Escola de Camins
Escola Tècnica Superior d'Enginyeria de Camins, Canals i Ports
UPC BARCELONATECH

TREBALL FINAL DE MÀSTER

Research about required temporal resolution of hydrological input data in hydrological simulation

Treball realitzat per:
Xavier Sellart Aldomà

Dirigit per:
Daniel Sempere Torres
Minjiao Lu

Màster en:
Enginyeria de Camins, Canals i Ports

Barcelona, 15 de juny de 2018

Departament d'Enginyeria Civil i Ambiental

NAGAOKA UNIVERSITY OF TECHNOLOGY

Abstract

Laboratory of Hydrometeorology
Department of Civil and Environmental Engineering

Master in Civil Engineering

Research about required temporal resolution of hydrological input data in hydrological simulation

by Xavier Sellart Aldomà

Temporal resolution of rainfall plays an important role in determining the hydrological response of river basins. Rainfall temporal variability can be considered as one of the most critical elements when dealing with input data of rainfall–runoff models and, therefore, its influence should be quantified. In this study, the extent to which the performance and outputs of a hydrological model are affected in basins of different sizes when using input data with different temporal resolutions is investigated. The set of study basins used consists of 5 catchments from the United States. Using rainfall data with hourly resolution obtained from the MOPEX datasets, several synthetic sets of data with different temporal resolutions were created and used as inputs of the Xinanjiang model to check the quality of the simulation in basins of different sizes using input data with low temporal resolution. The analysis of the simulation results showed significant differences in model performance at different time resolutions. On average, the temporal resolution of input data has a significantly higher impact on model performance during shorter and high-intensity storm events than during long periods with low precipitation. The results also showed that relation between the value of the performance indicators and the temporal resolution of the input data was strongly influenced by size of the study basin, and it was possible to quantify this influence. Finally, with all the knowledge obtained, a method to improve the quality of the simulation was sought, but the improvement reached was not very significant.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The foundation of hydrology is the understanding of the hydrological cycle, its processes and its influence on the environment (Gayathri, Ganasri, and Dwarakish, 2015). During the past few decades, due to the increase in computational power of the computers, mathematical models have had an increasing development in all the different branches of knowledge, including hydrology.

Hydrological models are a simplified version of a real complex system, which is represented in a mathematical form through an analytical expression (Wheater, Sorooshian, and Sharma, 2007). In a hydrological model, the real physical system that is usually represented is a hydrographic basin and each of the components of the hydrological cycle. A mathematical model can be helpful to make decisions regarding hydrology, so it is necessary to have knowledge of the inputs to the system and the outputs from the system to verify if the model is making a good representation of the real hydrographic basin.

Nowadays, hydrological models have become an important tool for managing water resources (Anand et al., 2018). The range of uses they have is really wide and they can be used as a prediction tool to predict flooding events and its magnitude and duration, to assess the impacts that climate change will have on the water cycle and to manage water resources. They can also be used in research as a support tool for helping in the understanding of the physical processes take part in the hydrological cycle. The computation of the stream-flow using precipitation data is complicated because it usually involves a large number of processes with various dynamics and characteristic time scales (Sun et al., 2015).

However, no matter how good a hydrological model can be, to be able to make accurate predictions it requires good parametrization and high quality hydrological data (Fredrik et al., 2011). The proper description and simulation of the hydrological processes may require data with short time steps and low observational errors. Each single hydrological model is unique and the accuracy of the results can vary greatly from model to model due to differences in the features of the model. Also, a single model, can produce several different outputs with different accuracies due to variations in the parameters used and in the quality of the input data (Nourani, Roughani, and Gebremichael, 2011). This generates the question of which is the most appropriate time step of the data that is used as input to hydrological models, which typically is precipitation.

The quality of the input data depends on the precision of the measurements and on their temporal resolution, which is the time interval at which those measurements are made. A high temporal resolution means that the time interval of the measurements is very small, and it is usually considered that data with a high temporal resolution is better than data with low temporal resolution because it contains more

information and therefore should contribute to better modeling of the rainfall–runoff relationship (Wilk and Hughes, 2002). In an ideal case, the hydrological data would not have observational errors and the temporal resolution of data set would be infinite. This is supported by studies that showed that runoff generation is highly affected by sub-hourly dynamics of precipitation, particularly where the infiltration-excess overland flow mechanism dominates the rainfall–runoff response (Kandel, Western, and Grayson, 2005). However, in most cases it is no possible to obtain data with sub-hourly temporal resolution, and temporal resolution of rainfall data usually available for practical applications is often lower than that required for the rainfall–runoff simulation, greatly compromising model accuracy (Aronica, Freni, and Oliveri, 2005). Therefore, it is necessary to find ways to deal with the uncertainty caused by the quality of the data set to be able assess the quality of the output data of the hydrological models.

1.2 Problem statement

Temporal resolution of rainfall plays an important role in determining the hydrological response of river basins. Rainfall temporal variability can be considered as one of the most critical elements when dealing with input data of rainfall–runoff models. Even though its importance, effects of the input data time resolution in rainfall–streamflow modeling is often overlooked. Different studies have been made about the effects of temporal resolution of the input data on the performance of hydrological models but none of them consider the relationship between the quality of the output and the size of the water basin to check the scale effect.

In some cases of temporal resolution of model input data and its impact on prediction of river discharge has been studied for only a single basin. For example, Wang, He, and Takase, 2009 did research about the effects of different time resolution on long and short-term runoff prediction in a small catchment of 21 ha in the mountainous part of Ozu City, Ehime prefecture, Japan. Their results demonstrated that high-resolution rainfall data is crucial for modeling runoff, especially for short-term runoff analysis. Bauwe et al., 2017 also carried out a study to evaluate the influence the temporal resolution of input data has on the simulated stream-flow using the SWAT model, but it was done using 5, 15, 30, and 60-min precipitation time steps only in one basin in northeastern Germany, near the city of Rostock. They concluded that the relevance of the precipitation time step depends on the catchment characteristics.

In the study carried out by Ficchì, Perrin, and Andréassian, 2016, they investigated the extent to which the performance of hydrological modeling can be improved by short time-step data. They selected a large set of 240 French catchments, for which 2400 flood events were selected and the simulations were carried out using GR4 rainfall–runoff model (Perrin, Michel, and Andréassian, 2003). The results confirmed that working on different catchments and time scales is necessary, since different model performance dependencies on time step were found over different catchments and across the range of reference time scales. However, they did not check the relationship between these dependencies and the size of the studied basins.

The literature review has shown that, despite the general knowledge of the importance of sub-daily variability of rainfall for flood volume modeling, the advantages of using rainfall data at fine temporal resolution for flow simulation are still

not well quantified. There is a need for further investigations to evaluate the usefulness of fine time-step information for hydrological model simulations, by comparing different model time-step outputs at common aggregated time scales, using a set of several different catchments.

1.3 Objective and content

The main objective of this thesis is to use the Xinanjiang model to do research works to check the scale effect of basin size in the temporal resolution of the hydrological input data of the model. This will be useful when doing simulations in basins where the available data has low temporal resolution because it will allow to assess the quality of the result.

In order to do so, several hydrological basins of different size will be simulated using different temporal resolution. The idea is to see how the performance indicators vary when changing the temporal resolution, and if this variation has some correlation with the basin size be able to quantify it. The variation of all the outputs of the XAJ model will also be analyzed to try to find out which parts of the model are affected the most by the temporal resolution of the input data to see if it is possible to improve the model and make it less sensitive to the temporal resolution of input data.

The structure of this thesis goes as follows:

- In Chapter 1 there is a introduction to the study carried out and its main objectives are explained.
- In Chapter 2 the main structure of the Xinanjiang model is described. The equations and the physical meaning of the parameters used by the model are discussed as well.
- In Chapter 3 the location, the statistics of precipitation, potential evaporation and runoff of the basins selected for the study are shown. The requirements of the characteristic of the basins and where the data series were obtained are also explained.
- In Chapter 4 the procedure followed for the calibration and validation of the 15 parameters of the Xinanjiang model in five different basins is described. There is also a description of the performance indicators used in this study and the values obtained for each one of them. Finally, the validation hydrographs for each one of the basins of the study are presented.
- In Chapter 5 the behavior of the Nash efficiency when using input data with different time resolution is analyzed under several different cases. Conclusions about the influence that the size of the study basin has on the required temporal resolution of the input are presented.
- In Chapter 6 the effects of the temporal resolution of the input data on the different outputs of the model are presented. There is also a discussion on how to improve the quality of the simulation by modifying some of the parameters obtained in the calibration.
- In Chapter 7 the main conclusions of this study are summarized.

Chapter 2

The Xinanjiang model

2.1 Introduction

The Xinanjiang (XAJ) model (Zhao, 1992) is a conceptual hydrological rainfall-runoff model to simulate runoff generation and concentration in a basin from a humid or semi-humid region. It is the most popular rainfall-runoff model in China and also widely used all over the world. It is mainly used for flood forecasting, but it can also be used for other purposes such as water resources estimation, flood design and field drainage, and water quality accounting.

The main feature of the model is the concept of runoff formation on repletion of storage, which means that runoff is not produced until the soil moisture content of the aeration zone reaches the field capacity, and thereafter runoff equals the rainfall excess without further loss. The runoff generated is separated into three different components: surface runoff, inter-flow and groundwater. Finally, the local runoff is transferred to the outflow in the outlet of the basin. A flowchart of the Xinanjiang model can be seen in Figure 2.1.

The Xinanjiang model is divided in the following main parts:

- The adjustment in data input
- The evaporation
- The runoff production
- The runoff separation
- The flow routing

In the Xinanjiang model, big catchments are divided into smaller sub-basins by Thiessen polygons, the outflow hydrograph from each of which is first simulated and then routed down the channels to the main basin outlet using the Muskingum method. However, since in this study only small catchments have been simulated and it is considered that the hydrological characteristics of the catchment are homogeneous. That is why the flow routing features of the model were not been used in this study.

2.2 Structure of the model

In this section of the study the different parts of the structure of the model used in this study will be explained. The flow routing part of the XAJ model was not used in this study, therefore, it will not be explained.

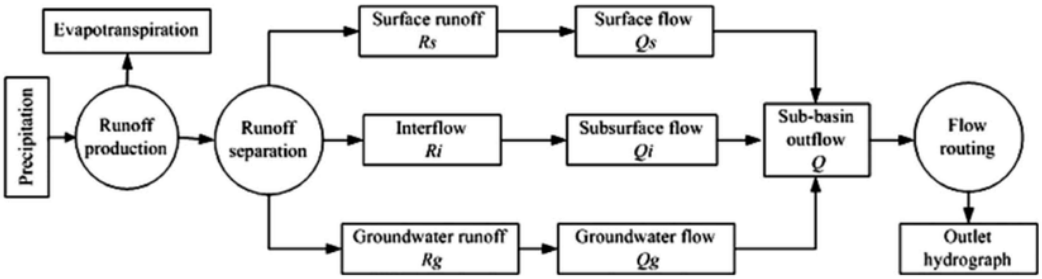


FIGURE 2.1: Flowchart of the Xinanjiang model for a single subbasin (Yao et al., 2009)

2.2.1 Adjustment in data input

The inputs of the XAJ model are the areal mean rainfall depth on the basin and the potential evapotranspiration data. The potential evaporation, E_p, as a direct measurement is difficult to get and that is why it is calculated from pan evaporation using the following equation:

$$E_p = C_{ep} \cdot E_{pan} \tag{2.1}$$

where E_{pan} is the pan evaporation observed at many hydrological stations, and C_{ep} is an empirical adjustment coefficient that represents the ratio of potential evapotranspiration to pan evaporation.

As it was explained before, in this study the XAJ model was be used in a pure lumped way because of the size of the catchments. The rainfall that is inputted into the model represents the mean areal rainfall in the whole catchment. To consider the possible errors caused by gauge losses, altitudinal distribution of rainfall, and the representativeness of the gauge network, an adjustment coefficient, C_p, which is the ratio of measured precipitation to actual precipitation, is introduced to derive adjusted rainfall, P, from areal average measured rainfall, P_a, using the following equation:

$$P = C_p \cdot P_a \tag{2.2}$$

2.2.2 Evapotranspiration

Evapotranspiration is related to potential evapotranspiration through a three-layer soil moisture model. The soil in a catchment is considered to have three layers, called upper layer, lower layer and deeper layer. Therefore, the total areal mean soil moisture capacity, WM, can be divided into upper, lower and deeper areal mean soil moisture capacity: WUM, WLM and WDM respectively. Until the storage WUM of the uppermost layer is exhausted, evaporation occurs at the potential rate equal to the potential evaporation rate:

$$EU = E_p \tag{2.3}$$

On exhaustion of the upper layer, with capacity WUM, any remaining potential evapotranspiration is applied to the lower layer, but with a modified efficiency. This modification consists in a multiplication of the remaining potential evapotranspiration by the ratio of the actual storage, WL, to the capacity storage WLM of the lower layer.

$$EL = (E_p - EU) \cdot \frac{WL}{WLM} \tag{2.4}$$

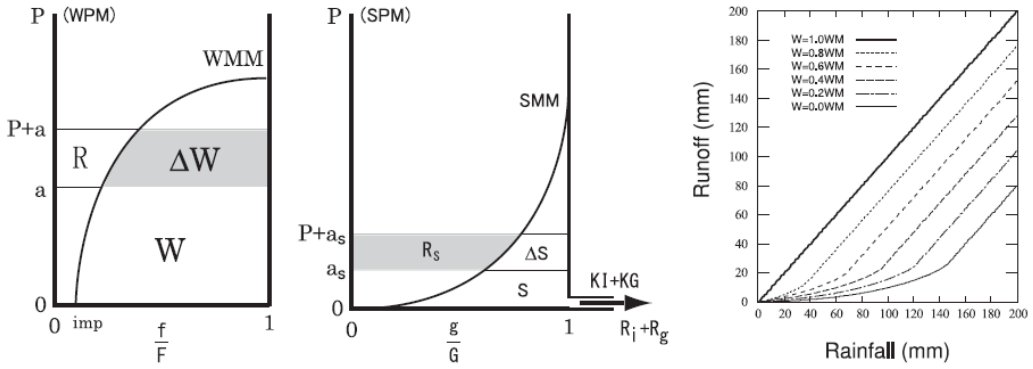


FIGURE 2.2: Schematics of the XAJ model (left: distribution of tension water over the catchment; center: distribution of free water; and right: rainfall-runoff relationship based on the tension water distribution) (Lu and Li, 2014)

When the lower layer actual storage is reduced to a specified proportion C, coefficient of deep evapotranspiration, of WLM, evapotranspiration is assumed to continue, but at a further reduced rate ED given by the following equation:

$$ED = (E_p - EU) \cdot C \tag{2.5}$$

Therefore the actual evapotranspiration $E = EU + EL + ED$, and the total storage of soil moisture W will be updated after calculation of evapotranspiration as $W = L + WU + WD$ to be used as initial storage in next step.

2.2.3 Runoff production

As it has been said in the introduction of this chapter, the main feature of the XAJ model is the concept of runoff formation on repletion of storage. Runoff production at a point, occurs only on repletion of the tension water storage at that point. To provide for a non-uniform distribution of tension water capacity throughout the basin, a tension water capacity curve, left graph in Figure 2.2 is introduced.

In Figure 2.2, $\frac{f}{F}$ represents the proportion of the pervious area of the basin whose tension water capacity is less than or equal to the value of the ordinate WPM. The tension water capacity at a point WPM, varies from zero to a maximum WMM (which is a parameter) according to the following a pure statistical distribution represented by a single parabolic curve:

$$1 - \frac{f}{F} = (1 - imp)\left(1 - \frac{WPM}{WMM}\right)^b \tag{2.6}$$

where the exponent of the tension water capacity curve b is a shape parameter, imp is the ratio of impervious area to the total area of the basin, WMM is the maximum possible value of WPM and $\frac{f}{F}$ is the proportion of the previous area of the basin whose tension water capacity is less than or equal to WPM. The areal mean tension water capacity constitutes an alternative parameter to the maximum value WMM. These are related through the shape parameter b. Integrating Equation 2.6 it can be proved that:

$$WM = (1 - imp)\,\frac{WMM}{1 + b} \tag{2.7}$$

The state of the catchment, at any time, is assumed to be represented by a point x on the curved line of the left plot in Figure 2.2. The area to the right and below the point x is proportional to the areal mean tension water storage W (not capacity). This assumption implies that each point in the sub-basin is either at capacity tension (points to the left of x) or at a constant tension (points to the right of x).

When rainfall exceeds evaporation the ordinate of Figure 2.2 is increased by the excess, x moves upwards along the curve and runoff is generated on the area where the tension water capacities are filled:

$$R = \int_a^{a+P} \frac{f}{F} dWPM \tag{2.8}$$

and the increment of the areal mean tension water storage, ΔW, will be

$$\Delta W = P - R = \int_a^{a+P} \left(1 - \frac{f}{F}\right) dWPM \tag{2.9}$$

The relationship between rainfall and runoff is shown in the right-hand side graph of Figure 2.2. For all initial storages, the line slope becomes steeper when rainfall increases and finally reaches 45 degree.

2.2.4 Runoff separation

The total runoff R generated in a wet period in accordance with the middle plot in Figure 2.2 and calculated with Equation 2.8, must be separated into its three components, RS surface runoff, RG the ground water contribution, and RI a contribution to interflow.

The free water storage capacity SPM is assumed to be distributed between zero and a point maximum SMM in a parabolic manner (middle plot in Figure 2.2), over the area of the sub-basin which is currently producing runoff. The equation of the parabolic curve is shown below:

$$\left(\frac{g}{G}\right) = 1 - \left(1 - \frac{SPM}{SMM}\right)^{EX} \tag{2.10}$$

where g is that portion of the catchment area for which the free water storage capacity is less than or equal to SPM and the exponent of the free water capacity curve EX is a parameter. The current state of free water storage in the basin can be represented by a point (ordinate a_s in the middle plot in Figure 2.2), implying that the portion of the catchment to the left of that point is at capacity storage and to the right the storage is constant, below capacity level.

The areal mean free water storage can be obtained by integration of equation 2.10:

$$S = \int_0^a \left(1 - \frac{g}{G}\right) dSPM = SM(1 - (1 - \frac{a_s}{SMM})^{1+EX}) \tag{2.11}$$

where

$$SM = \frac{SMM}{1 + EX} \tag{2.12}$$

is the areal mean free water capacity.

Using this distribution, the total surface runoff generated can be derived as follows:

$$R_s = imp \cdot P + \left(\left(\frac{f}{F}\right)_R - imp\right) \int_{a_s}^{a_s+P} \frac{g}{G} dSPM \tag{2.13}$$

The remainder of R becomes an addition, AS, to the freewater storage S, which in turn contributes RI laterally to inflow and RG vertically to ground water, according to the relations:

$$RI = S \cdot KI \cdot \left(\left(\frac{f}{F} \right)_R - imp \right) \tag{2.14}$$

$$RG = S \cdot KG \cdot \left(\left(\frac{f}{F} \right)_R - imp \right) \tag{2.15}$$

Where KI and KG are parameters that represent the outflow coefficient of the free water storage to the interflow and the outflow coefficient of the free water storage to groundwater.

After the total runoff R is produced and separated into three components RS, RI, and RG, the three runoff components may be routed from the place where they produced to the outlet of the basin by using three linear reservoirs as follows:

$$Q_s(t+1) = c_s \cdot Q_s(t) + (1 - c_s) \cdot R_s \tag{2.16}$$

$$Q_i(t+1) = c_i \cdot Q_i(t) + (1 - c_i) \cdot R_i \tag{2.17}$$

$$Q_g(t+1) = c_g \cdot Q_g(t) + (1 - c_g) \cdot R_g \tag{2.18}$$

where cs, ci and cg are three recession coefficients for channel routing, lower interflow storage and groundwater storage respectively.

2.3 Summary

In this chapter, the main structure of the Xinanjiang model is described. The equations and the physical meaning of the parameters used by the model are discussed as well.

Chapter 3

Study basins

3.1 Introduction

Five different basins were selected for this study. The requirements for the basins were that they had to be natural basins in humid regions with annual rainfall higher than 500 mm and they had to have a size between 100 and 4000 km^2 approximately. It was also necessary that the record length of the data was at least 12 years long and that its temporal resolution was of at least 1 hour.

3.2 Selected study basins

All the data from the selected study basins was obtained from the U.S. MOPEX data set (Schaake, Cong, and Duan, 2006), which has long series of daily and hourly data. The data of the U.S. MOPEX data set contains the following information: the date, the mean areal precipitation (MAP) processed by the NWS Hydrology Laboratory, the climatic potential evaporation (EP) based on the NOAA Freewater Evaporation Atlas (Farnsworth, Thompson, and Peck, 1982), which is simulated by a sinusoid in one year, and the streamflow discharge obtained from the USGS National Water Information System (NWIS) (available at http://water.usgs.gov/nwis). It also provides information about the maximum and minimum air temperature, but they were not necessary for this study.

In Table 3.1 there is a summary of the characteristics of the basins selected for the study and in Figure 3.1 there is a map showing the location of the study basins. As it can be seen in Table 3.1, all basins comply with the requirements explained in the introduction of this chapter. If the aridity index, which is ratio of annual potential evaporation over annual precipitation, is lower than 0.9 then the the basin is considered to be in a humid region (Rahman and Lu, 2015).

TABLE 3.1: Characteristics of the MOPEX basins used in this study

MOPEX Basin ID	Location			Area (km^2)	Average Precipitation (mm/year)	Average Potential Evaporation (mm/year)	Total Daily Streamflow Discharge (mm/year)	Data Aridity Index
	Long.	Lat.	State					
02387500	-84.94	34.58	GA	4149	1287.7	973.5	491.25	0.756
02456500	-86.98	33.71	AL	2292	1381.8	973.0	555.01	0.704
03443000	-82.62	35.30	NC	767	1971.4	839.1	1203.46	0.426
03504000	-83.62	35.13	NC	134	2015.9	761.8	1421.24	0.378
11532500	-124.05	41.79	CA	1577	2576.0	750.0	1926.64	0.291

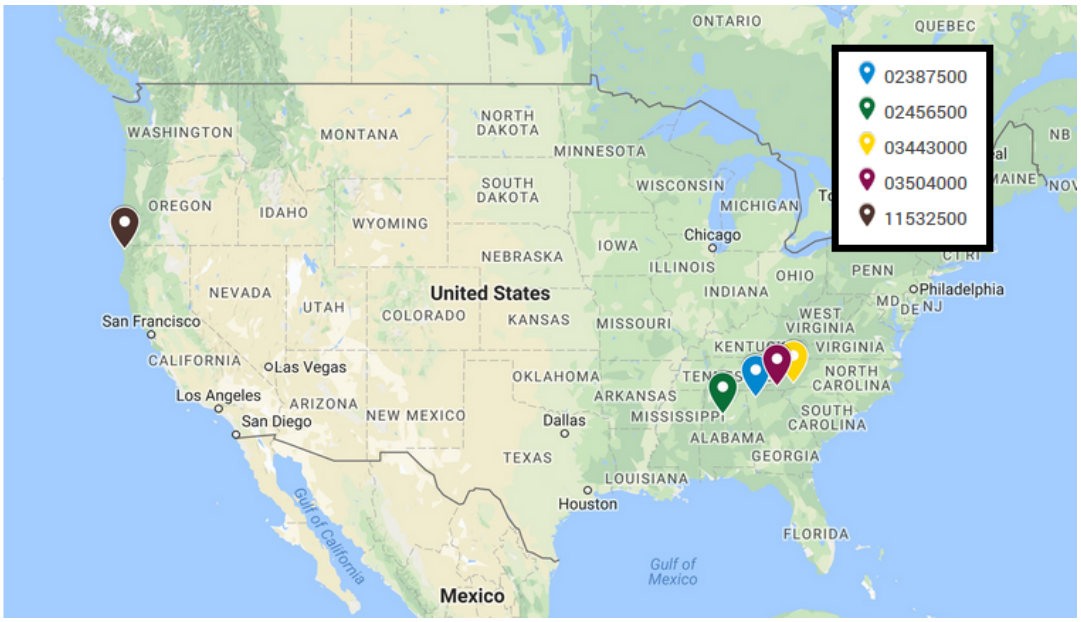
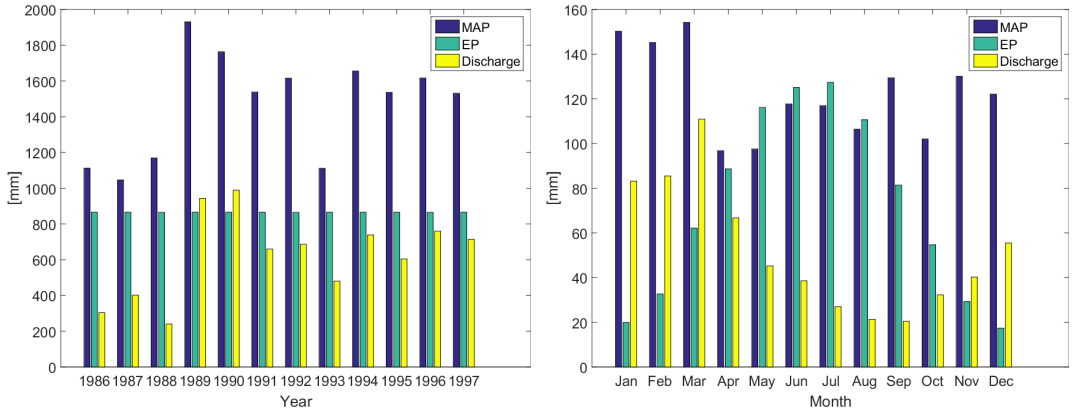


FIGURE 3.1: Map showing the location of the study basins

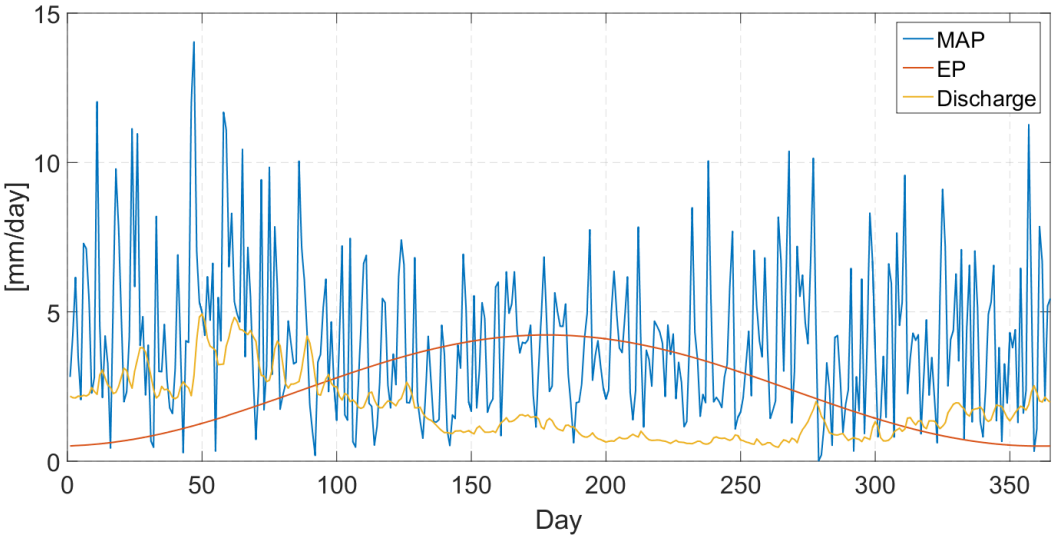
Finally, in Figures 3.2, 3.3, 3.4, 3.5 and 3.6, the annual, monthly and daily statistical values of rainfall, discharge and evaporation for each basin during 1986 to 1997, which will be the years used for calibration and validation, are illustrated.

