

**Université Pierre et Marie Curie, École des Mines de Paris
& École Nationale du Génie Rural des Eaux et des Forêts**

**Master 2 Sciences de l'Univers, Environnement, Ecologie
Parcours Hydrologie-Hydrogéologie**

**Flood forecasting at short and large lead times: How to choose the
best adapted model?**

**(Prévision des crues à brève et longue échéances: comment choisir
le modèle le mieux adapté ?)**

Adnan Ahmad TAHIR

Directeur(s) de recherche : Lionel BERTHET et Vazken ANDRÉASSIAN



**Groupement d'Antony
Parc de Tourvoie – BP 44,
92163 Antony Cedex**

Date : 28-01-2008 à 25-07-2008



Abstract:

This study analysis the sensitivity of the performance of the lumped flood forecasting GR3P model to the time resolution (i.e. the use of different short and large model's time steps) to reach different lead times. In this study, we test our model against a large set of watersheds to warrant the development and the evaluation of a reliable, robust and general procedure. Depending on the reaction times of the catchments, the selected lead times range from 2 to 72 h. The time steps go from 1 to 24 h. The model is optimized for each time step (but it keeps the very same structure). We estimate the efficiency of the GR3P model working on the different time steps to reach the given lead times with the persistence criterion and the persistence criterion on logarithms of discharge. When assessing the performances on high flows, the best performance is obtained when the model's time step equal to the assessment time step. Even using a null future precipitation scenario with a focus on high flows, it is still useful to choose a high time resolution model (i.e. model working on shortest time step) in order to get the most accurate forecast. Thus for operational forecasters, it is valuable to run the model GR3P at the hourly time step.

Résumé:

Ce stage porte sur une étude de sensibilité du modèle global de prévision des crues GR3P à la résolution temporelle (i.e. l'emploi d'un petit pas de temps ou d'un pas de temps plus grand) pour atteindre différents horizons de prévisions. Nous testons dans cette étude notre modèle sur un large échantillon de bassins versants pour assurer le développement et l'évaluation d'une procédure robuste, fiable et générale. Les horizons de prévisions sont choisis en fonction de notre échantillon de bassins versants et vont de 2 à 72 heures. Nous choisissons donc des pas de temps entre 1 et 24 heures. Le modèle est optimisé pour chaque pas de temps (tout en conservant la même structure). Les performances du modèle GR3P fonctionnant aux divers pas de temps pour atteindre les différents horizons sont évaluées à l'aide des critères de persistance sur les débits et sur les logarithmes des débits. Quand on évalue les modèles sur les hautes eaux, le modèle fonctionnant au pas de temps égal au pas de temps utilisé pour le calcul du critère de performance présente des performances significativement meilleures que les autres modèles. Ceci est également vrai lorsqu'on emploie un scénario de pluies futures parfaitement connues ou un scénario de pluies futures nulles. Aussi, en conditions opérationnelles, il est préférable de choisir le modèle fonctionnant à une haute résolution temporelle (le modèle avec le pas de temps le plus court) pour fournir les prévisions les plus précises.

ACKNOWLEDGEMENTS

God never spoils any effort. Every piece of work is rewarded according to the nature of devotion in it. I am humbly thankful to most gracious, merciful and almighty **Allah** Who gave me health, affectionate and talented teachers, helping friends and opportunities to complete my study.

I feel much honor to express my deep sense of gratitude and indebtedness to my honorable supervisors **Lionel BERTHET and Vazken ANDRÉASSIAN** for their mastery advice, keen interest, constructive criticisms and encouragement throughout the course of my studies.

I have no appropriate words that fully convey the wave of immense indebtedness and deep gratitude that I owe to **Charles PERRIN and Pierre JAVELLE** for their most co-operative attitude, inspiring guidance, critical review and valuable comments throughout the study period and preparation of this manuscript.

I feel my proud privilege to mention the feelings of obligations towards my colleagues **Florent LOBLIGEOIS, Baudouin SAINTYVES and Meggy HAU.**

Last but not least, I am most earnestly obliged to my dear Parents for their strenuous effects in enabling me to join the higher ideas of life and has always wished to see me glittering high on the skies of success like galaxies.

(Adnan Ahmad TAHIR)

TABLE OF CONTENTS

1.	ABSTRACT	3
2.	ACKNOWLEDGEMENTS	4
2.	LIST OF FIGURES	7
3.	LIST OF TABLES	8
CHAPTER 1	INTRODUCTION	9
1.1	Objective of study	9
1.2	Rainfall-runoff modelling and its importance	9
1.2.1	Lumped vs. Distributed Models	9
1.2.2	Classification of model structures	10
a)	Metric Model Structures (empirical)	10
b)	Parametric Model Structures (conceptual)	10
c)	Mechanistic Model Structures (physically based)	11
1.3	Operational application of rainfall-runoff modelling	11
1.4	Methods being used for flow forecasting	12
CHAPTER 2	REVIEW OF LITERATURE	13
2.1	Forecasting models with different updating techniques	13
2.2	Use of forecasting models on time resolution basis	14
CHAPTER 3	MATERIALS & METHODS	15
3.1	Description of the Model GR3P	15
3.2	Criteria	18
3.2.1	What criterion to use?	18
a)	Nash-Sutcliffe criterion	18
b)	Persistence criterion	19
3.2.2	How to use this criterion over a large sample of watersheds?	20
3.3	Calibration and validation	20
3.3.1	Calibration	20

3.3.2	Procedure of calibration-validation (split-sample test)	20
3.4	Catchments data base	21
3.4.1	178 catchments	21
3.4.2	Data available for each catchment	22
3.5	Method of our study	23
CHAPTER 4	RESULTS & DISCUSSION	28
4.1	First results	28
4.2	Can we see some differences for slow or fast catchments?	29
4.2.1	Analysis of the performances of different model's time steps on different watershed samples	29
4.3	How can we explain those first results?	30
4.4	Experiments to test our hypotheses	31
4.4.1	Analysis of the performance of model using different time steps when taking into account only the high flows	31
4.4.1.1	Does it depend on reaction time of our catchments?	32
4.4.1.2	Analysis of the performance of model using different time steps when taking into account the high flows on fast reacting catchments	33
4.4.2	Analysis of the performance of model using different time steps when taking into account only the large flow variation	34
4.4.3	Use of different performance criteria	35
4.5	Analysis of the performance of model using different time steps when precipitation scenario is taken as zero	36
4.5.1	When taking into account the whole periods of data	36
4.5.2	Analysis of the performance of model using different time steps when taking into account the high flows with a null future precipitation scenario	37
4.6	Discussion	38

4.6.1	Focusing on high flows or on large flow variations?	39
4.6.2	The best model's time step is the one used for performance assessment	40
4.6.3	Criteria of performance	40
4.6.4	Operational learning from our study	41
CHAPTER 5	CONCLUSIONS	42
REFERENCES		43
APPENDICES		
APPENDIX A	Some important notations used in this study	45
APPENDIX B	Details of the routing reservoir updating phase	46
APPENDIX C	Different values of fixed parameters tested on a set of time steps to choose best values of fixed parameters	47
LIST OF FIGURES		
3.1	General scheme/structure of the flow forecasting model GR3P and order in which operations are made in the model (meanings of main notations are given in the appendix A)	16
3.2	Cumulative unit hydrograph (CUH) and discrete unit hydrograph (UH) for $X_3 = 30$ hrs	17
3.3	Limit values for persistence criterion and C_{2MP}	19
3.4	Map showing the sample catchments in France	22
3.5	Aggregation of the available hourly discharge data over time steps of 12 h	23
3.6	Example of graphs plotted between different values of fixed parameters and performance criterions	24
3.7	Example of graphs plotted between the performance criterion and model working at different time steps to reach a given lead time	25
3.8	Example of comparison of forecasts using different time resolutions to	27

	reach a lead time of 24 h. Performances are assessed at time step of 6 h	
4.1	Examples of graphs showing the performance of different model's time steps to reach different lead times	28
4.2	Examples of graphs showing the performance of different model's time step on different samples of watersheds	30
4.3	Examples of graphs showing some of the better performing model's time steps to reach different lead times when taking into account the high flows	32
4.4	Example of graph showing the procedure to choose the fast reacting sub sample of a catchment	33
4.5	Examples of graphs showing some of the better performing model's time steps when taking into account the high flows on fast reacting catchments	34
4.6	Examples of graphs showing the better performing model's time steps to reach lead times of 2 h and 3 h on very slow reacting watershed	35
4.7	Examples of graphs showing the better performing model's time steps to reach lead times of 3 h, 12 h, 24 h and 48 h	36
4.8	Examples of graphs showing some of the better performing model's time steps when taking into account the high flows with null future precipitation scenario	38
4.9	Examples of graphs showing the samples and a zoom of discharge distribution and the gradient of discharge distribution	39

LIST OF TABLES

3.1	Statistical indexes of precipitation and streamflow	21
3.2	Model's time steps values on different assessment time steps to reach a given lead time	27

Chapter 1

INTRODUCTION

Due to the increasing vulnerability of our societies to floods and the need for better management of water resources, the demand for efficient streamflow forecasting methods is constantly increasing. Rainfall-Runoff modelling is a tool commonly used for flood forecasting. Much research is still being driven to improve the forecast. In fact, different types of floods (flash floods, slow and rapid-onset floods etc.) can occur. The lead time for which we can issue an acceptable forecast will depend on the type of flood. Therefore the modellers should consider very different lead times from a few hours to several days. This is why we are designing and assessing forecasting model on time resolution basis (i.e. different time steps) to reach different lead times.

1.1 Objective of study

Our main objective is to estimate the efficiency of the GR3P model (flood forecasting model) working at different time steps to reach different given lead times. In the operational application, this study will be helpful for the forecasters to choose the appropriate model's time step to reach the desired lead time.

For better understanding of our study, a short description of rainfall-runoff modelling, its importance, classification of rainfall-runoff models, their operational applications and different methods of flow forecasting, are presented later in this chapter.

1.2 Rainfall-runoff modelling and its importance

Rainfall-Runoff models are tools used for hydrological investigations in engineering and environmental science (Wagener, T., *et al.* 2004). The main objective of this modelling domain is to simulate the catchment response to rainfall in terms of streamflow. Various types of hydrological models are now developed by researchers.

1.2.1 Lumped vs. Distributed Models

Models can be classified depending on their spatial description of the catchment: they can be specified as lumped and distributed models (or semi-distributed models).

Lumped models do not explicitly take into account the spatial variability of inputs, outputs, or parameters but they consider the whole catchment as a single unit. They are usually structured to utilize average values of the watershed characteristics affecting runoff volume (e.g. the GR3P model used in this study).

Distributed models include spatial variation in inputs, internal variables, outputs, and parameters. In general, the watershed area is divided into a number of elements and runoff volumes are first calculated separately for each element (e.g. some implementations of TOPMODEL are distributed).

1.2.2 Classification of model structures

There are a large number of various model structures developed so far. Therefore, it is necessary to classify these structures, and one commonly applied classification uses three distinct classes (Wagener, T., Wheeler, H.S. and Gupta, H.V. 2004).

- a) Metric (data-based, **empirical** or black-box) model structures,
- b) Parametric (**conceptual** or grey box) model structures,
- c) Mechanistic (**physically based** or white box) model structures.

a) Metric Model Structures (**empirical**)

These model structures commonly use the available time-series to derive both the model structure and the corresponding parameter values. They are purely based on the information obtained from the data (hence also called data-driven models) and any prior knowledge about catchment behaviour and flow processes. e.g. Artificial Neural Networks (ANN), Nearest Neighbours Method (NNM) and Transfer Functions (TF) are empirical. Metric models are usually spatially lumped, i.e. they treat the catchment as a single unit.

b) Parametric Model Structures (**conceptual**)

In contrast to metric models, the structure is defined according to the concepts of the modeller about hydrological system (e.g. water balance, conservation of mass and the available data and processes that the modeller considers as dominant in the catchment), and hence such models are also commonly termed conceptual. However these models still depend on time-series of system output, mainly streamflow, to get the values of their parameters in a calibration procedure. The main part of these models is a storage element. These storages are

filled by rainfall, infiltration or percolation, and emptied by evapotranspiration, runoff, drainage etc. The parameters define the size of the storage elements or the distribution of flow between them. Most of these models are lumped. However, there is one common approach to divide the catchment into smaller sub-catchments, the so called *semi-distributed* approach.

c) Mechanistic Model Structures (physically based)

The basics of these models are the principles of physics (conservation of mass, momentum and energy). These became practically applicable in the 1980s, as a result of improvements in computer power. The expectation was that the extent of physical realism to which these models are based would be sufficient to relate their parameters, such as soil moisture characteristics and unsaturated zone hydraulic conductivity functions for subsurface flow or friction coefficients for surface flow, to the need for model calibration. However, these models faced the problems of extreme data demand, scale related problems and over-parameterization. One consequence is that the model parameters cannot be derived through measurements; mechanistic model structures therefore still require calibration, usually of a few key parameters, though applied to a large number of elements. The expectation that these models could be applied to ungauged catchments has therefore not been fulfilled. They are typically rather applied in a way that is similar to lumped conceptual models. These models use normally the smaller distributed units based on grids, hill slopes or some type of hydrologic response unit.

