Optimal Control of Water Systems Under Forecast Uncertainty

Robust, Proactive, and Integrated

Luciano Raso

Optimal Control of Water Systems Under Forecast Uncertainty

Robust, Proactive, and Integrated

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties

in het openbaar te verdedigen op maandag 16 September 2013 om 10:00 uur

door

Luciano Raso

Civiel ingenieur
geboren te Reggio Calabria (Italië)

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. N. van de Giesen

Samenstelling promotiecommissie:
Rector Magnificus, *voorzitter*
Prof. dr. ir. **N. van de Giesen**, *Technische Universiteit Delft, promotor*
Dr. ir. **P.O. Malaterre**, *Irstea Montpellier*
Prof. dr. ir. **A. Montanari**, *University of Bologna*
Dr. ir. **P.J. van Overloop**, *Technische Universiteit Delft*
Dr. ir. **D. Schwanenberg**, *University of Duisburg-Essen*
Prof. dr. ir. **R. Soncini-Sessa**, *Politecnico di Milano*
Prof. dr. ir. **W. Thissen**, *Technische Universiteit Delft*
Prof. dr. ir. **M. Stive**, *Technische Universiteit Delft, reservelid*

ISBN: 97890-6562-331-7

The research described in this dissertation was performed at the Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, and at Deltares, Delft and financially supported by Deltares.

Published by VSSD, Delft, the Netherlands, September 2013

Keywords: Operational Water Management, Optimal Control, Model Predictive Control, Ensemble Forecasting, Uncertainty, Multistage Stochastic Programming, Closed Loop Formulation, Robustness, Predictive Uncertainty, Heteroschedasticity

To professor Maria D'Africa, and to her faith in rationality.

Preface

This thesis is the main output of my research as PhD student, carried out at Delft University of Technology and funded by Deltares Foundation. The position I have occupied, halfway between these two institutions, is reflected in its content, having both application and theoretical value. It combines a practical usefulness, which is of value for a consulting company that sells innovative solutions, and attention to the formal correctness, which is of value for an university that delivers education and research.

From the application side, it can be seen as a Real Time Control (RTC) application to water systems. Complex systems can be integrated and harmonized by optimal control. Initially, after the Second World War, control was a warfare application only (missile control). Control applications were later extended to industry (powerplants), showing now their usefulness for water systems operation. If technologies had karma, the reincarnation of control in nobler and nobler systems means that it behaved well in previous applications.

Deltares applies RTC to different water systems all over the world. Deltares have recently won a research contract for the management of the Federal Columbia River Power System, by Bonneville Power Administration (USA). The method presented in this thesis is a constituent element of the Deltares proposal; even before the end of my PhD, the research investment already paid back.

From a more theoretical side, this thesis is about decisions under uncertainty, involving utility theory and probability. The work presented here is an attempt to bring decision theory "down from the platonic world of ideas", and probability "out of the urn".

Herbert Simon, nobel laureate, regarded in fact utility theory as an ideal method, good as mental framework, but with little practical applicability, difficult to extend to real world problems. Utility theory is indeed a rather formal method; it requires to define explicitly the system model and the objectives. However, in water systems operation, models of the physical process are available, and the decision set

is generally bounded. Therefore a complete problem setting, as required in utility theory, is a possible option. Defining all elements of the problem forces the analyst to make them explicit, thus verifiable and communicable. This matches the educative role that, in my opinion, university should have: providing instruments to frame problems. The adage of Charles Kettering, saying that "a problem well stated is a problem half solved" is, in this context, even more true. Nevertheless, the objective function definition deserves, in my opinion, further research. An objective function defines what I call "subjective" rationality, i.e. what I think it is rational. However, it does not ensure "objective" rationality, i.e. what it is actually rational.

A cornerstone of this thesis is uncertainty, and the mathematical tool to deal with it, probability theory. Probability theory is the science of partial information. Information is, in this framework, a constraint on uncertainty.

A paradigm shift is taking place within universities, that I am experiencing and promoting. I was taught to interpret probability as frequency for an experiment repeated sufficient times; I have turned to interpret probability as lack of knowledge. Uncertainty does not stem from some randomness of the system, as an ontological property of the real world [Popper, 1959], but from a systematic partial knowledge of it [Jaynes and Bretthorst, 2003]. Information is in itself an interesting entity. Information is costly to obtain and to store, measurable, as mass or electricity, and as mass or electricity, it propagates according to its own laws. Understanding the laws that explain the information flow in time is an exciting challenge.

My PhD has been "a pleasant Research Path on the equilibrium between the Valley of Chaos and the Desert of Boredom". I managed to go through it until the end. It was indeed pleasant, I never got bored and rarely got lost in the chaos. Doing research requires communication, problem solving, and analytical skills, emotional balance, concentration, intuition, curiosity, creativity, and hard work. Even if this thesis is an individual product, it was not a solitary walk. I thank here all the people with whom I shared part of this path and supported me, professionally or emotionally. They are recognized, one by one, in the Acknowledgments section.

Summary

Water systems consist of natural and man made objects serving multiple essential purposes. They are affected by many types of meteorological disturbances. In order to deal with these disturbances and to serve the desired objectives, infrastructures have been built and managed by societies for specific purposes. Given a water system, and its purposes, the control of the existing infrastructures is the subject of *operational water management*. The system controller, either a natural person or a mathematical algorithm, takes his recursive decisions observing the state of the system and trying to bring it to the desired condition. *Model Predictive Control* (MPC) is an advanced method for the control of complex dynamic systems. When applied to water systems operation, MPC provides *integrated* and *optimal* management.

If disturbance forecasts are available, this information can be integrated in the control policy and water management becomes *proactive*. Before the realization of the disturbance, the MPC controller sets the system to a state which is optimal to accommodate the expected disturbance. A typical example is lowering the water level of a reservoir before an expected storm event in order to avoid floods. However, MPC is a deterministic algorithm, therefore mismatches between the nominal and the real system can jeopardize its robustness.

In proactive control of open water systems, the main uncertainty is generally related to the difficulty of producing good forecasts. Weather and hydrological processes are difficult to predict, and meteorological or rainfall-runoff models can be wrong. Especially when using only one deterministic estimate, the control is more vulnerable to forecast uncertainty, running the risk of taking action against a predicted event that will not occur.

The research question of this thesis is

> How to use existing forecasting methods in optimal control schemes, thereby enhancing robustness in the face of forecasting uncertainty?

In open water systems, such as rivers, canals, or reservoirs, the available forecast is generally the natural inflow, which is the output of a deterministic rainfall-runoff

model. The model produces a point estimate, which is the expected value of the variable of interest. Nevertheless, the nonlinearity of the control problem requires the forecast of the entire probability distribution. When residuals are assumed independent, identically distributed, zero-mean, and Gaussian, then the variance is the only extra parameter required to build up the entire distribution, and its value can be estimated from the data. However, residuals of rainfall-runoff models are in fact heteroscedastic (i.e. the variance changes in time) and autocorrelated. In Chapter 2 it is shown how to deal with both deficiencies. Dynamic modelling of predictive uncertainty is built up by regression on absolute residuals, and applied to two test cases: the Rhone River, in Switzerland, and Lake Maggiore, at the border between Italy and Switzerland.

When the information on the catchment state does not offer sufficient anticipation, for example because the catchment dynamics are fast compared to the controlled system, it is necessary to include weather forecasts. Meteorological agencies produce not only a deterministic trajectory of the future state of the weather system, but a set of them, called ensemble, to communicate the forecast uncertainty.

The algorithm presented in Chapter 3, called *Tree-Based Model Predictive Control* (TB-MPC), exploits the information contained in the ensemble, setting up a Multistage Stochastic Programming (MSP) problem within the MPC framework. MSP is a stochastic optimization scheme that takes into account not only the present uncertainty, but its resolution in time as well. Going on in time along the control horizon, information will enter the system. Consequently, uncertainty will be reduced, and the control strategy after uncertainty reduction will change according to the occurring ensemble member. The key idea of TB-MPC is producing a tree topology from the ensemble data and using this tree in the following MSP optimization. A tree specifies in fact the moments when uncertainties are resolved.

Generating a tree from ensemble data is both difficult and of critical importance. It has been considered an open problem until now, especially regarding the tree branching structure, which also strongly affects control performance. Chapter 4 shows a new methodology that produces a tree topology from ensemble data. The proposed method models the information flow to the controller. This implies the explicit definition of the available observations and their degree of uncertainty.

Chapter 5 summarizes the contribution of my PhD and the research directions that, in my opinion, deserve more investigation.

Samenvatting

Watersystemen bestaan uit natuurlijke en kunstmatige objecten die meerdere essentiële doelen dienen. Ze worden benvloed door verschillende meteorologische verstoringen. Om met deze verstoringen om te kunnen gaan en om aan de gewenste doelstellingen te kunnen voldoen, is infrastructuur gebouwd die wordt beheerd volgens de maatschappelijke doelen. Operationeel waterbeheer betreft het regelen van de bestaande infrastructuur gegeven het watersysteem en haar doelstellingen. De systeemregelaar, zijnde een natuurlijke persoon of een wiskundig algoritme, observeert de toestand van het systeem en probeert het in de gewenste staat te brengen. Model Predictive Control (MPC) is een geavanceerde methodiek voor het regelen van complexe dynamische systemen. Toegepast op watersystemen kan met MPC een gentegreerde en optimaal beheer worden verkregen.

Als voorspellingen van verstoring beschikbaar zijn, kan deze informatie worden gentegreerd in het operationele beheer en wordt het waterbeheer anticiperend, dat wil zeggen: vóór de realisatie van de verstoring, brengt de MPC-regelaar het systeem naar een staat die optimaal is, anticiperend op de verwachte storing. Een voorbeeld is de verlaging van het waterpeil van een reservoir vóór een hoog afvoer gebeurtenis om zodoende een overstroming te voorkomen. Echter, MPC is een deterministisch algoritme, waarbij de onzekerheden tussen model en werkelijkheid de betrouwbaarheid in gevaar kunnen brengen. De grootste onzekerheid in het anticiperend regelen van open-watersystemen heeft vaak betrekking op de moeilijkheid om goede voorspellingen te produceren. Hydrologische en meteorologische processen zijn moeilijk te voorspellen. Meteorologische en neerslag-afvoer modellen kunnen onnauwkeurig zijn. Vooral bij het gebruik van slechts een deterministische puntschattingen is de regeling kwetsbaar voor onzekerheden in de voorspellingen met het risico om actie te ondernemen anticiperend op een voorspelde gebeurtenis die uiteindelijk niet zal gebeuren. De onderzoeksvraag van dit proefschrift is aldus

> hoe kunnen de bestaande voorspellingen in optimale regelmethodieken worden gebruikt om zodoende te zorgen voor robuustheid tegen de onzekerheden in deze voorspellingen.

