Alma Mater Studiorum - Università di Bologna

Analysis and Developments of Uncertainty Processors for Real Time Flood Forecasting

Presentata da: Gabriele Coccia

Coordinatore Dottorato:

Prof. Rolando Rizzi

Relatore:

Prof. Ezio Todini

ii

Contents

List of Figures

List of Tables

Executive Summary

Hydrologic risk has always been an extremely relevant issue, due to the severe consequences that may be generated by droughts and floods in terms of casualties and economic losses. More specifically, floods are natural phenomena, often catastrophic, and cannot be avoided, but their hazard and consequent damages can be reduced if they are predicted sufficiently in advance. Flood forecasting plays an essential role in the hydro-geological and hydrological risk prevention and reduction of losses. In the last decades, an increasing number of real time flood forecasting systems have been made operational. Such systems give important information on the evolution of future events, which helps reducing the associated uncertainty.

Dealing with future events means inevitably dealing with uncertainty issues. One can try to reduce such uncertainty by using *avatars* of the future, namely forecasting models; nonetheless, models are imperfect, which means that we are still left with a residual uncertainty on what will actually happen. The present work is aimed at defining what is meant with predictive uncertainty, at discussing its role in real time flood forecasting and at presenting a Bayesian methodology for its assessment.

After a brief introduction to the topic in question, Chapter 2 starts explaining which is the role of uncertainty in the decision making processes and giving the definition of Predictive Uncertainty (PU) according to Krzysztofowicz (1999) and Todini (2008). Afterwards, some important elements related to the predictive uncertainty, such as the need of identifying probabilistic alarm thresholds and making available probabilistic forecast within

a horizon time, are pointed out. At the end of the chapter a brief review of the existing approaches for predictive uncertainty assessment is done in which two statistical processor groups are presented. The first group is aimed at solving the discrete probability problems while the second group tackles the continuous probability case, which as will be discussed in the sequel is more relevant to the flood forecasting case. Among the existing continuous uncertainty processors, three will be analyzed more in detail: the Hydrological Uncertainty Processor of Krzysztofowicz (Krzysztofowicz, 1999), the Bayesian Model Averaging of Raftery (Raftery, 1993) and the Quantile Regression (Koenker, 2005; Weerts et al., 2010). Concerning the discrete uncertainty processors, a methodology to assess the flooding probability will be presented and it will be shown how it can be applied according to different methodologies.

In Chapter 3, a new Bayesian methodology for predictive uncertainty assessment will be introduced, the Model Conditional Processor (MCP). The basic ideas underlying the MCP will be evolved to reach a more complex structure able to combine several deterministic model forecasts in order to provide essential probabilistic knowledge like the predictive uncertainty distribution, the probability to exceed a prefixed level in a particular moment or within a time horizon together with the probability of the flooding time.

In Chapter 4, two applications of the MCP will be shown. The first one will concern the Multi-Model approach and the study case will be the Baron Fork River, in Oklahoma, USA. The data set provided by the NOAAs National Weather Service, within the DMIP 2 Project, allowed two distributed models, the TOPKAPI model (Todini and Ciarapica, 2001; Liu and Todini, 2002) and the TETIS model (Francés et al., 2007; Vélez et al., 2009), to be calibrated and a data driven model to be implemented using the Artificial Neural Network. The three model forecasts will be combined to reduce the PU and to improve the probabilistic forecast taking advantage of the different models capabilities. A comparison between the proposed methodology (MCP) and the Quantile Regression technique will be carried out. The sec-

ond application will concern the Multi-Temporal approach (Krzysztofowicz, 2008) and the study case will be the Po River, in Italy. In this application the forecasts provided by the flood forecasting system of the Department of Civil Protection of the Emilia Romagna region will be processed in order to evaluate the MCP capability to assign the correct flooding probability within a horizon time and the flooding time probability.

Finally, the results obtained in the two study cases mentioned above will be analyzed in order to proof the MCP good performance and its usefulness in the probabilistic forecasting assessment in terms of clear benefits towards the decision making process.

Chapter 1

Introduction

The present work concerns the real time flood forecasting, with particular attention to operational applications in flood emergency management.

The hydrologic risk (and the hydro-geologic one, closely related to it) is, and has always been, a very relevant issue, due to the severe consequences that may be provoked by a flooding or by waters in general in terms of human and economic losses. This is reflected by the Italian Civil Protection Department:

In Italy, in the 20th century, due to floods and landslides:

- *more than 29,500 areas affected by flooding*

- *more than 31,500 areas affected by landslides*

- *10,000 considering dead, injured and missing people*

- *350,000 evacuated*

- *thousands of houses have been destroyed or damaged*

- *thousands of bridges have been destroyed or damaged*

- *hundreds kilometers of roads and railways have been destroyed or damaged*

- *approximately 8.000 square kilometers (2.6% of the entire surface of Italy) defined as high and very high flood-risk areas. (after: Project AVI CNR)*

Floods are natural phenomena, often catastrophic, and cannot be avoided, but their damages can be reduced if they are predicted sufficiently in advance. For this reason, the flood forecast plays an essential role in the hydro-geological and hydrological risk prevention. Thanks to the development of sophisticated meteorological, hydrologic and hydraulic models, in recent decades the flood forecasting has made a significant progress. Real time flood forecasting systems, based on the complex modeling of the natural physical processes that trigger the flood events, have been made operational. Such systems give important information about the evolution of future events; nonetheless, models are imperfect, due to the structural and parametric approximations and simplifications of hydrological models together with the input data provided by the meteorological forecasts. This means that we are still left with a residual uncertainty on what will actually happen. In this thesis, this type of uncertainty is what will be discussed and analyzed.

Dealing with future events means inevitably dealing with the uncertainty issue. Hence, from an operational standpoint, flood forecasting cannot be tackled without taking into account the associated uncertainty issue. In this sense, it is possible to affirm that the ultimate aim of forecasting systems is not to reproduce the river behavior, but this is only a means through which reducing the uncertainty associated to what will happen as a consequence of a precipitation event. In other words, the main objective is to assess whether or not preventive interventions should be adopted and which operational strategy may represent the best option.

The main problem for a decision maker is to interpret model results and translate them into an effective intervention strategy. This conversion is a complex task, especially if more than one model is available providing different forecasts for the same problem. In fact, not always more information

or data helps to take a decision, on the contrary their interpretation can be more difficult, and so increase the confusion instead of providing more solid basis for decision-making. This issue often reflects the difficulties in communication between scientific and technical staffs. On one hand, the former are mainly focused on improving models predictive ability, tying to include into the models they develop all the available knowledge about the physical process as well as the most advanced modeling techniques. On the other hand, the latter have to use the information provided by these models to take decisions, therefore they will always have to deal with uncertain data, no matter the greatest improvements that can be made in hydrological modeling. Hence, the uncertainty assessment represents the key link between the two different teams. Namely, the scientist must try to reduce and assess the predictive uncertainty while the technician should understand and base their actions on the estimation of such uncertainty.

To make this possible, it is necessary to clearly define what is meant by uncertainty, since in the literature confusion is often made on this issue. Therefore, the first objective of this thesis is to clarify this concept, starting with a key question: is the decision maker interested in knowing the model uncertainty or the uncertainty of the reality? In other words, should be the choice of the intervention strategy to adopt based on the evaluation of the model prediction based on its ability to represent the reality or on the evaluation of what actually will happen on the basis of the information given by the model forecast?

Once the previous idea is made unambiguous and well explained, the other main concern of this work is related to the importance and responsibility of the role of those who must intervene to prevent human and economic losses caused by a flood event. The decision making process must start from an analysis as objective as possible of the deterministic model forecasts in order to estimate the probability of a certain event occurrence.

With this in mind, this work is finally aimed at developing a tool that can provide an effective decision support, making possible doing objective and

realistic risk evaluations. In particular, such tool should be able to provide an uncertainty assessment as accurate as possible. This means primarily three things: it must be able to correctly combine all the available deterministic forecasts, it must assess the probability distribution of the predicted quantity (level, discharge, volume) and it must quantify the flooding probability. Furthermore, given that the time to implement prevention strategies is often limited, the flooding probability will have to be linked to the time of occurrence. For this reason, it is necessary to quantify the flooding probability within a horizon time related to that required to implement the intervention strategy and it is also necessary to assess the probability of the flooding time.

Chapter 2

Research Framework

2.1 Decision Making Under Uncertainty

In the last decades, the interest in assessing uncertainty in models forecasts has grown exponentially within the scientific communities of meteorologists and hydrologists. In particular, the introduction of the Hydrological Uncertainty Processor (Krzysztofowicz, 1999; Krzysztofowicz and Kelly, 2000), aimed at assessing predictive uncertainty in hydrological forecasts, has created the basis for the estimation of flood predictive uncertainty.

Flood emergency management requires adopting operational decisions in real time that may lead to dramatic consequences (economical losses, casualties, etc.). The hardest obstacle the managers have to deal with is the uncertainty on the future evolution of events. Decision theory (De Groot, 1970; Raiffa and Schlaifer, 1961) studied this problem and provided the most appropriate solutions for taking decisions under uncertainty. This approach consists in minimizing the expected value of a utility function $U(y)$ representing the losses, or more in general the manager perception of them, as a function of a predictand that will occur at a future time (such as a future discharge or water stage in a cross section). This quantity is unknown at the time of the decision (t_0) and the aim of forecasting is to assess its probability of occurrence, in terms of a predictive uncertainty probability density

function.

In the case of flood forecasting, predictive uncertainty can be defined as the uncertainty that a decision maker has on the future evolution of a predictand that he uses to make a specific decision.

In order to fully understand and to appreciate what is actually meant by predictive uncertainty, it is necessary to realize that what will cause the flooding damages is the actual future realization of the discharge and/or the water level that will occur, not the prediction generated by a forecasting model; in other words the damages will occur when the actual water level y_t and certainly not if the prediction $\widehat{y}_t$ will overtop the dyke level y_D (Todini, 2009). Therefore a utility/damage function at any future time $(t > t_0)$ must be expressed as a function of the actual level that will occur at time t:

$$\begin{cases} U\left(y_t\right) = 0 & \forall y_t \leq y_D \\ U\left(y_t\right) = g\left(y_t - y_D\right) & \forall y_t > y_D \end{cases} \tag{2.1}$$

where $g(\cdot)$ represents a generic function relating the cost of damages and losses to the future, albeit unknown water stage y_t. In this case the manager, according to the decision theory (De Groot, 1970; Raiffa and Schlaifer, 1961), must take his decisions on the basis of the expected utility $E\left\{U\left(y_t\right)\right\}$. This value can be estimated only if the probability density function of the future event is known, and it can be written as:

$$E\left\{U\left(y_t\right)\right\} = \int_0^{+\infty} U\left(y_t\right)f\left(y_t\right)dy_t \tag{2.2}$$

where $f\left(y_t\right)$ is the probability density expressing our incomplete knowledge (in other words our uncertainty) on the future value that will occur. This density, which can be estimated from historical data, is generally too broad because it lacks the conditionality on the current events. This is why it is essential to improve this historical probability distribution function by more realistically using one or more hydrological models able to summarize all the available information (like the rain forecast, the catchment geomorphology,

the state of the river at the moment of the forecast, etc. . .) and to provide a more informative density $f\left(y_t|\widehat{Y}_{t|t_0}\right)$, which expresses our uncertainty on the future predictand value after knowing the models' forecasts issued at time t_0, namely $\widehat{Y}_{t|t_0} \triangleq \left[\widehat{y}_{1t|t_0}, \widehat{y}_{2t|t_0}, \ldots, \widehat{y}_{Mt|t_0}, \right]$, where M is the number of forecasting models. Equation 2.2 can now be rewritten as:

$$E\left\{U\left(y_t|\widehat{Y}_{t|t_0}\right)\right\} = \int_0^{+\infty} U\left(y_t\right)f\left(y_t|\widehat{Y}_{t|t_0}\right)dy_t \qquad (2.3)$$

The probability distribution function $f\left(y_t|\widehat{Y}_{t|t_0}\right)$ represents the PU, hereafter denominated $f\left(y|\widehat{y}\right)$ for sake of simplicity.

Summarizing, for a decision maker to take rational decisions it is necessary first of all to define his propensity to risk (eg the decision-maker could be risk-prone when the level of damage is low or risk-averse when the level of damage is high), which afterwards can be included in the chosen utility function, and then try to minimize the expected value of this risk on the basis of a predictive density function conditional on all the information he/she can gather and in particular on the available model forecasts.

2.2 Definition Of Predictive Uncertainty

Predictive uncertainty (PU) can be defined as the probability of any future (real) value conditional upon all the knowledge and information, available up to the present, we were able to acquire through a learning inferential process.

In clarifying to hydrologists the meaning of predictive uncertainty, Krzysztofowicz (1999) points out that "*Rational decision making (for flood warning, navigation, or reservoir systems) requires that the total uncertainty about a hydrologic predictand (such as river stage, discharge, or runoff volume) be quantified in terms of a probability distribution, conditional on all available information and knowledge.*" and that "*Hydrologic knowledge is typically embodied in a deterministic catchment model*".

These statements underline two aspects usually not fully and clearly un-

derstood by hydrologists. The first is that, as previously mentioned, the
objective of forecasting is the description of the uncertainty of actual future
values of water stage, discharge, runoff volume, etc. rather than the uncer-
tainty of predictions generated by the hydrological forecasting models. The
second is that this uncertainty, generally expressed in terms of a probability
density (or probability distribution) function, is "conditional" upon the hy-
drological forecasting model prediction, which is now seen as the available,
although uncertain, knowledge of the future. In other words, the forecasting
model prediction is now a way to reduce uncertainty in the decision making
process and not the provider of deterministic (and therefore "certain") future
levels, flows, etc.

To clarify these aspects, following Todini (2008) let us introduce the con-
cept of the joint probability distribution of the real quantity of interest, the
predictand (namely the discharge, the water level at a specific cross section,
etc.) y, and the model forecast $\widehat{y}$.
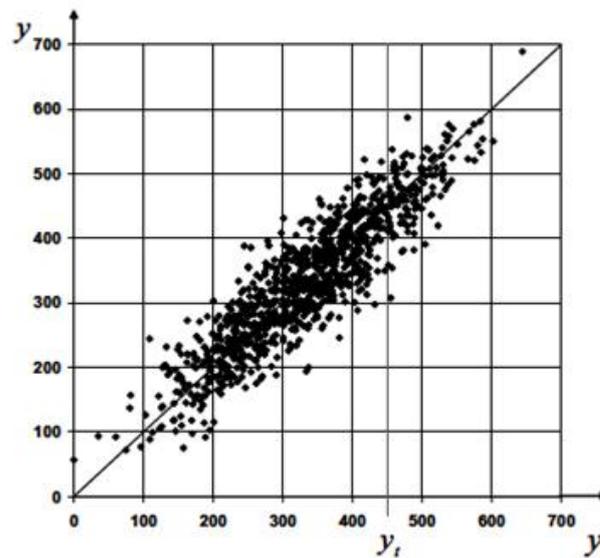
Figure 2.1: Joint $y-\widehat{y}$ sample frequency from which a joint probability density
can be estimated.The conditional density of y given $\widehat{y}$ is then obtained by
cutting the joint density for the given a value of $\widehat{y}$, namely $\widehat{y}_t$ (Liu et al.,
2005)

Unless the model is exceptionally accurate, thus perfectly matching the observations, a scatter will always be observed in the $y - \widehat{y}$ plane as in Figure 2.1. This scatter is a representation of the joint sample frequency of y and $\widehat{y}$ that can be used to estimate the joint probability density. For any given model, the model forecast $\widehat{y}_t$, will be a function of the specific value θ of the parameter set used and of the input forcing x_t (the covariate), thus the joint probability density can be expressed as in Equation 2.2:

$$f\left(y_t | \left(\widehat{y}_t | x_t, \theta\right)\right) \tag{2.4}$$

which, for the sake of clarity, is written more explicitly, than in classical statistical notation, by explicitly describing the conditionality of the model output $\widehat{y}_t$, on the covariate and the model parameters.

If there is no scatter and $y_t = \widehat{y}_t \} \forall \} t$, then and only then one can use $\widehat{y}_t$ as a forecast of y_t. In all the other cases there is an inherent uncertainty, and, in order to predict y_t, one must derive the conditional probability of y_t given $\widehat{y}_t$. This is easily done by cutting for a given $\widehat{y}_t$ the previously mentioned joint probability density (Figure 2.1) and renormalizing it. This can be formalized as:

$$f\left(y_t | \left(\widehat{y}_t | x_t, \theta\right)\right) = \frac{f\left(y_t, \left(\widehat{y}_t | x_t, \theta\right)\right)}{\int_0^\infty f\left(y_t, \left(\widehat{y}_t | x_t, \theta\right)\right) dy} \tag{2.5}$$

It is important that the reader understands that the conditional uncertainty of Equation 2.5 expresses the predictive uncertainty of a "given" model under a "given" input forcing, "given" initial and boundary conditions and a "given" set of parameter values.

This has nothing to do with the uncertainty induced by the model choice, and/or by the input and output measurement errors, and/or initial and boundary condition errors, and/or by the parameter value uncertainty. If we believe that these additional uncertainties may strongly affect the estimated value and its uncertainty, as in the case of parameter estimation, then all these uncertainties should be specifically assessed and marginalized. Note that when dealing with forecasting, it is fundamental to "marginalize" out

the uncertainties. It is not sufficient to express the uncertainty due to model, input forcing, initial and boundary conditions, parameters etc.: what is requested when forecasting in the presence of uncertainty is the expected value of the conditional density, taken on the basis of the assessed uncertainty.

For instance, following the Bayesian approach (De Finetti, 1975), all the previously mentioned additional uncertainty (namely model, measurements, initial and boundary conditions and parameters) can be concentrated in a number of "dummy" parameters, the uncertainty of which is described via a posterior probability density and successively marginalized out. Using the notation in Mantovan and Todini (2006), a formal definition of such predictive uncertainty is given as:

$$f\left(y_t | x_t, Y_{t_0}, X_{t_0}\right) = \int_\Theta f\left(y_t | x_t, \theta\right) g\left(\theta | Y_{t_0}, X_{t_0}\right) d\theta \qquad (2.6)$$

This can also be written more explicitly as:

$$f\left(y_t | \left(\widehat{y}_t | x_t, Y_{t_0}, X_{t_0}\right)\right) = \int_\Theta f\left(y_t | \left(\widehat{y}_t | x_t, \theta\right)\right) g\left(\theta | Y_{t_0}, X_{t_0}\right) d\theta \qquad (2.7)$$

where the predictand y_t is explicitly written conditionally upon the model output $\widehat{y}_t$, which is in turn conditional on the covariate and the parameters.

In Equations 2.6 and 2.7:

$f\left(y_t | x_t, Y_{t_0}, X_{t_0}\right)$ or $f\left(y_t | \left(\widehat{y}_t | x_t, Y_{t_0}, X_{t_0}\right)\right)$ is the probability density of the predictand conditional upon the historical observations and the covariate after marginalizing the uncertainty due to the parameters.

Y_{t_0} is the set of historical predictand observations (for instance water levels, discharges, etc.) and n is the record length; X_{t_0} is the set of historical covariates (for instance rainfall, upstream inflows, etc.); y_t is the predictand value of interest; x_t is the corresponding value of the covariate; θ is a given parameter vector; Θ is the ensemble of all possible parameter realizations; $f\left(y_t | x_t, \theta\right)$ or $f\left(y_t | \left(\widehat{y}_t | x_t, \theta\right)\right)$ is the probability density of the predictand value of interest conditional upon the covariate and a generic set of parameters θ; $g\left(\theta | Y_{t_0}, X_{t_0}\right)$ is the posterior density of parameters given the historical

observations.

Equation 2.7 shows that, in order to account for all the different (model, parameters and measurements) uncertainties, the predictive uncertainty must be derived by marginalizing the dummy parameters effect from the conditional probability density expressed by Equation 2.5. Therefore, the posterior probability density for the parameter vector θ, namely $g\left(\theta|Y_{t_0}, X_{t_0}\right)$, can be derived by means of a Bayesian inferential process (Mantovan and Todini, 2006) and plays an important role in the derivation of the predictive uncertainty. The Bayesian inference process allows this posterior probability density to be derived from the historical observations Y_{t_0}, X_{t_0}, starting from a prior density expressing our subjective knowledge about the parameters. The posterior density is used to marginalize out the conditionality on the parameters; this involves integrating over Θ, the entire domain of existence of the parameters, its product with the probability density function $f\left(y_t|\left(\widehat{y}_t|x_t, \theta\right)\right)$ of the predictand y_t conditional on the covariate x_t and the parameter vector θ that identifies the specific model. Equation 2.7 is then the basis for deriving the predictive probability density for the tth observation, and can be used to describe the predictive uncertainty both in "hindcast" mode, when $t \leq t_0$, and in "forecast" mode for $t > t_0$.

In "forecast" mode there is another issue to be discussed. If one uses a predicted input $\widehat{x}_t$ instead of a measured one, as for instance when using the meteorological quantitative precipitation forecasts as the forcing of a hydrological model, instead of the observed precipitation (which is obviously not available at a future time), both $f\left(y_t|\left(\widehat{y}_t|x_t, \theta\right)\right)$, and $g\left(\theta|Y_{t_0}, X_{t_0}\right)$, are no more valid. The first one becomes $f\left(y_t|\left(\widehat{y}_t|\widehat{x}_t, \theta\right)\right)$ and the second one $g\left(\theta|Y_{t_0}, \widehat{X}_{t_0}\right)$, which means that both the density of the predictand conditional on the parameters as well as the posterior parameter densities must be re-derived using the predicted values of the co-variate instead of the measured ones. There are alternative ways to obtain the parameter posterior density, but the use of the predicted co-variate values is unavoidable.

2.3   The Probabilistic Threshold Paradigm

Today, similarly to what was done for more than a century, in order to trigger their decisions, the majority of water authorities involved in flood emergency management prepare their plans on the basis of pre-determined water depths or thresholds ranging from the warning water level to the flooding level. Decisions, and consequent actions, are then taken as soon as a real time measure of the water stage overtops one of these thresholds. This approach, which is correct and sound in the absence of flood forecasting models is a way of anticipating events on the basis of water level measures (in the cross sections of interest or in upstream cross sections), but can only be effective on very large rivers where the time lag between the overtopping of the warning and the flooding levels is sufficiently large to allow for the implementation of the planned flood relief strategies and interventions (Todini and Coccia, 2010).

Given that all the water stage measures are affected by relatively small errors (1-2 cm), they can be, and have been, considered as deterministic; therefore in the sequel this approach will be referred to as the deterministic threshold paradigm.

Unfortunately, the advent and the operational use of real time flood forecasting models, has not changed this paradigm, which has been the cause of several unsatisfactory results. Today, the flood managers, instead of comparing the actual measurements to the different threshold levels, they compare the forecasts, namely the hydrologic or hydraulic models' outputs, which is obviously done in order to further anticipate decisions by taking advantage from the prediction time horizon. Unfortunately, by doing so the forecasts are implicitly assumed to be real and deterministic, which is not the case, given that the forecasts, by their nature are virtual reality and are affected by prediction errors, which magnitude is by far larger than that of the measurement errors.

More recently, the concept of predictive uncertainty has radically changed the deterministic thre-shold paradigm (Todini and Coccia, 2010). This in-

herent uncertain nature of forecasts, as opposed to the higher accuracy of measurements, requires the definition of a probabilistic threshold paradigm, defined in terms of the probability of flooding taken at different probability levels (20%, 50%, etc.) instead of the definition of deterministic threshold values. The probabilistic thresholds, as opposed to the deterministic water level thresholds, can result into improved tools in the hands of decision makers. As it will be shown in the sequel, using the probabilistic thresholds, the same predicted water level may have different meaning and different effects on decisions owing to the reliability of prediction. In other words the same forecast may or may not trigger the decision of issuing a warning or evacuating an area, conditionally to its assessed level of uncertainty. More uncertain forecasts need necessarily to be treated more cautiously than more reliable ones; uncertain lower water stage forecasts could then trigger a protective measure, whereas higher, albeit more accurate water stage forecasts, would not.

Particular attention must be given to the probability of exceeding an alert threshold (for example the dike level or the corresponding discharge), which for simplicity will be called alert level. Namely, the knowledge of the predictive uncertainty allows a probability alert threshold to be estimated instead of the commonly used deterministic alert level. As mentioned above, model forecast is a representation of the reality, but not the reality itself. Hence, the comparison between the deterministic model prediction and the actual alert level can be considered an incorrect operational approach, since one compares to the real threshold a virtual quantity such as the forecast instead of real quantity that will occur in the future. A more correct way to proceed would be to account for the probability of exceeding the alert level conditional to the knowledge of the model(s) forecast(s) in terms of a probabilistic threshold value, which must reflect the emergency manager's safety concept. With the probabilistic threshold concept the reliability of the different models can also be taken in account because it is the spread of the density that characterises the uncertainty, not the expected value. As can be

seen from the Figure 2.2, for the same expected value (the horizontal dashed line) a better forecast (Model A), characterised by a narrower predictive density, will show a smaller probability of exceedance of the flooding level when compared to a worse one (Model B).
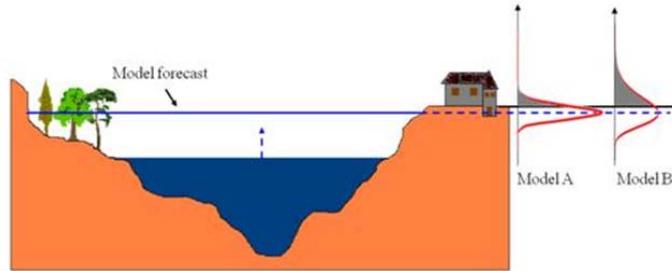
Figure 2.2: Probability of exceeding the dyke level for the same expected value, forecasted by models with different reliability.

This property can be looked at from an alternative perspective, giving scope to the definition of a probabilistic forecast paradigm. As shown in Figure 2.3 the same probability of flooding (exceedance of the flooding threshold level) corresponds to a lower expected value if the spread of PU is larger (Model B) than when it is narrower (Model A). This implies that if a probabilistic threshold is defined (for instance 15% probability of flooding) instead of a deterministic threshold level, when the PU is larger the decision maker must be more cautious and would be advised to issue an alert even when, looking at the expected value of the forecast, he would not think of issuing it, because he may regard it as being too low.

Nonetheless, the pre-requisite to implement the new probabilistic threshold paradigm is an accurate and effective estimate of predictive uncertainty. In the following sections the introduction of a new probabilistic thresholds paradigm and how this is conditioned upon a reliable estimate of predictive uncertainty will be discussed. The present work also aims at showing how the probabilistic threshold paradigm may lead to a dynamic application of the principle of precaution as a function of the degree of predictive uncertainty
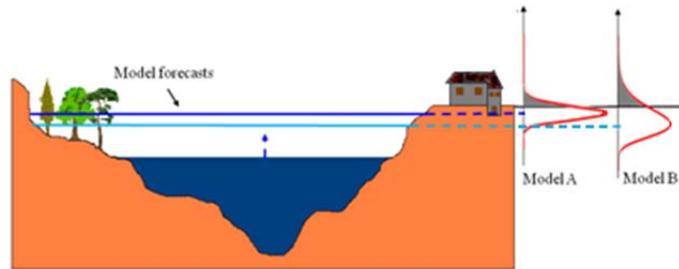
Figure 2.3: Comparison between the expected value provided by models with different reliability when the probability of exceeding the dyke level is the same for all the models.

with consequent benefits both in terms of increased reliability and robustness of decisions.

2.4 The need for probabilistic forecasts within a time horizon

As described above, the concepts of Predictive Uncertainty and Probabilistic Thresholds allow probabilistic sort of knowledge to be included in a flood warning response system, taking especially advantage of the future event probability distribution. In particular, in order to take a decision optimally, the system needs quantitative information about the predictive uncertainty, but also about the maximum river stage exceedance probability within a time interval (Krzysztofowicz, 2008).
A flood forecasting system usually provides forecasts within a time horizon which depends on catchment features and on the influence of precipitation uncertainty over the forecasted flow. From a probabilistic point of view, these forecasts provide important information concerning the exceeding probability of a prefixed river stage within the forecast lead time. For instance, the knowledge of forecasts until 48 hours in advance allows not only the predictive uncertainty to be assessed for each forecasted time step, but also it

allows the probability to exceed the dyke level and how it varies during the forecast time to be estimated within the entire time interval.

Following Rougier (2007) and Todini (2009), a simple question may help to clarify this concept:

Which is the probability that the river dykes will be overtopped in the next 24 hours?

This seems to be a well-posed question, and certainly a topical one. In fact, it is the kind of question a flood emergency manager might ask the technical staff. The Predictive Uncertainty assessment, as explained in the previous sections, aims at answering this kind of questions. When decisions about issuing or not a flood alarm must be taken, many variables have to be taken into account. Therefore, the best operational procedure should be applied on the basis of the knowledge they are able to provide. The Predictive Uncertainty assessment deal with the first problem, that is to combine all the available information in a probability value of the occurrence of an event, which means defining the *hazard*. The second question concerns the fact that an operational decision must be taken according to this estimated probability value. In this process the hazard must be converted into *risk* and other variables must be taken in account: the *vulnerability* of the area and its value (in human lives and economic terms). In the risk assessment, these variables are multiplied by the *hazard*, so it is evident that a small flooding probability may be enough to issue an alarm in highly vulnerable areas with high economic value. Hence, it is also clear the reason why in this framework, an accurate flooding probability estimation is very important.

Coming back to the previously mentioned question, there is an important aspect that ought to be highlighted. The question refers to the probability within a specific horizon time, not to the probability at a specific moment. It must be bore in mind that there is a significant difference between the two formulations. The latter does not account for what will happen during the whole forecast time, but only for the specific chose instant time, so it is

conditioned only to that specific model forecast. Otherwise, the former takes into account the whole forecast, hence, the PU must be conditioned to all the available forecasts, that is the forecasts at each step within the whole forecast lead time.

Most of the uncertainty processors available in literature can not answer the previous question, but they only can answer to the following one:

Which is the probability that the river dykes will be overtopped exactly at the hour 24th?

It is evident that if this probability is computed for each time forecast step it is possible to identify a variation of the probability of occurrence during the entire forecast time, but each probability is conditioned only to the correspondent prediction and the forecast is not considered on its whole. The problem is so converted from an univariate form to a multivariate one and the Predictive Uncertainty takes the form of a multivariate distribution of the actual future values of the predictand at each forecast time step. If the integration of this multivariate distribution is done for each variable below the value of the prefixed predictand threshold, it is possible to identify the probability that this threshold will never be exceeded during the forecast time, so its complementary to 1 is the probability to observe almost one exceeding occurrence within the forecast time horizon.

Additionally, another question concerning the probability of the exceeding time can be answered. Once the exceeding probability within the time interval is identified, a decision maker can also be interested in knowing when this exceeding event will occur. Even if this is not a crucial information in a decision making process, it is closely related to the previous concepts. Namely, if the probabilities of occurrence are computed within every time step, they identify a discrete cumulative exceeding probability within the entire time horizon, from which it is easy to identify the probability of the time exceeding as its discrete derivative (Krzysztofowicz, 2008).

2.5   Existing Approaches For PU Assessment

The prediction problem can be tackled with two different approaches, depending on the nature of the decision problem to be solved. The first approach relates to continuous processes, which require the estimation of the entire predictive probability function: for instance when dealing with flood damages, which vary with the water level reached, the expected value of these damages can only be estimated if the full probability density of water levels is available. There are other cases where only the integral above or below a threshold of the predictive density is required. This is the case for instance when one has to decide, based on model forecasts, whether the flood event will exceed an alert threshold or not; these cases can be described in discrete probability terms.

2.5.1   Continuous Probability Problem

Hydrological Uncertainty Processor (HUP)

Krzysztofowicz (1999) introduced a Bayesian processor, the Hydrological Uncertainty Processor (HUP) which aims at estimating the predictive uncertainty given a set of historical observations and a hydrological model prediction. The HUP was developed around the idea of converting both observations and model predictions into a Normal space by means of the NQT in order to derive the joint distribution and the predictive conditional distribution from a treatable multivariate distribution. In practice, as described in Krzysztofowicz (1999), after converting the observations and the model forecasts available for the historical period into the Normal space, the HUP combines the prior predictive uncertainty (in this case derived using an autoregressive model) with a Likelihood function in order to obtain the posterior density of the predictand conditional to the model forecasts. From the Normal space this conditional density is finally re-converted into the real space in order to provide the predictive probability density.

The introduction of HUP generated a positive impact into the hydrological

community, because it was the first time that predicting uncertainty was correctly formulated and used in hydrological forecasting. Nonetheless, HUP has three major limitations. The first one relates to the fact that only one model at a time can be used in HUP, which is hardly extendable to multi model forecasts. Moreover the used prior autoregressive (AR) model frequently tends to be inadequate to represent the predictand, as for instance in the case of a flood routing problem where the AR model is adequate for representing the recession but not the rising limb of the flood wave. Finally, the HUP procedure implies the independence of the AR model errors from those deriving from the used prediction model, which is not guaranteed due to the fact that both models tend to be highly correlated to the observations, which inevitably induces a level of correlation among them.

Bayesian Model Averaging (BMA)

Introduced by Raftery (1993), Bayesian Model Averaging (BMA) has gained a certain popularity in the latest years. The scope of Bayesian Model Averaging is correctly formulated in that it aims at assessing the mean and variance of any future value of the predictand conditional upon several model forecasts. Differently from the HUP assumptions, in BMA all the models (including the AR prior model) are similarly considered as alternative models. Raftery et al. (2005) developed the approach on the assumption that the predictand as well as the model forecasts were approximately Normally distributed, while Vrugt and Robinson (2007) relaxed this hypothesis and showed how to apply the BMA to Log-normal and Gamma distributed variables. In practice the Bayesian Inference problem, namely the need for estimating a posterior density for the parameters, is overcome in the BMA by estimating a number of weights via a constrained optimization problem. Once the weights have been estimated, BMA allows to estimate the mean and the variance of the predictand conditional upon several models at the same time.

The original BMA, as introduced by Raftery (1993), has shown several prob-

lems. First of all, as pointed out by Vrugt and Robinson (2007), the original assumption of approximately Normally distributed errors, is not appropriate for representing highly skewed quantities such as water discharges or water levels in rivers. Therefore one must either relax this hypothesis, as done by Vrugt and Robinson (2007) who applied the BMA to Log-normal and Gamma distributed variables or to convert the original in the Normal space once again using the NQT, as done in Todini (2008). Another problem, which emerges from the application of BMA is the use of the "expectation-maximization" (EM) algorithm (Dempster et al., 1977) proposed by Raftery et al. (2005), which was not found to properly converge to the maximum of the likelihood. To overcome this problem, one can either use sophisticated, complex optimization tools such as the SCEM-UA (Vrugt et al., 2003) or, as proposed by Todini (2008), a simple and original constrained Newton-Raphson approach, which converges in a very limited number of iterations.

The Error Heteroscedasticity Problem: Quantile Regression (QR)

The latest uncertainty processors (UP) approaches tackle the problem of the heteroscedasticity of the errors often present in hydrological modelling. All the previously described techniques imply homoscedasticity of the error variance, which is assumed to be independent from the magnitude of the observed or forecasted values. In real cases this assumption leads to a lack of accuracy, especially at reproducing high flows, because the NQT tends to increase the variance of the lower values. Moreover, the number of observed and computed low and medium flows is much larger than that of high flows with the consequence of a higher weight in the determination of the regression or the correlation coefficients used by the different approaches. As a consequence the estimation of high flows in the Normal Space will be affected by a distortion in the mean as well as an overestimation of the variance, which will inevitably increase when returning into the Real Space.

Recently, in order to overcome this problem, the Quantile Regression (Koenker, 2005) was used (Weerts et al., 2010).

The Quantile Regression (QR) approach tries to represent the error heteroscedasticity identifying a linear variation of the quantiles of the PU as a function of the model forecast magnitude. This technique allows all the desired quantiles of the PU to be assessed in the Normal Space and then reconverted by means of the Inverse NQT to the Real Space. The τth sample quantile is computed solving the Equation 2.8, from which is possible to identify the parameters a_τ and b_τ which defines the linear regression for the τth quantile.

$$\min_{a_\tau, b_\tau \in R} \sum_{i=1}^{n} \rho_\tau \left(\eta - a_\tau - b_\tau \cdot \widehat{\eta} \right) \tag{2.8}$$

where $\rho_\tau \left(x \right) = \begin{cases} x \cdot (\tau - 1) & if \ x < 0 \\ x \cdot \tau & if \ x \geq 0 \end{cases}$

The problem is correctly formulated and allows each quantile of the PU to be computed, but it requires the estimation of at least two parameters per quantile (in the linear case) and the number of parameters to be estimated may become quite large. Moreover, QR not always improves from assuming homoscedasticity: this depends on the actual distribution of the errors. Figure 2.4 and Figure 2.5 show two situations in which the use of QR leads to very different results.
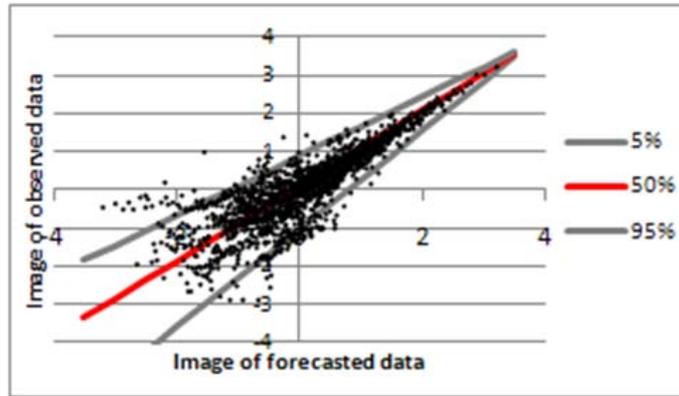
Figure 2.4: An optimal situation for using the QR.

Figure 2.4 is an optimal situation for using QR because the variation of error variance is linearly decreasing with the magnitude of the forecasts and the resulting quantiles well represent the real distribution of the data. On the contrary, in Figure 2.5 it is not possible to identify a linear variation of the error variance and the use QR does not provide improved assessments of PU, particularly for high forecast values.
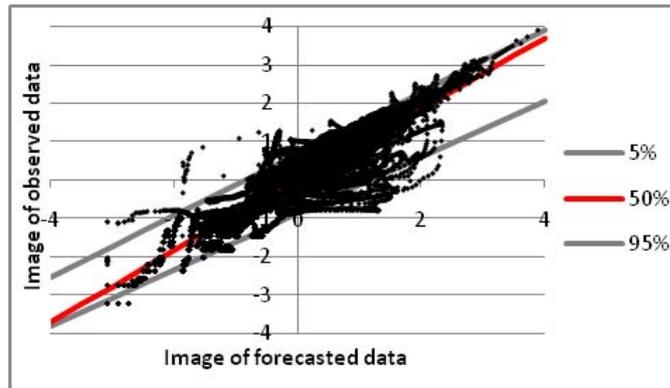
Figure 2.5: Poor results are obtained using QR in the situation represented here, which, by the way, is quite common in hydrological applications.

2.5.2   Discrete Probability Problem

As mentioned above, the prediction problem can be also faced adopting a discrete formulation, especially when only the integral above or below a threshold of the predictive density is needed. In practical problems, such as the flood warning management, the decision making process is based on the assessment of the probability to exceed the dyke level, that is the probability of a flooding occurrence. As described in Section 2.3, the continuous processors allow this probability to be computed, but also discrete probability processors (DPPs) can be used to estimate it. The main advantage given by these processors is their simple structure that can considerably reduce the computational time. Moreover, the DPPs can be coupled with continuous

processors to provide more information concerning the predicted event and the effects that it could cause.

When dealing with discrete probability problems, the predictive problem is generally simpler when both the predictand and the predictors are binary functions such as rain/no-rain or quantities above/below a threshold. Unfortunately, several problems, generally referred to as binary response, have binary predictands but continuous predictors. In this case the problem can be quite complex due to the need of converting the continuous functions into binary ones.

Let us consider a binary response variable, the predictand (y), taking values of 1 or 0, and a single explanatory variable, the predictor $(\widehat{y})$. The most commonly used statistical models for this type of data are the generalized linear models:

$$g\left(\pi_i\right) = \beta_0 + \beta_1 \cdot \widehat{y}_i \qquad (2.9)$$

where $\pi_i = P\left\{y_i = 1\right\}$ is the probability of positive response, namely y_i taking the value 1 when the $\widehat{y}$ value is $\widehat{y}_i$, while g is the link function (McCullagh and Nelder, 1989; Nelder and Wedderburn, 1972). In the present work, this kind of problem has been faced in the present work using two different link functions: the logistic one and the integral of the beta functions. The solution of the problem is obtained identifying the function parameters that maximize the probability of success, that is the probability to predict the level exceeding when it actually occurs and to do not predict it otherwise. If the alert level is called T, the probability of success is defined as:

$$P_s = \prod_{i=1}^{n} \pi_i{}^{r_i} \cdot \left(1 - \pi_i\right)^{1-r_i} \qquad (2.10)$$

where r_i is the observed value at step i, which can assume values $r_i = 1$ if $y \geq T$ and $r_i = 0$ if $y < T$. The solution of Equation 2.10 cannot usually be analytically obtained and a numerical algorithm is required. In literature different algorithms have bee proposed, such as a weighted iterative least

squares (Miller, 1992) or a Newton-Raphson based approach proposed by Todini et al. (2009); in the present work the SCE-UA algorithm (Duan et al., 1992) has been used.

Based on the same idea, another kind of function to determine the probability response has been introduced. In this case the parameter to be identified is a threshold in the virtual space of the predictors, a *virtual threshold*, on which the probability response depends; this threshold must be computed as the one that maximizes the probability of success. This kind of processor has been called *Beta Distribution Processor* because it is based on a feature of the beta distribution that will be explained below.