1.3 Operational application of rainfall-runoff modelling

Flood forecasting is a very important part of water resources management activities which relate to flood warning, flood control or reservoir operation. To the researchers working on hydro systems and water resources managers, operational applications such as flood forecasting is still an important existing demand because: i) this is a real time operational application, ii) this is a stressful situation and iii) hydrologists have no time for data quality control. This is why operational users ask for robust and easy to use tools. Garrote & Bras (1995) stated that flood forecasting is to be considered as one of the unsolved problems of operational hydrology.

It is also known that flow forecasting is different from flow simulation in a hydrological context.

- **Flow simulation** consists in running a hydrological model using as input variables only the past sequence of rainfall (and other inputs such as potential evapotranspiration) until the current time step to estimate the flow value at the current time step. Observed flows are not used in the simulation process (they are only used in the prior phase of model calibration and evaluation).

Flow simulation is used:

- To assess our understanding of Rainfall-Runoff process.
 - To estimate the missing information of streamflow in a long time series data.
 - To estimate the heavy floods occurred, if we only have a long data time series of rainfall.
- **Flow forecasting** consists in running a hydrological model to calculate future flow values over the forecast period using the same inputs as previously and a scenario of future rainfall up to the forecast lead time. In addition to these inputs, the past sequence of measured flows is also used up to the time when the forecast is issued.

1.4 Methods being used for flow forecasting

A lot of methods have been proposed to overcome this operational hydrological problem. Up to now the simulation models were being used with an updating procedure to forecast the flow and it is a common thinking that any type of simulation model with an independently chosen updating procedure can be used for forecasting (Tangara M., 2008). A research held at CEMAGREF has described that by choosing a simulation model with an updating scheme independently cannot ensure the efficiency of forecasting so there should be some specific characteristics to combine the simulation models and updating procedures. This issue led the researchers of CEMAGREF to develop a model, GR3P, which is specifically a forecasting model with a built-in updating technique. This study is to be achieved by using GR3P.

This report is organized as follows: chapter 2 provides a short background of the study. Chapter 3 describes the development of model used in this study, the methodology adopted for forecasting and for the estimation of performance. Chapter 4 presents the results and provides a discussion on the relative performances of different model's time steps used to reach a given lead time.

Chapter 2

REVIEW OF LITERATURE

2.1 Forecasting models with different updating techniques

Many approaches have been proposed in order to issue forecasts with different hydrological models. To improve the performance of these models, different updating techniques are used in combination with different lumped or distributed rainfall-runoff models. Updating is a procedure to improve the efficiency of forecasting model by continuously comparing simulations to observations at the time of forecast and by changing values of some variables or parameters to reduce the differences e.g. *Yang, X. and Michel, C. (2000)* introduced a parameter updating procedure that can be combined with conceptual rainfall-runoff models for flood forecasting purposes. *Refsgaard, J. C. (1997)* proposed a classification of updating procedures used for forecasting depending on the modification of variables during the feedback process:

- 1) Updating of input variables (precipitation, air temperature).
- 2) Updating of state variables (snowpack's water equivalent and water contents of reservoir).
- 3) Updating of model parameters (runoff coefficient and hydrograph).
- 4) Updating of output variables (observed stream flow).

Updating techniques can be used in two different ways. First, they can be used in combination with a simulation model which is calibrated separately and secondly, it can be used as a built-in technique of the model to use the last observed flows as an actual model input. *Tangara, M. et al. (2008)* have compared the performance of a model, GR3P, which has a built-in updating technique with a combination of a simulation model and an updating technique. Updating technique in second case was the same as used by the model GR3P. The results of this study have shown that the GR3P model with a built-in updating technique is a more simple and efficient model. *Toth, E. and Brath, A. (2007)* have shown that the forecasting ability of a conceptual rainfall-runoff model in combination with an output updating technique is better than a neural network model when focusing on the prediction of flood events and especially in case of a limited availability of calibration data.

2.2 Use of forecasting models on time resolution basis

Time resolution is also an important factor in hydrological modelling. The idea of studying a model on the basis of time resolution i.e. different short and large time steps and lead times is not widely studied. We found few articles which have described the importance of time resolution in hydrological modelling in comparison with the number of articles dealing with the question of spatial resolution. *Hughes, D. A. (1993)* discussed the advantages of incorporating variable length time steps into deterministic hydrological models and has presented a method based on the use of rainfall intensities to determine appropriate length time steps automatically. It differs from other approaches that have been reported (e.g. by *Dunsmore et al., 1986*), in that the same model functions are used regardless of the time interval. The performance of a model can be different on different lead times. *Tangara, M. (2005)* have reported that GR3P model proves to be an efficient forecasting tool for short lead times. *Toth, E. and Brath, A. (2007)* showed that the forecasting ability of neural network models is proved to be excellent over all lead times when simulating over continuous periods, provided that an extensive set of both stream flow and precipitation data is available for calibration purposes. The above references show that the time resolution has a significant effect on the hydrological model performance.

Chapter 3

METHODOLOGY

Here we introduce the main elements of our approach: the used model first, then its calibration and validation, the criteria of performance and the data we use. We finish with the proposed method to estimate the efficiency of the GR3P model working at different time steps to reach different given lead times.

3.1 Description of the Model GR3P

The GR3P Flood forecasting model used in this study belongs to a combination of metric (empirical) and parametric (conceptual) approaches i.e. *hybrid metric-conceptual models*. GR3P model is lumped and has only three free parameters (commonly called parameters) to calibrate:

- 1) X_1 : Maximum capacity of the quadratic routing store (level noted R, in mm)
- 2) X_2 : Adjustment coefficient of effective rainfall
- 3) X_3 : Base width of the unit hydrograph (UH)

There are also four fixed parameters in the GR3P model: (1) the maximum capacity of soil moisture accounting store, (2) the exponent of unit hydrograph, (3) the coefficient of percolation function and (4) the exponent of final correction. Fixed parameters differ from free parameters, in that their values do not change from one catchment to another whereas the values of free parameters of a model change from one catchment to another.

The model can be understood as the union of two parts: a production function and a routing function. Production function defines the amount of water to be delivered as discharge at the outlet in the next time steps while routing function distributes/allocates this amount of water to each of the next time steps.

A complete structure of the model GR3P and the way in which it operates is shown in figure 3.1.

Production function

In the functioning of the GR3P model, first comes the production function. Net precipitation and net potential evapotranspiration are calculated using the precipitation and potential

evapotranspiration at the current time step t , and the water content of the production store is updated.

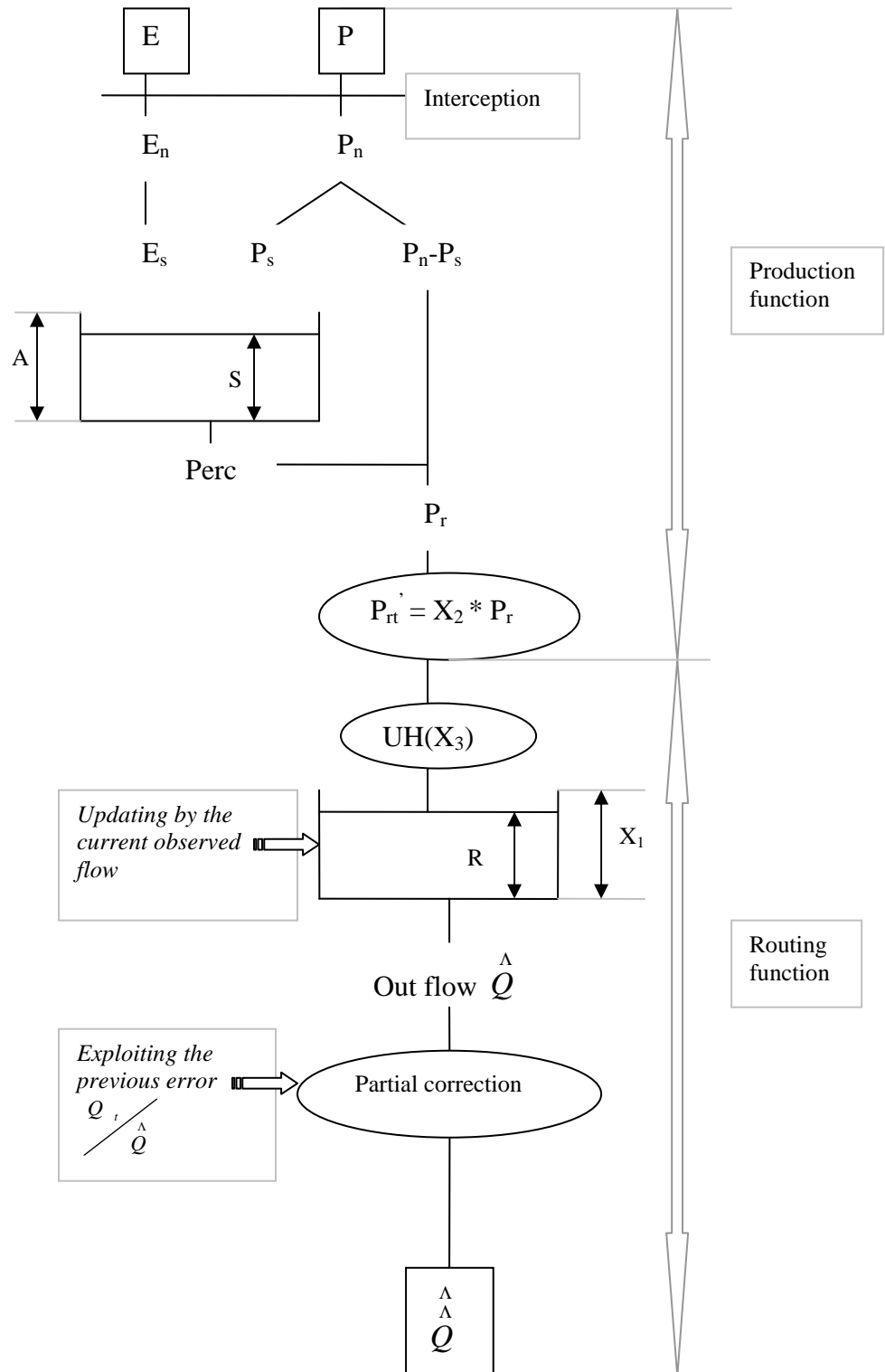


Figure 3.1: General scheme/structure of the flow forecasting model GR3P and order in which operations are made in the model (meanings of main notations are given in the appendix A).

A part of net precipitation directly goes to the routing function as direct flow while the remaining fills the production store. The level of this store defines how much water goes to the routing function.

A part of the stored water percolates and joins the routing system/function. The amount of water contents percolated from the store is most often much smaller than the direct flow when there is some precipitation. Water entering the transfer function is then multiplied by X_1 , the adjustment coefficient that plays the role of water balance adjustment for short time scale in the production module.

Routing function

The flux which leaves the production function is then transformed by the routing procedure, which acts in two steps.

a) First, there comes a unit hydrograph. The discrete ordinates $UH(k)$ of the unit hydrograph are calculated by the difference of the successive values of the cumulative hydrograph (S-curve) as follows: $UH(k) = CUH(k) - CUH(k-1)$

$$\text{If } k \leq 0 \text{ then } CUH(k) = 0$$

$$\text{If } 0 < k < X_3 \text{ then } CUH(k) = k^\alpha / [k^\alpha + (X_3 - k)^\alpha]$$

$$\text{If } k \geq X_3 \text{ then } CUH(k) = 1$$

where X_3 is a parameter (in hours) to be calibrated and α is the exponent of the unit hydrograph. Figure 3.2 shows cumulative unit hydrograph and discrete unit hydrograph.