Bij open-watersystemen zoals rivieren, kanalen of reservoirs, is de beschikbare voorspelling de natuurlijke afvoer, zijnde de uitput van een deterministische neerslag-afvoer model. Het model produceert een puntschatting, die de verwachtingswaarde van de variabele is. Toch vereist de niet-lineariteit van het regelprobleem de voorspelling van de gehele kansverdeling. Typisch worden residuen onafhankelijke, identiek verdeelde, nul-gemiddelde en normaal aangenomen. In dit geval is de variantie de enige extra parameter vereist om de gehele verdeling te construeren en deze waarde kan uit de data worden geschat. Fouten in neerslag-afvoer modellen zijn echter heteroscedastisch (de variantie verandert) en autogecorreleerde. In hoofdstuk 2 wordt getoond hoe met beide gebreken om te gaan. Dynamische modellering van predictieve onzekerheid wordt opgebouwd door regressie op de absolute afwijkingen en toegepast op twee testcases: de rivier Rhône, in Zwitserland, en het Maggioremeer, op de grens tussen Italië en Zwitserland.

Als de informatie over het stroomgebied niet voldoende anticipatietijd biedt, bijvoorbeeld omdat het stroomgebied dynamisch snel is in vergelijking met het geregelde systeem, is het noodzakelijk om weersvoorspellingen te gebruiken. Meteorologische instituten produceren niet alleen een deterministisch traject van de toekomstige toestand van het weersysteem, maar tevens een verzameling van trajecten om de voorspellingenonzekerheid te communiceren, genaamd ensemble.

Het algoritme dat wordt gepresenteerd in Hoofdstuk 3, genaamd *Tree-Based Model Predictive Control* (TB-MPC), maakt gebruik van de informatie beschikbaar in de ensemble en construeert Multistage Stochastic Programming (MSP) binnen het MPC kader. MSP is een stochastische optimalisatie methodiek die rekening houdt met de huidige onzekerheid en tevens de resolutie van onzekerheid in de tijd. Met het voortschrijden in de tijd over de regelhorizon, komt informatie het systeem binnen. Hierdoor zal de onzekerheid worden verminderd en de regelstrategie zal veranderen na deze reductie van onzekerheid afhankelijk van het optredende lid van het ensemble. Het idee van TB-MPC is het opbouwen van een boomstructuur uit de ensemblegegevens en de boom te gebruiken in de volgende MSP-optimalisatie. Een boom geeft de momenten aan dat onzekerheden worden opgelost.

Het opbouwen van een boom uit ensemble data is zowel moeilijk als van cruciaal belang. Tot nu toe is het een open probleem, vooral wat betreft de boomstructuur die ook grote invloed op de regeling heeft. Hoofdstuk 4 toont een nieuwe methode die de boomstructuur uit de ensemble genereert. De voorgestelde methode modelleert de informatiestroom richting de controller. Dit impliceert de expliciete definitie van de beschikbare waarnemingen en hun mate van onzekerheid.

Hoofdstuk 5 vat de bijdrage van mijn promotieonderzoek samen en de onderzoek-

srichtingen die, mijns inziens, meer aandacht verdienen.

Contents

Chapter 1

Integrated, Optimal, and Proactive Operational Water Resources Management

1.1 Integrated Water Resources Management, down to earth

Integrated Water Resources Management (IWRM) is defined by Cardwell et al. [2006] as

> a coordinated, goal-directed process for controlling the development and use of river, lake, ocean, wetland, and other water assets.

Although the concept of IWRM has been dominant for decades in the water resources management debate, it has been deeply criticized by authoritative experts, accused of vagueness and of being a "mere trendy label", whose success is due to the feeling of using a broader and holistic approach, suggesting the idea of a better management [Biswas, 2004]. The concept is not just charged with a naively innocent vagueness. Biswas [2008], criticizing specifically IWRM as defined by the Global Water Partnership (GWP), also accused the IWRM paradigm of lending itself to legitimize the centralization of competencies and power under the water authorities.

Indeed, the IWRM concept is very abstract. Its generality aims at encompassing the large heterogeneity of problems and situations you can meet in water systems; but then it seems to have little usefulness when brought to an operational level. Nonetheless, since the acceptance of the IWRM principle, many experts have worked to bring this concept into practice and this thesis can be considered part of this stream.

Conscious of the high level of connectivity among variables in water related problems, and the influence of water on other sectors, experts stress the need for a "multi-dimensional" approach. Under an organizational point of view, this requires the coordination and cooperation among the many different sectors and issues related to water management. However, this does not necessarily imply the simplistic (and generally wrong) solution of centralization under water authorities.

The word "integrated" suggests the thought of systems thinking. This all-connected-to-all approach was already present in the forefather of modern science, the Italian Galileo Galilei, who stated: *"you cannot pick up a flower without disturbing a star"*. Systems thinking can be defined as the process of understanding how things influence one another within a whole. In water management, this implies considering all water management dimensions (physics, economics, etc.) together, as opposed to the reductionist approach, in which the problem is broke down in pieces and solved by parts.

On the other hand, if systems are always interrelated, then there is no limit to integration, neither some threshold that defines whether a management is integrated or not. A shift of focus is needed here to avoid the fall of indistinguishability. In "the night where all the cows are black", to quote Hegel, and all the managements are integrated, then IWRM is indeed just a trendy label that "can be attached to the same thing people used to do" [Biswas, 2004].

A first answer to the critics of IWRM is that integrated is different from holistic [Mitchell, 2004]. Admitting that there is no limit to integration does not exclude that you can move to more integrated management. In this sense, integrated-integration is not an attribute of the management, but a process. Therefore, the focus should shift from the adjective "integrated", as a state of the management, to the verb "integrate", that draws attention to the action of integration.

The idea of integration as a process has been criticized for the absence of both a clear defined objective and a way to identify whether this has been reached or not [Biswas, 2004]. However, it is completely acceptable that IWRM has no content objectives, which depend from case to case and from stakeholder to stakeholder. The objective of IWRM is a "meta-objective". While the content objectives can change, IWRM states the permanency of some process-related objectives. Water professionals, working within the IWRM framework, have only process objectives. Similarly to a referee in a football match, who is not interested in the results itself. His role is to take care that the process (the match) follows some specific rules that ensure some "meta-objectives" (fairness of the result, show, and so on).

For this reason I reject the GWP definition of IWRM, stating that the scope of IWRM is "maximizing the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems". In this definition, which is one of the most cited ones, IWRM is framed within a specific value system. But integration is a technical issue, thus a process objective, whereas equity, sustainability, and efficiency are political values, thus content objectives. Objectives (thus values) should be derived from society. Scientist and technicians can partic-

ipate to a shift of values by providing new information, but have no legitimacy to define them.

In summary, integration is a tool, a process objective that is desirable because it enhances the achievement of better results, whatever they are. This statement, far from being a truism, is a strong assumption of the IWRM paradigm.

1.1.1 Models for IWRM

IWRM has been defined as "a journey, and not a destination" [Biswas, 2004]. Using the same metaphor, models are then the vehicles that help the Decision Maker to get the final destination that he desires to reach. The image of IWRM as a journey is in line with the stress on integration as a process, as just explained. The word *process* comes in fact from the latin *procedere*, which means to go forward, and models are a valuable instrument to go forward along the path of integration.

A model is a mathematical object that mimics reality in order to make predictions by quantifying some variables of interest. In the field of water management, these variables can be, for example, the water levels along a river or the concentration of some pollutant at a specific location. The power of models lies in their capacity to combine reductionist and systemic approaches. A physical system can be analyzed (the etymological meaning of analysis is *to cut apart, to divide in smaller pieces*). Once divided in smaller element, every sub-element is modeled by a mathematical object. Then, the sub-elements can be synthesized (whose etymology is *to put together, to combine*) by modeling the relations among the components, in order to reproduce the original system. If analysis is for science and understanding, synthesis is for integration. Integration can be enlarged to what we are able to model. For this reason mathematical models are a tool for integrated management.

Different water-related themes can be included in the same model and coordinated. For example, water quantity problems can be managed together with water quality, geomorphology, and/or ecological issues. *Integration in space* can be implemented by enlarging the model to encompass the up or downstream systems, as done by Overloop et al. [2010], or joining surface water and groundwater. Galelli et al. [2010] extended the system to encompass irrigation districts, and Harou et al. [2009] included the economic dimension. *Integration in time* can mean both integration among different moments, where *present and future* uses and conditions are analyzed, and integration between different time scales, where *long and short term* decisions are considered together. In the first case, the problem is balancing of resource use over time in order to find a satisfactory compromise, which is the concept of sustainability. The latter is the combination of decision at strategic and operational level.

For example, the strategic decision is whether to build up a new reservoir and the operational decisions regard its daily management. Both aspects are of paramount importance for the good management of any resource. This is further explored in section 1.3.

To represent water systems, we generally use dynamic models having the following general form.

$$x_t = f_t(x_{t-1}, u_t, d_t, \varepsilon_t) \tag{1.1a}$$

$$y_t = g_t(x_t, u_t, d_t, \varepsilon_t) \tag{1.1b}$$

$$c(x_t, u_t) \leq 0 \tag{1.1c}$$

In Equations (1.1), x are the states, holding the system "memory", u are the system inputs, or the controls, that we can choose and manipulate, and d the disturbances, on which we do not have any influence. ε is a vector of stochastic variables modelling the uncertainties. All the variables are vectors of the proper size, and t is the time index. Equation (1.6e) is the state equation and represents the system dynamics, written in an explicit, finite difference form. Equation (1.6f) is the output equation, in which y is the output vector. Inequalities (1.6h) are other physical constraints. Taking a reservoir as simple example, x is the water volume stored in the reservoir, u is the release, d the inflow from the upstream basin, y the reservoir water level, and t the daily time index. The constraints are in this case on state and control. The reservoir volume is in fact positively defined and the release capacity is limited between zero and some maximum value.

Models are used to build up knowledge providing the information on how the decisions u affects the output y, which are variables meaningful to the stakeholders. In this way they can assess the effects of alternative decisions u and value them [Liu et al., 2008; Loucks, 1992; Loucks et al., 1985]. This is, in summary, the added value of a model for decision making.