Finally, the last DPP presented is the Bayesian Univariate Binary Processor (Todini et al., 2009). This processor is based on the Bayes theorem and does not require a structural link model, it only requires the estimation of a *virtual threshold* in the space of the predictor.
The DPPs have been tested on a preliminary study case on the Po river in Italy, the obtained results highlighted what was expected. The DPPs require low computational times and their predictive ability is good, even if none of them could reach the performance of the continuous processors. In fact, only the continuous processors can identify the complete predictive distribution. Additionally, the continuous processors are based on a more informative process, which allows the probability to exceed a river stage to be more accurately estimated and the false and missed alarm rates obtained with the DPPs are almost always greater than the ones produced by the continuous processors. Nevertheless, the DPPs can provide a further support in operational alarm warning management and can be coupled to the continuous processors without a huge computational extra-effort. For this reason they have been analyzed and described in the following sections.

LOGIT Processor

When the link function has a logistic form, the probability of positive response is defined as:

$$\pi = \frac{e^{\beta_0 + \beta_1 \cdot \widehat{y}}}{1 + e^{\beta_0 + \beta_1 \cdot \widehat{y}}} \tag{2.11}$$

The historical available data are used in the calibration phase to estimate the values of $\beta_0{}^*$ and $\beta_1{}^*$ that maximize Equation 2.10. Afterward, these values are used in Equation 2.11 to validate the model.

The probability to exceed the level T is then computed by Equation 2.12:

$$P\left(y \geq T | \widehat{y_i}\right) = \frac{e^{\beta_0^* + \beta_1^* \cdot \widehat{y_i}}}{1 + e^{\beta_0^* + \beta_1^* \cdot \widehat{y_i}}} \tag{2.12}$$

The binary response is finally obtained verifying if $P\left(y \geq T | \widehat{y}\right)$ is greater or lower than 0.5, which indicates if the probability to exceed the level T is greater or lower than the one of not exceeding it. In the first case the response takes value 1 and in the second case 0.

$$\begin{cases} y_i > T & \forall P\left(y > T | \widehat{y_i}\right) > \frac{1}{2} \\ y_i \leq T & \forall P\left(y > T | \widehat{y_i}\right) \leq \frac{1}{2} \end{cases} \tag{2.13}$$

Beta Function Processor

If the link function is the integral of a beta function with parameters r and s, the probability function is defined as:

$$\pi_i = \int_0^{P(\widehat{y_i})} Be\left(p; r, s\right) dp = \int_0^{P(\widehat{y_i})} \frac{\Gamma\left(r + s\right)}{\Gamma\left(r\right) \cdot \Gamma\left(s\right)} \cdot p^{r-1} \cdot (1 - p)^{s-1} dp \tag{2.14}$$

Differently from the logistic function, Equation 2.14 does not relate directly the predictor value $\widehat{y_i}$ with the exceeding probability, but it is functional on the probability of occurrence of $\widehat{y_i}$. In other words, in Equation 2.14 $P\left(\widehat{y_i}\right)$ identifies the probability of occurrence of $\widehat{y_i}$, computed with the Weibull plotting position on the calibration data.

As done for the LOGIT model, during the calibration phase, in Equation 2.14 the optimal parameters r^* and s^* are obtained maximizing the success

probability of Equation 2.10.

Finally, in order to validate the model the probability $P\left(\widehat{y}_i\right) = \frac{i}{N+1}$ is assigned to the model predicted value $\widehat{y}_i$, where i is the position assumed by the prediction into the ordered calibration vector and N is the total number of calibration data. Then, the probability of occurence is used to obtain the probability to exceed the level T through Equation 2.14.

$$P\left(y \geq T | \widehat{y}_i\right) = \int_0^{P(\widehat{y}_i)} Be\left(p; r^*, s^*\right) dp \tag{2.15}$$

The binary response is computed according to the scheme of Equation 2.13.

Beta Distribution Processor

As mentioned above, the Beta Distribution Processor takes advantage of a Beta distribution according to which the probability distribution of the $i-esim$ element in an ordered vector composed of N elements is represented by a Beta distribution with parameters i and $N - i + 1$:

$$Be\left(i, N - i + 1\right) = \frac{N!}{(i-1)!\,(N-i)!} \cdot p^{i-1} \cdot (1-p)^{N-i} \tag{2.16}$$

If an unknown *virtual threshold* $\widehat{T}$ is defined in the domain $]0,1[$ and a model prediction assuming position i into the ordered calibration vector, $\widehat{y}_i$, is considered, it is possible to identify the probability to exceed the virtual threshold conditioned to the model predicted value as:

$$\pi_i = \int_{\widehat{T}}^{\infty} Be\left(i, N - i + 1\right) dp \tag{2.17}$$

The virtual threshold $\widehat{T}$ is the only model parameter, whose optimal value $\widehat{T}^*$ is identified by maximizing the probability of success of Equation 2.10, as done for the previous binary models. Hence, the probability to exceed the level T conditioned to the model prediction $\widehat{y}_i$ is computed as:

$$P\left(y > T|\widehat{y}_i\right) = \int_{\widehat{T}^*}^{\infty} Be\left(i, N - i + 1\right) dp \tag{2.18}$$

As done for the previous binary models, the binary response is obtained verifying if $P\left(y \geq T|\widehat{y}_i\right)$ is greater or lower than 0.5:

$$\begin{cases} y_i > T & \forall P\left(y > T|\widehat{y}_i\right) > \frac{1}{2} \\ y_i \leq T & \forall P\left(y > T|\widehat{y}_i\right) \leq \frac{1}{2} \end{cases} \tag{2.19}$$

Bayesian Univariate Binary Predictor (BUBP)

The use of a virtual threshold is an important feature also of the BUBP and concerning this threshold some considerations can be done. The need for a discrete binary response model lies in the fact that on the one hand the predictand is a binary quantity while, in general, the predictors are represented by continuous variables within a certain range. This is why a first step in the proposed model is required to convert the conditioning variables, the predictors, into binary quantities (below a threshold = 0, above a threshold =1). It is also necessary to clarify that the thresholds to which one must compare the predictors to generate a binary variable are not necessarily the same threshold used for the predictand. Usually the predictand is a *real* quantity which is compared to a specific *real* threshold: for instance the rain/no-rain event or a water level in a river which is above/below a warning level or the dyke height. On the contrary, the predictors must be considered as *virtual reality* representations. This is so not only when dealing with the output of a model but also when the predictor is an error corrupted direct or indirect measure of the predictand (Todini et al., 2009). With this in mind, we can easily understand that the thresholds we must compare the predictors to, are not the *real threshold* but rather *virtual thresholds* in the virtual space of the predictors. In other words if for instance we want to estimate the probability of a real water level being above a warning level conditional to our modeled water level being above a threshold, this threshold will not necessarily be the same real warning level, but rather a specific virtual warning level coher-

ent with the virtual world in which our model operates. Maybe this virtual reality concept be can better understood when the predictand is the actual water level in a river and the predictors are some estimated rainfall over the catchment and the soil moisture storage. In this case, one must estimate a virtual rainfall threshold and a virtual soil moisture threshold values in order to maximise their probability when observing the real water stage above the real warning level.

The proposed BUBP approach is valid for the case of a single predictor, anyway the extension of the approach to multiple predictors is quite easy to do, but it will not be described in this document.

Knowing the real threshold T, which is given as part of the problem, and one a priori unknown virtual threshold $\widehat{T}$, which must be estimated from the observations, the joint probability of the observation y and the prediction $\widehat{y}$ can be matched to the joint probability mass function of Figure 2.6, where, the binary variables r (real) and v (virtual) are defined as follows:

Figure 2.6: The four components of the joint probability mass function.

The four components of the joint probability mass function can be easily computed from observations conditionally to the knowledge of the threshold value $\widehat{T}$:

$$\begin{cases} P_{r=0,v=0} = P_{y \leq T, \hat{y} \leq \hat{T}} = \frac{n_{00}}{n} \\ P_{r=0,v=1} = P_{y \leq T, \hat{y} > \hat{T}} = \frac{n_{01}}{n} \\ P_{r=1,v=0} = P_{y > T, \hat{y} \leq \hat{T}} = \frac{n_{10}}{n} \\ P_{r=1,v=1} = P_{y > T, \hat{y} > \hat{T}} = \frac{n_{11}}{n} \end{cases} \tag{2.20}$$

where n is the total number of observations, and as in Figure 2.7, n_{00} is the number of observations for which $r = 0$ and $v = 0$; n_{01} is the number of observations for which $r = 0$ and $v = 1$; n_{10} is the number of observations for which $r = 1$ and $v = 0$; n_{11} is the number of observations for which $r = 1$ is the number of observations for which $r = 1$ and $v = 1$.

Figure 2.7: The simple estimator of the joint probability mass function.

Similarly one can compute the marginal probabilities:

$$\begin{cases} P_{r=0} = P_{y \leq T} = \frac{n_{00} + n_{01}}{n} = \frac{n_0}{n} \\ P_{r=1} = P_{y > T} = \frac{n_{10} + n_{11}}{n} = \frac{n_1}{n} \\ P_{v=0} = P_{\hat{y} \leq \hat{T}} = \frac{n_{00} + n_{10}}{n} \\ P_{v=1} = P_{\hat{y} > \hat{T}} = \frac{n_{01} + n_{11}}{n} \end{cases} \tag{2.21}$$

As opposed to the LOGIT and Beta Function and as described for the Beta Distribution, this representation does not require a link model. The only parameter to be estimated is the virtual threshold $\hat{T}$. The calibration, namely the estimation of this parameter, can be successfully achieved by maximising the Likelihood of successes and at the same time minimizing the

Likelihood of failures. These Likelihoods can be easily defined on the basis of the probabilities of the predictor conditional upon the observations, that can be derived by means of the Bayes theorem:

$$
\begin{cases}
P_{v=0|r=0} = P_{\widehat{y}\le\widehat{T}|y\le T} = \dfrac{P_{y\le T,\widehat{y}\le\widehat{T}}}{P_{y\le T,\widehat{y}\le\widehat{T}}+P_{y\le T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{00}}{n}}{\frac{n_{00}}{n}+\frac{n_{01}}{n}} = \dfrac{n_{00}}{n_{00}+n_{01}} \\[2ex]
P_{v=1|r=0} = P_{\widehat{y}>\widehat{T}|y\le T} = \dfrac{P_{y\le T,\widehat{y}>\widehat{T}}}{P_{y\le T,\widehat{y}\le\widehat{T}}+P_{y\le T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{01}}{n}}{\frac{n_{00}}{n}+\frac{n_{01}}{n}} = \dfrac{n_{01}}{n_{00}+n_{01}} \\[2ex]
P_{v=0|r=1} = P_{\widehat{y}\le\widehat{T}|y> T} = \dfrac{P_{y> T,\widehat{y}\le\widehat{T}}}{P_{y> T,\widehat{y}\le\widehat{T}}+P_{y> T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{10}}{n}}{\frac{n_{10}}{n}+\frac{n_{11}}{n}} = \dfrac{n_{10}}{n_{10}+n_{11}} \\[2ex]
P_{v=1|r=1} = P_{\widehat{y}>\widehat{T}|y> T} = \dfrac{P_{y> T,\widehat{y}>\widehat{T}}}{P_{y> T,\widehat{y}\le\widehat{T}}+P_{y> T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{11}}{n}}{\frac{n_{10}}{n}+\frac{n_{11}}{n}} = \dfrac{n_{11}}{n_{10}+n_{11}}
\end{cases}
\tag{2.22}
$$

When dealing with only one predictor the problem is easily solved by searching, in only one dimension, the optimal threshold value which maximises the following Likelihood function, which expresses the probability of successes given the threshold and the observations:

$$
\begin{aligned}
L_s\left(\widehat{T}\right) &= P_{v=0|r=0} \cdot P_{v=1|r=1} = P_{\widehat{y}\le\widehat{T}|y\le T} \cdot P_{\widehat{y}>\widehat{T}|y>T} = \\
&= \frac{P_{y\le T,\widehat{y}\le\widehat{T}}}{P_{y\le T}} \cdot \frac{P_{y>T,\widehat{y}>\widehat{T}}}{P_{y>T}} = \frac{n_{00}\left(\widehat{T}\right)\cdot n_{11}\left(\widehat{T}\right)}{n_0 \cdot n_1} \propto \\
&\propto n_{00}\left(\widehat{T}\right)\cdot n_{11}\left(\widehat{T}\right)
\end{aligned}
\tag{2.23}
$$

while, at the same time, minimises the Likelihood function which expresses the probability of failures given the threshold and the observations:

$$
\begin{aligned}
L_f\left(\widehat{T}\right) &= P_{v=1|r=0} \cdot P_{v=0|r=1} = P_{\widehat{y}>\widehat{T}|y\le T} \cdot P_{\widehat{y}\le\widehat{T}|y>T} = \\
&= \frac{P_{y\le T,\widehat{y}>\widehat{T}}}{P_{y\le T}} \cdot \frac{P_{y>T,\widehat{y}\le\widehat{T}}}{P_{y>T}} = \frac{n_{01}\left(\widehat{T}\right)\cdot n_{10}\left(\widehat{T}\right)}{n_0 \cdot n_1} \propto \\
&\propto n_{01}\left(\widehat{T}\right)\cdot n_{10}\left(\widehat{T}\right)
\end{aligned}
\tag{2.24}
$$

this can be formulated as follows:

$$Max_{\widehat{T}}\left\{L_s\left(\widehat{T}\right) - L_f\left(\widehat{T}\right)\right\} = n_{00}\left(\widehat{T}\right) \cdot n_{11}\left(\widehat{T}\right) - n_{01}\left(\widehat{T}\right) \cdot n_{10}\left(\widehat{T}\right) \quad (2.25)$$

The search of the threshold value can be in the positive direction if there is a positive correlation between reality and virtual reality, namely when both predictand and predictor generally grow or decrease at the same time, or it can be in the negative direction if the correlation is negative. A simple solution to this problem is to compute the two thresholds in the opposed directions and then select the one that produces the largest value of the objective function. Once the threshold value $\widehat{T}$ is found, it is now easy to perform a prediction. The predictive scheme is now:

$$\begin{cases} P_{y\leq T|\widehat{y}} = \begin{cases} P_{y\leq T|\widehat{y}\leq\widehat{T}} & \forall \widehat{y} \leq \widehat{T} \\ P_{y\leq T|\widehat{y}>\widehat{T}} & \forall \widehat{y} > \widehat{T} \end{cases} \\ P_{y>T|\widehat{y}} = \begin{cases} P_{y>T|\widehat{y}\leq\widehat{T}} & \forall \widehat{y} \leq \widehat{T} \\ P_{y>T|\widehat{y}>\widehat{T}} & \forall \widehat{y} > \widehat{T} \end{cases} \end{cases} \quad (2.26)$$

In this case, as opposed to what was done to calibrate the threshold (and generally what is also done to validate the model in terms of POD or FAR) the conditionality is no more on the observations, which being in a in predictive mode, are now assumed to be unknown. The conditional probabilities to be used are then the probabilities of the real event conditional upon the occurrence of the virtual one.

The probabilities appearing in 2.25 can now be derived using the Bayes theorem as follows:

$$\begin{cases} P_{r=0|v=0} = P_{y\leq T|\widehat{y}\leq\widehat{T}} = \dfrac{P_{y\leq T,\widehat{y}\leq\widehat{T}}}{P_{y\leq T,\widehat{y}\leq\widehat{T}}+P_{y>T,\widehat{y}\leq\widehat{T}}} = \dfrac{\frac{n_{00}}{n}}{\frac{n_{00}}{n}+\frac{n_{10}}{n}} = \dfrac{n_{00}}{n_{00}+n_{10}} \\ P_{r=0|v=1} = P_{y\leq T|\widehat{y}>\widehat{T}} = \dfrac{P_{y\leq T,\widehat{y}>\widehat{T}}}{P_{y\leq T,\widehat{y}>\widehat{T}}+P_{y>T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{01}}{n}}{\frac{n_{01}}{n}+\frac{n_{11}}{n}} = \dfrac{n_{01}}{n_{01}+n_{11}} \\ P_{r=1|v=0} = P_{y>T|\widehat{y}\leq\widehat{T}} = \dfrac{P_{y>T,\widehat{y}\leq\widehat{T}}}{P_{y\leq T,\widehat{y}\leq\widehat{T}}+P_{y>T,\widehat{y}\leq\widehat{T}}} = \dfrac{\frac{n_{10}}{n}}{\frac{n_{00}}{n}+\frac{n_{10}}{n}} = \dfrac{n_{10}}{n_{00}+n_{10}} \\ P_{r=1|v=1} = P_{y>T|\widehat{y}>\widehat{T}} = \dfrac{P_{y>T,\widehat{y}>\widehat{T}}}{P_{y\leq T,\widehat{y}>\widehat{T}}+P_{y>T,\widehat{y}>\widehat{T}}} = \dfrac{\frac{n_{11}}{n}}{\frac{n_{01}}{n}+\frac{n_{11}}{n}} = \dfrac{n_{11}}{n_{01}+n_{11}} \end{cases} \quad (2.27)$$

Using the predictive probability $P_y > T|\widehat{y}$ (only one of the two is needed since $P_{y \leq T|\widehat{y}} = 1 - P_{y > T|\widehat{y}}$) one can decide whether $y_i > T$ or $y_i \leq T$ according to:

$$\begin{cases} y_i > T & \forall P_{y > T|\widehat{y}} > \frac{1}{2} \\ y_i \leq T & \forall P_{y > T|\widehat{y}} \leq \frac{1}{2} \end{cases} \tag{2.28}$$

Chapter 3

Methodological Approach

3.1 The Model Conditional Processor (MCP)

3.1.1 Basic Ideas

The Model Conditional Processor (MCP) is a Bayesian methodology, proposed by Todini (2008), for estimating the predictive uncertainty. The derivation of the predictive distribution is essentially based on the estimation of a joint predictand-prediction distribution, computed by taking advantage of the model behaviour knowledge acquired through the available historical series. Since the multivariate distributions can be formulated and effectively analytically treated in a very limited number of cases, Krzysztofowicz (1999) suggested transforming the observations and model forecasts in a Gaussian or Normal space via a non parametric transformation known as the Normal Quantile Transform (NQT) (Van der Waerden, 1952, 1953a,b). The NQT allows the observation y and the model forecast $\widehat{y}$ to be converted into a Normal space using the quantiles associated to the order statistics, computed by means of the Weibull plotting position.

The original variables y and $\widehat{y}$ are so converted to their transformed values η and $\widehat{\eta}$ respectively, which are distributed with a Normal Standard Distribution, and the probability of each element is the same as its original

corresponding value. So the relation between the original variables and their transformed values is:

$$P\left(y < y_i\right) = \frac{i}{n+1} = P(\eta < \eta_i), for\}i = 1..n,$$

where n is the number of the historical available data and i the plotting position order.

In the Normal space the joint distribution of η and $\widehat{\eta}$ can be assumed as a Normal Bivariate, $f\left(\eta, \widehat{\eta}\right)$, with mean and variance:

$$\mu_{\eta,\widehat{\eta}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.1}$$

$$\Sigma_{\eta,\widehat{\eta}} = \begin{bmatrix} 1 & \sigma_{\eta\widehat{\eta}} \\ \sigma_{\eta\widehat{\eta}} & 1 \end{bmatrix} \tag{3.2}$$

Moreover, the covariance between η and $\widehat{\eta}$, due to the Normal Standard distribution of the two variables, is equal to the correlation coefficient $\rho_{\eta\widehat{\eta}}$. Hence, the Equation 3.2 can be written as the cross correlation matrix:

$$\Sigma_{\eta,\widehat{\eta}} = \begin{bmatrix} 1 & \rho_{\eta\widehat{\eta}} \\ \rho_{\eta\widehat{\eta}} & 1 \end{bmatrix} \tag{3.3}$$

Through the knowledge of the joint and marginal distributions it is easy to compute the predictive distribution according to the Bayes theorem. In fact, the predictive uncertainty, defined as the distribution of the predictand conditioned on the model forecast, can be obtained by calculating the ratio between the joint distribution and the forecast marginal distribution:

$$f\left(\eta|\widehat{\eta}\right) = \frac{f\left(\eta,\widehat{\eta}\right)}{f\left(\widehat{\eta}\right)} =$$

$$= \frac{\left[2\pi \left| \begin{array}{cc} 1 & \rho_{\eta\widehat{\eta}} \\ \rho_{\eta\widehat{\eta}} & 1 \end{array} \right| \right]^{-\frac{1}{2}} exp\left(-\frac{1}{2}\left[\begin{array}{cc} \eta & \widehat{\eta} \end{array}\right]\left[\begin{array}{cc} 1 & \rho_{\eta\widehat{\eta}} \\ \rho_{\eta\widehat{\eta}} & 1 \end{array}\right]^{-1}\left[\begin{array}{c} \eta \\ \widehat{\eta} \end{array}\right]\right)}{[2\pi]^{-\frac{1}{2}}exp\left(-\frac{1}{2}\widehat{\eta}^2\right)}$$

$$(3.4)$$

This equation leads to the definition of the predictive distribution in the Normal space as a Normal Distribution with moments:

$$\begin{aligned} \mu_{\eta|\widehat{\eta}} &= \rho_{\eta\widehat{\eta}} \cdot \widehat{\eta} \\ \sigma^2_{\eta|\widehat{\eta}} &= 1 - \rho_{\eta\widehat{\eta}}^2 \end{aligned}$$

$$(3.5)$$

Therefore, after obtaining the conditional probability in the normal space, the results have to be converted into the real world in order to compute the predictive probability $f(y|\widehat{y})$. To do so the predictive density has to be sampled in the Normal space and then the obtained quantiles have to be reconverted into the real space by a reverse process. This is due to the fact that the transformation is highly non linear, and, for instance, the mean value in the Normal space does not correspond to the mean value in the real world, in fact it corresponds to the median (50% probability) (Todini, 2009).

Finally, in order to introduce the probabilistic threshold paradigm described in Section 2.3 it is necessary to compute the probability to exceed a maximum river stage H, $P\left(y > H|\widehat{y},\widehat{y}^*\right)$. This can be done directly from the predictive uncertainty in the Normal Space. In fact, if η_H is the trasformation in the Normal Space of H, obtained by using the NQT considering that H belongs to the observed variable y, then it is possible to write the following equality:

$$P\left(y > H|\widehat{y},\widehat{y}^*\right) = P\left(\eta > \eta_H|\widehat{\eta},\widehat{\eta}^*\right)$$

$$(3.6)$$

The knowledge of the exact form of the predictive uncertainty in the Gaussian Space allows this probability to be analitically computed. It is equal to the integral of the Normal Distribution with moments of Equation 3.5 above the threshold value η_H:

$$P\left(\eta > \eta_H | \widehat{\eta}, \widehat{\eta}^*\right) = \int_{\eta_H}^{+\infty} f\left(\eta | \widehat{\eta}, \widehat{\eta}^*\right) \, \mathrm{d}\eta \qquad (3.7)$$

The Distribution Tails

In this process the use of the Weibull plotting position implies the need of using an additional model to be fitted to the tails of all the variables, namely the observations and the model forecast, in the real space, in order to adjust probability quantiles larger than $\frac{n}{n+1}$ or lower than $\frac{1}{n+1}$. To identify the best curve for fitting the distribution tails, several models have been tested, such as General Extreme Value, 3 parameters Log-Normal, Paretian, Exponential functions and others. This analysis led to two main conclusions, firstly was not possible to identify one model better than another one because the processor is very sensitive to the chosen function. In fact small variations of the probability value cause big differences in the correspondent value of the variable and the range of probability concerning the tails is usually lower than 0.001. Secondly, in order to avoid marked steps in the PU when the model tails are used, it is necessary to impose the continuity between the distribution obtained with the Weibull plotting position and the tails functions. The latter can be easily solved forcing the function to pass through the specific probability value above (or below for the lower tail) which the tail will be used. At the end of this analysis for the lower tail the following function has been chosen:

$$p\left(y\right) = p_{inf} \cdot \left[\frac{y}{y\left(p_{inf}\right)}\right]^a \qquad (3.8)$$

and for the upper tail:

$$p\left(y\right) = 1 - \left(1 - p_{sup}\right) \cdot \left[\frac{y_{max} - y}{y_{max} - y\left(p_{sup}\right)}\right]^{b} \tag{3.9}$$

where p_{inf} and p_{sup} are the limits below and above which the tails will be used; $y\left(p_{inf}\right)$ and $y\left(p_{sup}\right)$ are the values of the variable y correspondent to the probability limits; y_{max} is the maximum value for which the probability is assumed to be equal to 1 and it is assumed to be equal to twice the maximum value ever observed; a and b are the parameters to be estimated. Concerning the lower tail it is assumed that the null probability is assigned to the null value of the variable y, that is true when dealing with discharges, but not if y represents level values. In this case it is necessary to refer all the values to the bedstream level, in order that the null level is the lowest level possible. Moreover, using level values also y_{max} must be computed as the double of the maximum level observed referred to the bedstream level. The parameters a and b are estimated with the Least Squares method using all the data respectively lower or greater than $y\left(p_{inf}\right)$ and $y\left(p_{sup}\right)$. For sake of simplicity, Equations 3.8 and 3.9 are linearised using a logarithmic conversion and the Least Square problem assumes a linear form, in fact Equations 3.8 and 3.9 can be written as:

$$\ln\left[\frac{p\left(y\right)}{p_{inf}}\right] = a \cdot \ln\left[\frac{y}{y\left(p_{inf}\right)}\right] \tag{3.10}$$

$$\ln\left[\frac{p\left(y\right)}{1 - \left(1 - p_{sup}\right)}\right] = b \cdot \ln\left[\frac{y_{max} - y}{y_{max} - y\left(p_{sup}\right)}\right] \tag{3.11}$$

and the parameters a and b can be obtained as:

$$a = \frac{\sum_{i}\left\{\ln\left[\frac{p(y_i)}{p_{inf}}\right] \cdot \ln\left[\frac{y_i}{y\left(p_{inf}\right)}\right]\right\}}{\sum_{i}\left\{\ln\left[\frac{y_i}{y\left(p_{inf}\right)}\right]\right\}^{2}} \tag{3.12}$$

$$b = \frac{\sum_i \left\{ \ln \left[\frac{p(y_i)}{1-(1-p_{sup})} \right] \cdot \ln \left[\frac{y_{max}-y_i}{y_{max}-y(p_{sup})} \right] \right\}}{\sum_i \left\{ \ln \left[\frac{y_{max}-y_i}{y_{max}-y(p_{sup})} \right] \right\}^2} \qquad (3.13)$$

In addition to the Least Square method an alternative way to identify the parameters of the functions is using a numerical algorithm in order to maximize the correspondence between the expected value and the observed value. The SCE-UA algorithm (Duan et al., 1992) has been added to the MCP and the user can choose between that or the Least Square method. In the first case, the initial values assigned to the parameters (a' and b') are the ones obtained using the linear regression, as described above, and the search range varies between 0.5 and 2 times the initial value. It is necessary to note that the use of the SCE-UA algorithm has two disadvantages. Firstly the calibration phase can take long time because for each set of parameters tried by the algorithm the entire calibration series must be processed. Secondly, there is high risk of *overfitting* the calibration data.

3.1.2   Multi-Model Approach

The previously described MCP methodology has generated the idea of generalizing the procedure using a multi-Normal approach (Todini, 2008). Often, a real time forecasting system is composed by more than one model, or a chain of models, and the emergency manager has to take a decision on the basis of multiple forecasts of the same quantity that may also be very different from each other. It is very difficult to find an objective way to state that one model is better than another , or to assign a correct weight to each forecast in order to extrapolate from all the available information a stochastic forecast that allows the emergency to be managed in the best way.

In order to combine several model forecasts, the MCP can be improved by generalizing the Bivariate Normal approach to a Multivariate Normal approach (Mardia et al., 1979). In this case the Multivariate space is composed by *M+1* variables, that are the observed discharges (or water levels) y and

the M predictions $\widehat{y}_k, \ k = 1..M$. Using the NQT, all the variables are converted to their transformed values, η and $\widehat{\eta}_k, k = 1..M$, in the multi-Normal space.

All the variables in the Normal space have a Standard Normal Distribution and the predictive uncertainty, defined now as the distribution of the future event conditioned on the forecasts of the M models, can be expressed as $(y|\widehat{y}_1, \ldots, \widehat{y}_M)$, for simplicity abbreviated to $f(y|\widehat{y}_k)$ for the original variable and $f(\eta|\widehat{\eta}_k)$ in the normal space.

The joint distribution is a Multi-Normal Distribution with mean and variance:

$$\mu_{\eta,\widehat{\eta}_k} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.14}$$

$$\Sigma_{\eta,\widehat{\eta}_k} = \begin{bmatrix} 1 & \sigma_{\eta\widehat{\eta}_1} & \cdots & \sigma_{\eta\widehat{\eta}_M} \\ \sigma_{\widehat{\eta}_1\eta} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_{\eta\widehat{\eta}_{M-1}} \\ \sigma_{\widehat{\eta}_M\eta} & \cdots & \sigma_{\widehat{\eta}_{M-1}\eta} & 1 \end{bmatrix} \tag{3.15}$$

Moreover, all the covariances, due to the Normal Standard distribution of all the variables, are equal to the correlation coefficients. So Equation 3.15 can be written as the cross correlation matrix:

$$\Sigma_{\eta,\widehat{\eta}_k} = \begin{bmatrix} 1 & \rho_{\eta\widehat{\eta}_1} & \rho_{\eta\widehat{\eta}_2} & \cdots & \rho_{\eta\widehat{\eta}_M} \\ \rho_{\widehat{\eta}_1\eta} & 1 & \rho_{\widehat{\eta}_1\widehat{\eta}_2} & \ddots & \rho_{\widehat{\eta}_1\widehat{\eta}_M} \\ \rho_{\widehat{\eta}_2\eta} & \rho_{\widehat{\eta}_2\widehat{\eta}_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \rho_{\widehat{\eta}_{M-1}\widehat{\eta}_M} \\ \rho_{\widehat{\eta}_M\eta} & \rho_{\widehat{\eta}_M\widehat{\eta}_1} & \cdots & \rho_{\widehat{\eta}_M\widehat{\eta}_{M-1}} & 1 \end{bmatrix} \tag{3.16}$$

Defining:

$$
\begin{cases}
\Sigma_{\eta\eta} = 1 \\[2ex]
\Sigma_{\eta\widehat{\eta}} = \begin{bmatrix} \rho_{\eta\widehat{\eta}_1} & \rho_{\eta\widehat{\eta}_2} & \cdots & \rho_{\eta\widehat{\eta}_M} \end{bmatrix} \\[4ex]
\Sigma_{\widehat{\eta}\widehat{\eta}} = \begin{bmatrix}
1 & \rho_{\widehat{\eta}_1\widehat{\eta}_2} & \cdots & \rho_{\widehat{\eta}_1\widehat{\eta}_M} \\
\rho_{\widehat{\eta}_2\widehat{\eta}_1} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \rho_{\widehat{\eta}_{M-1}\widehat{\eta}_M} \\
\rho_{\widehat{\eta}_M\widehat{\eta}_1} & \cdots & \rho_{\widehat{\eta}_M\widehat{\eta}_{M-1}} & 1
\end{bmatrix}
\end{cases}
\tag{3.17}
$$

and substituting Equations 3.17 in Equation 3.16, the cross correlation matrix can also be written as:

$$
\Sigma_{\eta,\widehat{\eta}_k} = \begin{bmatrix} \Sigma_{\eta\eta} & \Sigma_{\eta\widehat{\eta}} \\ \Sigma_{\eta\widehat{\eta}}^T & \Sigma_{\widehat{\eta}\widehat{\eta}} \end{bmatrix}
\tag{3.18}
$$

Then the predictive uncertainty can be expressed as:

$$
f\left(\eta\,|\,\widehat{\eta}_k\right) = \frac{f\left(\eta,\widehat{\eta}_1,\ldots,\widehat{\eta}_M\right)}{f\left(\widehat{\eta}_1,\ldots,\widehat{\eta}_M\right)}
\tag{3.19}
$$

The solution of Equation 3.19 is easily obtained and leads to a Normal distribution with moments derived from Equation 3.18 as:

$$
\mu_{\eta|\widehat{\eta}_k,\widehat{\eta}_k^*} = \Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \cdot \begin{bmatrix} \widehat{\eta}_1^* \\ \vdots \\ \widehat{\eta}_M^* \end{bmatrix}
\tag{3.20}
$$

$$
\sigma^2_{\eta|\widehat{\eta}_k,\widehat{\eta}_k^*} = 1 - \Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \cdot \Sigma_{\eta\widehat{\eta}}^T
$$

Please note that Equations 3.20 do not differ from the classical Multiple Regression results.

As done for the univariate case, the predictive uncertainty in the real world, $f\left(y\,|\,\widehat{y}_k\right)$, is obtained by converting $f\left(\eta\,|\,\widehat{\eta}_k\right)$ by means of the Inverse

NQT.

3.1.3   Multi-Temporal Approach

Following the concepts described in Section 2.4, the multi-model approach can be further generalized. If T forecast time steps are provided by the M model chains of the flood forecasting system it is possible to include all them into the MCP. Hence, the Normal Multivariate Space is composed by $T \cdot (M + 1)$ variables, namely the observed values, $\eta_i, i = 1..T$, and the forecasts of the M models at each one of the T time steps, $\widehat{\eta}_{i,k}, k = 1..M$. Note that η_i contains T vectors equals but moved forward a time step. Also note that in the multi-model case, explained in the previous section, there is just one observed value, while in the multi-temporal approach there are T observed values as much as the number of considered time steps. Therefore, the predictive uncertainty in the Normal Space is represented by a normal multivariate distribution, composed by T variables, $f\left(\eta_1, .., \eta_T | \widehat{\eta}_{1,1}, ..., \widehat{\eta}_{T,M}\right)$, for simplicity abbreviated to $f\left(\eta_i | \widehat{\eta}_{i,k}\right)$.

The joint distribution of observed and forecasted variable is a Normal $T \cdot (N + 1)$-variate distribution with the following mean and variance, where the correlation coefficients are written instead of the covariances :

$$\mu_{\eta_i, \widehat{\eta}_{i,k}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.21}$$

$$\Sigma_{\eta_i, \widehat{\eta}_{i,k}} = \begin{bmatrix} 1 & \cdots & \rho_{\eta_1,\eta_T} & \rho_{\eta_1,\widehat{\eta}_{1,1}} & \cdots & \rho_{\eta_1,\widehat{\eta}_{T,M}} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \rho_{\eta_T,\eta_1} & \cdot & 1 & \rho_{\eta_T,\widehat{\eta}_{1,1}} & \cdots & \rho_{\eta_T,\widehat{\eta}_{T,M}} \\ \rho_{\widehat{\eta}_{1,1},\eta_1} & \ddots & \rho_{\widehat{\eta}_{1,1},\eta_T} & 1 & \ddots & \rho_{\widehat{\eta}_{1,1},\widehat{\eta}_{T,M}} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \rho_{\widehat{\eta}_{T,M},\eta_1} & \cdots & \rho_{\widehat{\eta}_{T,M},\eta_T} & \rho_{\widehat{\eta}_{T,M},\widehat{\eta}_{1,1}} & \cdots & 1 \end{bmatrix} \tag{3.22}$$

In this case, the components of the cross correlation matrix can be expressed

as:

$$
\left\{
\begin{aligned}
\Sigma_{\eta\eta} &=
\begin{bmatrix}
1 & \rho_{\eta_1,\eta_2} & \cdots & \rho_{\eta_1,\eta_{T-1}} & \rho_{\eta_1,\eta_T} \\
\rho_{\eta_2,\eta_1} & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \rho_{\eta_{T-1},\eta_T} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\rho_{\eta_T,\eta_1} & \rho_{\eta_T,\eta_2} & \cdots & \rho_{\eta_T,\eta_{T-1}} & 1
\end{bmatrix} \\[2em]
\Sigma_{\eta\widehat{\eta}} &=
\begin{bmatrix}
\rho_{\eta_1,\widehat{\eta}_{1,1}} & \rho_{\eta_1,\widehat{\eta}_{1,2}} & \cdots & \rho_{\eta_1,\widehat{\eta}_{T,M-1}} & \rho_{\eta_1,\widehat{\eta}_{T,M}} \\
\rho_{\eta_2,\widehat{\eta}_{1,1}} & \ddots & \ddots & \ddots & \rho_{\eta_2,\widehat{\eta}_{T,M}} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\rho_{\eta_{T-1},\widehat{\eta}_{1,1}} & \ddots & \ddots & \ddots & \rho_{\eta_{T-1},\widehat{\eta}_{T,M}} \\
\rho_{\eta_T,\widehat{\eta}_{1,1}} & \rho_{\eta_T,\widehat{\eta}_{1,2}} & \cdots & \rho_{\eta_T,\widehat{\eta}_{T,M-1}} & \rho_{\eta_T,\widehat{\eta}_{T,M}}
\end{bmatrix} \\[2em]
\Sigma_{\widehat{\eta}\widehat{\eta}} &=
\begin{bmatrix}
1 & \rho_{\widehat{\eta}_{1,1},\widehat{\eta}_{1,2}} & \cdots & \rho_{\widehat{\eta}_{1,1},\widehat{\eta}_{T,M-1}} & \rho_{\widehat{\eta}_{1,1},\widehat{\eta}_{T,M}} \\
\rho_{\widehat{\eta}_{1,2},\widehat{\eta}_{1,1}} & \ddots & \ddots & \ddots & \rho_{\widehat{\eta}_{1,2},\widehat{\eta}_{T,M}} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\rho_{\widehat{\eta}_{T,M-1},\widehat{\eta}_{1,1}} & \ddots & \ddots & \ddots & \rho_{\widehat{\eta}_{T,N-1},\widehat{\eta}_{T,M}} \\
\rho_{\widehat{\eta}_{T,M},\widehat{\eta}_{1,1}} & \rho_{\widehat{\eta}_{T,N},\widehat{\eta}_{1,2}} & \cdots & \rho_{\widehat{\eta}_{T,M},\widehat{\eta}_{T,M-1}} & 1
\end{bmatrix}
\end{aligned}
\right.
\tag{3.23}
$$

and substituting Equations 3.23 in Equation 3.22, the cross correlation matrix can be written as:

$$
\Sigma_{\eta_i,\widehat{\eta}_{i,k}} =
\begin{bmatrix}
\Sigma_{\eta\eta} & \Sigma_{\eta\widehat{\eta}} \\
\Sigma_{\eta\widehat{\eta}}^T & \Sigma_{\widehat{\eta}\widehat{\eta}}
\end{bmatrix}
\tag{3.24}
$$

Hence, the predictive uncertainty can be expressed as:

$$
f\left(\eta_i|\widehat{\eta}_{i,k}\right) = \frac{f\left(\eta_1,\ldots,\eta_T,\widehat{\eta}_{1,1},\ldots,\widehat{\eta}_{T,M}\right)}{f\left(\widehat{\eta}_{1,1},\ldots,\widehat{\eta}_{T,M}\right)}
\tag{3.25}
$$

The solution of Equation 3.25 is easily obtained and leads to a Normal dis-

tribution with moments derived from Equation 3.24 as:

$$\mu_{\eta_i|\widehat{\eta}_{i,k},\widehat{\eta}^*_{i,k}} = \Sigma_{\eta\widehat{\eta}} \cdot \Sigma^{-1}_{\widehat{\eta}\widehat{\eta}} \cdot \begin{bmatrix} \widehat{\eta}^*_{1,1} \\ \vdots \\ \widehat{\eta}^*_{T,M} \end{bmatrix} \tag{3.26}$$

$$\Sigma_{\eta_i|\widehat{\eta}_{i,k},\widehat{\eta}^*_{i,k}} = \Sigma_{\eta\eta} - \Sigma_{\eta\widehat{\eta}} \cdot \Sigma^{-1}_{\widehat{\eta}\widehat{\eta}} \cdot \Sigma_{\eta\widehat{\eta}}^T$$

Now, the predictive uncertainty describes the joint predictive distribution of the observed values for the T forecasted time steps. With respect to the previously described approach, the additional information that can be extrapolated from this distribution is the probability to exceed a maximum river stage H (e.g. the dyke level) within the horizon time $t = 1..T$ and the exceeding time probability. To compute the former it is possible to follow the idea of Equation 3.7, but considering that in this case the interest is focused on computing the probability within t time steps, $\forall t = 1..T$. This probability takes into account that the level H may be overtopped even just in one of the considered time steps. Hence, it must be computed with the following equation:

$$P\left(\eta_t > \eta_H|\widehat{\eta}_{t,k},\widehat{\eta}^*_{t,k}\right) = 1 - \int_{-\infty}^{\eta_H} \ldots \int_{-\infty}^{\eta_H} f\left(\eta_i|\widehat{\eta}_{i,k},\widehat{\eta}^*_{i,k}\right) \mathrm{d}\eta_1 \ldots \mathrm{d}\eta_t \tag{3.27}$$

where: $t = 1..T$

For the simplest case of one model and two time steps, the area computed by Equation 3.27 is shown in Figure 3.1. It represents the total probability of exceedance during just the first time step (red area), just the second time step (green area) and both time steps (grey area).

By the Equation 3.27 a sort of *cumulative probability*, function of the time, is derived and for sake of simplicity hereafter it will be called $P_H(t)$.
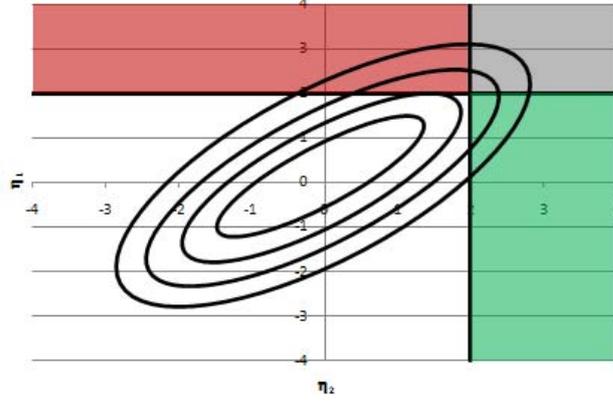
Figure 3.1: Schematic representation of Predictive Uncertainty with two fore-cast lead times in which the probability to exceed the value η_H within the total horizon time is highlighted.

