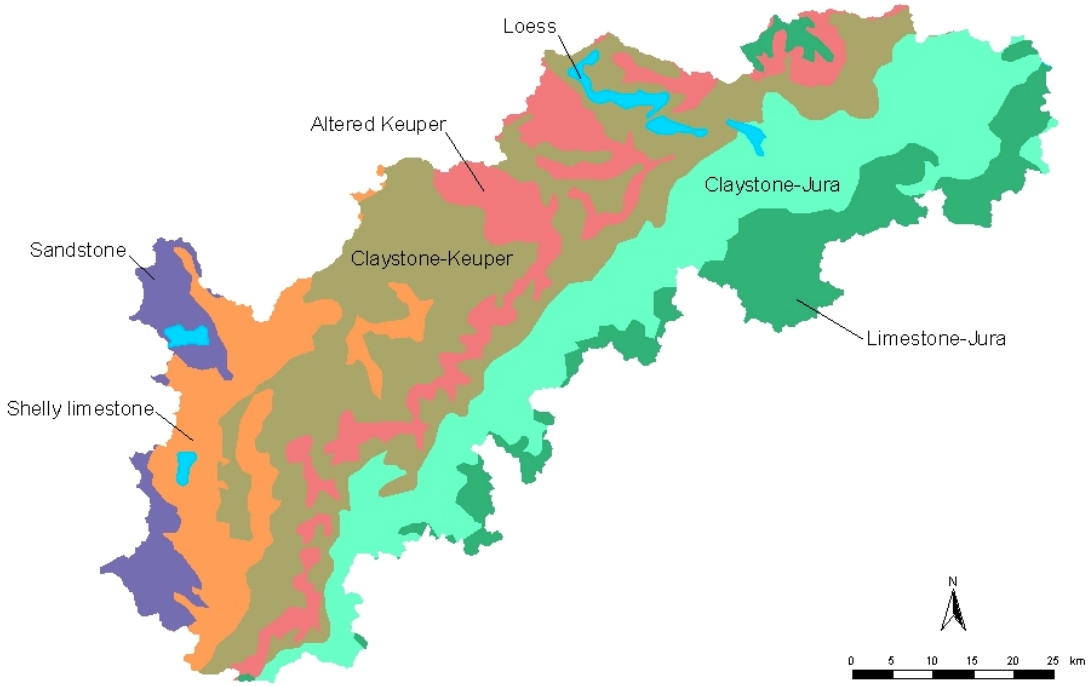

Figure 2.14 Main geological formations of the Study Area. The karstic formation corresponds to Limestone-Jura (LfU).

In general, a basin whose underground has large proportions of karstic formations would present huge abnormalities in its water budget, and thus in its discharge regime. Since such cases occur within the Study Area, the following indicator was developed to consider this fact, namely

$$x_{i16} = \frac{1}{N_i} \sum_{k \in \Omega_i} \mu_k \,, \tag{2.21}$$

where

$$\mu_k = \begin{cases} 1 & \text{if } k \in \text{karstic formation} \\ 0 & \text{otherwise} \end{cases} . \tag{2.22}$$

2.6 Land Use and Land Cover

From a hydrological point of view, land cover would be much more adequate than land use in the subsequent analyses since it is related with processes such as evapotranspiration, interception (canopy), and albedo; and with hydrologic parameters such as root zone depth and surface roughness that are governing processes such as the surface runoff and infiltration.

Due to these reasons, the land cover of the Study Area is to be estimated for successive points in time in the Study Area. Nevertheless, relationships between land use and land cover should be established in order to understand the system's behaviour under anthropogenic impacts.

In the present study, eight different land use categories disaggregated at a Municipality (*Gemeinde*) level have been obtained from the Statistical Office of Baden Württemberg (SLA) for the Study Area. They are called: forest, built-up and open areas, commercial use, industrial use, transportation, recreation areas, agricultural areas, water bodies and wetlands and 'other uses' (Flächenerhebung, SLA for the period 1981 to 1997). As a result of aggregating this information for the 216 municipalities comprised within the Upper Neckar Basin (i.e. the drainage area of the Plochingen gauging station; depicted with No. 3 in Figure 2.5) time series for each land use category have been obtained and depicted in Figure 2.15. These graphs clearly show the main land use changes that occurred within the Study Area in the last two decades, namely: agriculture has a steady negative average growth rate of about 0.48% per year whereas built-up, commerce, recreation, transportation, and water bodies and wetlands showed growth at average rates of 0.90%, 1.67%, 3.73%, 0.40% and 1.01% per year respectively. The category 'other-uses', on the contrary, abruptly declined with an average growth rate of about -2.32% per year. Moreover, forest, which had a period of fast growth from 1985 to 1993 (about 0.67% per year), reduced its pace to an insignificant 0.15% per year during the period 1993-1997.

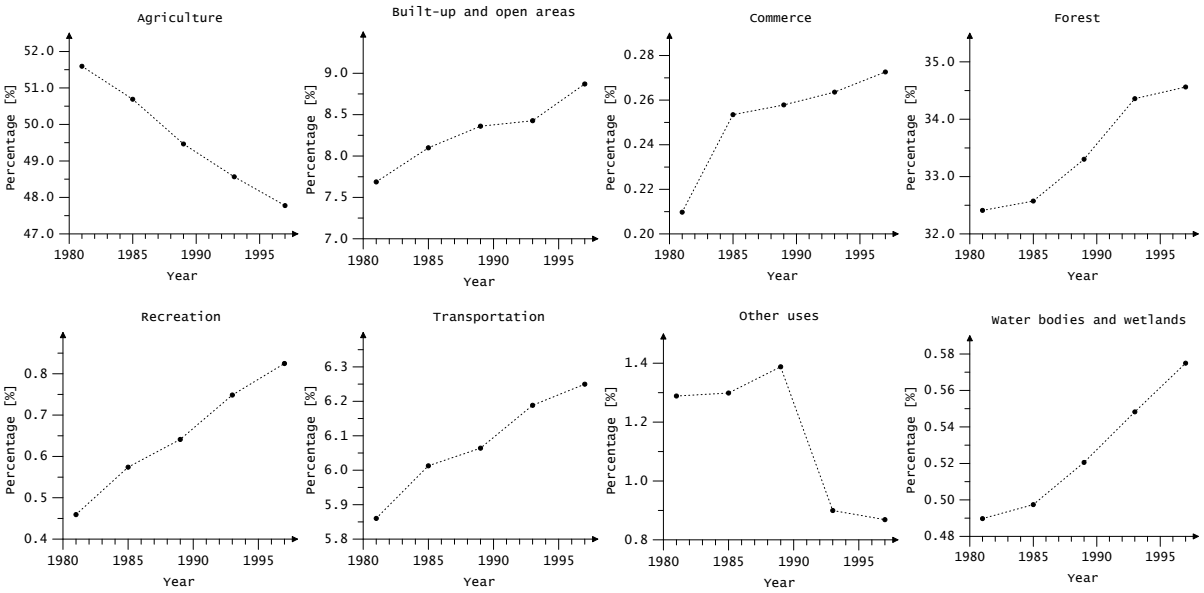

Figure 2.15 Land use time series in percentage for the Upper Neckar Basin from 1981 to 1997. These results were obtained by aggregating land use data for 213 Municipalities located within the catchment.

With regard to land cover categories, they were obtained from a classified LANDSAT TM scene for the State of Baden-Württemberg for the year of 1993 (IPF, 1995). This classification contains originally 64 sub-categories that were aggregated into 10 classes as is shown in Table 2.1 to ease the comparison between land cover and land use classes obtained from the mentioned data sources.

The land use and cover categories employed do not reveal a simple one to one relationship among them. This is due not only to the completely different gathering and surveying procedures, and spatial

resolutions, but also due to the different conventions of what a land use class represents. Therefore, to make these categories comparable, they had to be aggregated according to certain criterion. In the present case, according to the purpose of the study the following criterion was chosen: the aggregation procedure has to provide three land cover classes with remarkably different hydrologic responses, namely: forest, impervious areas and permeable areas (see Table 2.1).

Table 2.1 Correspondence of land use and land cover categories at the Municipal level.

Hydrologic Land Cover Category	**Land Use Survey (*Flächenerhebung*)**		**Classified LANDSAT TM 1993**	
Forest	F	Forest	DF CF MF	Deciduous forest Conifer forest Mixed forest
Impervious cover	BU C I T	Built-up area Commercial use Industrial use Transportation	DS SS I	Dense settlement Scattered settlement Industrial use
Permeable cover	R A O	Recreation areas Agricultural areas Others	G C V O	Grassland Cropland Vineyards Orchards

It was found that such aggregation fits very well at the municipal level as corroborated by Figure 2.16, especially with regard to forest and permeable cover. Quite the opposite, the area of impervious cover estimated by the land use survey often exceeds those values estimated by the LANDSAT. A plausible explanation of such a discrepancy comes from the fact that the land use survey data is based on the cadastral information, which has a very high spatial resolution whereas the estimates from the LANDSAT have a ground resolution of 30 m (only).

The definition of each land cover class adopted in this study is as follows:

Forest consists of areas covered by coniferous, deciduous, and mixed forest.

Impervious cover consists of areas covered by high and low density settlements, as well as industrial areas, airport runways, highways, and railway tracks. All of these categories have artificial drainage systems and their surface is totally or partially covered by asphalt, concrete, or any other sealing material.

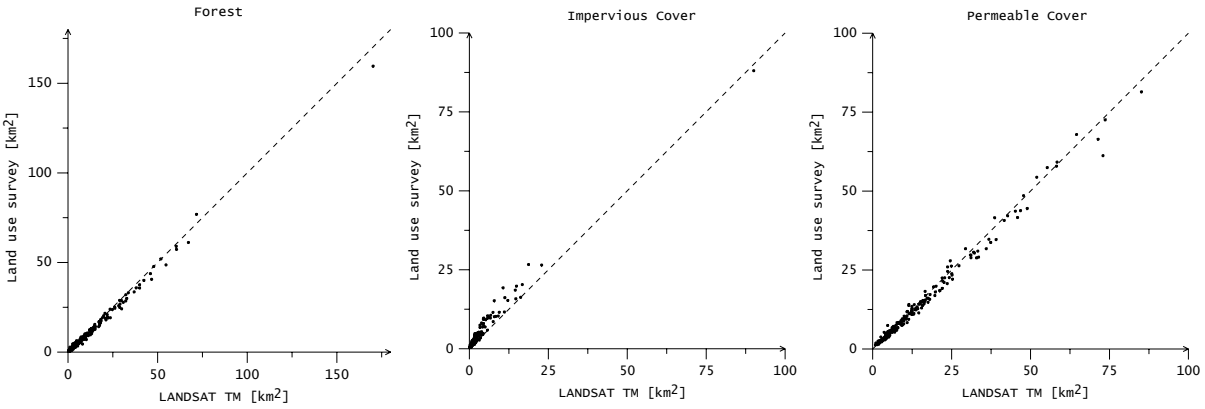

Figure 2.16 Comparison between the area of all municipalities within the Study Area for three different land cover types estimated from a classified LANDSAT TM image acquired in 1993 and data of the Land Use Survey in the same year (Flächenerhebung, SLA Baden-Württemberg 1993).