3.3 Summary

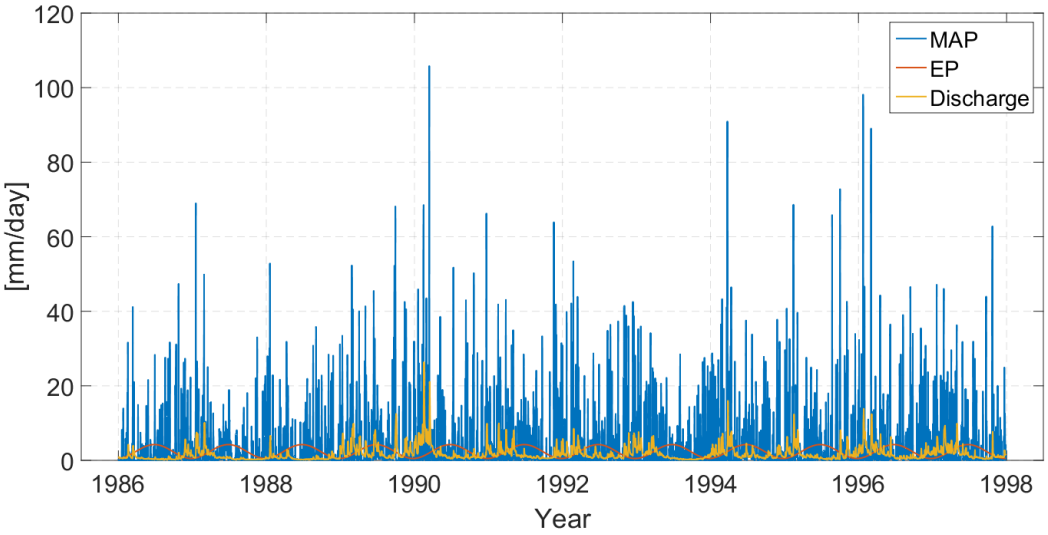
In this chapter, the location, the statistics of precipitation, potential evaporation and runoff of the basins selected for the study are shown. The requirements of the characteristic of the basins and where the data series were obtained are also explained.



(A) Annual statistics (B) Monthly statistics

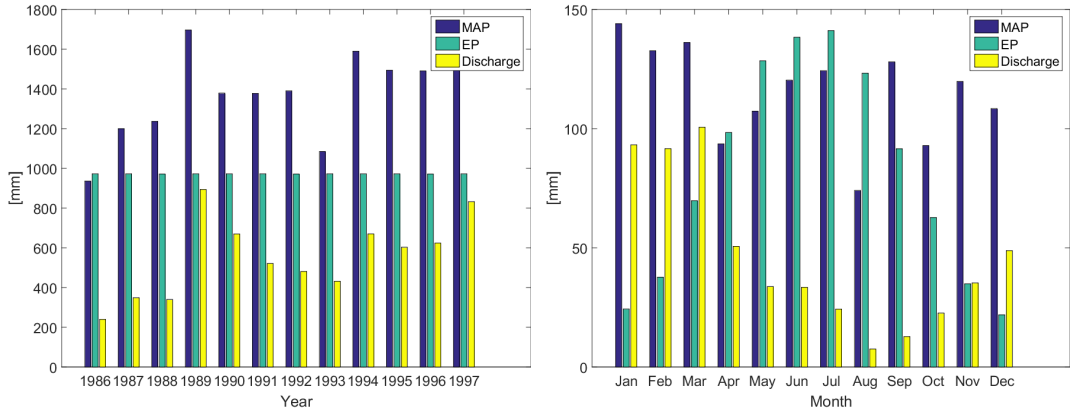


(C) Daily statistics

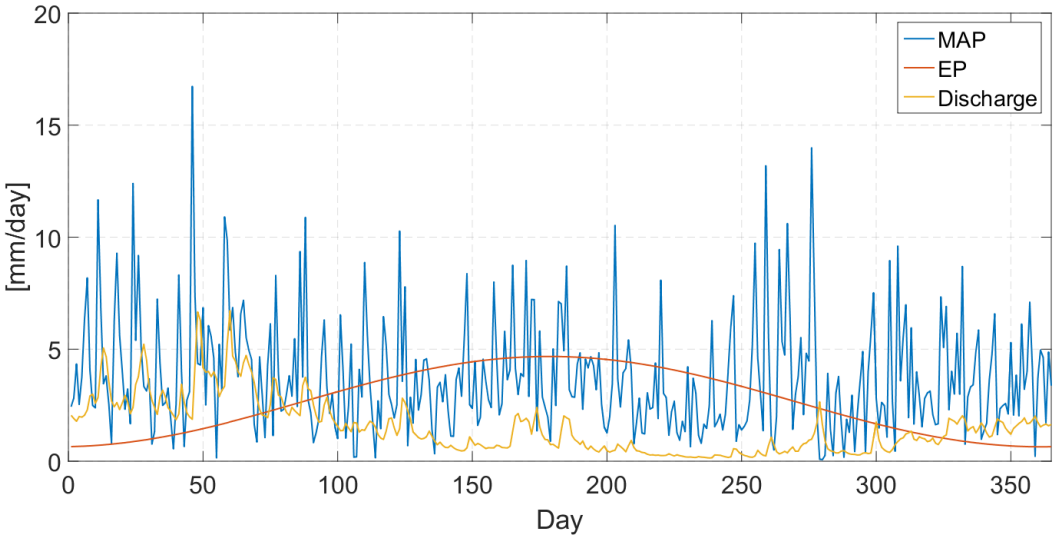


(D) Data record of 12 years (1986-1997)

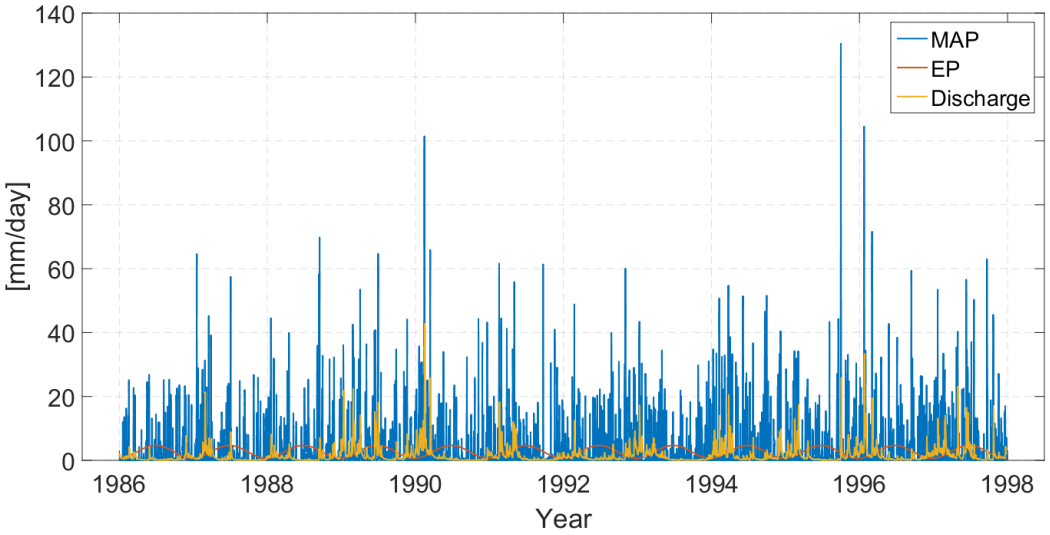
FIGURE 3.2: Values of rainfall, discharge and evaporation for basin 02387500



(A) Annual statistics (B) Monthly statistics

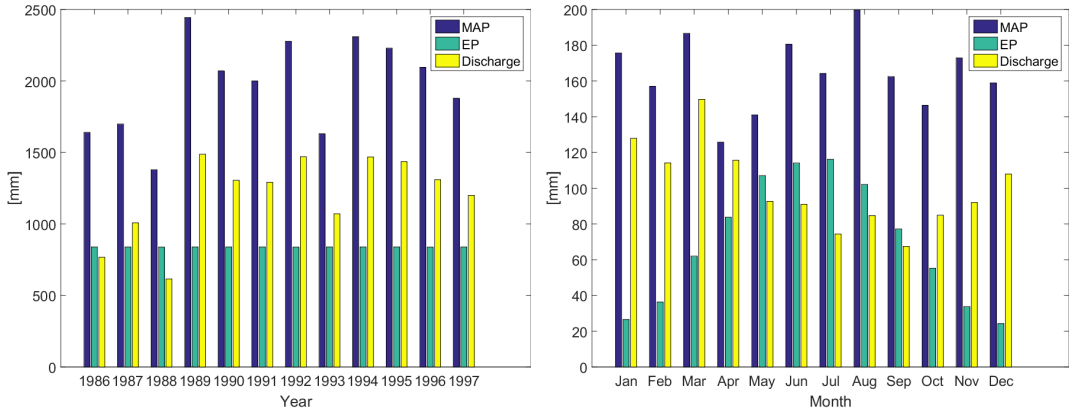


(C) Daily statistics

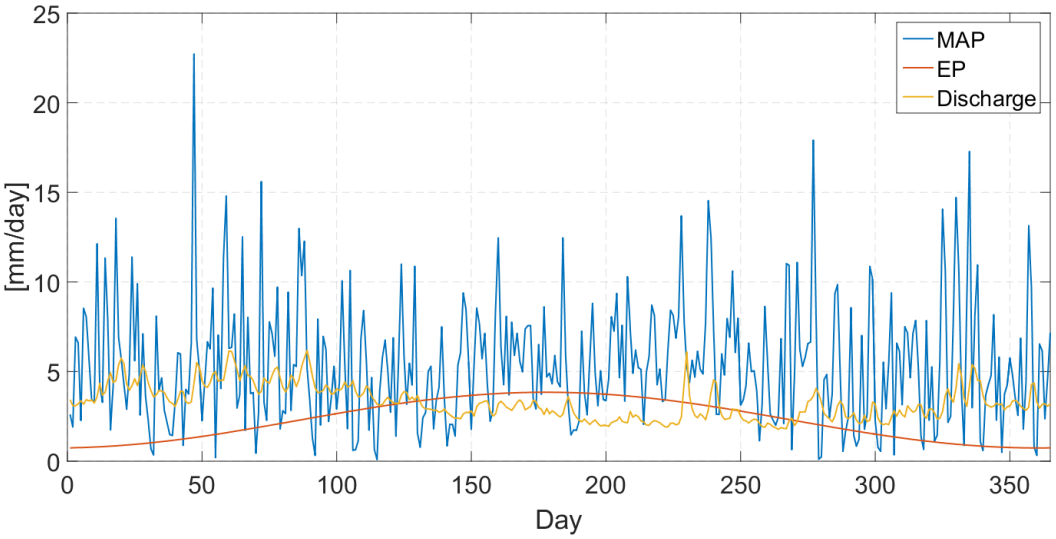


(D) Data record of 12 years (1986-1997)

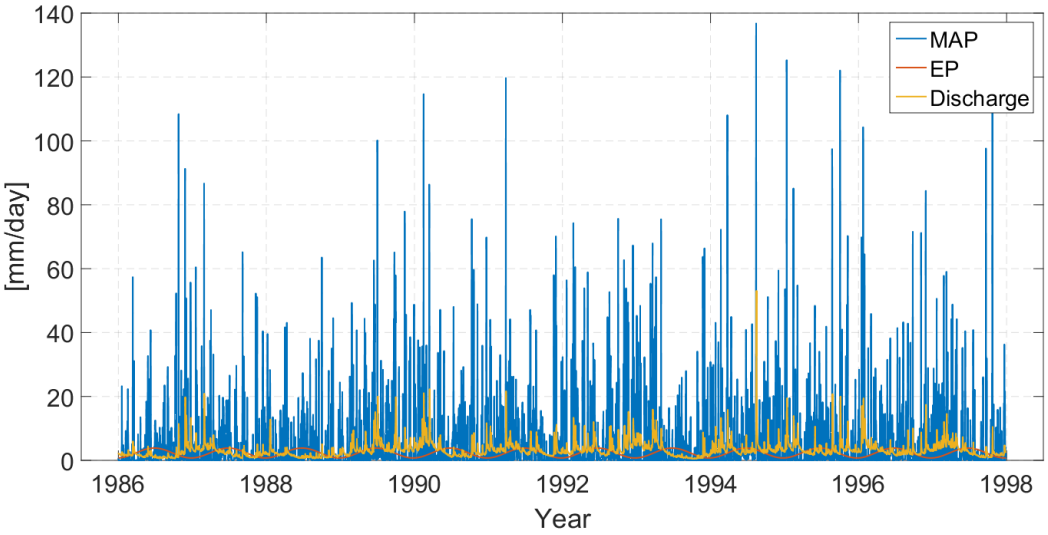
FIGURE 3.3: Values of rainfall, discharge and evaporation for basin 02456500



(A) Annual statistics (B) Monthly statistics

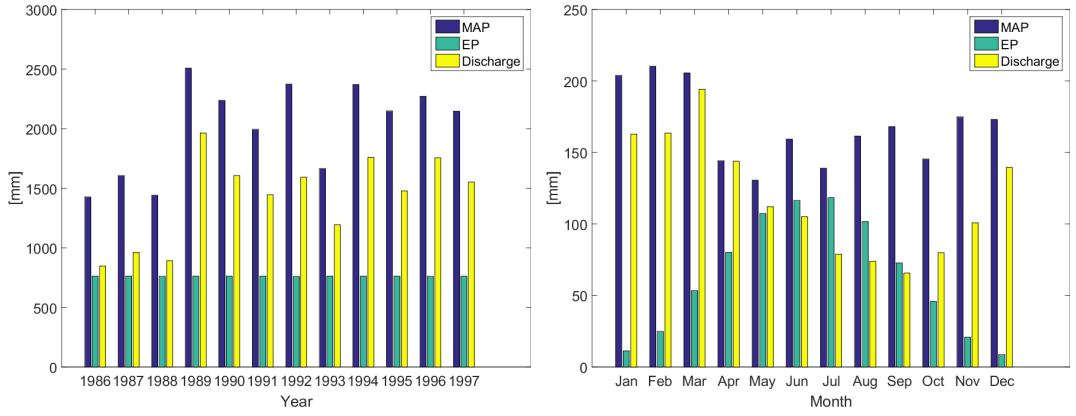


(C) Daily statistics

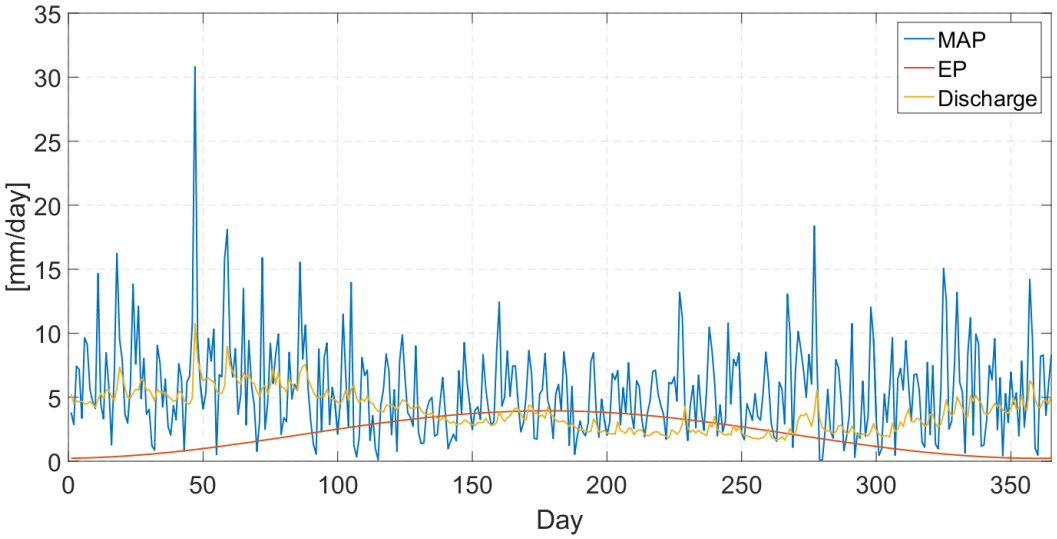


(D) Data record of 12 years (1986-1997)

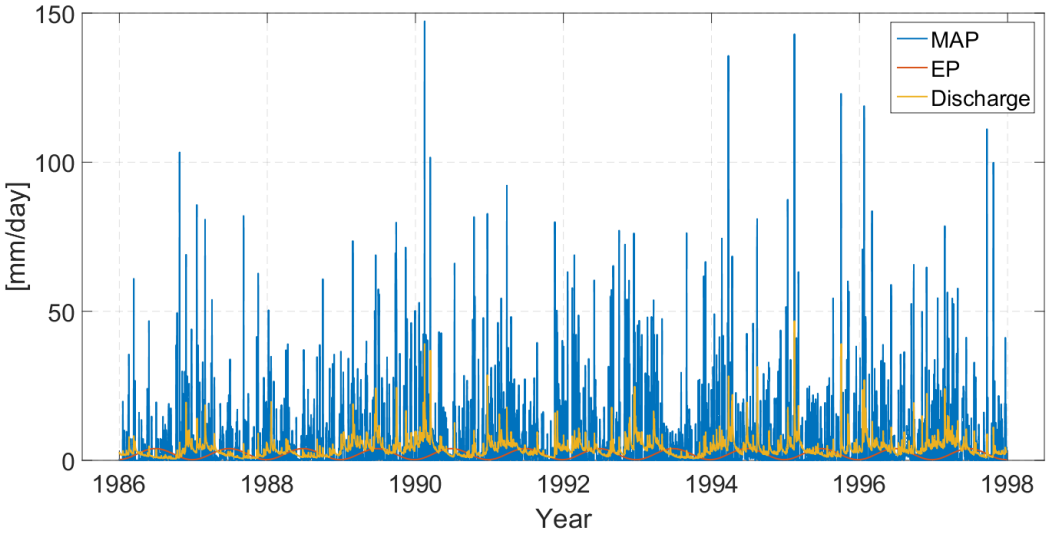
FIGURE 3.4: Values of rainfall, discharge and evaporation for basin 03443000



(A) Annual statistics (B) Monthly statistics

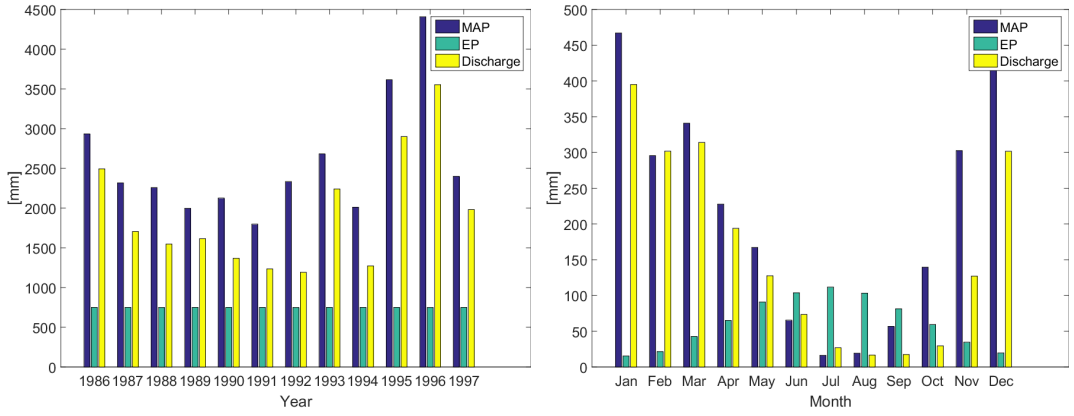


(C) Daily statistics

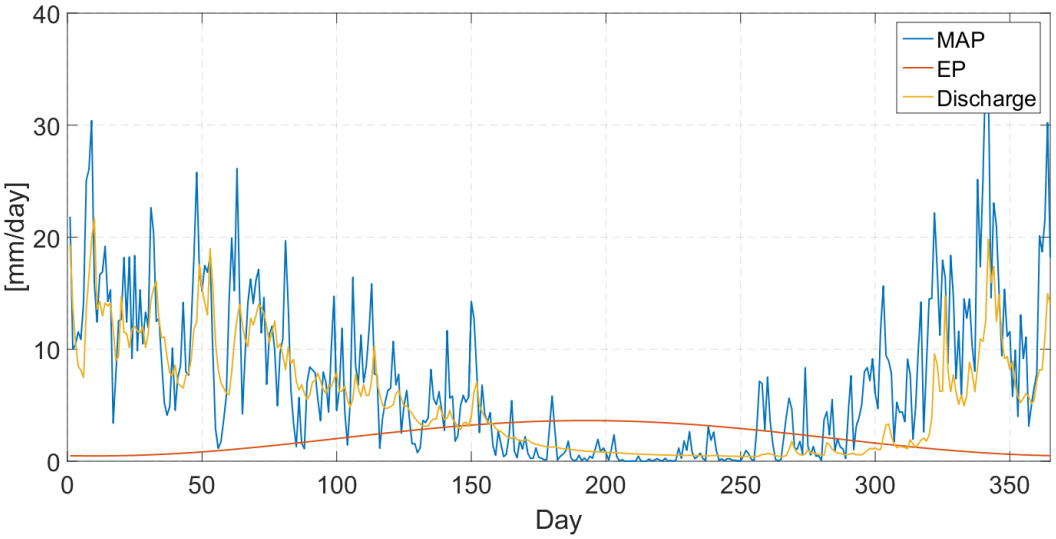


(D) Data record of 12 years (1986-1997)

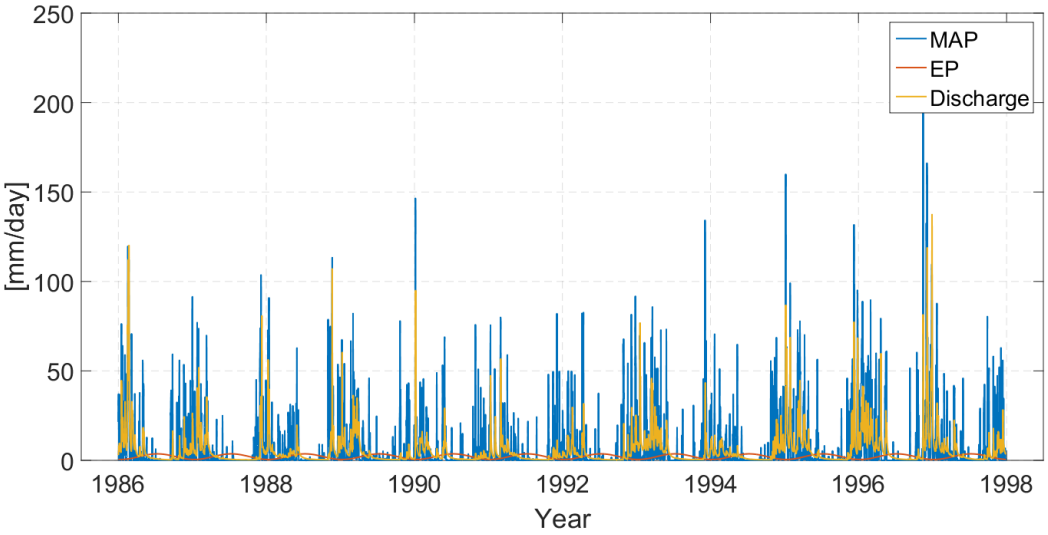
FIGURE 3.5: Values of rainfall, discharge and evaporation for basin 03504000



(A) Annual statistics (B) Monthly statistics



(C) Daily statistics



(D) Data record of 12 years (1986-1997)

FIGURE 3.6: Values of rainfall, discharge and evaporation for basin 11532500

Chapter 4

Calibration and validation

4.1 Introduction

As it was described on chapter 2, the XAJ model has 15 parameters that need to be calibrated and validated. In Table 4.1 there is a summary of the parameters of the XAJ model. Due to the difficulty of obtaining field measurements of these parameters, they usually have to be given fixed values in accordance with experience or optimized through a process of trial and error, which is what was done in this study.