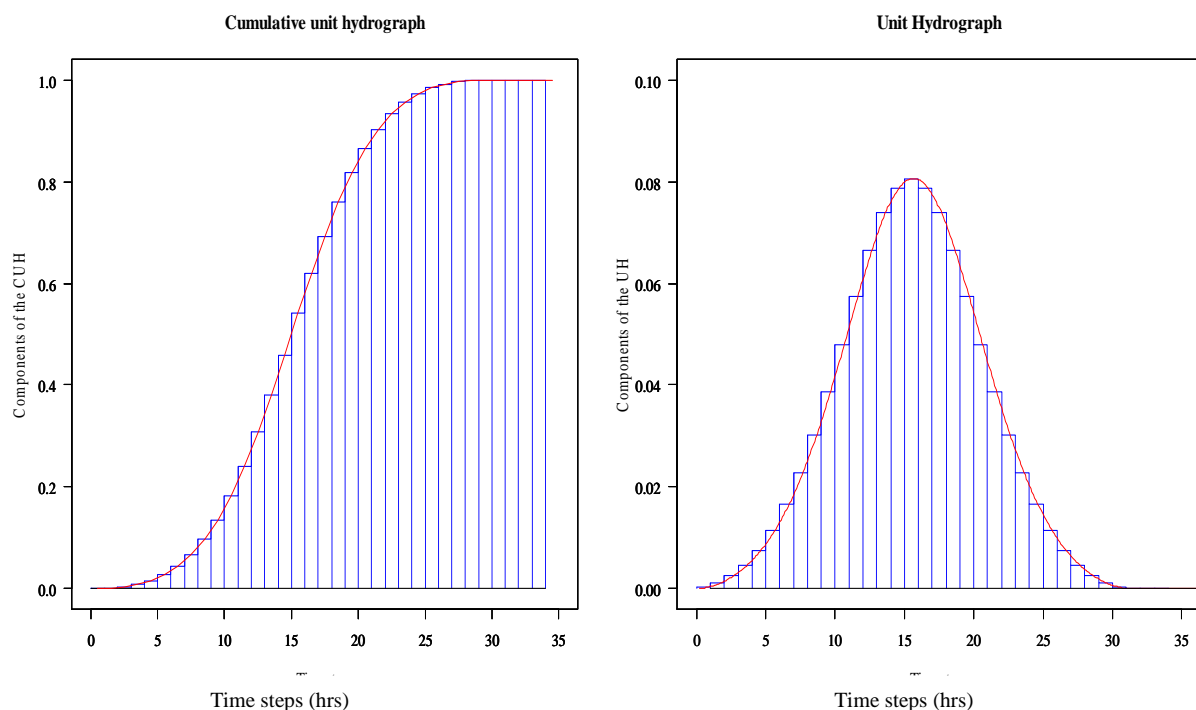


Figure 3.2: Cumulative unit hydrograph (CUH) and discrete unit hydrograph (UH) for $X_3 = 30$ hrs.

b) Then comes a quadratic routing store (capacity of X_1 in mm). The water content of the routing store at the end of the current time step t is directly calculated using the observed flow.

How does the model use the last observed flow for update?

The model has two updating techniques.

a) The water level in the store at the end of current time step 't' is given by

$$\hat{R}_{t|t} = \frac{\sqrt{Q_t^2 + 4X_2Q_t} - Q_t}{2} \quad \text{Eq. A}$$

Where the subscript ' \wedge ' indicates the update (i.e. use of the current observed flow), see appendix A for notations. The observed discharge Q_t is therefore an actual input of the updating procedure. The explanation of equation A is given in Appendix B.

b) Eventually forecast flows $\hat{Q}_{t+1|t}$ to $\hat{Q}_{t+i|t}$ are obtained by applying the forecast error made at the previous time step $\hat{Q}_{t|t-1}$ (i.e. forecast made at time step $t-1$ for time step t). Forecast flow at time step $t+i$ (i between 1 and L) is given by:

$$\hat{Q}_{t+i|t} = \hat{Q}_{t+i|t} \cdot \left[\frac{Q_t}{\hat{Q}_{t|t-1}} \right]^\beta$$

3.2 Criteria

Many different criteria are used to calibrate and validate the models. Some criteria for validation of the model are given here.

3.2.1 Which criterion to use?

a) Nash-Sutcliffe criterion

This model efficiency coefficient is often used to assess the prediction efficiency of the hydrological models (Nash and Sutcliffe, 1970). It is given by the formula:

$$E = 100 \cdot \left[1 - \frac{\sum_{t=1}^N (Q_{t+i} - \hat{Q}_{t+i|t})^2}{\sum_{t=1}^N (Q_{t+i} - \bar{Q})^2} \right]$$

This criterion is a comparison of the mean square error and of the variance of streamflows. It can also be understood as a comparison of our model to a naïve model which, in this case, is the mean flow over the test period. It is a classical as it is widely used for the simulation of stream flow.

These efficiencies have a value ranging from $-\infty$ to 1. The closer the model efficiency to 1, the more accurate the model is.

b) Persistence criterion

It is computed by the following formula:

$$Pers = 100 \cdot \left[1 - \frac{\sum_{t=1}^N (Q_{t+l} - \hat{Q}_{t+l|t})^2}{\sum_{t=1}^N (Q_{t+l} - Q_t)^2} \right]$$

The persistence index is a criterion similar to the Nash and Sutcliffe criterion but its naïve model is better adapted to forecasting. This criterion compares the error of the tested model to the error of a naïve model that assumes the future forecast streamflow values same as the last observed streamflow values.

Nash and Sutcliffe criterion is very coarse in a forecasting context because it uses the mean observed flow which may be too big or too small than on the forecast period and which therefore yields high (and often misleading) efficiency values.

C_{2MP} is a transformation of persistence criterion, given by the following formula:

$$C_{2MP} = \frac{100 \times Pers}{200 - Pers}, \text{ where } Pers \text{ is the persistence criterion. } C_{2MP} \text{ has been proposed by Mathevet (2005).}$$

If C_{2MP} has a positive value, i.e. between 0 and 100 then the tested model is better than the naïve model. Conversely if the value of C_{2MP} is negative i.e. between 0 and -100, then the performance of the naïve model is better than the tested model. Limits of values for the persistence as well as for C_{2MP} are shown in the figure 3.3.

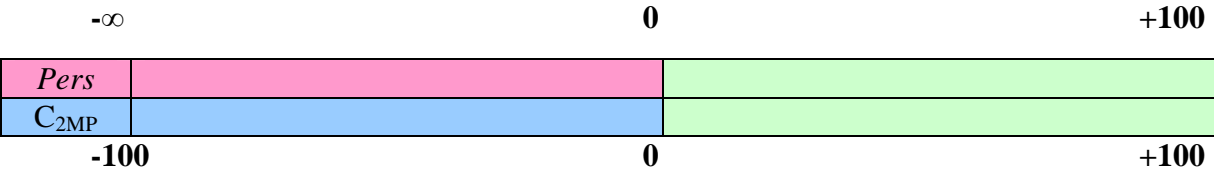


Figure 3.3: Limit values for persistence criterion and C_{2MP} .

We prefer using C_{2MP} because it has a limited lower value also in the case when naïve model is better than the tested model i.e. -1, which in the case of persistence criterion is unlimited i.e. $-\infty$. We can say that C_{2MP} is a more significant way to present the criterion values because it has limited values for both the best and worst performance of the model when we are interested in the average performance of large number of watersheds.

3.2.2 How to use this criterion over a large sample of watersheds?

To compare different versions of the model, we analyse the distribution of criterion values for the different versions. The best version is chosen by examining, first the median values (or averages) and then the whole distribution (especially the tails) of the parameter values.

3.3 Calibration and validation

3.3.1 Calibration

The quality of any calibration process is very much dependent on the quantity and quality of the time-series data (data of precipitation, evapotranspiration etc.) used. The necessary quantity of data depends on the amount of information in it (with respect to identifying the model parameters) i.e. the number of events (storms) rather than length of data series. The quality of the data relates to the errors which are present in the information.

The quantity of data required for calibration depends on the number of parameters of model structures to be estimated and on the quality and characteristics of the data. Franchini and Pacciani (1991) found that the required length of the calibration data was directly related to the number of parameters to be optimized. (Wagener, T., Wheeler, H.S. and Gupta, H.V., 2004).

3.3.2 Procedure of calibration-validation (split-sample test)

The available stream flow record should be split into two segments one of which should be used for calibration and the other for validation. If available stream flow record is long enough that one half of it may suffice for calibration, it should be split into two equal parts, each of them should be used in turn for calibration and validation, and the results from both arrangements should be compared. The model should be judged acceptable only if the two results are similar and the error in both validations runs acceptable (*V. Klemeš, 1986*).

A single calibration is not sufficient to test the model because it is only the estimation of the best combination of parametric values: validation over another period is also needed. A large

number of parameters increase the model performance over the calibration period because the number of degrees of freedom also increases and model yields a better fit of observed data. But in validation process, this trend may disappear and a model with a limited number of parameters may achieve the results as well as the models with large number of parameters (Perrin C., 2001). If we increase the number of parameters beyond a certain limit then performance of model starts decreasing in validation: this is due to over calibration.

3.4 Catchments data base

3.4.1 178 catchments

A data base of more than one thousand catchments of different regions (area ranging between 10 to 8900 km²) under different climatical conditions is available and 178 catchments are selected on the basis of similarity in the climatic conditions and behaviour of the catchments by using their discharge auto correlation. Area for these 178 catchments varies between 10 and 5940 km². Principal components analysis (PCA) was used with ten statistical indices of precipitation and streamflow (Table 3.1) describing the diversity of catchments.

Indices	Unit
Mean Annual Stream flow	mm
Seasonal streamflow variation	%
Low flows	dL.s ⁻¹ .km ⁻²
Modular low flow	%
2-year flood	L.s ⁻¹ .km ⁻²
Base flow index	%
Mean annual precipitation	mm
Seasonal precipitation	%
Rainfall of return-period 2 years	mm
Annual fraction of no-rainfall days	%

Table 3.1: Statistical indexes of precipitation and streamflow.

Why do we use so many catchments (178)?

In this study, we chose to test our model against a large set of watersheds encompassing widely different regions. Indeed we believe that only a large data base can warrant the development and the evaluation of a reliable, robust and general procedure (Tangara, M., 2005).

Why do not we use all of the available catchments?

We are going to estimate the efficiency of model on short time steps and if we will take into account all of the available catchments then the process for estimation of model performance will take a long time.

Sample watersheds (with their outlets) in France are shown in the figure 3.4. The small grey points show all catchments (more than 1000) whereas red squared blocks show the 178 catchments chosen for our study.

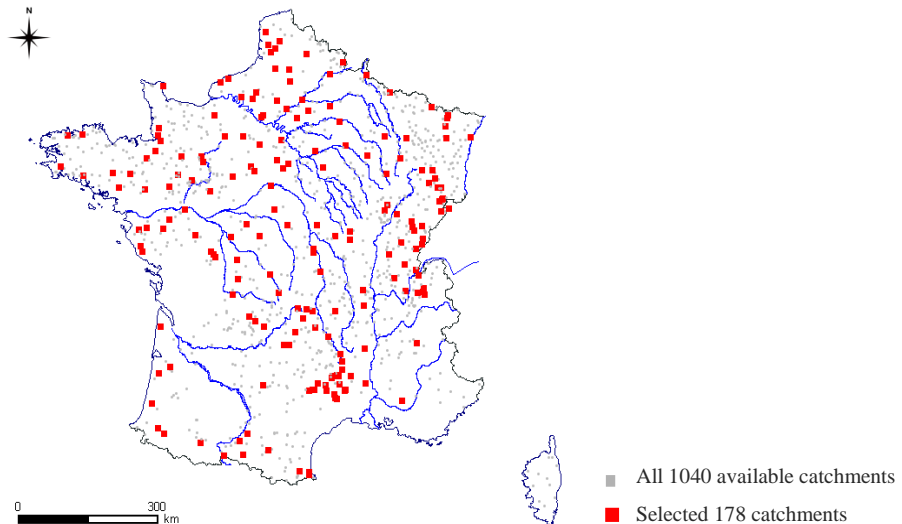


Figure 3.4: Map showing the catchments in France.

3.4.2 Data available for each catchment

In the data base we have data of precipitation, potential evapotranspiration and discharge at the hourly time step for the years 1995-2005. To measure the mean areal precipitation, the rainfall data was obtained from a network over 600 rain gauges in France. The areal precipitation is then assessed by using the Thiessen polygon method. The potential evapotranspiration was estimated by the method proposed by L. Oudin (2005). Discharge is evaluated through a rating curve.

In the case of model GR3P we used observed rainfall as scenario of future rainfall to test the flow forecasting approaches. This choice constitutes an optimistic scenario but it benefits equally to the compared procedures and has thus no impact on their ranking, and thus on our conclusions (Tangara, M., 2008).

3.5 Method of our study

Step 1

In the first step we chose different lead times and time steps (to reach these lead times). The response time of the catchment from our sample varies between a few hours to few days so consequently the lead times chosen go from 2 h, 3 h, 6 h, 12 h, 18 h, 24 h, 36 h, 48 h and 72 h. Then the time steps chosen to run the model are 1 h, 2 h, 3 h, 6 h, 12 h, and 24 h. A complete procedure to understand the functioning of our model is shown in the figure 3.8.

Step 2

In the next step we aggregate the available hourly data in the form of time steps 2 h, 3 h, 6 h, 12 h and 24 h for each catchment. An example of aggregation of discharge data is shown in the figure 3.5.