Obviously it is verified that:

$$P_H(t) \leq P_H(\tau) \quad \forall \, t < \tau \tag{3.28}$$

The exceeding time probability t^* is proportional to the derivative of this *cumulative probability*. Hence, it can be obtained by marginalzing the derivative function of $P_H(t)$:

$$f(t^*) \propto \frac{\Delta P_H(t)}{\Delta t} \tag{3.29}$$

3.1.4   Truncated Normal Joint Distribution

As described in Section 2.5.1 the problem of the heteroscedasticity of the errors, often present in hydrological modeling, should be taken in account and the Quantile Regression, even if is a valid methodology, not always can well represent the real error variance. For instance, in the situation represented in Figure 3.2 a different alternative approach can be used to improve results. Namely, within the MCP framework the entire Normal domain is divided into two (or more) sub-domains where Truncated Normal Distributions (TNDs)

can be used (Coccia and Todini, 2010). In this case, the MCP can be applied assuming that the joint distribution in the Normal Space is not unique, but can be divided into two (or more) TNDs. A threshold separating low flows form high flows in the forecast domain is relatively easy to be identified. Figure 3.2 shows the two TNDs that can be used in the example.
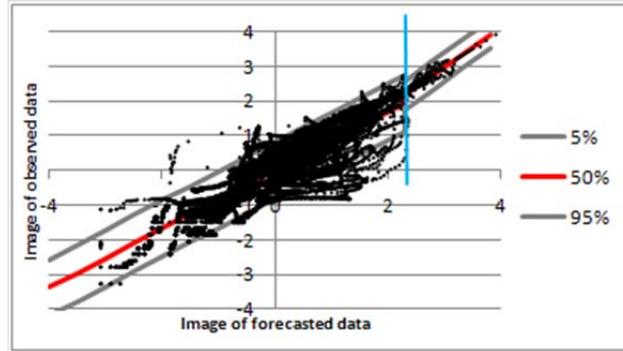
Figure 3.2: Truncated Normal Joint Distributions. The division of the Joint Distribution in the Normal Space into two Bivariate Truncated Normal Distributions is shown. The red line represents the modal value, while the grey lines represent the 5% and the 95% quantiles. The light blue line represents the threshold used in order to divide the two TNDs.

The identification of the two TNDs is not immediate, but can be obtained by the following procedure that depends on the number of available forecasting models.

TNDs With Only One Forecasting Model After converting the original variables y and $\widehat{y}$ to their transformed values η and $\widehat{\eta}$, the so obtained samples are assumed to belong to two unknown normal distributions truncated over $\widehat{\eta}$ by a threshold a. The moments of these truncated distributions can be estimated by equating them to the sampling moments.

For the sample that includes the high flows, the Truncated Normal distribution for $\widehat{\eta} > a$ is:

$$f\left(\widehat{\eta}|\widehat{\eta} > a\right) = \frac{f\left(\widehat{\eta}\right)}{\int_a^{+\infty} f\left(\widehat{\eta}\right) d\widehat{\eta}} = \frac{f\left(\widehat{\eta}\right)}{1 - F_{\widehat{\eta}}\left(a\right)} \tag{3.30}$$

whit $f(\widehat{\eta})$ defined as:

$$f(\widehat{\eta}) = \frac{1}{\sqrt{2\pi}s_{\widehat{\eta}}} exp\left\{-\frac{1}{2}\left(\frac{\widehat{\eta} - m_{\widehat{\eta}}}{s_{\widehat{\eta}}}\right)^2\right\} \qquad (3.31)$$

where $\widehat{m}$ and $\widehat{s}$ are the mean and the standard deviation of the non truncated, albeit unknown distribution.

Therefore, the joint distribution is the following Truncated Normal Bivariate distribution:

$$f(\eta, \widehat{\eta}|\widehat{\eta} > a) = \frac{f(\eta, \widehat{\eta})}{\int_{-\infty}^{+\infty}\left[\int_{a}^{+\infty} f(\eta, \widehat{\eta})\,d\widehat{\eta}\right]d\eta} = \frac{f(\eta, \widehat{\eta})}{1 - F_{\widehat{\eta}}(a)} \qquad (3.32)$$

Where $f(\eta, \widehat{\eta})$ is defined as:

$$f(\eta, \widehat{\eta}) = \frac{exp\left\{-\frac{1}{2}\left[\begin{array}{cc} \eta - m_\eta & \widehat{\eta} - m_{\widehat{\eta}} \end{array}\right] S^{-1}\left[\begin{array}{c} \eta - m_\eta \\ \widehat{\eta} - m_{\widehat{\eta}} \end{array}\right]\right\}}{\sqrt{2\pi\,|S|}} \qquad (3.33)$$

where: $S = \left[\begin{array}{cc} s_\eta{}^2 & s_{\eta\widehat{\eta}} \\ s_{\eta\widehat{\eta}} & s_{\widehat{\eta}}{}^2 \end{array}\right]$

In Equations 3.31 and 3.33, the values of $m_{\widehat{\eta}}$, $s_{\widehat{\eta}}$, m_η, s_η and $s_{\eta\widehat{\eta}}$ are unknown but can be derived from the sampling moments. Applying the Bayes theorem to the Truncated Normal, the predictive uncertainty (which in this case represents the probability distribution of η conditional on the model forecast $\widehat{\eta}^* > a$) becomes:

$$f(\eta|\widehat{\eta} > a, \widehat{\eta}^*) = \frac{f(\eta, \widehat{\eta}|\widehat{\eta} > a, \widehat{\eta}^*)}{f(\widehat{\eta}|\widehat{\eta} > a, \widehat{\eta}^*)} = \frac{f(\eta, \widehat{\eta}|\widehat{\eta}^*)}{f(\widehat{\eta}|\widehat{\eta}^*)} \qquad (3.34)$$

and it is normally distributed with mean and variance:

$$\begin{aligned} \mu_{\eta|\widehat{\eta}>a,\widehat{\eta}^*} &= m_\eta + \frac{s_{\eta\widehat{\eta}}}{s_{\widehat{\eta}}{}^2}\left(\widehat{\eta}^* - m_{\widehat{\eta}}\right) \\ \sigma^2_{\eta|\widehat{\eta}>a,\widehat{\eta}^*} &= s_\eta{}^2 - \frac{s_{\eta\widehat{\eta}}{}^2}{s_{\widehat{\eta}}{}^2} \end{aligned} \qquad (3.35)$$

Similarly for $\widehat{\eta}^* < a$, Equations 3.30 and 3.32 become respectively:

$$f\left(\widehat{\eta}|\widehat{\eta} < a, \widehat{\eta}^*\right) = \frac{f\left(\widehat{\eta}\right)}{\int_{-\infty}^{a} f\left(\widehat{\eta}\right) d\widehat{\eta}} = \frac{f\left(\widehat{\eta}\right)}{F_{\widehat{\eta}}\left(a\right)} \tag{3.36}$$

$$f\left(\eta, \widehat{\eta}|\widehat{\eta} < a, \widehat{\eta}^*\right) = \frac{f\left(\eta, \widehat{\eta}\right)}{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{a} f\left(\eta, \widehat{\eta}\right) d\widehat{\eta}\right] d\eta} = \frac{f\left(\eta, \widehat{\eta}\right)}{F_{\widehat{\eta}}\left(a\right)} \tag{3.37}$$

According to the procedure described in Appendix A, the previous equations allow the PU in the Normal Space to be defined as a Normal Distribution with mean and variance:

$$\mu_{\eta|\widehat{\eta}>a,\widehat{\eta}^*} = \mu_{\eta} + \frac{\sigma_{\eta\widehat{\eta}}}{\sigma_{\widehat{\eta}}^2}\left(\widehat{\eta}^* - \mu_{\widehat{\eta}}\right)$$
$$\sigma_{\eta|\widehat{\eta}>a,\widehat{\eta}^*}^2 = \sigma_{\eta}^2 - \frac{\sigma_{\eta\widehat{\eta}}^2}{\sigma_{\widehat{\eta}}^2} \tag{3.38}$$

for the case that the predicted value $\widehat{\eta}^*$ is greater than the threshold value a. Here μ_{η}, $\mu_{\widehat{\eta}}$ are respectively the sample means of $\eta|\widehat{\eta} > a$ and $\widehat{\eta}|\widehat{\eta} > a$ and σ_{η}, $\sigma_{\widehat{\eta}}$ are their sample standard deviations. These moments are obviously computed considering only the sample including the data belong to the upper sample.

If $\widehat{\eta}^*$ is lower than the threshold value a, the mean and variance of PU in Normal Space are:

$$\mu_{\eta|\widehat{\eta}=\widehat{\eta}^*<a} = \mu_{\eta} + \frac{\sigma_{\eta\widehat{\eta}}}{\sigma_{\widehat{\eta}}^2}\left(\widehat{\eta}^* - \mu_{\widehat{\eta}}\right)$$
$$\sigma_{\eta|\widehat{\eta}=\widehat{\eta}^*<a}^2 = \sigma_{\eta}^2 - \frac{\sigma_{\eta\widehat{\eta}}^2}{\sigma_{\widehat{\eta}}^2} \tag{3.39}$$

where μ_{η}, $\mu_{\widehat{\eta}}$, σ_{η} and $\sigma_{\widehat{\eta}}$ are computed taking in account only the data of the lower sample.

TNDs With More Than One Forecasting Model When dealing with more than one model, the procedure becomes a bit more difficult. The threshold should be identified for each model and the joint distribution would be represented by 2^M MTNDs (where M is the number of models) that include all the possible simultaneous combinations of each model overtopping or not

its respective threshold. The moments of each MTNDs should be obtained by means of the sampling moments computation, but unfortunately in real cases often the available data are not enough to identify representative samples and the MTNDs cannot be well assessed.

In order to avoid this situation the problem can be tackled with a different approach. The MCP can be applied in three phases. Firstly, each model is processed separately using the TNDs as described above. In this phase, for each model its threshold is identified. In the second phase, the series of expected values of each model simulation (previously obtained) are combined using two MTNDs indentified on the basis of the model that better represented the high flows. In other words, for each model the variances of the upper sample are computed and then they are compared each other in order to identify which model will be used in the second phase in order to split the multivariate joint distribution in two MTNDs. Finally, in the third phase the series of expected values computed in the second phase is processed using the TNDs as described above. The detailed description of the procedure is the following.

Considering M available models and applying to each model the methodology described in Section 2.4.2.1, the following parameters are computed:

a'_i = threshold used for identifying the TNDs of the model i=1,M

$\sigma^2_{\eta|\widehat{\eta}_i=\widehat{\eta}^*_i>a'_i}$ = conditioned variance of the upper TND for model i=1,M

In the second phase the joint MTNDs are identified on the basis of the model k, which is the model that better represents the high flows:

$$\sigma^2_{\eta|\widehat{\eta}_k>a'_k,\widehat{\eta}^*_k} < \sigma^2_{\eta|\widehat{\eta}_i>a'_i,\widehat{\eta}^*_i} \ \forall \ i \neq k$$

Considering the upper sample, for sake of simplicity let's define the vector a, such as:

$$\begin{cases} a_i = -\infty & \forall \ i \neq k \\ a_k = a'_k \end{cases}$$

the vector $\widehat{\eta}$ represents the variables related to the model simulations, $\widehat{\eta} = \begin{bmatrix} \widehat{\eta}_1 & \cdots & \widehat{\eta}_M \end{bmatrix}^T$.

The joint distribution of the simulated variables $\widehat{\eta}_i > -\infty \ \forall \ i \neq k$ and $\widehat{\eta}_k > a_k$ is:

$$f(\widehat{\eta}|\widehat{\eta}_k > a_k) = \frac{f(\widehat{\eta})}{1 - F_{\widehat{\eta}_k}(a_k)} \tag{3.40}$$

Where $f(\widehat{\eta})$ is defined as:

$$f(\widehat{\eta}) = \frac{exp\left\{-\frac{1}{2}[\widehat{\eta} - \widehat{m}]S_{\widehat{\eta}\widehat{\eta}}^{-1}[\widehat{\eta} - \widehat{m}]^T\right\}}{(2\pi)^{1/M}\sqrt{|S_{\widehat{\eta}\widehat{\eta}}|}} \tag{3.41}$$

where $\widehat{m} = \begin{bmatrix} m_{\widehat{\eta}_1} \\ \vdots \\ m_{\widehat{\eta}_M} \end{bmatrix}$ is the vector containing the means of the marginal distributions of $\widehat{\eta}$ and $S_{\widehat{\eta}\widehat{\eta}}$ is the covariance matrix between the variables $\widehat{\eta}$:

$$S_{\widehat{\eta}\widehat{\eta}} = \begin{bmatrix} s_{\widehat{\eta}_1}^2 & s_{\widehat{\eta}_2\widehat{\eta}_1} & \cdots & s_{\widehat{\eta}_M\widehat{\eta}_1} \\ s_{\widehat{\eta}_1\widehat{\eta}_2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & s_{\widehat{\eta}_{M-1}\widehat{\eta}_1} \\ s_{\widehat{\eta}_1\widehat{\eta}_M} & \cdots & s_{\widehat{\eta}_1\widehat{\eta}_{M-1}} & s_{\widehat{\eta}_M}^2 \end{bmatrix} \tag{3.42}$$

Therefore, the joint distribution of all the variables is the following MTND:

$$f(\eta, \widehat{\eta}|\widehat{\eta} > a) = \frac{f(\eta, \widehat{\eta})}{1 - F_{\widehat{\eta}_k}(a_k)} \tag{3.43}$$

Where $f(\eta, \widehat{\eta})$ is defined as:

$$f(\eta, \widehat{\eta}) = \frac{exp\left\{-\frac{1}{2}\begin{bmatrix} \eta - m & \widehat{\eta} - \widehat{m} \end{bmatrix}S^{-1}\begin{bmatrix} \eta - m \\ \widehat{\eta} - \widehat{m} \end{bmatrix}\right\}}{(2\pi)^{\frac{1}{M+1}} \cdot \sqrt{|S|}} \tag{3.44}$$

where:

m is the mean of the marginal distributions of η

$$S = \begin{bmatrix} S_{\eta\eta} & S_{\eta\widehat{\eta}} \\ S_{\eta\widehat{\eta}}^T & S_{\widehat{\eta}\widehat{\eta}} \end{bmatrix} \tag{3.45}$$

$$S_{\eta\eta} = \begin{bmatrix} s_\eta^2 \end{bmatrix}$$

$$S_{\eta\widehat{\eta}} = \begin{bmatrix} s_{\eta\widehat{\eta}_1} & \cdots & s_{\eta\widehat{\eta}_m} \end{bmatrix}$$

In Equations 3.41 and 3.44, $\widehat{m}$, $S_{\widehat{\eta}\widehat{\eta}}$ m, $S_{\eta\eta}$ and $S_{\eta\widehat{\eta}}$ are unknown but can be derived from the sampling moments.

Applying the Bayes theorem to the joint MTND, the predictive uncertainty, namely the probability distribution of η conditional on the realization of the model forecasts $\widehat{\eta}^*$, becomes:

$$f\left(\eta | \widehat{\eta}_k > a_k, \widehat{\eta}^*\right) = \frac{f\left(\eta, \widehat{\eta} | \widehat{\eta}_k > a_k, \widehat{\eta}^*\right)}{f\left(\widehat{\eta} | \widehat{\eta}_k > a_k, \widehat{\eta}^*\right)} = \frac{f\left(\eta, \widehat{\eta} | \widehat{\eta}^*\right)}{f\left(\widehat{\eta} | \widehat{\eta}^*\right)} \tag{3.46}$$

Please, note that Equation 3.46 is conceptually equal to Equation 3.32. In other words, being M the number of models considered, $f\left(\eta, \widehat{\eta}\right)$ is a *(M+1)-variate* and $f\left(\widehat{\eta}\right)$ is *M-variate*, and in Equation 3.32 *M=1*.

The conditional distribution of Equation 3.46 is normally distributed with mean and variance:

$$\begin{aligned} \mu_{\eta|\widehat{\eta}_k>a_k,\widehat{\eta}^*} &= m + S_{\eta\widehat{\eta}} S_{\widehat{\eta}\widehat{\eta}}^{-1} \left(\widehat{\eta}^* - \widehat{m}\right) \\ \sigma^2_{\eta|\widehat{\eta}_k>a_k,\widehat{\eta}^*} &= S_{\eta\eta} - S_{\eta\widehat{\eta}} S_{\widehat{\eta}\widehat{\eta}}^{-1} S_{\eta\widehat{\eta}}^T \end{aligned} \tag{3.47}$$

Following the procedure described in Appendix A, the previous equations lead to define PU in the Normal Space as a Normal Distribution with mean and variance:

$$\begin{aligned} \mu_{\eta|\widehat{\eta}_k>a_k,\widehat{\eta}^*} &= \mu + \Sigma_{\eta\widehat{\eta}} \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \left(\widehat{\eta}^* - \widehat{\mu}\right) \\ \sigma^2_{\eta|\widehat{\eta}_k>a_k,\widehat{\eta}^*} &= \Sigma_{\eta\eta} - \Sigma_{\eta\widehat{\eta}} \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \Sigma_{\eta\widehat{\eta}}^T \end{aligned} \tag{3.48}$$

if the predicted value of the model k, $\widehat{\eta}_k^*$, is greater than the threshold value a_k. Here μ, $\widehat{\mu}$ are respectively the sample means of $\eta | \widehat{\eta}_k > a_k$ and $\widehat{\eta} | \widehat{\eta}_k > a_k$ and $\Sigma_{\eta\eta}$, $\Sigma_{\eta\widehat{\eta}}, \}\Sigma_{\widehat{\eta}\widehat{\eta}}$ are the components of the covariance matrix of $\eta, \widehat{\eta} | \widehat{\eta}_k > a_k$.

Similarly for the sample below the threshold and taking into account that the vector a is defined as:

$$\begin{cases} a_i = +\infty & \forall \; i \neq k \\ a_k = a'_k \end{cases}$$

Equations 3.40 and 3.43 become, respectively:

$$f\left(\widehat{\eta}|\widehat{\eta} < a\right) = \frac{f\left(\widehat{\eta}\right)}{F_{\widehat{\eta}_k}\left(a_k\right)} \tag{3.49}$$

$$f\left(\eta, \widehat{\eta}|\widehat{\eta} < a\right) = \frac{f\left(\eta, \widehat{\eta}\right)}{F_{\widehat{\eta}_k}\left(a_k\right)} \tag{3.50}$$

Hence, if $\widehat{\eta}_k^*$ is lower than the threshold value a_k, the mean and variance of PU in Normal Space are:

$$\begin{aligned} \mu_{\eta|\widehat{\eta}_k < a_k, \widehat{\eta}^*} &= \mu + \Sigma_{\eta\widehat{\eta}}\Sigma_{\widehat{\eta}\widehat{\eta}}^{-1}\left(\widehat{\eta}^* - \widehat{\mu}\right) \\ \sigma^2_{\eta|\widehat{\eta}_k < a_k, \widehat{\eta}^*} &= \Sigma_{\eta\eta} - \Sigma_{\eta\widehat{\eta}}\Sigma_{\widehat{\eta}\widehat{\eta}}^{-1}\Sigma_{\eta\widehat{\eta}}^{T} \end{aligned} \tag{3.51}$$

where μ, $\widehat{\mu}$, $\Sigma_{\eta\eta}$, $\Sigma_{\eta\widehat{\eta}}$ and $\Sigma_{\widehat{\eta}\widehat{\eta}}$ are computed taking in account only the data of the lower sample.

PU derived from a Truncated Normal Joint Distribution

Considering M available forecasting models and starting from the hypothesis that the data divided over $\widehat{\eta}_k$ by the threshold a_k, belong to two Multivariate Truncated Normal Distributions (MTNDs) and considering the upper sample, the marginal distributions of $\eta|\widehat{\eta} > a$ and $\widehat{\eta}|\widehat{\eta} > a$ are respectively a Truncated Normal (TN) and a Multivariate Truncated Normal (MTN) called

$$f\left(\eta|\widehat{\eta}_k > a_k\right) = TN\left(\mu, \Sigma_{\eta\eta}\right) \tag{3.52}$$

and

$$f\left(\widehat{\eta}|\widehat{\eta}_k > a_k\right) = MTN\left(\widehat{\mu}, \Sigma_{\widehat{\eta}\widehat{\eta}}\right) \tag{3.53}$$

Their Joint Truncated Distribution is called

$$f\left(\eta, \widehat{\eta} | \widehat{\eta}_k > a_k\right) = MTN \left(\begin{bmatrix} \mu \\ \widehat{\mu} \end{bmatrix}, \begin{bmatrix} \Sigma_{\eta\eta} & \Sigma_{\eta\widehat{\eta}} \\ \Sigma_{\widehat{\eta}\eta} & \Sigma_{\widehat{\eta}\widehat{\eta}} \end{bmatrix} \right) \qquad (3.54)$$

All the parameters μ, $\Sigma_{\eta\eta}$, $\widehat{\mu}$, $\Sigma_{\widehat{\eta}\widehat{\eta}}$ and $\Sigma_{\eta\widehat{\eta}}$ are known, because they are assumed to be equal to the sample ones.

The distributions

$$f\left(\eta\right) = N\left(m, S_{\eta\eta}\right) \qquad (3.55)$$

$$f\left(\widehat{\eta}\right) = N\left(\widehat{m}, S_{\widehat{\eta}\widehat{\eta}}\right) \qquad (3.56)$$

$$f\left(\eta, \widehat{\eta}\right) = N \left(\begin{bmatrix} m \\ \widehat{m} \end{bmatrix}, \begin{bmatrix} S_{\eta\eta} & S_{\eta\widehat{\eta}} \\ S_{\widehat{\eta}\eta} & S_{\widehat{\eta}\widehat{\eta}} \end{bmatrix} \right) \qquad (3.57)$$

are the Multivariate Complete Normal Distributions (MCNDs) to which the MTNDs, respectively represented by Equations 3.52, 3.53, 3.54, are supposed to belong.

All the parameters of the MCNDs, m, S, $\widehat{m}$, $S_{\widehat{\eta}\widehat{\eta}}$ and $S_{\eta\widehat{\eta}}$, are unknown and they must be identified in order to define the conditioned distribution, that is the PU in the Normal Space conditioned to the model forecasts that transformed using the NQT gives a value for the model k greater than the threshold value a_k.

In fact, as described by Equation 3.46, the conditioned distribution is

$$f\left(\eta | \widehat{\eta}_k > a_k, \widehat{\eta}^*\right) = \frac{f\left(\eta, \widehat{\eta}\right)}{f\left(\widehat{\eta}\right)} = N\left(\mu_{\eta | \widehat{\eta}_k > a_k, \widehat{\eta}^*} . \sigma^2_{\eta | \widehat{\eta}_k > a_k, \widehat{\eta}^*}\right) \qquad (3.58)$$

Hence, the mean and variance of the conditioned distribution are (see Equations 3.47:

$$\mu_{\eta | \widehat{\eta}_k > a_k, \widehat{\eta}^*} = m + S_{\eta\widehat{\eta}} \cdot S^{-1}_{\widehat{\eta}\widehat{\eta}} \cdot \left(\widehat{\eta}^* - \widehat{m}\right) \qquad (3.59)$$

$$\sigma^2_{\eta|\widehat{\eta}_k > a_k, \widehat{\eta}^*} = S_{\eta\eta} - S_{\eta\widehat{\eta}} \cdot S_{\widehat{\eta}\widehat{\eta}}^{-1} \cdot S_{\eta\widehat{\eta}}^T \tag{3.60}$$

The parameters of the MCNDs can be derived from the following equations, provided by the Truncated Multi-Normal Distribution Theory (Tallis, 1961), which relate the moments of the MTNDs to the ones of the MCNDs.

$$m = \mu - \frac{\Sigma_{\eta\widehat{\eta}_k}}{\sqrt{\Sigma_{\widehat{\eta}_k\widehat{\eta}_k}}} \cdot \lambda\left(\alpha_k\right) \tag{3.61}$$

$$\widehat{m} = \widehat{\mu} - \frac{\Sigma_{\widehat{\eta}\widehat{\eta}_k}}{\sqrt{\Sigma_{\widehat{\eta}_k\widehat{\eta}_k}}} \cdot \lambda\left(\alpha_k\right) \tag{3.62}$$

$$S_{\eta\eta} = \Sigma_{\eta\eta} + \frac{\Sigma_{\eta\widehat{\eta}_k}^2}{\Sigma_{\widehat{\eta}_k\widehat{\eta}_k}} \cdot \delta\left(\alpha_k\right) \tag{3.63}$$

$$S_{\widehat{\eta}\widehat{\eta}} = \Sigma_{\widehat{\eta}\widehat{\eta}} + \frac{\Sigma_{\widehat{\eta}\widehat{\eta}_k} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}_k}^T}{\Sigma_{\widehat{\eta}_k\widehat{\eta}_k}} \cdot \delta\left(\alpha_k\right) \tag{3.64}$$

$$S_{\eta\widehat{\eta}} = \Sigma_{\eta\widehat{\eta}} + \frac{\Sigma_{\eta\widehat{\eta}_k} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}_k}}{\Sigma_{\widehat{\eta}_k\widehat{\eta}_k}} \cdot \delta\left(\alpha_k\right) \tag{3.65}$$

where:

$$\begin{aligned}
\alpha_k &= \frac{a_k - m_k}{\sqrt{S_{\widehat{\eta}_k\widehat{\eta}_k}}} \\
\lambda\left(\alpha_k\right) &= \frac{\phi(\alpha_k)}{1 - \Phi(\alpha_k)} \\
\delta\left(\alpha_k\right) &= \lambda\left(\alpha_k\right) \cdot \left[\lambda\left(\alpha_k\right) - \alpha_k\right]
\end{aligned} \tag{3.66}$$

and ϕ and Φ respectively represent the pdf and the cdf of the Normal Standard Distribution.

The equality between Equations 3.47 and Equations 3.48 (or between Equations 3.35 and 3.38 for the bi-dimensional case), leads to:

$$\mu_{\eta|\widehat{\eta}^* > a} = m + S_{\eta\widehat{\eta}} S_{\widehat{\eta}\widehat{\eta}}^{-1} \left(\widehat{\eta}^* - \widehat{m}\right) = \mu + \Sigma_{\eta\widehat{\eta}} \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \left(\widehat{\eta}^* - \widehat{\mu}\right) \tag{3.67}$$

$$\Sigma^2_{\eta|\widehat{\eta}^* > a} = S_{\eta\eta} - S_{\eta\widehat{\eta}} S_{\widehat{\eta}\widehat{\eta}}^{-1} S_{\eta\widehat{\eta}}^T = \Sigma_{\eta\eta} - \Sigma_{\eta\widehat{\eta}} \Sigma_{\widehat{\eta}\widehat{\eta}}^{-1} \Sigma_{\eta\widehat{\eta}}^T \tag{3.68}$$

For sake of simplicity, these equalities will be demonstrated for one available forecast model. In this case only two variables are taken in account, η e $\widehat{\eta}$, and their joint distribution is truncated over the variable $\widehat{\eta}$ by the threshold value *a*. Hence, by substituting Equations 3.61, 3.62, 3.64 and 3.65, adapted for the specific case, in Equation 3.67 the following equation is obtained:

$$\mu - \frac{\Sigma_{\eta\widehat{\eta}}}{\sqrt{\Sigma_{\widehat{\eta}\widehat{\eta}}}} \cdot \lambda\left(\alpha\right) + \frac{\Sigma_{\eta\widehat{\eta}} + \frac{\Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}}}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right)}{\Sigma_{\widehat{\eta}\widehat{\eta}} + \frac{\Sigma_{\widehat{\eta}\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right)} \cdot \left[\widehat{\eta}^* - \widehat{\mu} + \frac{\Sigma_{\widehat{\eta}\widehat{\eta}}}{\sqrt{\Sigma_{\widehat{\eta}\widehat{\eta}}}} \cdot \lambda\left(\alpha\right)\right] =$$
$$= \mu + \frac{\Sigma_{\eta\widehat{\eta}}}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \left(\widehat{\eta}^* - \widehat{\mu}\right)$$

$$(3.69)$$

Which can be rewritten as:

$$\mu - \frac{\Sigma_{\eta\widehat{\eta}}}{\sqrt{\Sigma_{\widehat{\eta}\widehat{\eta}}}} \cdot \lambda\left(\alpha\right) + \frac{\Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}} \cdot \left[1 + \delta\left(\alpha\right)\right]}{\Sigma_{\widehat{\eta}\widehat{\eta}}^2 \cdot \left[1 + \delta\left(\alpha\right)\right]} \cdot \left[\widehat{\eta}^* - \widehat{\mu} + \sqrt{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \lambda\left(\alpha\right)\right] =$$
$$= \mu + \frac{\Sigma_{\eta\widehat{\eta}}}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \left(\widehat{\eta}^* - \widehat{\mu}\right)$$

$$(3.70)$$

By developing the Equation 3.70 the following equation is obtained:

$$\mu + \frac{\sigma_{\eta\widehat{\eta}}}{\sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \left(\widehat{\eta}^* - \widehat{\mu}\right) = \mu + \frac{\sigma_{\eta\widehat{\eta}}}{\sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \left(\widehat{\eta}^* - \widehat{\mu}\right) \tag{3.71}$$

Taking in account the Equation 3.68, it can be rewritten as:

$$\Sigma_{\eta\eta} + \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right) - \frac{\left[\Sigma_{\eta\widehat{\eta}} + \frac{\Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}}}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right)\right]^2}{\Sigma_{\widehat{\eta}\widehat{\eta}} + \frac{\Sigma_{\widehat{\eta}\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right)} = \Sigma_{\eta\eta} - \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \tag{3.72}$$

By developing it the following equation is obtained:

$$\Sigma_{\eta\eta} + \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right) - \frac{\left\{\Sigma_{\eta\widehat{\eta}} \cdot \Sigma_{\widehat{\eta}\widehat{\eta}} \cdot \left[1 + \delta\left(\alpha\right)\right]\right\}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}} \cdot \left\{\Sigma_{\widehat{\eta}\widehat{\eta}}^2 \cdot \left[1 + \delta\left(\alpha\right)\right]\right\}} = \Sigma_{\eta\eta} - \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \tag{3.73}$$

Which can be rewritten as:

$$\Sigma_{\eta\eta} + \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \delta\left(\alpha\right) - \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \cdot \left[1 + \delta\left(\alpha\right)\right] = \Sigma_{\eta\eta} - \frac{\Sigma_{\eta\widehat{\eta}}^2}{\Sigma_{\widehat{\eta}\widehat{\eta}}} \tag{3.74}$$

Now the equality is obtained:

$$\sigma_{\eta\eta} - \frac{\sigma_{\eta\widehat{\eta}}^{\mathbf{2}}}{\sigma_{\widehat{\eta}\widehat{\eta}}} = \sigma_{\eta\eta} - \frac{\sigma_{\eta\widehat{\eta}}^{\mathbf{2}}}{\sigma_{\widehat{\eta}\widehat{\eta}}} \tag{3.75}$$

If considering the lower sample, only the second of the Equations 3.66 changes, while the other two expressions are still the same:

$$\begin{aligned}
\alpha_k &= \frac{a_k - m_k}{\sqrt{S_{\widehat{\eta}_k \widehat{\eta}_k}}} \\
\lambda\left(\alpha\right) &= -\frac{\phi(\alpha)}{\Phi(\alpha)} \\
\delta\left(\alpha\right) &= \lambda\left(\alpha\right) \cdot \left[\lambda\left(\alpha\right) - \alpha\right]
\end{aligned} \tag{3.76}$$

The change in the form of $\lambda\left(\alpha\right)$ does not modify the previous procedure, which remains valid also for the lower sample and leads to the same result, with the only obvious difference that the sample moments are computed on the lower sample.

Chapter 4

Applications of the MCP

4.1 Baron Fork River, Multi-Model Application

4.1.1 Study Case

The NOAA's National Weather Service, has provided a long series of observed discharge and precipitation data for the Baron Fork River, OK, USA within the frame of the DMIP 2 Project which aims at comparing distributed hydrological models. Using this data set three models were developed: two physically based hydrological models, the TOPKAPI model (Todini and Ciarapica, 2001; Liu and Todini, 2002) and TETIS model (Francés et al., 2007; Vélez et al., 2009), and an additional data driven model based on Artificial Neural Networks. The catchment has a drainage area of about 800 km^2 at the measurement station of Eldon with a mean slope around 0.25%, while some kilometres downstream Eldon the river flows into the Illinois river. The simulations provided by the three models have been processed using the MCP firstly separately and then combined each other.
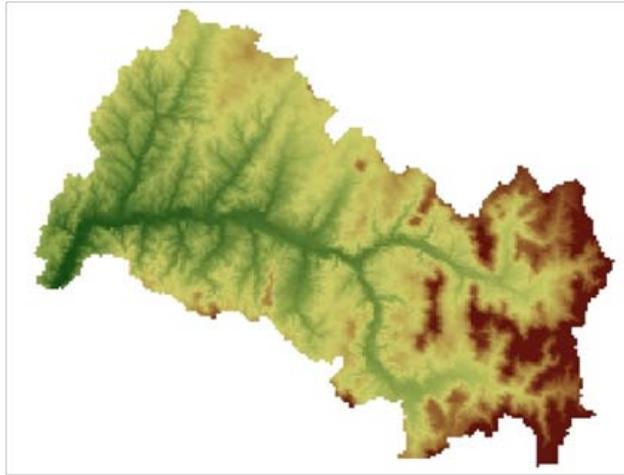
Figure 4.1: Digital Elevation Model of the Baron Fork basin at Eldon.

4.1.2 Available Data

Available meteorological data were hourly rain and temperature grids included between 10/01/1995 and 09/30/2002, with a 4 km resolution. During the same period the observed discharges in the measurement station of Eldon, OK, were available, too. Summarizing, the available data allow the basin behaviour to be simulated during a long period of about 7 years with a time step of 1 hour.

4.1.3 TOPKAPI Model Application

The TOPKAPI model has been developed at the University of Bologna (Todini and Ciarapica, 2001; Liu and Todini, 2002), it is composed of six components, which take into account the surface, sub-surface and deep flows, the routing in the channel, the snow accumulation/melt and the evapotranspiration. The application domain is divided in cells where the mass and momentum balances are solved at every time step. A more detailed model description can be found in Appendix A.

The model has been calibrated by a trial and error procedure applied to

the data included between 10/01/1996 and 09/30/2002; the year included between 10/01/1995 and 09/30/1996 has been used as 'warm up' period, allowing the model to reach a reasonable initial state.

Soil Type

The NOOA's National Weather Service provided the maps of 11 soil layers, from the surface to a depth of approximately 2 meters. The combination of these maps allowed identifying areas with the same soil type succession and for each area the parameter mean values required by the TOPKAPI model have been extrapolated. The mean initial parameter values have been computed with a weighted average on the basis of the percentage of each layer on the entire depth. The total depth of each soil type was identified with respect to the bedrock level. The resulting map is shown in Figure 4.2 while Table 4.1 describes the calibrated parameter values. During the calibration phase, especially the hydraulic conductivities and the soil depths have been changed, in particular the former ones have been increased of approximately one decimal order and the latter ones have been generally decreased.
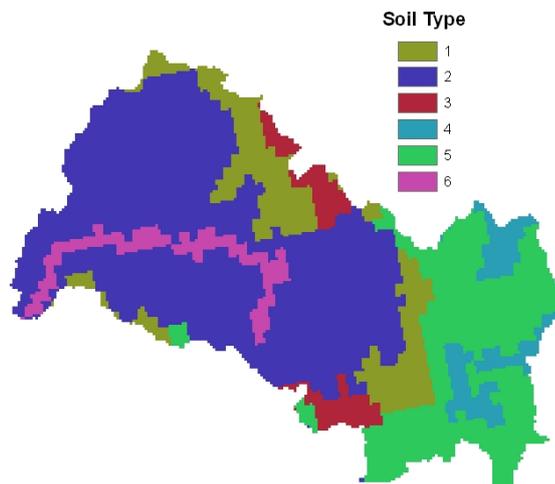
Figure 4.2: Soil Type map of the Baron Fork basin at Eldon.

Code	$K_{sh}[ms^{-1}]$	θ_r	θ_s	$Depth[m]$	α_s	$K_{sv}[ms^{-1}]$	α_s
1	$4.22 \cdot 10^{-4}$	0.458	0.084	0.60	1.41	$4.65 \cdot 10^{-8}$	19.83
2	$4.56 \cdot 10^{-4}$	0.464	0.083	0.70	1.45	$8.12 \cdot 10^{-8}$	19.48
3	$2.46 \cdot 10^{-4}$	0.460	0.080	0.70	1.43	$5.55 \cdot 10^{-8}$	20.02
4	$3.14 \cdot 10^{-4}$	0.441	0.072	0.40	1.49	$8.53 \cdot 10^{-8}$	17.36
5	$8.86 \cdot 10^{-4}$	0.458	0.081	0.42	1.45	$4.01 \cdot 10^{-8}$	19.69
6	$3.25 \cdot 10^{-4}$	0.444	0.060	0.70	1.69	$7.33 \cdot 10^{-8}$	13.78

Table 4.1: Calibrated values for the soil type parameters. Saturated horizontal hydraulic conductivity (K_{sh}), residual water content (θ_r), saturated water content (θ_s), depth, the exponent of the horizontal flow law (α_s), saturated vertical conductivity (K_{sv}) and the exponent of the percolation law (α_s).

Land Use

The TOPKAPI model requires two parameters concerning the land use, which are the Surface Manning Coefficient (n) and the Crop Factors (K) for each month.
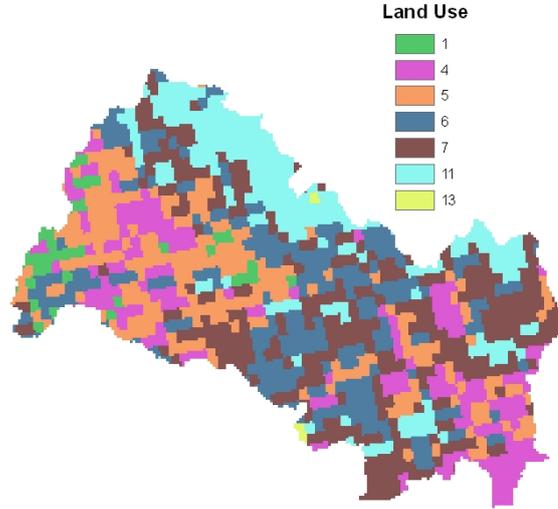
Figure 4.3: Land Use map of the Baron Fork basin at Eldon.

The Manning coefficient is used to solve the surface component, where the momentum equation is approximated by the Manning's formula. The crop factors are necessary to compute the reference evapo-transpiration accord-

ing to Thornthwaite and Mather (1955), as described in Appendix A. The map of land use (Figure 4.3) has been derived from the data provided by the Corine Land Cover Project. Table 4.2 shows the calibrated parameters.

ID	1	4	5	6	7	11	14
n	0.22	0.28	0.28	0.16	0.10	0.10	0.10
Jan	0.90	0.60	0.75	0.75	0.80	0.70	0.20
Feb	0.90	0.70	1.00	1.00	1.05	1.00	0.20
Mar	0.90	0.95	0.90	0.90	1.10	1.10	0.20
Apr	0.90	1.05	0.95	0.95	1.10	1.20	0.20
May	0.90	1.05	1.20	1.20	1.10	1.35	0.20
Jun	0.90	0.80	0.85	0.85	0.80	1.20	0.20
Jul	0.90	0.80	0.85	0.85	0.90	1.10	0.20
Aug	0.90	0.80	0.85	0.85	0.80	0.90	0.20
Sep	0.90	0.80	0.85	0.85	0.80	1.30	0.20
Oct	0.90	1.20	1.05	1.05	1.00	1.25	0.20
Nov	0.90	1.10	1.00	1.00	1.00	1.20	0.20
Dec	0.90	0.60	0.75	0.75	0.90	0.75	0.20

Table 4.2: Calibrated values for the land use parameters: Surface Manning coefficient (n) and Crop Factors for each month.

Channel Network

The channel network is derived according to the Strahler Orders. TOPKAPI automatically computes them and draws the network (Figure 4.4). The parameters to solve the channel component are assigned for each Strahler order (Table 4.3). A triangular cross section has been adopted for each channel class, hence the the slope of the river banks (α) is the only required parameter to approximately describe the geometry. Moreover, the routing model requires the definition of the Manning coefficient (n) for the channel roughness.

ID	1	2	3	4	5	6	7
n	0.1	0.09	0.08	0.075	0.07	0.065	0.06
α	1.5	2.2	2.7	3.2	3.7	4.5	6

Table 4.3: Calibrated values for the channel parameters. Channel Manning coefficient (n) and the tangent of the river banks angle (α).
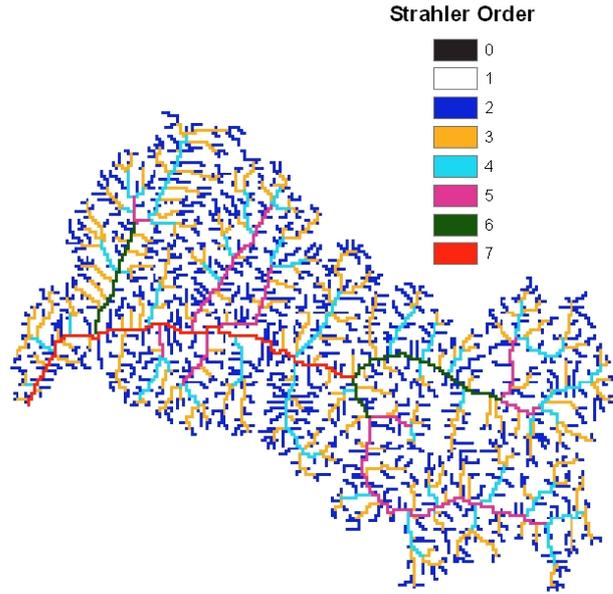
Figure 4.4: Channel network of the Baron Fork basin at Eldon.