The calibration was carried out using a web-based modified version of the XAJ model which allows the user to run the model in a user-friendly environment. The calibration was performed using a 8-year long (1986-1993) data record and it was validated with a 4-year long (1994-1997) data record. These are data lengths commonly used for the calibration and validation of hydrological models (Loliyana and Patel, 2015), and that is why they were used.

The issue of how the available data should be used in the model calibration and evaluation process is important regardless of the type of hydrological model being considered. There have been few studies that have investigated this issue, which is generally ignored in the vast majority of papers on hydrological modeling. There are even fewer studies that have attempted to develop approaches for addressing this issue (Zheng et al., 2018).

4.2 Sensitivity analysis

Before starting the calibration of any hydrological model, it is important carry out some sort of sensitivity analysis to better understand how a complex hydrological model works. A sensitivity analysis (SA) aims to identify the key parameters that affect model performance and it also plays important roles in model parameterization, calibration, optimization, and uncertainty quantification. The more complex a model is and the more number of parameters it has, the more necessary it is to carry out a good SA.

Many articles have been written about the sensitivity analysis in hydrological modeling, discussing about the methods, theoretical framework, and applications that SA have in hydrological modeling (Song et al., 2015). In this study it was decided to carry out a sensitivity analysis similar to the one done by Song et al., 2013 using the Morris method (Morris, 1991).

4.2.1 The Morris method

The Morris method (Morris, 1991) is a screening method used to identify qualitatively important parameters for mathematical models. The Morris method is sometimes referenced to as a qualitative method: it gives rough estimations with a limited

TABLE 4.1: Summary of the parameters of the Xinanjiang model

Parameter	Physical meaning	Common values
Cp	Ratio of measured precipitation to actual precipitation	0.8-1.2
Cep	Ratio of potential evapotranspiration to pan evaporation	0-2.0
b	Exponent of the tension water capacity curve	0.1-0.3
imp	Ratio of the impervious to the total area of the basin	0-0.005
WUM	Water capacity in the upper soil layer (mm)	5-20
WLM	Water capacity in the lower soil layer (mm)	60-90
WDM	Water capacity in the deeper soil layer (mm)	10-100
C	Coefficient of deep evapotranspiration	0.1-0.3
SM	Areal mean free water capacity of the surface soil layer (mm)	1-50
EX	Exponent of the free water capacity curve	0.5-2.5
KI	Outflow coefficient of the free water storage to inter-ow	KI+KG=0.7
KG	Outflow coefficient of the free water storage to groundwater	KI+KG=0.7
cs	Recession constant for channel routing	0.5-0.9
ci	Recession constant for the lower inter-flow storage	0.5-0.9
cg	Daily recession constant of groundwater storage	0.9-0.998

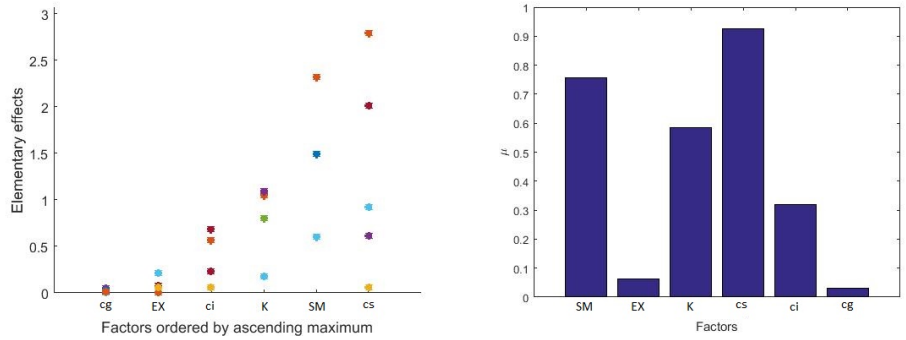
number of calculations. It can be used to simplify a function by identifying the factors that have a low influence on the output.

This method derives measures of global sensitivity from a set of local derivatives, or elementary effects, sampled on a grid throughout the parameter space. It is based on one-at-a-time (OAT) methods, in which each parameter x_i is perturbed along a grid of size Δ_i to create a trajectory through the parameter space. For a given model with k parameters, one trajectory will contain a sequence of k such perturbations. Each trajectory yields one estimate of the elementary effect for each parameter (i.e., the ratio of the change in model output to the change in that parameter). The equation to do the calculation of a single elementary effect for the i-th parameter goes as follows:

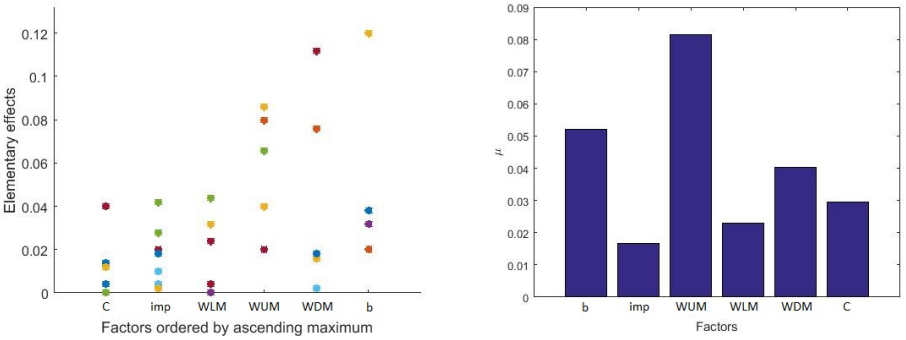
$$EE_i = \frac{f\left(x_1, \ldots, x_i + \Delta_i, \ldots, x_p\right) - f\left(x\right)}{\Delta_i} \tag{4.1}$$

where $f\left(x\right)$ represents the prior point in the trajectory. Using the single trajectory shown in Equation 4.1, one can calculate the elementary effects of each parameter with only $p + 1$ model evaluations. However, by using only a single trajectory, this OAT method is highly dependent on the location of the initial point x in the parameter space and does not account for interactions between parameters. For this reason, the Morris method performs the OAT method over R trajectories through the parameter space. In this study, the sampling approach used was the one originally proposed by Morris, 1991, in which trajectories through the parameter space are generated by perturbing one factor at a time, beginning at a randomly sampled point.

The reason to carry out the sensitivity analysis using the Morris method is that the required number of simulations N necessary to perform the analysis is $N = R(k + 1)$, where k is the number of input variables and R is the number of points to be sampled. Previous studies have demonstrated that using $R = 10$ produces satisfactory results (Campolongo, Tarantola, and Saltelli, 1999), so, for example, in a case of $k = 15$, only 160 model simulations are required for the Morris method, while other kind of methods like variance-based methods would require approximately



(A) Elementary effects on group 2 (B) Barplot of μ group 2 parameters



(C) Elementary effects on group 3 (D) Barplot of μ group 3 parameters

FIGURE 4.1: Results of the sensitivity analysis in basin 02387500

10.000 or more simulations. Therefore, the Morris method is very efficient compared to more demanding methods for sensitivity analysis.

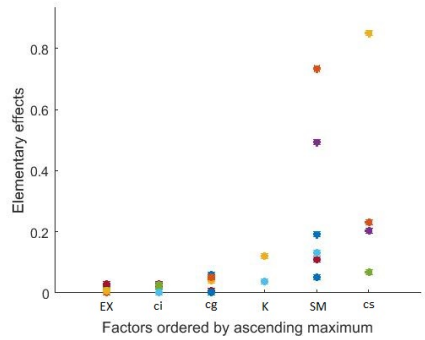
4.2.2 Results

The calibration of the model was carried out following multi-step optimization scheme. Based on the research from Li and Lu, 2012, the total 15 parameters of the XAJ model can be divided into three groups:

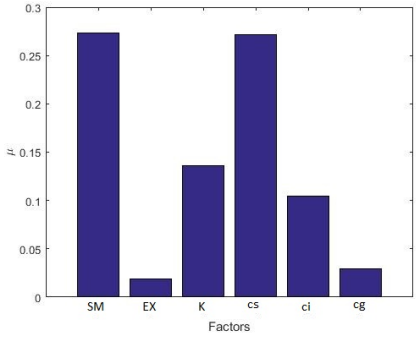
- Group 1: parameters for data adjustment, C_p and C_{ep}, which are sensitive at annual scale.
- Group 2: parameters controlling runoff component separation and routing, SM, EX, KI, KG, cs, ci and cg, which are sensitive at daily scale.
- Group 3: parameters controlling runoff generation, imp, b, WDM, WUM, WLM and C, which are sensitive at annual scale when C_p and C_{ep} are set to be constants.

Therefore, two different sensibility analysis were carried out on each basin, one for the parameters in group 2 and another one for the parameters in group 3. It was not necessary to do a sensitivity analysis for the parameters in group 1 because its values were obtained using Li's equation (Li and Lu, 2014).

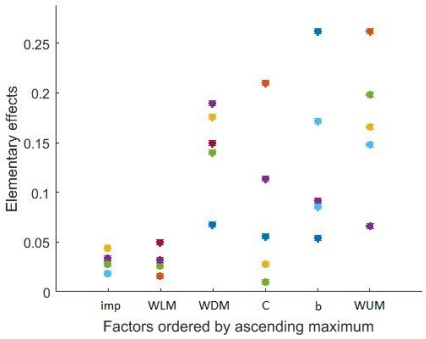
In Figures 4.1, 4.2, 4.3, 4.4 and 4.5, the results of the sensitivity analysis obtained using $R = 10$ are presented. As it can be seen, the most important parameters in group 2 are SM, K and cs because are the ones that have the highest μ in all basins,



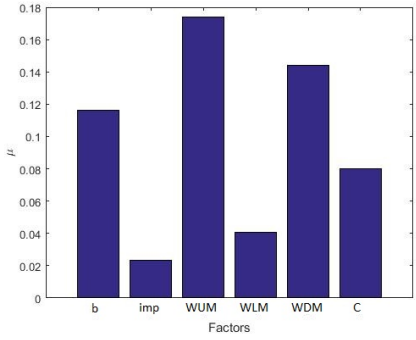
(A) Elementary effects on group 2



(B) Barplot of μ group 2 parameters

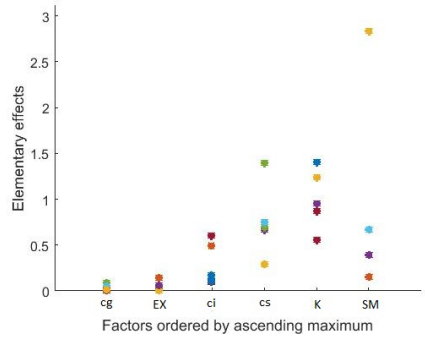


(C) Elementary effects on group 3

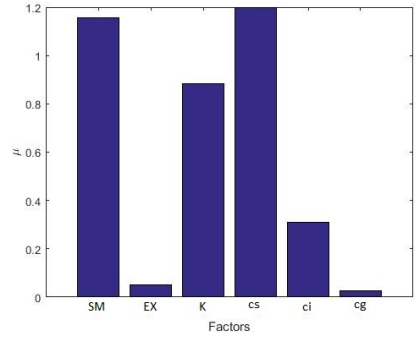


(D) Barplot of μ group 3 parameters

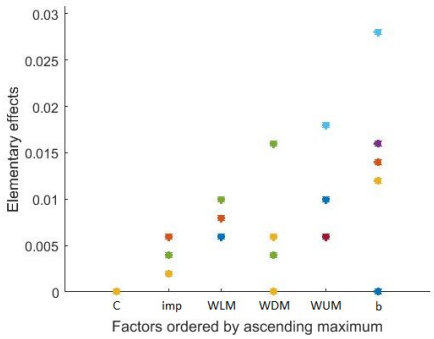
FIGURE 4.2: Results of the sensitivity analysis in basin 02456500



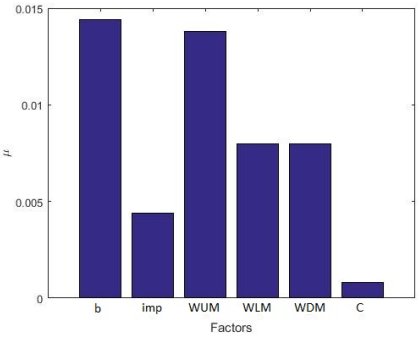
(A) Elementary effects on group 2



(B) Barplot of μ group 2 parameters

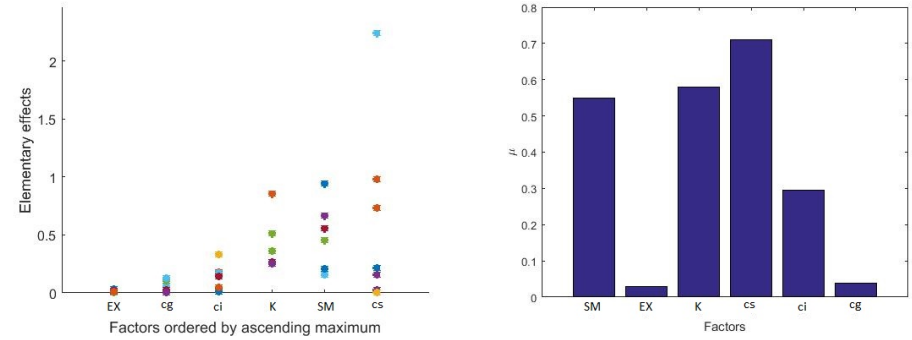


(C) Elementary effects on group 3

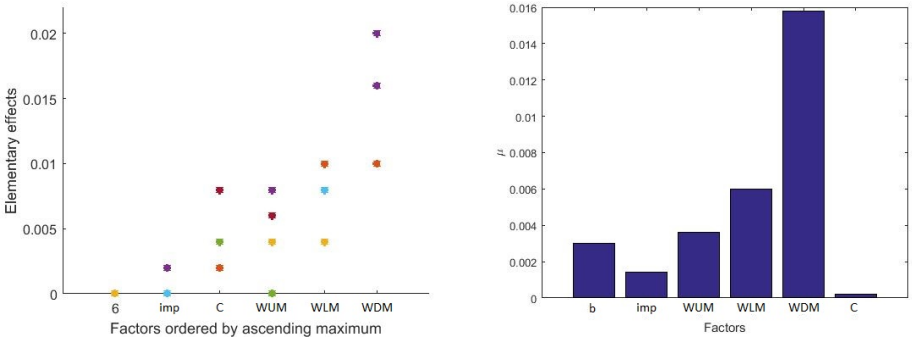


(D) Barplot of μ group 3 parameters

FIGURE 4.3: Results of the sensitivity analysis in basin 03443000

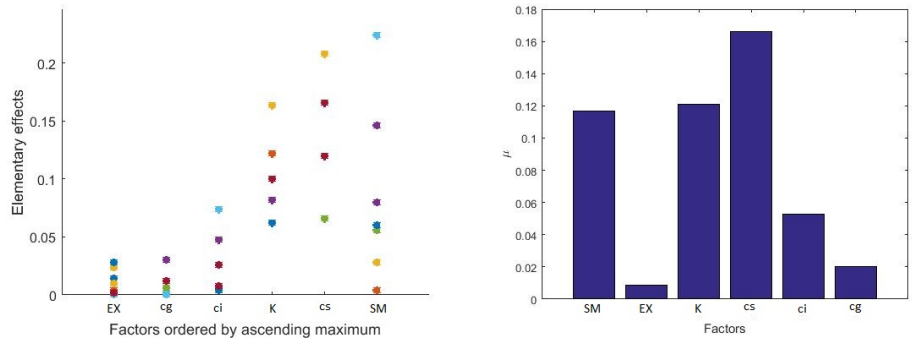


(A) Elementary effects on group 2 (B) Barplot of μ group 2 parameters

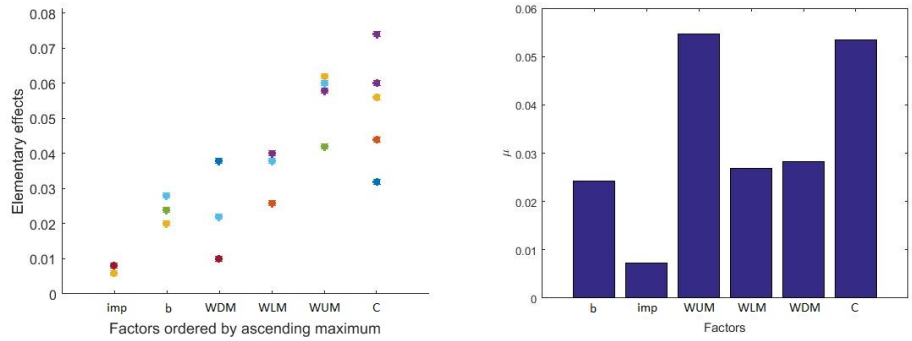


(C) Elementary effects on group 3 (D) Barplot of μ group 3 parameters

FIGURE 4.4: Results of the sensitivity analysis in basin 03504000



(A) Elementary effects on group 2 (B) Barplot of μ group 2 parameters



(C) Elementary effects on group 3 (D) Barplot of μ group 3 parameters

FIGURE 4.5: Results of the sensitivity analysis in basin 11532500

and therefore are the ones that will have to be calibrated with more care. On the other hand, EX, ci and cs have a lower value of μ, and that means that their influence on the output of the model is less significant than the influence of the rest of parameters. For the parameters of group 3, the most important parameter is usually WUM, because it has the highest μ on three of the five study basins. The rest of parameters have different influence on the output of the model depending on the study basin. When performing the calibration, the first parameters of each group to be calibrated will be the ones with a higher μ and then the ones with a smaller μ.

4.3 Calibration

The Xinanjiang model was calibrated with the aid of a web-based application, accessible at https://xaj.nagaokaut.ac.jp/ (last accessed on June 13, 2018) (Kyi, Lu, and Li, 2016). This application allows the user to run a modified version of the XAJ model in a user-friendly environment, providing useful support for better calibration by suggesting parameter settings based on observed hydro-climatic data, calculating Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) at daily, monthly and annual scales using time series of observed and simulated discharge data for every successful model run and hydrograph visualization. The user can perform repeated model runs with different parameter sets until a satisfactory accuracy is achieved. The user can download the result and parameter files for all or any specific model run. The main difference of the modified XAJ model used with respect of the original XAJ model is that the flow routing features of the model are not used and the model works in a lumped way. However, since in this study only small catchments have been simulated and it was considered that the hydrological characteristics of the catchments were homogeneous, it was not necessary to use the flow routing part of the model.

The calibration of the parameters of the model was done through a manual trial-and-error process. This method needs long time for calibration and the success is strongly dependent on the modeler's experience. The number of parameters in the XAJ model is high and to facilitate the calibration the multi-step optimization scheme proposed by Li and Lu, 2012 was used. Compared with optimizing all the parameters directly, the multi-step optimization scheme makes the calibration easier because instead of having to optimize 15 at once, it divides them into three groups and the parameters are optimized group by group.

The procedure followed to optimize each group of parameters was to start with the values of C_{ep}, which is the ration of potential evapotranspiration to pan evaporation, that was obtained using Li's equation (Li and Lu, 2014) by considering the relationship of runoff coefficient and pan aridity index. To calibrate the rest of the parameters, a set of random initial values was created and then, using the knowledge acquired from the sensitivity analysis, different sets of values for the first group of parameters were tried until a high Nash-Sutcliffe efficiency was achieved and then proceeded to the calibration of the following groups. Once all three groups of parameters had been optimized, the optimization procedure will start again with the first group using as initial values the values obtained in the previous iteration. This was done until a satisfactory Nash-Sutcliffe efficiency was achieved. To evaluate the quality of the simulation, the following classification for model performance evaluation was used: good simulation if $0.65 < NSE < 0.75$ and very good if $0.75 < NSE$ (Moriasi et al., 2007).

To perform the calibration, the model was set to run at a daily interval and the initial condition used was that the soil was fully saturated, but in order to avoid the effect from initial conditions, the data in the first year was not considered when calculating the performance indicators.

4.3.1 Performance indicators

There is a large number of performance criteria that can be used for the identification of meaningful parameter values (Guse et al., 2017). In this study, to check the quality of the calibration, three of the most popular performance indicators in hydrology were used.

Nash-Sutcliffe

The Nash–Sutcliffe model efficiency coefficient (Nash and Sutcliffe, 1970) is used to assess the predictive power of hydrological models. It is defined as:

$$NASH = 1 - \frac{\sum_{t=1}^{n}[Q_o(t) - Q_s(t)]^2}{\sum_{t=1}^{n}[Q_o(t) - \overline{Q}_0]^2} \tag{4.2}$$

were Q_o, Q_s and $\overline{Q}_0$ are the observed streamflow, the simulated streamflow, and the average observed streamflow, respectively. In this study the daily, monthly and annual Nash–Sutcliffe efficiency were calculated.

RMSE and NMRSE

The root-mean-squared error represents the sample standard deviation of the differences between predicted values and observed values. It is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q_{o,i} - Q_{s,i})^2}{n}} \tag{4.3}$$

Normalizing the RMSD facilitates the comparison between datasets and in this study it was normalized by dividing it by the average of the measured data:

$$NRMSE = \frac{RMSE}{\overline{Q}_0} \tag{4.4}$$

In this study the daily RMSE and NRMSE were calculated.

R-squared

The R^2 parameter is the proportion of the variance in the dependent variable that is predictable from the independent variables.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}[Q_s(t) - \overline{Q}_0]^2}{\sum_{i=1}^{n}[Q_o(t) - \overline{Q}_0]^2} \tag{4.5}$$

In this study the daily R^2 was calculated.

TABLE 4.2: Value of the parameters obtained in the calibration

Parameter	02387500	02456500	03443000	03504000	11532500
C_p	1.00	1.00	1.00	1.00	0.95
C_{ep}	1.200	1.036	0.9771	0.794	0.907
b	0.1	0.3	0.3	0.3	0.3
imp	0.00	0.05	0.03	0.05	0.05
WUM	20	20	20	20	20
WLM	90	90	70	90	90
WDM	100	100	90	100	100
C	0.25	0.30	0.30	0.30	0.30
SM	40	50	25	35	10
EX	1.5	0.5	1.5	2.5	0.5
KI	0.20	0.70	0.10	0.15	0.30
KG	0.5	0.0	0.6	0.55	0.40
c_s	0.80	0.80	0.85	0.75	0.70
c_i	0.70	0.75	0.50	0.90	0.90
c_g	0.990	0.998	0.990	0.990	0.990

TABLE 4.3: Value of the performance indicators obtained in the calibration

Indicator	02387500	02456500	03443000	03504000	11532500
DNash	0.803	0.757	0.812	0.806	0.752
MNash	0.922	0.908	0.910	0.886	0.927
ANash	0.989	0.858	0.973	0.925	0.827
RMSE	0.912	1.307	1.066	1.335	3.658
NRMSE	0.529	0.906	0.330	0.353	0.858
R^2	0.807	0.764	0.815	0.809	0.753

4.3.2 Calibration results

The sets of parameters obtained for each basin after performing the calibration, can be seen in Table 4.2. Also, in Table 4.3 there is a summary of the values obtained for each one of the performance indicators explained in this chapter when considering a spin-up time of 1 year. It can be seen that all basins have a daily Nash-Sutcliffe efficiency higher than 0.75, which is classified as a very good result, and that the rest of the indicators also have reasonably good values.

4.4 Validation

After calibrating the parameters using data from 1986 to 1993, the model was validated using data from 1994 to 1997. During the validation the model was also running at a daily interval. All the analyzed basins obtained a daily Nash-Sutcliffe efficiency higher than 0.7 in the validation period, which falls in the good or very good range of values (Moriasi et al., 2007), and that ensures a good performance of the model. In Table 4.4 there is a summary of the values obtained for each one of the performance indicators after the validation when considering a spin-up time of 1 year.

TABLE 4.4: Value of the performance indicators obtained in the validation

Indicator	02387500	02456500	03443000	03504000	11532500
DNash	0.756	0.713	0.752	0.779	0.874
MNash	0.895	0.860	0.777	0.826	0.977
ANash	0.572	0.377	0.937	0.556	0.999
RMSE	0.938	1.603	1.285	1.627	4.484
NRMSE	0.495	0.853	0.357	0.372	0.582
R^2	0.756	0.718	0.757	0.779	0.877

Figures 4.6, 4.7, 4.8, 4.9 and 4.10, show the observed and the simulated streamflow for each year of the validation period. As it can be seen, there is a good correlation between the observed streamflow and the simulated streamflow.