Permeable cover consists of areas covered by crops, grass, orchards, or a mixture of them, as well as wetlands and fallow lands. This category also includes the category water bodies (areas covered by rivers, lakes and reservoirs) which are negligible in the Study Area.

The land cover changes endured by the Study Area were estimated by means of a time series of land cover scenes consisting of one rasterised land cover map for 1960 and three LANDSAT scenes acquired in 1975, 1984 and 1993 respectively.

2.6.1 The Land Cover Time Series

The land cover in 1961

For the decade from 1961 to 1970 there is only one source of information that allows estimating the land cover in the Study Area, namely the topographic maps at scale 1 : 25 000 surveyed by the LVA during 1961 to 1963. These maps depict many kinds of vegetation and impervious covers that were aggregated and digitised in accordance with the definitions stated above. Forest corresponds in this case to three vegetation categories depicted in the topographic maps as deciduous forest, coniferous and mixed forest. Permeable cover consists of orchards, parks, vineyards, tree nursery, meadows, pasture, wetlands, moorland, quarries, and swampy areas. The rest, impervious cover, corresponds to what is categorised as settlements and transportation corridors. The resulted rasterised map is shown in Figure 2.17 (upper left). A subdivision in sub-categories is doubtful in this case since the reference maps do not allow such refinement.

The land cover in 1975

The state of the land cover in the Study Area during the decade from 1971 to 1980 was obtained from a LANDSAT TM5 scene acquired for 1975. This image was originally classified into 10 different land cover classes using a standard Maximum Likelihood Classifier with a spatial resolution of 50×50 m (LfU-IER-ILPÖ). In order to make compatible the land cover classes considered in this study with those of the original classification, the following reclassification and aggregation was conducted: a) forest comprises those areas originally classified as deciduous, conifer and mixed forest; b) impervious cover consists of scattered and dense settlements, industrial areas; and c) permeable cover contains grassland, arable land, vineyards-orchards and water bodies. The final land cover map obtained as a result of the reclassification is depicted in Figure 2.17 (upper right).

The land cover in 1984

The condition of the land cover in the Study Area during the decade between 1981 and 1990 is represented by a LANDSAT TM7 scene acquired for July 1984, with 7 spectral bands and a spatial resolution of 30×30 m (INS). In order to derive a land cover map from the raw data, it was first necessary to georeference it to the Gauss-Krueger coordinate system (UTM), and then to classify it into four land cover classes, namely: the three classes mentioned above plus water bodies. The reason for proceeding in this way is the diametric differences regarding the reflectance of the land cover categories defined before and water bodies. Had this not been done (i.e. performing the classification with three classes), the classifying rules would have been distorted, i.e. the land cover classes would have contained a number of misclassified pixels, thus affecting the overall classification accuracy.

After the classification was performed, the category -water bodies- was aggregated to permeable areas because of both its compatibility and insignificant overall share. In order to classify a satellite image, training and validation sites, as well as a classification algorithm are required. In this case, the unique sources for the former were the topographic maps and orthophotos available for the region (scale 1 : 25 000, LVA). Regarding the latter, a fuzzy rule-based classifier proposed by Bárdossy and Samaniego (2001) was used because its result, the land cover map for 1984 shown in Figure 2.17 (lower left), got the highest index of agreement (i.e. the lowest number of misclassified pixels) during the validation phase of the classification as compared with results obtained by standard methods such as maximum likelihood or unsupervised cluster-type classifiers.

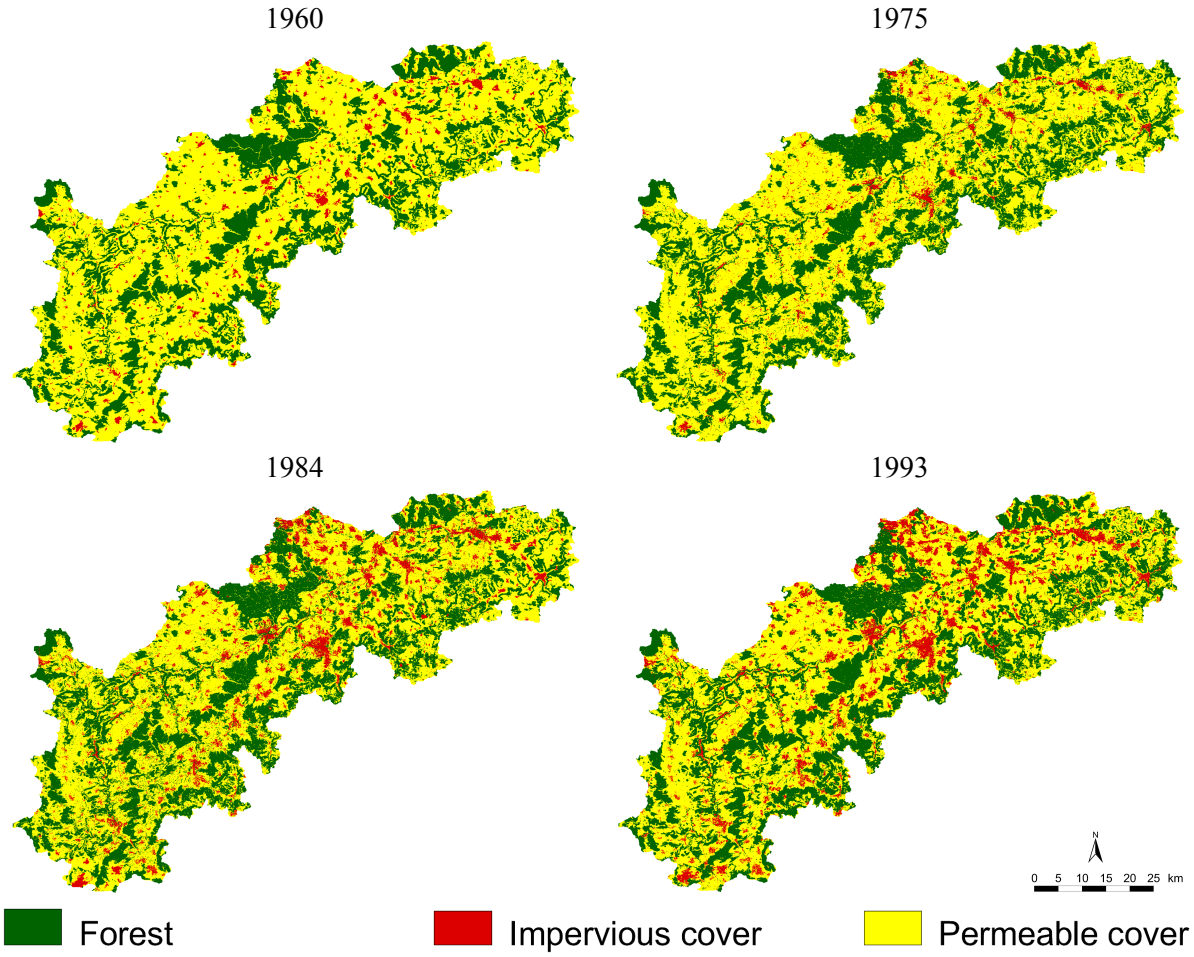

Figure 2.17 Time series of the land cover of the Study Area from 1960 to 1993. (LVA, LANDSAT).

The land cover in 1993

The last land cover map available for this study was a LANDSAT TM scene (IPK, 1995) with a spatial resolution of 30×30 m which originally classified its information into 16 classes containing 64 sub-categories. In order to fulfil the needs of the present study, the 16 original classes were aggregated as follows: a) forest, composed of conifer, deciduous, and mixed forest; b) impervious cover, composed of dense and scattered settlement, as well as of sealed industrial areas; and c) permeable cover, composed of the rest, namely: arable land, vineyards, intensive fruit production, fallow land, open areas (i.e. not sealed with no vegetation), intensive grassland, wetlands, extensive grassland

(dry), traditional orchards, and water bodies. The final result of this aggregation is shown in Figure 2.17 (lower right).

2.6.2 Fraction of a Given Land Cover [%]

In the present study, the proportion of a land cover category l in percent, in a given spatial unit i, and for the time t will be used as an indicator of the intensity of the land-atmosphere interactions (Section 2.6). Hence, the proportion of the spatial unit covered by forest ($l = 1$) can be computed by

$$x_{i17}^{t} = \frac{100}{N_i} \sum_{k \in \mathcal{L}_i} \mu_V^t(k), \tag{2.23}$$

where

$$\mu_V^t(k) = \begin{cases} 1 & \text{if} \quad V_k^t = l = 1 \equiv \text{forest} \\ 0 & \text{otherwise} \end{cases} \tag{2.24}$$

$\mathbf{V}^t$ Land cover image at point in time $t = 1961, 1975, 1984, 1993$.

$\mathcal{L}_i \subseteq \Omega_i$ A subset of the spatial unit Ω_i considered appropriate for the analysis.

In the same way, the fraction of impervious cover ($l = 2$) is estimated by

$$x_{i18}^{t} = \frac{100}{N_i} \sum_{k \in \mathcal{L}_i} \mu_V^t(k), \tag{2.25}$$

with

$$\mu_V^t(k) = \begin{cases} 1 & \text{if} \quad V_k^t = l = 2 \equiv \text{impervious} \\ 0 & \text{otherwise} \end{cases} \tag{2.26}$$

Finally, the fraction of permeable cover ($l = 3$) can be easily estimated by

$$x_{i19}^{t} = 100 - x_{i17}^{t} - x_{i18}^{t}, \tag{2.27}$$

since -by definition- the sum of the areas of the three land cover categories within a spatial unit i at any time t is always equal to A_i.

For the remaining points in time, i.e. for all t different from $1961, 1975, 1984,$ and 1993, a reasonable estimate for the proportion of land cover of a spatial unit i can be calculated by a linear interpolation of the closer upper and lower land cover observations since these variables are changing slowly over time.

The domain $\mathcal{L}_i$, where the previous functions are to be evaluated, will be chosen according to the necessities of the subsequent analyses. For instance, if $\mathcal{L}_i \equiv \Omega_i$, then previous variables will estimate the fraction of a given land cover type at basin level; whereas if $\mathcal{L}_i \equiv \mathcal{B}_i \subset \Omega_i$, these variables will estimate the same categories of land cover on the riparian zones of the stream network $\mathcal{N}_i$.

2.7 Climatic or Meteorological Factors

Climatic and meteorological factors are the most dynamic components of the basin's water cycle because of the intrinsic complexity and chaotic behaviour of the atmospheric processes governing the weather at the macroscale. In general, a time series of a climatic factor (i.e. an observable such as precipitation or temperature) is characterized by a combination of: 1) periodic variations (e.g. seasonal or cyclic oscillations); 2) gradual changes of the average (i.e. a trend); 3) sudden changes (e.g. a change in the climatic regime); 4) serial dependence (i.e. persistence or serial correlation); and 5) pure random component or noise (Chow 1964, Chatfield 1989). All these characteristics are to be considered during the estimation of the climatic factors as will be explained afterwards.

Since the water cycle exhibits an intrinsic seasonality -i.e. it behaves differently during summer and winter- and considering that the aim of the study is to filter out the likely effects of the land cover change from the climatic variations, the estimation of the climatic factors will be carried out for both "water seasons", namely: summer and winter. In the Study Area, summer, from a hydrological point of view, starts on the 1st of May and ends on the 31st of October, whereas winter spans from the 1st of November to the 30th of April.