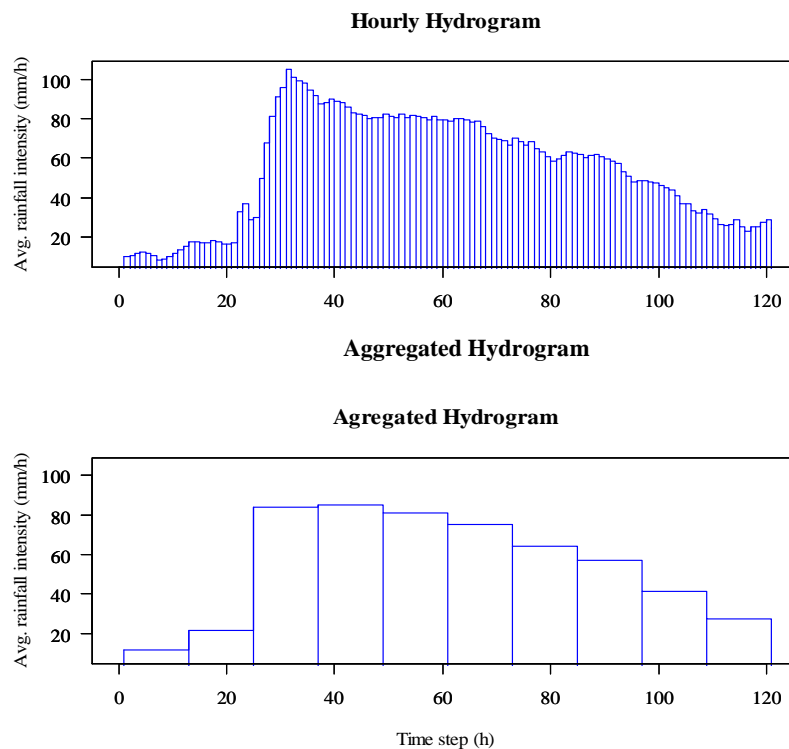


Figure 3.5: Aggregation of the available hourly discharge data over time steps of 12 h.

Step 3

Then we optimize the model for each time step (with respect to the model structure - identical for all time steps): we estimate the best values of each of the fixed parameters of the model running at each of the mentioned time steps (Appendix C).

To estimate their best values we try different values of parameters and estimate the performance of the model at these values of parameters. The values with best performance at each time step to reach given lead times are chosen. An example of the graphs plotted for different values of each of the parameter at time step 6 h are given in the figure 3.6.

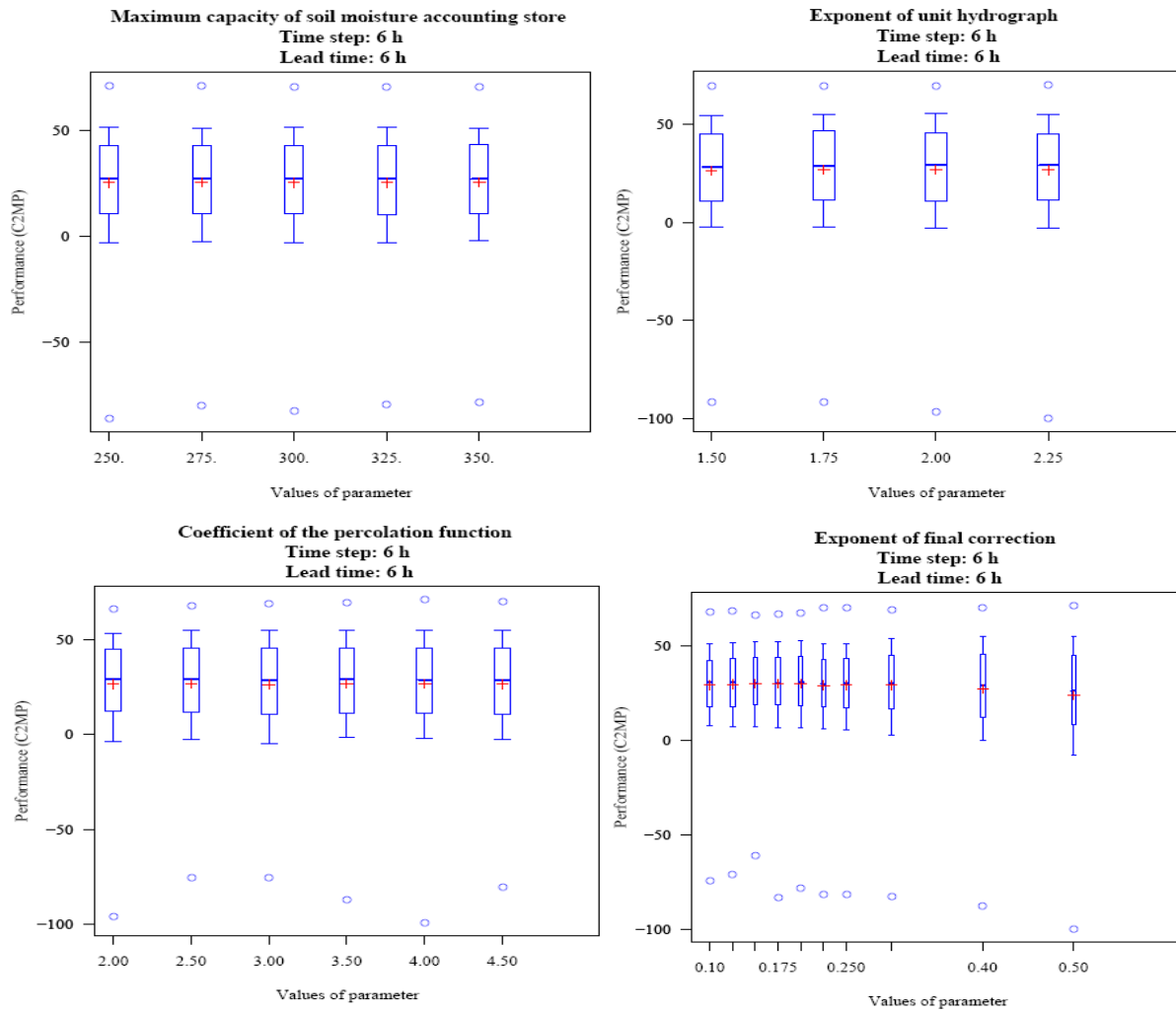


Figure 3.6: Example of graphs plotted between different values of fixed parameters and performance criterion.

Step 4

Finally we estimate the efficiency of the GR3P model working on different time steps to reach given lead times. The criterion used to assess the performance of our method is C_{2MP} (i.e. the transformation of the persistence criterion).

We want to assess first, punctual forecast (e.g. what would the discharge be at 6 h as shown in the figure 3.8) and then also we want to make further forecasts so we use different assessment time steps. If we want to assess the performance of a model working with a small time step on a larger assessment time step, we aggregate its results. Conversely if we want to assess the performance of a model working with a large time step on a smaller assessment time step then we disaggregate its results according to the required assessment time step. The different assessment time steps used in this study are given in table 3.2.

Model working at a time step uses the best values of fixed parameters for this time step. The analysis is carried out for each lead time. We compare for a lead time, the performance of models working at different time steps, assessed at the time step. An example of the graphs showing the performance of each of model's time steps is given in the figure 3.7.

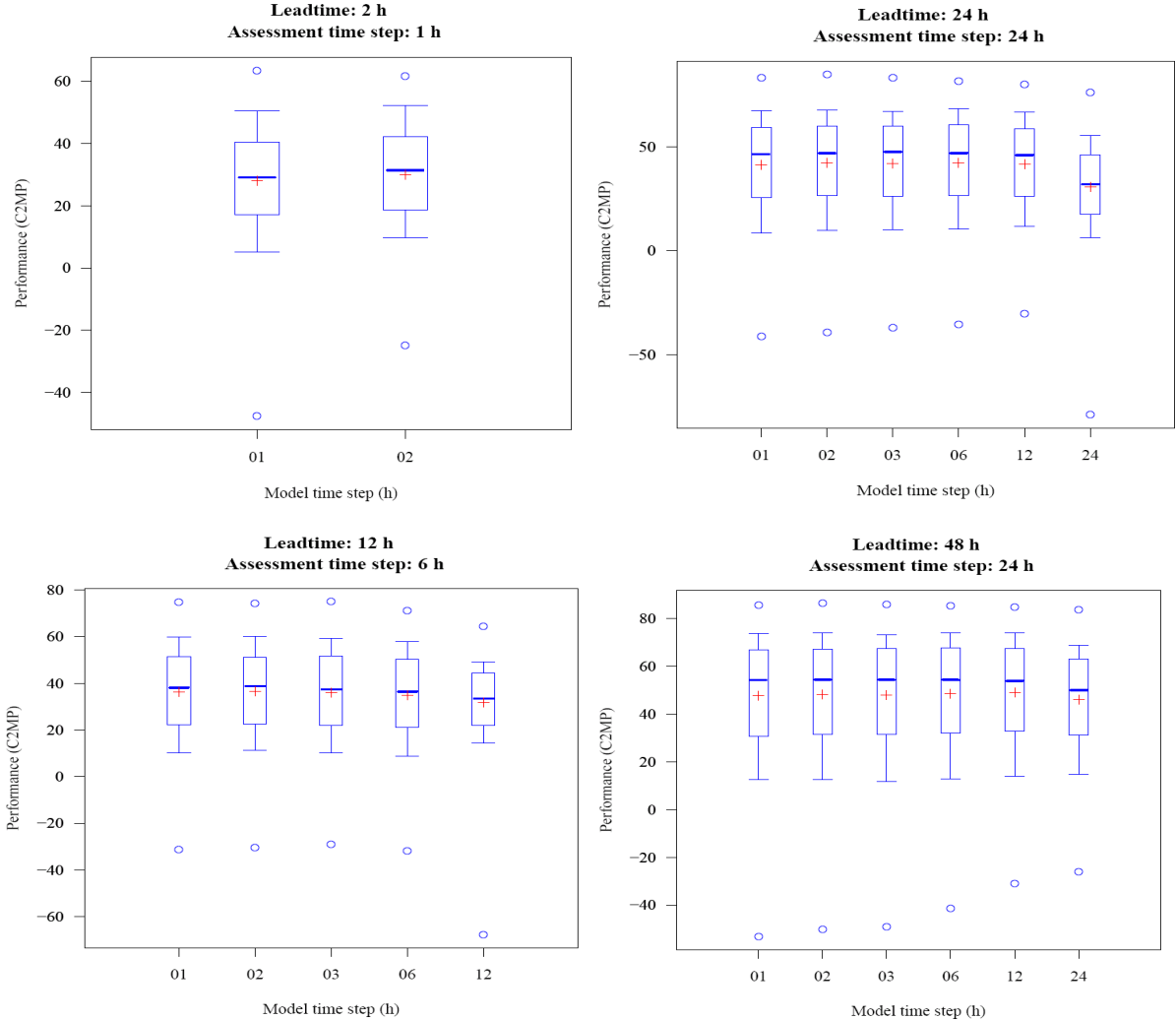


Figure 3.7: Example of graphs plotted between the performance criterion and model working at different time steps to reach a given lead time.

LT	2		3		6				12				18				24						36						48						72							
ATS	1	2	1	3	1	2	3	6	1	2	3	6	12	1	2	3	6	1	2	3	6	12	24	1	2	3	6	12	1	2	3	6	12	24	1	2	3	6	12	24		
MTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	2	2	3	3	2	2	3	2	2	2	3	2	2	2	2	3	2	2	2	3	2	2	2	2	2	3	2	2	2	2	3	2	2	2	2	2	3	2	2	2		
					3	6	6	3	3	6	6	3	3	3	6	6	3	3	6	6	3	3	3	3	6	6	3	3	3	6	6	3	3	3	3	6	6	3	3	3		
					6			6	6	12	12	6	6	6			6	6	12	12	6	6	6	6	12	12	6	6	6	12	12	6	6	6	6	12	12	6	6	6		
									12			12	12						12	24	24	12	12	12				12	12	12				12	12	12				12	12	12
																		24				24	24																			

LT Lead times
 ATS Assessment time steps
 MTS Model's time steps assessed for each time step on a given lead time.

Table 3.2: Model's time steps values on different assessment time steps to reach a given lead time.