4.5 Summary

In this chapter the procedure followed for the calibration and validation of the 15 parameters of the Xinanjiang model in five different basins is described. There is also a description of the performance indicators used in this study and the values obtained for each one of them. Finally, the validation hydrographs for each one of the basins of the study are presented.

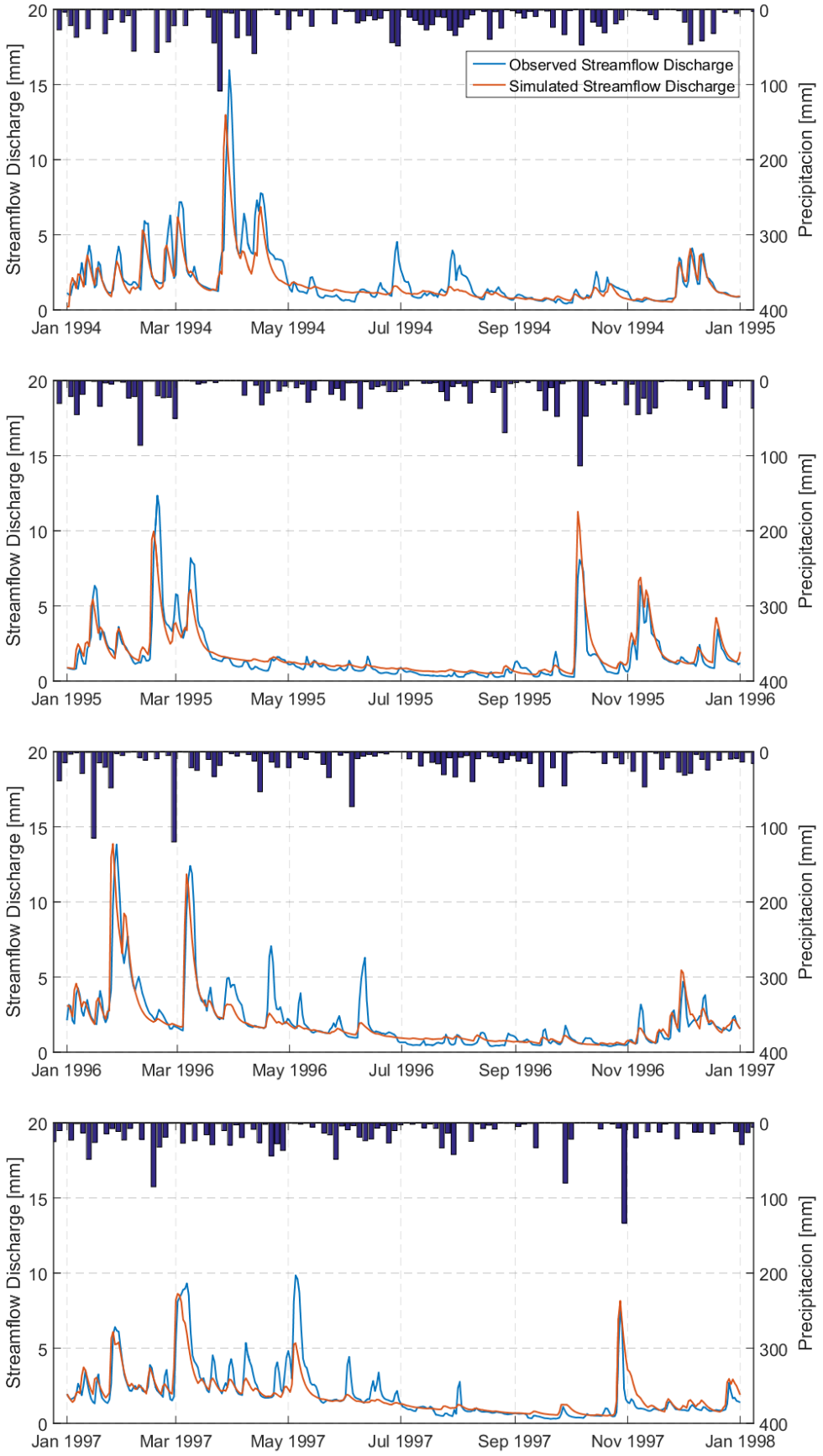


FIGURE 4.6: Validation hydrograph for basin 02387500

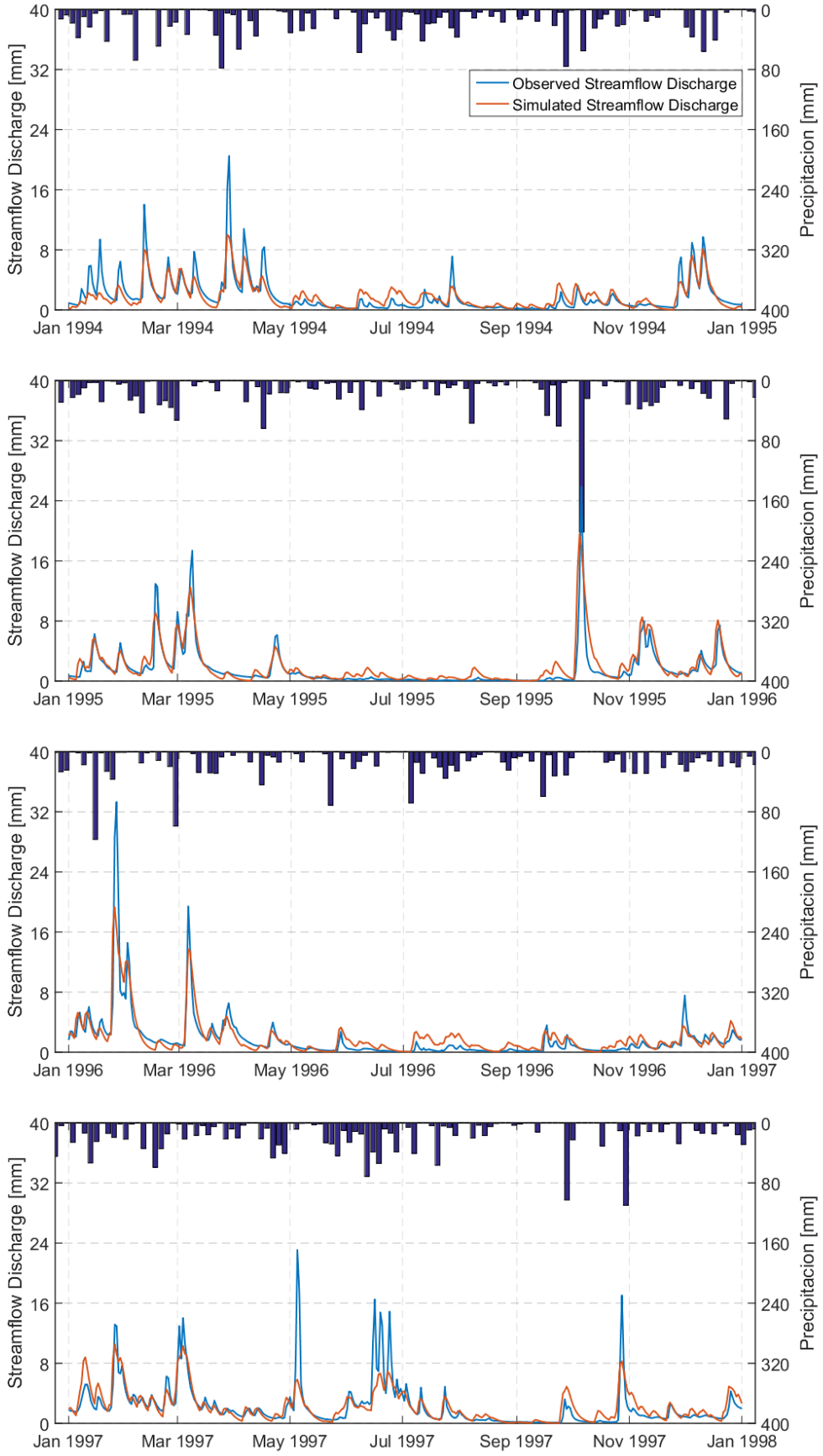


FIGURE 4.7: Validation hydrograph for basin 02456500

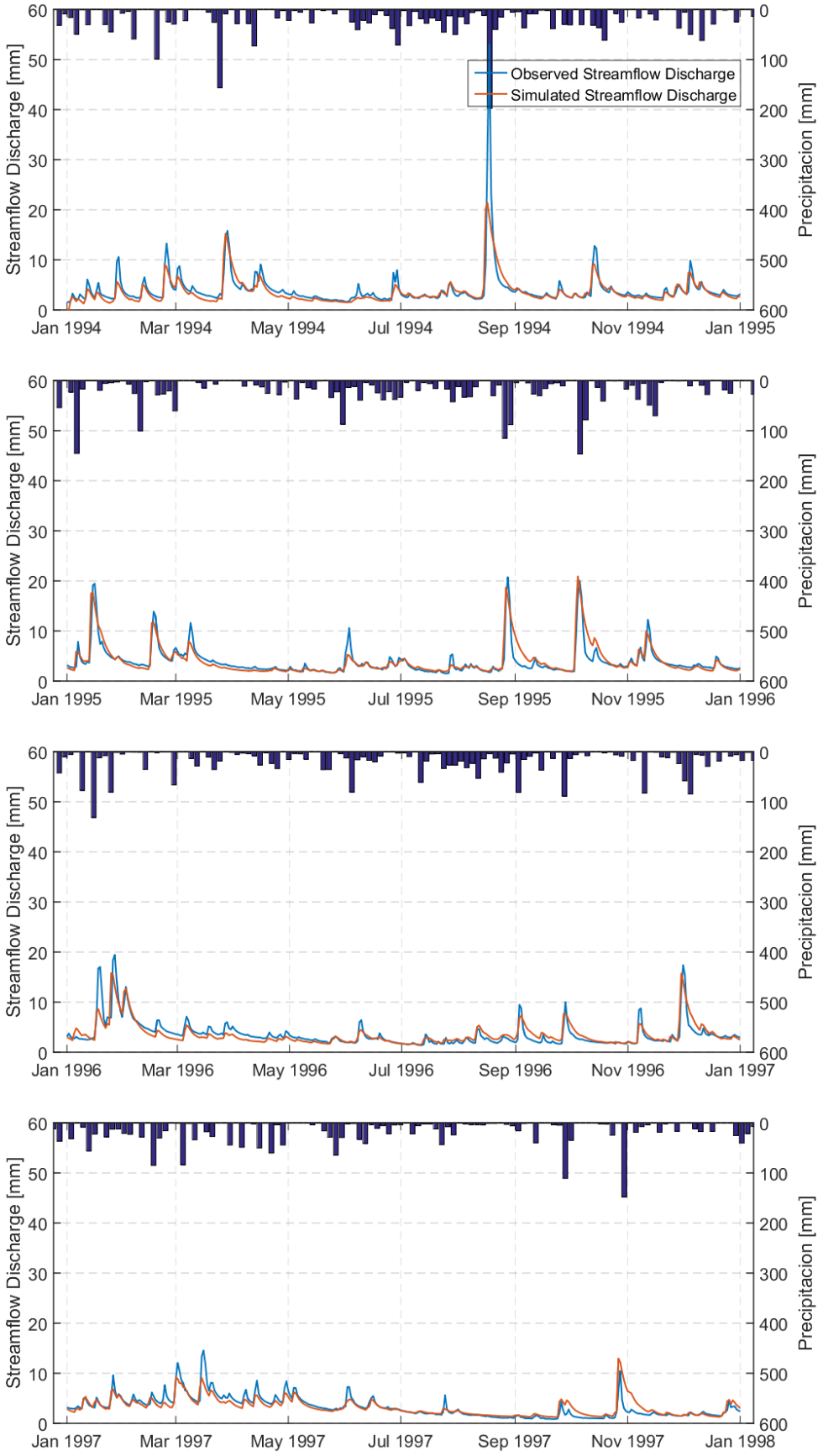


FIGURE 4.8: Validation hydrograph for basin 03443000

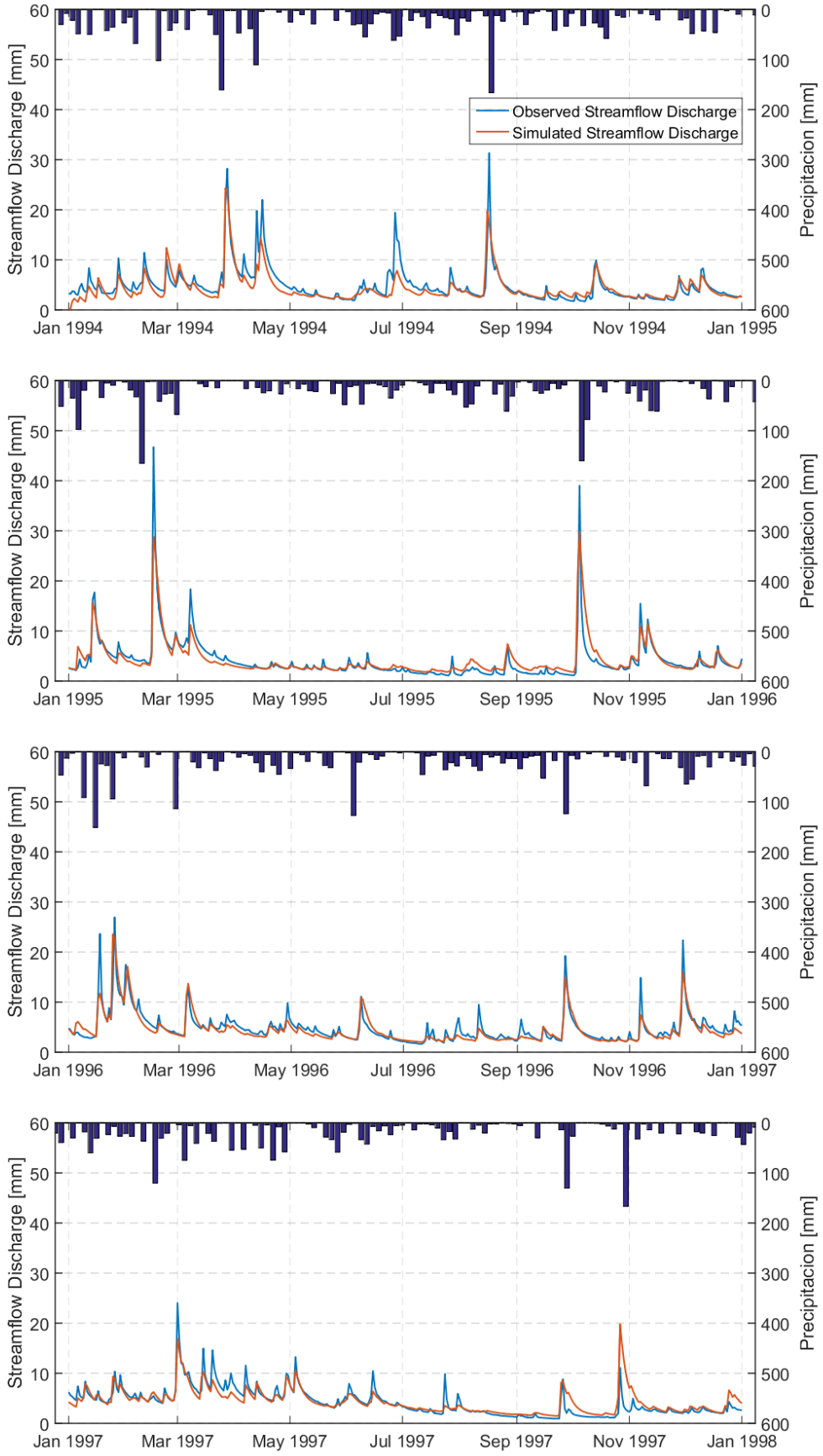


FIGURE 4.9: Validation hydrograph for basin 03504000

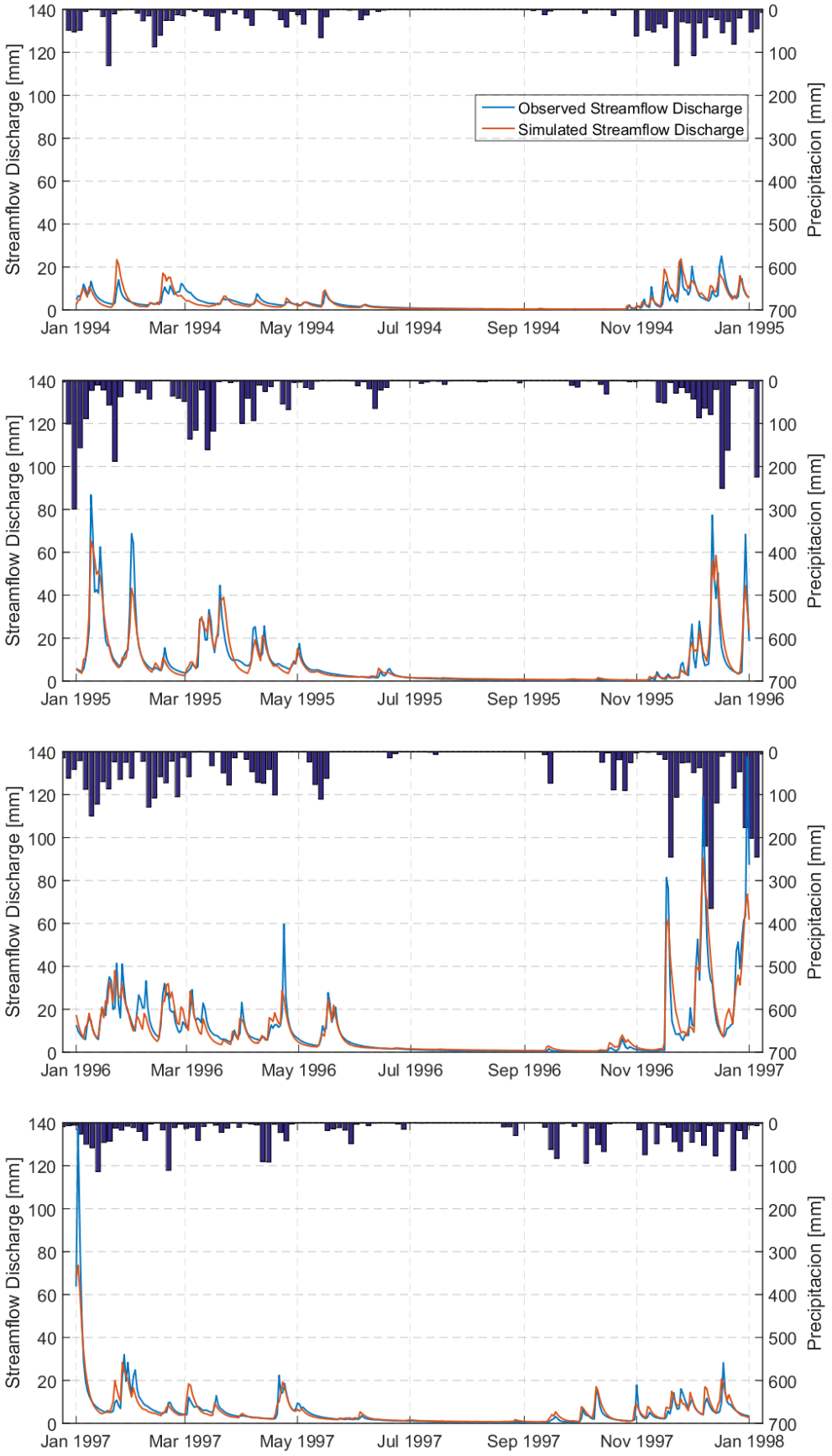


FIGURE 4.10: Validation hydrograph for basin 11532500

Chapter 5

Impact of temporal resolution of input data on model performance

5.1 Introduction

In the previous chapters, the process of calibration and validation of the XAJ model for five different study basins has been explained. The data available in the study basins had hourly resolution, which allowed to do a precise simulation, but in most basins, specially the ones in developing countries, the data available usually has a lower temporal resolution. The main objective of this chapter is to find out how the temporal resolution of the input data affects the output of the model and which temporal resolutions can be used without compromising very much the quality of the simulation. Another objective is to assess the influence that the size of the study basin has on the required temporal resolution of the input data of the model.

In order to do so, different input data sets were created from the initial data set with hourly resolution to create synthetic data with different time resolution. These new data sets were introduced into the XAJ model to check the values of the NASH efficiency when the temporal resolution of the input data was decreased.

Finally, when all the simulations were carried out, it was possible to extract some conclusions about the influence that the time scale of the simulation and the size of the basin have on the NASH efficiency.

5.2 Creation of new input data with different temporal resolution

The data available had a temporal resolution of 1 hour and in order to carry out this research it was necessary to do simulations pretending that the data had a lower temporal resolution. The temporal resolutions used in this study were 1, 2, 3, 4, 6, 8, 12 and 24 hours.

The procedure followed to create the low resolution data sets was to add the precipitation that had fallen during a time interval and then equally distribute it through the time interval. Using this procedure it is possible to get the data sets that would have been obtained if the measurements had been taken at a lower time resolution. In Table 5.1 there is an example of the different input data that can be created for 1 single day from the original data that has hourly resolution. Even though the total rainfall during the day is exactly the same, its distribution changes when using different time resolutions. Having data with a low time resolution usually implies that the peaks are reduced because they are distributed along the interval, but this also means that it is considered that rainfall occurs during times when there was no rain. For example, during the first two hours of the example from Table 5.1 there was

TABLE 5.1: Example of input data with different temporal resolution

	1h	2h	3h	4h	6h	8h	12h	24h
0:00	0.0000	0.0000	0.0033	0.0075	0.0133	0.0150	0.0167	0.0342
1:00	0.0000							
2:00	0.0100	0.0150						
3:00	0.0200		0.0233					
4:00	0.0200	0.0250		0.0225				
5:00	0.0300							
6:00	0.0200	0.0200	0.0267		0.0200			
7:00	0.0200							
8:00	0.0400	0.0300		0.0200		0.0575		
9:00	0.0200		0.0133					
10:00	0.0200	0.0100						
11:00	0.0000							
12:00	0.0200	0.0350	0.0900	0.0950	0.0733		0.0517	
13:00	0.0500							
14:00	0.2000	0.1550						
15:00	0.1100		0.0567					
16:00	0.0300	0.0300		0.0175		0.0300		
17:00	0.0300							
18:00	0.0100	0.0050	0.0133		0.0300			
19:00	0.0000							
20:00	0.0300	0.0400		0.0425				
21:00	0.0500		0.0467					
22:00	0.0000	0.0450						
23:00	0.0900							
Total rainfall	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82

no rain, but when using a time interval of 3 hours or higher it is considered that there is rainfall during the first hours of the day. Also, the peak in rainfall experienced at 14 o'clock is reduced when the temporal resolution decreases. In Figure 5.1 the hydrographs created using different time resolutions can be seen together to be able to appreciate better how the input data varies when changing the time resolution.

The XAJ model was set to run always on a hourly timescale. It is important that the model always does the simulation using the same timescale because this way the timescale of the model and the temporal resolution of the data do not couple with each other because the objective of this study only focuses on the influence of the temporal resolution of the input data, not the influence of the timescale of the model.

5.3 Results and discussion

In this section, the results obtained after simulating a 4 year period (from 1994 to 1997) of the five study basins using 8 different time resolutions are presented. The whole 4 years were simulated and then different time periods where separated in order to see better the results and to be able to extract more conclusions. The performance indicator chosen to analyze the quality of the simulation was the hourly Nash-Sutcliffe efficiency. The graphs of the results show the normalized

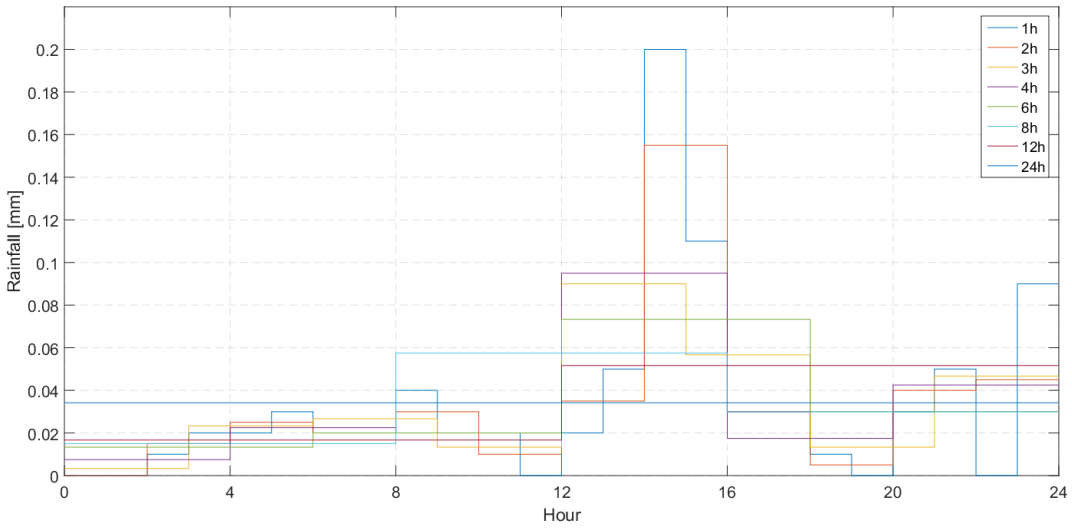


FIGURE 5.1: Hydrographs created using different temporal resolutions

Nash-Sutcliffe efficiency to be able to compare the different study basins and then they show the absolute value of the Nash-Sutcliffe efficiency in order to know how good the simulation is. The color of the legend used in the graphs is related to the size of the basin: the darker the color, the bigger the basins. This makes it easier to extract conclusions related to the impact of size of the study basin on the required temporal resolution of the input data.