1.2 Optimization for IWRM

Models can be used to assess the effect of a set of alternative decisions for decision making. The set of alternatives U is made of all possible u. When this set is small, it can be assessed by a simple "what-if" analysis. For example, if the decision is whether building up a fixed weir or not, the decision variable is a binary one; studying the

effects of different configurations boils down to analyzing the two situations: with or without the structure.

When the decision space is large, then assessing all alternatives can become extremely complex. For example, the release from a reservoir at a specific moment cannot be decided alone but it must be coordinated with all the other releases along a horizon. In such a problem, there are as many decision variables as time steps in the horizon. The set of all possible alternatives is the cartesian product of the decision variables. This implies that the alternative dimension growths exponentially with the length of the horizon. It is very easy, in such a case, to run into a decision space larger than the human capacity to handle it.

At such level of complexity, optimization is a valuable tool. By setting an optimization problem, a large number of alternatives can be evaluated and coordinated [Soncini-Sessa et al., 2007].

Optimization for water management has also been criticized [Reed and Kasprzyk, 2009]. However, this was referred to the deterministic, single-criterion optimality, as a sort of "automatic" way to find a simple solution for a complex problem. It is unlikely that the authors neglect the added value in terms of meaningful information that optimization (and decision theory, in general) can provide as aid to the decision [Tsoukias, 2008].

Beyond the model of the physical process, Equation (1.1), optimization for decision making requires an *objective function* J, modelling the interests at stake.

$$J(\{y_t\}_{t=1}^{h}, x_h) = G(\{y_t\}_{t=1}^{h}) + g_h(x_h) \tag{1.2}$$

J is a functional that quantifies the "satisfaction" of the stakeholder, assigning a value to the trajectory of the system output, $\{y_t\}_{t=1}^{h}$. The second term, $g_h(\cdot)$, is the terminal condition, that accounts for the consequences of leaving the system at state x_h at the end of the horizon h. As an example, in a reservoir used for hydroelectric production, G is the value of produced energy, and g_h the value that can be produced after time h. Objectives are separable if they can be split in "sub-objectives" g_t, one per each time step [Soncini-Sessa et al., 2007]. In this case, Equation (1.2) can be written in the following form.

$$J(\cdot) = \sum_{t=1}^{h} g_t(y_t) + g_h(x_h) \tag{1.3}$$

Choosing a proper objective function is, in general, not an easy task. It requires the

rationalization of stakeholders' values. Defining J is in fact a participatory modeling process [Soncini-Sessa et al., 2007], in which human preferences are translated into a mathematical object.

A loss function, quantifying the "dissatisfaction" of the stakeholder, is assumed, in a rational framework, to be equivalent to the negative of the objective function. I stressed here the word *rational* because an important research line investigated how people tend to overestimate losses on gains, having a higher sensitivity to negative events [Kahneman et al., 1982]. This is a cognitive bias, though; for a rational decision maker there is no difference between a gained euro and a not-lost one. As a simple example of loss function, for a stakeholder interested in avoiding floods, $-J$ can be the probability of a flood event per year. In the following of this thesis, if J is to be minimized, it is a loss function, otherwise it is an objective function.

J has dimension R^M, where M is the number of the objectives included in the project. When the objectives are more than one, J is a vector having at least one element per group of stakeholders, and the optimization is a multi-objective one. This is a common case in water management, in which many stakeholders contend the resource on a shared system, each willing to manage the common resource for its own interest [Castelletti et al., 2008; Draper et al., 2003; Castelletti and Soncini-Sessa, 2007]. The inclusion of multiple interests in order to strike a balance[1] among the involved stakeholders is probably the most relevant dimension of integration.

1.3 Operational Water Management as Optimal Control Problem

Decisions can be divided in *strategic*, u_p, and *operational*, u_t. They differ in the time span of their effects, or, equivalently, the frequency at which they are taken. The first is a one-off decision, whereas the second is a recursive one. Their difference is not just in frequency, but also in nature. Recursive decisions can balance each other. If one has been slightly wrong, the next decision can compensate the error. This cannot be done, or it is much harder, for strategic decisions.

The relation between these two classes is that strategic decisions require the definition of an operational decision law, $m(\cdot)$. An operational decision law, or control law, is a function that, at each decision instant t, gives the decision u_t depending on the present state x_t.

---

1. Cultural note: When dealing with conflicting objectives that must be considered and balanced, Dutch has a more proper word, "overleggen", without equivalent in English. "Overleggen" conveys the idea of the process of meeting, discussing, consulting, thinking (carefully) about a solution, negotiating, agreeing on a compromise, and deliberating.

The optimal operation problem is nested within the optimal strategic decision, as described in Equation (1.4).

$$\min_{u_p} \left[\min_{m(\cdot)} J(\cdot) \right] \tag{1.4}$$

Modelling and optimization can be used for both strategic and operational decisions, and their coordination [Soncini-Sessa et al., 1999]. However, this thesis is about operational water management only, for example operating hydraulic structures, thus dealing with the optimal use of existing structures when strategic decisions are fixed. Operational water management is about solving a control problem, that is the recursive selection of control u_t to steer the system and bring it to the desired state. Optimal control is the optimal selection of u_t that minimizes the objective function J.

Solving the optimal control problem

Solving an optimal control means finding a control law for a given system (Equations 1.1) such that a certain optimality criterion (Equation 1.2) is achieved, i.e. solving the optimal control problem (1.5).

$$\min_{m(\cdot)} J(\cdot) \tag{1.5a}$$

$$\text{System dynamic (Equations 1.1)} \tag{1.5b}$$

$$\text{Other constraints} \tag{1.5c}$$

The method to solve problem (1.5) depends on the specific problem characteristics [Labadie, 2004; Sahinidis, 2004; Reznicek and Cheng, 1991]. Dynamic Programming (DP) (and its stochastic version, Stochastic Dynamic Programming, SDP) [Bertsekas, 1995] offers a brilliant and elegant solution for the optimal control of dynamic systems. DP is an off-line optimization; its solution gives a optimal control law, indicating the optimal control value to apply to the system in function of the present state. Based on the Bellman equation, DP splits costs in present-costs and cost-to-go, and finds a functional solution solving iteratively the equation backwards.

While DP is, from a theoretical point of view, a superb object, practical applications reveal its limits. It requires in fact variables discretization and separability of the objective function, thus written in form of Equation (1.3). The main limit of DP, however, is that the calculation time increases exponentially with the number

of variables. Using a evocative name, this drawback has been called the "curse of dimensionality" [Bellman and Dreyfus, 1966]. Practical applications of DP are thus limited to simple systems. To find a solution using DP, large systems must be reduced to a model having a small number of states, and these necessary approximations lead then to suboptimal solutions.

In summary, even if the optimal control law obtained with DP has the form of a perfect theoretical framework, another formulation of the optimal control problem, much leaner and simpler, has become popular for practical real time applications. This is the argument of the next section.

1.4 Model Predictive Control for operational IWRM

Model Predictive Control (MPC) solves an open-loop optimal control problem at every control instant along an finite horizon, applying to the real system only the first control value. At the following instant the horizon is shifted and the optimal control problem reformulated. For this reason MPC is also called *receding horizon control*, which is less used but conveys better the essence of MPC characteristics [Morari et al., 1999; Camacho and Bordons, 2004; van Overloop, 2006].

At each decision time step, MPC solves the following optimization.

$$\min_{\{u_t\}_{t=1}^{h}} J(\{y_t\}_{t=1}^{h}, x_h) \tag{1.6a}$$

subject to

$$\text{Initial conditions} \tag{1.6b}$$

$$x_0 \tag{1.6c}$$

$$\text{System model} \tag{1.6d}$$

$$x_t = f_t(x_{t-1}, u_t, d_t) \tag{1.6e}$$

$$y_t = g_t(x_t, u_t, d_t) \tag{1.6f}$$

$$\text{Other Constraints} \tag{1.6g}$$

$$c(x_t, u_t) \leq 0 \tag{1.6h}$$

MPC solves the open-loop optimal control problem as in (1.6) in order to find an optimal control signal $\{u_t^*\}_{t=1}^{h}$ on a finite control horizon h. From this control signal, only u_1^* is applied to the real system. In operational use, MPC works as follows:

1. obtain measurements/estimates of the system states x_0 and, when available, disturbances $\{d\}_{t=1}^{h}$
2. solve optimization problem as in problem (1.6)
3. implement the first part the optimal control signal until new measurements/estimates are available
4. repeat from 1

MPC generally uses a linear model and a quadratic loss function (distance from a desired trajectory) [Camacho and Bordons, 2004]. More recently, nonlinear models have been employed [Henson, 1998], though, the objective function remains almost always a quadratic distance from a trajectory. The research activity on MPC is still very active, and the existing literature is vast. The reader interested in the argument can read the good summary written by Morari et al. [1999], or the good introduction by Maciejowski [2002].

Compared to simple feedback controllers, such as Proportional Integral Delay (PID) [Astrom, 1995], MPC has the advantage to explicitly include constraints, even if it requires the modeling of the physical process and the "tuning" of an objective function definition. MPC shows clear advantages in dealing with systems with delay, where reactive rules can lead to instability. According to Qin and Badgwell [2003], MPC success is mainly due to its explicit constraint handling capability, including constraints in the optimization problem, enabling the controller to predict future constraint violations and respond accordingly [Henson, 1998].

Mayne et al. [2000] points out that, under some conditions, MPC control law is implicitly equivalent to the optimal feedback control law, obtained using DP. The difference is that, in MPC, $m(\cdot)$ is not pre-computed off-line for all possible values of x, but it is found in real time at each control instant, for the present x_0 only, where the hypothetical control law for MPC is $m_{MPC}(x_0) = u_1^*|x_0$, and $u_1^*|x_0$ is the first time step of the optimal control signal given x_0.

In problem 1.6, the argument of the optimization is a control series. Therefore problem (1.6) is much simpler to solve than problem (1.5), where the argument of the optimization is the entire control law, that is a function. This allows the control of larger systems. According to Mayne et al. [2000], the ability to handle control problems where off-line computation of a control law is difficult or impossible to compute is the *raison d'etre* of MPC.

The finite control horizon is at the same time the strength and the weakness of MPC. The horizon finiteness makes it easier to find a solution in real time. On the other hand, all the effects from the end of the horizon h onwards are not taken into account. Whether these effects are relevant to the the first controls that will be applied

depends on the specific physical system and objective function. Experience shows that, for water systems, this can be an issue, especially for reservoir management. This is further discussed in the conclusions, in Chapter 5.