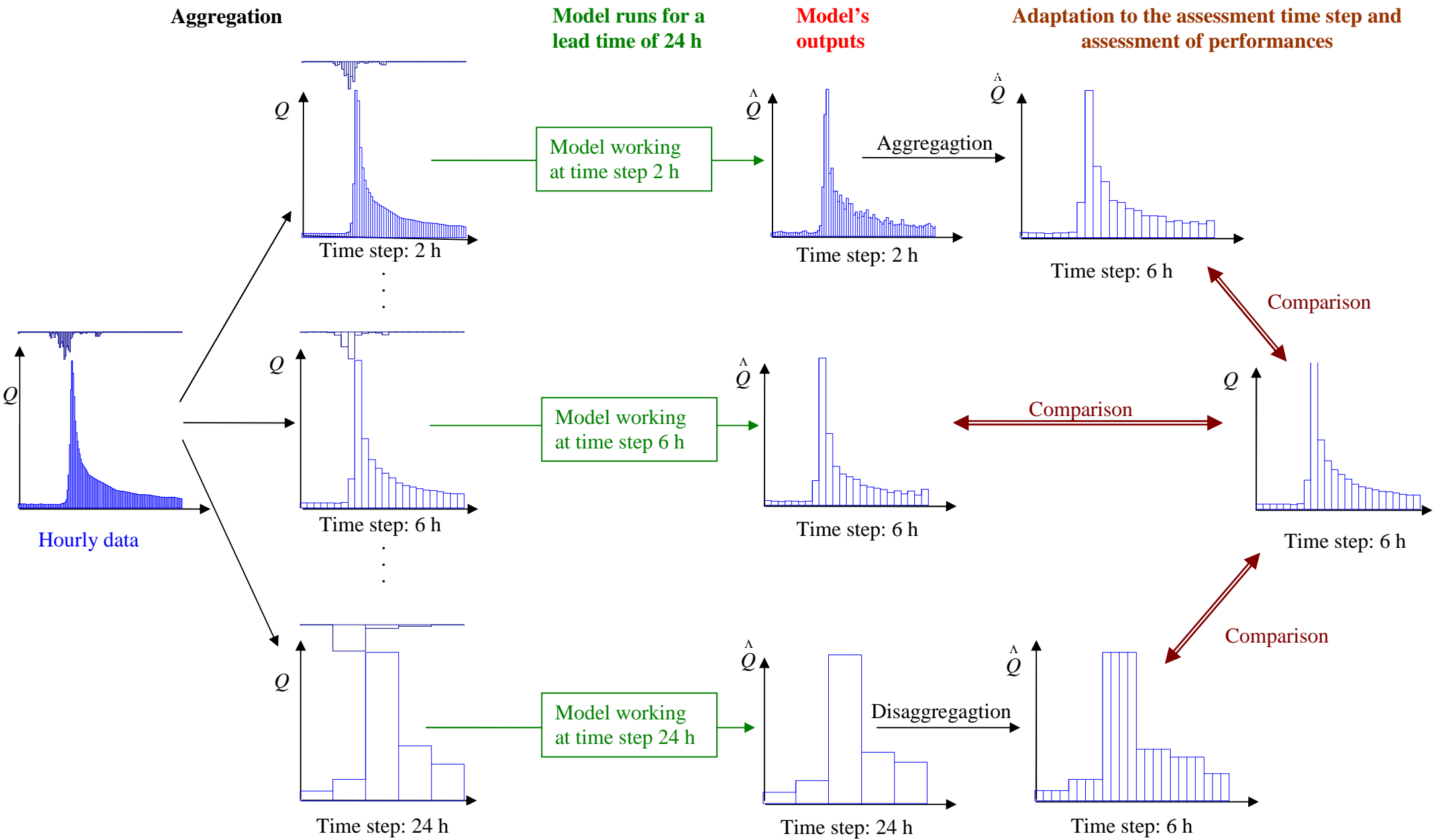


Figure 3.8: Example of comparison of forecasts using different time resolutions to reach a lead time of 24 h. Performances are assessed at time step of 6 h. Q is the observed streamflow while \hat{Q} is the calculated streamflow.

Chapter 4

RESULTS AND DISCUSSION

4.1 First results

In our first experiment, we assess the efficiency of the GR3P model working at different time resolutions (i.e. the different time steps we defined previously in chapter 3), by using persistence criterion over the whole available time series of our set of 178 catchments (i.e. on the whole periods) to reach the different lead times, we defined previously.

First results indicate that there is a group of time steps for which the model shows no significant difference of performance for most watersheds of our sample. The group consists of available model's time steps (TS) which are lower than or equal to the lead time and 12 h: if we call $G_{BMTS}(L)$ this group of time steps for lead time L, then $TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq \min(L, 12)$. Performance for model's time steps larger than these groups of model's time steps decreases. Examples of the graphs showing comparison of the performance are given in the figure 4.1.

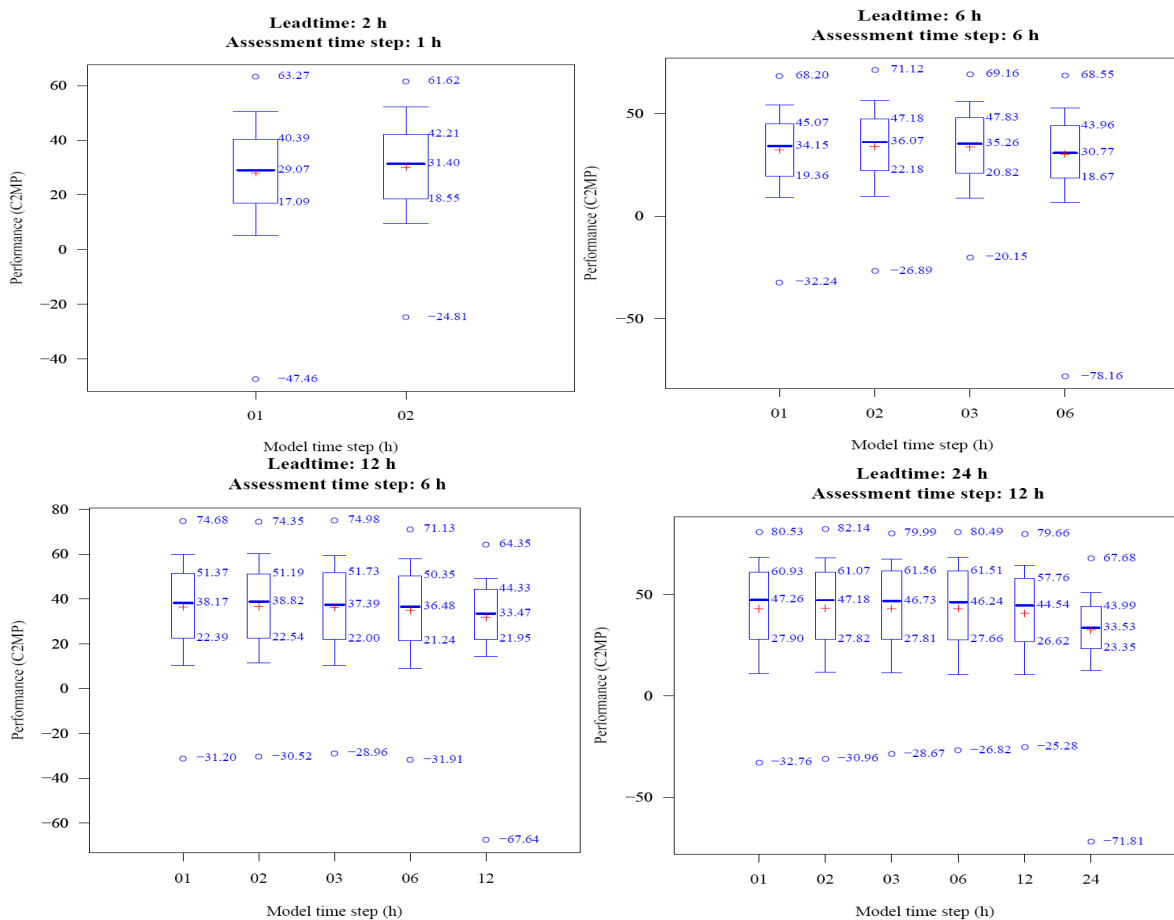


Figure 4.1: Examples of graphs showing the performance of different model's time steps to reach different lead times.

However we observe some different behaviours of the extreme values of the performance (i.e. minimum and maximum values) for different time steps within the group $G_{BMTS}(L)$. The situation becomes worst for the minimum values of performance when the model's time step equals to the given lead time.

The fact that a model shows better performance than another for watersheds where performances are bad for any model, indicates that the first model is more robust than the second model. Here the models working at time steps lower than the lead time are more robust than the model working at a time step equal to the lead time.

These results are the same whatever the assessment time step we use.

4.2 Can we see some differences for slow or fast catchments?

By using their discharge auto correlation we differentiated slow responding catchments to fast responding catchments. We have classified qualitatively the watersheds into four classes i.e. very peaky, peaky, smooth and very smooth, depending upon the values of discharge auto-correlation. If the value of auto-correlation is near to 0 then it means that watershed reacts very fast / peaky, if these values go near to 1 the reaction of watershed will tend to become much slower / smooth.

4.2.1 Analysis of the performances of different model's time steps on different watershed samples

To reach the lead times of 2 h and 3 h, the model's time steps equal to the lead time performs better for very slow reacting watersheds than a model using 1 h time step. For other watersheds there is no significant difference of performance between the models using different time steps.

The results also indicate that there is a group $G_{BMTS}(L)$ with $TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq 6h$ for lead times greater than or equal to 3 h, whatever the classes of watersheds we consider: in most cases, we can see no difference depending on catchment reaction. Examples of the performances are shown in the graphs given in figure 4.2.

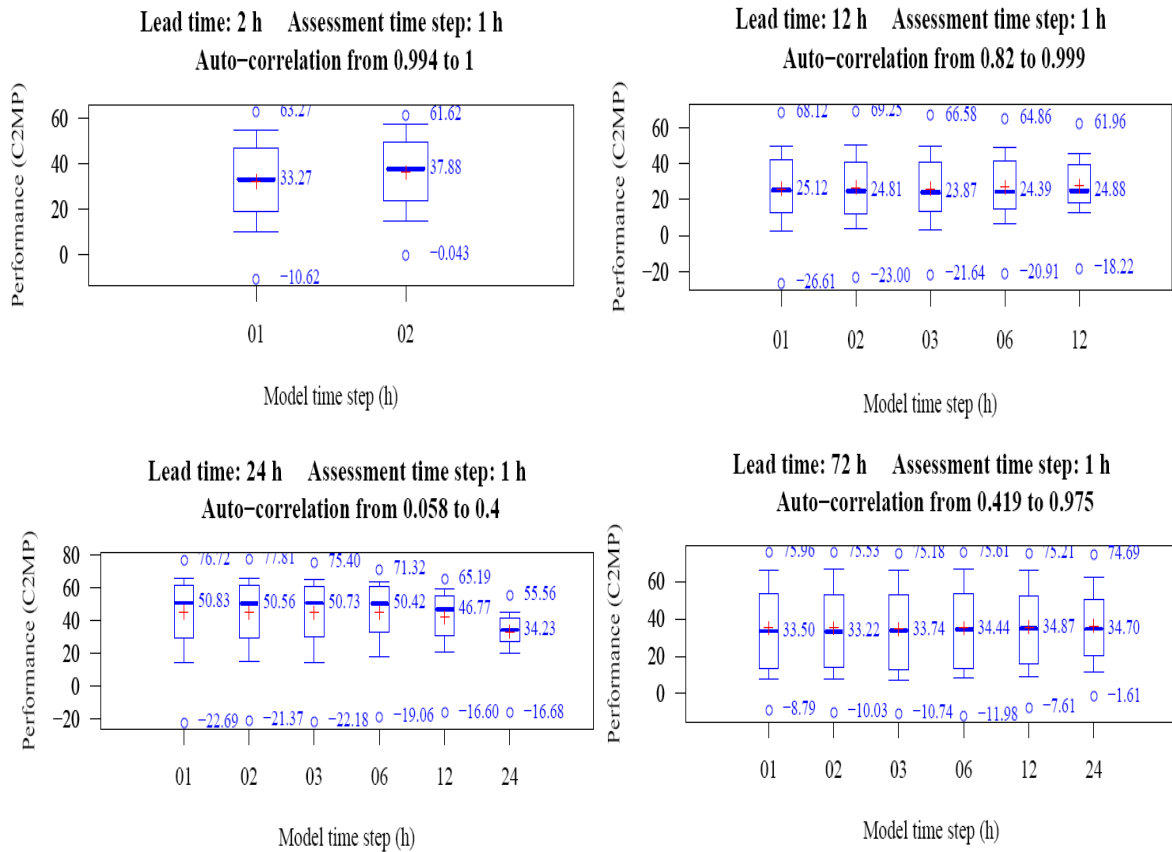


Figure 4.2: Examples of graphs showing the performance of different model's time step on different samples of watersheds.

4.3 How can we explain those first results?

Different hypotheses can be proposed to explain our first results.

- 1) There is no more information in the shortest time step data than in the larger time step data because:
 - a) For a larger model time step we have almost as much information of stream flows as in shorter model time steps for periods of little variation of stream flows, i.e. most periods.
 - b) The information in the smaller time step data is not large than in larger time step data because there are a lot of errors in these data (especially in precipitation data). In larger time step data these errors compensate each other so the information in the data for both shorter and larger time step are almost equal.
- 2) The model is not adapted to work with very precise information (shorter time step data).
- 3) Our criterion is not adapted to see those differences of information in data at different time steps.

4.4 Experiments to test our hypotheses

We estimate the performance of different model's time steps by taking into account only the periods of flow which are of high interest in this experiment.