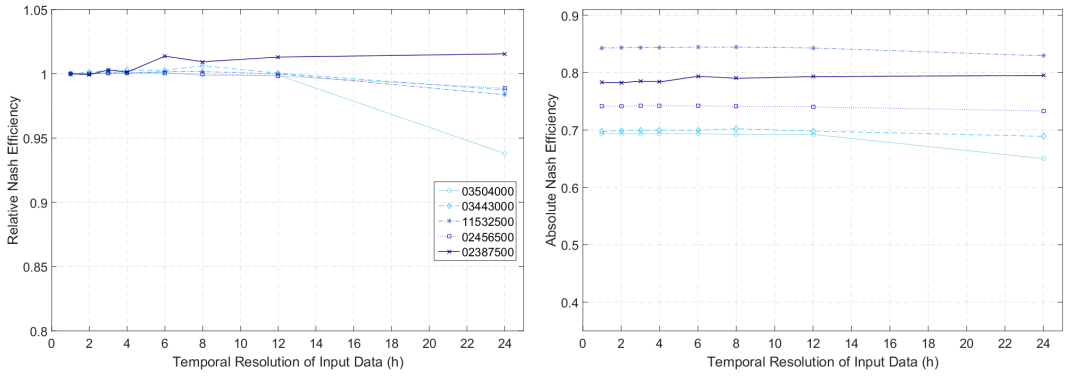
5.3.1 4-year period results

After the simulation of the 4 year period was carried out, the different Nash efficiencies were calculated for the whole period. The values obtained for the Nash efficiency when the model is running at a hourly timescale are quite good and very similar to the ones obtained when when the model was running at a daily interval. The range of values of the Nash efficiency goes from 0.69 to 0.85, which can be considered satisfactory results.

To be able to analyze better the behavior of the Nash efficiency, its values were plotted. In Figure 5.2 the normalized and absolute Nash efficiencies can be seen. When computing the Nash efficiency for the 4 year period, three different behaviors can be observed depending on the size of the study basin.

First of all, the Nash efficiency of basin 03504000, which was the smallest basin, did not considerably vary when using time resolutions ranging from 1 to 12 hours but when the time resolution of the input data was lower than 12 hours, then there was a high decrease of the Nash efficiency. When a daily resolution of the input data was used for basin 03504000, then the value obtained for the Nash efficiency was around 0.65, which means that there was a decrease of approximately of 6 % of the Nash efficiency. This decrease was the highest experienced by any of the basins.

Secondly, the behavior of basins 03443000, 11532500 and 02456500, which were the medium size study basins with sizes ranging from 700 km^2 to 2300 km^2, was pretty similar and could be considered that all 3 basins had the same behavior. In



(A) Normalized Nash efficiency (B) Absolute Nash efficiency

FIGURE 5.2: Nash efficiency for 1994-1997

this case the Nash efficiency suffered a slightly decrease of its value as the time resolution of the input data decreased. When the using data with a 24-hour resolution, the decrease of the Nash efficiency was between 1 and 2 % with respect of the Nash efficiency obtained when the simulation was carried out using input data with hourly resolution.

Finally, it could be seen that the Nash efficiency of basin 02387500, which was the biggest one with an area slightly above 4000 km^2, has a strange and counter-intuitive behavior. When the time resolution of the input data of the model was decreased, then the Nash efficiency increased. Even though this increase was not very high, in the more favorable case was below 2%, it was an unexpected result because what was expected was that Nash efficiency would decrease when increasing the time interval of the input data.

The conclusion that could be extracted from this part of the study was that the behavior of the Nash efficiency when using input data with different time resolutions is different for basins with different size. The behavior of the Nash efficiency in the basins could be classified into three different groups, which are the same groups that the basins were classified when looking at its size (small, medium and large basins). No relation was found between the variation of the normalized Nash efficiency and the initial value of the Nash efficiency obtained because even though basins 03443000, 11532500 and 02456500 have different initial values of the Nash efficiency (ranging from 0.7 to 0.85), their normalized Nash efficiency has a very similar behavior. The unexpected result obtained for basin 02387500 might have been due to the fact that the size of the basin was very big compared to the rest of the basins. If the size of the basin is too big, then the rain that happens at a hourly scale is not reflected in the stream-flow until some time later because it takes more time for the water to travel through the surface of the basins.

5.3.2 Yearly results

In this section, the results obtained when computing the NASH efficiency of only 1 year of the 4 years simulated are presented and discussed. This way it is possible to see how the Nash efficiency changes during each year of the study period. Also, it will be possible to see the behavior of the Nash efficiency when a shorter period of time is considered (one year instead of four).

Like in the previous case, the values of the normalized and absolute Nash efficiency were plotted and they can be seen in Figure 5.3. The range of values of the Nash efficiency for each year when the input data has hourly resolution ranges from 0.61 to 0.87. The Nash efficiency obtained in some study basins for the year 1997 was not as high as it would be desirable, but during the other years all basins had values of the Nash efficiency higher or very close to 0.7, which is what is recommended in most cases.

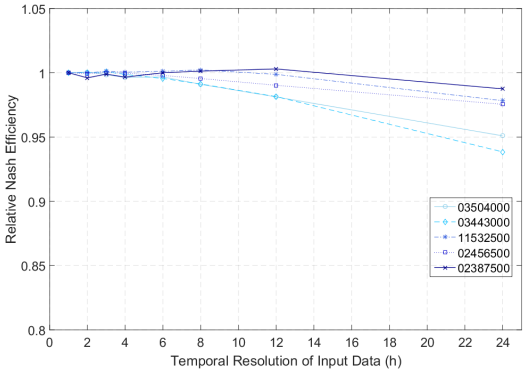
Looking at the graphs it can be seen that the trend followed by the Nash efficiency is very similar to the one followed in the previous case and that the basins can be divided in 3 different groups based on their Nash efficiency behavior.

To begin with, basin 03504000, which was the smallest one, it was the one that usually suffered the highest decrease in its Nash efficiency. The years 1994 and 1997 other basins had a lower Nash efficiency than basins 0350400, but their values were very close. When the using data with a 24-hour resolution, the decrease of the Nash efficiency was between 2 and 9 % for any of the years in the study period with respect of the Nash efficiency obtained when the simulation was carried out using input data with hourly resolution. Also, in this basin, the loss of quality of the simulation happened before than in other basins. For example, in Figures 5.3a and 5.3e it can be seen that when using input data with temporal resolution of 6, 8 or 12 hours, most basins have a normalized Nash efficiency value very close to one, whereas the normalized Nash efficiency in basin 03504000 has already started to decrease.

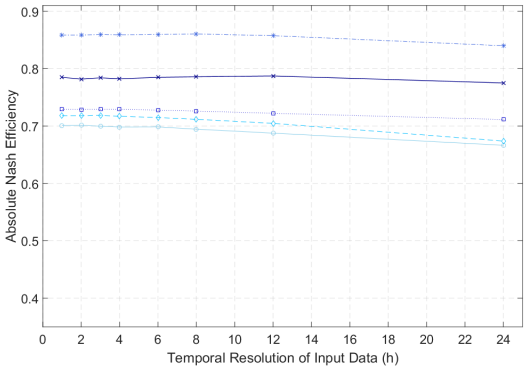
On the other hand, the behavior of the biggest basin, ID number 02387500, was quite the opposite. Not only was the basin in which the decrease in the Nash efficiency was the smallest, but during some years it even increased its value when the temporal resolution is of the input data decreased. Usually, this increase was not very high but during 1996 it reached a value of 4 %, which was the highest increase experienced in any year.

In between these two behaviors there were basins 3443000, 11532500 and 02456500, whose Nash efficiency usually decreased, in contrast with the behavior of basin 02387500, but this decrease was not as high as the one experienced by basin 03504000. The only time that a basin of this group experienced an increase of the Nash efficiency was basin 02456500 during the year 1995, but this increase was very small compared to the ones experienced by basin 02387500 because its value did not even reach a 1 %. Not considering this odd case, the range of values of the decrease of the Nash efficiency in these three basins was between 0 and 6 %.

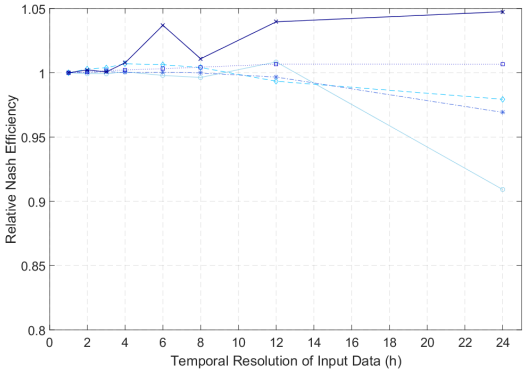
Once the results of these part of the study have been commented, it can be concluded that the behavior of the Nash efficiency when using input data with different time resolutions considering only one year was very similar to when considering four years. This similarity was more noticeable in some years, like 1995 and 1996, than others, but it was always there. In 1994, even though basin 03504000 was not the basin which experienced the highest lost of quality in the simulation, it is the smallest second smallest basin the one the suffer the highest decrease of the Nash efficiency, and the value of this decrease was very similar in both basins. Moreover, in 1997, even though the difference between small and medium-size basins could not be observed, it was possible the see the completely different result obtained of the biggest basin. The range of values of the decrease experienced in the Nash efficiency in this case was slightly higher than in the previous case: the maximum decrease in the smallest basin was 9 % instead of 6 % and the range of values of decrease in the medium-size basins went from 1 and 2 % obtained in the previous case to 0 to 6 % in this case. This seems to indicate that the temporal resolution of the input data is more important when the period of time analyzed gets shorter.



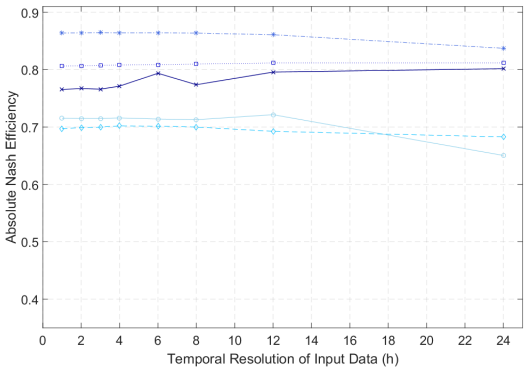
(A) Normalized Nash efficiency for 1994



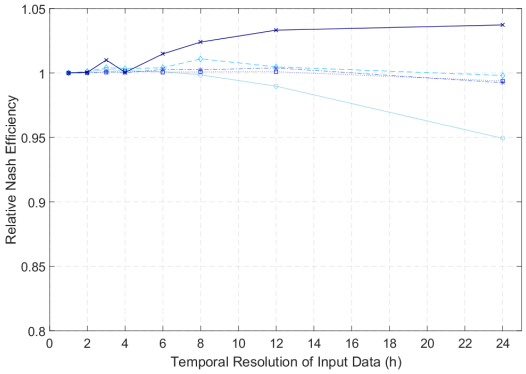
(B) Absolute Nash efficiency for 1994



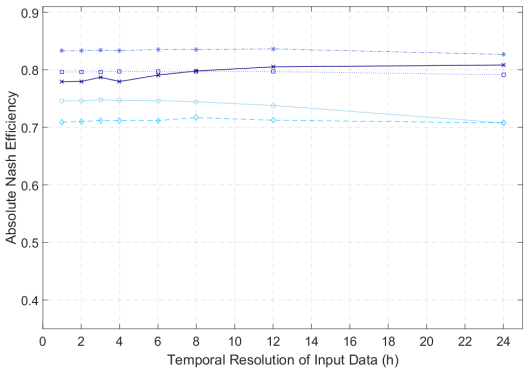
(C) Normalized Nash efficiency for 1995



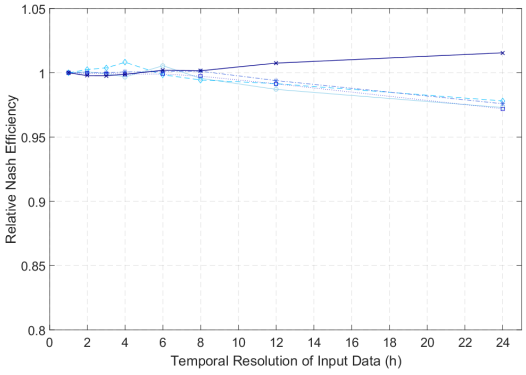
(D) Absolute Nash efficiency for 1995



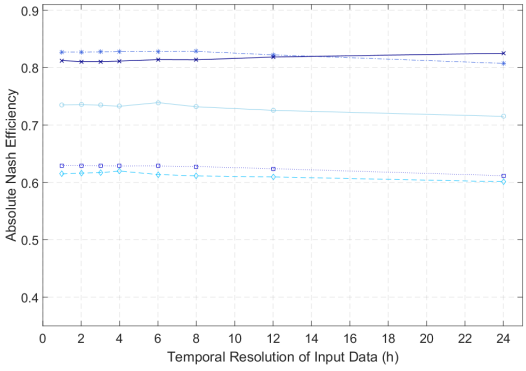
(E) Normalized Nash efficiency for 1996



(F) Absolute Nash efficiency for 1996



(G) Normalized Nash efficiency for 1997



(H) Absolute Nash efficiency for 1997

FIGURE 5.3: Nash efficiency for each year

5.3.3 Seasonal results

In this part of the study the results obtained when computing the Nash efficiency when only considering data from on particular season are presented. The selected study basins, as can be seen in Figures 3.2, 3.3, 3.4, 3.5 and 3.6, had very different behaviors depending on the season. During a year, usually the months of highest rainfall are the winter months and during the summer the rainfall is less intense. It most cases the rainfall during hot months is around 60% of the rainfall during cold months, but in some basins this difference was even higher. Since the rainfall has different behaviors during the year, it was also interesting to analyze if the Nash efficiency behaved differently depending on the season that is being simulated.

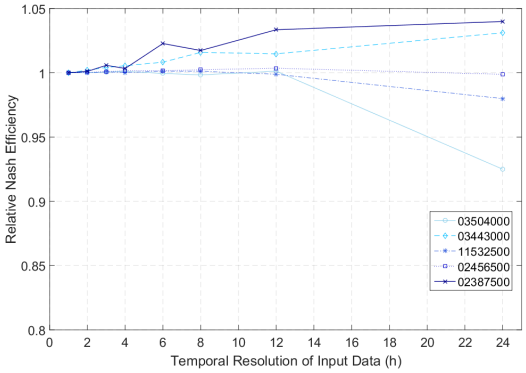
To be able to analyze the evolution of the Nash efficiency, its values were plotted against the time resolution of the input data in Figure 5.4. The values obtained for the Nash efficiency when only one season was considered and the model was running at a hourly timescale were very similar to the ones obtained when when the model was running at a daily interval. The range of values of the NASH efficiency went from 0.69 to 0.85, which can be considered satisfactory results. By looking at Figure 5.4, it can be seen that, like in the case of the 4-year period, the basins can be divided in 3 different groups based on their Nash efficiency behavior.

The first group was the smallest basin, 03504000, that when the time resolution of the input data was lower than 12 hours the Nash efficiency was stable, but there was a rapid decrease of the Nash efficiency when data with daily resolution was used. The value of the Nash efficiency when using a 24-hour time interval in this basin was between 92 and 94% during all seasons. The highest decrease in the Nash efficiency happened during the spring season and the smallest decrease happened during winter season, but the difference of the decrease between these seasons was very small, around 1%.

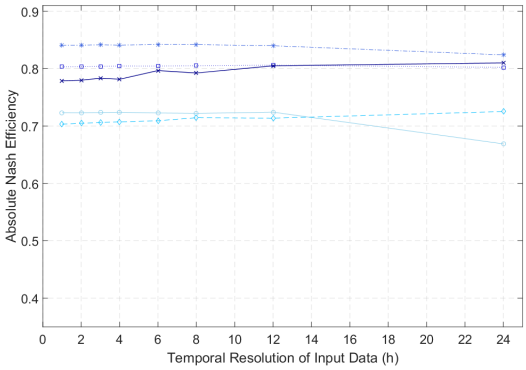
The behavior of the Nash efficiency experienced in basin 02387500 when the time resolution of the input data decreased was that there was an improvement in the NASH efficiency. This improvement happened in all seasons and its values ranged from 2% to 4% when the time resolution of the input data was 24 hours. The season with the highest improvement in the Nash efficiency in basin 02387500 was winter and the season with the least improvement was spring.

In the group of the medium size basins (basins 03443000, 11532500 and 02456500) the behavior experienced by the Nash efficiency was that it suffered a slightly decrease when the time interval of the input data was increased. When using data with a 24-hour resolution, the decrease of the Nash efficiency was between 0 and 2 % with respect of the Nash efficiency obtained when the simulation was carried out using input data with hourly resolution. This happened for all fours seasons in basins 11532500 and 02456500 and during spring and summer in basin 03443000. The during winter and autumn, the Nash efficiency in basin 03443000 had a similar behavior than the one experienced by basin 02387500, increasing its value as the time resolution of the input data increased.

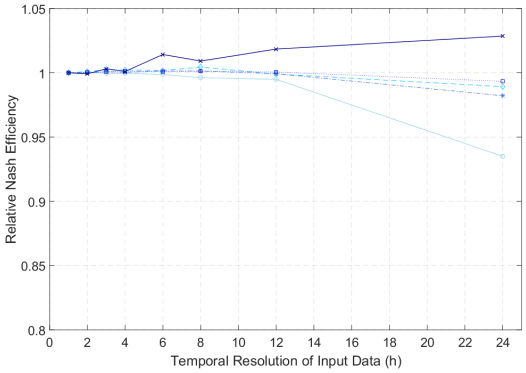
Having analyzed the variation of the Nash efficiency in the five study basins during different seasons, the main conclusion extracted was that the impact of temporal resolution of input data on model performance was not related to the season of the simulation. Looking at Figure 5.4 it can be seen that almost all basins have the same behavior, no matter what season is being considered. The only basins in which there was a different behavior of the Nash efficiency depending on the season was basin 02387500. This different behavior is not related to the the absolute value of the NASH efficiency because its value was very similar during any of the 4 seasons. Usually,



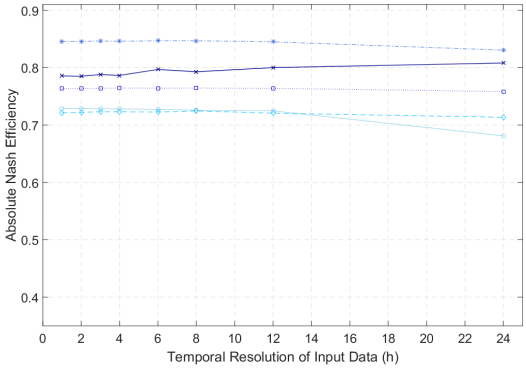
(A) Normalized Nash efficiency for winter



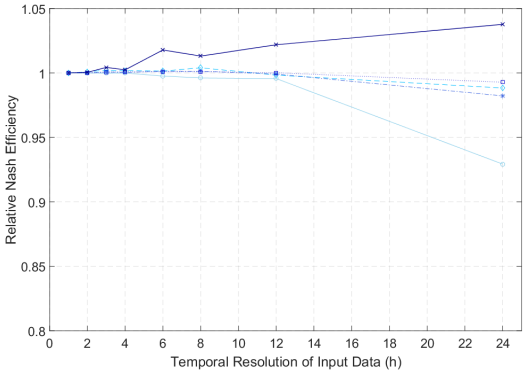
(B) Absolute Nash efficiency for winter



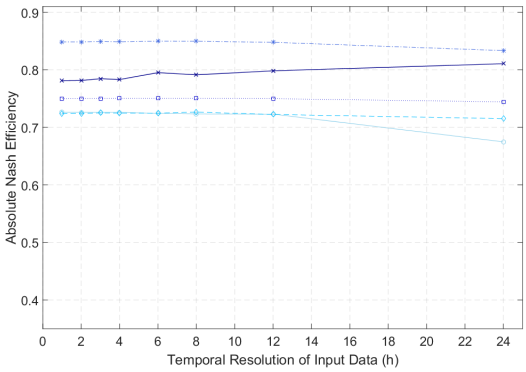
(C) Normalized Nash efficiency for spring



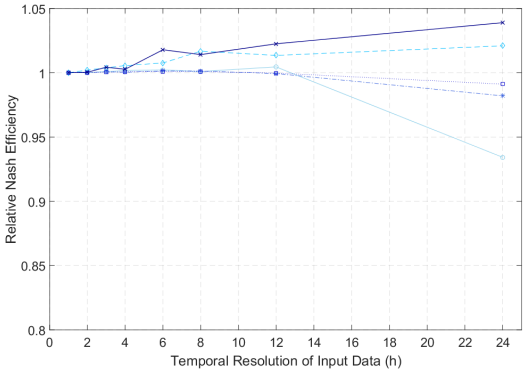
(D) Absolute Nash efficiency for spring



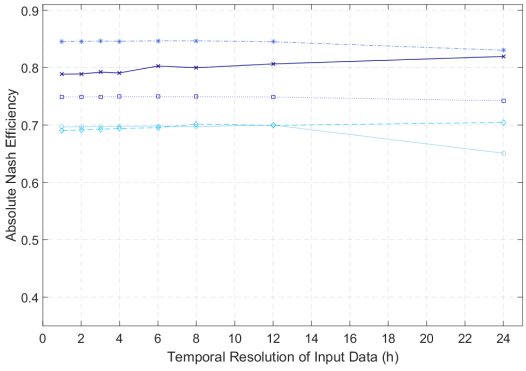
(E) Normalized Nash efficiency for summer



(F) Absolute Nash efficiency for summer



(G) Normalized Nash efficiency for autumn



(H) Absolute Nash efficiency for autumn

FIGURE 5.4: Nash efficiency for each season

the season in which the Nash efficiency had the smallest decrease or the highest increase was the winter season and the season in which the Nash efficiency has the highest decrease or the smallest increase is the spring season, but the difference with others seasons was negligible. In this case, like in the previous cases, the behavior of the Nash efficiency could be classified into three different groups, which are the same groups in which the basins can be classified when looking at its size (small, medium and large basins). This indicates the existence of a relation between the size of the study basin and the required temporal resolution of hydrological input data in hydrological simulation.

5.3.4 Monthly results

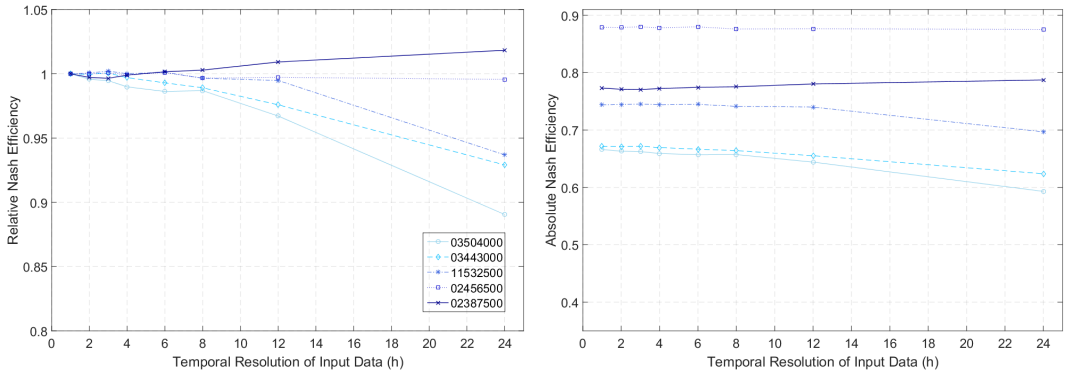
In this section of the study, there are the results and conclusions obtained when computing the Nash efficiency when only considering data from a single month. The months selected for this study were the months of highest intensity of rainfall because the Xinanjiang model it is mainly used for flood forecasting, so it is interesting to see the impact of temporal resolution of input data on the XAJ model performance during flood events. The months selected for this study for each basin can be seen in Table 5.2. The values presented in Figures 5.5a and 5.5b are for the months of highest intensity rainfall of each basin and the values presented in Figures 5.5c and 5.5d are for the months of second highest intensity rainfall.

In this case, the values obtained for the Nash efficiency when the model was running at a hourly timescale were lower than the ones obtained in previous cases. This was due to the fact that the Nash efficiency coefficient is sensitive to extreme values and during months of high rainfall is when the streamflow discharge takes the most extreme values. The range of values of the Nash efficiency went from 0.44 to 0.91, but this was because the performance of basin 03504000 during the month with second highest intensity rainfall was not very good. If this low value had not been not considered, then the range of values of the Nash efficiency would have been between 0.65 to 0.91, which can be considered a value that indicates a good performance (Moriasi et al., 2007).

Looking at Figure 5.5 it can be seen that, like in the previous cases, there was a decrease in the quality of the simulation when the time interval of the input data was increased. The difference with the previous cases was that when only the results of one month were considered, the decrease of the Nash efficiency was faster. In some cases, the decrease of the Nash efficiency when using data with a 24-hour interval reached values close to 80% of the Nash efficiency when using input data with a 1-hour interval.