Results

The TOPKAPI model has been calibrated by a trial and error procedure looking at all the available data and watching out for both high and low flows. Nevertheless, obtaining the best parameter set was not the main objective of this application. In fact, the main aim of the application was to test the MCP, rather than the TOPKAPI model. Although the parameter set obtained in the present study are not the best ones, the result obtained are adequate to allow discussion of the merits of the MCP.

Briefly, analyzing the TOPKAPI model simulation, the main flood peaks are generally underestimated, especially when the discharge is around 1000

$m^3 s^{-1}$, as visible in Figures 4.6 and 4.7. Figure 4.7 represents the main event occurred during the available period, the peak flow is underestimated of approximately 450 $m^3 s^{-1}$ over a maximum observed value of approximately 1550 $m^3 s^{-1}$. Moreover, the recession curves are usually not steep enough and the low flow is generally underestimated. This behavior lead to think that the soil emptying is too slow and, when a low soil moisture percentage is reached, the subsurface flow is too small. The same behavior is shown in other events (Figures 4.5 and 4.6) and it is due to the presence of only one soil response in the TOPKAPI soil conceptualization; last modifications in the model presented by Coccia et al. (2010) solved this problem. Nevertheless, the time peak is almost always correctly reproduced and the other minor events are well simulated also in terms of peak value, as shown in Figures 4.5 and 4.8.
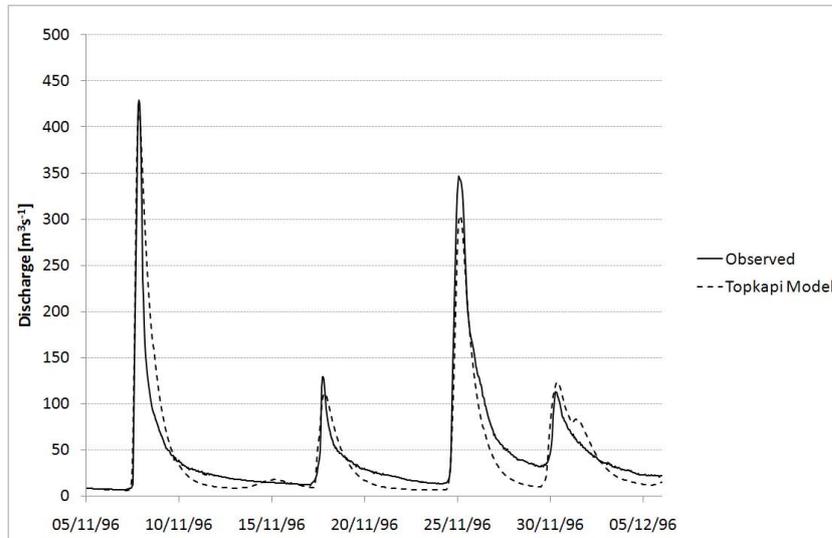
Figure 4.5: TOPKAPI simulation of events occurred in November 1996. Observed discharges (black line); TOPKAPI simulation (dashed line).

Finally, in order to evaluate the model calibration some evaluation indexes have been computed on the entire simulated period (Table 4.4).

The overall evaluation indexes in Table 4.4 show that the calibration
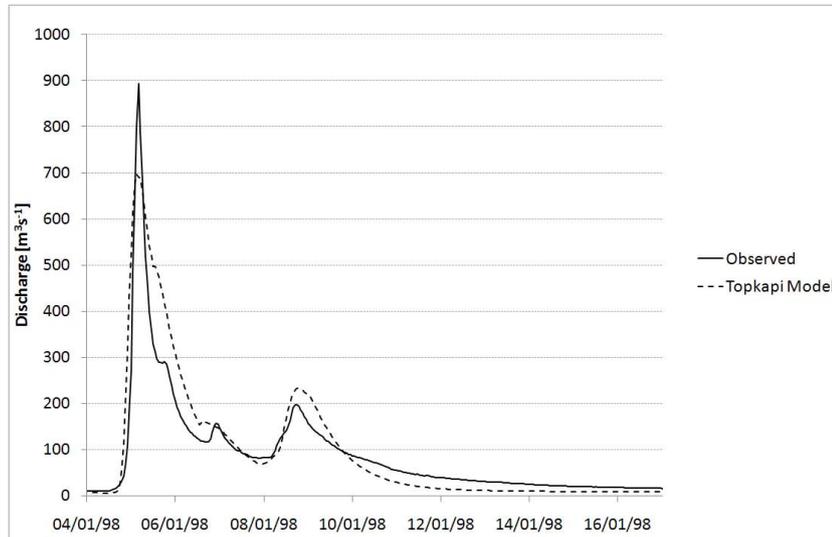
Figure 4.6: TOPKAPI simulation of events occurred in April 1998. Observed discharges (black line); TOPKAPI simulation (dashed line).
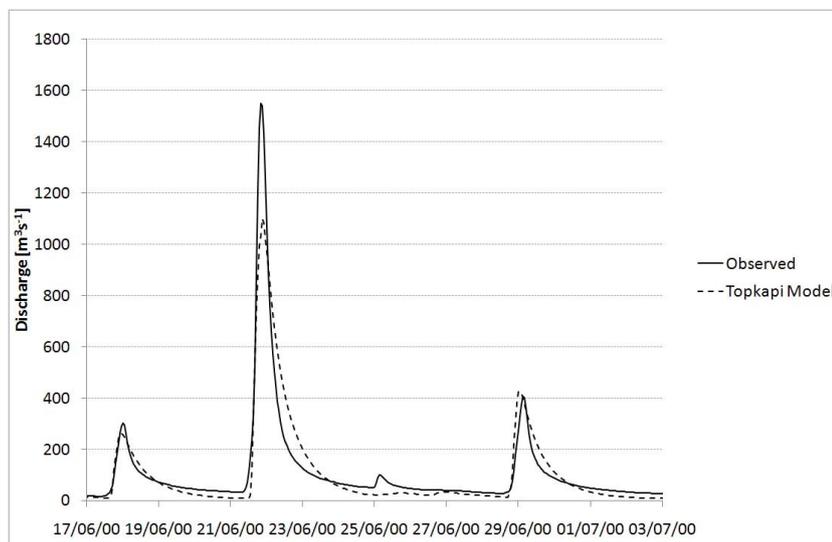
Figure 4.7: TOPKAPI simulation of the main event occurred in June 2000. Observed discharges (black line); TOPKAPI simulation (dashed line).
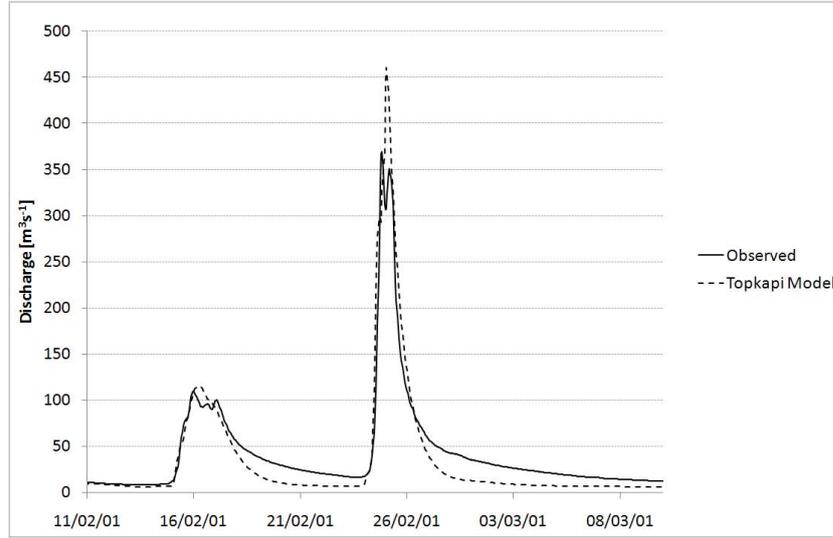
Figure 4.8: TOPKAPI simulation of events occurred in November 2001. Observed discharges (black line); TOPKAPI simulation (dashed line).

Max Q Obs	Max Q TPK	% Bias	RMSE	R^2	E
1549 m^3s^{-1}	1101 m^3s^{-1}	-1.38 %	13.29 m^3s^{-1}	0.91	0.82

Table 4.4: Overall evaluation indexes for the TOPKAPI simulation. Maximum observed and simulated discharges; Percent Bias; Root Mean Square Error; Correlation Coefficient (R^2); Nash-Sutcliffe efficiency (E).

results are good, especially the percent bias is very low, approximately -1.4% that corresponds to a bias of about $0.14\ m^3 s^{-1}$, and the correlation coefficient is high, approximately 0.91. The Nash-Sutcliffe coefficient is quite high, 0.82, but lower than the common model efficiency, which points out the possibility to improve the calibration even if it has not be done for the reasons described above.

4.1.4 TETIS Model application

In the TETIS model, developed by the Polytechnic University of Valencia (Francés et al., 2007; Vélez et al., 2009), the conceptual scheme, at each cell, consists of a series of 5 connected tanks, each one of them representing different water storages in the soil column. The vertical connections between tanks describe the precipitation, evapotranspiration, infiltration and percolation processes, whereas, the horizontal flows represent the main hydrological processes as: snowmelt, overland runoff, interflow and base flow. The routing along the channel network couples its geomorphologic characteristics with the kinematic wave approach. The TETIS model automatic calibration procedure was applied to the hydrological year included between October 2001 and September 2002. As done for the TOPKAPI model, the first year of data has been used as 'warm up' period and with the remaining data the model has been validated. The configuration of the TETIS model for the Baron Fork river has been implemented by J. Camilo Munera from the Polytechnic University of Valencia and he provided all the following maps, tables, images and comments.

Before analyzing the main model parameters it is necessary to highlight that the automatic calibration does not modify the initial parameter values (described in the following sections), but it modifies 9 correction factors by which the parameters are multiplied when are used in the different model components, as described in (Vélez et al., 2009). The calibrated correction factor values are shown in Table 4.5.

Maximum Static Storage

The Maximum Static Storage is the most important parameter in TETIS. This parameter represents the maximum water quantity that can be stored in the static tank, which accounts for initial abstractions and the capillary water storage in the upper part of the soil. Further information about this parameter can be found in Francés et al. (2007) and Vélez et al. (2009). In Figure 4.9 the Maximum Storage Capacity map is depicted.

Correction Factor	Involved Process	Calibrated Value
R-1	Static Tank	0.38130
R-2	Evapotranspiration	0.72315
R-3	Infiltration	1.25944
R-4	Overland Flow	2.00000
R-5	Percolation	0.33292
R-6	Interflow	30.0015
R-7	Groundwater outflow	0.00000
R-8	Base flow	114.4292
R-9	Channel Routing	0.20089

Table 4.5: Calibrated corrector factors for the TETIS model.
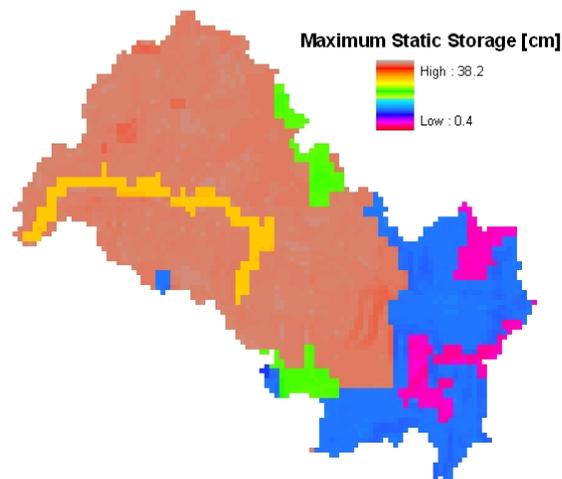
Figure 4.9: Maximum Static Storage map of the Baron Fork basin at Eldon.

Infiltration and Percolation Capacities

The infiltration and percolation capacities represent the hydraulic conductivities respectively of the surface and gravitational tanks. These parameters are fundamental in assessing the overland and subsurface flows and the vertical flows among the tanks. The initial parameter values, which have been estimated on the basis of the soil texture maps provided by the NOAA's National Weather Service, are shown in Figures 4.10 and 4.11.
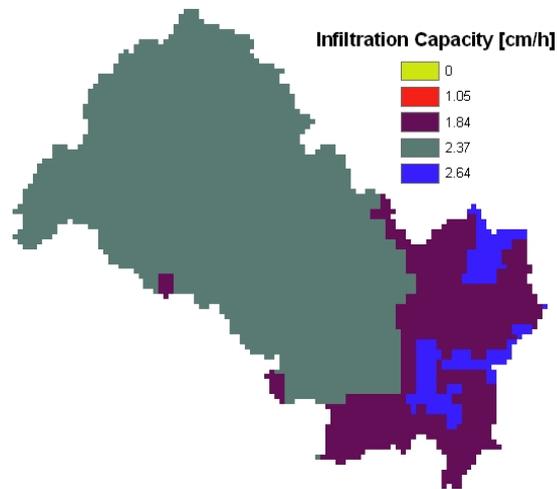
Figure 4.10: Infiltration Capacity map of the Baron Fork basin at Eldon.
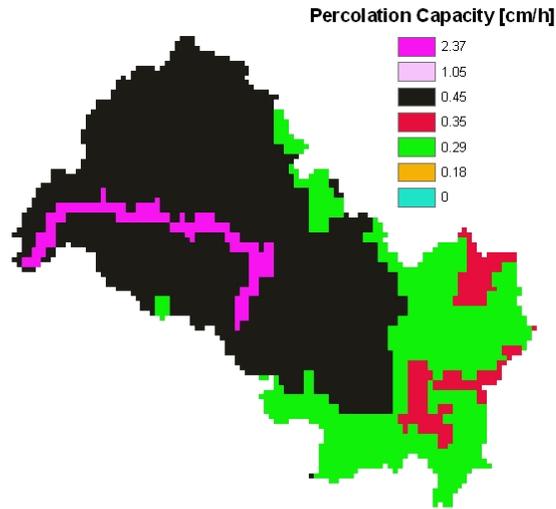
Figure 4.11: Percolation Capacity map of the Baron Fork basin at Eldon.

Results

The period chosen to calibrate the model is the year included between October 2001 and September 2002 that is the last available year. The selection was done considering a compromise between the computation time and the calibration efficiency. Model calibration attempts considering longer time-series if data have been done, however despite the fact that using 2 or 3 years of data the evaluation indexes for the calibration period increase and the computational time becomes very high. Hence, the gain in the simulation of the entire period is not enough to justify it and the use of too much data for the automatic calibration can lead to the overfitting problem.

In Figure 4.12 the observed data are compared with the TETIS simulation during the calibration year. Two questions about the results of the calibration are worth to be mentioned, firstly that the model can well reproduce the observed data, especially with particular accuracy for the recession curves and the base flow. The presence of three soil responses in the model conceptualization allows the subsurface and base flows to be more realistic than those produced by the TOPKAPI model, which has just one soil response;

this concurs with the considerations done above concerning the need to identify two different soil layers in the TOPKAPI model. The second question concerns the underestimation of the two main peaks; the TOPKAPI model showed the same behavior that lead to think to some kind of error in the observed data, maybe in the rating curve or in the precipitations. The good model performance in the calibration period is also outlined by the evaluation indexes shown in Table 4.6; the value of the percent bias (-13.6%) reflects the peaks underestimation.
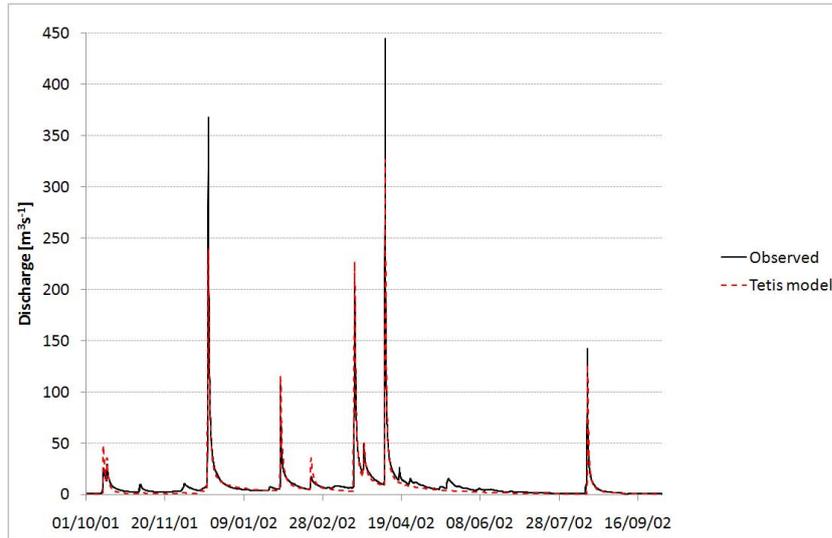
Figure 4.12: TETIS simulation during the entire calibration period, from October 2001 to September 2002. Observed discharges (black line); TETIS simulation (dashed line).

Max Q Obs	Max Q TPK	% Bias	RMSE	R^2	E
445 m^3s^{-1}	326 m^3s^{-1}	-13.6 %	7.13 m^3s^{-1}	0.96	0.91

Table 4.6: Evaluation indexes for the TETIS simulation computed during the calibration period. Maximum observed and simulated discharges; Percent Bias; Root Mean Square Error; Correlation Coefficient (R^2); Nash-Sutcliffe efficiency (E).

Figures from 4.13 to 4.16 depict some events occurred during the validation period and confirm the behavior seen during the calibration, that is the correct representation of the recession curves and the base flow, but a systematic underestimation of the peaks, which is accentuated with respect to the calibration.

Table 4.7 shows the evaluation indexes computed on the entire period, including validation and calibration data. The reduction of accuracy when the model is used in validation has been pointed out. Considering the entire set of data, the evaluation indexes does not differ much to the ones obtained with the TOPKAPI model, in fact they show a similar behavior, even if the TETIS model better represents the base flow while the TOPKAPI the peak flows.

Max Q Obs	Max Q TPK	% Bias	RMSE	R^2	E
1549 $m^3 s^{-1}$	807 $m^3 s^{-1}$	-15.4 %	13.92 $m^3 s^{-1}$	0.90	0.80

Table 4.7: Evaluation indexes for the TETIS simulation computed during the entire available period. Maximum observed and simulated discharges; Percent Bias; Root Mean Square Error; Correlation Coefficient (R^2); Nash-Sutcliffe efficiency (E).
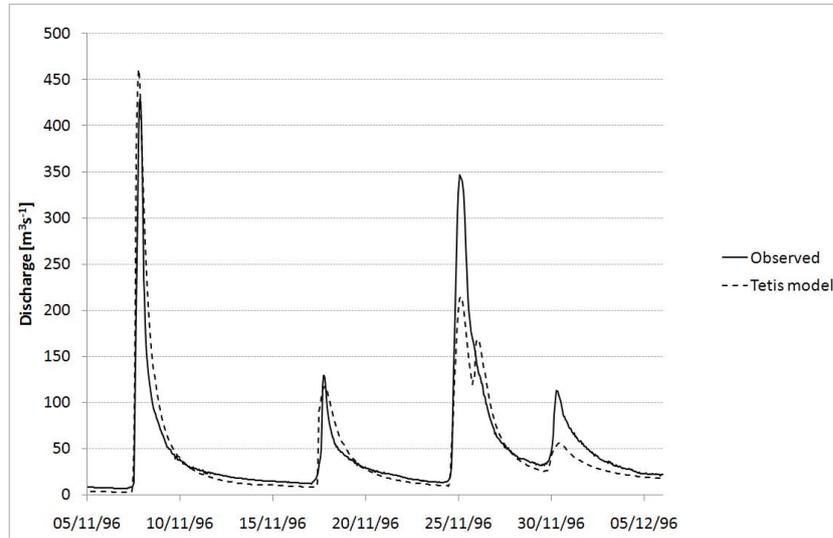
Figure 4.13: TETIS simulation of events occurred in November 1996. Observed discharges (black line); TETIS simulation (dashed line).
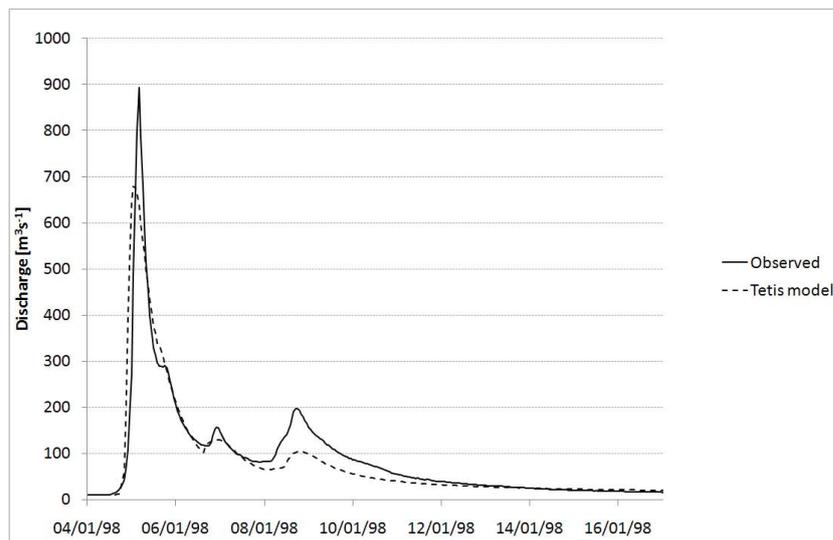
Figure 4.14: TETIS simulation of events occurred in April 1998. Observed discharges (black line); TETIS simulation (dashed line).
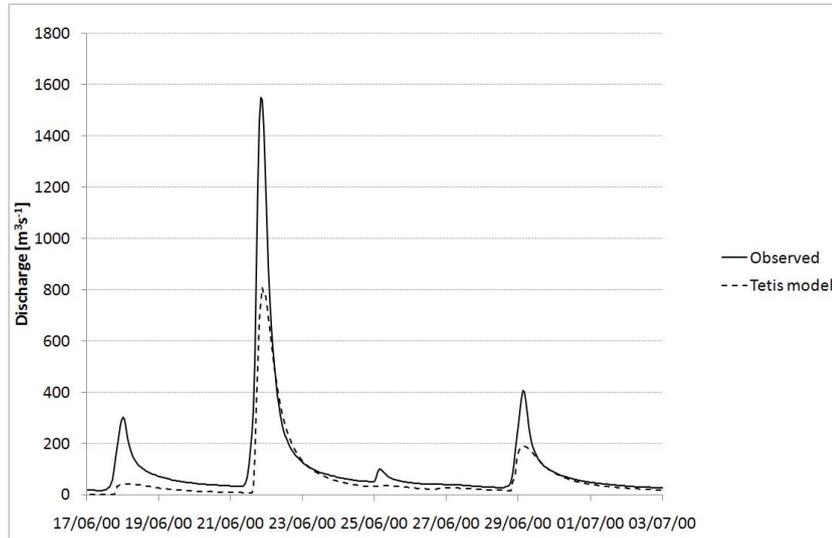
Figure 4.15: TETIS simulation of the main event occurred in June 2000. Observed discharges (black line); TETIS simulation (dashed line).
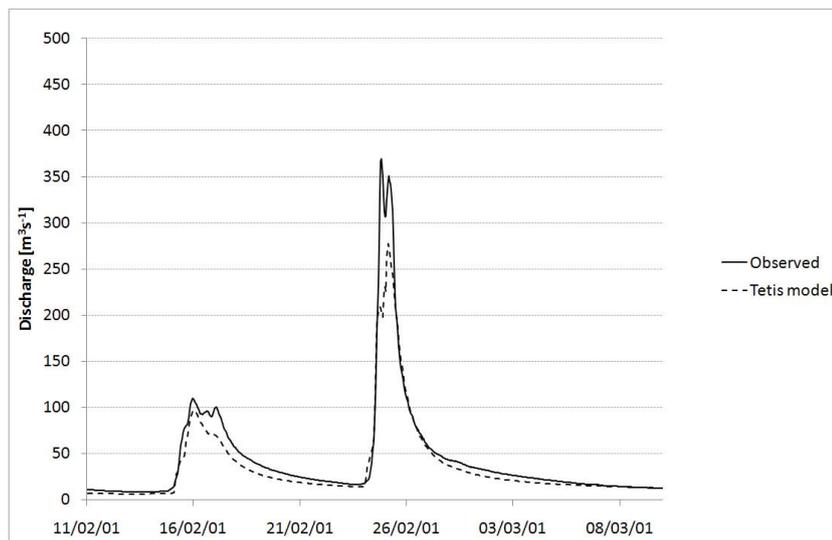
Figure 4.16: TETIS simulation of events occurred in November 2001. Observed discharges (black line); TETIS simulation (dashed line).

4.1.5   ANN Model Application

The Artificial Neural Network model was developed and applied in two main phases, as previously done by Toh and Brath (2008) and Pujol (2009). Firstly the data were divided in three groups through a Self Organizing Map (SOM) network (Kohonen, 1990) that automatically classifies the input data according to specific criteria established by the user. The SOM network has been calibrated using the data included in the period between 10/01/1995 and 05/31/1997, the remaining data until 09/30/2002 have been used for the validation process. The three sets of data classified by the SOM network have been used separately in order to calibrate three different Multi Layer Perceptron (MLP) networks (Werbos, 1974; Parker, 1987; Werbos, 1988, 1990; Pujol, 2009), which are able to predict discharges with an horizon time of 6 hours. Summarizing, using the SOM network the data have been divided in three groups, which represent three different hydrological states of the system; then each group has been calibrated with a Feed Forward Network in order to forecast the discharge 6 hours in advance. Moreover, to avoid the risk of overfitting the calibration data, the *early stopping procedure* has been used introducing a verification set of data, form 06/01/1997 to 01/31/1998. The data from 02/01/1998 to 09/30/2002 have been used for validating the model.

Self Organizing Map (SOM) Network

The SOM network, introduced by (Kohonen, 1990), is an unsupervised classification methodology that allows data to be classified on a statistic basis, without an *a-priori* class definition. In the structure proposed by (Kohonen, 1990) the neurons are topologically ordered and they compete against each other until just one of them is activated. The learning objective is to link topologically similar neurons and input patterns with similar features. The SOM network is composed of two neuron layers, the input one and the output one. The output layer is composed by a matrix with dimensions equals

to the number of classes.

In this application, the SOM network has been fed with pattern composed of 4 variables. If the time at which the prediction is done is called t_0, the input data to the SOM network are the accumulated precipitation of 2 days before t_0 ($AccP$), the discharge observed at t_0 (Q_{t_0}) and the gradient of the discharge during 2 hours before t_0 (ΔQ_{t_0-1} and ΔQ_{t_0-2}). The output layer contains a vector with three elements and each element corresponds to one of the three classes in which the data will be classified. Hence, the network output is a number included between 1 and 3 that represents the class to which the input pattern belongs. The flooding process non-linearity is mainly due to the soil saturation state therefore the classes' identification identification procedure should respect this state. In fact, this approach aims at identifying whether the precipitation event occurs when the soil is dry, saturated or almost saturated, so that each one of these conditions will be then processed with a specific network in order to reduce the process non-linearity.
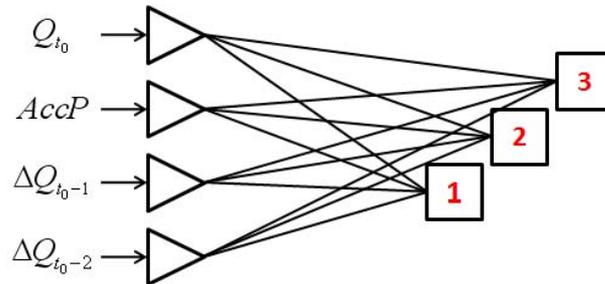
Figure 4.17: Schematic scheme of the SOM network.

Figure 4.17 depicts the simple conceptual SOM scheme above described and Figures 4.18 and 4.19 show two examples of the obtained classification. In these figures the depicted discharge is the forecast target, which is the discharge 6 hours in advance; hence, the discharge values belonging to the same group will be then predicted by the same network. The flood events belong to the first class, the low flows to the third one and the second class contains uncertain situations. The classification results are quite satisfactory,

even if outliers are present in each class. For example, some low flows have been included in the first group, so this may lead to predict an event that did not actually occur. Another example concerns the event occurred approximately at the end of May 2000 (Figure 4.18), in which data corresponding to some hours before the flood event were included in the first group which will probably bring forward the event prediction.
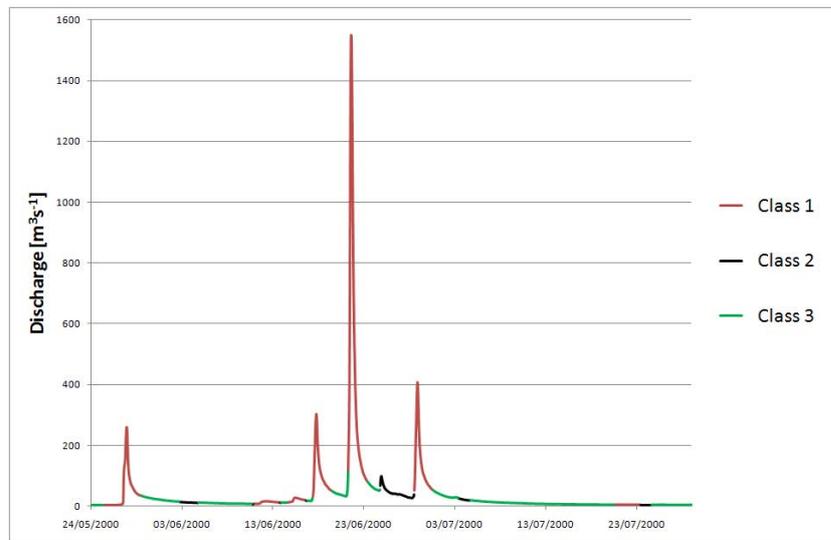
Figure 4.18: Classification obtained with SOM network for the events occurred during summer 2000.

Multi Layer Perceptron (MLP) Networks

The second step of the ANN model application was the processing of the previously classified data through the Multi Layer Perceptron networks (Werbos, 1974; Parker, 1987; Werbos, 1988, 1990; Pujol, 2009). This kind of network is one of the most used in hydrology, especially coupled with a Back-Propagation algorithm for its training. Its structure is composed of at least three neuron layers: input, output and one (or more) hidden layer. The input and output layers contain respectively as many neurons as the number of input and output variables; the number of the neurons in the hidden layer
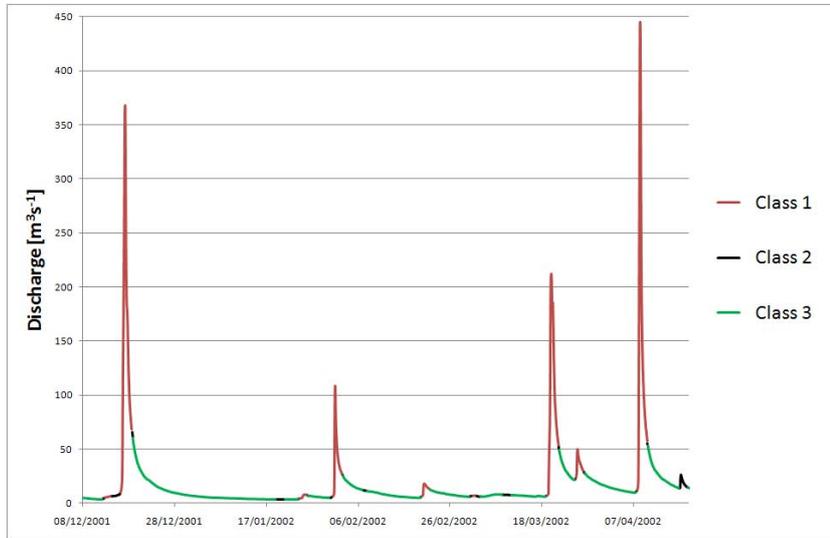
Figure 4.19: Classification obtained with SOM network for the events occurred in the first months of 2002.

is one of the parameters that have to be set concerning the network structure. The weights of each neuron connection are automatically identified during the calibration process, in this phase the Back-Propagation algorithm adapts them to a set of pattern proposed to the network with both the input and output variables. Iteratively the algorithm modifies the weights in order to minimize the error function until the error tolerance is reached.

In this case three MLP networks, one for each class identified by the SOM network, have been set up to forecast discharges 6 hours in advance starting from the available observed variables at the forecast time (t_0). The variables used as input to the network are the observed precipitation and discharges at t_0 and, respectively, during 12 and 2 hours before. The output of the networks is the discharge 6 hours after the t_0. Hence, the input layers of each network contain 16 neurons and the output layers 1. After several tests, simple networks' structures have been chosen; every network has been calibrated with different numbers of hidden neurons and with 100 different parameters' initial states. For each network, the number of hidden neurons that gave the best verification results has been selected; the model performance was eval-

uated computing the Nash coefficient and the RMSE and taking in account their mean values and their variances on the 100 calibrations with different initial states.

Figures from 4.20 to 4.25 represent the evaluation indexes resulting from this test. For the first class, containing 2195 data in the calibration set, the network configuration that led to the best results was composed of 6 hidden neurons, for the second class (1708 data), the best results have been obtained with 3 neurons and for the third class (containing 10690 data) with 5 neurons. As was expected, the third class gave the best performances because it represents the low flows while the first class, which represents the high flows, was more difficult to be well reproduced.
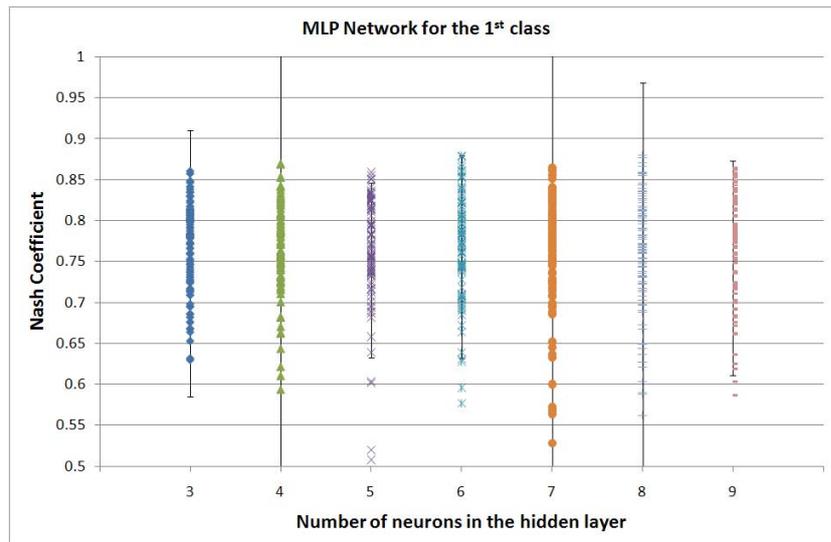
Figure 4.20: Nash efficiency computed for different numbers of neurons and different initial weights' value for the first class.

Results

The available data have been divided in three sets as described above and depicted in Figure 4.32. The verification set has been used to check the error function at the same time that the calibration in order to stop it when the
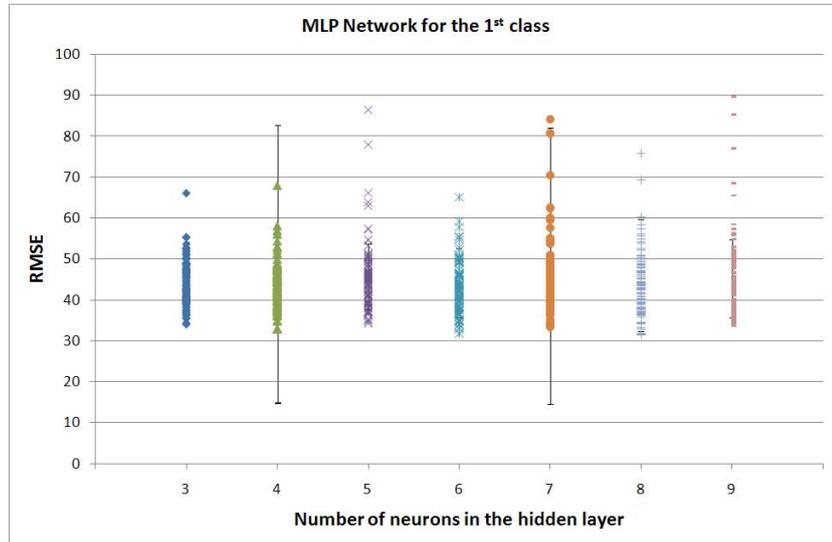
Figure 4.21: RMSE computed for different numbers of neurons and different initial weights' value for the first class.
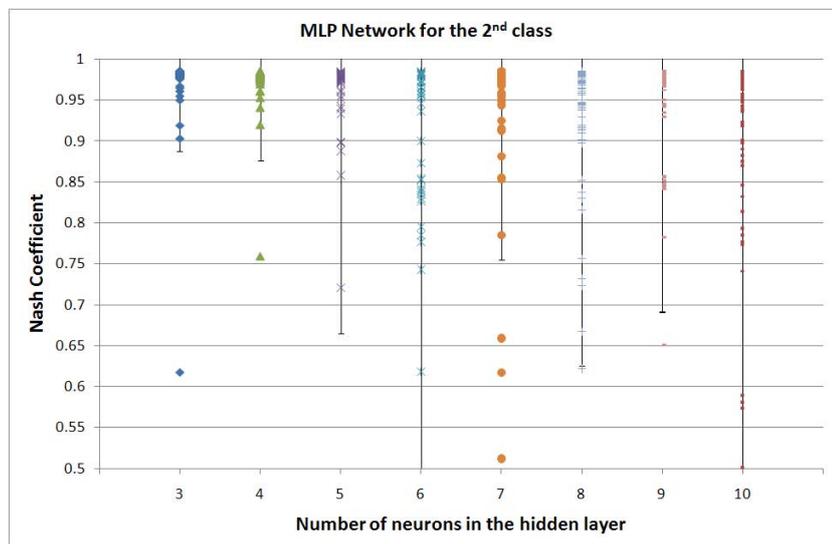
Figure 4.22: Nash efficiency computed for different numbers of neurons and different initial weights' value for the second class.
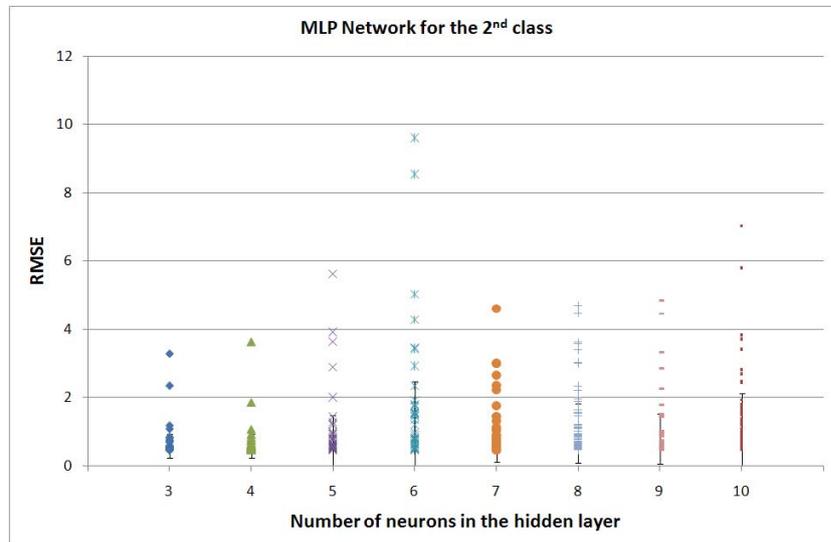
Figure 4.23: RMSE computed for different numbers of neurons and different initial weights' value for the second class.
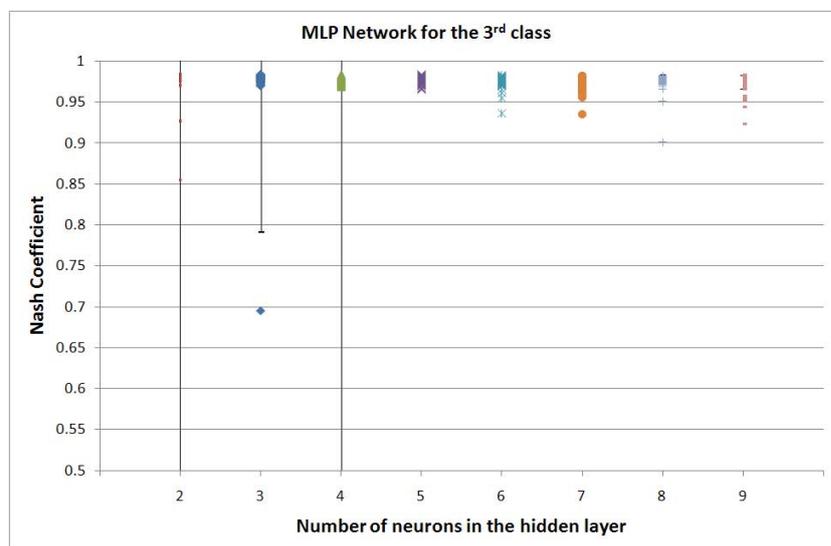
Figure 4.24: Nash efficiency computed for different numbers of neurons and different initial weights' value for the third class.
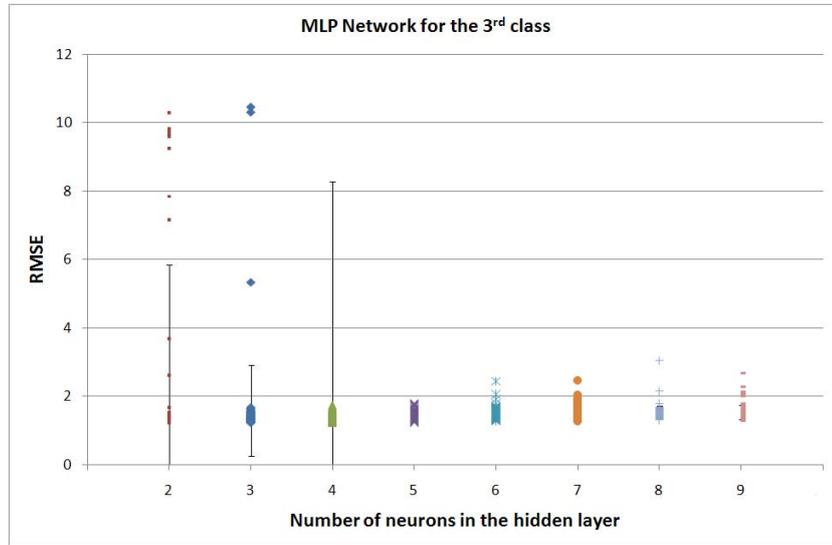
Figure 4.25: RMSE computed for different numbers of neurons and different initial weights' value for the third class.

efficiency in verification begins to decrease, even if the error tolerance has not been reached. This procedure is called *early stopping* and it is used to avoid the calibration data overfitting and the lack of generalization ability. Table 4.8 shows some evaluation indexes computed during the calibration, verification and validation periods and highlights that the system of networks is able to reproduce data different from those seen during the calibration. The calibration, verification and validation evaluation indexes do not differ much one from each other. In fact, the RMSE is even greater in validation than in calibration, the Nash efficiency decreases from 0.91 to 0.88 in validation and only the percent bias has a significant increase from 0.95 % to 4.15 %, even if the absolute value just changes from 0.11 m^3s^{-1} to 0.4 m^3s^{-1}.