- Only the high flows (when stream flow is higher than a certain proposed value) because the highest errors are made on high flows and also because it is an important operational demand.
- Only the large flow variation periods. Flow variations, most often are low and errors are small on these periods. Significant errors only happen on large stream flow variation time steps.

4.4.1 Analysis of the performance of model using different time steps when taking into account only the high flows

An experiment is done by taking into account for calibration and performance assessment (validation) only the flows higher than quantile 0.98 of discharge values: $q_{98}(Q)$.

The results differ from those of the previous experiment. They indicate that to reach a lead time of 2 h or 3 h, the model's time steps that are equal to the lead times perform significantly better. For lead times of 6 h, there is a group $G_{BMTS}(L)$ with $TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq 6h$ for which the model shows no significant difference of performance. To reach lead times greater than or equal to 18 h, the model's time step equal to assessment time step shows significantly better performance. For lead time of 12 h the model's time step equal to the assessment time step gives slightly better performance.

Thus the way we assess the model's performances (which assessment time step?) influences the choice of time step leading to the best performances: when we consider only high flows, the best performances are achieved with a time step equal to the time resolution used to assess the performances in most cases. Examples of the results are shown by the graphs in figure 4.3.

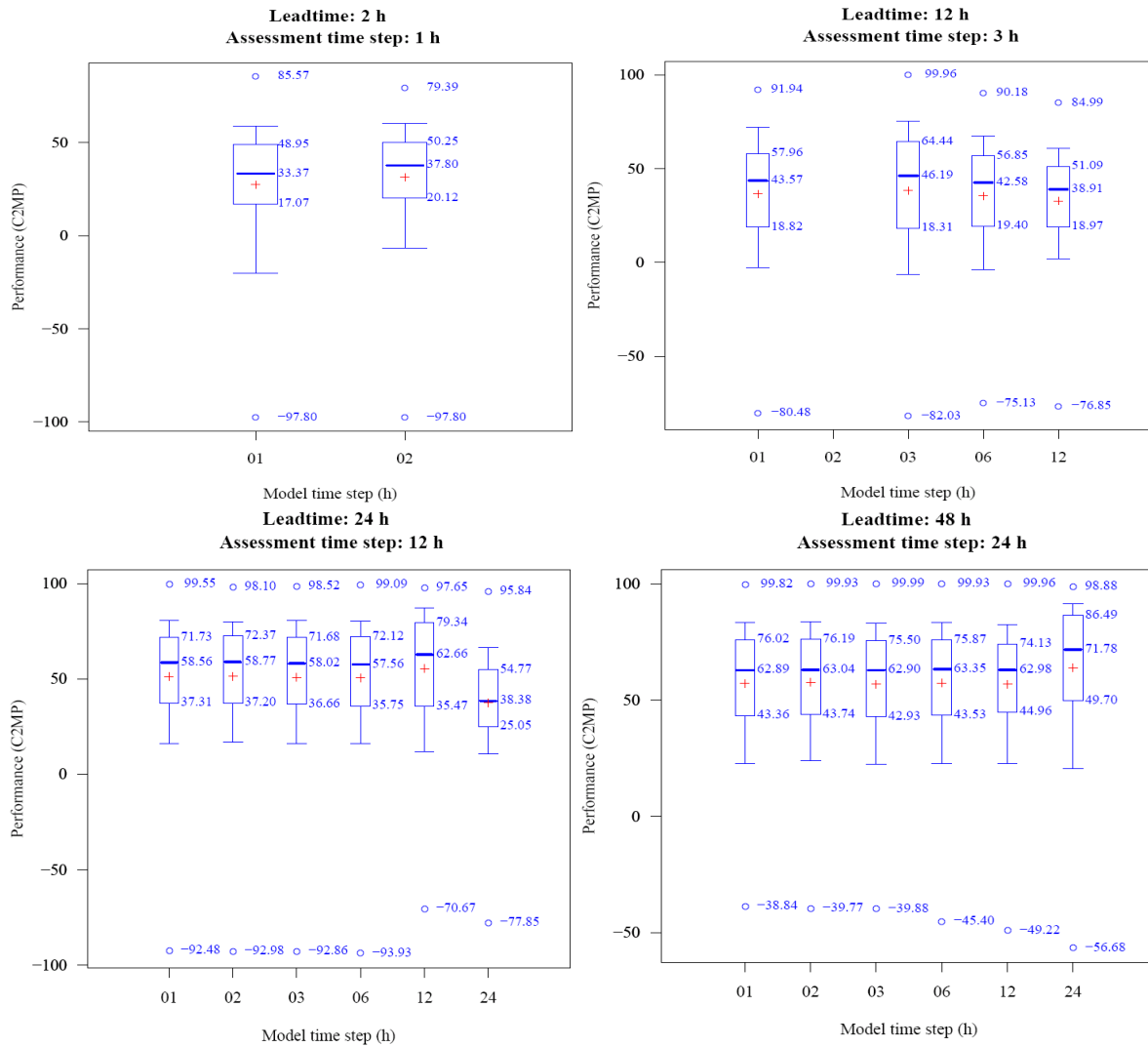


Figure 4.3: Examples of graphs showing some of the better performing model's time steps to reach different lead times when taking into account the high flows.

4.4.1.1 Does it depend on the reaction times of our catchments?

We also analyse the performance by dividing the catchments into different classes as discussed in section 4.2. The results are almost similar to the above results (i.e. results of section 4.4.1) for lead times 2 h and 3 h at slow reacting catchments and for the lead times between 12 h and 24 h on mostly fast and very fast reacting catchments. For lead times greater than 24 h, above results (i.e. model's time step equal to the assessment time step performs better than other model's time steps) are true on all classes of catchments.

4.4.1.2 Analysis of the performance of model using different time steps when taking into account the high flows on fast reacting catchments

Here we use another way to select fast reacting catchments. These are chosen by assessing model performance for different lead times with two different scenarios of future rainfall (i.e. perfect and null future precipitation) for all the 1040 catchments of our complete sample. Example for one catchment is shown in figure 4.4.

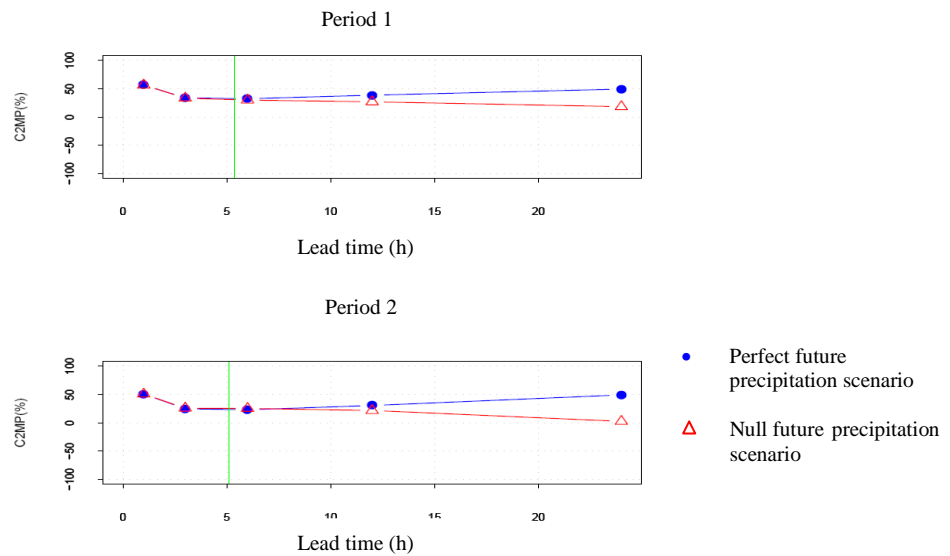


Figure 4.4: Example of graph showing the procedure to choose the fast reacting subsample of a catchment. L_3 is equal to 5 h for this catchment.

When the lead time is much smaller than the reaction time of the catchment then the future precipitation scenario is of little importance. Conversely when the lead time is much larger than the reaction time of the catchment then the future precipitation scenario is of significant importance and the performances obtained with two different scenarios are significantly different.

Lead time (L_3) corresponding to a difference of performance equal to 3 points between both future precipitation scenarios is then assessed for each catchment. Catchments whose L_3 is lower than 2 h are considered as our fast catchments.

An experiment is done by taking into account for calibration and performance calculation (validation) the flows higher than quantile 0.98 of discharge values: $q_{98}(Q)$, on these fast

reacting catchments. The results are almost the same as in the experiment of high flows on our sample of 178 catchments.

The results indicate that to reach a lead time of 2 h or 3 h, the model's time steps that are equal to the lead times perform significantly better. To reach lead times greater than or equal to 12 h, the model's time step equal to the assessment time step shows significantly better performance. Even for the 6 h lead time the model's time step equal to the assessment time step gives slightly better performance than other model's time steps. Examples of the results are shown by the graphs in the figure 4.5.

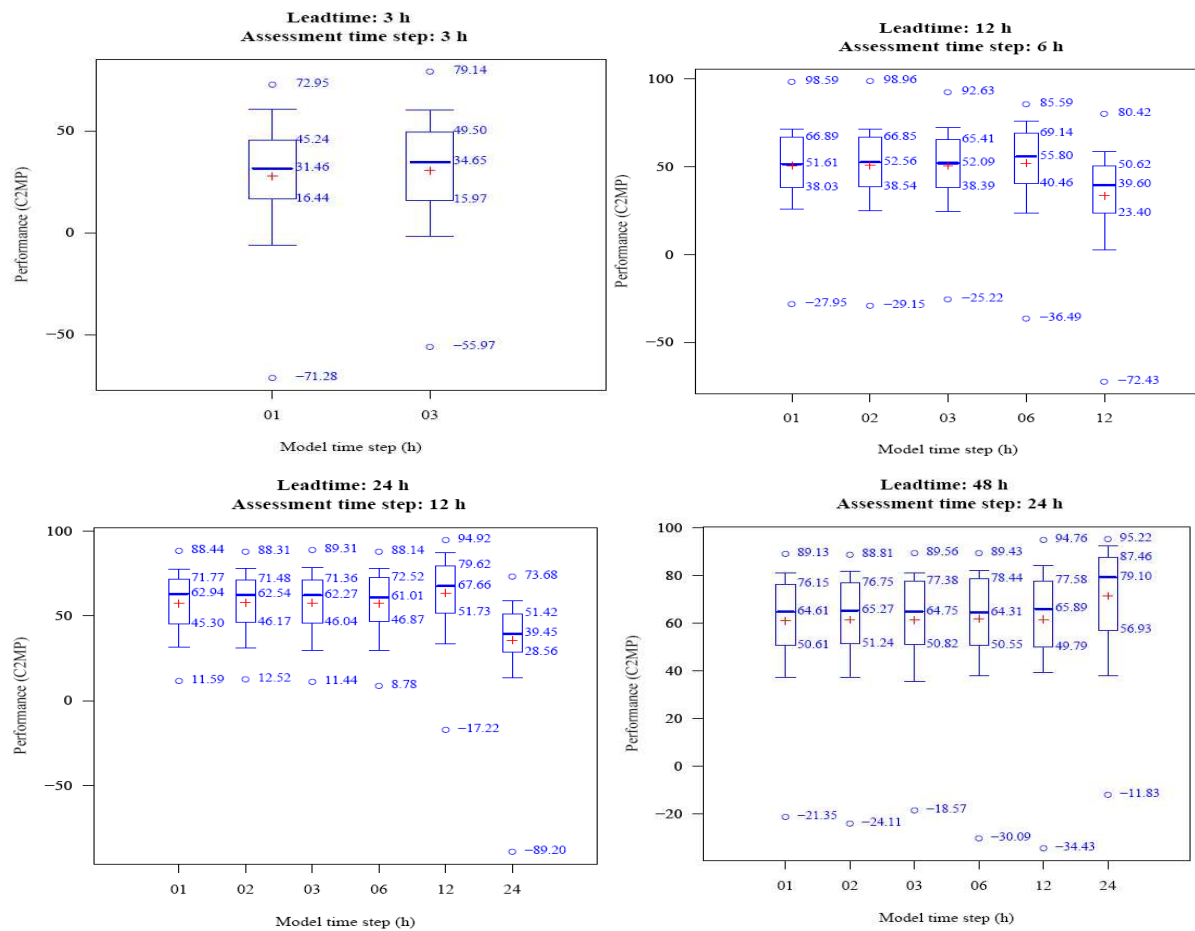


Figure 4.5: Examples of graphs showing some of the better performing model's time steps when taking into account the high flows on fast reacting catchments.

4.4.2 Analysis of the performance of model using different time steps when taking into account only the large flow variation

In the case of large flow variation we consider only time steps with flow variation higher than quantile 0.98 of flow variation: $q_{98}(\Delta Q)$. We observe by the results that to reach a lead time of 2 h or 3 h, the performance of model using time step equal to the lead time is better than

other model's time step. We observe a group $G_{BMTS}(L)$ with $TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq 6h$ for which the model shows no significant difference of performance for most watersheds of our sample. Thus we find similar results to the experiment realized over the whole periods of data.