1.4.1 Proactive operational water management using forecasts and the problem of uncertainty

Scientist and engineers have worked to build up meteorological and hydrological models to predict the system dynamic ahead of time. The output of such models are forecast trajectories of the variable of interest, $\{d_t\}_{t=1}^{h_f}$, where h_f is the forecasting horizon, spanning from some hours to a few days ahead. Forecasts of the upstream system (including weather forecasts) that are inputs to the controlled system are valuable information for control.

Use of forecasts makes the control *proactive* [Zavala et al., 2009], which is defined as the capacity to set the system to a state that best accommodates the incoming disturbances. Adding a deterministic input does not add any complexity to the MPC scheme, then forecasts can easily be integrated in it. This is, in general, an advantage, but it can bring along some risks. Forecasts introduce in fact hydrological and meteorological uncertainty in the control problem. These uncertainties are typically much larger than those of the controlled hydraulic systems. The scope of this thesis is finding methods to enhance the control robustness using existing forecasts.

Uncertainty affects negatively the control performance. Disregarding the presence of uncertainty, thus treating the problem as deterministic, leads to a departure from the optimal results. The distance from optimality depends on the specific problem and grows with the distance from *certainty equivalence* conditions. Certainty equivalence is guaranteed when i) costs are evaluated by quadratic functions; ii) system dynamics are linear; iii) there are no inequality constraints; and iv) uncertain inputs are independent and normally distributed. Under the certainty equivalence condition, deterministic and stochastic optimization give the same results. Systems departing from certainty equivalent conditions are more sensitive to wrong controls [Van de Water and Willems, 1981]. Philbrick and Kitanidis [1999] explored how much, departing from certainty equivalency, deterministic solutions give worse results, noticing that few actual water systems are close to certainty equivalence.

Uncertainty introduces the issue of robustness, which, in control theory, has been intensively studied. A system is robust when "stability is maintained and that the performance specifications are met for a specified range of uncertainty range" [Morari et al., 1999] [2]. In water systems, dealing generally with stable or meta-stable systems,

2. Also within the water management community, the word "robust" implies the explicit treatment

stability is generally not a problem. But performances are, being the system used for human purposes. Robustness is not a property of the system itself; it requires, by definition, the selection of some performance specifications under some uncertainty range.

MPC is a deterministic optimization that does not deal with uncertainty in an explicit way. Despite this, MPC has some "inherent robustness" [Mayne et al., 2000; De Nicolao et al., 1996; Magni and Sepulchre, 1997] linked to the feedback mechanism given by iterative reformulation of the optimal control problem. MPC is, in fact, an open loop optimization, but only the first control values are applied. This limits the performance loss due to uncertainty. Therefore, for some systems, under defined uncertainties and robustness requirements, solving the nominal control problem is sufficient for satisfactory results.

In other cases, effects of uncertainty are larger and unacceptable. In these cases, MPC's inherent robustness is not sufficient. Control theory research has developed many ways for synthetic robustness [Bemporad and Morari, 1999], which means enhancing the performance, given a range of uncertainty. Different works are dedicated to this task [Muñoz de la Peña, 2005; Löfberg, 2003; Batina, 2004], and solutions are generally tailored to the specific problem at hand.

This thesis presents some new methods for synthetic robustness that have been developed specifically for the operation of water systems when uncertainty lies not within the controlled system but in the forecasted inflow. The thesis can be divided in two parts. The first one treats a case where the uncertainty model was not defined, where we show an innovative method to build up a distribution around the punctual output of a deterministic hydrological model. This is the content of Chapter 2. The second part contains a method to use the forecast in MPC when it comes in form of an ensemble [Leutbecher and Palmer, 2008]. This is shown in Chapters 3 and 4.

---

of uncertainty [Watkins Jr and McKinney, 1997].

Chapter 2

Predictive uncertainty by absolute errors dynamic modelling

Setting up a stochastic optimal control problem requires a probabilistic hydrological forecasts. Nevertheless, many existing hydrological models are deterministic and provide point estimates of the variable of interest. Often, the model residual error is assumed to be homoscedastic, however practical evidence shows that the hypothesis usually does not hold.

We propose a simple and effective method to quantify predictive uncertainty of deterministic hydrological models affected by heteroscedastic residual errors. It considers the error variance as a hydrological process separate from that of the hydrological forecast, and therefore predictable by an independent model. The variance model is built up using time series of model residuals and, under some conditions on the same residuals, it is applicable to any deterministic model. Tools for regression analysis applied to the time series of residual errors, or better their absolute values, combined with physical considerations upon the hydrological features of the system, can help identifying the most suitable input to the variance model and the most parsimonious model structure, including dynamic structure if needed. The approach has been called Dynamic Uncertainty Modelling By Regression on Absolute Errors (DUMBRAE) and is demonstrated by application to two test cases, both affected by heteroscedasticity but with a very different dynamics of uncertainty. Modeling results and comparison with other approaches, i.e a constant, a cyclostationary, and a static model of the variance, confirm the validity of the proposed method.

2.1 Introduction

Deterministic models have been widely used in hydrology for both forecasting and simulation purpose. The assessment of their uncertainty is a major research issue. For

managers and decision-makers, quantification of uncertainty associated to model estimates is a valuable information for both alarm and operational purposes. From the modeler point of view, it provides indications for model diagnosis and improvement [Gupta et al., 2008; Reichert and Mieleitner, 2009], and targeted data collection. Theoretically, since the forecast value by a deterministic model will never be exact, associating it with some kind of characterization of its error is the only way to assess its quality [Weijs et al., 2010]. The lower the error the better the model: this is what we implicitly do in practice every time we use a deterministic model, since we regard the model prediction as the expected, or most probable value of the estimated variable. Quantification of model uncertainty makes this assumption explicit and provides the model user with a more formal and accurate evaluation of the residual error.

Sources of uncertainty in deterministic models are often classified in measurement errors, both in the input and output, uncertainty in the parameters and uncertainty in the model structure, including selection of the input and the mathematical relation between input, state and outuput variables. Many approaches to uncertainty assessment rely on such decomposition. One or more sources of uncertainty are given a statistical description and uncertainty is propagated in the model via random sampling and simulation, to obtain a sample or distribution of model predictions in place of a single value (see for example Thyer et al. [2009]; Kuczera and Parent [1998]; Kavetski et al. [2006]). The application of these methods may require collecting several information for the statistical characterization of the different uncertainty sources (e.g. the accuracy of the measurement devices, the analysis of the error induced by data pre-processing), and can be limited by the computational cost of model simulation.

On the other hand, "model residual" approaches skip any distinction of uncertainty sources and directly analyze the time series of model residuals to build a model of the global predictive uncertainty, which is often sufficient for practical purposes. The drawback of model residual approaches is that, while they do not require assumptions about the different sources of uncertainty, they usually do for the characterization of the model residual. Historically, and especially in model calibration, the most common approach is to assume that the residual be an independent identically distributed process, usually zero-mean and Gaussian. The approach has been widely criticized because most of these assumptions are violated in hydrological applications, especially autocorrelation and homoscedasticity of the model residuals [Sorooshian and Dracup, 1980]. Many methods have been proposed either to manipulate model residuals so that they satisfy such assumptions (e.g. using Box-Cox

transformations [Box et al., 1970], see [Kuczera and Parent, 1998; Bates and Campbell, 2001], or the normal quantile transform as in Montanari and Brath [2004]) or to relax some of these assumptions (e.g. Romanowicz et al. [2006]; Schaefli et al. [2007]; Schoups and Vrugt [2010]).

We present here a novel "model residual" approach for estimating the predictive uncertainty of deterministic hydrological models. In our approach, we will assume that the residual error of the model be uncorrelated in time or that it can be described by an autoregressive model with uncorrelated residual. The assumption is very useful because the statistical description of the residual process reduces to providing a sequence of marginal probability distribution functions (pdfs). We use the same distribution type at all time steps while allowing for the residual variance to change in time and reproduce the heteroscedastic behaviour that is often observed in hydrological time series. For several practical reasons that will be clarified throughout the chapter, we will use Gaussian distributions, although the approach can be extended to other distributions if Gaussian proved to be unsatisfactory, provided that they are symmetric. Under these hypotheses, the identification of the residual error pdf is reduced to the estimation of the error variance (or standard deviation). The latter can be a function of time (the season) or other hydro-meteorological inputs, depending on the case study under exam.

The novelty of our approach is that we will not assume a priori the input variables of the variance model nor the type of relation between these inputs and the variance, but rather we will infer such information from data analysis and consideration of the features of the hydrological system under exam. This is possible if one regards the variance model identification as a regression analysis over the time series of the model residual errors. Further, we will show that under the Gaussian assumption it is possible (and numerically more efficient) to identify a model of the error standard deviation from the time series of absolute errors, rather than a model of the variance from the time series of squared errors. Another contribution of our work is that we will show the effectiveness of introducing past absolute errors among the input of the standard deviation model, which means that the error standard deviation is modelled as a dynamic process. For all these reasons, we named the proposed approach *Dynamic Uncertainty Model By Regression on Absolute Error* (DUMBRAE).

This chapter is organized as follows. In the following section, the DUMBRAE approach is fully described from the methodological standpoint. Then the issue of how to evaluate the quality of an uncertainty model is discussed, from visual inspection of the inferred confidence bounds to formal methods. Relying on these results we

can demonstrate the effectiveness of the proposed approach through the application to two case studies.

2.2 Methodology

We consider a deterministic model

$$\hat{y}_t = f(\mathbf{x}_t, \boldsymbol{\lambda}) \tag{2.1}$$

that provides the flow forecast $\hat{y}_t$ as a function of several inputs collected in the vector $\mathbf{x}_t$ and a parameter set $\boldsymbol{\lambda}$. Since the forecast is affected by multiple sources of error, including measurement error in the input $\mathbf{x}_t$, error in the model parameters $\boldsymbol{\lambda}$ and structure $f(\cdot)$, the actual flow will be given by

$$y_t = \hat{y}_t + r_t \tag{2.2}$$

where r_t is the model residual error. Equation (2.2) implicitly assumes that the model output can be univocally split into two mutually exclusive components, deterministic and random. This dichotomy is questionable (see discussion in Koutsoyiannis [2009]) but it is very useful from the operational standpoint because it allows one to separate the identification of the hydrological model (2.1) and that of the uncertainty model, and define one common strategy to uncertainty modelling that can be applied to any pre-calibrated hydrological model independently from its structure.