Figures 4.26 and 4.27 depict some events occurred during the calibration period and Figures from 4.28 to 4.30 during the validation and verification periods. It can be observed that the ANN model produces good simulations, even if in some cases it is instable, especially for small events (Figure 4.31).

Indexes	Periods			
	Entire	Calibration	Verification	Validation
Max Q Obs $[m^3s^{-1}]$	1549	893	894	1549
Max Q ANN $[m^3s^{-1}]$	1706	876	617	1706
Percent Bias [%]	3.06	0.95	2.14	4.15
RMSE $[m^3s^{-1}]$	10.46	10.92	12.19	10.00
R^2	0.94	0.952	0.94	0.94
E	0.89	0.91	0.88	0.88

Table 4.8: Evaluation indexes for the ANN simulation computed during entire, calibration, verification and validation periods. Maximum observed and simulated discharges; Percent Bias; Root Mean Square Error; Correlation Coefficient (R^2); Nash-Sutcliffe efficiency (E).
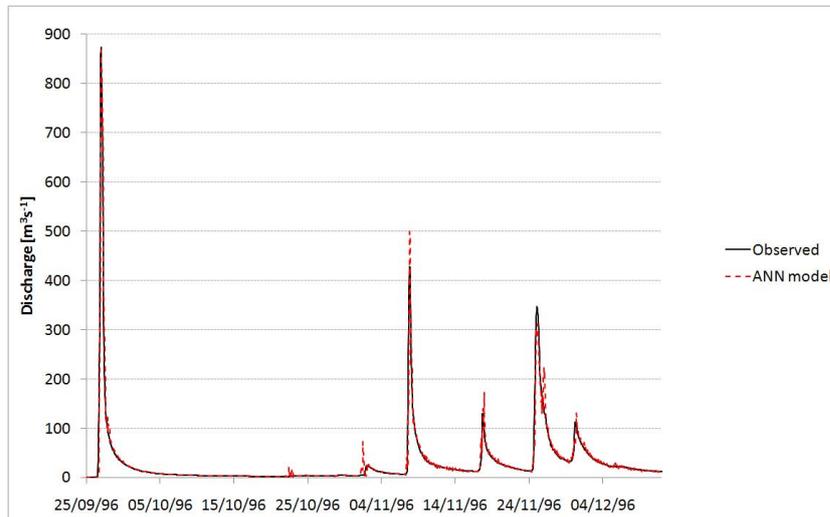
Figure 4.26: ANN simulation of calibration events. Observed discharges (black line); ANN simulation (dashed line).
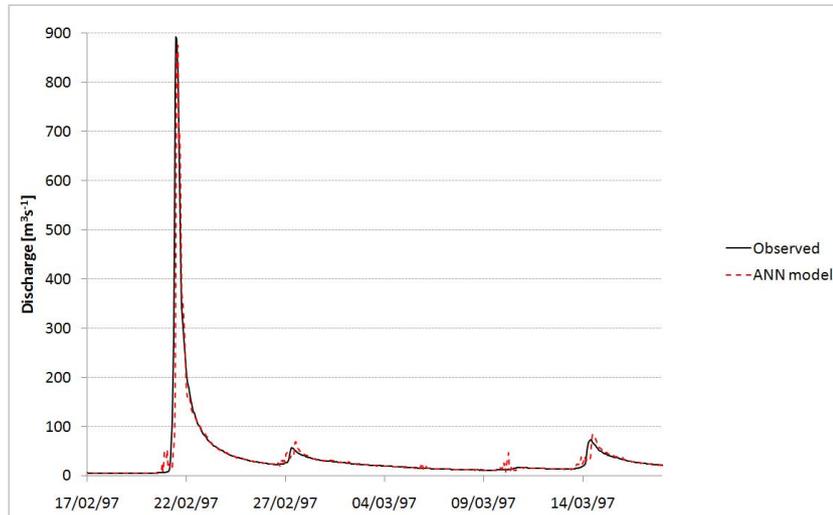
Figure 4.27: ANN simulation of calibration events. Observed discharges (black line); ANN simulation (dashed line).
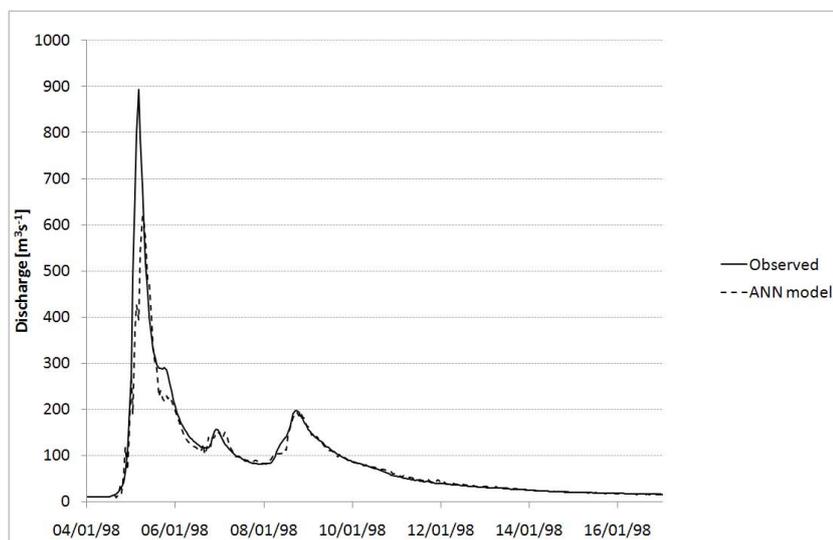
Figure 4.28: ANN simulation of validation events. Observed discharges (black line); ANN simulation (dashed line).
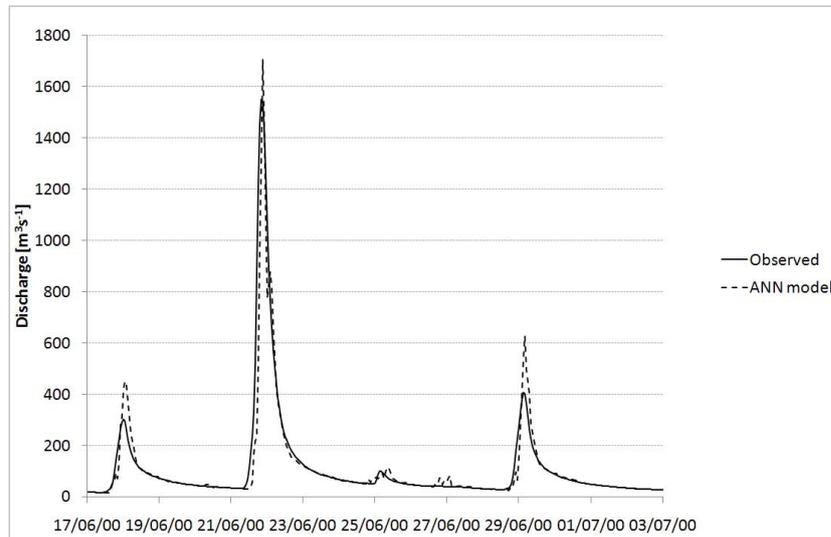
Figure 4.29: ANN simulation of validation events. Observed discharges (black line); ANN simulation (dashed line).
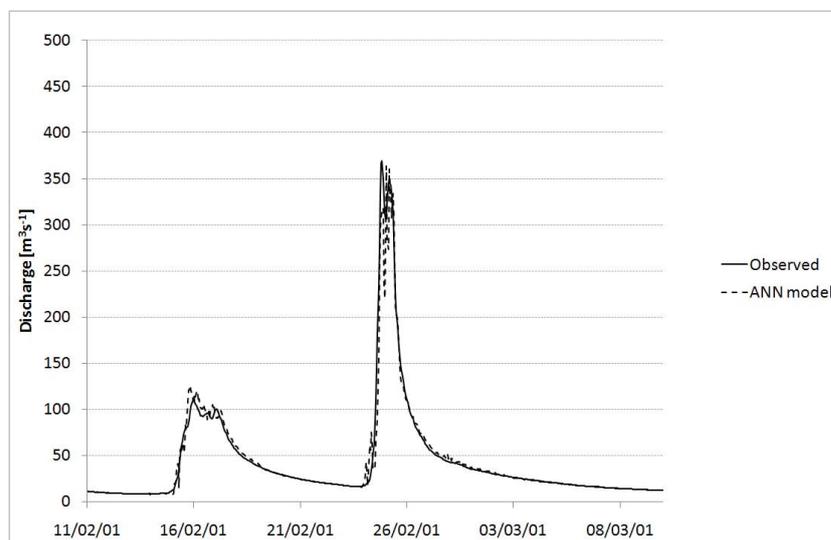
Figure 4.30: ANN simulation of validation events. Observed discharges (black line); ANN simulation (dashed line).
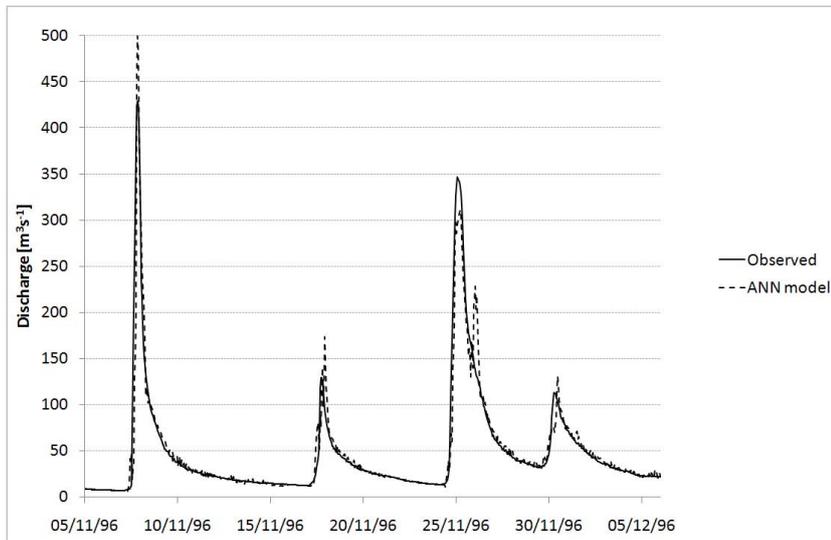
Figure 4.31: Instability of ANN in small validation events. Observed discharges (black line); ANN simulation (dashed line).

4.1.6   MCP Application

Data set

In Figure 4.32 a schematic summary of the division of the data used for calibrating and validating each model is depicted.

The two physically based models are conceptually quite similar; it can be highlighted that the TOPKAPI model tends to underestimate the highest flood events, to overestimate the smallest ones and to reproduce the flood events of medium magnitude quite well. The TETIS model also generally underestimates the highest events and often underestimates the small events too. The ANN model, due to its nature of data driven model, is not able to well reproduce the peack flows, which are often underestimated and predicted with late of 1 or 2 hours, but it has a perfect behaviour in reproducing the low flows.
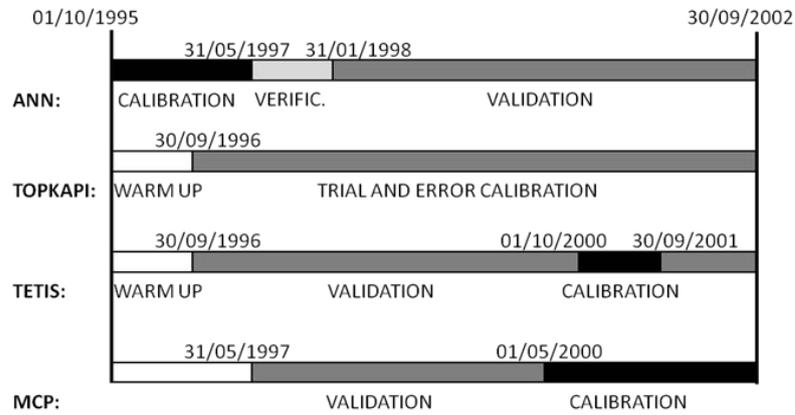
Figure 4.32: Schematization of the available data division for calibrating and validating the models and the MCP

Predictive Uncertainty Assessment

The MCP has been applied in three phases and in every phase the Joint TNDs have been used.

1. Firstly the models simulations have been processed separately. All the historical data have been processed and the expected value of the predictive distribution has been computed at each time step. Figures 4.33(a), 4.33(b) and 4.33(c) represent schematically the predictive distribution computed separately with each model. For the ANN model it was not necessary to divide the data in two samples because the joint distribution of observed and forecasted transformed values was well represented by just one bi-variate normal distribution. The TNDs have been used for the other two models and both of them give a lower uncertainty for the upper sample. Concerning the TOPKAPI model it is necessary to note that the threshold used for the division in two samples seems to be too low, because a threshold of about 2.4 would have further reduced the uncertainty in the upper sample. The processor found that threshold because the seeking is lower and upper limited in order to count with significant samples for computing the moments

of the truncated distribution.

2. In the second phase, the series of the expected values of each model simulation have been processed with the MCP multivariate approach and the combined expected value has been computed from the predictive uncertainty of each time step.

3. Finally, in the third phase, this series of expected values has been processed. Figure 4.34 shows the normal space obtained in this phase, also in this case the use of the TNDs was not necessary.

Figures 4.35 and 4.36 summarize the obtained results with regard to the models combination computed by means of the expected value of the predictive distribution. Figure 4.35 represents the Error Standard Deviation and Figure 4.36 represents the Nash-Sutcliffe coefficient.

In Figures 4.37 - 4.38 two examples of models combination are shown, one during the calibration period and the other one during the validation period. In both events the uncertainty band is narrower as the number of models increases and in the calibration event the expected value computed with the combination of all the models well matches the observed series. In the validation event, the peack flow is quite better represented when only the TOPKAPI model is used, probably due to its better forecast in this specific case, but also in this event the uncertainty band is reduced combining all the models.

The knowledge of the uncertainty distribution also allows the probability of exceeding an alert threshold to be estimated, that is a stochastic way to predict the flooding risk. In section 4 a way to identify the alert threshold, different to the deterministic method commonly used, will be discussed. The threshold has been set at 350 $m^3 s^{-1}$. In Figures 4.39 and 4.40, the comparison between the deterministic and stochastic discharge forecasts and the correspondent probability of overtopping the threshold is shown.
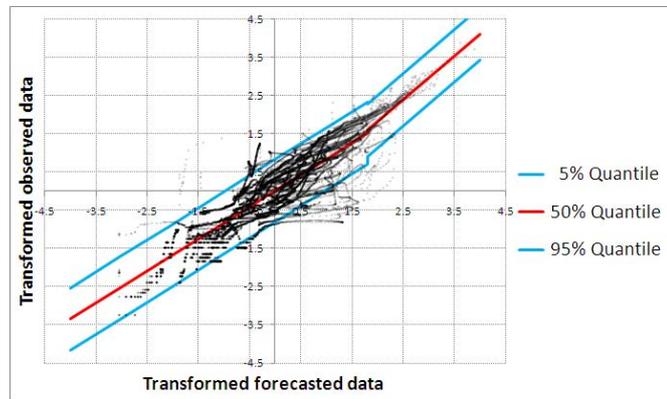
(a)
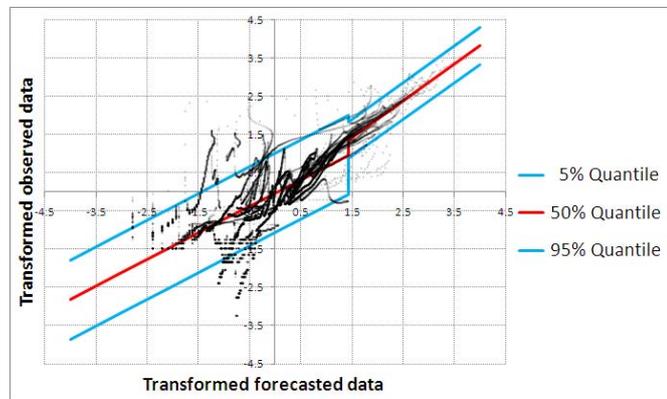
(b)
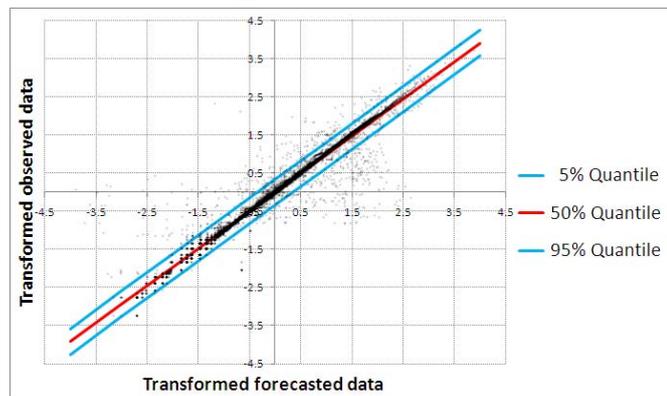
(c)

Figure 4.33: Representation of the Normal Space obtained using the MCP with the TOPKAPI (a), TETIS (b) and ANN (c) forecasts.
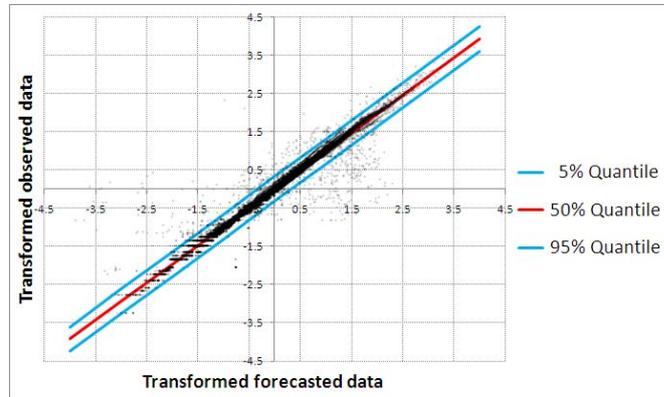
Figure 4.34: Representation of the Normal Space obtained using the MCP with the combination of the 3 models.

Comparison Between MCP with TNDs and QR

In this paragraph a comparison between the results obtained using the TNDs with the MCP and the Quantile Regression methodology (QR) will be presented. For sake of consistency the QR has been applied in three phases, as in the case of the MCP, combining all the deterministic models. Firstly each model has been processed independently and then the results of this first phase have been combined with a Multiple Quantile Regression. For 19 quantiles $\tau = 5, 10, ..., 95$ the regression parameters have been computed following the Equation 2.8. The expected value in the real space has been computed as the average of the 19 quantiles transformed using the inverse NQT.

Figure 4.43 shows that the uncertainty band given by the QR is much narrower than the one obtained with the MCP, even if some clues let think this uncertainty band is not realistic. Namely, first of all it is too influenced by the ANN model, which is the cause of the oscillations during the rising limb. Secondly, is the closeness between the expected value and the 95th quantiles, especially visible for high discharges. Looking at the uncertainty band obtained for the validation event in Figure 4.44 the inconsistent behavior of the QR is confirmed. In fact, most of the main event is out of the
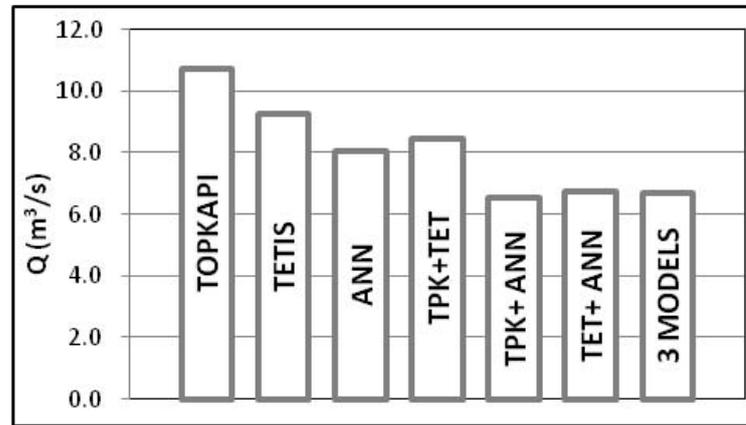
Figure 4.35: Error Standard Deviation for TOPKAPI model (TPK), TETIS model (TET), ANN model and their combinations during the entire validation period of the MCP.
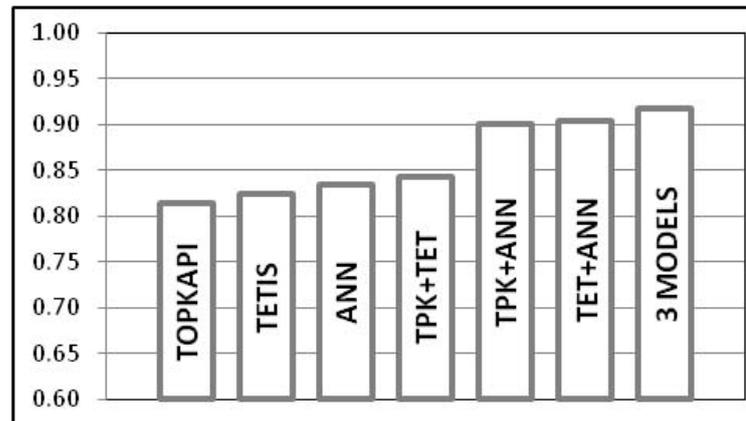
Figure 4.36: Nash-Sutcliffe coefficient for TOPKAPI model (TPK), TETIS model (TET), ANN model and their combinations during the entire validation period of the MCP.
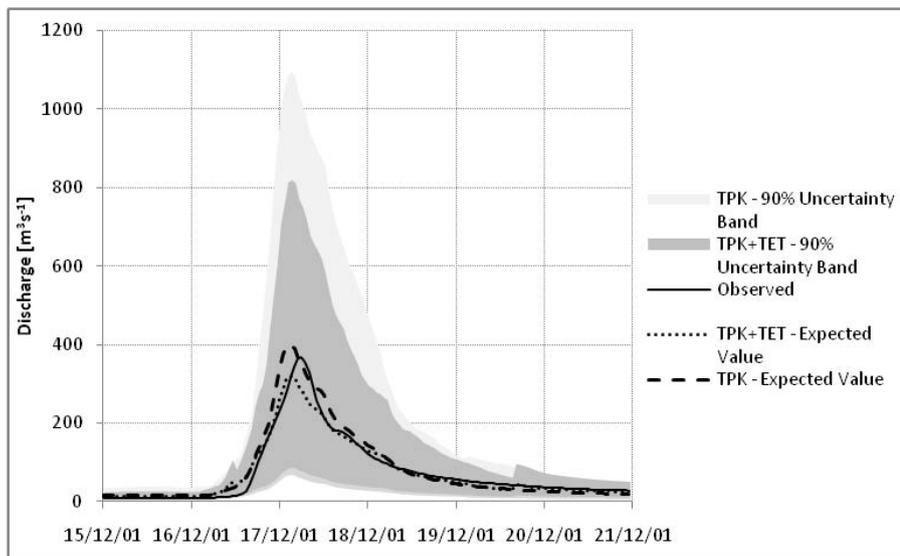
Figure 4.37: Comparison between the PU computed with one or two models on a flood event during calibration period. Observed discharges (black line); expected value conditioned only to the TOPKAPI forecast (dashed line); expected value conditioned to the TOPKAPI and TETIS forecasts (dotted line); 90% Uncertainty Band conditioned to the TOPKAPI forecast (light grey band); 90% Uncertainty Band conditioned to the TOPKAPI and TETIS forecasts (grey band).
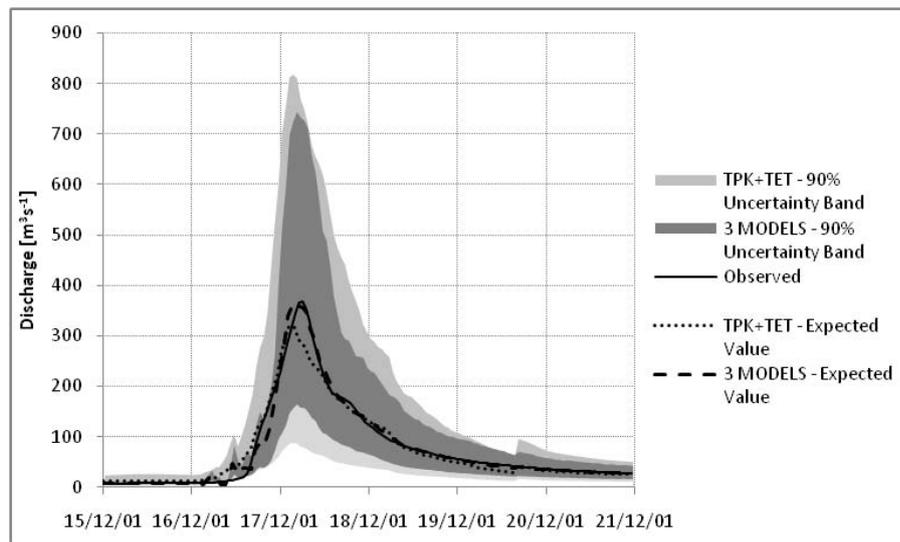
Figure 4.38: Comparison between the PU computed combining, two or three models on a flood event during calibration period. Observed discharges (black line); expected value conditioned only to the TOPKAPI and TETIS forecasts (dotted line); expected value conditioned to the TOPKAPI, TETIS and ANN forecasts (dashed line); 90% Uncertainty Band conditioned to the TOPKAPI and TETIS forecasts (light grey band); 90% Uncertainty Band conditioned to the TOPKAPI, TETIS and ANN forecasts (grey band).
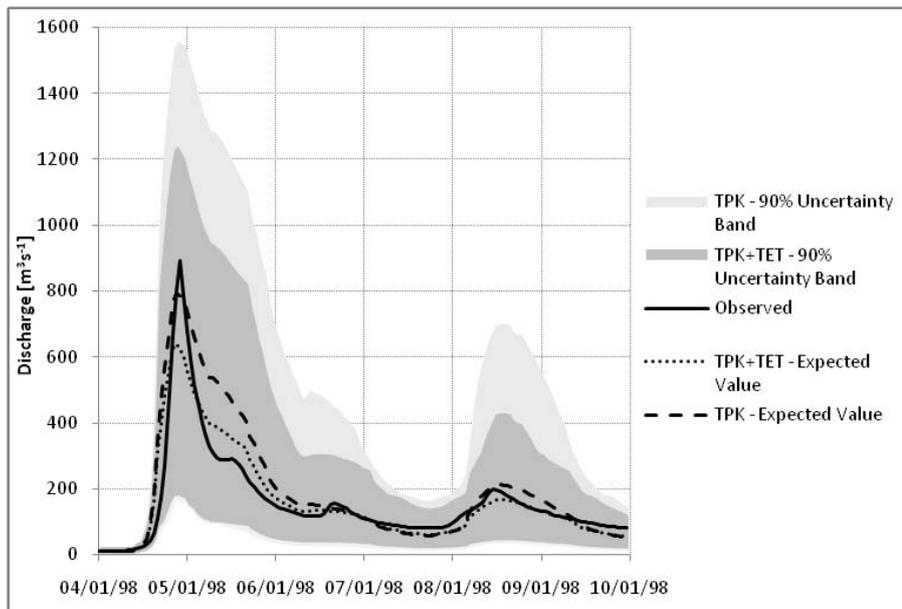
Figure 4.39: Comparison between the PU computed with one or two models on a flood event during validation period. Observed discharges (black line); expected value conditioned only to the TOPKAPI forecast (dashed line); expected value conditioned to the TOPKAPI and TETIS forecasts (dotted line); 90% Uncertainty Band conditioned to the TOPKAPI forecast (light grey band); 90% Uncertainty Band conditioned to the TOPKAPI and TETIS forecasts (grey band).
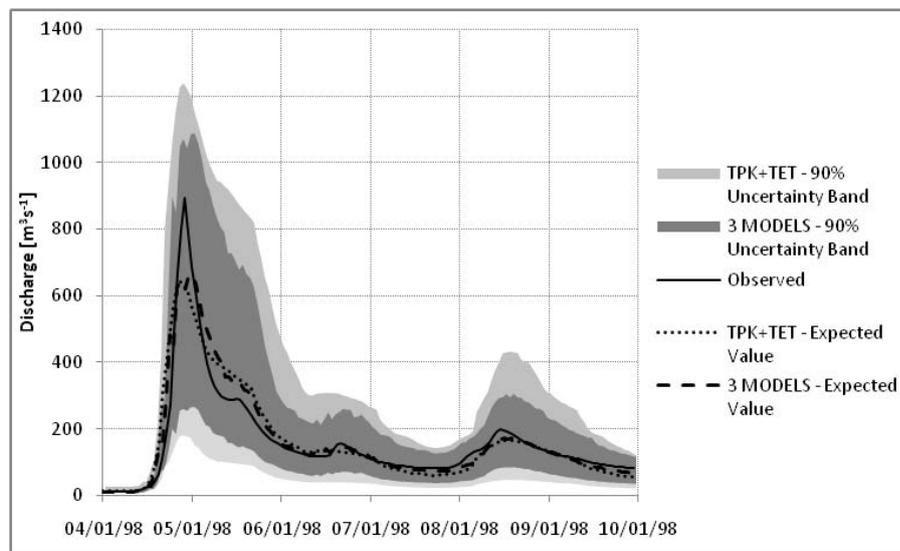
Figure 4.40: Comparison between the PU computed combining, two or three models on a flood event during validation period. Observed discharges (black line); expected value conditioned only to the TOPKAPI and TETIS forecasts (dotted line); expected value conditioned to the TOPKAPI, TETIS and ANN forecasts (dashed line); 90% Uncertainty Band conditioned to the TOPKAPI and TETIS forecasts (light grey band); 90% Uncertainty Band conditioned to the TOPKAPI, TETIS and ANN forecasts (grey band).
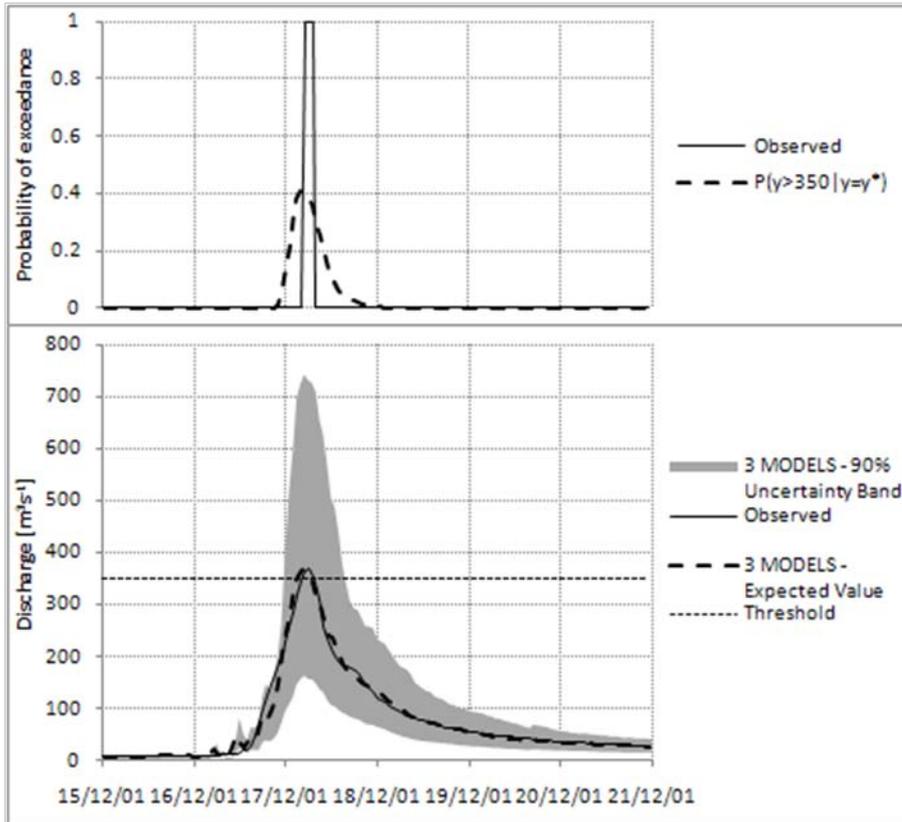
Figure 4.41: Flood event during calibration period. The lower part represents the discharge forecast; observed values (continuous line); expected value conditioned to the TOPKAPI, TETIS and ANN forecasts (dashed line); 90% Uncertainty Band (grey area); alarm threshold of 350 m^3s^{-1} (small dashed line). The upper part represents the probability of exceeding the alarm threshold; observed binary response (continuous line) and Probability of exceeding the threshold computed by the MCP (dashed line).
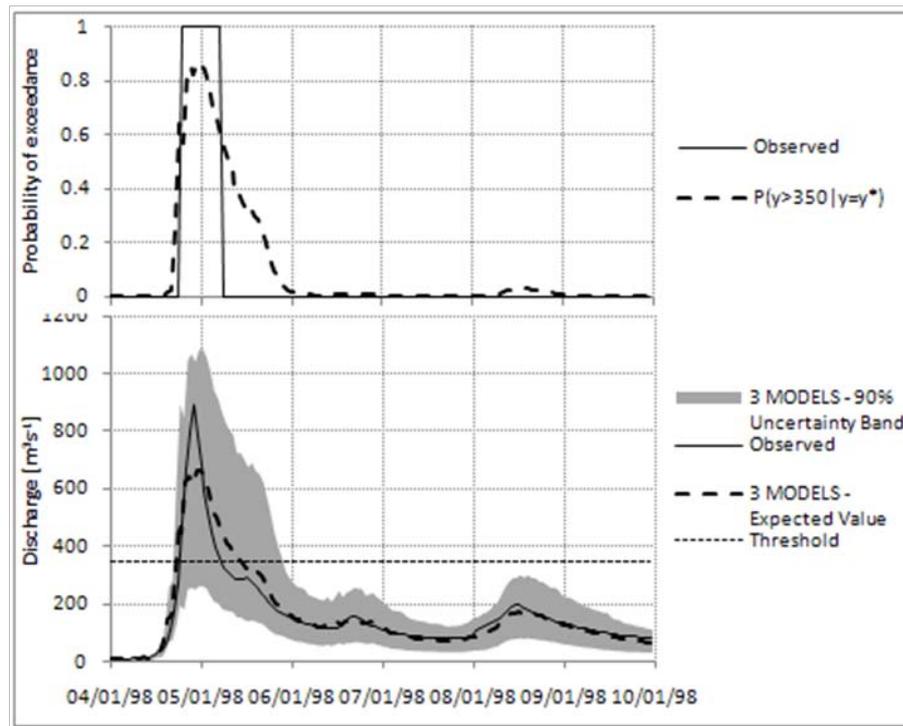
Figure 4.42: Flood event during validation period. The lower part represents the discharge forecast; observed values (continuous line); expected value conditioned to the TOPKAPI, TETIS and ANN forecasts (dashed line); 90% Uncertainty Band (grey area); alarm threshold of 350 m^3s^{-1}(small dashed line). The upper part represents the probability of exceeding the alarm threshold; observed binary response (continuous line) and Probability of exceeding the threshold computed by the MCP (dashed line).

band and it is delayed of some hours. Comparing Figures 4.43 and 4.28 it
is possible to note that the band given by the QR follows the ANN wrong
prediction, hence a little weight is assigned to the forecasts of the distributed
models, which are better in this case.
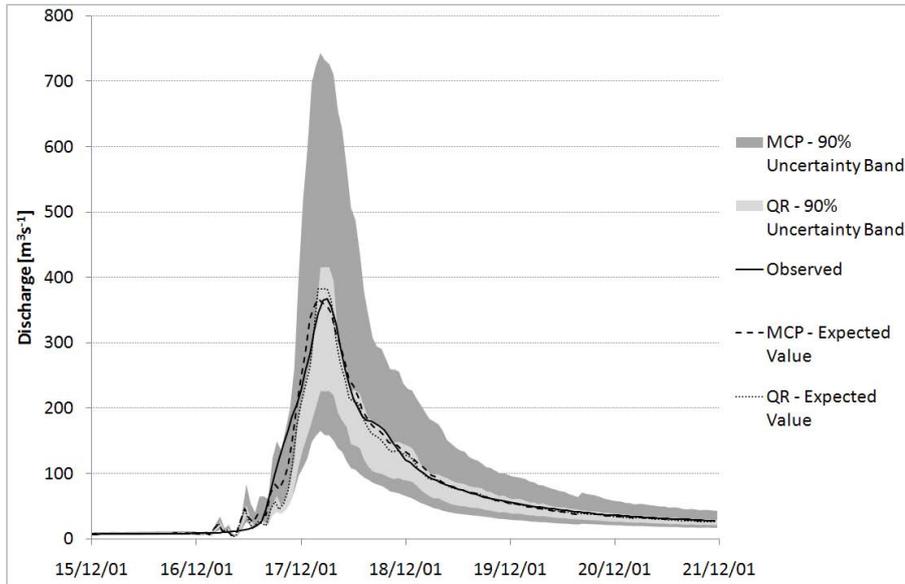
Figure 4.43: Comparison between the PU computed with the MCP and the
QR on a flood event during calibration period. Observed discharges (black
line); expected value obtained with the MCP (dashed line); expected value
obtained with the QR (dotted line); 90% Uncertainty Band computed by the
MCP (dark grey band); 90% Uncertainty Band computed by the QR (light
grey band).

Also the representation of the analysis of the Normal Space obtained us-
ing the QR can help to understand why the Predictive Uncertainty is not
well estimated. For sake of simplicity, Figure 4.45 shows the 5% and 95%
quantiles computed only with the prediction of the TOPKAPI model, even
if it would have been more representative, unfortunately the Multi-Normal
Space is not easy to be depicted. In Figure 4.45 is possible to note as the
lower quantile does not well represent the data, especially for the highest
values; due to the joint distribution of the observed and predicted discharges
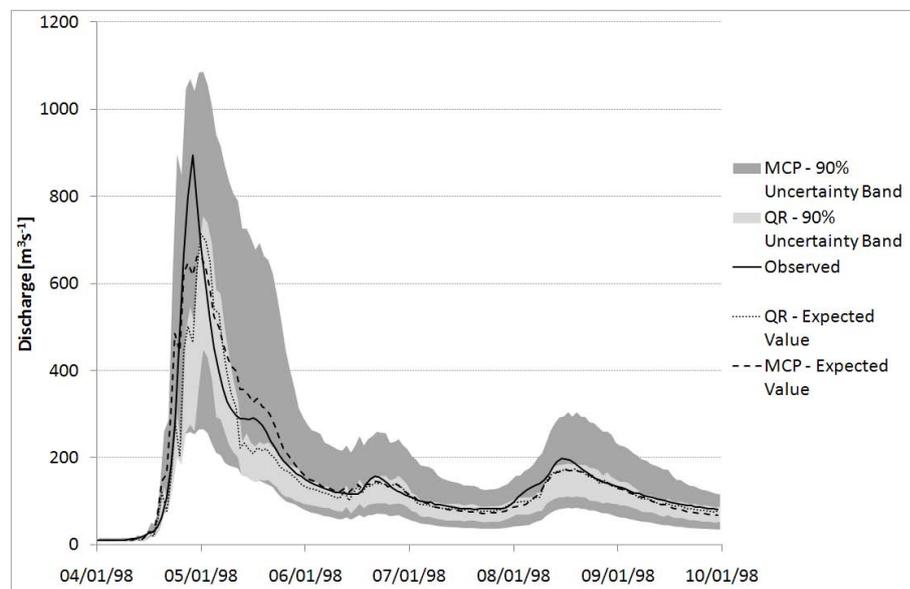
Figure 4.44: Comparison between the PU computed with the MCP and the QR on a flood event during validation period. Observed discharges (black line); expected value obtained with the MCP (dashed line); expected value obtained with the QR (dotted line); 90% Uncertainty Band computed by the MCP (dark grey band); 90% Uncertainty Band computed by the QR (light grey band)).

the QR is not able to give a realistic estimation of the PU in this application.
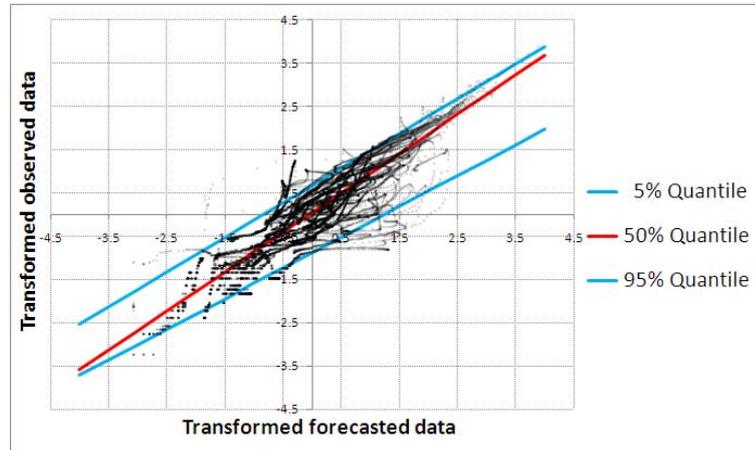
Figure 4.45: Representation of the Normal Space obtained using the Quantile Regression with the TOPKAPI forecasts.