Another reason why the climatic factors should be computed at annual or semi-annual intervals is related with the serial dependence among the observations of the climatic time series during the analysis. Serial dependence can be avoided if the autocorrelation $r(k)$ lies between the 95% confidence intervals given approximately by $\pm 2/\sqrt{T}$. Within such limits a time series can be considered as a random one (Chatfield 1989). Figure 2.18 shows the correlograms for the annual and daily precipitation time series for two catchments with different sizes in the Study Area. In the present study, the autocorrelation between observations a k apart from each other, of a time series $\{x_{ij}^1,\ldots,x_{ij}^t,\ldots,x_{ij}^T\}$ of an observable j in a given spatial unit i, during the period $t = 1,\ldots,T$, is given by

$$r_{ij}(k) \cong \frac{\sum_{t=1}^{T-k}\left(x_{ij}^t - \overline{x}_{ij}\right)\left(x_{ij}^{t+k} - \overline{x}_{ij}\right)}{\sum_{t=1}^{T}\left(x_{ij}^t - \overline{x}_{ij}\right)^2}, \tag{2.28}$$

where

$\overline{x}_{ij} = \frac{1}{T}\sum_{t=1}^{T} x_{ij}^t$ is the overall mean, and T is the number of observations.

2.7.1 Precipitation

Precipitation is a major factor governing the water balance of a region (Ward and Robinson 2000), hence its occurrence and spatial distribution are key elements to understand the water cycle at the catchment level. Since the complex mechanisms governing this phenomenon are a science of their own, they can not be covered in this study; instead of that, a statistical analysis will be used to predict the precipitation $p_k^t\ [mm]$ on each cell $k \in \Omega_i$ for every day during the interval $t \in [01.11.1960,\ldots,31.10.1993]$. For this analysis the spatial resolution of every cell is 300×300 m.

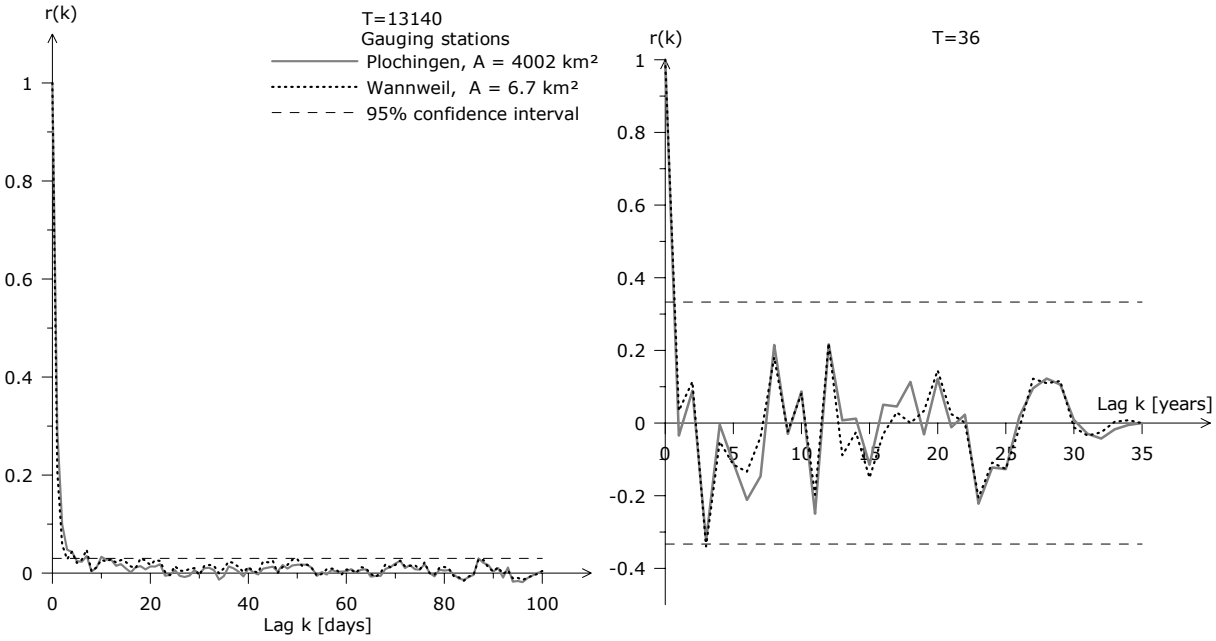

Figure 2.18 Correlograms of daily (left panel) and annual (right panel) precipitation for two basins with different sizes in the Study Area.

Thus, the goal is to estimate the spatial distribution of p_k^t based on daily raingauge records available for the 288 meteorological stations located both inside and in the surroundings of the Study Area (see Figure 2.19). Traditional interpolation methods such as the arithmetic mean, Thiessen-polygons, isohyetal, and inverse distance among others can be used, but they have the following shortcomings: either a) the spatial continuity is ignored or lost at longer distances, or b) their estimation error is not unbiased. Furthermore, additional information, for instance the variation of either the precipitation or the surface temperature according to the elevation (i.e. the orographic effect) can not be thoroughly considered.

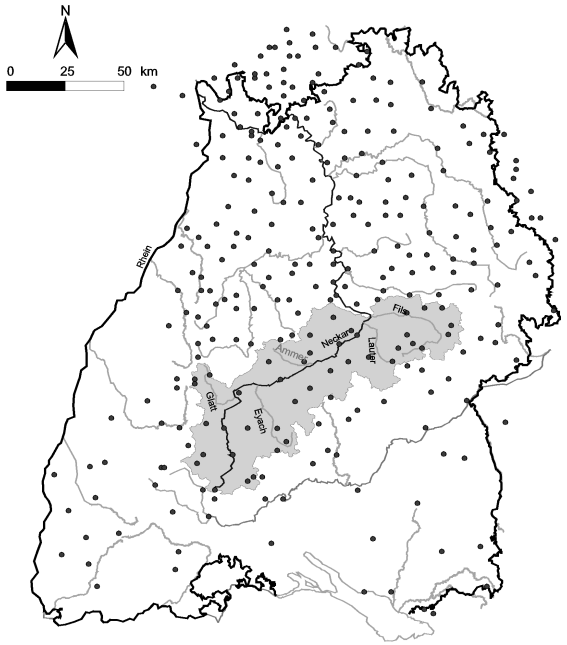

Figure 2.19 Meteorological stations located in the State of Baden-Württemberg and its neighbouring States used in this study (Source LfU and DWD).

Because of that, a geostatistical technique called External Drift Kriging[5] (EDK) (Ahmed and de Marsily 1987) using a DEM as additional information is to be used. This method estimates the value of a variable at an unsampled point, say p_k^t, from the neighbouring ones within the field of values. The variability within the field is assumed to be function of the distance and direction (Wilby, 1997), which is represented by a variable's variogram.

In practice, the variogram for a given variable has to be estimated from the available data (i.e. a sample), hence it is called experimental or sample variogram. In order to calculate it, the following definitions have to be introduced.

Let the regionalized variable precipitation $\{p_k^t : k \in \Omega_i \; t = 1,\ldots,\mathcal{T}\}$ be a realization of a random function $\{P_k : k \in \Omega_i \subset \mathbb{R}^3\}$. Based on this definition, and assuming that the "*intrinsic hypothesis"* holds, i.e. the expectation of this random function is constant all over the domain Ω_i and the variance of the increment of the random function P_k at two different locations, say $\vec{k}_l$ and $\vec{k}_{l'}$, depends only on the vector separating them, i.e. $\vec{h}$ (Bárdossy 1997, Chilès and Delfiner 1999), the variogram function can be defined as follows

$$\hat{\gamma}^t(h) = \frac{1}{2N_h} \sum_{\vec{k}_l - \vec{k}_{l'} \approx \vec{h}} \left(p_{k_l}^t - p_{k_{l'}}^t\right)^2, \tag{2.29}$$

where

$\vec{k}_l, \vec{k}_{l'}$ is the position vectors of raingauge stations l and l' [m],

$p_{k_l}^t$ is the observed precipitation in the raingauge station l in time t expressed in [mm],

N_h the number of raingauge stations separated by a distance $h \pm \varepsilon\,[\mathrm{m}]$,

ε a tolerance value in [m],

$\hat{\gamma}^t(h)$ the sample variogram for daily (or annual) precipitation in time t in [mm^2]

l $= 1,\ldots,\mathcal{L}$, an index denoting those points of the domain with known values of the random function (i.e. the raingauge stations in Baden-Württemberg and its surroundings, $\mathcal{L} = 288$),

t a time index. Only for the sake of simplicity, it may represent in this case either daily or annual precipitation.

As an example, Figure 2.20 shows the sample variogram for the annual average precipitation in the State of Baden-Württemberg obtained from the 288 raingauge stations depicted in Figure 2.19 during the period 1961-1995.

---

[5] Word coined in recognition of D.G. Krige, a South African mining engineer, who used for first time statistical methods to assess ore reserves in the early 1950s. Kriging, as a Geostatistical procedure was later formalized by G. Matheron in 1970.

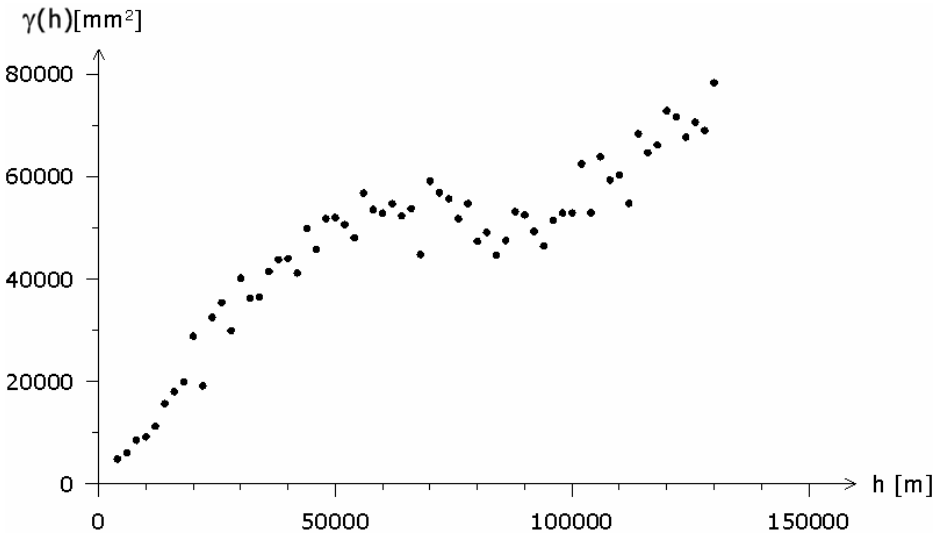

Figure 2.20 Experimental variogram for the annual average precipitation for the State of Baden-Württemberg.

Based on the sample variogram, a theoretical one $\gamma(h)$ suitable for the further analysis can be fitted using a robust method, e.g. L1 norm (Chilès and Delfiner 1999). In the present study, an isotropic theoretical variogram composed of a pure nugget effect and a spherical variogram was adopted for the daily precipitation based on a previous work carried out by Bárdossy et al. (CC-HYDRO, 1999). Its equation is:

$$\gamma(h) = \begin{cases} 0 & \text{if} \quad h = 0[m] \\ 0.02 & \text{if} \quad 0 < h \le 1[m] \\ 0.98\left(\frac{3}{2}\frac{h}{r_p} - \frac{1}{2}\frac{h^3}{r_p^{\,3}}\right) & \text{if} \quad 1 < h \le r_p = 40\,000[\mathrm{m}] \\ 0.98 & \text{if} \quad h > r_p \end{cases}, \tag{2.30}$$

where r_p is a constant denoting the range of the theoretical variogram for daily precipitation in [m].