The residual error r_t can be described by its probability distribution function (pdf) and the predictive uncertainty of model (2.1) is derived from such pdf. For instance, confidence bounds can be obtained by adding the quantiles of the error pdf to the flow forecast $\hat{y}_t$. The most common approach is to introduce some hypothesis about the pdf shape, and identify the pdf parameters from the residual time series. Alternatively, Solomatine and Shrestha [2009] propose a method to derive the error quantiles for given degree of confidence (and ideally the entire error pdf) without making any a priori assumption.

Traditionally, the residual error is assumed to be independent, identically distributed, zero-mean and Gaussian, i.e. it is assumed that the error pdf be $N(0, \sigma^2)$ for all t. The approach is often unsatisfactory since in hydrological time series all these assumptions are rarely satisfied, and many works in the literature aim at relaxing some of them.

Violation of the assumption of independence means that the deterministic model (2.1) produces systematic error, for instance because it neglects some of the processes (snowmelt, evapotranspiration, ...) occurring in the system. However, when model (2.1) is used in prediction mode, the problem can be easily overcome by modelling the model residual r_t as an autoregressive process of order q,

$$r_t = \boldsymbol{\alpha} \cdot \mathbf{v}_{t-1}^T + e_t = [\alpha_1, ..., \alpha_q] \cdot [r_{t-1}, ..., r_{t-q}]^T + e_t \tag{2.3}$$

whose residual e_t satisfies the independence assumption provided that a sufficiently large value of q is used. Identification of model (2.3) is straigthforward as ordinary least squares can be used to estimate the parameters α_i. Once model (2.3) is available, the flow forecast is corrected as

$$\hat{y}_t' = \hat{y}_t + \boldsymbol{\alpha} \cdot \mathbf{v}_{t-1}^T \tag{2.4}$$

and its predictive uncertainty is given by projecting the pdf of e_t, which is independent by construction.

Independence is a valuable property because it allows describing e_t by a marginal distribution $p_t(\cdot)$, independently of the error distribution at previous or following time step, i.e.

$$e_t \sim p_t(\cdot; \boldsymbol{\eta}_t) \qquad t = 1, 2, ...$$

where $\boldsymbol{\eta}_t$ is the parameter vector for the distribution family $p_t(\cdot)$. However, this means that the distribution $p_t(\cdot)$ must be identified and evaluated based on one data only, since we have one observation per time step. Sometimes, the error histogram is used to infer the shape of the error pdf. However, this is not justified if we assume that the pdf be different at each time step, because each datapoint should be considered as an extraction from a different distribution. The QQ plot (or probability plot, see section below) is a tool to assess the overall fit of a sample of data $e_1, e_2, ..., e_N$ against a sample of different distributions $p_1, p_2, ...p_N$, however it can be used only after all the distributions have been identified. Therefore, in our analysis we will assume a given distribution family for each time t, identify all the distributions (i.e. estimate the parameters $\boldsymbol{\eta}_t$ for all t), and finally we will test our choice a posteriori by means of the probability plot. With this approach, the most flexible distribustion possible should be assumed so that the shape of the pdf is not constrained a priori.

In the absence of other information, the error is assumed to be zero mean and to follow the same distribution at all time steps, while its standard deviation σ_t is let

vary in time, thus accounting for the heteroscedasticity observed in the error time series,

$$e_t = \sigma_t \tilde{e}_t, \qquad \tilde{e}_t \sim p(\cdot; \boldsymbol{\eta}) \tag{2.5}$$

In general, a sequence of random variables is said to be heteroscedastic if the random variables have different variances. Regression analysis in the presence of heteroscedasticity has been widely studied in econometrics [Engle, 1982]. In the hydrological context, variability of the standard deviation σ_t is usually related to flow conditions: the higher the flow, the higher the variance. For instance, Schaefli et al. [2007] assume that the error distribution $p(\cdot)$ be Gaussian with zero mean and that the standard deviation takes one out of two possible values corresponding to two different hydrological condition, low or high flow. Schoups and Vrugt [2010] use a Skew Exponential Power (SEP) density as error pdf $p(\cdot)$, and assume that the error standard deviation σ_t be linearly related to the flow forecast $\hat{y}_t$, based on the evidence that predictive uncertainty increases at higher flow.

In our approach we will demonstrate that the variability of the standard deviation σ_t can be effectively related also to other hydrological inputs, not necessarily flow, and we will provide a general method to infer what variables are most significant and how to estimate their relation to σ_t. By application of our method to the proposed case studies we will show that improving the model of the standard deviation can significantely enhance the uncertainty model, even under simple assumptions about the error pdf $p(\cdot)$ (e.g. Gaussian). In general, we will assume that σ_t be given by some relation of the form

$$\sigma_t = g(\mathbf{z}_t, \boldsymbol{\phi}) \tag{2.6}$$

where $\mathbf{z}_t$ is a vector of suitable input variables, not necessarily flow but possibly also other variables like, for instance, precipitation; and $\boldsymbol{\phi}$ is a parameter vector.

The model identification thus encompasses (i) modelling the variance, i.e. identifying relation (2.6); and (ii) estimating the (stationary) pdf $p(\cdot)$ appearing in (2.5). The model identification criterion will be the maximization of the likelihood function L which, under the independence assumption and equations (2.5),(2.6), takes up the form

$$L(\ \mathbf{e}\ |\ \boldsymbol{\eta}, \boldsymbol{\phi}\ ) = \prod_{t=1}^{N} p\left(\frac{e_t}{\sigma_t}; \boldsymbol{\eta}\right) = \prod_{t=1}^{N} p\left(\frac{e_t}{g(\mathbf{z}_t, \boldsymbol{\phi})}; \boldsymbol{\eta}\right) \tag{2.7}$$

Notice that, when Gaussian distribution is used, the error pdf $p(\cdot)$ is fully described by its mean and standard deviation and the parameter vector $\boldsymbol{\eta}$ disappears from equation (2.7). On the other hand, the problem can be enlarged to encompass also

the calibration of the original hydrological model (2.1), or (2.4). In this case, the likelihood function becomes

$$L(\ \mathbf{e}\ |\ \boldsymbol{\eta}, \boldsymbol{\phi}, \boldsymbol{\lambda}, \boldsymbol{\alpha}\ \ ) = \prod_{t=1}^{N} p\left(\frac{\bar{y}_t - f(\mathbf{x}_t, \boldsymbol{\lambda}) - \boldsymbol{\alpha} \cdot \mathbf{r}_t^{\boldsymbol{T}}}{g(\mathbf{z}_t, \boldsymbol{\phi})}; \boldsymbol{\eta}\right) \tag{2.8}$$

Notice, however, that Equation (2.8), requiring the past observed residuals r_t, can be used in real time forecasting mode only, which is the use that, in this thesis, we are interested in. In the remaining of this chapter we will focus on the case when the hydrological model has already been calibrated, and use Gaussian distribution as the error pdf $p(\cdot)$. Under this assumption, the likelihood function is (2.7), where the pdf parameter vector $\boldsymbol{\eta}$ is dropped. Then, the problem boils down to identifying the standard deviation model (2.6), that is, selection of the input variables $\mathbf{x}_t$, choice of the class function $g(\cdot)$ and estimation of the parameters $\boldsymbol{\phi}$. These topics will be discussed in the following section.

2.2.1 Identification of the standard deviation model

In our approach we propose not to fix a priori the relation between the error standard deviation and the hydrological inputs, but rather to infer it from data analysis and consideration of the specific features of the case study under exam. This is possible because the dynamics of σ_t is revealed by the time series of the residual error. In fact, by definition the error variance σ_t^2 is given by $E[(e_t - \mu_t)^2]$ and thus, if the error is zero mean, $\sigma_t^2 = E[e_t^2]$. Identifying the model of the variance σ_t^2 can be viewed as a regression analysis problem where the time series to be modelled is the time series of squared errors.

However, modelling the time series of squared error can be difficult because the operation of squaring emphasizes high values and increases the distance between high and low errors, so that the resulting time series is difficult to model. If the error pdf is Gaussian, it is possible to directly relate the standard deviation σ_t to the absolute error $|e_t|$, which is a much smoother time series. In fact, under assumption of symmetric distribution, it can be demonstrated that the standard deviation is linearly proportional to the absolute error

$$\sigma_t = k\ E[|e_t|] \tag{2.9}$$

where the proportionality coefficient equals $k = \sqrt{2\pi}/2$ in the case of Gaussian distribution (see Appendix for proof). Therefore we can directly identify model (2.6) by regression analysis over time series of absolute error. Selection of the model input $\mathbf{z}_t$ can be based on analysis of the autocorrelation function between absolute error

and candidate input variables like flow forecast, past observations of flow or meteorological variables (precipitation, temperature, etc.). Other data analysis tools that may prove useful for input selection include scatterplots, cluster analysis, mutual information, etc. For a review of available methods for input variable selection and their application in the hydrological context, see Bowden et al. [2005]. As it will be shown in the application case studies, data analysis is also guided and supported by the knowledge of the system characteristics and the model features.

Once the model input have been chosen, model identification requires choosing the class function $g(\cdot)$ and parameter estimation. In this work, we will follow a parsimonious approach and start by simple relations, e.g. linear, and move to more complex ones only if the modelling results are not satisfactory.

As for parameter estimation, the problem is spontaneously formulated as maximizing the likelihood function (2.7). Notice that even if the error pdf $p(\cdot)$ is Gaussian and the standard deviation model (2.6) is linear, both the likelihood function and its logarithm are nonlinear in the parameters $\boldsymbol{\phi}$, which can make the maximum likelihood approach computationally demanding. The problem is further complicated by the fact that parameter values should be constrained to guarantee that the resulting standard deviation be positive for all possible values of the input vector $\mathbf{z}_t$.

Alternatively, a more straightforward approach to parameter estimation is to minimize the mismatch between the output of model (2.6) and the observed absolute error, e.g. using Euclidean distance

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \sum_{t=1}^{N} (k\ |e_t| - g(\mathbf{z}_t, \boldsymbol{\phi}))^2 \tag{2.10}$$

The advantage is that when $p(\cdot)$ is Gaussian and $g(\cdot)$ is linear, the fast and efficient linear least squares solution can be used. However, the criterion underlying the solution (2.10) does not reflect the true modelling scope, which is to identify the error pdf and not to interpolate the absolute error time series. Nonetheless, the solution (2.10) can be effectively used as the initialization of the recursive nonlinear optimization approach used to solve the maximum likelihood problem.