This is confirmed also by Figure 4.46 where the evaluation indexes computed on the entire validation period are shown.

The Nash-Sutcliffe coefficient (Figure 4.46(b)) obtained with the QR is similar to the ones computed with the deterministic models and it is much lower than the one obtained with the MCP. The Error Standard Deviation (Figure 4.46(a)) is even greater than the one computed with the ANN model deterministic forecast and much greater than the one computed with the MCP.

An alternative way to use the QR in cases like this is to divide the data in representative samples, similar to what has be done introducing the Truncated Normal Distributions in the MCP. With this change the QR can achieve good results in the PU assessment also in cases where the error eteroschedasticity is high (Figure 4.47) and the evaluation indexes are improved, even if
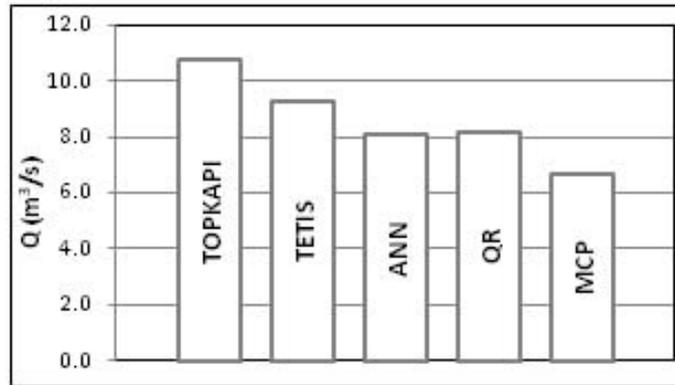
(a)
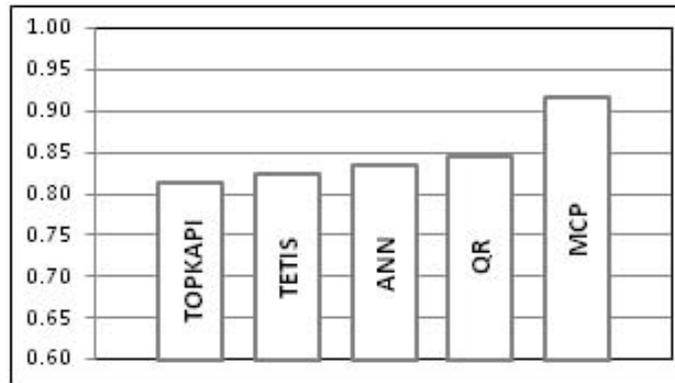
(b)

Figure 4.46: Error Standard Deviation (a) and Nash-Sutcliffe coefficient (b) for TOPKAPI model, TETIS model, ANN model and their combination with the MCP and the QR during the entire validation period.

they are still worse than the ones obtained with the MCP (Figure 4.48).

Nevertheless, this methodology has three disadvantages that must be considered. The identification of different samples requires a great number of parameters to be estimated for applying the QR; referring to the present application, in order to combine three model forecasts, dividing the data in two samples and computing 19 quantiles, it was necessary to compute 152 parameters, 8 for each quantile. Another risk in using the QR and dividing the data in several samples is the *overfitting* of the calibration data, which
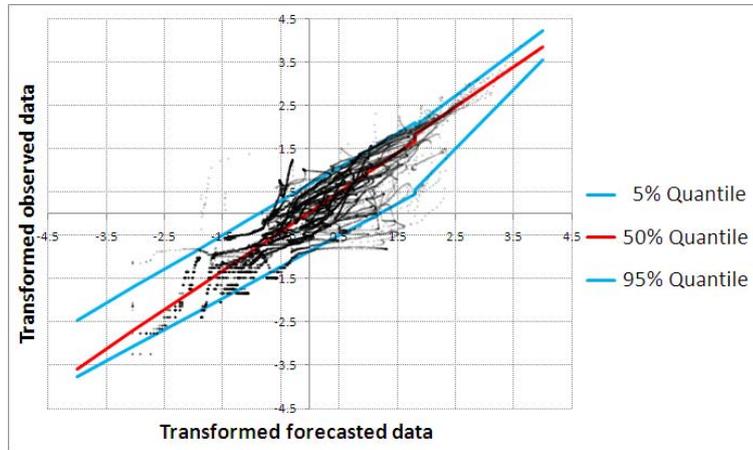
Figure 4.47: Representation of the Normal Space obtained using the Quantile Regression with the TOPKAPI forecasts divided in 2 samples.

can lead to a lack of generalization ability. Moreover, a common problem in QR is the quantile crossing, which can be solved but it requires an additional computational effort.

Probabilistic Thresholds Analysis

The obtained results allow also an analysis of the correctness of probability of exceeding an alert threshold, estimated using MCP together with the improvement obtainable using the combination of models, to be performed.

As can be seen from the Figure 4.49, apart from a small bias in the lower part mainly due to the larger error variance of the lower Truncated Normal, there is a relatively good agreement between the actual threshold exceedances and the probability of exceedance estimated from the PU density obtained through the MCP combination of the three models (TOPKAPI, TETIS and ANN). This agreement allows for the change of paradigm discussed in Section 1.2, which would not be possible in case of incorrect estimate of the quantiles.

In addition, Tables 4.9 and 4.10 allow to exemplify the improvements obtainable by the Bayesian combination of the different models. Table 4.9 confirms the behaviour represented in Figure 2.2 showing the probability that
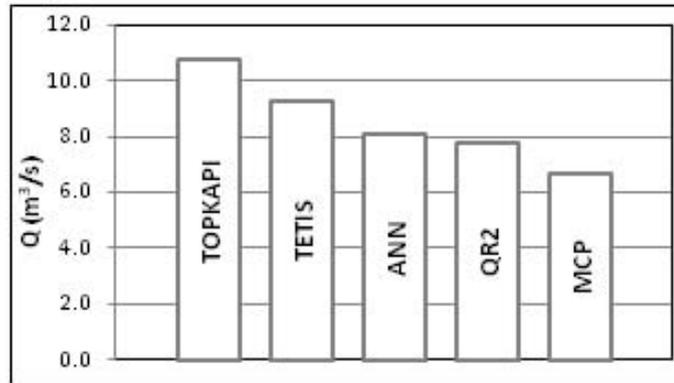
(a)
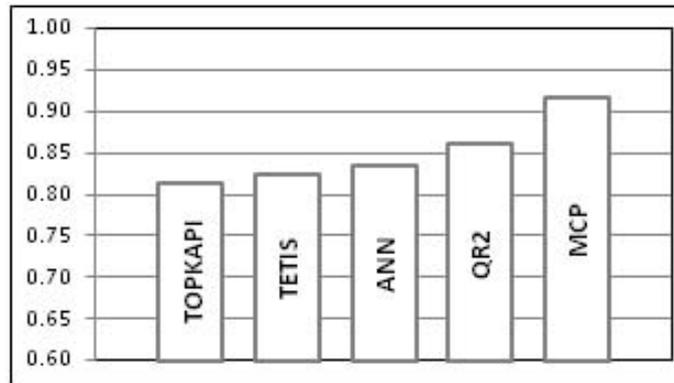
(b)

Figure 4.48: Error Standard Deviation (a) and Nash-Sutcliffe coefficient (b) for TOPKAPI model, TETIS model, ANN model and their combination obtained with the MCP and the QR with the division in 2 samples during the entire validation period.

the true value exceeds the 350 m^3s^{-1} threshold when the expected value of prediction equals 250 m^3s^{-1}, computed for each model and their Bayesian combination. One can see the reduction of exceedance probability as a function of the quality of the forecast. Finally, the effect of the introduction of the new probabilistic forecast paradigm can be appreciated in Table 4.10 that shows, similarly to what is qualitatively displayed in Figure 2.3, the expected value of the prediction corresponding to the probability of 20% that the true value will exceed the 350 m^3s^{-1} threshold, computed for each model
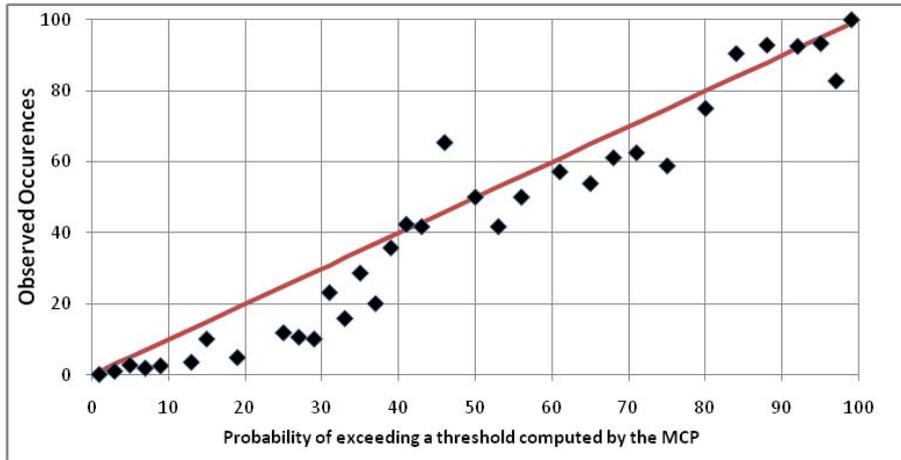
Figure 4.49: Frequency of actual threshold exceedances vs the probability estimated using the MCP Bayesian combination of the three models. The red line represents the perfect behaviour.

and their Bayesian combination. As can be seen better models allow to wait until the expected value of prediction is closer to the flooding level, while worse models require earlier action corresponding to lower levels on the basis of the principle of precaution, which corresponds to the fact that the decision maker is more uncertain.

$P\left(y > 350 m^3 s^{-1} | \widehat{y} = 250 m^3 s^{-1}\right)$			
TOPKAPI	*TETIS*	*ANN*	*3 MODELS*
0.25	0.34	0.16	0.15

Table 4.9: Probability that the true value exceeds the 350 $m^3 s^{-1}$ threshold when the expected value of prediction equals 250 $m^3 s^{-1}$, computed for each model and their Bayesian combination.

$E\left[y|\widehat{y}\right]|\left[P\left(y > 350m^3s^{-1}\mid\widehat{y}\right) = 0.2\right]$			
TOPKAPI	*TETIS*	*ANN*	*3 MODELS*
$217\ m^3\mathrm{s}^{-1}$	$138\ m^3s^{-1}$	$270\ m^3s^{-1}$	$284\ m^3s^{-1}$

Table 4.10: Expected value of prediction corresponding to the probability of 20% that the true value will exceed the 350 m^3s^{-1} threshold, computed for each model and their Bayesian combination.

4.2 Po River, Multi-Temporal Application

4.2.1 Study Case

The Po river is the main italian river, both in terms of length (652 km) and discharge (it can reach 10.500 $\mathrm{m}^3\mathrm{s}^{-1}$ near Pontelagoscuro) (Figure 4.50). It starts from the Monviso mountain, during its course it is feed by 141 tributaries end eventually it ends into the Adriatic Sea with a delta of about 380 km^2. Moreover, the Po basin is the largest italian catchment with a drainage area of approximately 71.000 km^2, which represents a quarter of the entire national area. The average discharge at the gauging station of Pontelagoscuro, located just upstream of the river delta, is about 1500 $\mathrm{m}^3\mathrm{s}^{-1}$, for an average annual water volume of 47.3·10^9 m^3.

The Civil Protection of Emilia Romagna Region has implemented a flood forecast system based on the hydraulic model PAB (Todini and Bossi, 1986), which makes possible to obtain forecasts with a lead time of 72 hours at the Pontelagoscuro station.

Differently than in the case of the Baron Fork River application, in this case it was not necessary to calibrate the forecast model because it is already operative in a flood forecasting system. This makes interesting the Po river application; since the uncertainty due to the model calibration and to the meteorological prediction is the actual one with which the decision makers have to deal and on which basis they have to take an operational decision about the emergency management.

Taking advantage of the availability of these data, this application objective was to verify the behavior of the MCP in assessing the *probability of*
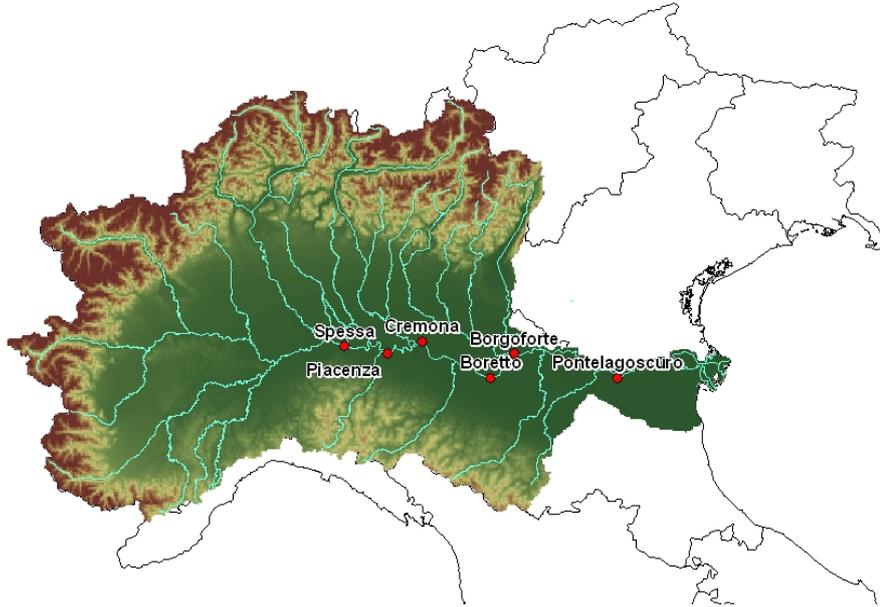
Figure 4.50: Digital Elevation Model of the Po basin.

exceeding a maximum river stage within a time horizon and the *exceeding time probability*, as described in Section 3.1.3.

4.2.2   Available Data

The Civil Protection of Emilia Romagna Region provided a long series of observed discharges corresponding to 6 measurement stations (Ponte Spessa, Cremona, Piacenza, Borgoforte, Boretto and Pontelagoscuro) and the hourly forecasts of the PAB model within a time horizon of 36 hours. The available hourly data includes the period from May 2000 to January 2009. Four years have been used to calibrate the MCP, whilst five years were used for the validation process. The forecasted data have been extracted considering a 3 hours time step, therefore for each hourly time step, within the 36-time horizon, 12 forecasts were taken into account.

The entire period of available data includes nine significant flood events; in particular, the two main events occured in October 2000 and November

2002 belong to the calibration data set, while the validation period includes minor events, such as those occured in November 2004, June, November and December 2008 (Figure 4.51).
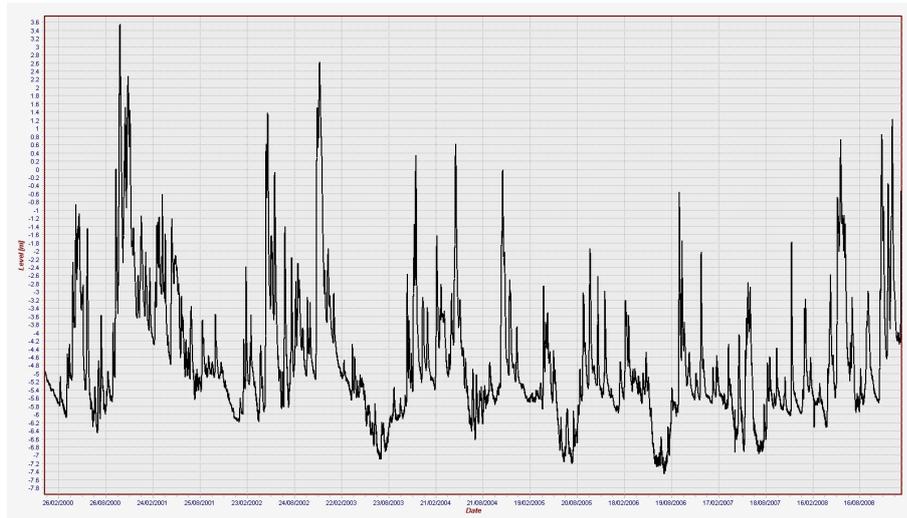
Figure 4.51: Entire series of observed data provided by the Civil Protection of Emilia Romiagna Region.

In addition, for each station the Civil Protection identified three alarm levels, which are used in the operational flood alarm process; in this application the lowest of these levels has been regarded as the maximum river stage for which the probability of exceedance has been analyzed.

4.2.3 MCP Application

To pursue the intent of the application, the MCP has been applied in four phases, of which the first three are the same described for the multi-model approach, whilst the fourth one is specific for this approach.

1. The first phase consists in the separate processing of each model prediction for each forecast lead time; for each forecasted time series the MCP bivariate approach has been applied using the TNDs. For each

model, the result of this phase is a series of forecasted expected values for each forecast lead time. More exactly, for each horizon time and each available model, the expected values of the predictive uncertainty conditioned to the model forecast are computed.

2. In this phase the series of expected values are processed in order to combine the different models predictions. However, for the present application only one model is available, therefore this phase must be skipped. Nevertheless, in case more models were available, this phase result would have been the expected values of the predictive uncertainty conditioned to the forecast of all the models for the same horizon time.

3. The third phase is the separate processing of each one of the series obtained in the previous phase, with the aim of computing the correct predictive uncertainty distribution.

4. The last phase is the most important for this application; the series obtained in the third phase (one for each forecast lead time) are combined together, in order to compute the multivariate predictive uncertainty described in Section 3.1.3. Following the procedure described in that section, the *cumulative probability* to exceed the alarm level is computed for each horizon time. Afterwards, the exceeding time probability is obtained as the discrete derivative of the cumulative probability to exceed the alarm level.

Among the 6 stations where observed data were available, 3 stations have been chosen in order to analyze the application of the Multi-Temporal version of the MCP: Ponte Spessa, Borgoforte and Pontelagoscuro. Ponte Spessa is the first station in the Emilia Romagna stretch of the Po river. The forecasting system uses the observed data some kilometers upstream this station, so the forecast lead time for which the forecasts can be considered reliable is not greater than approximately 24 hours. Pontelagoscuro is the last station of the forecasting system, it is close to the end of the river,

some kilometers far from the Adriatic Sea; in this station the forecasts can be considered reliable several hours in advance, hence in this application a forecast time horizon of 36 hours has been used. Borgoforte is located in the middle between the two previous stations, here the forecast lead time is 36 hours also in this case.

In the next sections the obtained results will be shown, firstly the probability to exceed the alarm levels within a horizon time will be analyzed and then the probability of the exceeding time.

Probability to exceed the alarm level within a time horizon

The probability to exceed the alarm level within a time horizon is the key information given by the Multi-Temporal approach. The knowledge of such probability is fundamental not just in order to quantify the hazard concerning a future flood event, but also because it allows combining the information given by all the forecasts during the whole forecast lead time, and so to almost eliminate, or significantly reduce the time bias of the prediction. In fact, if a forecast reproduces the peak flow systematically late or in advance, the Multi-Temporal approach can identify this error and help in reducing it.

This is the case of Ponte Spessa station, where the deterministic model prediction is often delayed with respect to the observed data. This is due to the forecasting system structure, which is based on the propagation of the flow observed few kilometers upstream Ponte Spessa. Since the forecast quality tends to rapidly decrease as the horizon time increases and because the time of flow propagation until Ponte Spessa is approximately 12 hours, it has been observed that any horizon time greater than 15 hours leads to a delay in the forecasts (Figure 4.52).

Figures 4.53 and 4.54, which concern the validation events, show the benefit obtainable using the Multi-Temporal MCP on this station. In these figures the lower panel represents the model prediction (blue line) respectively considering 12 and 18 hours forecast lead time. In the middle panel,
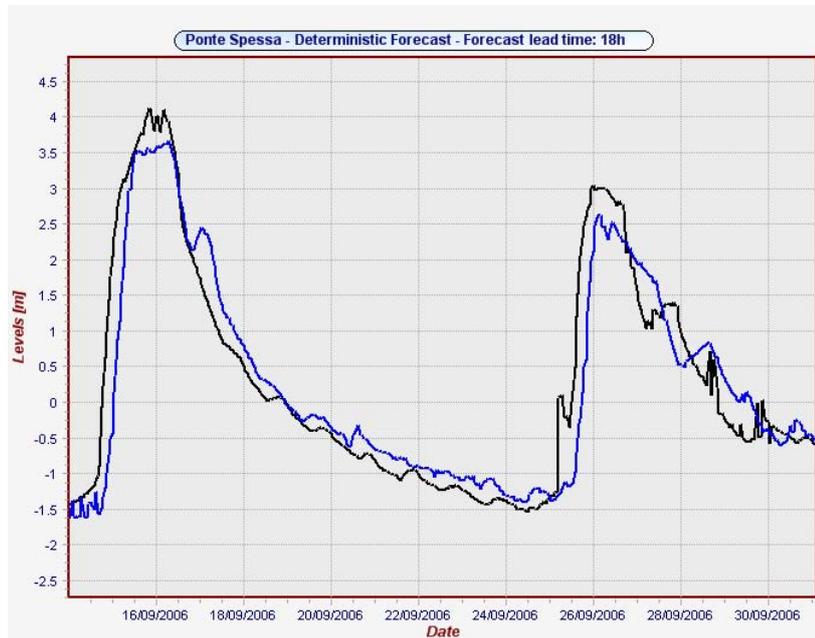
Figure 4.52: Delayed prediction from the deterministic model at Ponte Spessa. The forecast lead time is 18 hours.

the probability to exceed the level of 3.5 meters within the total horizon time is depicted, the light red line is the MCP response, the black line is the observed binary value and the blue line the deterministic forecast.

In Figure 4.53 the deterministic model makes a good prediction and the binary deterministic prediction perfectly matches the actual one. The MCP reflects this behavior and it reaches an exceeding probability of 80% when the level exceeding actually occurs, starting few hours before to predict increasing probability values. The MCP response has been depicted also in a more effective and clearer way: in the upper panel of this figure the probability values computed by the MCP are grouped in three classes according to different probability levels: the green group includes probabilities less than 25%, the red group includes the probability values greater than 75% and finally the yellow group contains the values within these two thresholds. These thresholds can be chosen on the basis of the vulnerability and value of the area, as mentioned in Section 2.4.
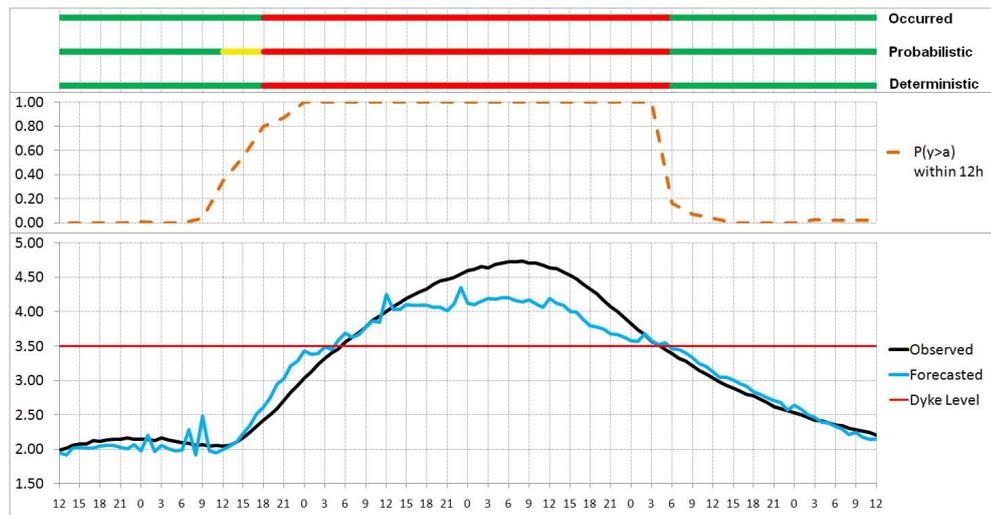
Figure 4.53: Comparison between deterministic and probabilistic predictions at Ponte Spessa. The forecast lead time is 12 hours.

Figure 4.54 represents an event predicted 18 hours in advance. This 18-hours horizon time causes often the deterministic model forecast to be delayed. In fact, in the depicted event the model predicts the level exceeding several hours later than the actual occurrence. The MCP can correct this error assigning a high exceeding probability to a certain number of hours preceding the event. In particular, it assigns a probability greater than 80% to the 15 hours before the model prediction as well as to the time interval in which the exceeding occurs. The benefit of using the Multi-Temporal approach is clearly visible in the upper panel of this figure, where the probabilistic response provides the highest alarm level right on time with the observed behavior.

In the other two stations, Borgoforte e Pontelagoscuro, the deterministic forecasts are reliable up to several hours in advance, rarely the predictions are delayed and the actual behavior is well reproduced.

Concerning Borgoforte, Figures 4.55 and 4.56 depict two validation events in which the deterministic model predicts the level exceeding few hours in advance. Also in this case the MCP is able to recognize and correct this error.
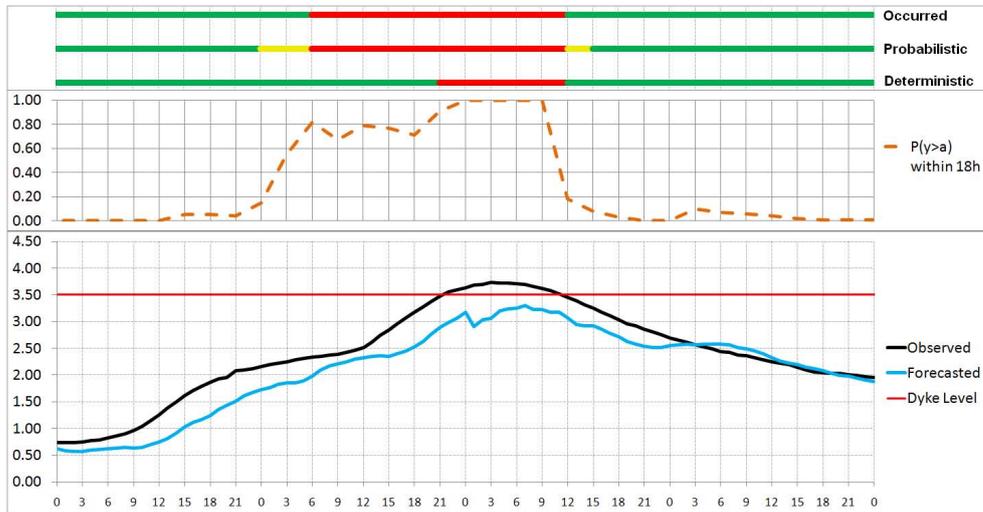
Figure 4.54: Ponte Spessa, delayed deterministic prediction corrected by the MCP. The forecast lead time is 18 hours.
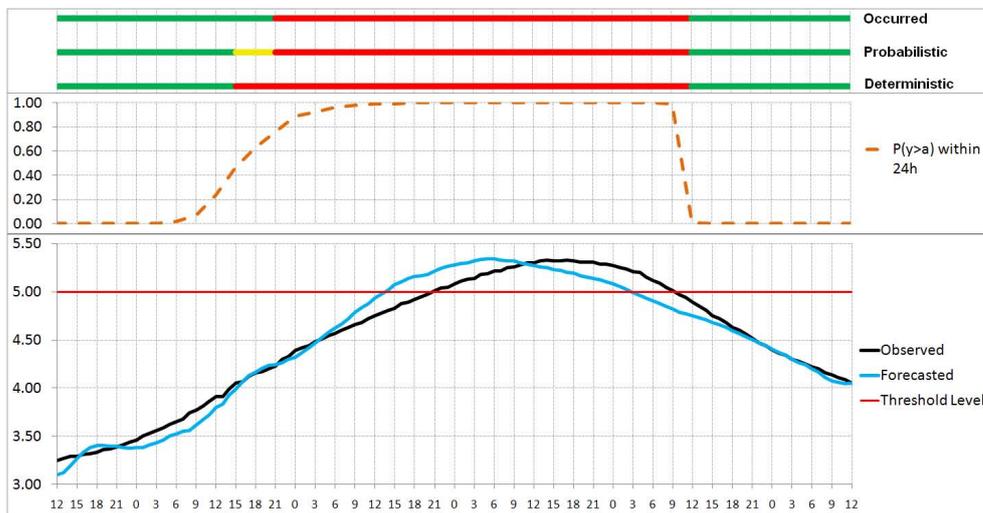
Figure 4.55: Comparison between deterministic and probabilistic predictions at Borgoforte. The forecast lead time is 24 hours.
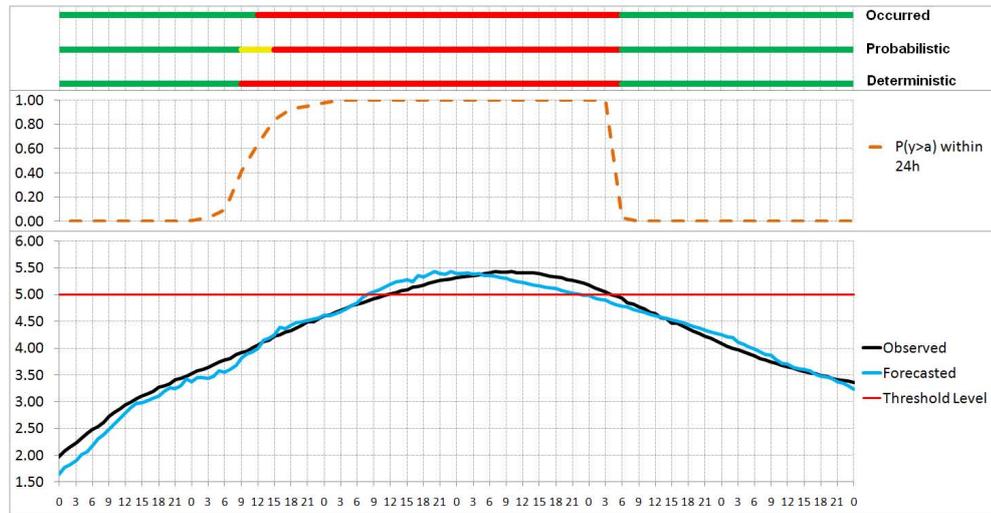
Figure 4.56: Comparison between deterministic and probabilistic predictions at Borgoforte. The forecast lead time is 24 hours.

At Pontelagoscuro, the deterministic forecast is very good and the exceeding time is often well reproduced. Figures 4.57 and 4.58 represent two validation events for this station. In the first case the MCP reflects the optimal behavior of the deterministic model, which is on time with the actual observation (Figure 4.57). The second event represents a situation in which the alarm level is not exceeded, but the maximum level registered is very close to this threshold; the MCP computes exceeding probability values always lower than 45%, issuing only a yellow response, correctly reflecting the actual occurrence (Figure 4.58).
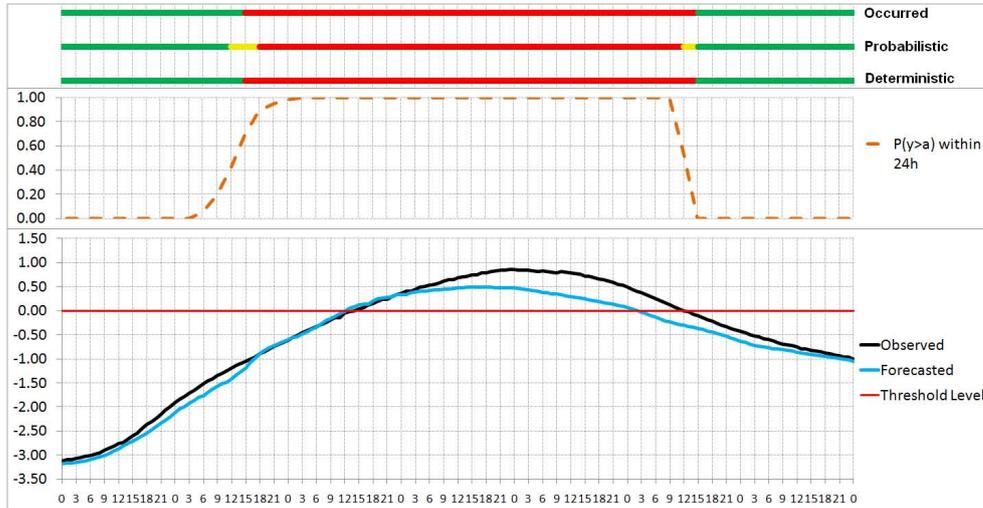
Figure 4.57: Comparison between deterministic and probabilistic predictions at Pontelagoscuro. The forecast lead time is 24 hours.
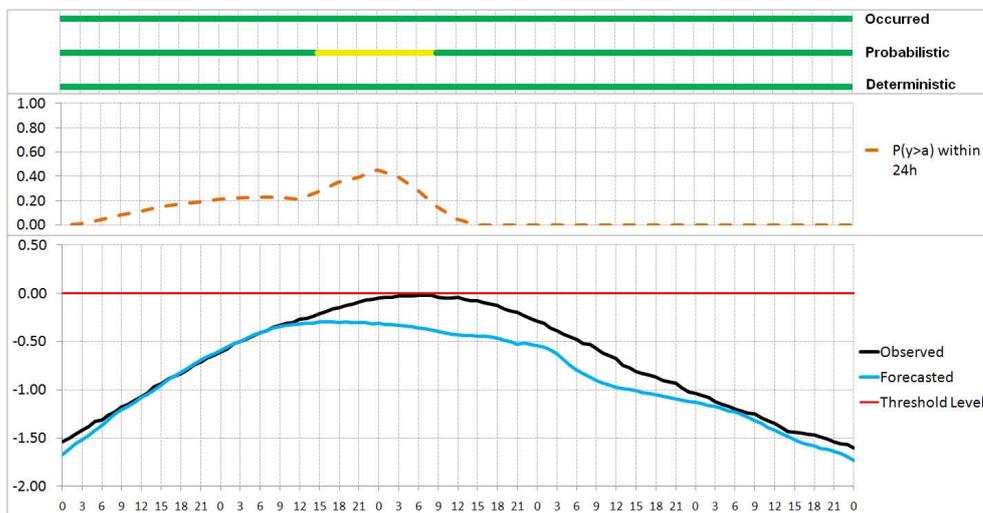
Figure 4.58: Comparison between deterministic and probabilistic predictions at Pontelagoscuro. The forecast lead time is 24 hours.

Verification of the exceeding probability evaluation correctness

As mentioned for the Multi-Model application to the Baron Fork River, the correctness of the exceeding probability value computed by the MCP can be verified comparing the observed exceeding event frequency to the probability value given by the MCP. For instance, considering all the cases in which the MCP provides an exceeding probability of approximately 0.2, to be correct, the correspondent observed occurrences should be approximately the 20%. Unfortunately, it is not possible to make this verification for each probability value, but it is necessary to make discrete the domain. In the present application the discretization has been done with a 0.05 probability interval.

Figures from 4.59 and 4.64 represent the verification made with the calibration and validation data for the three considered stations and a time horizon equal to 18 hours for Ponte Spessa, while a 24-hour time horizon was considered for Borgoforte and Potelagoscuro.
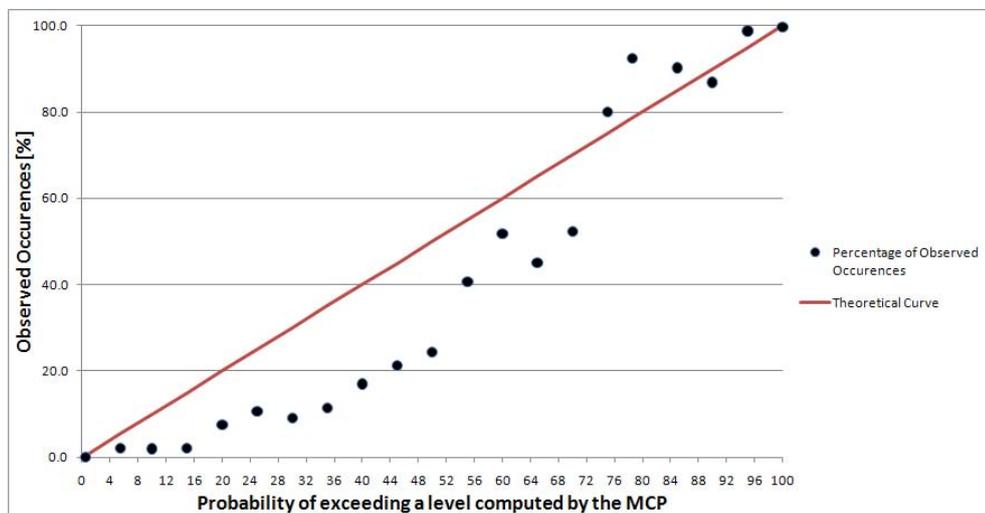
Figure 4.59: Frequency of actual exceedances vs the probability estimated using the MCP with the calibration data at Ponte Spessa. The red line represents the perfect behavior.

At Ponte Spessa (Figures 4.59 and 4.60) the verification results highlight
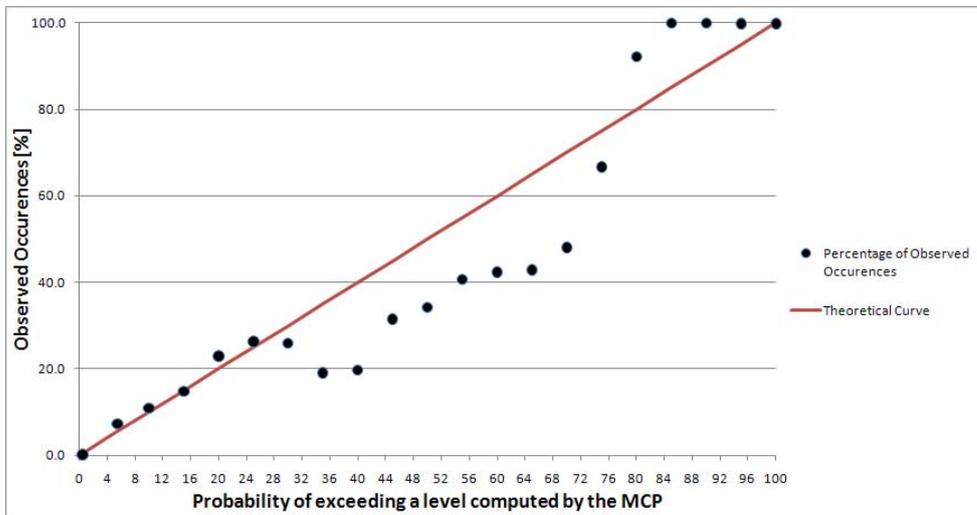
Figure 4.60: Frequency of actual exceedances vs the probability estimated using the MCP with the validation data at Ponte Spessa. The red line represents the perfect behavior.

a positive bias, especially for the lower values of the calibration period, while considering the validation data, the MCP tends to subestimate the actual probability when dealing with high probability values. The positive bias observed during the calibration period can be explained considering that inevitably approximations are used in the joint distribution representation of observed and predicted data, which is higher when the correlation between these two variables is lower. This behavior is much less evident at Borgoforte or Pontelagoscuro than Ponte Spessa and it is due to the fact that in these stations the correlation between actual and predicted levels is higher, especially for forecast lead time greater than 15 hours.

The subestimation of the high probability values obtained with the validation data at Ponte Spessa may be probably due to the upper distribution tail that is the other main assumption into the MCP structure.