4.4.3 Use of different performance criteria

Until now we used the persistence criterion. The persistence criterion puts a large emphasis on the highest errors. Persistence on logarithms of discharge is more sensitive to smaller errors present in the time series. Using the logarithmic form of the persistence criterion is another significant estimation of model performance.

We now make an experiment to assess the performance of our model using the logarithmic form of the persistence criterion.

Results indicate that, to reach the lead times lower than or equal to 3 h, the model's time steps equal to the lead time (i.e. 2 h and 3 h respectively), performs better for very slow reacting catchments. To reach the lead times between 6 h and 24 h there is a group $G_{BMTS}(L)$ of time steps lower than or equal to 6 h ($TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq 6h$) which shows no significant difference of performance but to reach lead times greater than or equal to 24 h, the group $G_{BMTS}(L)$ consists of time steps lower than or equal to 12 h ($TS_j \in G_{BMTS}(L) \Leftrightarrow TS_j \leq 12h$). Examples of graphs plotted for lead time 2 h and 3 h are shown in the figure 4.6.

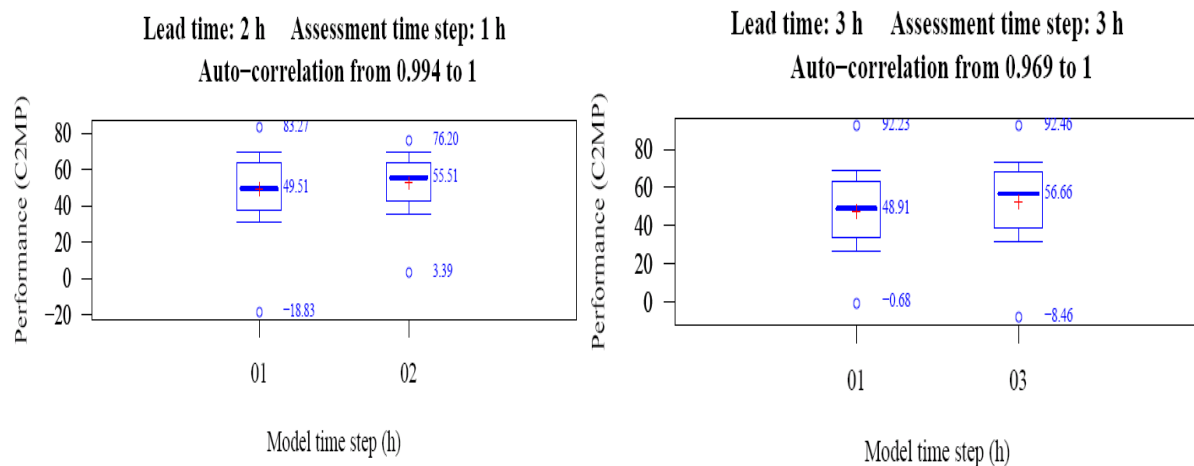


Figure 4.6: Examples of graphs showing the better performing model's time steps to reach lead times of 2 h and 3 h on very slow reacting watershed.

The results are almost similar to the experiment of large flow variation as described in section 4.4.2. Using a criterion which focuses less on floods leads to the same conclusion, so our first criterion may be considered to be adapted for our study.

4.5 Analysis of the performance of model using different time steps when precipitation scenario is taken as zero

We make also an experiment to compare the performances of different model's time steps by using the future precipitation scenario as zero/null. We do this experiment because the operational services people may have no forecast of rainfall so a perfect future precipitation scenario is much too optimistic.

4.5.1 Analysis of the performance of model using different time steps when taking into account the whole periods of data with a null future precipitation scenario

We observe by the results that to reach a lead time of 2 h or 3 h, the performance of model using time step equal to the lead time is better than other model's time step. The results also indicate that model's time step of 6 h shows significantly better performance than all other model's time steps to reach the lead times between 6 h and 18 h and the performance of model's time step of 12 h is significantly better to reach the lead times between 24 h and 36 h. To reach the lead times of 48 h and 72 h, the model's time step equal to 24 h shows significantly better performance than all other model's time steps. These results are the same whatever the time steps used for assessment.

So we can say that to reach our lead times there is not a group of model's time steps leading to the same best performances and that a model time step slightly lower than the lead time is the best solution (in terms of C_{2MP}).

We also analyse the performance on different classes of catchments and we observe that the results are almost similar to the above results (i.e. the performance of model using time step equal to the lead time is significantly better than other model's time step) for lead times 2 h and 3 h on all catchments classes. The results are also similar to above results (i.e. results of section 4.5.1) for the lead times greater than 6 h on all classes of catchments. Examples of some of the results are shown by the graphs in the figure 4.7.

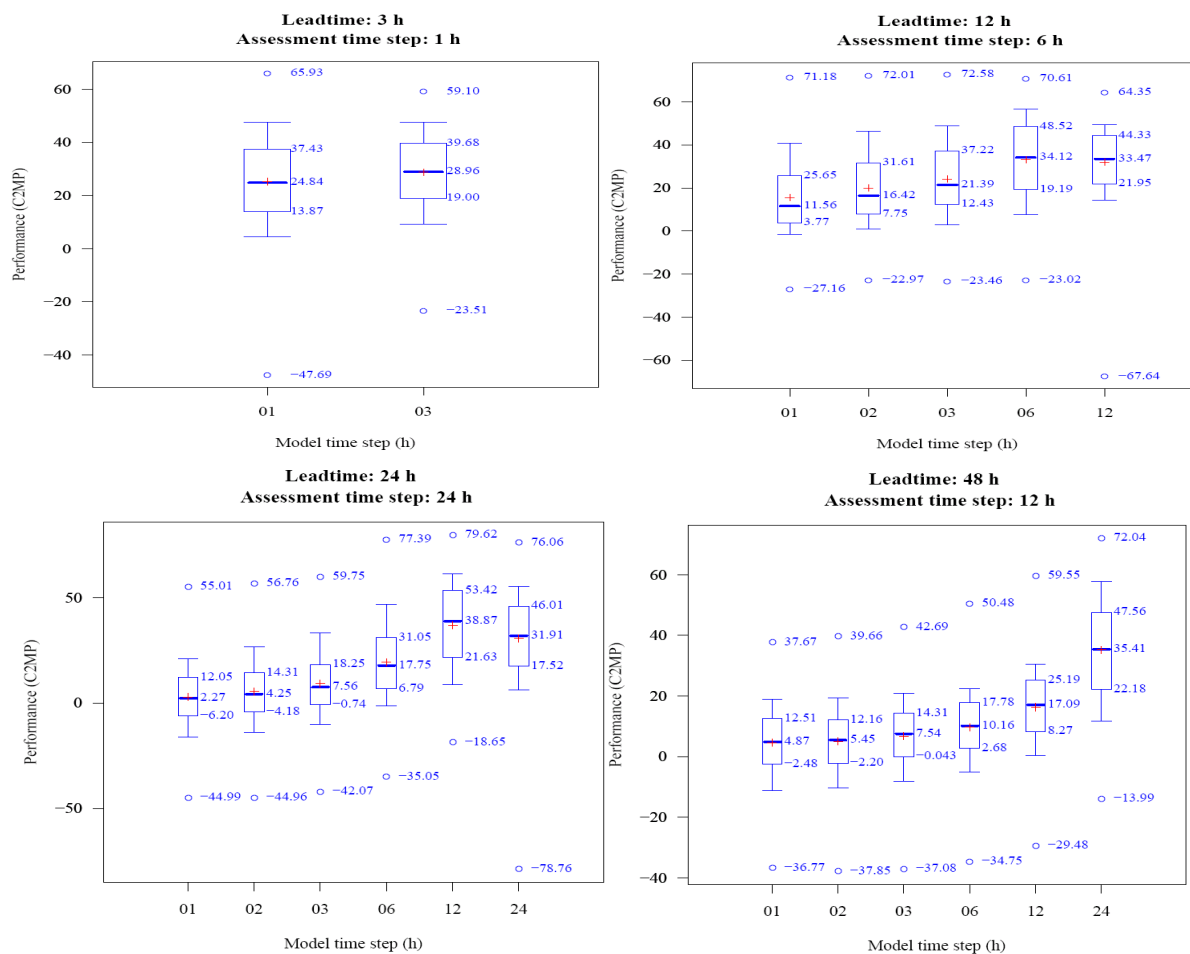


Figure 4.7: Examples of graphs showing the better performing model's time steps to reach lead times of 3 h, 12 h, 24 h and 48 h.

4.5.2 Analysis of the performance of model using different time steps when taking into account the high flows with a null future precipitation scenario

An experiment is done by taking into account for calibration and performance calculation (validation) the flows higher than quantile 0.98 of discharge values with a null future precipitation scenario. The results are almost the same as in both of the earlier experiments of high flows.

The results indicate that to reach a lead time of 2 h or 3 h, the model's time steps that are equal to the lead times perform significantly better. To reach lead times greater than or equal to 18 h, the model's time step equal to the assessment time step shows significantly better performance. There is a group $G_{BMTS}(L)$ with model's time steps 3 h and 6 h for lead time 6 h and a group with model's time steps 6 h and 12 h for lead time 12 h, for which the model

shows no significant difference of performance. The results are almost similar on all the classes of catchments.

Here also the way we assess the model's performances (which assessment time step?) influences the choice of time step leading to the best performances. Examples of the results are shown by the graphs in the figure 4.8.

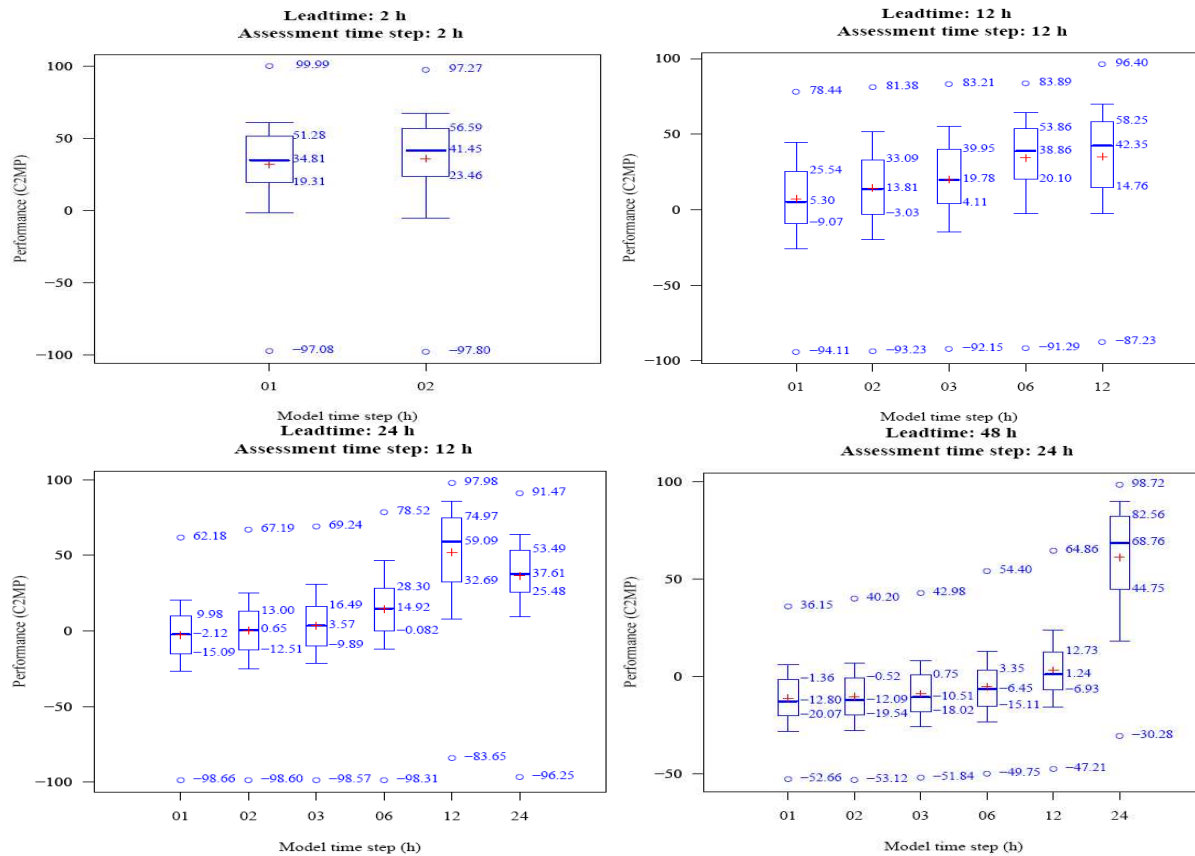


Figure 4.8: Examples of graphs showing some of the better performing model's time steps when taking into account the high flows with null future precipitation scenario.