2.2.2 Model evaluation

The identified model of the standard deviation will be evaluated in terms of likelihood value L, or equivalently, negative log-likelihood $-\log(L)$, evaluated over a validation dataset different from the one used for model calibration. The use of L

as an evaluation indicator is theoretically justified by the likelihood principle, which states that, for given model structure, all information that the data contains about the model parameters is in the likelihood function [Jaynes and Bretthorst, 2003]. Also, the likelihood is the only scoring rule which is unambiguous, local and proper [Weijs et al., 2010]. Other evaluation indicators that will be used are the Akaike Information Criterion AIC= $2n_\phi - \log(L)$, which balances the model log-likelihood and the model complexity, as measured by the number n_ϕ of model parameters; and the proportion of observations within the confidence limits (or prediction interval coverage probability, PICP [Solomatine and Shrestha, 2009]).

A closer insight into the behaviour of the standard deviation model will be given by visual inspection of the associated confidence bound, as well as the QQ plot. The QQ plot, or probability plot [Laio and Tamea, 2007; Thyer et al., 2009], is a graphical tool for assessing the goodness-of-fit of a sample of probability distributions against a sample of data. In our application, since the error is assumed additive and Gaussian, the conditional probability distribution of the flow forecast $\hat{y}'_t$ is Gaussian with standard deviation equal to the error standard deviation computed by model (2.6). The goodness-of-fit of such distribution cannot be assessed by conventional statistical tests, because only one extraction from that distribution is available, i.e. the measurement y_t. However, from the probability integral transform it follows that if the estimated cumulative distribution function (cdf) $\hat{F}_{y_t}(\cdot)$ of y_t coincides with the true cdf $F^0_{y_t}(\cdot)$, then the value $u_t = \hat{F}_{y_t}(y_t)$ is an extraction from a uniform distribution over $[0, 1]$. Since this is true for any $t = 1, ..., N$, N being the number of flow measurements, one can evaluate the goodness-of-fit of the N cdfs $\hat{F}_{y_t}(\cdot)$ by checking if $U = \{u_1, ..., u_N\}$ is a sample of mutually independent, uniformly distributed observations. Independence of the sample can be checked by looking at the autocorrelation function. As for the uniformity hypothesis, we compute the value of the empirical cdf of u_t as $F^0_{u_t} = R_t/N$, where R_t is the number of elements in U lower than u_t, and compare it with the value of the uniform cdf, $u^0_t = F^0_{u_t}(u_t) = u_t$.

2.3 Application to the simulation model of Rhone River, Switzerland

Rhone River is located in a high mountainous catchment, where the hydrological regime is strongly affected by glacier and snowmelt, with peak flows in summer (July-August) and a low flow period in winter (February-March). The simulated discharge is generated using a semi-lumped conceptual glacio-hydrological model, described in details in Schaefli et al. [2005]. The model has two levels of discretization. The

first one distinguishes between catchment areas covered and not covered by ice. Accumulation and melting of snow and ice in the ice-covered area are modelled by two parallel linear reservoirs. The not-covered area is modelled by a linear reservoir for the slow contribution of the soil underground water and a nonlinear one for the direct runoff. The second level of discretization is among a set of elevation bands: runoff discharge is computed separately for each band and then aggregated. No routing among the components of the model is considered because the runoff delay is much smaller than the modeling time step, which is 24 hours. Time series for model identification (precipitation, temperature and potential evapotranspiration) cover the period from 1981 to 1984. Data from 1990 to 1994 are used for validation.

The error of this conceptual model is correlated in time, as already discussed in Schaefli et al. [2007], and it can be described by a first order autoregressive model. The residual error after adding the autoregressive component is

$$e_t = y_t - \hat{y}_t - \alpha r_{t-1} \tag{2.11}$$

where y_t is the measured runoff, $\hat{y}_t$ is the hydrological model prediction at time t, and r_{t-1} is the residual of the hydrological model at previous time step, $r_{t-1} = y_{t-1} - \hat{y}_{t-1}$.

Figure 2.3 shows that although the residual e_t is independent from its previous values, its squared value e_t^2 have a significant autocorrelation. This is an evidence of heteroscedasticity [Engle, 1982]. The absolute error $|e_t|$ shows an even higher autocorrelation. As anticipated, squaring the error reduces the smoothness and thus the autocorrelation of the time series. For this reason we will focus on the absolute error time series and use it to derive a model of the error standard deviation.

2.3.1 Model identification

Potential inputs of the standard deviation model are all the inputs and the output of the hydrological model, i.e. precipitation p_t, temperature T_t and runoff forecast $\hat{y}_t$, as well as the absolute value of the last observed residual $|e_{t-1}|$ and, since temperature mainly affects the flow through snowmelt, temperature value when positive, $T_t \cdot H(T_t)$, where $H(\cdot)$ is the unit step function.

Table 2.1 reports the crosscorrelation value between different candidate input variables and the absolute error. It shows that the precipitation is weakly related with the absolute value of the residual, whereas all the other inputs have a significative correlation. Weak correlation between precipitation and error can be justified by the fact that precipitation includes both snow and rain undistinguished. Predicted runoff has a slightly higher correlation than the other variables.

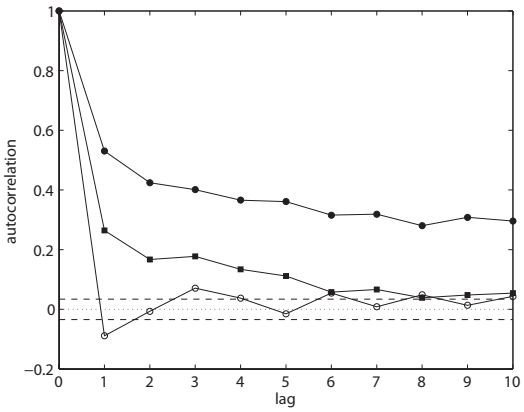

Figure 2.1: Autocorrelation of residual error e_t (white circles), squared error e_t^2 (squares) and absolute error $|e_t|$ (black circles) for the Rhone River case study (calibration data set).

Following this data analysis, we selected as the input variable the lagged absolute error, which accounts for the slow dynamics of the standard deviation; an exogenous component identified in the predicted runoff $\hat{y}_t$; and a constant term, to avoid too small variance values during low flow periods. According to the parsimonious modelling approach, we started from a linear relation between these variables, while relaxing the lag-one assumptions and considering also absolute error and flow forecast values in previous time intervals. The model is thus an ARX (Auto-Regressive eXogenous input) model and takes the following form:

$$\hat{\sigma}_t = a + \sum_{i=1}^{n} b_i |e_{t-i}| + \sum_{i=1}^{m} c_i \hat{y}_{t-i} \tag{2.12}$$

The model order n and m is chosen by trial-and-error. Parameters a, b_i and c_i are estimated using criterion (2.10) because $g(\cdot)$ is linear and (2.10) becomes numerically efficient. However, the optimization should be constrained such that $\hat{\sigma}_t$ can take up positive values only. If the domain of all inputs is real positive, as in this case,

Table 2.1: Crosscorrelation value between different candidate input variables and the absolute error (Rhone River, calibration period 1981-1984).

	input				
	p_t	T_t	$T_t \cdot H(T_t)$	$\hat{y}_t$	$\lvert e_{t-1} \rvert$
corr(input,$\lvert e_t \rvert$)	0.116	0.456	0.511	0.588	0.530

positivity of all the parameters is a sufficient condition for the positivity of the model output. The calibration procedure thus starts by applying unconstrained linear least squares and, if the solution does not satisfy the positivity constraint, uses such solution as the starting point for an iterative constrained least-squares procedure (*lsqnonlin* function in the Matlab Optimization Toolbox)

2.3.2 Model evaluation

The Dynamic Uncertainty Model By Regression on Absolute Error (DUMBRAE) of equation (2.12) is here compared with a constant, a mixture of normal distributions (here called *two-mixed*) and a periodic model. The constant model assumes that the standard deviation of the prediction error be constant in time and equal to the sample standard deviation. For this test case, this model was already recognized as unsatisfactory by [Schaefli et al., 2007], who proposed (ibidem) a model mixture of two normal distributions to treat heteroscedasticity using two different values of the variance, one smaller variance for the more predictable low flow and one higher variance for the high flow. The rule for switching from one to another is based on some conditions on the predicted runoff. For a more complete benchmark we also developed a simple periodic model that estimates the standard deviation as a function of time and is identified using a Fourier series expanded until a (small) finite number of harmonics. The model is described in details in appendix B.

Table 2.2 reports the values of the evaluation indicators over the calibration and validation period by the DUMBRAE model (2.12) with $n = 1$ and $m = 1$, and the three benchmarks. It shows that DUMBRAE is the best of the evaluated set, since it has the minimum value of $-\log(L)$ and AIC, and the PICP closest to 0.95. As expected the constant variance model is extremely poor. The periodic model proves better over the calibration data set but in validation the negative log-likelihood goes to infinity. A closer insight into data shows that the periodic model performs well in many cases (this is also showed later in the QQ plot) but for two events that fall in the tail of the estimated pdf and thus are associated with extremely small probability value. This is because the periodic model follows the seasonal trend but

Table 2.2: Comparison of different standard deviation models (Rhone River): n_ϕ is the number of model parameters; $-\log(L)$ is the negative log-likelihood; AIC is the Akaike Information Criterion; PICP is the prediction interval coverage probability at level of confidence $\alpha = 0.95$.

Model	n_ϕ	$-\log(L)$		AIC		PICP	
		cal	val	cal	val	cal	val
Constant	1	6.02×10^3	3.34×10^3	10.9×10^3	6.02×10^3	0.946	0.940
Periodic	7	2.26×10^3	Inf	4.54×10^3	Inf	0.929	0.919
Two-Mixed	3	4.68×10^3	2.05×10^3	8.42×10^3	4.25×10^3	0.952	0.962
DUMBRAE	3	1.71×10^3	1.36×10^3	3.60×10^3	2.57×10^3	0.953	0.958

it does not use real-time information. DUMBRAE instead, even if traced by a simple ARX relation, can dynamically adjust to the increase and decrease of the variance.

Figure 2.3.2 shows the 95% confidence interval in time of the four benchmark models for the first year of the calibration period (1981). It can be seen that the periodic and the DUMBRAE model have more or less the same trend, which follows relatively well the observed values of the residuals (black points). The two-mixed and the constant model provide an unnecessary large confidence band. The two-mixed model, having only a binary configuration (high-low uncertainty), can only partially modulate the variance. The constant variance model is completely inflexible, and the estimation of its average value is dominated by few, large error values. Notice that because the error distribution is assumed to be Gaussian and thus have an infinite support, negative flows are given a non-zero probability and the lower quantile of the flow can be negative. In principle this is not acceptable and may be a reason for abandoning the Gaussian assumption. In practice, however, the problem may be overcome by simply setting to zero the negative quantiles. In this application, the DUMBRAE model did not produce any negative flow quantile over the calibration horizon nor the validation horizon while all the other benchmark models did, especially the constant model. The reason for this good property of DUMBRAE is that it reduces the standard deviation in correspondence to small forecast values (see equation 2.12), thus keeping the lower quantile very close to the forecast when the latter is close to zero.