A linear estimator for the precipitation $\hat{p}_k^t$ at an unsampled location k in time t can be obtained as a linear combination of all sampled locations by

$$\hat{p}_k^t = \sum_{l=1}^{L} \lambda_l^t p_{k_l}^t \,. \tag{2.31}$$

In order to determine the weights λ_l^t, the fitted variogram (2.30) as well as the supplementary knowledge available, i.e. that precipitation is related to the terrain's elevation (Gilman in Chow 1964), are to be used. The inclusion of this additional information implies a modification of the first assumption of the intrinsic hypothesis, namely: the expectation of p_k^t is in a linear relationship with the elevation Z_k, in other words

$$E\left[p_k^t \middle| Z_k\right] = c_1 + c_2 Z_k \quad \forall\, k \in \Omega_i \,, \tag{2.32}$$

where c_1 and c_2 are unknown constants.

Hence, the expectation of the linear estimator is

$$E\left[\hat{p}_k^t\right] = E\left[\sum_{l=1}^{\mathcal{L}} \lambda_l^t p_{k_l}^t\right] = \sum_{l=1}^{\mathcal{L}} \lambda_l^t E\left[p_{k_l}^t\right] = \sum_{l=1}^{\mathcal{L}} \lambda_l^t \left[c_1 + c_2 Z_{k_l}\right] = c_1 + c_2 Z_k\,, \tag{2.33}$$

which implies the following constraints for the weights, namely

$$\sum_{l=1}^{\mathcal{L}} \lambda_l^t = 1\,, \tag{2.34}$$

and

$$\sum_{l=1}^{\mathcal{L}} \lambda_l^t Z_{k_l} = Z_k\,. \tag{2.35}$$

There are an infinite number of combinations for the weights λ_l^t that fulfil the previous constraints, but only one of them (i.e. "*the best linear unbiased estimator*") will minimize the error $\hat{p}_k^t - p_k^t$ characterized by its expected mean square $E\left[\left(\hat{p}_k^t - p_k^t\right)^2\right]$ (Chilès and Delfiner 1999).

Therefore, the minimisation problem can be stated as follows

$$\begin{cases} \min E\left[\left(\hat{p}_k^t - p_k^t\right)^2\right] \\ \text{subject to} \quad \sum_{l=1}^{\mathcal{L}} \lambda_l^t = 1 \quad \text{and} \quad \sum_{l=1}^{\mathcal{L}} \lambda_l^t Z_{k_l} = Z_k \end{cases}. \tag{2.36}$$

Lagrange multipliers μ_1 and μ_2 can be used to solve (2.36) since it is a constrained optimisation, which will lead to the following system of equations (Ahmed and De Marsily 1987)

$$\begin{cases} \sum_{l'=1}^{\mathcal{L}} \lambda_{l'}^t \gamma(k_l - k_{l'}) + \mu_1 + \mu_2 Z_{k_l} = \gamma(k_l - k) \qquad l', l = 1, \ldots, \mathcal{L} \\ \sum_{l'=1}^{\mathcal{L}} \lambda_{l'}^t = 1 \\ \sum_{l'=1}^{\mathcal{L}} \lambda_{l'}^t Z_{k_l} = Z_k \end{cases}. \tag{2.37}$$

This system, which is a function of the theoretical variogram, has to be solved for each cell k and each point in time t within the domain Ω_i. As an example, Figure 2.21 shows the results obtained with this method for the spatial distribution of the annual cumulated precipitation in 1963, 1973, 1983 and 1993.

Cumulative precipitation [mm]

Based on the daily interpolated precipitation $\hat{p}_k^{t(d)}$ for a day d of the water year t and considering that the area of each cell $k \in \Omega_i$ is constant; the annual precipitation in a given catchment Ω_i at the end of a "water year" t can be obtained by

$$x_{i20}^t = P_i^t = \left\langle \hat{p}_k^{t(d)} \right\rangle_i = \iint_{\Omega_i} \int_t \hat{p}_k^{t(d)} d\Omega dt \cong \frac{1}{N_i} \sum_{k \in \Omega_i} \sum_{d=1}^{365} \hat{p}_k^{t(d)}\,. \tag{2.38}$$

The beginning of the water year t is the 1st of Nov. of the Julian year $t-1$ (i.e. $d = 1$), whereas its ending is on the 31st Oct. of the Julian year t (i.e. $d = 365$). For instance, the "water year" 1966 begins on 1st of Nov. of the Julian year 1965 and ends on the 31st of October of the Julian year 1966.

Similarly, the cumulative winter precipitation (from the 1st of Nov. to the 30th of Apr.) occurred in a basin i during a water year t is

$$x_{i21}^{t} = \frac{1}{N_i} \sum_{k \in \Omega_i} \sum_{d=1}^{181} \hat{p}_k^{t(d)} , \tag{2.39}$$

and hence, the cumulative summer precipitation (from the 1st of May. to the 31st of Oct.) can be obtained as

$$x_{i22}^{t} = P_i^t - x_{i21}^{t} . \tag{2.40}$$

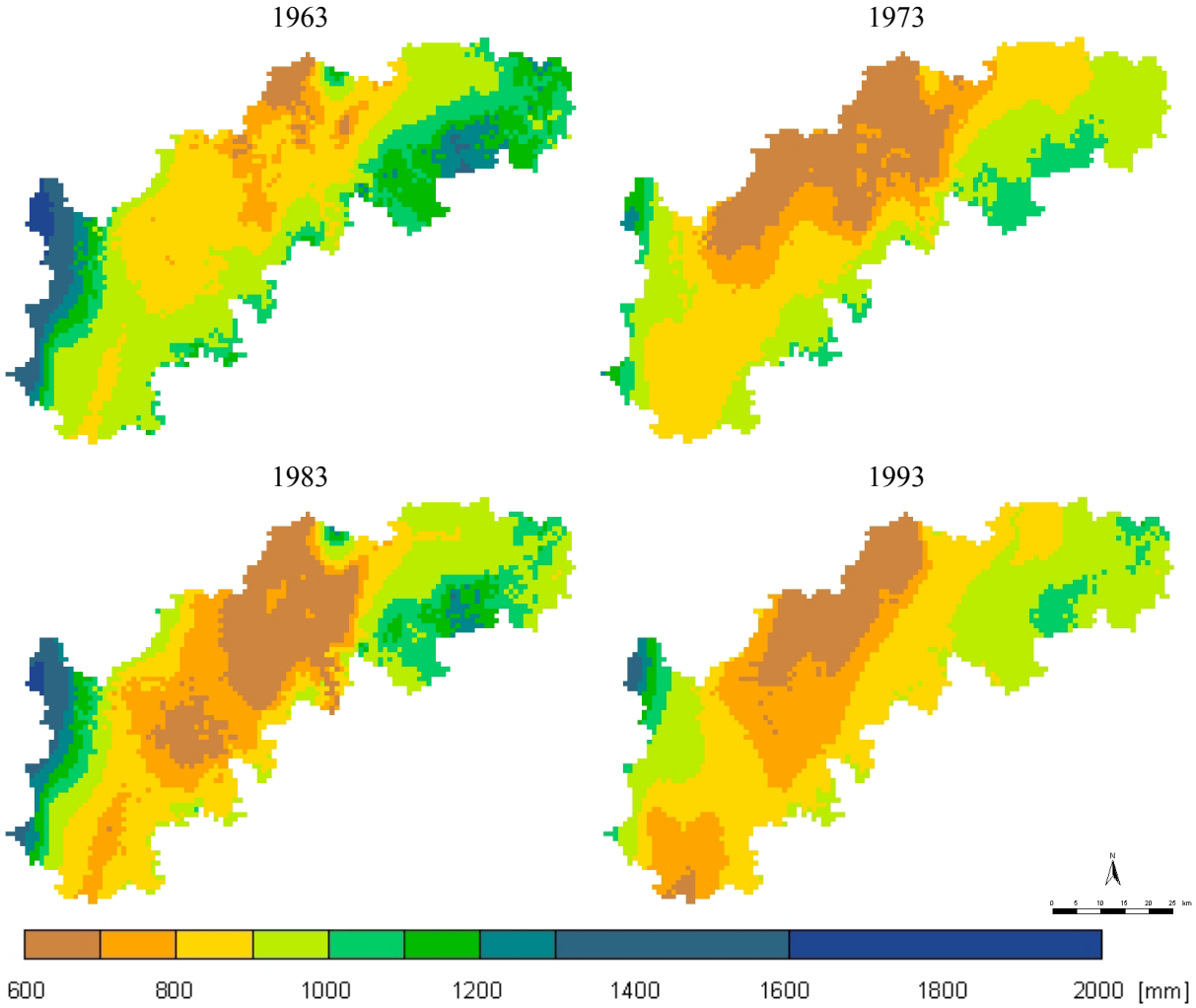

Figure 2.21 Spatial distribution of annual precipitation with in the Study Area for the years 1963, 1973, 1983 and 1993.

Mean precipitation [mm]

In order to estimate the mean precipitation at the spatial unit Ω_i in the "water year" t the following procedure is to be applied. Firstly, a daily expected value at this spatial level can be obtained by averaging daily precipitation values obtained before [see (2.31)] at the cell level for all those cells contained within a given basin i. This procedure will generate a time series of expected daily precipitation at the basin level, say $\overline{p}_i^{t(d)}$, which can be estimated by

$$\overline{p}_i^{t(d)} = \frac{1}{N_i} \sum_{k \in \Omega_i} \hat{p}_k^{t(d)} \,. \tag{2.41}$$

Once this time series has been estimated, the mean or expected precipitation [mm] in basin i in the "water year" t is

$$x_{i23}^t = \frac{1}{365} \sum_{d=1}^{365} \overline{p}_i^{t(d)} \,. \tag{2.42}$$

Using the same reasoning, the mean precipitation during winter and summer in the "water year" t can be calculated by

$$x_{i24}^t = \frac{1}{181} \sum_{d=1}^{181} \overline{p}_i^{t(d)} \,, \tag{2.43}$$

and

$$x_{i25}^t = \frac{1}{184} \sum_{d=182}^{365} \overline{p}_i^{t(d)} \,, \tag{2.44}$$

respectively. Precipitation time series have been provided by the DWD, which consider the length of the "water year" equal to 365 days.

Antecedent precipitation index (API)

After a rainfall, an affected basin's topsoil layer will dry up due to combined effects of evapotranspiration and infiltration. A simple approximation to this complex problem is to assume that a percentage of the precipitation $(1-\kappa)\%$ is lost every day. Then a gross indicator for the residual soil moisture or antecedent precipitation index can be estimated as follows (Linsley, Kohler and Paulhus 1982)

$$x_{i26}^{t(d)} = \sum_{c=0}^{C} \kappa^c \overline{p}_i^{t(d-c)} \,, \tag{2.45}$$

where

κ recession constant, commonly ranging within the interval $0.85 < \kappa < 0.98$ (Chow, 1964).

c $0,\ldots,C$, time index denoting the precipitation occurred c days before the event $t(d)$. t represents in this case a water year and d a given day of that year. C could be 15, 30, 60, 90 or 120 days.