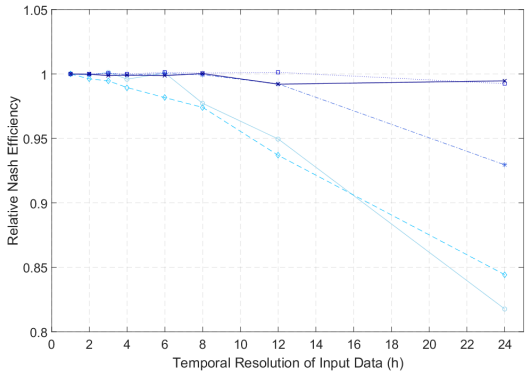
Another difference with previous cases, was that in this case the decrease of the Nash efficiency started to happen at a higher time resolution. When analyzing simulated data for a period of several months of years, the lost of quality experienced in the result of the simulation when using data with low temporal resolution was not very significant until using data with a 24-hour interval. In the previous cases, the values obtained for the quality indicators were pretty similar when data with resolutions of 1, 2, 3, 4, 6, 8 and 12 hours were used and an important decrease of the Nash efficiency was not experienced until data with a 24-hour interval was used, whereas in this case the value of the Nash efficiency started to experience a noticeable decrease around the 8 or 12-hours time interval.

The decrease in the quality of the simulation was strongly correlated with the size of the study basin. The behavior experienced in this part of the study was that

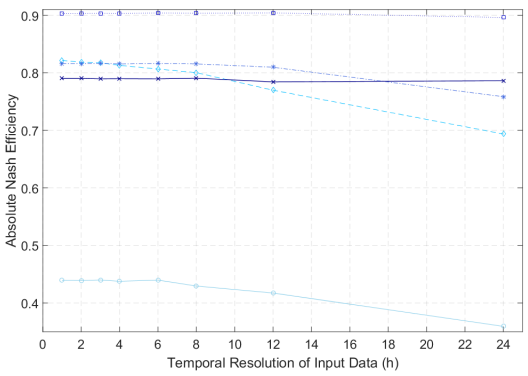


(A) Normalized Nash efficiency

(B) Absolute Nash efficiency



(C) Normalized Nash efficiency



(D) Absolute Nash efficiency

FIGURE 5.5: Nash efficiency for months of intense rainfall

the smaller the basin was, the more important was the decrease of the Nash efficiency. The biggest basins studied experienced a low decrease in the quality of the simulation, usually around 1 and 5%, the Nash efficiency of medium size basins decreased around 5 and 10% and with respect of the original value and the smallest basins experienced a lost of quality of the simulation between 10 and 20% in the worst cases.

Analyzing the results obtained it was been seen that the smallest study basins experienced a noticeable decrease in the quality of the simulation, whereas in the biggest basins this decrease was not as important. This means that the smaller the basin is, the more important is to have data with good temporal resolution in order to be able to do a good simulation. If it is necessary to carry out simulations for a period of several months or years, then it may be acceptable to use data with a temporal resolution of 12 hours, but when the period studied is of one months, then the temporal resolution of the input data has to be higher because the decrease of the Nash efficiency starts to be noticeable for data with 8-hour temporal resolution, especially in small basins with a size smaller than 800 km^2.

5.3.5 Weekly results

In this section, the results obtained when computing the Nash efficiency when only using the data from a single week with high intensity rainfall will be presented and discussed. As in the previous case, the goal was to analyze the performance of the XAJ model during flood events, and that is why weeks with high intensity rainfall were selected. The weeks used for each basin in this study can be seen in Table 5.3.

TABLE 5.2: Months with highest intensity rainfall of each basin

	Month with highest intensity rainfall	Month with second highest intensity rainfall
03504000	October of 1995	February of 1995
03443000	August of 1994	December of 1996
11532500	December of 1996	December of 1995
02456500	October of 1995	March of 1996
02387500	February of 1996	February of 1995

The values of the Nash efficiency that are shown in Figures 5.6a and 5.6b are for the weeks of the study period with the highest intensity rainfall of each basin and the values presented in Figures 5.6c and 5.6d are for the weeks with the second most intense rainfall of each basin.

When the results obtained for a single week using input data at a hourly timescale were analyzed, the range of values of the Nash efficiency obtained was between 0.45 and 0.90. As it was explained before, even though this range of values usually could be considered too low in some cases, it is necessary to consider that data from intense flood events is being analyzed and that the Nash efficiency coefficient is sensitive to extreme values. Almost all values of the Nash efficiency are above 0.5, and some researchers, like the ones at USGS, feel that 0.5 and above represents a good fit for stream-flow conditions (Christiansen, 2012). When data from only a week is considered, if the values of the simulated peak stream-flow and the observed peak stream-flow are not similar, even though if the rest of the result has a high correlation with the observed stream-flow, the Nash efficiency will have a low value.

The Nash efficiencies obtained in this part of the study are presented in Figure 5.6. In can be seen that in these part of the study the highest decreases in the value of the Nash efficiency were experienced, even more than in the monthly case. In the previous case some basins experienced a very small decrease of the Nash efficiency, in some it even increased, but when doing a weekly analysis, the quality of the performance of the model was considerably reduced in all basins during the weeks analyzed in this study compared to the performance obtained using input data with a 1-hour interval.

Similarly to what was seen when analyzing the monthly results, in this case the decrease of the Nash efficiency happened earlier than in other cases. The difference with the monthly case was that when analyzing the flood event of just a week, the decrease happened even with a higher temporal resolution and it was more considerable than in previous cases. In many of the study basins, the value of the normalized Nash efficiency when analyzing the results from just a week using input data with a 8-hour resolution was the same that was obtained when analyzing the results from a whole month using input data with a 12-hour resolution. In some cases even a greater lost of quality was experienced. For example, there was a case (basin 03504000 in Figure 5.6c) in which a decrease of 5 % happened even when using input data with a 6-hour resolution, when in previous cases this decrease didn't happen until input data with a 12-hour or 24-hour resolution was used.

As it has been seen in all previous cases, the smaller the study basin is, the more considerable is the lost of quality when the time resolution of the input data is reduced. In both of the weeks analyzed, the basins that experienced the highest decrease of the Nash efficiency were the ones with an area of smaller than 800 km^2. In both cases, the normalized value of the Nash efficiency was between 0.8 and 0.85,

TABLE 5.3: Weeks with highest intensity rainfall of each basin

	Week with highest intensity rainfall	Week with second highest intensity rainfall
03504000	01/10/1995 - 07/10/1995	11/02/1995 - 17/02/1995
03443000	15/08/1994 - 21/08/1994	04/10/1995 - 10/10/1995
11532500	05/01/1997 - 11/01/1997	08/01/1995 - 14/01/1995
02456500	01/03/1996 - 07/03/1996	03/10/1995 - 10/10/1995
02387500	03/10/1995 - 10/10/1995	04/03/1996 - 10/03/1996

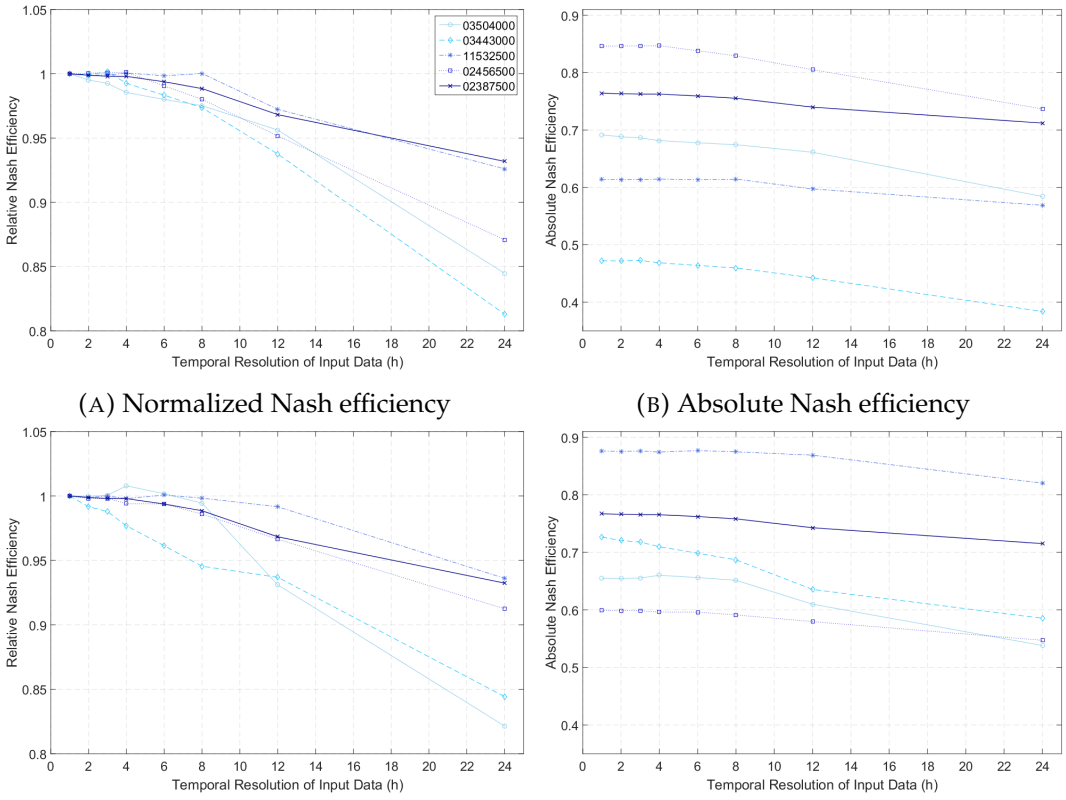
which are the lowest values obtained in this study. Also, the biggest basin was the one that usually has the most stable Nash efficiency when input data with different time resolutions was used. The main difference with respect the previous cases was that now the Nash efficiency always decreased and it reached values below 0.95 in both of the weeks analyzed, whereas in previous cases it didn't reach values this low, or it even sightly increased its value.

The variation of the range of values of the normalized Nash efficiency in this case compared with the previous case goes as follows: for the small basin the range of values went from 0.99 to 0.95 in the previous case to 0.95 to 0.90 in the current case, for the medium basins went from 0.95 to 0.90 in the previous case to 0.95 to 0.85 in the current case and for the biggest basin went from 0.90 to 0.80 in the previous case to 0.85 to 0.80 in the current case.

Like in previous cases, the main conclusion that can be extracted from this part of the study was that the size of the study basin plays an important role on the impact of temporal resolution of input data on model performance. After analyzing how the temporal resolution of the input affects the value of the Nash efficiency during a 1-week long flood event it has been seen that a lower time resolution of the input data always results in a decrease of the quality of the simulation. This decrease is strongly correlated to the size of the study basin. The smallest study basins are always the ones that have the lowest value of the Nash efficiency when the temporal resolution is decreased. Also, the decrease of the Nash efficiency happens faster in the smallest basins. This can be seen in the fact that the value of the normalized Nash efficiency that big basins have when data with a 24-hour time resolution is used is similar to the one that small basins have when data with a 8-hour time resolution is used. This means that is important to have data with good temporal resolution when carrying out the simulation of short flood events, specially if the basin small.

5.3.6 Low intensity rainfall results

Finally, it was decided to also analyze the impact of temporal resolution of input data on model performance during periods of low intensity rainfall. Even tough the Xinanjiang model is mainly used for flood forecasting, it also has several other applications like water resources estimation, field drainage and water quality accounting, and that is why it is interesting to see the impact of temporal resolution of input data on the XAJ model performance during periods of low intensity rainfall. Two different dry periods were analyzed in each basin, the length of which was 3 months. The periods selected in each basin for this part of the study are presented in Table 5.4. The values shown in Figures 5.7a and 5.7b are for the first dry period analyzed in each basin and the values presented in Figures 5.7c and 5.7d are for the months of the second dry period analyzed.



(A) Normalized Nash efficiency

(B) Absolute Nash efficiency

(C) Normalized Nash efficiency

(D) Absolute Nash efficiency

FIGURE 5.6: Nash efficiency for weeks of intense rainfall

The range of values obtained for the Nash efficiency when the model was using data with a 1-hour temporal resolution during periods of low intensity rainfall goes from 0.80 to 0.45. These values are lower than it would be desirable, but this was due to the fact that during the calibration stage the main goal was to obtain a good correlation between the peaks of the observed and simulated streamflow discharge and not as much attention was paid to the months when the streamflow discharge was low. This was done in order to obtain a high global value of the Nash efficiency because it is sensitive to extreme values.

Like in the previous cases, analyzing Figure 5.7, in which the evolution of the Nash efficiency when input data with different temporal resolutions was used is presented, it can be seen that when data with a low temporal resolution was used the results of the simulation are worse than the ones obtained when data with a high temporal resolution was used. The difference with previous cases was that now it was not possible to see the relation between the decrease of the Nash efficiency and study basin size.

Basin 03504000, like in almost all previous cases, is the one that experienced a higher decrease of the Nash efficiency when the time resolution of the input data was 24 hours, but opposite to what it would be expected, the basin with the second lowest value of the normalized Nash efficiency was basin 02387500, which was the biggest one. Even though basin 03504000 had a poorer performance than basin 02387500 when using input data with a 24-hour time resolution, the basin with the lowest values of the Nash efficiency when input data with resolutions of 1, 2, 3, 4, 6, 8 and 12 hours were used was basin 02387500, which was the opposite that happened

TABLE 5.4: Periods with low intensity rainfall studied in each basin

	First dry period analyzed	Second dry period analyzed
03504000	May of 1995 - July of 1995	July of 1996 - September of 1996
03443000	July of 1996 - September of 1996	July of 1997 - September of 1997
11532500	July of 1996 - September of 1996	July of 1997 - September of 1997
02456500	September of 1994 - November of 1994	August of 1997 - October of 1997
02387500	June of 1995 - August of 1995	September of 1996 - November of 1996

in previous cases.

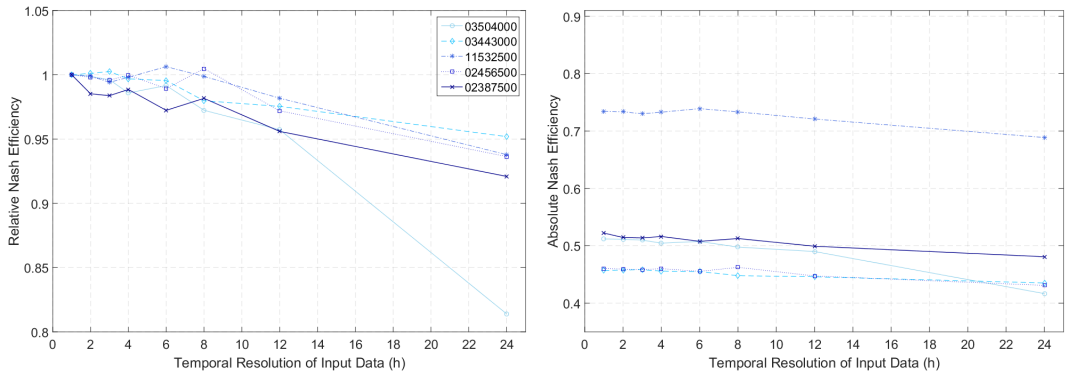
Out of all the study basins, the one that had a better simulated results when low time resolution data was used was basin 11532500, which was the one whose size was the median, 1577 km^2. The other two medium size basins also show a similarly good behavior of the Nash efficiency when input data of any time resolution is used.

After commenting the results obtained for the low intensity rainfall periods, it was difficult to extract a conclusion related to the size of the study basins. It is true that the three medium size basins present a similar behavior and can be grouped, but the fact that basins 03504000 and 02387500 show a similar behavior was not expected and it is a result that does not fit with all the results previously obtained. Also, the fact that basin 02387500 had the worst performance out of all basins when input data with resolutions of 1, 2, 3, 4, 6, 8 and 12 hours were used was an unexpected result. Since, unlike previous cases, the behavior of the Nash efficiency could not be classified into different groups according to the size of the study basin, this seems to indicate that during periods of low intensity rain the impact of temporal resolution of input data on model performance of the size of the study basin is low. Still, the impact of temporal resolution has to be considered because in some of the basins the normalized Nash efficiency decreased to values around 0.90, and a decrease of 10% in the quality of the performance of the model is not negligible, so the temporal resolution of hydrological input data is still important and necessary to be considered during periods of low intensity rainfall.

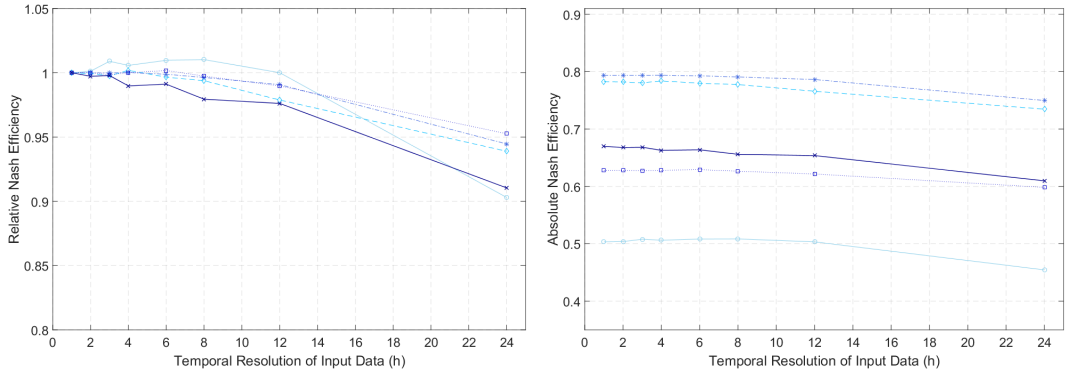
5.3.7 Different starting time results

All the results shown previously were obtained using input data with a low temporal resolution created using the procedure explained in Section 5.2. When using this procedure, it was necessary to chose at which hour of the day the low temporal resolution data set will begin because different starting times will produce different hydrographs, as it can be seen in Figure 5.8. The results shown until now were obtained using data whose starting time was midnight (00:00), but in this part of the study, the influence of the hour at which low temporal resolution data is measured will be analyzed. In order to do so, some of the simulations previously shown will be carried out with data with different starting times in order to be able to compare their accuracy.

In Figure 5.9, the results obtained for the absolute Nash efficiency in basin 03504000, the smallest one, using input data created with different starting times are presented. The starting times used were 00:00, 01:00, 02:00, 03:00 and 04:00. As it can be seen, all the results obtained were pretty similar. The difference between the starting time that yielded the best Nash efficiency and the one with the worst Nash efficiency was

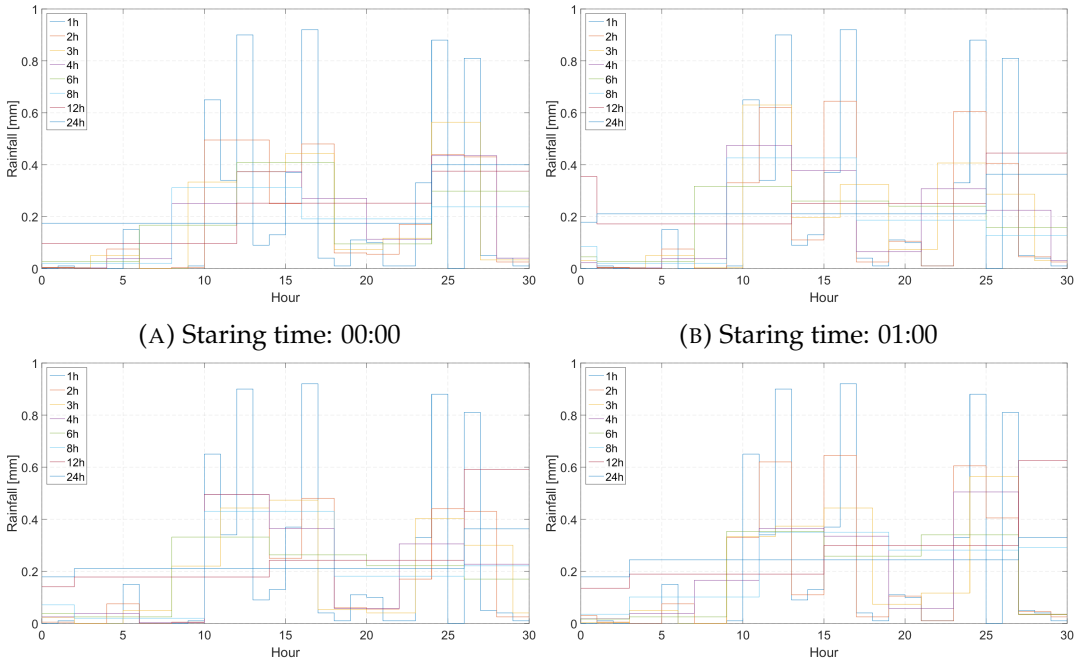


(A) Normalized NASH efficiency (B) Absolute NASH efficiency



(C) Normalized NASH efficiency (D) Absolute NASH efficiency

FIGURE 5.7: NASH efficiency for periods of low-intensity rainfall



(A) Staring time: 00:00 (B) Staring time: 01:00

(C) Staring time: 02:00 (D) Staring time: 03:00

FIGURE 5.8: Hydrographs obtained using different starting times

usually around 0.01. During the period of low intense rainfall was when the biggest difference was reached, and its value was 0.025.

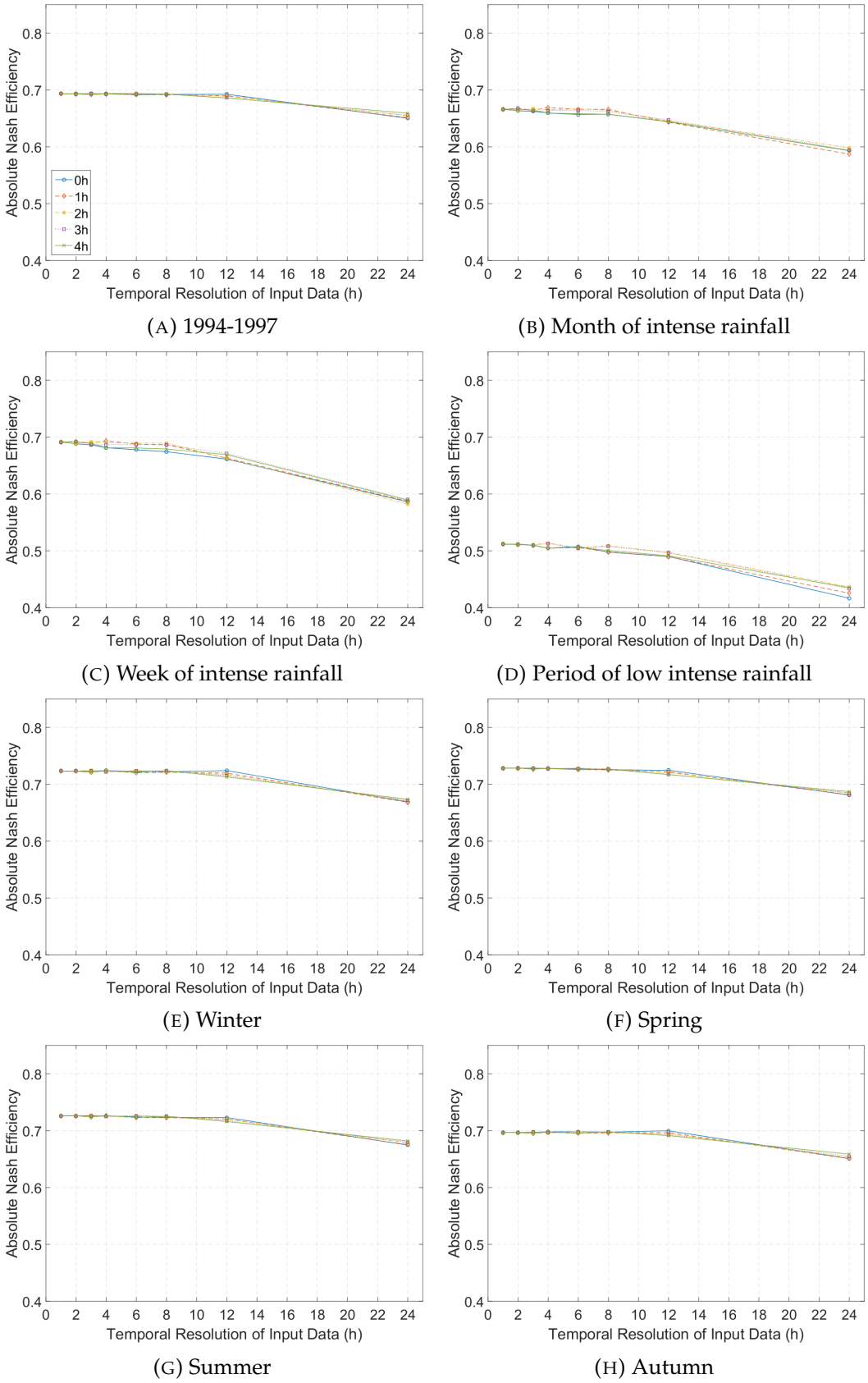
Also, it was not possible to see a trend that indicated which starting time was the best because depending on the case that was being analyzed and the time resolution of the input data used, the starting time that yielded a better result changed. For example, when looking at the results from 1994 to 1997 in Figure 5.9a, with a time resolution of 12 hours the best result was obtained with the starting time 00:00, but with a time resolution of 24 hours the starting time 00:00 yielded the lowest value of the Nash efficiency.