Figures 4.63 and 4.64, concerning the station of Pontelagoscuro, highlight the same problem for the u tail, but confirm the general good behavior shown by all the three stations in representing the exceeding probability in
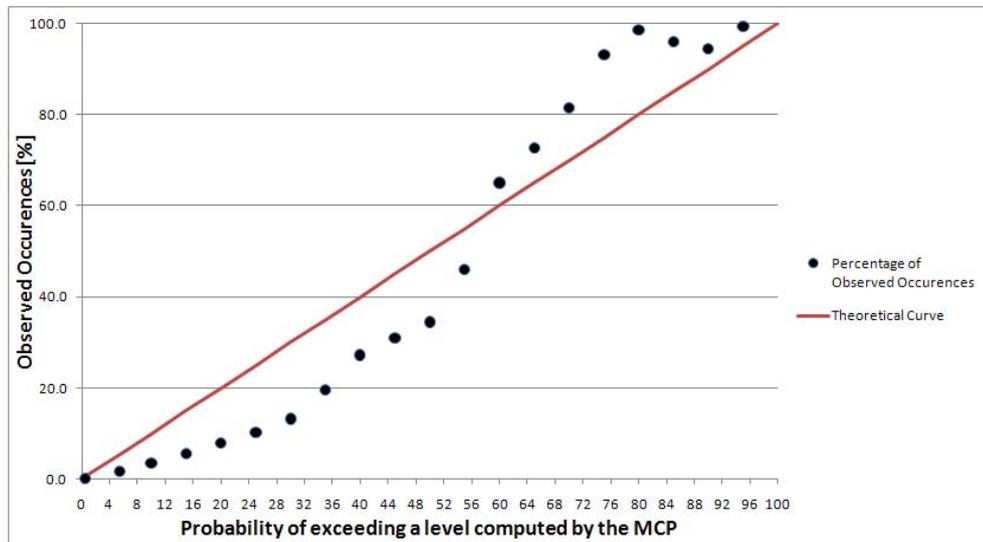
Figure 4.61: Frequency of actual exceedances vs the probability estimated using the MCP with the calibration data at Borgoforte. The red line represents the perfect behavior.
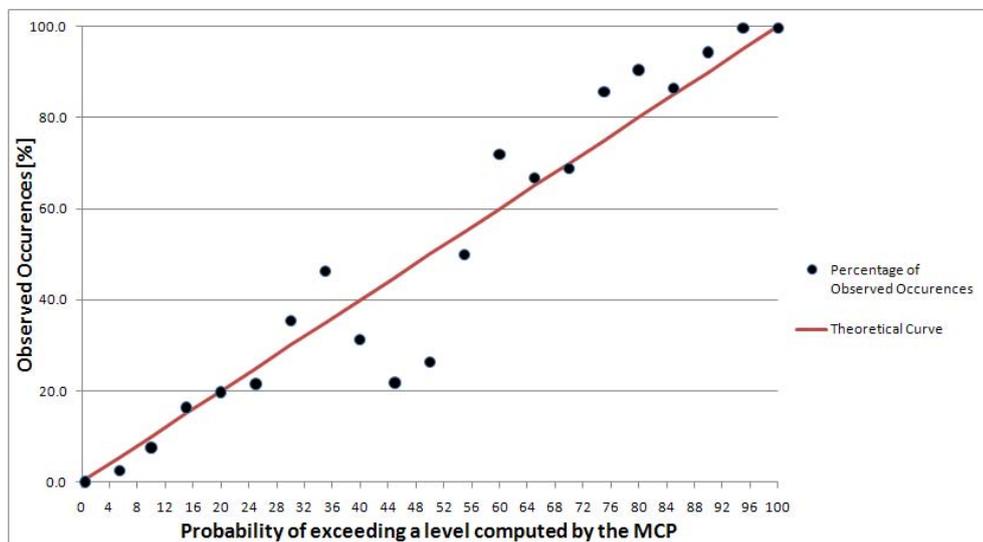
Figure 4.62: Frequency of actual exceedances vs the probability estimated using the MCP with the validation data at Borgoforte. The red line represents the perfect behavior.
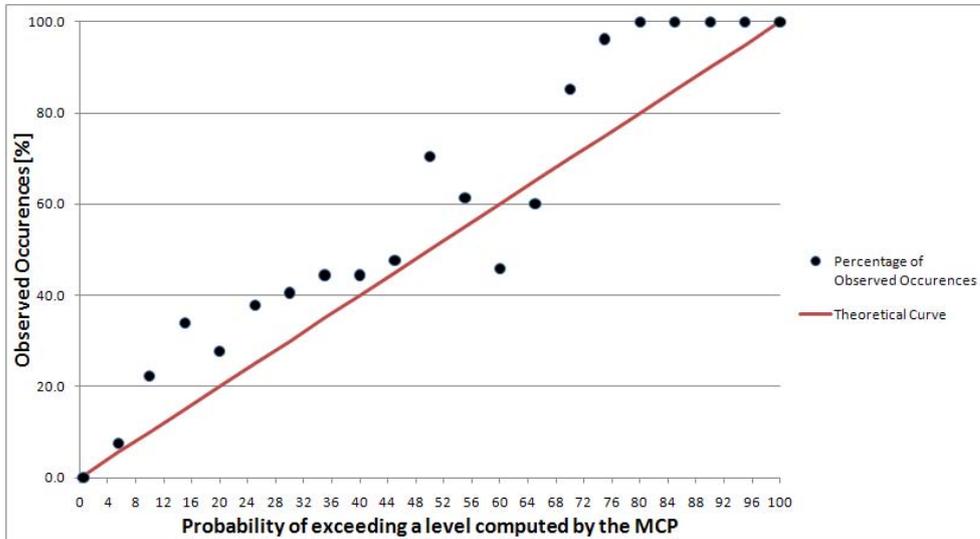
Figure 4.63: Frequency of actual exceedances vs the probability estimated using the MCP with the calibration data at Pontelagoscuro. The red line represents the perfect behavior.
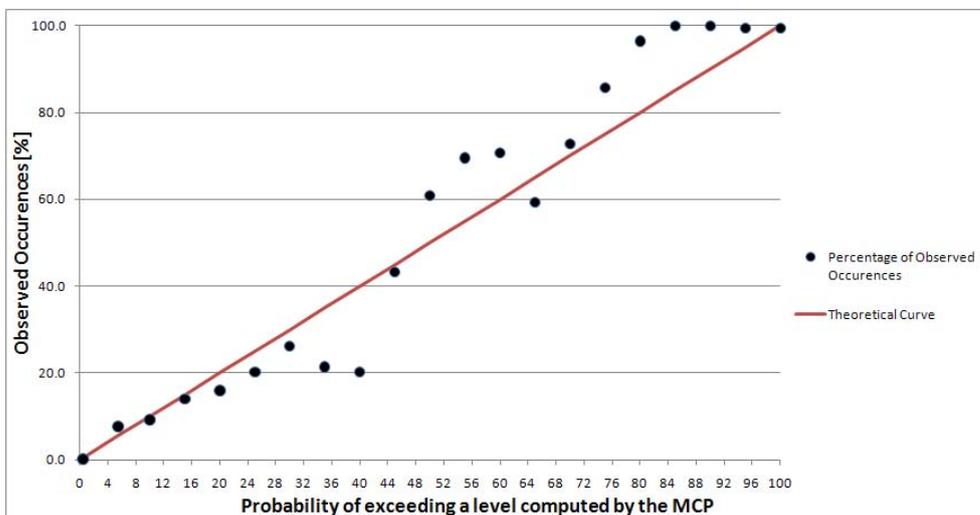
Figure 4.64: Frequency of actual exceedances vs the probability estimated using the MCP with the validation data at Pontelagoscuro. The red line represents the perfect behavior.

its whole.In fact, despite these few points, it can be certainly said that there is a relatively good agreement between simulated probabilities and actual occurrences and the MCP can be defined as a reliable tool to estimate the exceeding probability.

Exceeding Time Probability

The Multi-Temporal approach provides also essential information concerning the exceeding time probability, which represents the discrete probability of the instant when the exceeding event occurs. Such probability is computed as the discrete derivative of the cumulative probability of exceeding the alarm level (described in the previous section) and it is represented in the following figures applying a linear interpolation between the obtained discrete values.
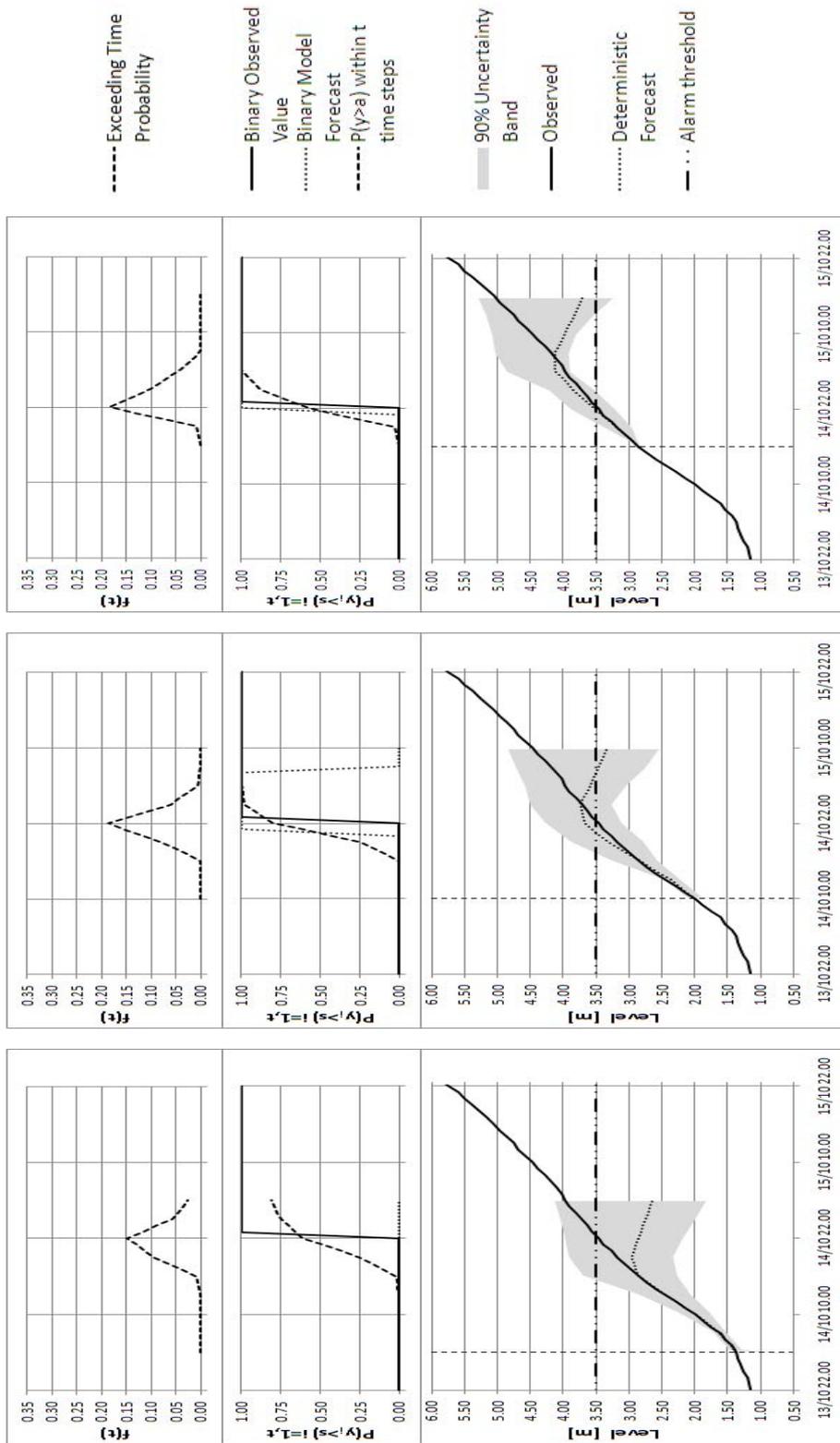
Figure 4.65: Calibration event occured in October 2000 at Ponte Spessa station

Figures from 4.65 to 4.70 depict for each station two events: one occurred during the calibration period and the other one during the validation period. Each figure is composed of 3 panels: in the lowest panel the predictive uncertainty is represented by the grey area; in the middle one the observed binary response of exceeding the threshold is compared with the modeled one as well as the cumulative probability computed by the MCP; finally, the upper panel represents the probability of the exceeding time. The forecast horizon taken into account is 24 hours and the alarm level (e.g. the river stage used for computing the exceeding probability and the exceeding time probability) is, for each station, the lowest between those provided by the Civil Protection of Emilia Romagna.
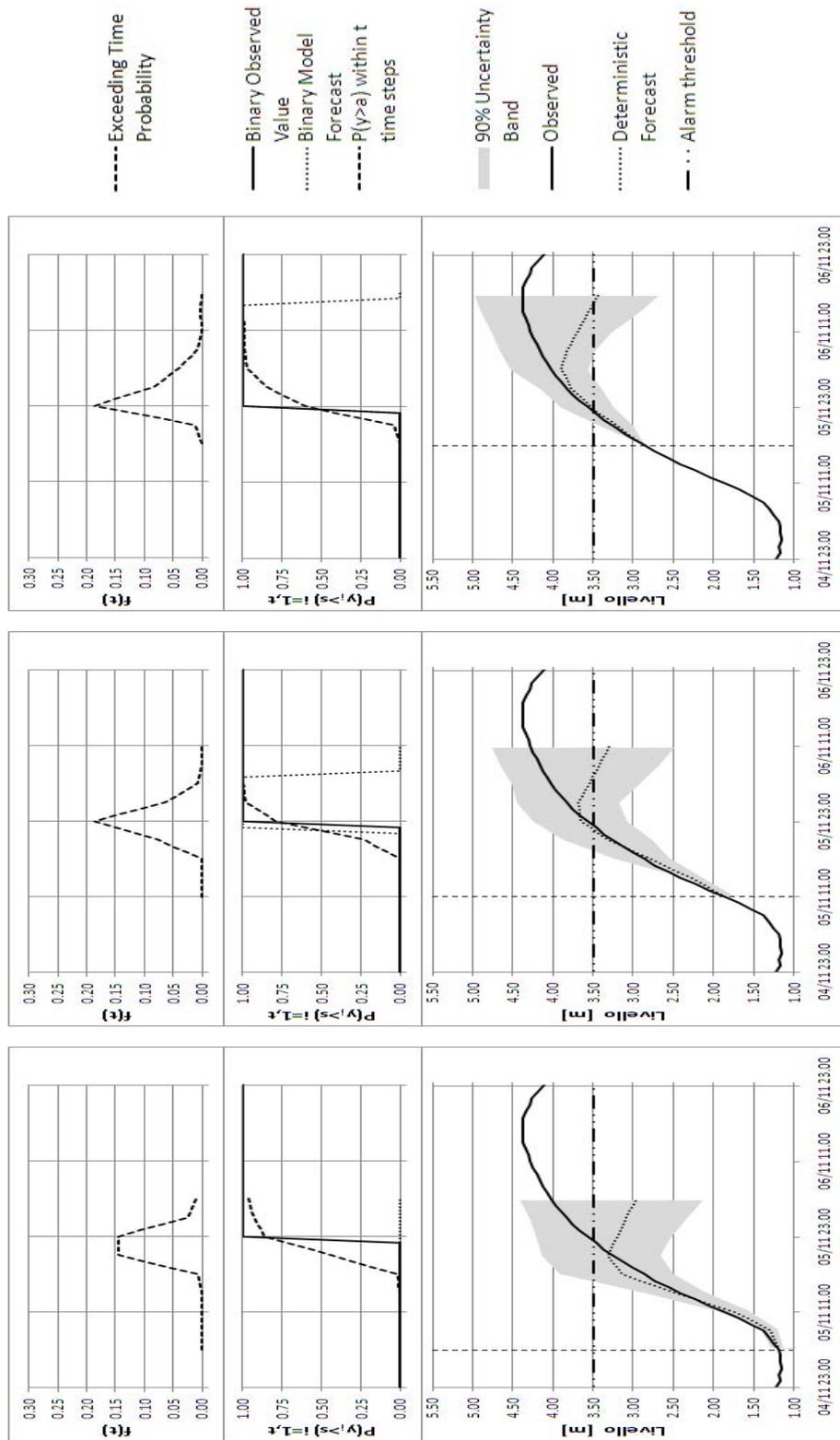
Figure 4.66: Validation event occurred in November 2008 at Ponte Spessa station

Figure 4.65 represents the event occurred at Ponte Spessa in October 2000, starting on 14th October at 4.00, while Figure 4.66 represents an event occurred also occurred at Ponte Spessa during the validation period, specifically on 5th November 2008. In these figures three consecutive forecasts are depicted, with a time interval of 6 hours between them.

Concerning the application to Ponte Spessa station, it is interesting to note that the MCP computes a high probability to exceed the threshold within the horizon forecast time even if the deterministic prediction does not reach the alert level. Moreover, the exceeding time probability well represents the observed behavior and the modal value is always really close to the actual instant when the exceedance occurs. As expected, the comparison between the three consecutive forecasts depicted for each event shows that the uncertainty of the exceeding time decreases when the forecast is updated. In fact, the exceeding time probability distribution obtained linearly interpolating the discrete values becomes tighter when the starting forecast time increases and the modal value tends to the actual exceeding time value.
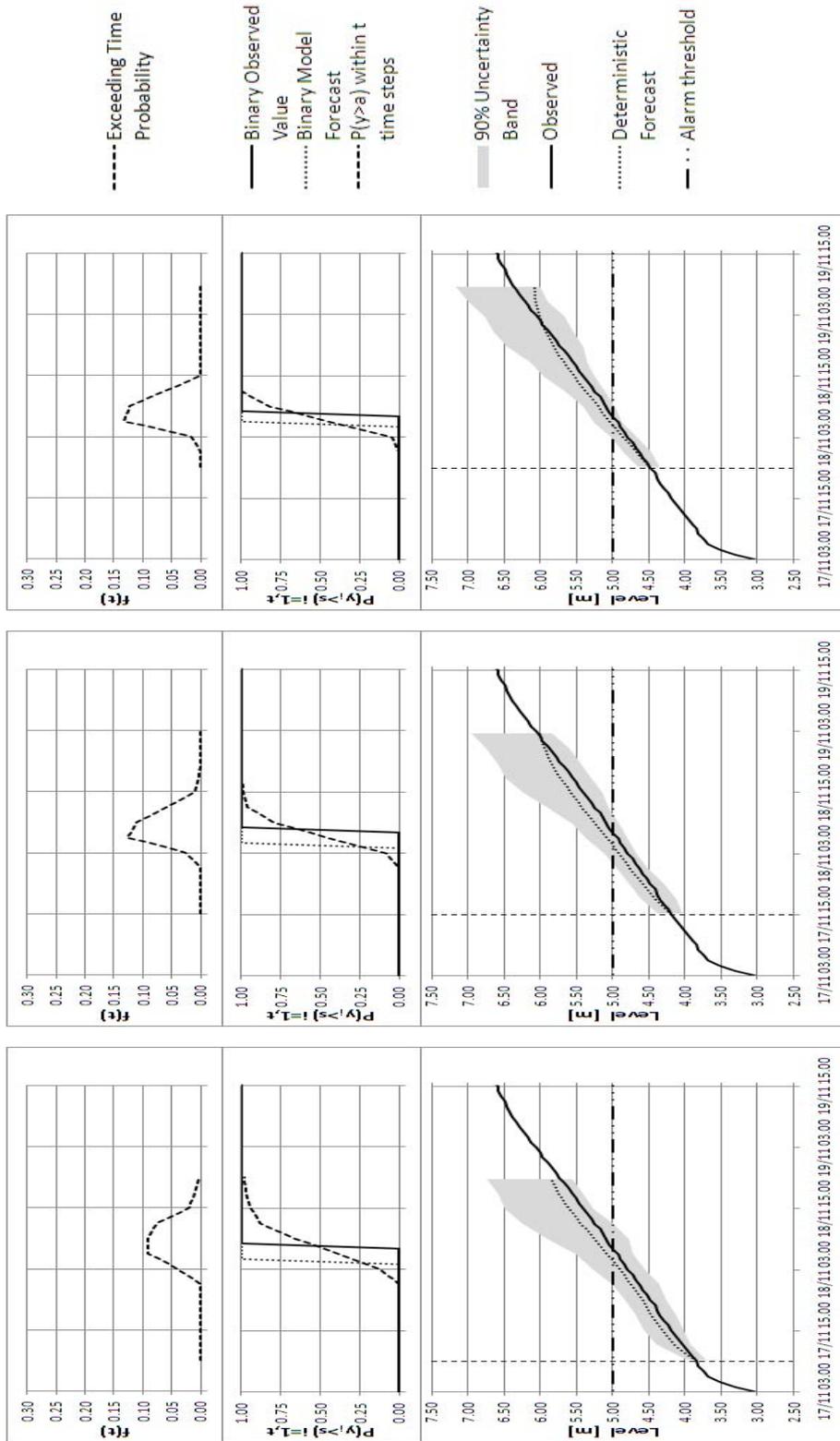
Figure 4.67: Calibration event occurred in November 2002 at Borgoforte station
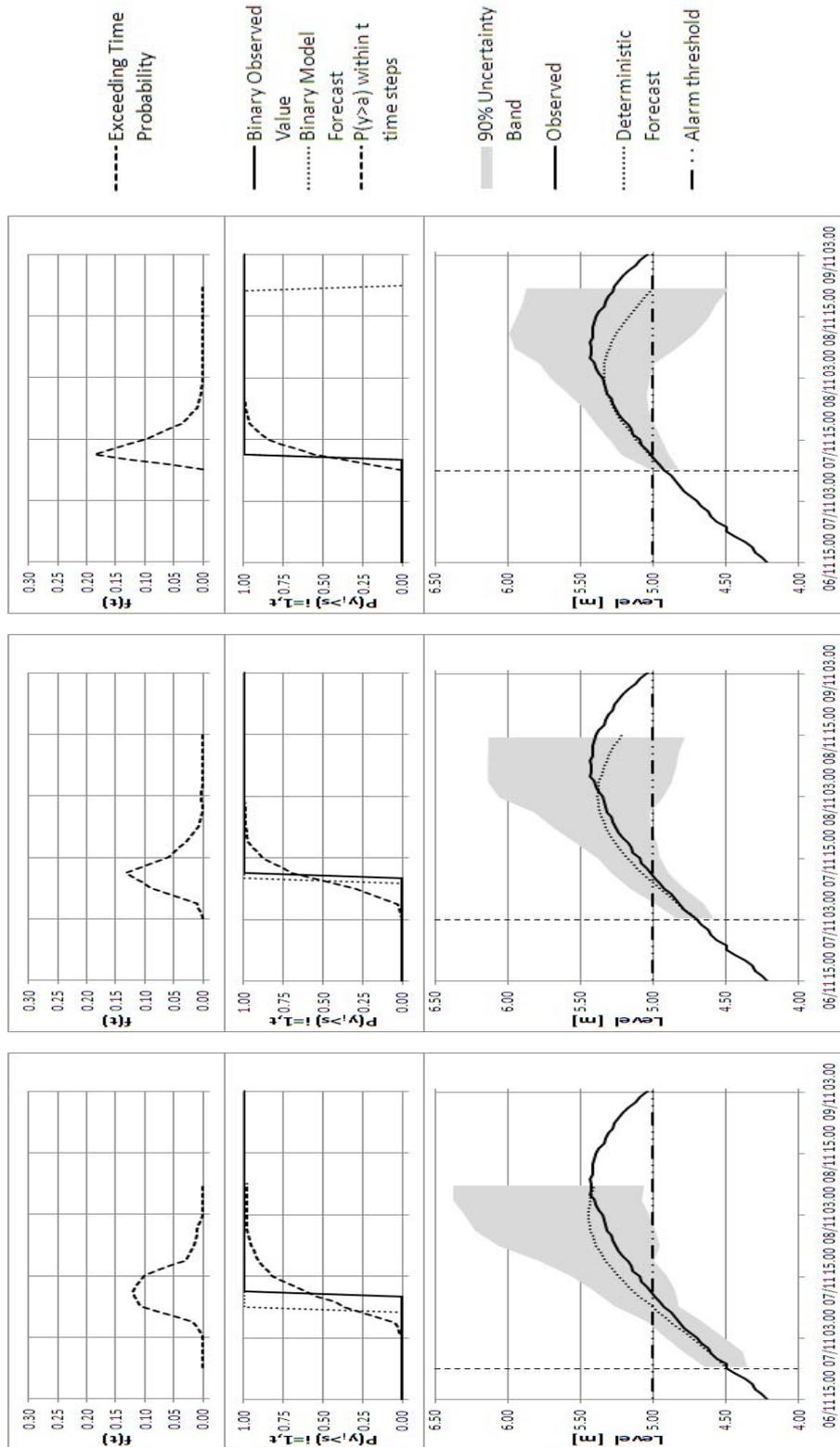
Figure 4.68: Validation event occurred in November 2008 at Borgoforte station

At Borgoforte and Pontelagoscuro stations the considered events (Figures 4.67 and 4.69 concern calibration events while Figures 4.68 and 4.70 show validation ones) point out the same behavior observed at Ponte Spessa. Also in these cases the MCP application allows the time of flood occurrence to be well estimated. In fact, the exceeding time probability often reaches the highest value closer to the actual exceeding instant than the deterministic forecast, correcting the fact that the exceeding event is predicted with certain delay or few hours in advance. This behavior is visible for both calibration and validation events. Finally, also in these stations the forecast updating leads to the reduction of the prediction uncertainty, as it is expected to do.
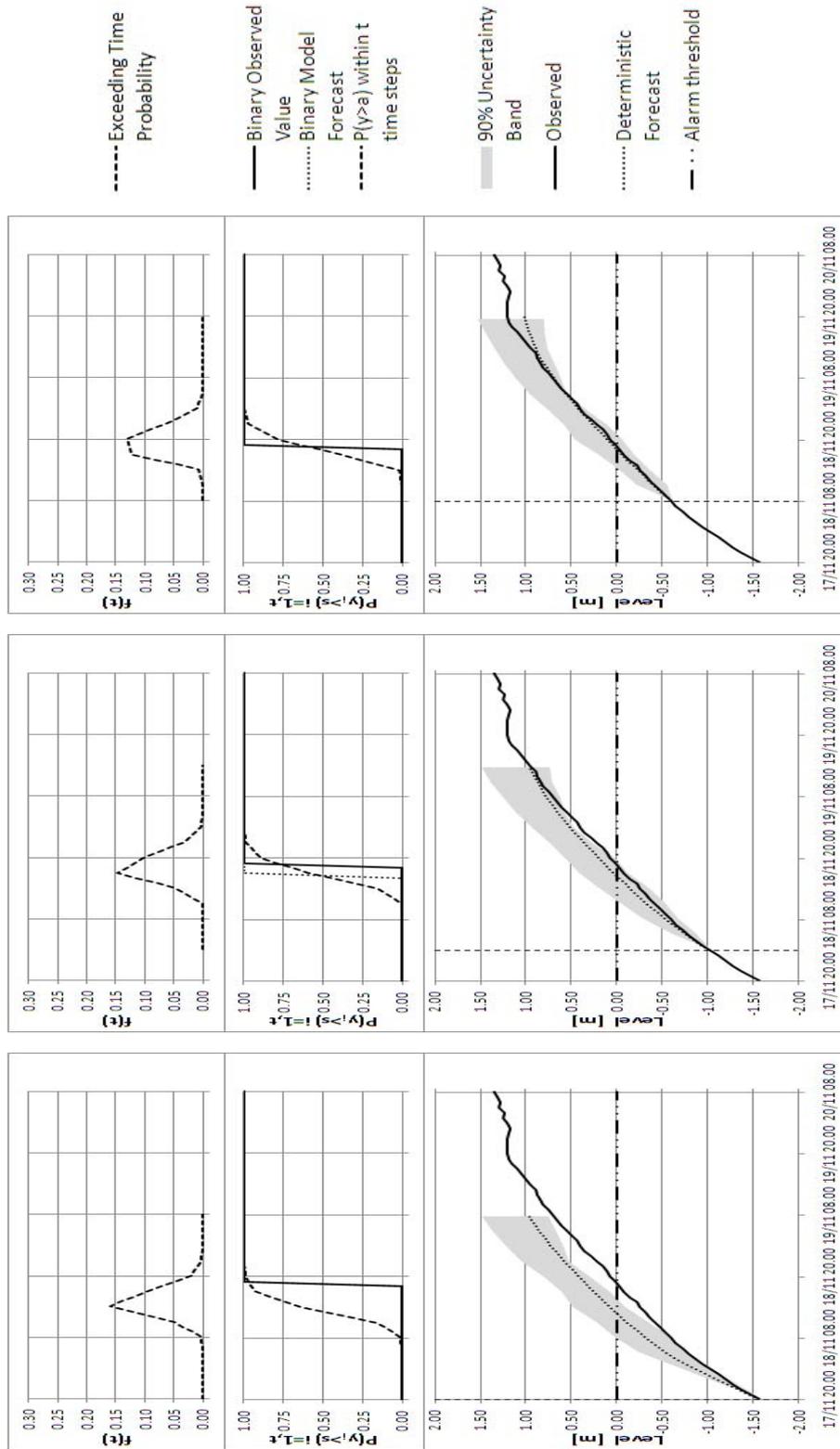
Figure 4.69: Calibration event occurred in November 2002 at Pontelagoscuro station
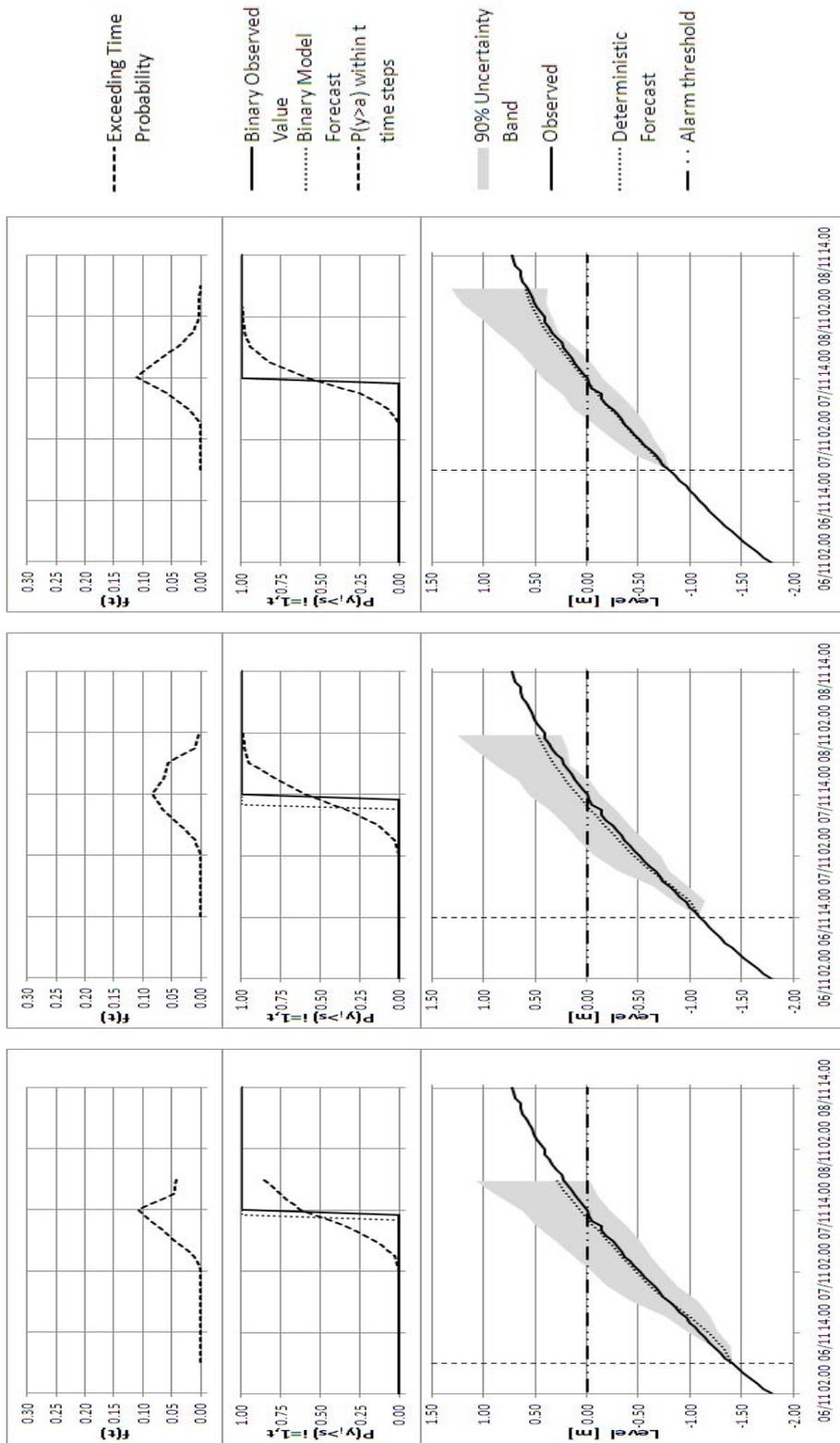
Figure 4.70: Validation event occurred in November 2008 at Pontelagoscuro station

Chapter 5

Conclusions and Future Research Lines

5.1 Discussions and Conclusions

The analysis presented in this thesis allowed to obtain significant results as well as fulfill the objectives planned at the outset. First of all, predictive uncertainty has been defined, in agreement with Krzysztofowicz (1999) and Todini (2008), as the probability of the occurrence of a real future value of the predictand (level, discharge, volume, ...) conditional upon all the knowledge and information available up to the present, which is usually embodied in observations and in deterministic forecasting models and can be acquired via a learning inferential process. It was also highlighted that the final aim of forecasting systems is to reduce predictive uncertainty through a representation of the physical processes that condition the hydrological phenomena.

This research work was mainly focused on the development of the Model Conditional Processor development for assessing predictive uncertainty. Two applications, the first one to the Baron Fork River (OK, USA) and the second one to the Po River (Italy) allowed to draw several important conclusions, which are summarized below.

- The predictive uncertainty assessment starts with the identification of the marginal distributions of the observed and predicted data as well as their joint distribution. Such marginal distributions are often unknown in the untransformed observation space, and moreover it is extremely difficult to make hypotheses on the shape of their joint distribution. Several works in the literature (Krzysztofowicz, 1999; Montanari and Brath, 2004; Todini, 2008) suggested to use to use a non-parametric approach based on order statistics, namely to use the Weibull Plotting Position as an estimate of the probability of an ordered vector. Accordingly, a nonlinear transformation, the Normal Quantile Transform, is used to move from the original observation space to the Normal one, where by construction the marginal distributions assume a Standard Normal shape and the joint distribution can be reasonably approximated by a Multivariate Normal distribution. Nonetheless, this approach has some disadvantages. First of all, it implies to identify additional models to adjust the quantiles outside the range of the historical available data. The proposed technique is very sensitive to the shape and to the parameters of these models and some precautions in the choice of the subset of observations used for calibrating the tails data must be taken. They must contain a large variety of cases, as required by any Bayesian approach, and in order to reduce the uncertainty on the marginal distribution tails the calibration data must include the highest number of extreme cases.

- The assumption of a Normal Multivariate joint distribution in the transformed space implies unavoidable approximations and it does not account for the heteroscedasticity error. In order to reduce it a non-linear regression model could be used. In this thesis a piecewise linear approach has been preferred to a fully non-linear model. The piecewise linear approach allows for the use of Truncated Multivariate Normal joint distributions. This technique can be easily developed and applied and good results have been obtained for both the study cases where it

has been used.

- The comparison between the use of the Truncated Normal Distributions technique (TNDs) and the Quantile Regression (QR) showed that the first technique, requires a lower computational cost and, in the application proposed in this work, leads to better results. This is also true when the QR technique is applied splitting the Normal space in two parts. Nonetheless, the alternative use of the two techniques should be approached on the basis of a detailed error analysis. After this analysis the methodology that better adapts to the specific case should be used. Anyway, the lower computational cost and the higher flexibility of TNDs make this technique adaptable to a wider number of cases.

- Multiple predictions originated by several models, as discussed in the introduction, is of difficult understanding and interpretation by the decision makers, particularly when these predictions are in contrast one to another. The application of the MCP on Baron Fork has shown that the proposed technique allows the correct combination of different forecasts into a unique probability of the event, which is of much easier interpretation and use in the decision making process. In fact, the combination of the three models predictions, obtained by assigning different weights to each model according to the Bayesian theory, allows the forecast quality to be improved as it is shown by the evaluation indexes in Figures 4.35 and 4.36. In particular, the two distributed model taken into account in this work (TETIS and TOPKAPI models) have very similar structures and this leads to just a marginal gain in terms of forecast improvement, which is shown by the standard deviation of the errors and the Nash-Sutcliffe coefficient (Figures 4.35 and 4.36). Instead, the combination of one physically based model with the data driven model leads to greater improvements in forecast and, in particular, the combination of all the three models gives the best values of the analyzed indexes in Figures 4.35 and 4.36. These results show that

the combination of models with different nature allows the probabilistic forecast to improve the deterministic forecast of each model.

- The validity of the methodology on which is based the Model Conditional Processor is also verified by the analysis of the probability to exceed a fixed level (such as the level of the embankment, which corresponds to the probability of flooding). Figure 4.49 shows that the probability values computed by the processor well represent the actual realizations concerning the application to the Baron Fork River. The same is shown by Figures from 4.59 to 4.64 about the application to the Po River, apart from some minor situations in which the probability is slightly underestimated or overestimated. This verification not only demonstrates the validity of the method, but also it makes possible to assess with good accuracy the actual hazard of the event occurrences, on which basis it is possible to make a proper risk analysis that is an essential element of the decision-making process.

- The analysis of the application to the Po river, where the Multi-Temporal approach has been tested, highlights the ability of the MCP to link the hazard of a predicted event to its time of occurrence. In an operational phase, the processor provides clear and easily interpretable estimates of the probability that an event will happen within a given time horizon. This information is essential to assess the available time to actuate intervention procedures and properly organizing the available resources.

- The results obtained at the Ponte Spessa station point out that the Multi-Temporal approach allows taking into account the possible presence of a systematic model prediction time displacement. In fact, at Ponte Spessa, due to structural features of the forecasting system, the predictions are often delayed when the forecast lead time is greater than 15 hours. As shown in Figures 4.54, 4.65 and 4.66, the MCP was able to recognize and drastically reduce this delay and to assess with good accuracy the exceeding time probability, systematically improving the

deterministic forecasts.

Other conclusions concerning the deterministic models, in particular the TOPKAPI and Artificial Neural Networks ones, were pointed out. Their application, although uniquely addressed to provide data for testing the MCP, involved a significant part of the work and suggested the following considerations.

- The TOPKAPI model application to the Baron Fork River showed a good predictive ability and provided reliable simulations of flood events as well as of the general catchment behavior. However, deficiencies were highlighted concerning the TOPKAPI capacity to differentiate the 'wetting up period' from the 'wet period'. The inclusion of just one soil layer in the TOPKAPI conceptual scheme produces a single subsurface flow response. In fact, the more surface layer is generally characterized by higher permeability and faster responses to the rain input, while the deeper layer has lower conductivity and lower and slower responses. Therefore, the consequences of the simplified conceptual scheme of the TOPKAPI had to be solved through model calibration, forcing the related parameters to assume values that represent a compromise solution between the two soil layers mentioned above. A soil representation composed of two different layers vertically interacting between them further improves the model predictive ability, as demonstrated by (Coccia et al., 2010).

- Artificial Neural Networks have been used as a data driven model, in order to complement the distributed models' forecast. Following Toh and Brath (2008) and Pujol (2009), the model has been developed combining a Self Organizing Map network (SOM) to three Multi Layer Perceptron ones (MLP). The former (SOM) has the task of classifying the data into groups that are expected to represent different catchment conditions. This will help the latter (MLP) to produce better forecasts. The SOM network automatic classification ability is highly dependent

on the calibration data selection and in this work it has to be pointed out that the network was not always able to differentiate wet conditions from dry ones. Nevertheless, the MLP networks were able to provide good forecasts, even if it showed high instability and some problems in reproducing the main flood events. The MLP predictive ability is closely related to the last available discharge (or level) observation and it rapidly decreases as the forecast lead time increases. Hence, models application to a small basin, such as the Baron Fork River, where process non-linearity is more evident, must be preceded by an accurate analysis of data and catchment features. The data selection in order to feed the SOM network is the crucial step to obtain reliable results, because the correct data assignment to the right group representing the catchment soil moisture state allows the process non-linearity to be well represented, which is essential to obtain reliable forecasts. Anyway, as pointed out by Pujol (2009), it is not advisable to use the Artificial Neural Networks without the support of a more robust forecast provided by a distributed or semi-distributed physically based model.

5.2 Future Research Lines

Nowadays, predictive uncertainty is an extremely relevant topic and it is still open to discussions. There are still wide research fields open for developments and improvements that may be undertaken to tackle several issues pointed out in this thesis. Three of them are here highlighted, because closely related to the present work.

- The tail models choice for the marginal distributions of the variables used in the MCP is one of the points that require to be deepened. Alternative approaches to the NQT should be studied using parametric probability distribution functions for the transformation into the Normal space.

- The choice of the truncation threshold in the TNDs is carried out in this work according to the criterion of maximizing the upper sample data correlation, given that the main objective is to better reproduce the probability of higher floods. The variation range for this threshold value must be limited in order to avoid the upper sample to include just few data, which could not be sufficient to significantly represent the joint distribution for higher values. Different methods of threshold selection must be evaluated, also including the possibility to differentiate them according to specific cases.

- The applications presented in this thesis show a rather good behavior of the MCP at estimating the elements associated with the predictive uncertainty (such as the flooding and the flooding time probabilities), but it would be useful to conduct further tests on different case studies and with other deterministic models.

Bibliography

Beven K.J.: Kinematic subsurface stormflow, Water Resour. Res., 17(5), 1419-1424, 1981.

Beven K.J.: On subsurface stormflow: prediction with simple kinematic theory for saturated and unsaturated flows, Water Resour. Res., 18(6), 1627-1633, 1982.

Blöschl G., and Sivapalan M.: Scale issues in hydrological modelling: a review, Hydrol. Proc., 9, 251-290, 1995.

Borah, D. K., Prasad, S. N., Alonso, C. V.: Kinematic wave routing incorporating shock fitting, Water Resources Research, 16(3), 529-541, 1980.

Cash, J.R., and Karp, A. H.: A variable order Runge-Kutta method for initial value problems with rapidly varying right hand sides, ACM Transactions on Mathematical Software, 16(3), 201-222, 1990.

Clapp R.B., and Hornberger G.M.: Empirical equations for some soil properties, Water Res. Res., 14, 601-604, 1978.

Coccia, G., and Todini E.: Application of a Bayesian processor for predictive uncertainty estimation in real time flood forecasting, BHS Third International Symposium, Role of hydrology in managing consequences of a changig global environment, Newcastle, 19th-23th July, 2010.

Coccia G., Mazzetti C., Ortiz E.A. and Todini E.: A different soil conceptualization for the TOPKAPI model application within the DMIP 2,

Abstract H21H-07 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec 2010.

Cunge J.A.: On the subject of flood propagation method (Muskingum method), Jour. Hydraulic Research, ASCE, 7(2), 205-230, 1969.

De Finetti, B.: Theory of Probability, Vol. 2.,Wiley, Chichester, UK, 1975.

De Groot, M. H.: Optimal Statistical Decision, McGraw-Hill, New York, NY, 1970.

De Marsily, G.: Quantitative hydrology: Groundwater Hydrology for Engineers, Geostatistic and stochastic approach in hydrogeology, Chapter 11, p286-329, Academic Press, California, 1986.

Dempster, A. P., Laird N. M., and Rubin D. B.: Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc., Series B, 39, 1-39, 1977.

Doorembos J., Pruitt W.O., Aboukhaled A., Damagnez J., Dastane N.G., van den Berg C., Rijtema P.E., Ashford O.M., and Frere M.; Guidelines for predicting crop water requirements, FAO Irrig. Drainage Pap., 24, 1984.