Based on (2.45), the maximum antecedent precipitation of a given spatial unit i in the "water year" t is

$$x_{i27}^t = \max\left(x_{i26}^{t(d)}\right) \quad \forall d = 1,\ldots,365 \,. \tag{2.46}$$

Maximum values of API at seasonal basis can also be estimated using (2.45). Important values for further analysis may be the maximum API during winter and summer, such indicators are given by

$$x_{i28}^t = \max\left(x_{i26}^{t(d)}\right) \quad \forall d = 1,\ldots,181 \,, \tag{2.47}$$

and

$$x_{i29}^{t} = \max\left(x_{i26}^{t(d)}\right) \quad \forall d = 182,\ldots,365\,, \tag{2.48}$$

respectively. An example of the yearly development of variable x_{i26}^{t} can be seen in Figure 2.23.

2.7.2 Temperature

Air temperature at surface level constitutes another basic indicator of the system because of its tight relationships with the potential evapotranspiration and snowmelt, which, in turn, will determine the amount of discharge produced within a basin during a period of time.

The spatial distribution of temperature υ_k^t [°C] for all cells $k \in \Omega_i$ in a given time t is also strongly related with the terrain's elevation (Petterssen in Chow, 1964). Hence, EDK (see Section 2.7.1) can also be used to determine the temperature for all unsampled cells within the domain based on the information available, namely daily average temperatures measured in every station shown in Figure 2.19 during the period $\{01.11.1960,\ldots,31.10.1993\}$. For this analysis the spatial resolution of every cell is 300×300 m, too.

In this case also, a theoretical variogram fitted by Bárdossy et al. (CC-HYDRO, 1999) for the State of Baden-Württemberg will be used, namely

$$\gamma(h) = \begin{cases} 0 & \text{if} \quad h = 0[\mathrm{m}] \\ 0.201231 & \text{if} \quad 0 < h \le 1[\mathrm{m}] \\ 1.185569\left(\dfrac{3}{2}\dfrac{h}{r_\upsilon} - \dfrac{1}{2}\dfrac{h^3}{r_\upsilon^{\,3}}\right) & \text{if} \quad 1 < h \le r_\upsilon = 39\,000[\mathrm{m}] \\ 1.185569 & \text{if} \quad h > r_\upsilon \end{cases}, \tag{2.49}$$

where r_υ is a constant denoting the range of the theoretical variogram for daily temperature in [m].

Figure 2.22 depicts a sample of the time series of mean temperature in January in the Study Area obtained by EDK.

Higher temperatures during certain months of the year may be important indicators of the behaviour of the system. For instance, during January they could be related to snowmelt and lower snow accumulation, which in turn would contribute to winter flooding. Conversely, higher temperatures would imply long drought periods (i.e. low flows in streams) during the summer season, and eventually, large amounts of water will be evaporated since the ground surface has been heated up. This warm-moist air would rise until it cools down, condensates, and finally triggers a local intensive precipitation (Ward and Robinson 2000). This phenomenon is called convective precipitation. Such intensive rainfall would, in turn, generate high peak flows and possibly flooding events in small catchments. Due to these reasons, the following indicators will be calculated based on the daily interpolated temperature $\hat{\upsilon}_k^{t(d)}$ for each cell k and in time $t(d)$.

Mean temperature [°C]

The mean temperature for the spatial unit Ω_i in time $t(d)$ is estimated by

$$\overline{v}_i^{t(d)} = \frac{1}{N_i} \sum_{k \in \Omega_i} \hat{v}_k^{t(d)} , \tag{2.50}$$

then, the mean temperature of January and July in the "water year" t for the spatial unit Ω_i are

$$x_{i30}^t = \frac{1}{31} \sum_{d=62}^{92} \overline{v}_i^{t(d)} , \tag{2.51}$$

and

$$x_{i31}^t = \frac{1}{31} \sum_{d=243}^{273} \overline{v}_i^{t(d)} , \tag{2.52}$$

respectively. Temperature time series provided by the DWD consider the length of the "water year" equal to 365 days.

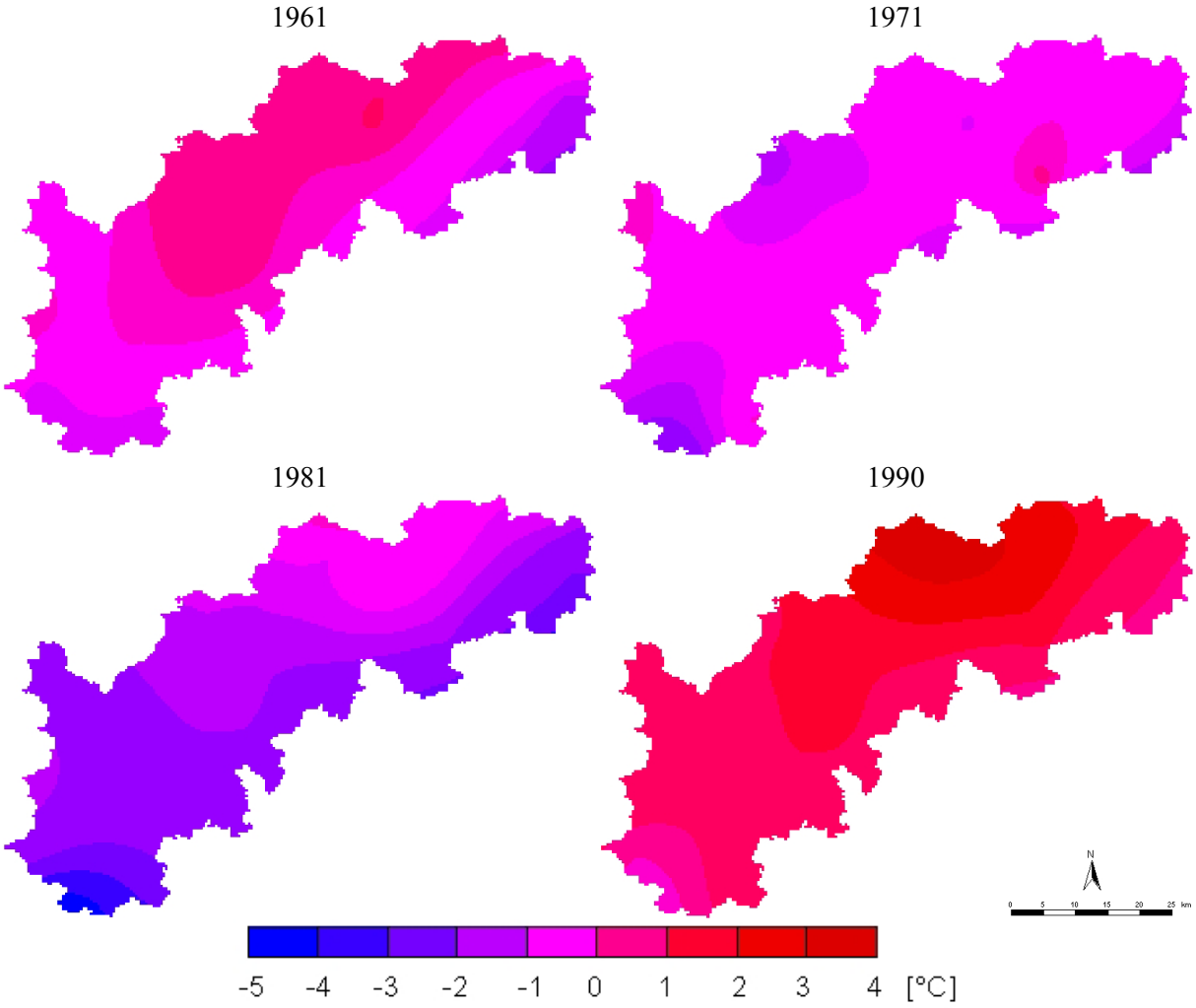

Figure 2.22 Sample of the spatial distribution of the mean temperature in January in the Study Area for the years 1961, 1971, 1981 and 1990 (Source data DWD).

Maximum temperature [°C]

The maximum temperature occurring in January and July in the "water year" t for the spatial unit Ω_i are

$$x_{i32}^{t} = \max\left(\overline{v}_i^{t(d)}\right) \quad d = 62,\ldots,92\,, \tag{2.53}$$

and

$$x_{i33}^{t} = \max\left(\overline{v}_i^{t(d)}\right) \quad d = 243,\ldots,273\,, \tag{2.54}$$

respectively.

Antecedent temperature index (ATI) [K]

As explained before the amount of heat cumulated on the ground surface would originate many meteorological phenomena closely related with the runoff of a catchment. Hence, Hopkins and Hackett (1961) devised an indicator proportional to the cumulated seasonal gain or loss of heat in the system in a given point in time. An example of the yearly development of this index can be seen in Figure 2.23. This indicator is proportional to the difference between the current air temperature and antecedent temperature index (Melloh 1999). It can be estimated recursively as follows

$$x_{i34}^{t(d)} = x_{i34}^{t(d-1)} + \zeta\left(\overline{v}_i^{t(d)} - x_{i34}^{t(d-1)}\right), \tag{2.55}$$

where

ζ temperature weighting multiplier for the previous day period. Commonly ranging within the interval $0.1 < \zeta \leq 1$ (Melloh, 1999).

$x_{i34}^{t(d)}$ antecedent temperature index for the day d of year t. $x_{i34}^{t(d)} = 0$ if $d = 1 \wedge t = 1961$, which corresponds to the beginning of the winter in year 1961.

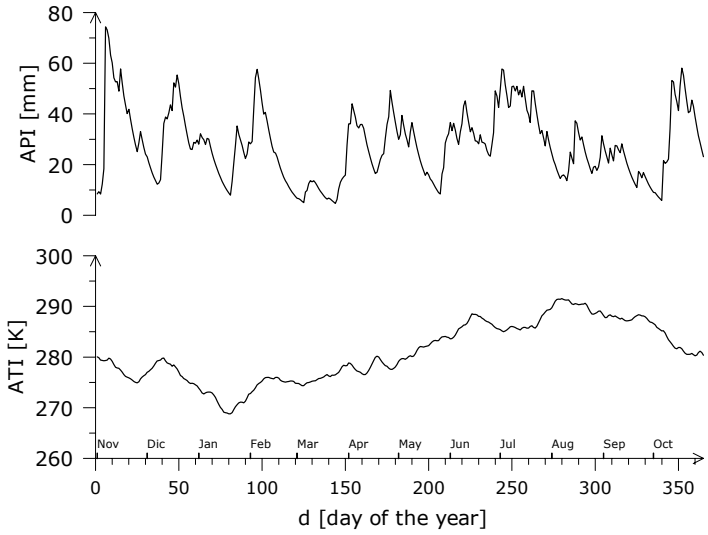

Figure 2.23 Comparison of the daily development of the API and ATI for catchment No. 3 during the water year 1980.