In Figure 5.10, the results obtained for the absolute Nash efficiency in basin 02387500, the biggest one, are presented in order to be able to see if the difference in basin size had an influence on the results. The results obtained and conclusions that can be extracted from them are very similar to the ones from basin 03504000. In this case, the difference between the best and the worst Nash efficiency was usually less than 0.01, and the biggest difference was 0.018 and it happened during the week of most intense rainfall. Like in the previous case, it was not possible to see which starting time was the best because it depended on the case that was being analyzed and the time resolution of the input data used.

Having analyzed the variation of the absolute Nash efficiency using different starting times for the input data, the main conclusion extracted is that the impact of the initial hour of the input data does not have a big influence on the model performance. Looking at the results obtained in the basins analyzed, the values of the absolute Nash efficiency are very similar no matter which starting time is used, and therefore it can be concluded that it has a very small influence on the result.

5.4 Summary

In this chapter, the behavior of the Nash efficiency when using input data with different time resolution is analyzed under several different cases. Conclusions about the influence that the size of the study basin has on the required temporal resolution of the input are presented.



(A) 1994-1997

(B) Month of intense rainfall

(C) Week of intense rainfall

(D) Period of low intense rainfall

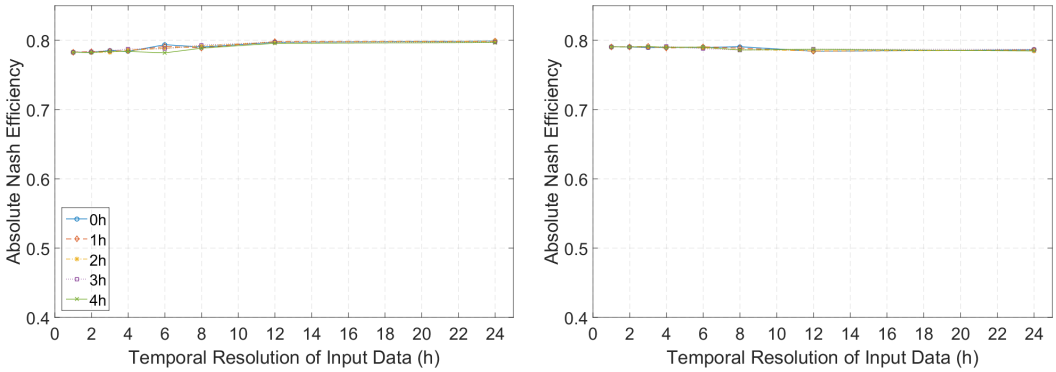
(E) Winter

(F) Spring

(G) Summer

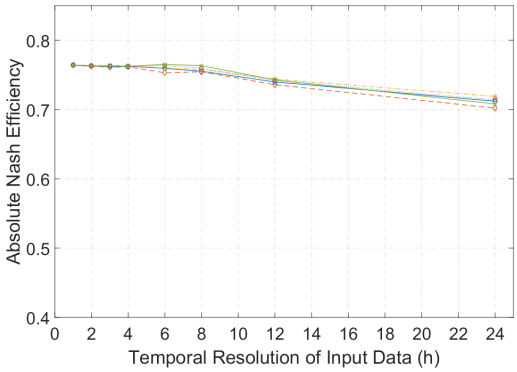
(H) Autumn

FIGURE 5.9: Absolute Nash efficiency obtained in basin 03504000 using different starting times

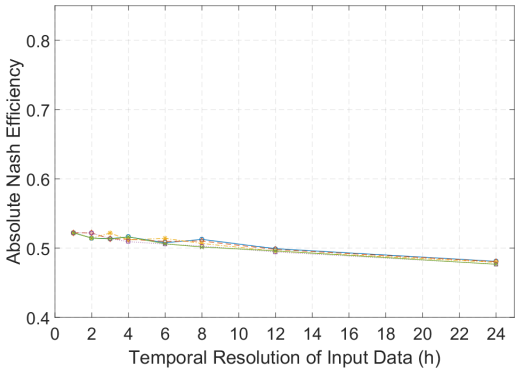


(A) 1994-1997

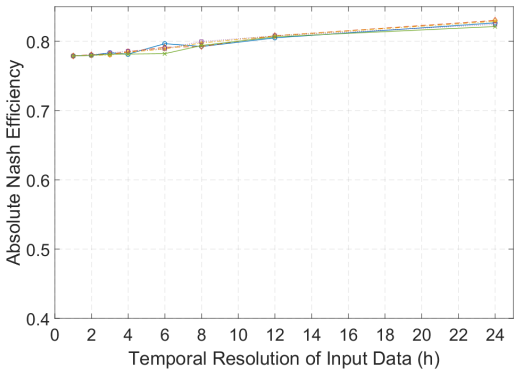
(B) Month of intense rainfall



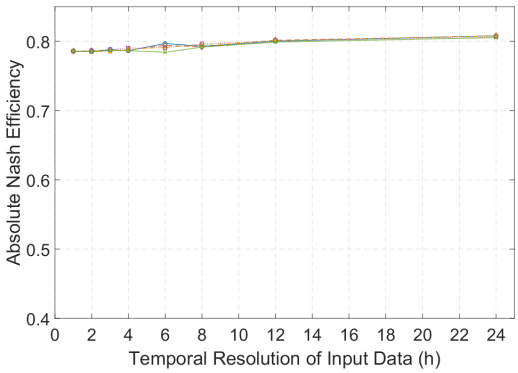
(C) Week of intense rainfall



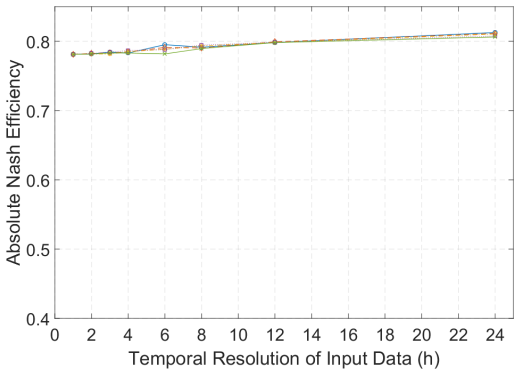
(D) Period of low intense rainfall



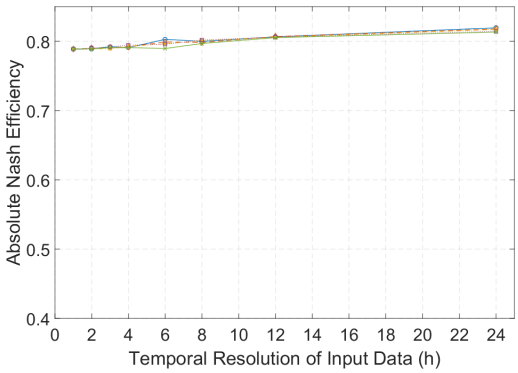
(E) Winter



(F) Spring



(G) Summer



(H) Autumn

FIGURE 5.10: Absolute Nash efficiency obtained in basin 02387500 using different starting times

Chapter 6

Impact of temporal resolution of input data on model outputs

6.1 Introduction

Once the influence of temporal resolution of the input data on the performance of the model has been observed and quantified, it is necessary to understand why this happens and which of the parts of the model described in Chapter 2 are affected the most by the temporal resolution of the input data.

Knowing how the temporal resolution of the input data affects the structure of the XAJ model can be useful when trying to improve the model to make it less sensitive to the temporal resolution of the input data.

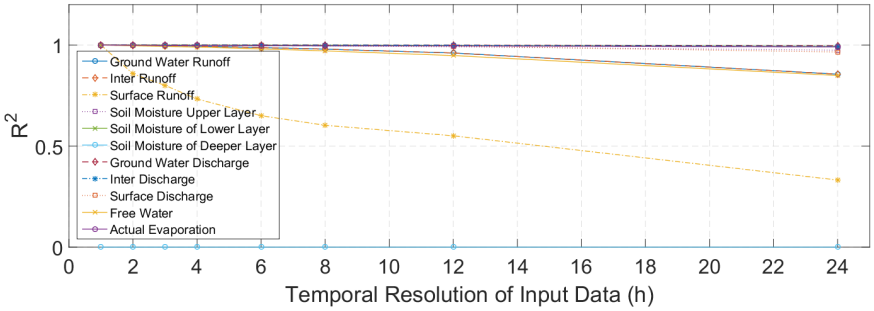
6.2 Temporal resolution influence on the outputs of the XAJ model

In the results file of the XAJ model, apart from the streamflow, there are also several other values that are also outputs of the model. In this part of the study the objective was to analyze how this outputs changed in order to check which parts of the XAJ model were more affected by the temporal resolution of the input data and were responsible for the decrease of the performance of the model when using data with low temporal resolution.

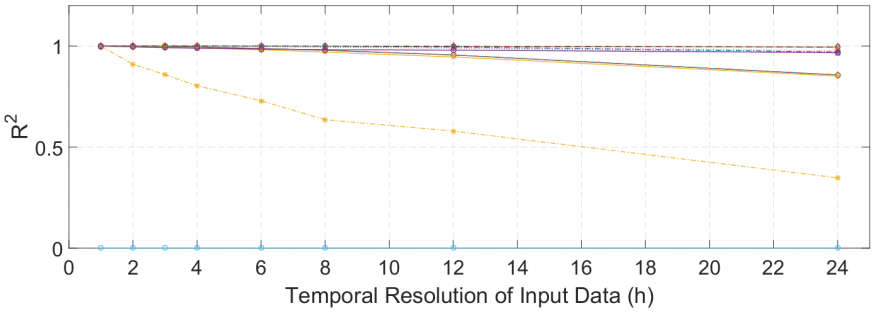
The values of the outputs obtained using input data with 1-hour temporal resolution were compared with the values of the outputs obtained using data with a lower temporal resolution to check which values changed the most when input data with different temporal resolution was used. The performance indicator used to check how different were the results was R^2. In Figure 6.1 the decrease in the quality of the performance indicator R^2 is plotted for all the output values of the 5 study basins.

Looking at Figure 6.1, it can be seen that the behavior of the different outputs of the model can be classified in three different groups: virtually no variation of the value (decrease of R^2 less than 0.05), small variation of the value (decrease of R^2 usually between 0.1 and 0.2) and high variation of the value (decrease of R^2 higher than 0.5). The classification depends on the size of the basin, but is almost the same for all basins. In the two smallest basins the classification is the one presented in Table 6.1 and the classification of the intermediate values in the 3 biggest basins is presented in Table Table 6.2. The only difference in both classifications is that the quality of the simulation of the actual evaporation is worse in the biggest basins than in the smallest basins.

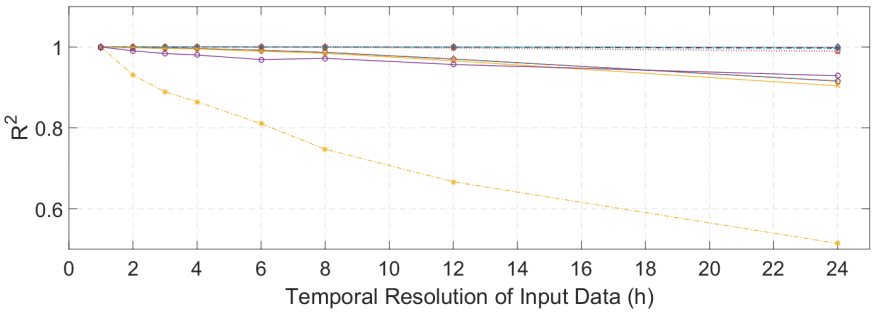
The fact that the values that were affected the most by the time resolution of the input data were the ones related to the runoff (Ground Water Runoff, Inter Runoff



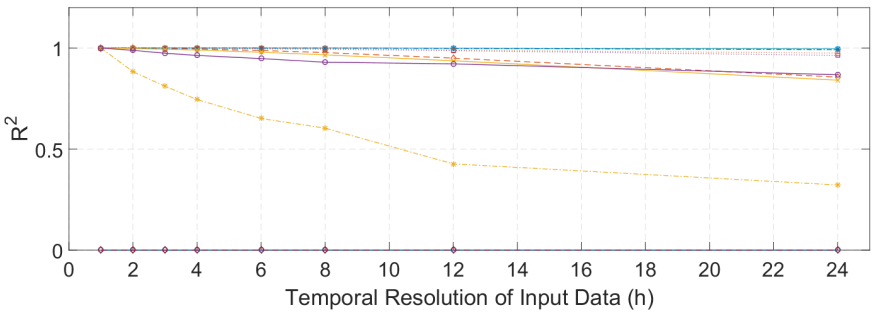
(A) Basin 03504000



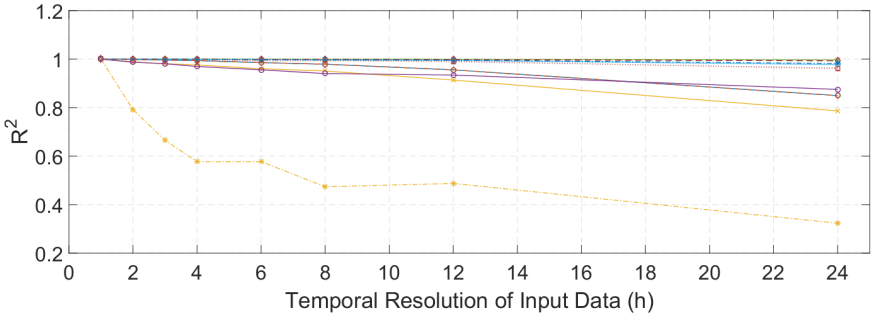
(B) Basin 03443000



(C) Basin 11532500



(D) Basin 02456500



(E) Basin 02387500

FIGURE 6.1: R^2 of the outputs of the model using different temporal resolutions

TABLE 6.1: Classification of the intermediate values in the smallest basins

No variation	Soil Moisture Upper Layer, Soil Moisture Lower Layer, Soil Moisture Deeper Layer, Ground Water Runoff, Inter Runoff, Surface Runoff Actual Evaporation
Small variation	Ground Water Runoff, Inter Runoff, Free Water
High variation	Surface Runoff

TABLE 6.2: Classification of the intermediate values in the biggest basins

No variation	Soil Moisture Upper Layer, Soil Moisture Lower Layer, Soil Moisture Deeper Layer, Ground Water Runoff, Inter Runoff, Surface Runoff
Small variation	Ground Water Runoff, Inter Runoff, Free Water, Actual Evaporation
High variation	Surface Runoff

and Surface Runoff) was a sign that the runoff generation part of the model is the part most affected by the temporal resolution of the input data. Once that the most affected part was known, it was decided to investigate further on how this part of the model was affected by the temporal resolution of the input data.

6.3 Partitioning of surface and subsurface runoff from pervious area

To check the relationship between surface and subsurface runoff, a new parameter α was introduced. Knowing that the formulas used to compute the surface and subsurface runoff are:

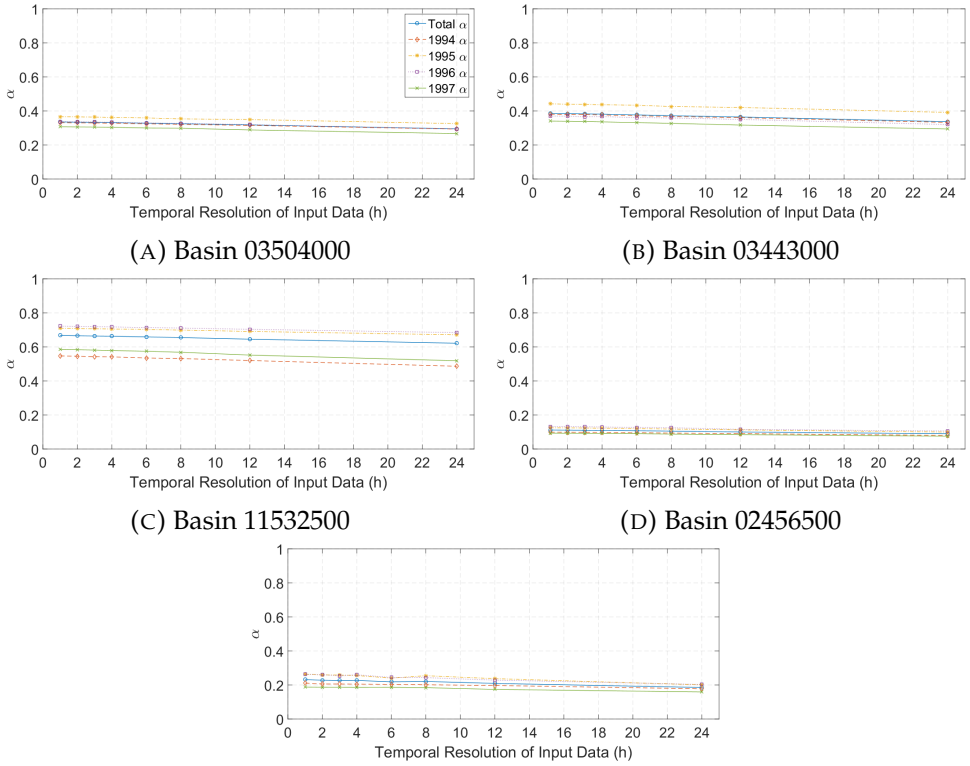
$$R_s = P \cdot imp + P \cdot \left(\frac{f}{F} - imp\right) \cdot \frac{g}{G} \tag{6.1}$$

$$R_{ss} = P \cdot \left(\frac{f}{F} - imp\right) \cdot \left(1 - \frac{g}{G}\right) \tag{6.2}$$

where P is precipitation, imp is fractional area of impervious area, $\frac{f}{F}$ fractional area where tension water storage is filled and $\frac{g}{G}$ is the fraction of $\frac{f}{F}$ where free water capacity is filled. Defining $\alpha = \frac{g}{G}$ and $R_s^{'} = R_s - P \cdot imp$, it is possible to derive:

$$\alpha = \frac{R_s^{'}}{R_s^{'} + R_{ss}} \tag{6.3}$$

It was necessary to compute α in this way to take into account the fact that each basin had a fraction of impervious surface on which all rain becomes runoff. In Figure 6.2, the values of α obtained on each basin for each year using input data with different temporal resolution are presented. It was possible to see that all basins followed a similar trend: there was a decrease of the value of α, which indicated that the total volume of Surface Runoff decreased while the total volume of Subsurface Runoff (Ground Water Runoff and Inter Runoff) increased. Even though the values of α were different depending on the year and on the study basin, the value of the decrease of α was similar for every year and it was around 10%.



(A) Basin 03504000

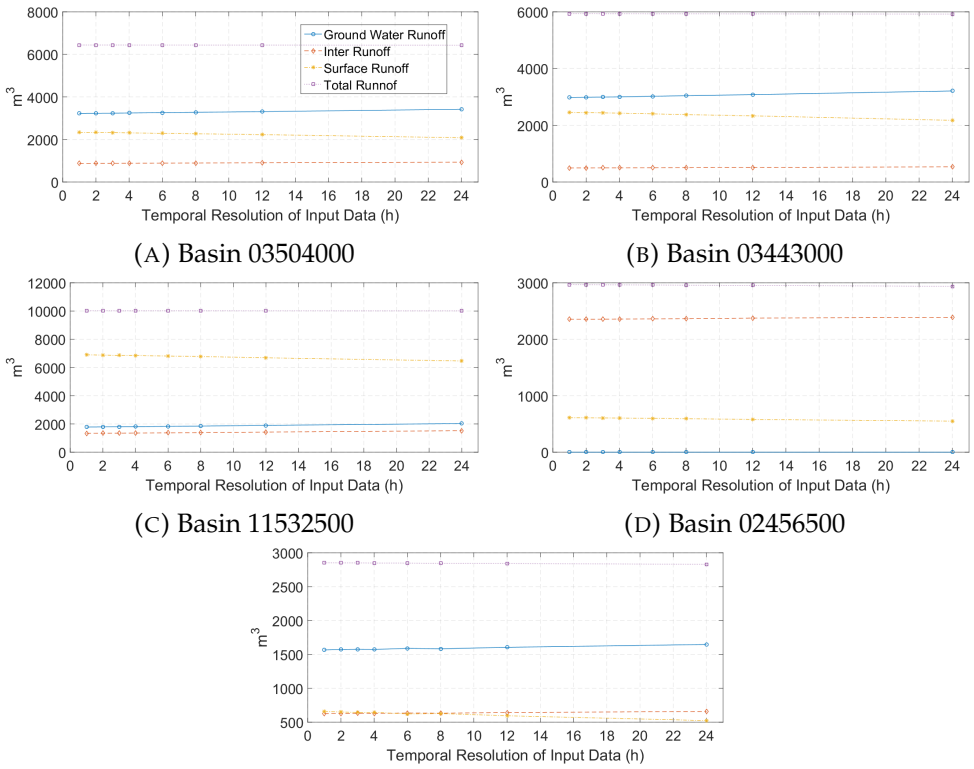
(B) Basin 03443000

(C) Basin 11532500

(D) Basin 02456500

(E) Basin 02387500

FIGURE 6.2: Values of α using different temporal resolutions



(A) Basin 03504000

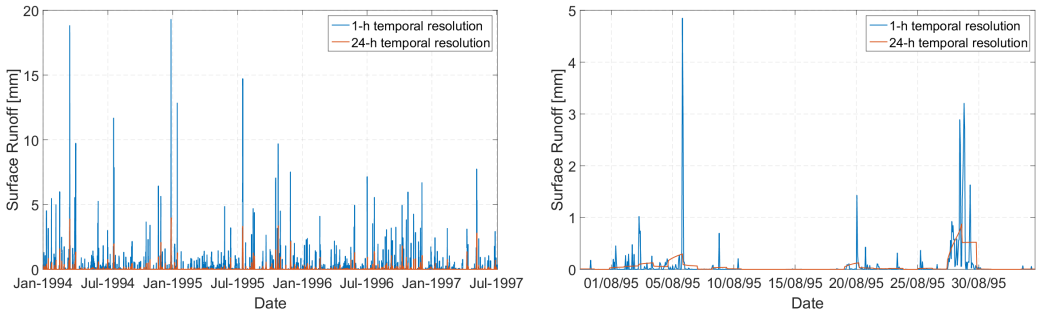
(B) Basin 03443000

(C) Basin 11532500

(D) Basin 02456500

(E) Basin 02387500

FIGURE 6.3: Values of the runoff using different temporal resolutions



(A) From 1994 to 1997 (B) August of 1995

FIGURE 6.4: Evolution of the surface runoff in basin 03504000

In Figure 6.3 the total volume of the different types of runoff is presented. It can be seen that the Ground Water Runoff and the Inter Runoff increase and the Surface Runoff decreases when the temporal resolution of the input data is reduced. The total volume of runoff is not affected by the temporal resolution because the reduction of Surface Runoff is compensated by the increase of Subsurface Runoff in all basins.

In Figure 6.4 it can be seen why the total volume of surface runoff decreases when using data with low temporal resolution. The peaks in precipitation that are present in data with high temporal resolution cannot be absorbed by the soil and they all become surface runoff, whereas the precipitation in the data with low temporal resolution is averaged and doesn't have those peaks and, therefore, it can be absorbed better by the soil, becoming subsurface runoff instead of surface runoff.

Once it was known that using input data with low temporal resolution causes a decrease in the quality of the simulation because it underestimates the quantity of surface runoff generated, it was decided to check if it was possible to compensate this decrease of the surface runoff during the calibration stage to obtain results of the same quality as the ones obtained using data with high temporal resolution, but using input data with low temporal resolution.

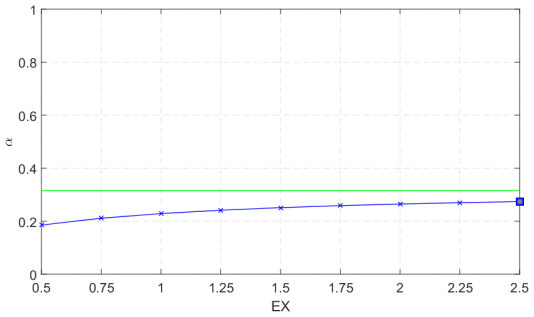
6.4 Recalibration of the parameters of the model

In this section of the study was decided check if it was possible to modify the parameters obtained in the calibration to make the result of the simulation using input data with a 24-hour temporal resolution more similar to the result obtained using input data with a 1-hour temporal resolution.

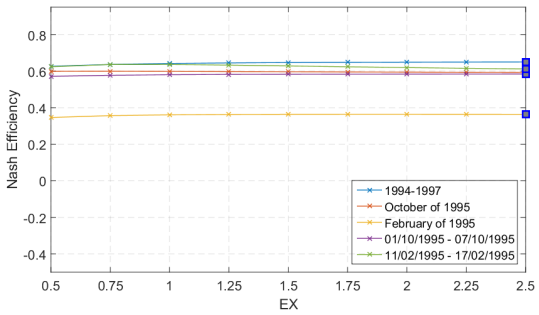
6.4.1 Recalibration of EX

The first parameter that was thought that should be modified was EX because it is the exponent of the free water capacity curve and therefore it is the parameter that controls the runoff separation.

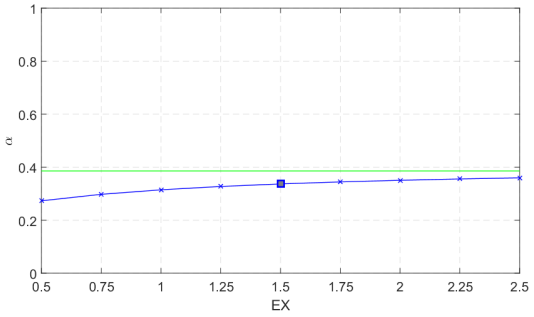
In the previous section it has been seen that the most affected part of the model by the temporal resolution of input data is the runoff separation part. Therefore, the objective of modifying EX was to obtain a similar runoff separation that the one that was obtained when using input data with 1-hour temporal resolution, hoping that if a similar runoff separation was obtained then the quality of the simulation would increase.