Finally, Figure 2.3) shows the QQ plots corresponding to the four standard deviation models over the calibration and validation dataset. It shows that none of the empirical cdfs, as estimated by DUMBRAE and the benchmark models, lays exactly on the bisector, i.e. none of them coincides with the theoretical cdf, however the empirical cdf corresponding to the periodic and DUMBRAE model are the closest

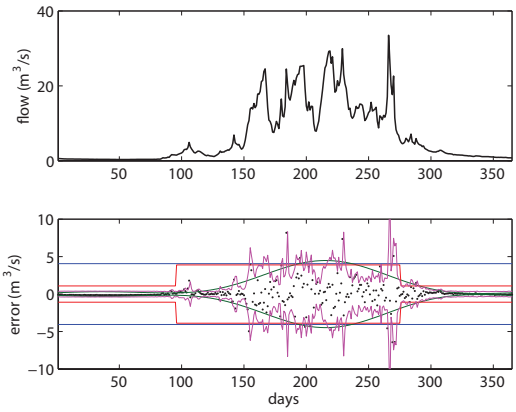

Figure 2.2: Top panel: Rhone River discharge. Bottom panel: residual error (black dots) and 95% confidence bounds based on the constant variance model (blue), two-mixed (red), periodic (green), and DUMBRAE (magenta). First year (1981) of the calibration dataset.

ones. Notice that these two models exhibit very similar results in the QQ plot, while having very different log-likelihood values. Specifically, the periodic model outperforms DUMBRAE in the QQ-plot over the validation dataset although it is definetly outperformed in terms of log-likelihood (see Table 2). This is because the periodic model strongly underestimates the variance in a small number of events, which are not visible in the QQ plot but are heavily penalized in the likelihood score.

In conclusion, the analysis indicates that the DUMBRAE model (2.12), although simple, provides an effective tool for estimating the standard deviation error. It outperforms the other three benchmark models in terms of the log-likelihood, AIC and PICP indicators; it provides an error cdf closer to the theoretical one than that of the constant and two-mixed model, although the s-shaped QQ-plot indicates that the standard deviation is sometimes overestimated; and it produces a confidence interval that more closely adjusts to the variation in the model (absolute) error.

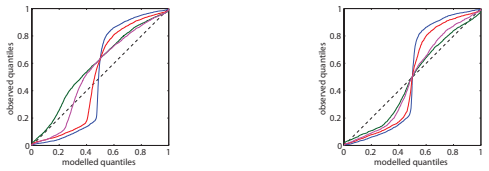

Figure 2.3: QQ plot of the constant variance model (blue), two-mixed (red), periodic (green), and DUMBRAE (magenta) for the Rhone River case study over the calibration dataset (left) and validation data set (right).

2.4 Application to inflow forecasting in Lake Maggiore catchment

Lake Maggiore is a regulated lake at the border between Italy and Switzerland. The lake catchment covers about 6600 km^2, with 17% of the watershed area above 2000 m above sea level. Climate conditions are extremely variable, with higher precipitation in autumn and spring and significant contribution from snowmelt in late spring and summer. A flow forecasting model was developed to support real-time operation of the lake [Pianosi and Ravazzani, 2010]. It is a data-driven, lumped model that provides the total inflow $\hat{y}_t$ to the lake in the next 24 hours, as a function of the data available at the time of forecast: measured precipitation in the catchment (spatial average), the observed inflow and the forecasting error in previous time intervals. The model is composed of (i) an autoregressive component employing the logarithm of the flow, so that the recession curve follows a more than exponential decay; (ii) an exogenous component that weights the precipitation input by a periodic function estimated from data; and (iii) a moving average component based on previous forecasting errors. Although it is a data driven model, its parameters can be given a physically-sound interpretation.

The model was calibrated using time series of inflow and precipitation over the period 1993-1997. The time series of residual error over the same period is used to identify the error variance model, whereas data from 1998 to 2000 are used for model validation.

Figure 2.4 shows the autocorrelation function of the model residuals. It can be seen that while the residual error is almost uncorrelated, the squared error and the absolute error are significantly autocorrelated, which is an evidence of heteroscedasticity

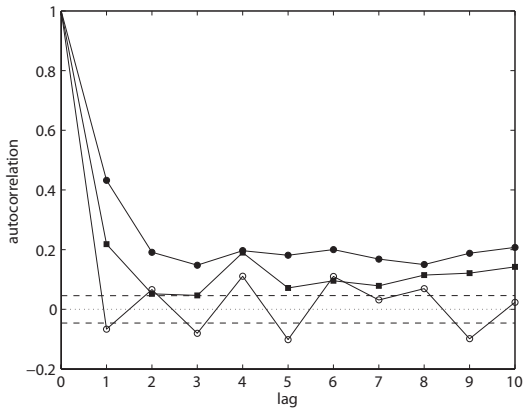

Figure 2.4: Autocorrelation of residual error e_t (white circles), squared error e_t^2 (squares) and absolute error $|e_t|$ (black circels) for the Lake Maggiore catchment case study (calibration data set).

of the error process [Engle, 1982].

2.4.1 Model identification

Potential inputs of the standard deviation model are first searched among the hydrological model inputs, that is, precipitation, residual error and flow observation in previous time intervals.

Figure 2.4.1 shows the correlation between past precipitation p_{t-k} and error e_t (white circles), squared error e_t^2 (squares) and absolute error $|e_t|$ (black circles). It can be noticed that the error is highly correlated with precipitation in the same time interval ($k = 0$), because the forecasting error is often due to a simultaneous and thus unpredictable precipitation event, while the correlation with past precipitation records ($k > 0$) is negligible, which confirms that all the information available at the time of forecast is correctly exploited by the hydrological model. However, the

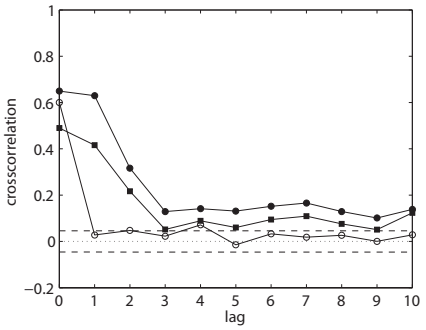

Figure 2.5: Cross-correlation between past precipitation p_{t-k} and error e_t (white circles), squared error e_t^2 (squares) and absolute error $|e_t|$ (black circles) for the Lake Maggiore catchment case study (calibration data set).

correlation between precipitation and squared error is high also for $k > 0$. This means that while the error value does not depend on past precipitation, the error variance does; or, in physically-sound terms, uncertainty in the inflow forecast increases after rainfall events. Finally, the correlation between precipitation and the absolute error is even stronger than between precipitation and squared error. Therefore, it can be expected that identifying the standard deviation model from time series of the absolute error be easier than identifying the variance model from squared errors.

Since the correlation of the absolute error $|e_t|$ with predicted flow $\hat{y}_t$ (0.44) and with observed flow in the previous time interval y_{t-1} (0.35) are both lower than correlation with the observed precipitation, the latter will be used as the exogenous input of the standard deviation model, which takes the form

$$\hat{\sigma}_t = a + \sum_{i=1}^{n} b_i |e_{t-i}| + \sum_{i=1}^{m} c_i p_{t-i} \tag{2.13}$$

Table 2.3: Comparison of different standard deviation models (Lake Maggiore catchment): n_ϕ is the number of model parameters; $-\log(L)$ is the negative log-likelihood; AIC is the Akaike Information Criterion; PICP is the prediction interval coverage probability at level of confidence $\alpha = 0.95$.

Model	n_ϕ	$-\log(L)$		AIC		PICP	
		cal	val	cal	val	cal	val
Constant	1	12.1×10^3	7.32×10^3	24.1×10^3	14.6×10^3	0.965	0.959
Periodic	5	11.7×10^3	7.20×10^3	24.0×10^3	14.3×10^3	0.921	0.911
State dependent	2	11.8×10^3	7.18×10^3	23.6×10^3	14.4×10^3	0.942	0.940
DUMBRAE	5	10.8×10^3	6.60×10^3	21.8×10^3	13.4×10^3	0.945	0.923

Following the uncertainty decomposition based on hydrological causes and presented by Götzinger and Bardossy [2008], the first two components (constant and autoregressive) in (2.13) can be interpreted as uncertainty due to the process description, whereas the third component is due to the precipitation input. Model calibration follows the same approach as described in the Rhone River application.

2.4.2 Model Evaluation

Just as in the Rhone River application, a constant and a periodic model of the standard deviation will be used as benchmarks. Furthermore, a state dependent model, as used for instance by Thyer et al. [2009]; Schoups and Vrugt [2010], will be assessed. It assumes that the standard deviation is linearly proportional to the inflow forecast, i.e.

$$\hat{\sigma}_t = a + b\ \hat{y}_t$$

Table 2.4.2 reports the evaluation indicator values by these three models and the DUMBRAE model (2.13) with $n = 2$ and $m = 2$. It shows that DUMBRAE has the lowest negative log-likelihood and thus the highest skill in reproducing the error probability. The AIC value is also minimum and the PICP is quite close to the theoretical value of 0.95 (even if the constant and state dependent model are slightly better over the validation dataset). Contrary to the Rhone river application, the lower quantile of the flow estimated by DUMBRAE is negative in correpondence to some heavy rainfall events that produce a very high standard deviation values (see equation 2.13), even if the frequency of such negative values is lower than with the constant and periodic model. As discussed in the previous section, this is not a problem for operational purposes (it will be sufficient to replace negative flow quantiles by zero) but it is a conceptual weakness in the proposed approach. Nonetheless, our

opinion is that for application-oriented scopes this weakness is acceptable since, on the other hand, the Gaussian assumption provides several important computational advantages (the variance model can be identified separately from the mean; time series of absolute errors can be used in place of squared errors thanks to equation 2.9).