The maximum ATI occurring either during the "water year" or during its winter or its summer t for the spatial unit Ω_i can be calculated by

$$x_{i35}^{t} = \max\left(x_{i34}^{t}\right) \quad d = 1,\ldots,365\,, \tag{2.56}$$

$$x_{i36}^{t} = \max\left(x_{i34}^{t}\right) \quad d = 1,\ldots,181\,, \tag{2.57}$$

and

$$x_{i37}^{t} = \max\left(x_{i34}^{t}\right) \quad d = 181,\ldots,365\,, \tag{2.58}$$

respectively.

2.7.3 Circulation Patterns [-]

Many anomalies of the water cycle (e.g. low flows or extraordinary floods) that occur at a mesoscale depend on the occurrence of particular macroscale (or continental) atmospheric circulation patterns. In recent studies, for instance, Redmont and Koch (1991) have linked the Pacific North America Index (PNA) to precipitation, temperature, and annual stream flow. Bárdossy and Plate (1992) have modelled the spatial distribution of precipitation related to the occurrence of atmospheric circulation patterns (CPs). Duckstein et al. (1993) have linked daily CP occurrences and partial duration of floods. Dracup and Kahya (1994) and then Piechota and Dracup (1996) have related streamflow anomalies and the occurrence of indices based on sea level pressure such as El Niño Southern Oscillation (ENSO). Shorthouse and Arnell (1997) have correlated the North Atlantic Oscillation Index (NAOI) and the average monthly runoff. Stahl and Demuth (1999) have successfully linked streamflow drought and CPs for mesoscale basins in southern Germany.

In the present case, the European atmospheric circulation patterns (CPs) or "*Großwetterlagen*" according to Hess and Brezowsky (1969) will be used to downscale their effects to mesoscale catchments, especially with regard to total drought duration and total duration of peak flows. The CP is a synoptic meteorological classification used by the German Weather Service [*Deutscher Wetterdienst* (DWD)], which is based on mean air pressure distribution over Europe and the northern Atlantic Ocean. This index is defined for a big spatial domain comprised between the coordinates 40°W, 30°N and 60°E, 80°N (Bárdossy, 1993).

The CPs proposed by Hess and Brezowsky distinguish between three major circulation types, namely: zonal, mixed, and meridional circulation types. These types are further subdivided according to the direction of movement of frontal zones, location of high and low pressure areas, and cyclonic and anti-cyclonic rotation. As a result 29 CPs plus one unclassified CP are obtained (Hess and Brezowsky, 1969), (see Appendix 7).

For the present study daily CPs occurrences for the period 1.1.1960 to 31.10.1993 were obtained from DWD. It has been observed and reported by many authors (e.g. Bárdossy, 1993, Stahl and Demuth, 1999) that the occurrence of certain types of CPs can be linked with the amount of precipitation in a given basin, and this, in turn, with the occurrence of wet and dry spells.

In order to cluster the CPs onto three groups, e.g. wet, normal and dry periods, a seasonal wetness index W_j can be estimated. This index estimates the ratio between the relative amount of precipitation occurring when a given CP-type j takes place, and the relative frequency of such CP. This index can be explicitly written as

$$W_j = \frac{\frac{1}{P}\sum_{d=1}^{\mathrm{T}} p_{\Omega j}^{d}}{\frac{1}{\mathrm{T}}\sum_{d=1}^{\mathrm{T}} \mu_j^{d}}, \tag{2.59}$$

where,

$$\mu_j^d = \begin{cases} 1 & \text{if } \mathrm{CP}(d) = j \\ 0 & \text{otherwise} \end{cases} \tag{2.60}$$

$$p_{\Omega j}^d = \begin{cases} \overline{p}_\Omega^d & \text{if } \mathrm{CP}(d) = j \\ 0 & \text{otherwise} \end{cases} \tag{2.61}$$

$$P = \sum_{d=1}^{\mathrm{T}} \overline{p}_\Omega^d \tag{2.62}$$

d Daily time index $\in \{1.11.1960, \ldots 31.10.1993\}$.

j CP-type index, $j = 1, \ldots, 30$; The equivalence of the CP indexes and the CP description can be found in Appendix 7.

T Total number of days either in summer or in winter season during the period $\{1.11.1960, \ldots 31.10.1993\}$.

P Total summer or winter precipitation in [mm] occurred at the Study Area (Ω) during the period $\{1.11.1960, \ldots 31.10.1993\}$.

$\mathrm{CP}(d)$ Atmospheric circulation pattern index according to Hess and Brezowsky for a given day d.

Using the index W_j the CPs were grouped into three categories by applying the following rules:

$$\text{if } \begin{cases} W_j \leq 0.6 & \Rightarrow \text{Dry period} \\ 0.6 < W_j \leq 1.0 & \Rightarrow \text{Normal period} \\ W_j > 1.0 & \Rightarrow \text{Wet period} \end{cases}. \tag{2.63}$$

The results of the above classification are shown in Table 2.2.

Table 2.2 Classification of circulation patterns (CPs) for winter and summer seasons according to the wetness index W for the Study Area.

Category	Circulation patterns winter	Circulation patterns summer
Dry	BM, HB, HFa, HM, HNa, HNFa, NEa, NWa, Sa, SEa, SEz, SWa, Wa	BM, HB, HFa, HM, HNa, NEa, NWa, Sa, SEa, SWa, Wa
Normal	HNFz, NEz, Sz, TB	HNFa, Na, Sz, U
Wet	HFz, HNz, Na, NWz, Nz, SWz, TM, TRM, TRW, U, Ws, WW, Wz	HFz, HNFz, HNz, NEz, NWz, Nz, SEz, SWz, TB, TM, TRM, TRW, WS, WW,Wz

In order to study possible relationships between the occurrence of "dry" circulation patterns and low flows occurring mainly during summer (based on the water budget of the Study Area it is clear that the main impact of low flows will occur during this season), two indicators were considered: 1) one that tallies the total number of occurrences of CPs clustered as "dry periods" for a given spatial unit i during the summer season of a year $t \in [1961, \ldots, 1993]$, that have a decreasing antecedent

precipitation index (x_{26}) for this spatial unit; and 2) one that reckons the total number of occurrences of CPs categorized as "dry periods" for each summer for the period 1961 to 1993. This second indicator is not spatial-unit specific because of the nature of the input data employed. Formally, these indicators can be written as

$$x_{i38}^{t} = \sum_{d=d_w+1}^{d_e} \mu_i^{t(d)} \quad , \tag{2.64}$$

$$x_{39}^{t} = \sum_{d=d_w+1}^{d_e} \mu^{t(d)} \quad , \tag{2.65}$$

where

$$d_e = \begin{cases} 365 & \text{if } t \text{ is a normal year} \\ 366 & \text{if } t \text{ is a leap year} \end{cases} \tag{2.66}$$

$$d_w = \begin{cases} 181 & \text{if } t \text{ is a normal year} \\ 182 & \text{if } t \text{ is a leap year} \end{cases} \tag{2.67}$$

$$\mu_i^{t(d)} = \begin{cases} 1 & \text{if } \begin{cases} \mathrm{CP}(t(d)) \in \{1\ 5\ 7\ 9\ 10\ 14\ 16\ 18\ 20\ 24\ 26\} \\ \wedge \\ x_{i26}^{t(d)} - x_{i26}^{t(d-1)} < 0 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{2.68}$$

$$\mu^{t(d)} = \begin{cases} 1 & \text{if } \mathrm{CP}(t(d)) \in \{1\ 5\ 7\ 9\ 10\ 14\ 16\ 18\ 20\ 24\ 26\} \\ 0 & \text{otherwise} \end{cases} . \tag{2.69}$$

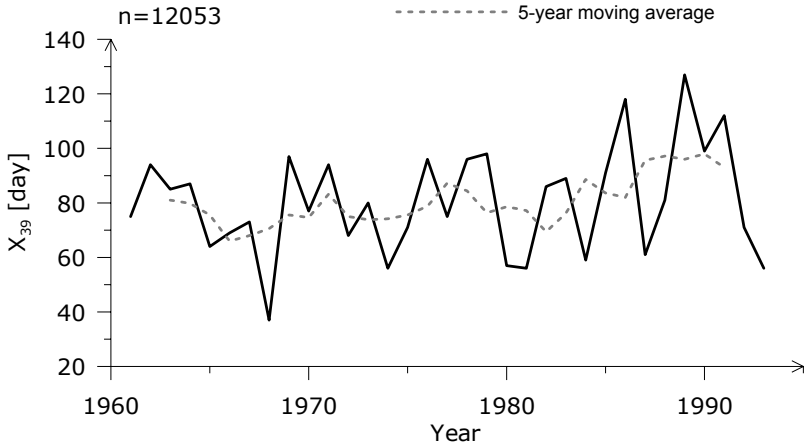

Figure 2.24 Time series showing the absolute frequency of occurrence of CPs classified as "dry periods" within the Study Area during summer.

By plotting the indicator x_{39} vs. time in Figure 2.24 it can be concluded that there exists a marked trend within the Study Area to have circulation patterns prone to cause droughts. It is also interesting to note that the relative increment of this indicator for a given year with its immediate subsequent

($x_{39}^{t+1} - x_{39}^{t}$) tends to grow along the time axis. Such developments are clear indications of macroclimatic changes, which in the present study can be considered as exogenous climatic variables.

For the analysis of peak flows two variables have been devised. Firstly, a variable that reckons the number of days in summer on which both the occurrence of "wet" circulation periods and an antecedent precipitation index greater than a given threshold occur simultaneously. In other words, a flood may be expected if a certain climatic condition and a given amount of precipitation during a continuous period have happened. Secondly, a variable that counts the number of occurrences of wet circulation periods during winter.

The former of the two variables mentioned above can be explicitly written as

$$x_{i40}^{t} = \sum_{d=d_w+1}^{d_e} \mu_i^{t(d)} \quad , \tag{2.70}$$

with

$$\mu_i^{t(d)} = \begin{cases} 1 & \text{if} \quad \begin{cases} \mathrm{CP}(t(d)) \in \{2\ 3\ 4\ 6\ 8\ 11\ 13\ 15\ 17\ 19\ 21\ 23\ 25\ 28\ 29\} \\ \wedge \\ x_{i26}^{t(d)} \geq \mathrm{F}_i\,(0.80) \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{2.71}$$

where

$\mathrm{F}_i\,(0.80)$ is a threshold value representing an antecedent precipitation index (API) in $[\mathrm{mm}]$ equal to the 80th percentile of variable x_{26}, or in other words, the API to be equalled or exceeded 20% of the time in basin i. Values above this threshold can be regarded as days on which heavy rainfall occurred just before the day d, and thus, saturation of the upper soil horizons had been reached,

and the latter variable as

$$x_{i41}^{t} = \sum_{d=1}^{d_w} \mu_i^{t(d)} \ , \tag{2.72}$$

with

$$\mu_i^{t(d)} = \begin{cases} 1 & \text{if} \quad \{\mathrm{CP}(t(d)) \in \{2\ 3\ 4\ 6\ 8\ 11\ 12\ 13\ 15\ 17\ 21\ 29\ 30\} \\ 0 & \text{otherwise} \end{cases} . \tag{2.73}$$

2.8 The System's Output: Runoff

Runoff is the *throughput* of the water cycle that appears at the outlet of a catchment during a given period (see Figure 2.4). In other words, it "is that part of the precipitation, as well as any other flow contributions, which appears in surface streams of either perennial or intermittent flow" (Chow, 1964). In accordance with these definitions, runoff can be measured as the volume of water per unit of

time $[\mathrm{m^3 s^{-1}}]$, but in order to ease comparison among catchments with different sizes, runoff can be expressed also in $[\mathrm{m^3 s^{-1} km^{-2}}]$, or simply in $[\mathrm{mm}]$, i.e. *specific runoff*.