4.6 Discussion

Our results show that there are two groups of lead times with different behaviours of model's time steps: 2 h and 3 h as the first group, lead times greater than or equal to 6 h as the second group. This can be linked with the fact that a lot of catchments of our sample react in a very few hours.

For the first group of lead times, the model's time steps equal to the lead times perform significantly better in almost all of the experiments made to answer the questions raised by the first results. The results differ from one experiment to another for the second group of lead times but most often the model working at a time step equal to the assessment time step

performs significantly better than other models when assessing the performances on high flows.

4.6.1 Focusing on high flows or on large flow variations?

When taking into account only the high flows, we observe that to reach the lead times greater than or equal to 12 h on all classes of catchments and also often for 6 h lead time on fast reacting catchments, the model's time steps equal to assessment time steps show significantly better performance than others. This is not the case when we consider large flow variation.

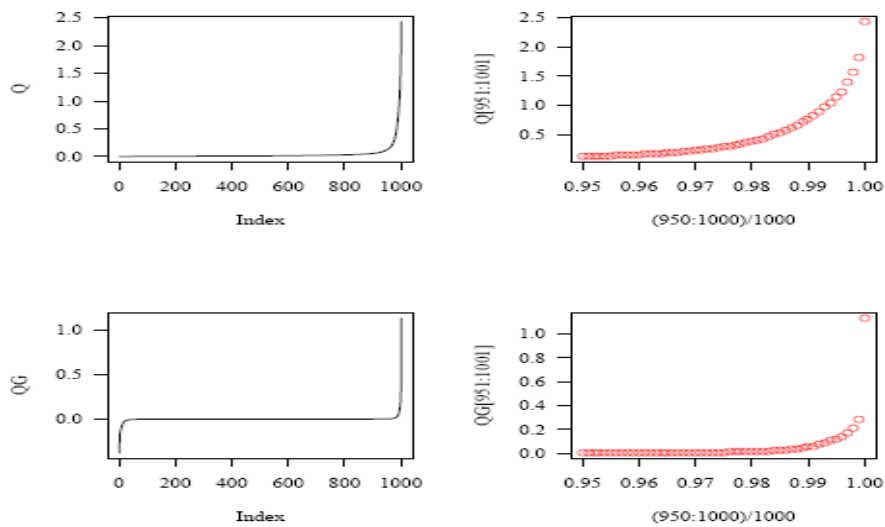


Figure 4.9: Examples of graphs showing the samples (on left) and a zoom (on right) of discharge distribution (top) and the gradient of discharge distribution (bottom).

To explain such a difference we compare the upper tails of discharge and discharge variation distributions. Figure 4.9 shows an example of discharge and discharge variation distributions for one catchment. In first graphs (on left) we have complete discharge and discharge variation distributions while in the graphs (on right) we have a zoom over quantiles 0.95 to 1.00. If we consider a subsample of dates with flow higher than quantile 0.98: $q_{98}(Q)$ then this subsample really differs from the complete sample. It is not the case for subsample defined by flow variations higher than quantile 0.98: $q_{98}(\Delta Q)$. This is true for a huge majority of our catchments.

This can explain why our results for high flows are different from the first experiment results (with the whole periods) while the results of the same experiment with high flow variations are analogous to the first experiment results.

We learn from these experiments that on the highest flows (i.e. events of major operational interests), it is necessary to use a short time step if we want the most accurate forecast at a high time resolution (shorter assessment time step).

4.6.2 The best model's time step is the one used for performance assessment

The fact that the best model's time step is most often the assessment time step can be explained by the loss of information in aggregation and disaggregation processes. When we aggregate data, we may assume that we lose some information (in figure 3.8, for data from time step 1 h to time step 6 h). But if the aggregation comes after the model's run, it aggregates the information with errors within the input and with errors made by the model (in its output). It may lead to worse aggregated output than if the aggregation could have been done before the model's run (in figure 3.8, aggregation of output data from time step 2 h to time step 6 h). When we disaggregate the model's output, we have no information on how to make smaller resolution data from coarser resolution data. Thus we certainly introduce much error in this process (in figure 3.8, disaggregation of output data from time step 24 h to time step 6 h). This is why using a time step for the model larger than the assessment time step leads to worse results.

When focussing on the high flows (flows higher than quantile 0.98 of discharge values: $q_{98}(Q)$), the shortest model's time step gives significantly better performances for high time resolution (for the shortest assessment time step) than other time steps for lead times greater than or equal to 12 h on all the classes of catchments and for lead time 6 h on fast reacting catchments. It proves that our model is adapted to work with precise information (shorter time step data).

4.6.3 Criteria of performance

Persistence criterion and persistence criterion on logarithms of discharge values are used to calculate the performances of different model's time steps. We see the same results by using two different criteria even if the second one is known to be more sensitive to the smaller errors than the persistence criterion. We learn that using a criterion which focuses less on floods leads to the same conclusion, so our first criterion may be considered to be adapted for our study.

4.6.4 Operational learning from our study

We observe the similar results when focusing on high flows with a null future precipitation scenario as some operational forecasters may have to do. Even using a null future precipitation scenario with a focus on high flows, it is still useful to choose a high time resolution model (i.e. model working on shortest time step) in order to get the most accurate forecast. Thus our conclusions are still of interest for operational forecasters.

Chapter 5

CONCLUSIONS

This study is an analysis of the performance of a forecasting model, GR3P, on the basis of time resolution (i.e. when using different time steps) to reach lead times varying from a few hours to a few days. The analysis is achieved on a large sample of data of 178 catchments to ensure the generality of our conclusions. Performances are also assessed through different time resolutions.

When calculating the performances by the criterion of persistence on continuous basis (i.e. over a whole period of data), it appears that we may use any of model's time steps within a group, which shows no significant difference of performances, to reach the larger lead times. This group consists of available model's time steps which are lower or equal than the lead time and than 12 h. These results are valid for any type of catchment (fast or slow reacting catchments). Some more experiments were designed to test some hypotheses to explain these questioning results. These experiments are made with a focus on high flows, large flow variations, different precipitation scenarios and different performance assessment criteria. The results obtained by these experiments show that most often the model working at a time step equal to the assessment time step performs significantly better than other models when assessing the performances on high flows. Consequently assessing at hourly time step proves that the shortest time step data e.g. hourly time step, has more precise information in it than a larger time step data for high flow periods and that our model can benefit from this precise information.

According to the findings of our study, the forecasters should use the high time resolution (shortest assessment and model's time step) to issue an accurate forecast. Even if the forecasters work in difficult situation (no future precipitation knowledge), the high time resolution model is still better.

REFERENCES

- Dunsmore, S. J., et al. 1986 "Antecedent soil moisture in design stormflow estimation." ACRU Rep. 23, Dept. Agric. Eng., Univ. of Natal, Pietermaritzburg, South Africa, 114 pp.
- Garrote, L. and Bras, R.L., 1995. "A distributed model for real-time flood forecasting using digital elevation models." Journal of Hydrology 167(1-4) : 279-306.
- Hughes, D. A. (1993). "Variable time intervals in deterministic hydrological models." Journal of Hydrology 143(3-4): 217-232.
- Kitanidis, P. K. and R. L. Bras (1980). "Real-time forecasting with a conceptual hydrologic model 2. Applications and results " Water Resources Research 16(6): 1034-1044.
- Klemeš, V. (1986). "Operational testing of hydrological simulation models." Hydrological Sciences Journal 31(1): 13-24.
- Nalbantis, I. (1995). "Use of multiple-time-step information in rainfall-runoff modelling." Journal of Hydrology 165: 135-159.
- Perrin, C., *et al.* (2001). "Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments." Journal of Hydrology 242(3-4): 275-301.
- Refsgaard, J. C. (1997). "Validation and intercomparison of different updating procedures for real-time forecasting." Nordic Hydrology 28(2): 65-84.
- Shamseldin, A. Y. and K. M. O'Connor (1996). "A nearest neighbour linear perturbation model for river flow forecasting." Journal of Hydrology 179(1-4): 353-375.
- Tangara, M. (2005). Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global. Ecole doctorale: Sciences de la Vie et de la Terre, Laboratoire: Hydrologie et Environnement. Paris, France, Ecole pratique des hautes études de Paris. Doctorate: 374.

Toth, E. and A. Brath (2007). "Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modelling." Water Resources Research 43(11).

Wagener, T., *et al.* (2004). Rainfall-Runoff modelling in gauged and ungauged catchments. London, Imperial college press. Book: 374.

Yang, X. and C. Michel (2000). "Flood forecasting with a watershed model: a new method of parameter updating." Hydrological Sciences Journal 45(4): 537-546.

APPENDIX A

Some important notations used in this study

t	Current time step at which the forecast is issued
l	Lead time
N	Total number of time steps
$q_x(X)$	Quantile x of X (x between 0 and 100%)
P_n	Net precipitation,
E_n	Net potential evapotranspiration
P_s	Part of precipitation going to the production store
E_s	Potential evapotranspiration from the production store
P_r	Direct flow
Perc	Part of the stored water contents which percolates
R_t	Level of the water content in the routing store
Q_t	Observed streamflow at time step t
X_1	Maximum capacity of the quadratic routing store (level noted R , in mm)
X_2	Adjustment coefficient of effective rainfall
X_3	Base width of the unit hydrograph (UH)
α	Exponent of the unit hydrograph
β	Exponent of the final correction (last update)
$\hat{R}_{t t}$	Water level in the store at the end of time step t
\hat{Q}	Calculated or estimated streamflow
$\hat{\hat{Q}}_{t+1 t}$	Forecasted discharge at $t+1$ knowing discharge at time step t
$\hat{\hat{\hat{Q}}}_{t+1 t}$	Estimation of the streamflow value at time step $t+1$, done at time step t ,
\bar{Q}	Mean observed streamflow

APPENDIX B
Details of the routing reservoir updating phase

Equation A is obtained as follows. Let us assume that

- R_t , the store content prior to the draining process,
- Q_t , is the output of the routing store over the time step, due to the draining process,
- $\hat{R}_{t|t}$, is the resulting reservoir content at the end of the time step (see appendix A for complete notations).

The chosen reservoir is such that the output, Q_t , is related to the reservoir water content, R , according to the following relationship:

$$Q_t = f(R_t) = \frac{R_t^2}{R_t + X_1} \text{ Where } X_1 \text{ is the capacity of the routing store.}$$

At the end of the present time step, updated R_t is: $\hat{R}_{t|t} = f^{-1}(Q_t) - Q_t$ (the exact value giving the observed discharge minus this discharge)

$$Q_t = \frac{(\hat{R}_{t|t} + Q_t)^2}{\hat{R}_{t|t} + Q_t + X_1} \text{ or equivalently:}$$

$$Q_t^2 + \hat{R}_{t|t} Q_t + X_1 Q_t = Q_t^2 + 2 \hat{R}_{t|t} Q_t + \hat{R}_{t|t}^2$$

This can be simplified into:

$$\hat{R}_{t|t}^2 + Q_t \hat{R}_{t|t} - X_1 Q_t = 0$$

The root for this quadratic equation is:

$$\hat{R}_{t|t} = \frac{\sqrt{Q_t^2 + 4X_1 Q_t} - Q_t}{2}$$

APPENDIX C

Different values of fixed parameters tested on a set of time steps to choose best values of fixed parameters

Name of parameter	TS	Values of Parameters																	
Storage	1	225	250	275	300														
	2		250	275	300		350												
	3		250	275	300		350												
	6		250	275	300	325	350												
	12		250	275	300	325	350												
	24			275	300	325	350	375											
Exponent of unit hydrograph	1	1.00	1.25	1.50															
	2		1.25	1.50	1.75	2.00													
	3		1.25	1.50	1.75	2.00													
	6			1.50	1.75	2.00	2.25												
	12			1.50	1.75	2.00	2.25												
	24			1.50	1.75	2.00	2.25	2.50	2.75										
Coefficient of percolation	1							4.50	4.75	5.00	5.25	5.50							
	2		2.00	2.50	3.00	3.50	4.00	4.50	4.75	5.00		5.50							
	3		2.00	2.50	3.00	3.50	4.00	4.50											
	6		2.00	2.50	3.00	3.50	4.00	4.50											
	12		2.00	2.50	3.00	3.50	4.00												
	24	1.50	2.00	2.50	3.00	3.50													
Coefficient of final correction	1												0.375	0.40	0.425	0.450	0.475	0.50	
	2		0.100				0.200			0.300	0.325	0.350	0.375	0.40	0.425	0.450	0.475	0.50	
	3		0.100				0.200			0.300	0.325	0.350	0.375	0.40				0.50	
	6		0.100	0.125	0.150	0.175	0.200	0.225	0.25	0.300				0.40				0.50	
	12		0.100	0.125	0.150	0.175	0.200	0.225	0.25										
	24	0.075	0.100	0.125	0.150	0.175	0.200												

Red= best chosen value of fixed parameter

TS= Set of tested time steps