Duan Q., Sorooshian S. and Gupta, V.: Effective end efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res., 28, 4, 1015-1031, 1992

Edwards, A. L.: The Correlation Coefficient, Ch. 4 in An Introduction to Linear Regression and Correlation, W. H. Freeman, San Francisco, CA, 33-46, 1976.

Francés, F., Vélez, J. I., and J. J. Velez,: Split-parameter structure for the automatic calibration of distributed hydrological models, J. Hydrol., 332, 226-240, 2007

Henderson F.M., Wooding R.A.: Overland flow and groundwater flow from a steady rainfall of finite duration, J. Geophys. Res., 69(6), 1531-1540, 1964.

Hurley D.G., and Pantelis G.: Unsaturated and saturated flow through a thin porous layer on a hillslopes, Water Resour. Res., 21(6), 821-824, 1985.

Kelly, K. S., and Krzysztofowicz, R.: A bivariate meta-Gaussian density for use in hydrology, Stoch. Hydrol. Hydraul., 11, 17–31, 1997.

Kohonen T.: The self-organizing map. Proceedings of the IEEE, 78, 9, 1464-1480, doi: 10.1109/5.58325, 1990.

Koenker, R.: Quantile Regression, Econometric Society Monographs, Cambridge University Press, 2005.

Koussis A.D.: Comparison of Muskingum method difference schemes, Jour. Hydraulics Division, ASCE, 106(5), 925-929, 1980.

Koussis A.D.: Accuracy criteria in diffusion routing: a discussion, Jour. Hydraulic Engineering, ASCE, 109(5), 803-806, 1983.

Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, Water Resour. Res., 35, 2739-2750, 1999.

Krzysztofowicz, R.: Probabilistic flood forecast: exact and approximate predictive distributions, Research Paper RK0802, University of Virginia, September 2008.

Krzysztofowicz, R., and Kelly, K. S.: Hydrologic uncertainty processor for probabilistic river stage forecasting. Water Resour. Res., 36 (11), 3265-3277, 2000.

Liu, Z., and Todini, E.: Towards a comprehensive physically based rainfall-runoff model, Hydrol. Earth, Syst. Sc., 6 (5), 859–881, 2002.

Liu, Z., Martina, M.V.L. and Todini, E.: Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment, Hydrol. Earth Syst. Sci., 9, 347-364, 2005

Mantovan, P. and Todini, E: Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE Methodology, J. Hydrol., 330, 368381, 2006.

Mardia, K. V., Kent, J. T., and Bibby, J. M.: Multivariate Analysis. Probability and Mathematical Statistics. Academic Press, London, UK, 1979.

Matheron, G.: La théorie des variables regionaliséés et ses applications, Cah. Cent. Morphol. Math., 5, 1970.

McCullagh, P. and Nelder, J.A.: Generalized Linear Models (second edition). Chapman and Hall, London, UK, 1989.

Miller, A. J.: Algorithm AS 274: Least squares routines to supplement those of Gentleman. Appl. Statist., 41, 458478, 1992.

MillerW.A., and Cunge J.A.: Simplified equations of unsteady flow, Unsteady flow in open channels, by K. Mahamood and V. Yevjevixh, Water Resource Publication, Ft. Collins, Colo, 1975.

Montanari A. and Brath A.: A stochastic approach for assessing the uncertainty of rainfall-runoff simulation, Water Resour. Res., 40, W01106, doi:10.1029/2003WR002540, 2004

Nelder, J.A. and Wedderburn R.W.N.: Generalised linear models. Journal of the Royal Statistical Society, A, 135, 370-384, 1972.

Parker D. B.: Optimal algorithms for adaptive networks: Second order backpropagation, second order direct propagation and second order Hebbian learning, IEEE 1st Int. Conf. on Neural Networks, 2, 593-600, 1987.

Ponce V.M., and Yevjevich V.: Muskingum-Cunge methods with variables parameters, Jour. Hydraulics Division, ASCE, 104(HY12), 1663-1667, 1978.

Prince R.K.: Flood routing, Development in Hydraulics Engineering, 3, P. Novak (eds), Elsevier, 129-173, 1995.

Pujol, L.: Predicción de caudales en tiempo real en grandes cuencas utilizando redes neuronales artificiales, Ph.D. dissertation, Polytechnic University of Valencia, Department of Hydraulic Engineering and Environment, 34-39, 126-134, 2009.

Raftery, A. E.: Bayesian model selection in structural equation models, Testing Structural Equation Models, Bollen, K.A., and Long, J. S. (Eds.), Newbury Park, CA, 163-180, 1993.

Raftery, A. E., Balabdaoui, F., Gneiting, T., and Polakowski, M.: Using Bayesian model averaging to calibrate forecast ensembles, Monthly Weather Review, 133, 1155-1174, 2005.

Raiffa H., and Schlaifer, R.: Applied statistical decision theory, The MIT Press, Cambridge, MA, 1961.

Rougier J.: Probabilistic inference for future climate using an ensemble of climate model evaluations. Climatic Change, 81, 247-264, 2007.

Sloan P.G., and Moore I.D.: Modeling subsurface stormflow on steeply sloping forested watersheds, Water Resour. Res., 20(12), 1815-1822, 1984.

Smith A.A.: Generalized approach to kinematic flood routing, Jour. Hydrol., 45(1-2), 71-89, 1980.

Stagnitti F., Parlange M.B., Steenhuis T.S., and Parlange J.Y.: Drainage from a uniform soil layer on a hillslope, Water Resour. Res., 22(5), 631-634, 1986.

Steenhuis T.S., Parlange J.Y., Parlange M.B. and Stagnitti F.: A simple model for flow on hillslopes, Agric. Water Manage., 14(5), 153-168, 1988.

Tallis, G. M.: The Moment Generating Function of the Truncated Multi-Normal Distribution, J. R. Stat. Soc., Series B, 23 (1), 223-229, 1961.

Tang X., Samuels P.G.: Variable parameter Muskingum-Cunge method for flood routing in a compound channel, Jour. Hydraulic Research, 37, 591-614, 1999.

Thornthwaite, C.W., and Mather, J.R.: The water balance, Publ. in Climatology, 8(1), Lab. of Climatology, Centerton, N.J, 1955.

Todini E.: New trends in modeling soil processes from hill-slope to GCMS Scales - The role of water and the hydrological cycle in global change, (eds) H. R. Oliver, S. A. Oliver, NATO ASI Series I: Global Environmental Change, 31: 317-347, 1995.

Todini, E.: Hydrological Modelling: past, present and future, Hydrol. Earth Syst. Sci.,11(1), 468-482, 2007a.

Todini, E.: A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach, Hydrol. Earth Syst. Sci., 11, 1645-1659, 2007b.

Todini, E.: A model conditional processor to assess predictive uncertainty in flood forecasting, Intl. J. River Basin Management, 6 (2), 123-137, 2008.

Todini, E.: Predictive uncertainty assessment in real time flood forecasting, Ph. C. Baveye, M. Laba and J. Mysiak (eds.), Uncertainties in Environmental Modelling and Consequences for Policy Making. NATO Science for Peace and Security Series C: Environmental Security. Springer Netherlands, Amsterdam, NL, DOI: 10.1007/978-90-481-2636-1_9, ISBN 978-90-481-2635-4, pp. 205-228, 2009.

Todini, E. and Bossi, A.: PAB (Parabolic and Backwater) an unconditionally stable flood routing scheme particularly suited for real time forecasting and control, Journal of Hydraulic Research, 24(5), 405-424, 1986.

Todini, E., and Coccia G.: From deterministic to probabilistic thresholds: a better way to take advantage of predictive uncertainty in flood emergency

management, BHS Third International Symposium, Role of hydrology in managing consequences of a changig global environment, Newcastle, 19th-23th July, 2010.

Todini, E., and Ciarapica L.: The TOPKAPI model. Mathematical Models of Large Watershed Hydrology, Chapter 12, Singh. V.P. et al. (eds), Water Resources Publications, Littleton, CO, 2001.

Todini G., Rizzi R. and Todini E.: Detecting Precipitating Clouds over Snow and Ice Using a Multiple Sensors Approach, Journal of Applied Meteorology and Climatology, 48, 9, 1858-1867, doi: 10.1175/2009JAMC2151.1, 2009.

Toth, E. and Brath, A.: Ret neurali per la previsione di portata: un modello modulare basato sulla separazione degli idrogrammi, 31 Convegno Nazionale di Idraulica e Costruzioni Idrauliche, Perugia, 9-12 Settembre, 2008.

Van der Waerden, B.L.: Order tests for two-sample problem and their power I. Indagationes Mathematicae, 14, 453-458, 1952.

Van der Waerden, B.L.: Order tests for two-sample problem and their power II. Indagationes Mathematicae, 15, 303-310, 1953a.

Van der Waerden, B.L.:Order tests for two-sample problem and their power III. Indagationes Mathematicae, 15: 311-316, 1953b.

Vélez, J. J., Puricelli, M., López Unzu, F. and Francés, F.: Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework, Hydrol. Earth Syst. Sc., 13, 229-246, 2009.

Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis Algorithm for optimization and uncertainty assessment of hydrological model parameters. Water Resour. Res., 39, 1201, doi:10.1029/2002WR001642, 2003.

Vrugt, J. A., and Robinson, B. A: Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging, Water Resour. Res., 43, W01411, doi:10.1029/2005WR004838, 2007.

Weerts, A. H., Winsemius, H. C., and Verkade, J. S.: Estimation of predictive hydrological uncertainty using quantile regression: examples from the national flood forecasting system (England and Wales), Hydrol. Earth Syst. Sci. Discuss., 7, 5547-5575, doi:10.5194/hessd-7-5547-2010, 2010.

Werbos P.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science, Ph. D. dissertation, Harvard University, Cambridge, 1974.

Werbos P.: Generalization of backpropagation with application to a recurrent gas model, Neural Networks, 1, 339-356, 1988.

Werbos P.: Backpropagation through time: What it does and how to do it. Proceedings of the IEE, 78, 10, 1550-1560, 1990.

Wienmann P.E., and Laurenson E.M.: Approximate flood routing methods: a review, Jour. Hydraulics Division, ASCE, 105(12), 1521-1536, 1979.

Wooding, R.A.: A hydraulic modeling of the catchment-stream problem, Journal of Hydrology, 3, 254-267, 1965.

Appendix A

The TOPKAPI Hydrological Model

The TOPKAPI (TOPographic Kinematic APproximation and Integration) model is a fully-distributed physically-based hydrologic model with a simple and parsimonious parameterization which simulates the rainfall runoff transformation using data collected by a network of rain-gauges.

The model is based on the idea of combining the Kinematic approach and the topography of the basin. Spatial distribution of catchment parameters, precipitation input and hydrologic response is achieved horizontally by an orthogonal grid network and vertically by soil layers at each grid pixel.

Three *structurally similar* non-linear reservoir differential equations characterize the TOPKAPI approach and are used to describe subsurface flow, overland flow and channel flow. Moreover the TOPKAPI model includes components representing the primary processes of the hydrologic cycle: infiltration, percolation, evapo-transpiration and snowmelt, plus a lake/reservoir component, a parabolic routing component and a groundwater component.

Being a physically based model, the values of the model parameters can be easily derived from digital elevation maps, soil type and land use maps in terms of topology, slope, soil permeability, soil depth and superficial roughness. A calibration based on observed streamflow data is then necessary for

fine tuning the model to reproduce the behaviour of the catchment.

Thanks to its physically based parameters, the TOPKAPI model can be successfully implemented also in un-gauged catchments where the model cannot be calibrated using measured data. In this case the model parameters can be derived from thematic maps, literature and experience.

A.1   The Soil Water Component

A.1.1   Basic Assumptions

The fundamental assumptions on which the TOPKAPI model is based, can be described as follows:

1. Precipitation is assumed to be constant over the integration domain (namely the single cell), by means of suitable averaging operations on the local rainfall data, such as Thiessen polygons techniques, Block Kriging (De Marsily, 1986; Matheron, 1970) or others;

2. All the precipitation falling on the soil infiltrates into it, unless the soil is already saturated in a particular zone (namely the single cell); this is equivalent to adopting the saturation mechanism from below as the sole mechanism for the formation of overland flow, ignoring on the other hand the possible activation of the Hortonian mechanism due to infiltration excess. This decision is justified by the fact that the infiltration excess mechanism is characteristic of a local modeling scale, whereas the saturation excess mechanism, being linked to a cumulative phenomenon and conditioned by a lateral redistribution movement of the water in the soil, becomes dominant as the scale of the modeling increases (Blöschl and Sivapalan, 1995).

3. The slope of the water table is assumed to coincide with the slope of the ground, unless the latter is very small (less than 0.01%); this constitutes the fundamental assumption of the approximation of the kinematic

wave in the De Saint Venant equations, and it implies the adoption of a kinematic wave propagation model with regard to horizontal flow, or drainage, in the unsaturated area (Henderson and Wooding, 1964; Beven, 1981, 1982; Borah et al., 1980; Sloan and Moore, 1984; Hurley and Pantelis, 1985; Stagnitti et al., 1986; Steenhuis et al., 1988);

4. Local transmissivity, like local horizontal flow, depends on the total water content of the soil, i.e. it depends on the integral of the water content profile in a vertical direction;

5. Saturated hydraulic conductivity is constant with depth in a surface soil layer but much larger than that of deeper layers; this forms the basis for the vertical aggregation of the transmissivity, and therefore of the horizontal flow, as it will be described in details in the following section.

A.1.2 The Vertical Lumping

The transmissivity of a soil layer in non-saturated condition is given by the following expression:

$$T = \int_0^L k\left(\tilde{\theta}\left(z\right)\right) dz \tag{A.1}$$

Where: L = soil thickness of the layer affected by the horizontal flow.
$k\left(\tilde{\theta}\left(z\right)\right)$ = hydraulic conductivity in non-saturated conditions.
$\tilde{\theta} = \frac{\theta - \theta_r}{\theta_s - \theta_r}$ = reduced water content.
θ_r, θ_s = residual and saturated water content
θ = actual water content in the soil.

In accordance with the hypotheses 4) and 5) the transmissivity given by Equation A.1 can be replaced by the following approximated expression:

$$T\left(\widetilde{\Theta}\right) = k_s L \widetilde{\Theta}^\alpha \tag{A.2}$$

where k_s = saturated hydraulic conductivity.

$\widetilde{\Theta} = \frac{1}{L} \int_0^L \tilde{\theta}\,(z)\,dz$ = mean value along the vertical profile of the reduced water content.

α = parameter depending on the characteristics of the soil.

The horizontal flux is calculated as follows, by means of an approximation of the Brooks and Corey's formula $k\left(\tilde{\theta}\right) = k_s \tilde{\theta}^\alpha$:

$$q = tan\,(\beta)\,k_s L \widetilde{\Theta}^\alpha \tag{A.3}$$

where β = slope angle [rad].

α = parameter which depends on the soil characteristics.

$$[q] = \left[m^3 s^{-1}\right]$$

A.1.3   Kinematic Wave Formulation for Sub-Surface Flow

The analysis of a generic hydraulic system is usually addressed using the continuity equation and the dynamic equation. In the TOPKAPI model, the dynamic equation is represented by an approximate form expressed by Equation A.3. Combining Equation A.3 with the equation for continuity of mass, the following system is obtained:

$$\begin{cases} (\theta_s - \theta_r)\,L\frac{\partial \widetilde{\Theta}}{\partial t} + \frac{\partial q}{\partial x} = p \\ q = tan\,(\beta)\,k_s L \widetilde{\Theta}^\alpha \end{cases} \tag{A.4}$$

Where p is the intensity of precipitation [ms^{-1}].

The model is written in just one direction since it is assumed that the flow along the slopes is characterized by a preferential direction, which can be described as the direction of maximum slope.

Equation A.4 can be rewritten in terms of the actual total water content in the soil η:

$$\eta = (\theta_s - \theta_r)\,L\widetilde{\Theta} \tag{A.5}$$

and making the following substitution:

$$C = \frac{tan\,(\beta)\,k_s L}{(\theta_s - \theta_r)^\alpha L^\alpha} \tag{A.6}$$

The term C represents in physical terms a *local conductivity coefficient*, since it depends on soil parameters for a point position, which encompasses the effects of hydraulic conductivity and slope, to which it is directly proportionate, and storage capacity, to which it is inversely proportionate.

Equation A.4, rewritten in terms of actual total water content in the soil, along the vertical profile, leads to the following Kinematic equation:

$$\frac{\partial \eta}{\partial t} = p - C\frac{\partial \eta^\alpha}{\partial x} \tag{A.7}$$

A.1.4 Non-Linear Reservoir Model for the Soil Water in a Generic Cell

By integrating Equation A.7 in the soil over the *ith* DEM grid cell, whose space dimension is X, gives:

$$\frac{\partial v_{s_i}}{\partial t} = pX - (C_{s_i}\eta_i^{\alpha_s}) \tag{A.8}$$

where: v_{s_i}= volume of water per unit of width $[m^2]$.

X = grid cell dimension [m].

The subscript s is introduced here to distinguish this soil water equation from the ones relevant to the overland and the drainage network flows and will be kept from now on. The subscript i is introduced to highlight that the equation is referred to the *ith* cell and it will be omitted from now on.

In the TOPKAPI model, the grid cells are connected by a tree shaped network; water moves down slope along this tree shaped flow pathway starting from the initial cells (without upstream contributing areas) representing the 'sources' towards the outlet. According to this procedure, and assuming that in each cell the variation of the vertical water content along the cell is negligible, the volume of water stored in each cell (per unit width) can

be related to the total water content, which is equivalent to the free water volume in depth, by means of the simple expression:

$$v_{s_i} = X\eta \tag{A.9}$$

Substituting for η in Equation A.8 and writing it for a generic cell, given the total inflow to the cell, the following non-linear reservoir equation is obtained:

$$\frac{\partial V_s}{\partial t} = \left(pX^2 + Q_o^u + Q_s^u\right) - \frac{C_s X}{X^{2\alpha_s}} V_s^{\alpha_s} \tag{A.10}$$

where: V_s = volume of water stored in the *ith* DEM grid cell [m³].

pX^2 = precipitation on the *ith* DEM grid cell [$m^3 s^{-1}$].

Q_o^u = streamflow entering the active cell i as overland flow from the upstream contributing area [$m^3 s^{-1}$].

Q_s^u = streamflow entering the active cell i as sub-surface flow from the upstream contributing area [$m^3 s^{-1}$].

α_s = parameter which depends on the soil characteristics.

The volume of water stored in a cell can be related to the actual total water content by means of the following equation:

$$V_s = Xv_s = X^2\eta \tag{A.11}$$

Substituting Equation A.11 into Equation A.10 the differential equation for the soil component can be written as:

$$\frac{\partial \eta}{\partial t} = \frac{1}{X^2}\left(pX^2 + Q_o^u + Q_s^u\right) - \frac{C_s}{X}\eta^{\alpha_s} \tag{A.12}$$

In general Equation A.12 can be written as:

$$\frac{\partial \eta}{\partial t} = a - b\eta^c \tag{A.13}$$

where: $a = \frac{pX^2 + Q_o^u + Q_s^u}{X^2}$

$$b = \frac{C_s}{X}$$

$$c = \alpha_s$$

Equation A.13 can be solved analytically or numerically by means of the Runge-Kutta method.

A.1.5 Soil Water Balance

For the *ith* cell at each time step, the soil water balance can be calculated as follows:

$$Q_s^d = \left(pX^2 + Q_o^u + Q_s^u\right) - \frac{V_s\left(t_0 + dT\right) - V_s\left(t_0\right)}{dT} \tag{A.14}$$

where: Q_s^d = outflow from the *ith* cell during the time interval dT $[m^3 s^{-1}]$.
pX^2 = water falling on the *ith* cell during the time interval dT $[m^3 s^{-1}]$.
V_s = volume of water stored in the soil $[m^3]$.

In case of saturation of the soil cell the volume of water that exceeds the soil can be computed as follows:

$$V_{exf_s} = V_s\left(t_0 + dT\right) - V_{sat_s}$$

where: V_{exf_s} = saturation excess volume for the *ith* cell $[m^3]$.
V_{sat_s} = saturated soil water storage for the *ith* cell $[m^3]$.

A.1.6 Subsurface Flow in a Cell with General Inclination

If we consider a pixel with slope equal to $tg\beta_1$ in x direction and slope equal to $tg\beta_2$ in y direction the Equation A.4 should be modified in the following way:

$$q = tan\left(\beta_1\right)\left(1 + \frac{tan\left(\beta_2\right)}{tan\left(\beta_1\right)}\right)k_s L\widetilde{\Theta}^\alpha \tag{A.15}$$

As a consequence also the local conductivity coefficient C_s will be modified:

$$C_s = tan\left(\beta_1\right) \frac{\Sigma_s k_s L}{\left(\theta_s - \theta_r\right)^\alpha L^\alpha} \tag{A.16}$$

with:

$$\Sigma_s = 1 + \frac{tan\left(\beta_2\right)}{tan\left(\beta_1\right)} \tag{A.17}$$

The coefficient σ_S will be called *soil drainage coefficient.* Equation A.13 representing the non-linear reservoir for the subsurface flow component will be modified in the following way:

$$\frac{\partial \eta}{\partial t} = a - \Sigma_s b \eta^c \tag{A.18}$$

From Equation A.18 the total outflow Q_{out} from the soil is computed. Then the outflow is partitioned between the downstream cell and the channel network, according to the flow partition coefficient.

A.2   The Surface Water Component

The input to the surface water model is the *precipitation excess* resulting from the saturation of the surface soil layer. In addition, water in the soil can exfiltrate on the surface as return flow due to a sudden change in hill slope or soil properties, and thus it can also feed the overland flow. The subsurface flow and the overland flow together feed the channel along the drainage network.

Overland flow routing is described similarly to the soil component, according to the kinematic approach (Wooding, 1965), in which the momentum equation is approximated by means of the Manning's formula. For a general cell, the kinematic wave approximation for overland flow is described as:

$$\begin{cases} \frac{\partial h_o}{\partial t} = r_o - \frac{\partial q_o}{\partial x} \\ q_o = \frac{1}{n_o} tan\left(\beta\right)^{\frac{1}{2}} h_o^{\frac{5}{3}} = C_o h_o^{\alpha_o} \end{cases} \tag{A.19}$$

where: h_o = water depth over the ground surface [m].

r_o = saturation excess resulting from the solution of the soil water bal-

ance either as precipitation or exfiltration from the soil in absence of rainfall $[\text{ms}^{-1}]$.

q_o = horizontal flow on the ground surface, corresponding to a streamflow per unit of width $[\text{m}^2\text{s}^{-1}]$.

n_o = Manning's friction coefficient for the surface roughness $[m^{-1/3}s]$.

α_o = exponent that derives from using Manning's formula, equal to $5/3$.

$C_o = \frac{tan(\beta)^{\frac{1}{2}}}{n_o}$ = coefficient relevant to Manning's formula for overland flow.

A subscript o denotes the overland flow. Equation A.19, leads to the following kinematic equation:

$$\frac{\partial h_o}{\partial t} = r_o - C_o \frac{\partial \left(h_o^{\alpha_o} \right)}{\partial x} \tag{A.20}$$

By analogy with what was done for the soil, assuming the surface water depth constant over the cell and integrating the kinematic equation over the longitudinal dimension, the non-linear reservoir equation for the overland flow for the *ith* cell can be obtained as:

$$\frac{\partial V_{o_i}}{\partial t} = r_{o_i} X W_{o_i} - \frac{C_o W_{o_i}}{(X W_{o_i})^{\alpha_o}} V_{o_i}^{\alpha_o} \tag{A.21}$$

where V_o = surface water volume in the cell $[\text{m}^3]$.

W_o = width of the surface (free of the channel) $[\text{m}]$.

The subscript i is introduced here to highlight that Equation A.21 was written for the *ith* DEM grid cell and it will be omitted from now on. The volume of water stored on the surface of each cell can be written through a simple expression:

$$V_o = X W_o h_o \tag{A.22}$$

Substituting Equation A.22 into Equation A.21 the differential equation for the surface component can be written as:

$$\frac{\partial h_o}{\partial t} = r_o - \frac{C_o}{X} h_o^{\alpha_o} \tag{A.23}$$

In general Equation A.23 can be written as:

$$\frac{\partial h_o}{\partial t} = a - b h_o^c \qquad (A.24)$$

whit: $a = r_o = \frac{1}{XW_o} \frac{V_{exf}}{dT}$

$$b = \frac{C_o}{X} = \frac{\tan{(\beta)}^{\frac{1}{2}}}{n_o X}$$

$$c = \alpha_o$$

where: V_{exf} = precipitation excess $[\text{m}^3]$

Equation A.23 can be solved numerically (Runge-Kutta) or analytically.

A.2.1 Surface Water Balance

For the *ith* cell at each time step, the surface water balance can be calculated as follows:

$$Q_o^d = (r_o X W_o) - \frac{V_o\,(t_0 + dT) - V_o\,(t_0)}{dT} \qquad (A.25)$$

where: Q_o^d = outflow from the *ith* cell during the time interval T $[m^3 s^{-1}]$.

$r_o X W_o$ = inflow into the *ith* cell during the time interval dT $[ms^{-1}]$.

V_o= volume of water on the surface $[m^3]$.

Up to this point it has been implicitly assumed that the entire overland flow from a cell flows into the downstream cell immediately. However, this is not entirely true since note has to be taken of the depletion caused by the drainage network. Thus, for the cells in the channel network, the overland flow is still evaluated by Equation A.23, but it is then partitioned between the channel and the downstream cell. This allows determination of the amount of overland flow feeding the drainage channel network.

A.2.2   Overland Flow in a Cell with General Inclination

If we consider a pixel with slope equal to $tg\beta_1$ in x direction and slope equal to $tg\beta_2$ in y direction the Equation A.19 should be modified in the following way:

$$q_o = \frac{1}{n_o}(tan\beta_1)^{\frac{1}{2}} \left[1 + \left(\frac{tan\beta_2}{tan\beta_1} \right)^{\frac{1}{2}} \right] h_o^{\frac{5}{3}} \tag{A.26}$$

$$\Sigma_o = 1 + \left(\frac{\tan{(\beta_2)}}{\tan{(\beta_1)}} \right)^{\frac{1}{2}} \tag{A.27}$$

The coefficient σ_O will be called *surface drainage coefficient*. Equation A.13 representing the non-linear reservoir for the overland flow component will be modified in the following way:

$$\frac{\partial h_o}{\partial t} = a - \Sigma_o b h_o^c \tag{A.28}$$

From Equation A.28 the total outflow Q_{out} from the overland flow is computed. Then the outflow is partitioned between the downstream cell and the channel network according to the flow partition coefficient.

A.3   The Channel Component

In the TOPKAPI model, different kinds of channel cross section geometries can be set; following, a rectangular cross section will be used as an example to describe the channel component structure.

A.3.1   Channels with Rectangular Cross Sections

The channel flow is described similarly to the surface component, although in this case the channel is assumed to be tree shaped with reaches having rectangular cross sections.

The kinematic wave approximation for the channel flow is described ac-

cording to the kinematic approach in which the momentum equation is approximated by means of the Manning's formula:

$$\begin{cases} \frac{\partial V_c}{\partial t} = (r_c + Q_c^u) - q_c \\ q_c = \frac{1}{n_c}\sqrt{s_0}\left(\frac{A_x}{C_x}\right)^{\frac{2}{3}} B_x y_c^{\frac{5}{3}} \end{cases} \tag{A.29}$$

where: y_c = water depth in the channel reach [m].

r_c = lateral drainage input, including the surface runoff and the soil drainage reaching the channel $[m^3 s^{-1}]$.

Q_c^u = inflow from the channel reach of the upper cell $[m^3 s^{-1}]$.

q_c = horizontal flow in the channel $[m^3 s^{-1}]$.

n_c = Manning's friction coefficient $[m^{-1/3}s]$.

s_0 = bed slope.

A_x = wet area $[m^2]$

C_x = wet contour $[m]$

B_x = width of the channel reach $[m]$.

A subscript c denotes the channel flow. Equation A.29, rewritten in terms of water depth in the channel reach, y_c, leads to the following equation:

$$\frac{\partial V_c}{\partial t} = (r_c + Q_c^u) - \frac{\sqrt{s_0}}{n_c}\left(\frac{A_x}{C_x}\right)^{\frac{2}{3}} B_x y_c^{\frac{5}{3}} \tag{A.30}$$

With simple substitutions we obtain the following equation that describes the non-linear reservoir equation for the channel flow for the *ith* cell:

$$\frac{\partial V_c}{\partial t} = (r_c + Q_c^u) - \frac{\sqrt{s_0}}{n_c}\left(\frac{1}{C_x}\right)^{\frac{2}{3}} \frac{1}{X^{\frac{5}{3}}} V_c^{\frac{5}{3}} \tag{A.31}$$

In general Equation A.31 can be written as:

$$\frac{\partial V_c}{\partial t} = a - bV_c^c \tag{A.32}$$

whit: $a = r_c + Q_c^u$

$$b = \frac{\sqrt{s_0}}{n_c}\left(\frac{1}{C_x}\right)^{\frac{2}{3}} \frac{1}{X^{\frac{5}{3}}}$$

$$c = \frac{5}{3}$$

where: $A_x = By_{c_0}$ = wet area at the beginning of the computation time step $[m^2]$.

$C_x = 2y_{c_0} + B$ = wet contour at the beginning of the computation time step $[m]$.

The channel width B is increasing as a function of the area drained by the *ith* cell on the basis of geo-morphological considerations.

A.3.2 Channel Water Balance

For the *ith* cell at each time step, the channel water balance can be calculated as follows:

$$Q_c^d = (r_c XW + Q_c^u) - \frac{V_c(t_0 + dT) - V_c(t_0)}{dT} \tag{A.33}$$

where: Q_c^d = outflow $[m^3 s^{-1}]$.

$r_c XW$ = inflow from the lateral cells $[m^3 s^{-1}]$.

Q_c^u = inflow from the upper cell $[m^3 s^{-1}]$.

V_c = volume of water in the channel $[m^3]$.

A.3.3 Analytical Solution of the Non-Linear Reservoir Ordinary Differential Equation (ODE)

As described in the previous subsections, the TOPKAPI model formulation leads to three tree-shaped cascades of non-linear reservoirs, each of which is described by a *structurally similar'* ordinary differential equation (ODE) to be solved in time. In the first version of the TOPKAPI model (Todini and Ciarapica, 2001), the solution of the ODE for each single reservoir representing the soil, the surface and the channel network, was based upon a variable step fifth order Runge-Kutta numerical algorithm due to Cash and Karp (1990). Nowadays, it has been found that the non-linear reservoir equation

can be solved analytically based on an appropriate approximation (Liu and Todini, 2002).

A.3.4   The Muskingum-Cunge-Todini Routing Method

In the TOPKAPI model it is possible to use the Muskingum-Cunge-Todini (MCT) (Todini, 2007b) routing method as an alternative to the Kinematic non-linear reservoir for channels with slope smaller than 0.1%, namely channels where the Kinematic approximation of De Saint-Venant equations does not hold.

The Muskingum-Cunge (MC) routing method (Cunge, 1969; Ponce adn Yevjevich, 1978; Koussis, 1980, 1983; Miller and Cunge, 1975; Wienmann and Laurenson, 1979) is actually a lumped Kinematic wave routing method, in which the Kinematic wave equation is transformed into an equivalent diffusive wave equation by matching the physical diffusion to the numerical diffusion resulting from the imperfectly centered finite difference scheme (Smith, 1980; Tang and Samuels, 1999). Thus the MC method accounts for both the convection and diffusion of the flood wave. The routing parameters can be linked to physical channel properties and flow characteristics (Cunge, 1969), and when these parameters are recalculated and updated as a function of local flow values for each computational cell, the routing parameters are varying in time (Prince, 1995). The MCT algorithm, is basically a variable parameter MC corrected for its typical mass balance error (Todini, 2007b).

A.4   The Evapo-Transpiration Component

The evapo-transpiration is taken into account as water loss, subtracted from the soil's water balance. A simplified technique is used to calculate evapo-transpiration starting from air temperature and from other topographic, geographic and climatic information. The effects of the vapour pressure and wind speed are explicitly ignored. In the TOPKAPI model, the evapo-transpiration is evaluated at the DEM grid scale.

A.4.1 Empirical Equation for Computing the Reference Potential Evapo-Transpiration

An empirical equation, that relates the reference potential evapo-transpiration ET_{0m}, to the compensation factor *Wta*, to the mean recorded temperature of the month T and the maximum number of hours of sunshine N of the month, was developed. The reference potential evapo-transpiration is computed on a monthly basis using one of the available simplified expressions such as for instance the one due to Thornthwaite and Mather (1955). The developed relationship is linear in temperature (and hence additive) and allows the unbundling of the monthly results on daily or hourly basis, while most other empirical equations are ill-suited for time intervals shorter than one month.

The relation used, which is structurally similar to the radiation method formula (Doorembos et al., 1984) in which the air temperature is taken as an index of radiation, is:

$$ET_{0m} = \alpha + \beta N W_{ta} T_m \qquad (A.34)$$

Where: ET_{0m} = reference evapo-transpiration for a monthly time step (computed using Thornthwaite's formula) [mm]

α, β = regression coefficients to be estimated

T_m = area mean air temperature averaged over a month [$^\circ C$]

N = monthly mean of the maximum number of daily hours of sunshine (tabulated as a function of latitude)

W_{ta} = weighting factor, it can be either obtained from tables or approximated by a fitted parabola:

$$W_{ta} = \overline{T}^2 + B\overline{T} + C$$

A, *B*, *C* = coefficients to be estimated

$\overline{T}$ = mean monthly temperature [$^\circ C$]

For a given time step Δt and a given crop culture, the potential evapo-transpiration value is computed as:

$$ET_0 = K_c \left(\alpha + \beta N W_{ta} T_{\triangle t} \right) \frac{\triangle t}{30 \cdot 24 \cdot 3600} \tag{A.35}$$

where: ET_0 = reference evapo-transpiration for a specified time step Δt [mm].

K_c = crop factor.

$T_{\triangle t}$ = pixel mean air temperature averaged over Δt [°C].

A.4.2   Estimation of the Average Monthly Potential Evapo-Transpiration According to Thornthwaite

The values of the potential evapo-transpiration can be computed for a given DEM grid according to Thornthwaite and Mather (1955), by means of the following formula:

$$ET_{0m}(i) = 16a(i) \left[10\frac{T(i)}{b} \right]^c \tag{A.36}$$

whit: $a = \frac{n(i)}{30}\frac{N(i)}{12}$

$$b = \sum_{i=1}^{12} \left[\frac{T(i)}{5} \right]^{1.514}$$

$$c = 0.49239 + 1792 \times 10^{-5}b - 771 \times 10^{-7}b^2 + 675 \times 10^{-9}b^3$$

where: $ET_{0m}(i)$ = average monthly potential evapo-transpiration [mm/month].

$T(i)$ = monthly-average air temperature for ith month [∘ C].

$n(i)$ = number of days in month i.

$N(i)$ = Mean Daily Duration of Maximum Possible Sunshine Hours (in 'Crop Water Requirements' FAO Irrigation and Drainage Paper 24).

A.4.3   Computation of the Actual Evapo-Transpiration

The potential evapo-transpiration is corrected as a function of the actual soil moisture content, to obtain the actual evapo-transpiration (ETa):

$$
\begin{cases}
ET_a = 0 & for\ V \leq \beta_1 V_{sat} \\
ET_a = ET_0 \frac{V}{V_{sat}} & for\ \beta_1 V_{sat} \leq V \leq \beta_2 V_{sat} \\
ET_a = ET_0 & otherwise
\end{cases}
$$

where: V, V_{sat} = actual and saturation volume of water into the soil [m³].
β_1, β_2 = parameters to be set.

A.5   The Snow Accumulation and Snow Melting Component

The snowmelt module of the TOPKAPI model is driven by a radiation estimate based upon the air temperature measurements; in practice, the inputs to the module are the precipitation, the temperature, and the same radiation approximation which was used in the evapo-transpiration module.

The snowmelt module consists of the following steps.

A.5.1   Estimation of Solar Radiation

The estimation of the solar radiation at the DEM is performed by re-converting the latent heat and the sensible heat, assumed equals to the reference evapo-transpiration back into radiation, by means of a conversion factor C_{er} (Kcal Kg^{-1}):

$$C_{er} = 606.5 - 0.695\,(T - T_0) \tag{A.37}$$

where: C_{er} = conversion factor $[KcalKg^{-1}]$.
T_0 = fusion temperature of ice $[273°K]$.
T = air temperature $[°\ K]$.

In addition, to account for albedo, which plays an extremely important role in snowmelt, it is necessary to apply an efficiency factor which will be assumed approximately as $\eta = 0.6$ for clear sky and $\eta = 0.8$ for overcast conditions. Moreover, a coefficient η_{rad} is used to take in account the radiation efficiency; it depends on the sun height with respect to the terrain slope. This leads to the following estimate for the driving radiation term:

$$Rad = 2\eta_{al}\eta_{rad}\left[606.5 - 0.695\left(T - T_0\right)\right]ET_0 \tag{A.38}$$

where: Rad = radiation term.

η_{al} = efficiency factor for albedo.

η_{rad} = radiation efficiency factor.

ET_0 = potential evapo-transpiration.

A.5.2   Computation of the Solid and Liquid Percentage of Precipitation

The percentage of liquid precipitation is calculated by means of a function of the air temperature:

$$F\left(T\right) = \frac{1}{1 + e^{-\frac{T - T_s}{\Sigma}}} \tag{A.39}$$

where σ is equal to 0.3 (derived by experimental data) and the value of T_S (which generally ranges between 271 and 275°K) must be derived, as previously mentioned, by plotting the frequency of the status of historically recorded precipitation as a function of air temperature.

A.5.3   Estimation of the Water and Energy Budgets on the Hypothesis of Zero Snowmelt

The water equivalent mass (Z) is estimated with the following simple mass balance equation:
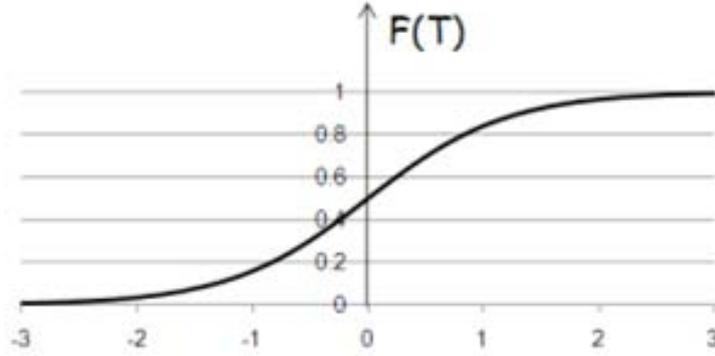
Figure A.1: Percentage of liquid precipitation for $T_S = 0$

$$Z^*_{t+\Delta t} = Z_t + P \tag{A.40}$$

where P is the precipitation.

The water equivalent at the end of the time step is identified with a star, because it is a tentative value which does not yet account for the eventual snowmelt. Similarly to the mass, the energy is estimated in the following way, by computing the increase (or decrease) of total energy (E):

$$E^*_{t+\Delta t} = E_t + Rad + C_{si}T \cdot [1 - F(T)] \, P + [C_{si}T_0 + C_{lf} + C_{sa}(T + T_0)] \, P \cdot F(T) \tag{A.41}$$

where: C_{si} = specific heat of ice

C_{lf} = latent heat of fusion of water

C_{sa} = specific heat of water

A.5.4 Estimation of Snowmelt and Updating of Mass and Energy Budgets

If the total available energy is smaller or equal to that required to maintain the total mass in the solid phase at the temperature T_0 i.e: $C_{si}Z^*_{t+\Delta t}T_0 \geq$

$E_{t+\Delta t}^*$, it means that the available energy is not sufficient to melt part of the accumulated snow, and therefore:

$$
\begin{cases}
R_{sm} = 0 \\
Z_{t+\Delta t} = Z_{t+\Delta t}^* \\
E_{t+\Delta t} = E_{t+\Delta t}^*
\end{cases}
\tag{A.42}
$$

where: R_{sm} = snowmelt [mm]

If the total available energy is larger than that required to maintain the total mass in the solid phase at the temperature T_0, it means that part of the accumulated snow will melt, and therefore the following energy balance equation holds:

$$
C_{si} \left(Z_{t+\Delta t}^* - R_{sm} \right) T_0 = E_{t+\Delta t}^* - \left(C_{si} T_0 + C_{lf} \right) R_{sm}
\tag{A.43}
$$

from which the snowmelt and the mass and energy state variables can be computed as:

$$
\begin{cases}
R_{sm} = \frac{E_{t+\Delta t}^* - C_{si} T_0 Z_{t+\Delta t}^*}{C_{lf}} \\
Z_{t+\Delta t} = Z_{t+\Delta t}^* - R_{sm} \\
E_{t+\Delta t} = E_{t+\Delta t}^* - \left(C_{si} T_0 + C_{lf} \right) R_{sm}
\end{cases}
\tag{A.44}
$$

A.6 The Percolation Component

For the deep aquifer flow, the response time related to the vertical transport of water through the thick soil above this aquifer is so large that horizontal flow in the aquifer can be assumed to be almost constant with no significant response on one specific storm event in a catchment (Todini, 1995). Nevertheless, the TOPKAPI model accounts for water percolation towards the deeper subsoil layers even though it does not contribute to the streamflow.

It is assumed that percolation starts if the soil moisture content of the upper soil layer exceeds its field capacity. The percolation rate from the upper soil layer is assumed to increase as a function of the soil water content, ac-

cording to an experimentally determined power law (Clapp and Hornberger, 1978; Liu et al., 2005).

$$P_r = k_{sv} \left(\frac{v}{v_{sat}} \right)^{\alpha_p} \tag{A.45}$$

where: P_r = percolation [mm]

k_{sv} = vertical soil saturated hydraulic conductivity

v = volume of water [m^3]

v_{sat} = local saturation volume [m^3]

α_p = exponent depending on the type of the soil (α_p 11 for sand; α_p 25 for clay)