Consequently, the runoff cycle is inherently dependent on the nature of two main cyclic processes: precipitation and air surface temperature. The latter, as opposed to the former, has a marked annual seasonality. As a result, runoff will exhibit some annual fluctuations and persistence along the time axis. As shown in Figure 2.25 (left panel), the bigger the size of the catchment, the higher the serial correlation coefficients $r(k)$ turn out to be. But in any case, when the lag is greater than 100 days the serial correlation is not significant anymore (at a p-value of 5%). This means that those values to be obtained for the subsequent analyses should be separated at least by 100 days to avoid autocorrelation. In the present study specific runoff, or any of its characteristics, will be evaluated annually or semi-annually.

Besides that, as it is depicted in Figure 2.25 (right panel), cumulated specific runoff exhibits a clear seasonality almost independent of catchment size. This means that cyclic fluctuations have a macroclimatic origin. Additionally, the correlogram also shows that cumulated annual runoff has a non-significant serial correlation.

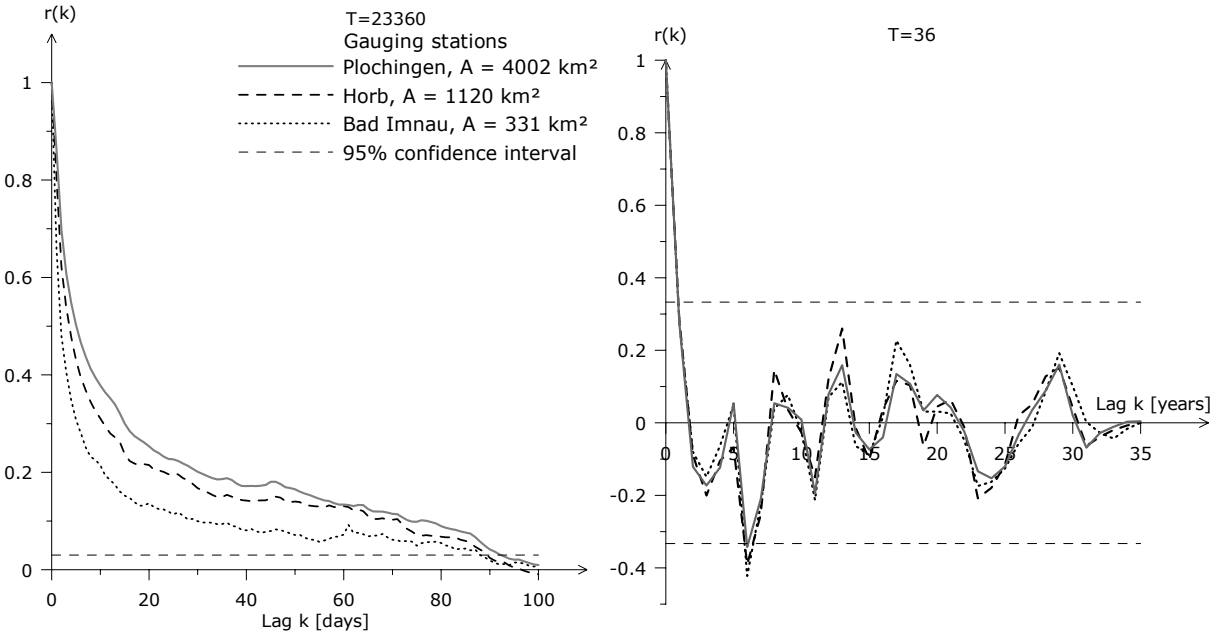

Figure 2.25 Autocorrelation functions of daily (left panel) and annual (right panel) discharge for three basins within the Study Area.

The basic information employed in this study regarding runoff has been obtained from LfU and DWD as time series of mean daily flows from midnight to midnight $q_i^{t(d)}\,[\mathrm{m^3 s^{-1}}]$ for each station i (see Figure 2.5) during the period $t \in [01.11.1960, \ldots, 31.10.1993]$. This data considers leap years.

2.8.1 Specific Runoff [mm]

The annual specific discharge [mm] for the spatial unit Ω_i cumulated at the end of the water year t is computed as follows

$$Q_{i1}^{t} = \frac{1}{A_i} \int_{d=1}^{d_e} q_i^{t(d)} dt = 86.4 \frac{1}{A_i} \sum_{d=1}^{d_e} q_i^{t(d)}, \tag{2.74}$$

where

$$d_e = \begin{cases} 365 & \text{if } t \text{ is a normal year} \\ 366 & \text{if } t \text{ is a leap year} \end{cases}. \tag{2.75}$$

The cumulated specific runoff in the winter season is then

$$Q_{i2}^{t} = 86.4 \frac{1}{A_i} \sum_{d=1}^{d_w} q_i^{t(d)}, \tag{2.76}$$

where

$$d_w = \begin{cases} 181 & \text{if } t \text{ is a normal year} \\ 182 & \text{if } t \text{ is a leap year} \end{cases}. \tag{2.77}$$

Finally, the cumulated specific runoff during the summer season is

$$Q_{i3}^{t} = Q_{i1}^{t} - Q_{i2}^{t}. \tag{2.78}$$

2.8.2 Characteristics of High Flows

In water management and planning it is always a very pertinent task to estimate the magnitude and the period of return (i.e. the inverse of the frequency of recurrence) of peak flows for a given location. Such characteristics of the runoff will govern the design of physical infrastructure along or across a stream (e.g. bridges, dams, river ports, flood walls, culverts) (Chow, 1964) on the one hand, and will guide planners to delimit either floodplains for protection or flooding prone areas where certain land uses must be avoided, on the other hand. Additionally, for other design purposes (e.g. embankments) it will be necessary to know the total duration that a stream's flow persists at flood stage.

For the reasons mentioned above, it is very important to determine whether these characteristics of the runoff would endure alteration in the future due to continuous changes of the land cover upstream of a place of interest. Because of that in this study the following characteristics of the high flows will be analysed.

Specific peak discharge [mm]

The specific peak discharges for a spatial unit Ω_i occurring during both the winter and summer seasons of a water year t expressed in [mm] are

$$Q_{i4}^{t} = 86.4 \frac{1}{A_i} \max\left(q_i^{t(d)}\right) \quad d = 1, \ldots, d_w, \tag{2.79}$$

and

$$Q_{i5}^{t} = 86.4 \frac{1}{A_i} \max\left(q_i^{t(d)}\right) \quad d = d_w + 1, \ldots, d_e, \tag{2.80}$$

respectively. A graphical representation of a peak flow can be seen in Figure 2.26.

Specific volume of the annual peak event [mm]

Runoff is composed of surface runoff and baseflow. The latter comprises all those long term sources entering the streams such as groundwater, whereas the former is the proportion contributed by precipitation. The estimation of the baseflow for each basin is not attempted here because it accounts for less that 10% of total flow in peaks (Laenen 1980). Moreover, if the basin is largely covered by impervious surfaces, this percentage is even lower, and it is often neglected (Bedient and Huber, 1992). Hence, the specific volume of the annual peak event occurring a day d_P can be approximated by

$$Q_{i6}^{t} = 86.4\frac{1}{A_i}\sum_{d=d_P-\Delta_0}^{d_P+\Delta_1} q_i^{t(d)} \quad \wedge \quad q_i^{t(d_P)} = \max\left(q_i^{t(d)}\right) \quad d \in [1, d_e]. \tag{2.81}$$

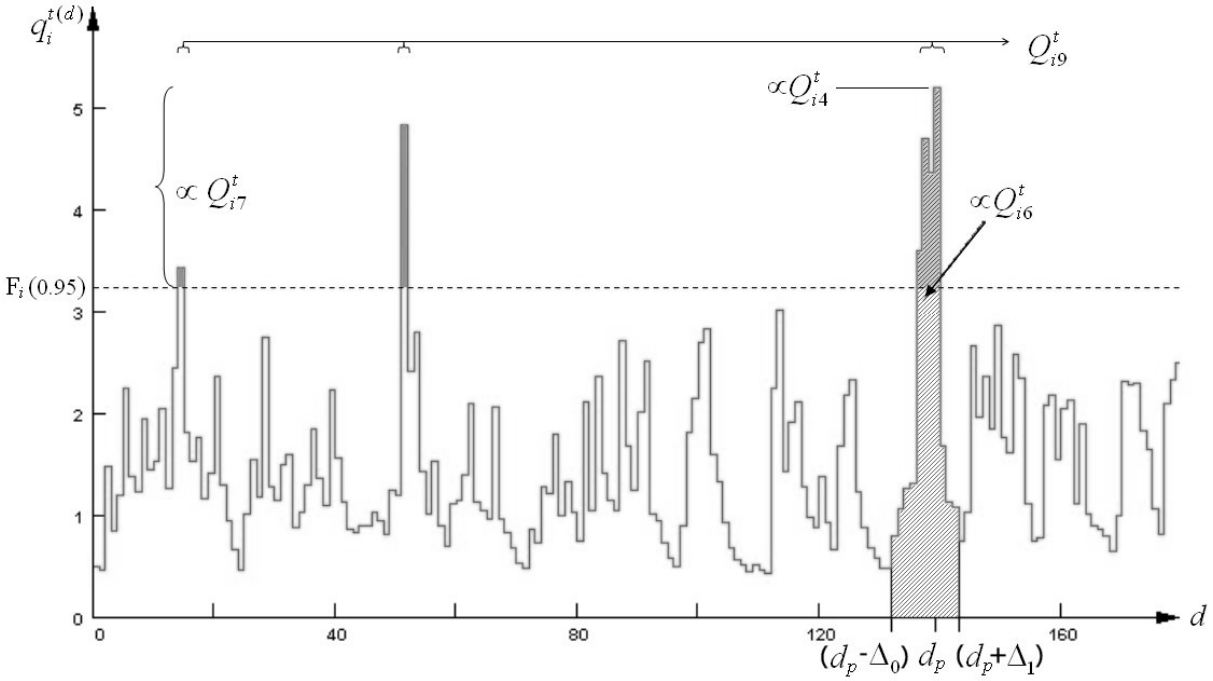

Figure 2.26 Schematic representation of a hydrograph depicting a peak flow occurring in winter, the specific volume of this event, the seasonal specific volume of all high flows, and their total duration in winter for a basin *i* during the water year *t*.