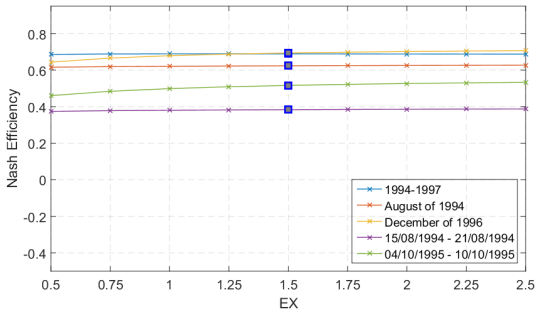
(A) α variation in 03504000



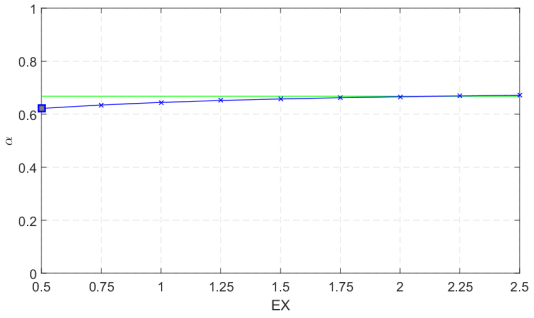
(B) Nash efficiency variation in 03504000



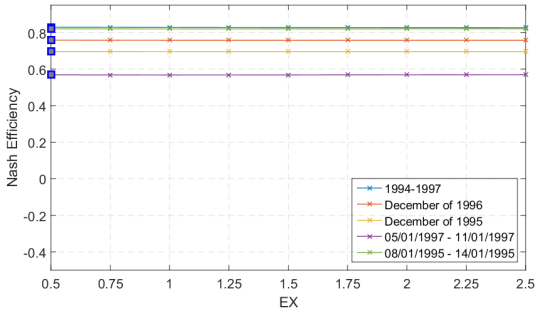
(C) α variation in 03443000



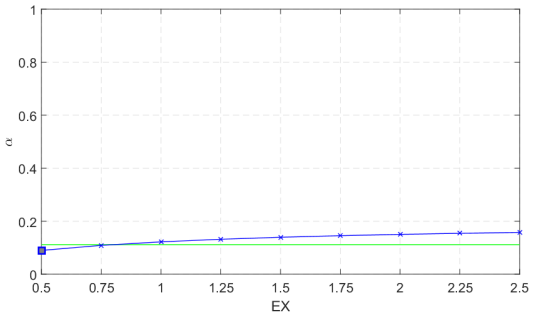
(D) Nash efficiency variation in 03443000



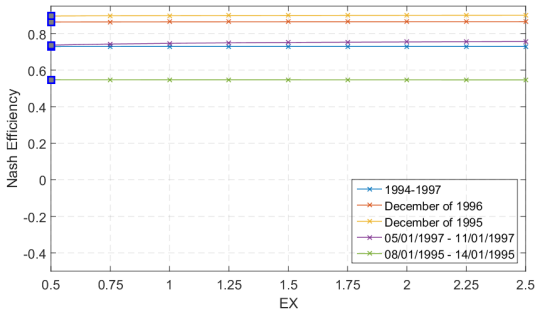
(E) α variation in 11532500



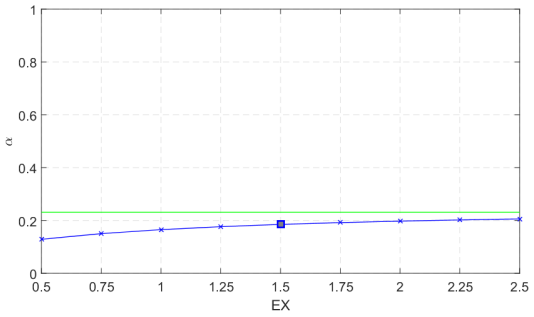
(F) Nash efficiency variation in 11532500



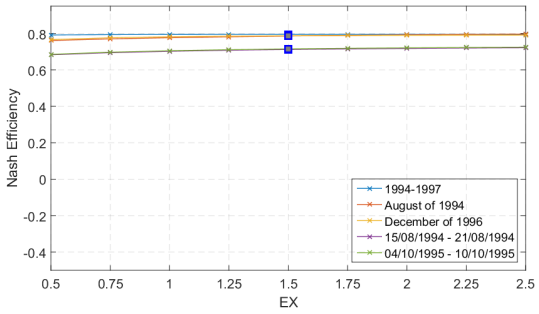
(G) α variation in 02456500



(H) Nash efficiency variation in 02456500



(I) α variation in 02387500



(J) Nash efficiency variation in 02387500

FIGURE 6.5: Recalibration of the model using different values for EX

The results obtained after running the model with different values of EX using input data with daily temporal resolution can be seen in Figure 6.5. The blue dot in Figure 6.5 indicates the value of EX that was obtained in the calibration using data with 1-hour temporal resolution and that was used in the validation. The green line in Figures 6.5a, 6.5c, 6.5e, 6.5g and 6.5i indicates the value of α obtained using data with 1-hour temporal resolution and therefore it is the value that α that was thought to improve the simulation.

As it can be seen in Figure 6.5, using higher values EX yield higher values of α, but in some basins it was not possible to obtain the desired value of α using the recommended range of values of EX, but increasing the value of EX always results in a α closer to the desired one. As it can be seen in Figures 6.5b, 6.5d, 6.5f, 6.5h and 6.6j the value of the Nash efficiency is not very sensitive to EX and the biggest difference of the Nash efficiency obtained using different values of EX is below 0.08.

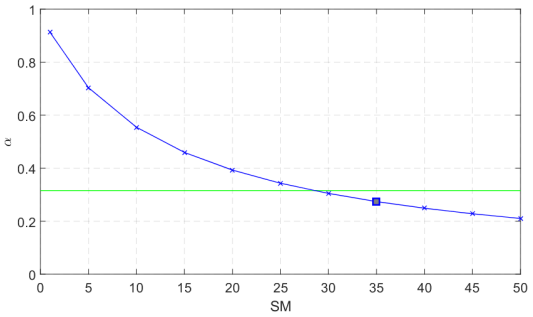
After performing this part of the study it was seen that it is difficult to improve the quality of the simulation using different values of EX because of the insensitivity of the Nash efficiency, and also because some of the results obtained contradicted the initial hypothesis. For example, in basin 02456500, the EX that was expected to improve the simulation the most was 0.8, but has it can be seen in Figure 6.5h, the best results are obtained when EX is equal to 2.5. Another problem is what happens in basin 03504000, that whereas in some cases the quality of the simulation is increased, in the others it decreases when using a new value for EX. This also happens in basins 11532500, but it is not as noticeable because in that basin the results are very insensitive to EX. On the other hand, in basin 02387500, using a value of EX of 2.5, which is the one that yielded a value of α closest to the desired one, improved the quality of the simulation in all cases. This different behavior of the results makes it difficult to obtain general conclusions on how to improve the quality of the simulation using different values of EX when the input data has a low temporal resolution.

6.4.2 Recalibration of SM

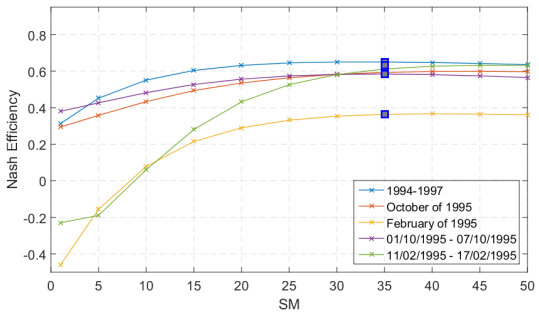
After trying to improve to quality of the simulation by changing the value of EX, the same procedure was followed with the parameter SM, which is the areal mean of the free water capacity of the surface soil layer.

The results obtained after running the model with different values of SM, while keeping EX constant, using input data with daily temporal resolution can be seen in Figure 6.6. Comparing the results obtained in this case with the ones from Figure 6.5, it can be seen that values of α and the Nash efficiency are more sensitive to SM than to EX. In this case, decreasing the value of SM is necessary to get the same value of α that was obtained using data with a temporal resolution of 1 hour. The amount in which should be decreased to reach the desired value of α is between 5 and 10 mm in all the study basins.

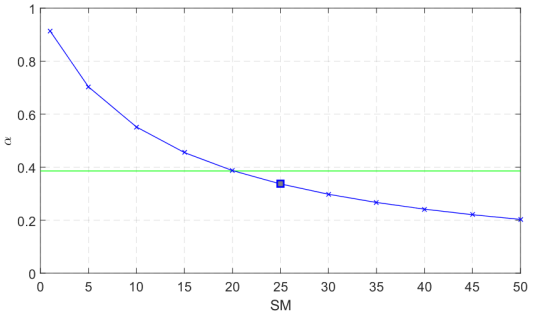
Updating the value of SM to the one that yields the same α that was obtained using input data with 1-hour temporal resolution usually increased the value of the Nash efficiency, but the improvement was very small. In some basins like 03443000 and 11532500, as can be seen in Figures 6.6d and 6.6f, even though the updated value of SM improved the simulation for the flood events, the value of the Nash efficiency decreased for the 4-year period. Like in this previous case, the fact that the values of the Nash efficiency didn't increase very much makes it difficult to significantly improve the quality of the simulation using different values of SM when the input data has a low temporal resolution.



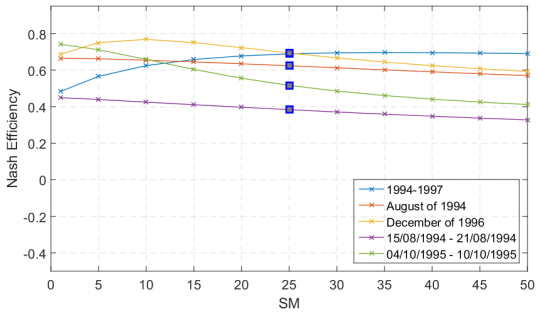
(A) α variation in 03504000



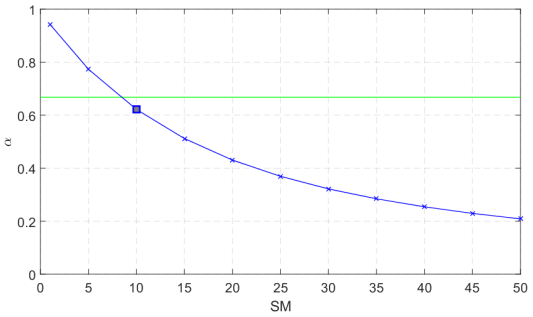
(B) Nash efficiency variation in 03504000



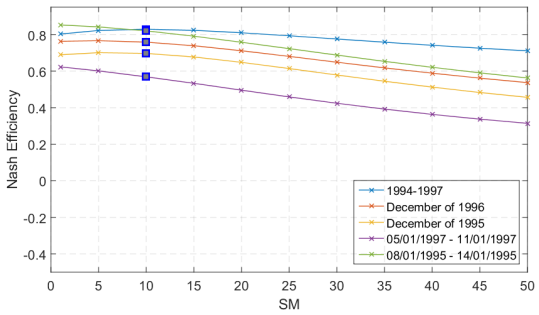
(C) α variation in 03443000



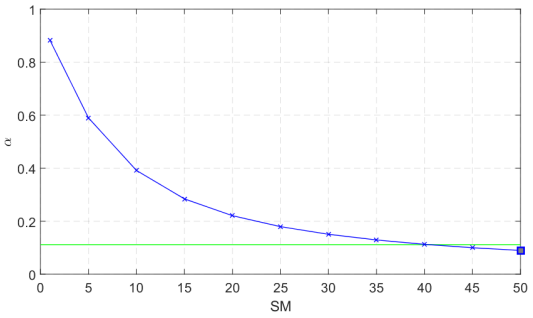
(D) Nash efficiency variation in 03443000



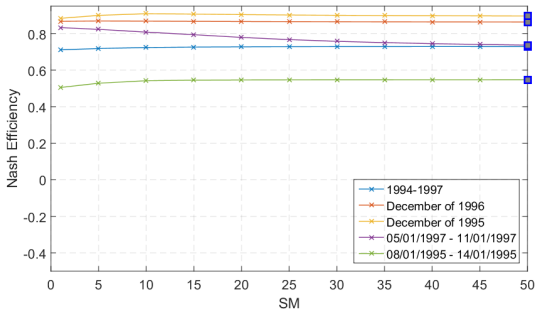
(E) α variation in 11532500



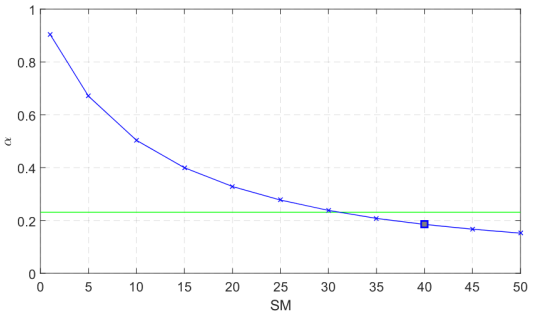
(F) Nash efficiency variation in 11532500



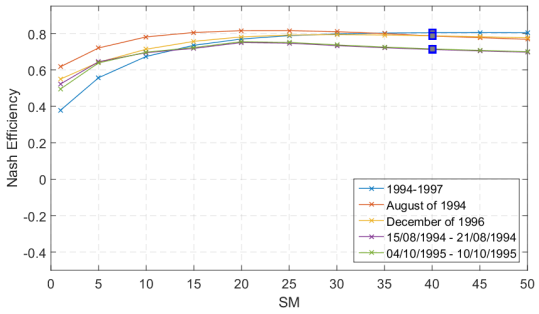
(G) α variation in 02456500



(H) Nash efficiency variation in 02456500



(I) α variation in 02387500



(J) Nash efficiency variation in 02387500

FIGURE 6.6: Recalibration of the model using different values for *SM*

6.5 Summary

In this chapter, the effects of the temporal resolution of the input data on the different outputs of the model are presented. There is also a discussion on how to improve the quality of the simulation by modifying some of the parameters obtained in the calibration.

Chapter 7

Conclusions

The main objective of this research project was to, using the Xinanjiang model, understand better the influence that the temporal resolution of the input data has on the quality of the result when streamflow is being simulated. In order to do so, the XAJ model was run using data with different temporal resolutions ranging from 1 hour to 1 day in 5 different basins. It was possible to see significant differences in the performance indicators of the model when using data with different temporal resolutions.

It was possible to observe that, when long periods of time are simulated, the quality of the result obtained when using data with a temporal resolution of 1 hour was not so different that the quality obtained when data with a daily temporal resolution was used. In most basins, the difference of the Nash efficiency in these 2 cases was less than 5%. Only in one of the study basins, in the smallest one, the decrease in quality was higher than 5% and it usually happened around the 20-hour temporal resolution.

On the other hand, when simulating shorter periods of high intensity rainfall there were bigger differences in the quality of the simulation when using input data with different temporal resolutions, especially in the smallest basins. In the smallest basins, 1000 km^2 or less, were the ones in which the highest decrease in the quality of the simulation was experienced, up to 18% in the worst cases. If a Nash efficiency higher than 95% of the Nash efficiency obtained using hourly input data is wanted, data with a temporal resolution of 10 hours or less should be used in these basins. In the biggest basins, the decrease was lesser but still considerable as it was around 10% and even higher in some cases. In these basins a lost of 5% of the quality of the simulation usually happened between the 15-hour and the 20-hour temporal resolution of the input data. The shorter the time length of the simulation, the sooner it happened.

It was also checked if there was a relationship between the season of the simulation and the decrease in quality of the simulation when using input data with different temporal resolutions. It was thought that because the rainfall and the streamflow had very different behaviors depending on the season maybe that could have and influence on the quality of the simulation. After carrying out the analysis, it was seen that the results obtained were very similar during all the seasons and therefore it was concluded that the seasonality was not an important factor to be considered.

The impact of temporal resolution of input data on model performance during periods of low intensity rainfall was also analyzed. The decrease in quality was more or less the same in all basins and it was between 5 and 10%, but it was not related to the size of the study basin.

Next, it was checked if the time at which the data is measured has an influenced on the quality of the result, and at which time the data should be measured to obtained the best possible results. In order to do so, the XAJ model was run with data

with different starting times. The main conclusion extracted in this part of the study is that the impact of the initial hour of the input data does not have a big influence on the model performance because the quality of the simulation was very similar for all the different data sets used.

Finally, the effect of temporal resolution of the input data on all the outputs of the model was checked. It was seen that the outputs that were affected the most by the temporal resolution of the input data were the ones related to the runoff separation. After this realization, it was tried to improve the quality of the simulation by modifying the parameters of the model to obtain the same runoff separation using data with different temporal resolutions. The parameters that were modified were EX and SM, and it was seen that the value of α was more sensitive to SM than to EX. Nevertheless, improving the quality of the simulation by modifying some of the parameters was proved to be difficult because even if improved values of α were obtained after modifying the parameters, the quality of the simulation did not increase very much, and in some cases it even decreased.

Bibliography

Anand, J. et al. (2018). "Regional scale hydrologic modeling for prediction of water balance, analysis of trends in streamflow and variations in streamflow: The case study of the Ganga River basin". In: *Journal of Hydrology: Regional Studies* 16, pp. 32–53. URL: https://doi.org/10.1016/j.ejrh.2018.02.007.

Aronica, G., G. Freni, and E. Oliveri (2005). "Uncertainty analysis of the influence of rainfall time resolution in the modelling of urban drainage systems". In: *Hydrological Processes* 19.5, 1055– 1071. URL: https://doi.org/10.1002/hyp.5645.

Bauwe, A. et al. (2017). "Does the Temporal Resolution of Precipitation Input Influence the Simulated Hydrological Components Employing the SWAT Model?" In: *JAWRA Journal of the American Water Resources Association* 53.5, pp. 997–1007. URL: https://doi.org/10.1111/1752-1688.12560.

Campolongo, F., S. Tarantola, and A. Saltelli (1999). "Tackling quantitatively large dimensionality problems". In: *Computer Physics Communications* 117.1, pp. 75–85. URL: https://doi.org/10.1016/S0010-4655(98)00165-9.

Christiansen, D.E. (2012). "Simulation of daily streamflows at gaged and ungaged locations within the Cedar River Basin, Iowa, using a Precipitation-Runoff Modeling System model". In: *U.S. Geological Survey Scientific Investigations Report* 5213, pp. 7–26. URL: https://pubs.usgs.gov/sir/2012/5213/.

Farnsworth, R.K., E.S. Thompson, and E.L. Peck (1982). "Evaporation atlas for the contiguous 48 United States". In: *NOAA Technical Report NWS 33*. URL: http://www.nws.noaa.gov/oh/hdsc/Technical_reports/TR33.pdf.

Ficchì, A., C. Perrin, and V. Andréassian (2016). "Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events". In: *Journal of Hydrology* 538, pp. 454–470. URL: https://doi.org/10.1016/j.jhydrol.2016.04.016.

Fredrik, W. et al. (2011). "Effects of temporal resolution of input precipitation on the performance of hydrological forecasting". In: *Advances in Geosciences* 29, pp. 21–25. URL: https://doi.org/10.5194/adgeo-29-21-2011.

Gayathri, K. D., B.P. Ganasri, and G.S. Dwarakish (2015). "A Review on Hydrological Models". In: *Aquatic Procedia* 4, pp. 1001–1007. URL: https://doi.org/10.1016/j.aqpro.2015.02.126.

Guse, B. et al. (2017). "Identifying the connective strength between model parameters and performance criteria". In: *Hydrology and Earth System Sciences* 21.1, pp. 5663–5679. URL: https://doi.org/10.5194/hess-21-5663-2017.

Kandel, D. D., A. W. Western, and R. B. Grayson (2005). "Scaling from process timescales to daily time steps: A distribution function approach". In: *Water Resources Research* 41.2, pp. 1–16. URL: https://doi.org/10.1029/2004WR003380.

Kyi, K.H., M. Lu, and X. Li (2016). "Development of a user-friendly web-based rainfall-runoff model". In: *Hydrological Research Letters* 10.1, pp. 8–14. URL: http://www.nws.noaa.gov/oh/hdsc/Technical_reports/TR33.pdf.

Li, X. and M. Lu (2012). "Multi-step optimization of parameters in the Xinanjiang model taking into account their time scale dependency". In: *Annual Journal of*

Hydraulic Engineering, JSCE 68.4, pp. 145–150. URL: https://doi.org/10.2208/jscejhe.68.I_145.

Li, X. and M. Lu (2014). "Application of aridity index in estimation of data adjustment parameters in the Xinanjiang model". In: *Annual Journal of Hydraulic Engineering, JSCE* 70.4, pp. 163–168. URL: https://doi.org/10.2208/jscejhe.70.I_163.

Loliyana, V.D. and P.L. Patel (2015). "Lumped conceptual hydrological model for Purna riverbasin, India". In: *Sadhana* 40.8, 2411–2428. URL: https://doi.org/10.1007/s12046-015-0407-1.

Lu, M. and X. Li (2014). "Time scale dependent sensitivities of the XinAnJiang model parameters". In: *Hydrological Research Letters* 8.1, pp. 51–56. URL: https://doi.org/10.3178/hrl.8.51.

Moriasi, D. N. et al. (2007). "Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations". In: *American Society of Agricultural and Biological Engineers* 50.3, pp. 885–900. URL: http://dx.doi.org/10.13031/2013.23153.

Morris, M. (1991). "Factorial Sampling Plans for Preliminary Computational Experiments". In: *Technometrics* 33.2, pp. 161–174. URL: https://doi.org/10.1080/00401706.1991.10484804.

Nash, J. and J. V. Sutcliffe (1970). "River flow forecasting through conceptual models part I — A discussion of principles". In: *Journal of hydrology* 10.3, pp. 282–290. URL: https://doi.org/10.1016/0022-1694(70)90255-6.

Nourani, V., A. Roughani, and M. Gebremichael (2011). "TOPMODEL capability for rainfall-runoff modeling of the Ammameh watershed at different time scales using different terrain algorithms". In: *Journal of Urban and Environmental Engineering* 5.1, pp. 1–14. URL: https://doi.org/10.4090/juee.2011.v5n1.001014.

Perrin, C., C. Michel, and V. Andréassian (2003). "Improvement of a parsimonious model for streamflow simulation". In: *Journal of Hydrology* 279.1, pp. 275 –289. URL: https://doi.org/10.1016/S0022-1694(03)00225-7.

Rahman, M. and M. Lu (2015). "Model Spin-Up Behavior for Wet and Dry Basins: A Case Study Using the Xinanjiang Model". In: *Water* 7.8, pp. 4256–4273. URL: https://doi.org/10.3390/w7084256.

Schaake, J., S. Cong, and Q. Duan (2006). "The US MOPEX data set". In: *IAHS publication* 307, pp. 9–28. URL: https://iahs.info/uploads/dms/13600.04-9-28-SCHAAKE.pdf.

Song, X. et al. (2013). "Parameter identification and global sensitivity analysis of Xin'anjiang model using meta-modeling approach". In: *Water Science and Engineering* 6.1, pp. 1–17. URL: https://doi.org/10.3882/j.issn.1674-2370.2013.01.001.

Song, X. et al. (2015). "Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications". In: *Journal of Hydrology* 523, pp. 739–757. URL: https://doi.org/10.1016/j.jhydrol.2015.02.013.

Sun, W. et al. (2015). "Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications". In: *Environmental Research* 139, pp. 36–45. URL: https://doi.org/10.1016/j.envres.2015.01.002.

Wang, Y., B. He, and K. Takase (2009). "Effects of temporal resolution on hydrological model parameters and its impact on prediction of river discharge". In: *Hydrological Sciences Journal* 54.5, pp. 886–898. URL: https://doi.org/10.1623/hysj.54.5.886.

Wheater, H., S. Sorooshian, and K. Sharma (2007). "Hydrological Modelling in Arid and Semi-Arid Areas". In: *Cambridge University Press*, pp. 1–223. URL: https://doi.org/10.1017/CBO9780511535734.

Wilk, J. and D.A. Hughes (2002). "Calibrating a rainfall-runoff model for a catchment with limited data". In: *Hydrological Sciences Journal* 47.1, pp. 3–17. URL: https://doi.org/10.1080/02626660209492903.

Yao, C. et al. (2009). "Application of a Developed Grid-Xinanjiang Model to Chinese Watersheds for Flood Forecasting Purpose". In: *Journal of Hydrologic Engineering* 14.9, pp. 923–934. URL: https://doi.org/10.1061/(ASCE)HE.1943-5584.0000067.

Zhao, R.J. (1992). "The Xinanjiang model applied in China". In: *Journal of Hydrology* 135.1, pp. 371–381. URL: https://doi.org/10.1016/0022-1694(92)90096-E.

Zheng, F. et al. (2018). "On lack of robustness in hydrological model development due to absence of guidelines for selecting calibration and evaluation data: Demonstration for data-driven models". In: *Water Resources Research* 54, 1013–1030. URL: https://doi.org/10.1002/2017WR021470.