As it is written in (2.81), the volume of a peak event is defined by taking into account its duration. Hence, in order to have a reliable estimate of this volume, it would be very important to define the duration of the peak event (i.e. $\Delta_0 + \Delta_1$), or in other words, its beginning and its end. There are a number of possibilities with varying degrees of complexity to estimate them. The simplest one has been used by Potter (1991). It consists in defining $\Delta_0 = \Delta_1 = 1$. The main drawback of this method is that the catchment's size is not considered and, thus, the flow concentration time is underestimated. More sophisticated techniques would involve the estimation of the baseflow using the instant hydrograph (Dawdy et al. 1972, Dingman 1994, McCuen, 1998). Since the available data is the mean daily discharge this method can not be applied. A compromise between these possibilities would be to take the beginning of the peak event as the first rise of stream flow calculated as follows

$$\Delta_0 = \max\big(t(d_P) - t(d)\big) \quad \forall d \in \left\{ \frac{d^2}{dt^2} q_i^{t(d)} > 0 \ \wedge \ t(d) < t(d_P) \right\}, \tag{2.82}$$

where the second derivative can be estimated by finite differences as

$$\frac{d^2}{dx^2} q_i^{t(d)} \cong \frac{q_i^{t(d-1)} - 2q_i^{t(d)} + q_i^{t(d+1)}}{\Delta d^2} = q_i^{t(d-1)} - 2q_i^{t(d)} + q_i^{t(d+1)}. \tag{2.83}$$

According to Chow (1964), the end of the event can be estimated approximately by

$$\Delta_1 = 0.8267\, A_i^{0.2}, \tag{2.84}$$

where the A_i is the area of the basin in [km²].

Seasonal specific volume of high flows [mm]

Given the original set $\{q_i^{1960(1)}, q_i^{1960(2)}, \ldots, q_i^{1993(364)}, q_i^{1993(365)}\}$ of mean daily discharge for the spatial unit i, the order statistic of this sample would be the same figures but sorted in ascending order, i.e. $\{q_i^{(1)}, q_i^{(2)}, \ldots, q_i^{(\mathrm{O}_i)}\}$, where O_i is the number of valid observations available for basin i different from zero. Here, the j^{th} smallest of the O_i data values is denoted by $q_i^{(j)}$.

In this study, high flows are by definition those values equal to or higher than the 95th percentile. Based on the ordered sample, this statistic is estimated as follows

$$\mathrm{F}_i(0.95) = q_i^{(\iota)}, \tag{2.85}$$

where ι is an integer rounding of the product $\frac{95}{100}\mathrm{O}_i$, and $\mathrm{F}_i(\bullet)$ is the distribution function of daily mean discharge for catchment i.

Then, the winter and summer specific volumes of high flows can be calculated by

$$Q_{i7}^t = 86.4 \frac{1}{A_i} \sum_{d=1}^{d_w} \mu_{95_i}^{t(d)} q_i^{t(d)}, \tag{2.86}$$

and

$$Q_{i8}^t = 86.4 \frac{1}{A_i} \sum_{d=d_w+1}^{d_e} \mu_{95_i}^{t(d)} q_i^{t(d)}, \tag{2.87}$$

respectively. Where

$$\mu_{95_i}^{t(d)} = \begin{cases} 1 & \text{if} \quad q_i^{t(d)} \geq \mathrm{F}_i(0.95) \\ 0 & \text{otherwise} \end{cases}. \tag{2.88}$$

The graphical representation of these indicators is depicted in Figure 2.26.

Total duration of high flows [day]

This indicator, as shown in Figure 2.26, accounts for the total number of days in a given season, either summer or winter, that have a daily mean discharge greater than or equal to a given threshold, for instance the 95th percentile. Using the membership function shown in (2.88), this indicator can be calculated for winter and summer as

$$Q_{i9}^{t} = \sum_{d=1}^{d_w} \mu 95_i^{t(d)} , \tag{2.89}$$

and

$$Q_{i10}^{t} = \sum_{d=d_w+1}^{d_e} \mu 95_i^{t(d)} , \tag{2.90}$$

respectively.

Frequency of recurrence of high flows [year^{-1}]

This indicator quantifies the total number of high flow events $\mathcal{H}_i^t$ that occur in the basin Ω_i during the period from $t(d_0)$ to $t(d_1)$, where d_0 and d_1 represent the beginning and the end of a season in the water year t. The k high flow event $\mathcal{H}_i^{t(k)}$ can be defined as a set given by

$$\mathcal{H}_i^{t(k)} = \left\{ q_i^{t(d)} \middle| \ \forall\, q_i^{t(d)} \geq \mathrm{F}_i(0.95) \ \wedge \ t(d_0) \leq d = t(d_0^k), \ldots, t(d_1^k) \leq t(d_1) \right\}, \tag{2.91}$$

where $t(d_0^k)$, and, $t(d_1^k)$ are the beginning and the end of the high flow event k during the year t respectively. Based on this definition, the frequency of high flows occurring during winter and summer can be estimated as the cardinality of the following sets

$$Q_{i11}^{t} = \left| \left\{ \mathcal{H}_i^{t(k)} \middle| \ \ d_0 = 1 \ \wedge \ d_1 = d_w \right\} \right|, \tag{2.92}$$

and

$$Q_{i12}^{t} = \left| \left\{ \mathcal{H}_i^{t(k)} \middle| \ \ d_0 = d_{w+1} \ \wedge \ d_1 = d_e \right\} \right|. \tag{2.93}$$

In the previous equations d_w and d_e are determined according to (2.77) and (2.75) respectively.

2.8.3 Characteristics of Low Flows

The *minimum* low flow that may occur in most small basins and some larger ones is obviously zero. Hence, in these cases the frequency of occurrence of zero flow would be a suitable drought indicator. This is, however, not the case in the present study. Often, the disastrous effects associated with low flows (e.g. degradation of water quality, reduction of the amount of water available for supply) are experienced long before a flow of zero is reached, hence a more useful index should be defined. Such an index should consider the frequency and duration of spells of low flows or droughts as represented in Figure 2.27. Many low flow definitions can be found in the literature, for instance in Pirt (1983), Correira et al. (1987), Demuth and Külls (1997), and Moore (1997). In this study, however, low flows are those discharge observations corresponding to a basin i that are smaller than or equal to a threshold value, for example, the 10$^{\text{th}}$ percentile of the daily mean discharge $\mathrm{F}_i(0.10)$ [see (2.85)].

Based on this definition, the water *deficit* $D_i^{t(d)} \left[\mathrm{m^3 s^{-1}}\right]$ occurring a day d in the water year t is defined by (Correira et al. in Duckstein and Plate 1987) as

$$D_i^{t(d)} = \begin{cases} \mathrm{F}_i(0.10) - q_i^{t(d)} & \text{if} \quad q_i^{t(d)} \leq \mathrm{F}_i(0.10) \\ 0 & \text{otherwise} \end{cases}. \tag{2.94}$$

Then, let a low flow spell k occurring during the year t within the basin i, say $\xi_i^{t(k)}$ be represented by the set

$$\xi_i^{t(k)} = \left\{ D_i^{t(d)} \middle| \ \forall \ D_i^{t(d)} \neq 0 \ \wedge \ d = t(d_0^k),\ldots,t(d_1^k) \right\}, \tag{2.95}$$

where $t(d_0^k)$, and, $t(d_1^k)$ are the beginning and the end of the low flow spell k during the year t respectively. Based on this definition, the *duration* [day] of such low flow event is

$$\Lambda_i^{t(k)} = t(d_1^k) - t(d_0^k) + 1 \qquad \forall k = 1,\ldots,\mathrm{K}_i^t, \tag{2.96}$$

where K_i^t is the number of low flow spells or droughts occurring in basin i during the water year t.

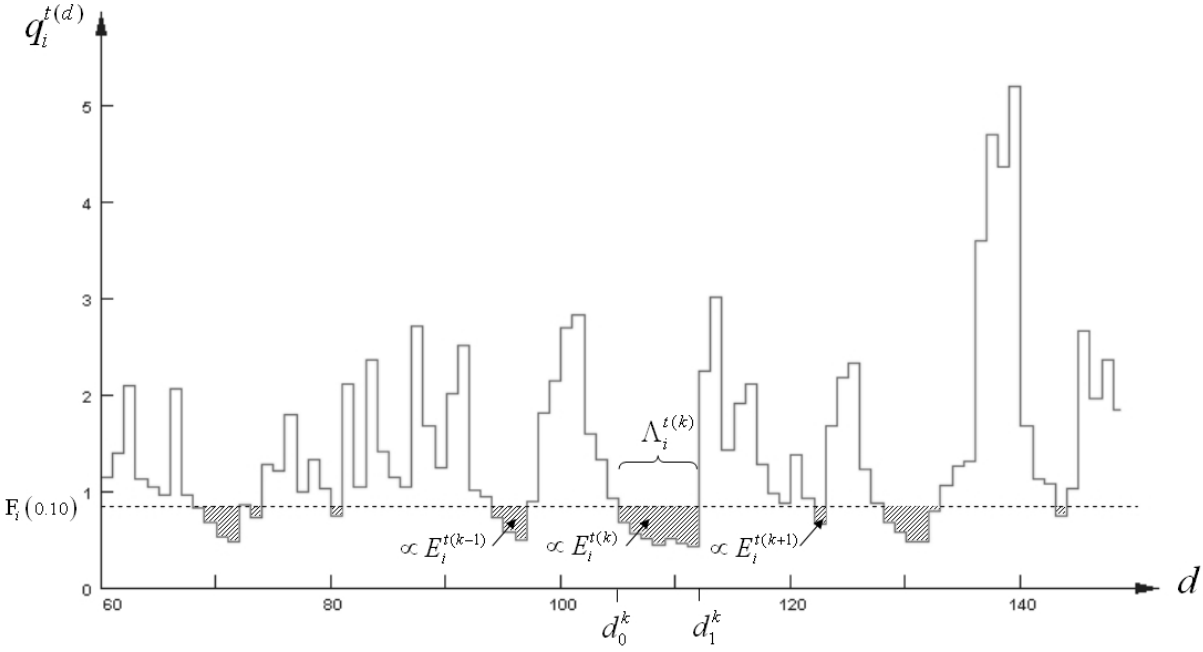

Figure 2.27 Schematic hydrograph depicting the occurrence of low flow spells, their respective deficit, and the duration of the k^{th} low flow spell of a given basin *i* during the water year *t*.

Furthermore, *total specific deficit of a drought* k [mm] can be defined as

$$E_i^{t(k)} = 86.4 \frac{1}{A_i} \sum_{d=d_o^k}^{d_1^k} D_i^{t(d)} \,. \tag{2.97}$$

Finally, the *intensity* of a drought k expressed in [mm/day] can be calculated as the ratio of its total deficit to its duration, namely

$$\Xi_i^{t(k)} = \frac{E_i^{t(k)}}{\Lambda_i^{t(k)}} \,. \tag{2.98}$$

Maximum drought duration [day]

The maximum drought duration occurring at the basin i during the year t is estimated using (2.96) as follows

$$Q_{i13}^t = \max\left(\Lambda_i^{t(k)}\right) \quad \forall k = 1,\ldots,\mathrm{K}_i^t \,. \tag{2.99}$$

Total drought duration [day]

The total number of days that a given basin i in a given water year t has endured a low flow regime can be calculated by

$$Q_{i14}^{t} = \sum_{k=1}^{\mathrm{K}_i^t} \Lambda_i^{t(k)} . \tag{2.100}$$

Maximum drought intensity [mm/year]

The maximum drought intensity referred to yearly basis based on (2.98) is given by

$$Q_{i15}^{t} = 31536 \left(\max \left(\Xi_i^{t(k)} \right) \right) \quad \forall k = 1, \ldots, \mathrm{K}_i^t . \tag{2.101}$$

Cumulative specific deficit [mm]

The total water deficit endured by the basin i during a water year t is

$$Q_{i16}^{t} = \sum_{k=1}^{\mathrm{K}_i^t} E_i^{t(k)} . \tag{2.102}$$