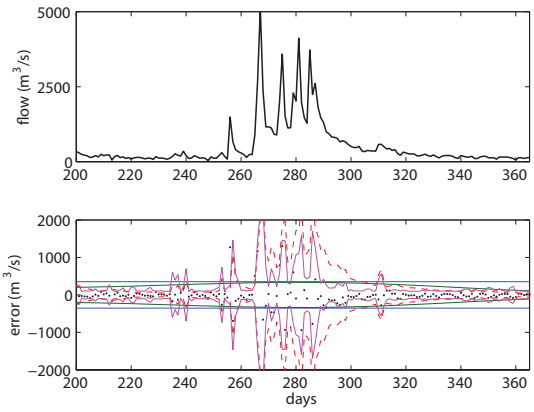

Figure 2.6: Top panel: Daily inflow to lake Maggiore. Bottom panel: residual error (black dots) and 95% confidence bounds based on the constant variance model (blue), state-dependent (dashed red), periodic (green), and DUMBRAE (magenta). Days 200-360 in the calibration dataset.

Figure 2.4.2 compares the 95% confidence intervals of the flow based on the DUMBRAE model and the three benchmark standard deviation models. It shows that if a constant standard deviation is used, the confidence interval (blue lines) is too large for low flow events and too narrow for flood events. Since floods generally occur in autumn, the periodic model properly produces a larger standard deviation value in that period, and thus a wider confidence interval (green lines). Still, the periodic model is not completely satisfactory since it provides an average assessment of the seasonal uncertainty but it cannot distinguish high uncertainty periods

caused by precipitation events. The confidence intervals based on DUMBRAE and the state dependent model (red and magenta lines) are narrow in correspondence to low flow values and wider in correspondence to floods, when unpredictability increases. However, an important difference in behaviour between those two models can be detected observing their trends during the recession process, i.e. from day 290 to day 300 in Figure 2.4.2. Forecasting errors in this phase are small because the hydrological model can accurately reproduce the flow dynamics in the recession phase. The DUMBRAE model can quickly adapt to this situation and properly reduce the confidence intervals because it is updated by information on absolute errors in previous time intervals (see equation (2.13)). The state dependent model, instead, by simply relating the standard deviation to the flow forecast, cannot distinguish if the high flow is due to an ongoing precipitation event (high uncertainty) or to a recession process (low uncertainty), and in this latter case overestimates the confidence interval. From a model diagnostics perspective, DUMBRAE indicates that the forecast error is mainly related to precipitation input, suggesting that either the precipitation observations (or their spatial aggregation) is a major source of uncertainty or that the hydrological model is not adequate in reproducing the fast response of the catchment to rainfall events (while properly reproducing the recession curve).

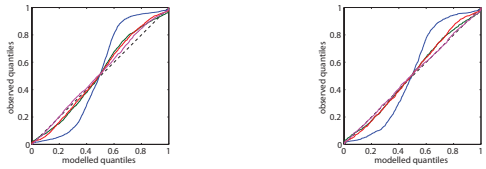

Figure 2.7: QQ plot of the constant variance model (blue), state-dependent (red), periodic (green), and DUMBRAE (magenta) for the lake Maggiore catchment over the calibration dataset (left) and validation data set (right).

Finally, Figure 2.7 reports the QQ plot generated by the four standard deviation models over the calibration and validation dataset. The line generated by the DUMBRAE model (magenta) is very close to the bisector (black line), while the curve of the other standard deviation models is s-shaped, which means that the z_t points are concentrated towards the center of the interval $[0, 1]$. This indicates that the confidence interval is frequently wider than needed, or in other terms, that the standard

deviation is often overestimated.

2.5 Conclusion

We have presented here a method for identifying a simple but effective uncertainty model to be associated with the predictions of a hydrological model with heteroscedastic errors. The method is very general and only requires the time series of residual errors over a historical period as an input, regardless of the structure of the hydrological model. However, it relies on the assumption that such error be independent and zero mean, and thus an appropriate transformation of the hydrological model output may be required to meet this condition. Also, we show that under the further assumption that the residual error is Gaussian, the error standard deviation is linearly proportional to the error absolute value, which means that modelling the error standard deviation boils down into a regression analysis of the absolute error time series. Based on these considerations, we show that a proper data analysis combined with physical considerations on the case study at hand can help identifying the most suitable input to the standard deviation model and the most parsimonious model structure. The approach has been called Dynamic Uncertainty Modelling By Regression on Absolute Errors (DUMBRAE).

We also provide some theoretical background and practical approaches for model identification, based on the maximum likelihood principle. The effectiveness of the method was demonstrated by application to two case studies and comparison with other approaches presented in the literature for modelling the error variance, as well as a simple constant and periodic model. Modelling results are assessed in terms of the formal likelihood measure, plus other evaluation indicators and a graphical tool, the QQ plot, which provides more insight about the estimated probability distribution. The analysis shows that the proposed method can effectively reproduce the heteroscedasticity of the residual errors. A suitable choice of the inputs to the standard deviation model can improve the model accuracy and compensate for the simplifying assumption of Gaussian errors. Using precipitation measurements and past observations of the error itself, and not only the flow forecast as usually done in the literature, significantly enhance the uncertainty description especially in the recession phase of the hydrograph. Finally, although DUMBRAE is an application-oriented approach and it was originally conceived for the practical goal of effectively associating prediction bounds to flow forecasts, it can also contribute to model diagnostics. In fact, as demonstrated in the proposed case studies (especially the lake Maggiore catchment), the analysis of the time series of residual errors help to identify what variables and processes mostly contribute to the forecasting error, and

thus what components of the hydrological model should be improved. One limitation of the proposed approach is that it depends on the a priori assumption of an adequate function family for the error standard deviation model. Further research will also concentrate on extending the approach to other more flexible probability distribution like for instance Gamma distributions.

Chapter 3

Tree-Based Model Predictive Control for Short-Term Management of Reservoir Systems using Ensemble Forecasts

Model Predictive Control (MPC) is a method for controlling complex, constrained dynamic systems. Applied to water systems operations, MPC provides integrated, optimal, and proactive management, when hydrological forecasts are available. Notwithstanding these properties, uncertainties can jeopardize MPC reliability. In water resources systems, the main uncertainty is generally introduced by the hydrological forecast. To communicate the forecast uncertainty, meteorological and hydrological forecasts are often produced in the form of ensembles, i.e. a set of possible future trajectories of the system.

We present an adaptive control method, called Tree-Based MPC (TB-MPC), which is able to exploit the information contained in an ensemble forecast. TB-MPC introduces the feedback mechanism of closed-loop in the MPC control scheme limited to the uncertainty covered by the ensemble, reducing the sensitivity to wrong forecasts and enhancing the control performances. In TB-MPC, ensemble members branch out when they diverge from each other. A tree, made through repeated out-branching, is used to set up a Multistage Stochastic Programming (MSP) problem, finding a different optimal control strategy for each branch. To show its potential advantages, TB-MPC is applied to the operational management of the Salto Grande reservoir, located at the border between Argentina and Uruguay, and compared to other control methods.

3.1 Introduction

Management of water systems can be considered as an optimal control problem [Labadie, 2004; Soncini-Sessa et al., 2007; Castelletti et al., 2008]. Optimal control

problem can be solved using dynamic programming (DP) [Bellman and Dreyfus, 1966], which is an off-line method, resulting in a series of optimal control laws.

Drawback of DP is that it is generally difficult to solve, running into the so-called "curse of dimensionality" [Bellman and Dreyfus, 1966]. In water management, DP has been used mainly for reservoir management, and yet for systems made of a small number of reservoirs [Stedinger et al., 1984; Trezos and Yeh, 1987]. Much research effort has been devoted to bypass the curse of dimensionality by some approximations in the solving methods or in the model of the physical system. See, for example, [Pereira and Pinto, 1991; Rotting and Gjelsvik, 1992; Faber and Stedinger, 2001; Castelletti et al., 2010; Turgeon, 1980; Castelletti et al., 2007; Tilmant et al., 2008].

Model Predictive Control is another way to solve the control problem, alternative to off-line solutions solved with DP. MPC resulted effective for optimal centralized control of complex, constrained systems [Maciejowski, 2002], and it is a promising technology for water systems [van Overloop, 2006; Raso and van Overloop, 2011].

Differently from DP, where a optimal control law is calculated off-line for all states, in MPC the current control action is obtained by solving on-line a finite horizon open-loop optimal control problem, where the initial state is is the current state of the system being controlled. Key elements of MPC are [Morari et al., 1999]: (1) a model of the physical process to predict future trajectories of the controlled variables over a finite horizon, (2) the calculation of a control sequence that optimizes an objective function, and (3) a receding horizon strategy. The receding horizon strategy means that, at each time instant, the first signal of the control sequence is applied and the horizon is shifted ahead. In MPC, constraints on inputs, states and outputs are explicitly considered [Schwanenberg et al., 2011a].

Solving the open-loop optimal control requires a mathematical programming problem, much less computationally costly than DP. For this reason, MPC is best fitted for all cases where an off-line control law is difficult or even impossible to compute [Mayne et al., 2000]. For water systems, this means that MPC can enlarge the controlled system beyond reservoirs only, such as including the routing along a downstream river [Ficchì et al., 2013], or combining water quantity and quality [Xu et al., 2010, 2013].

In MPC, available disturbance forecasts can be directly used in the control scheme, bringing along advantages and threats. The advantage is that the control strategy becomes *proactive* [Zavala et al., 2009]; before the realization of a forecasted disturbance, the control sequences set the system to a state optimal to accommodate it, for example by lowering the water level in a reservoir before an expected flood event. The threat is that use of forecasts can jeopardize the control robustness [Be-

mporad and Morari, 1999]. Forecasts are in fact the output of a model generally affected by large uncertainties, whereas MPC is a deterministic controller. For this reason, a deterministic open-loop proactive controller runs the risk of taking action in anticipation of an expected disturbance that, eventually, will not occur.

Even MPC is a deterministic open-loop control problem, the on-line iterations ensure the feedback mechanism to MPC, guaranteeing an "inherent robustness" [Mayne et al., 2000]. For some systems, solving the nominal control problem produces satisfactory results. Other systems, specifically when their characteristics depart from certainty equivalence [Philbrick and Kitanidis, 1999; Van de Water and Willems, 1981], are more sensitive to wrong controls. The more a system is away from the certainty equivalence condition, the more the explicit recognition and treatment of uncertainty has an advantage on deterministic solutions [Weijs et al., 2007].

In open water systems, uncertainty is often related to the difficulty of producing good meteorological and hydrological forecasts. Different studies show how streamflow forecast uncertainty can affect water systems operation [Zhao et al., 2011; Pianosi and Soncini-Sessa, 2009]. To communicate forecast uncertainty, meteorological and hydrological centers produce ensemble forecasts (EF). An ensemble is a set of representative trajectories of the possible future outputs of the meteorological or hydrological system [Gneiting and Raftery, 2005].

EFs have been employed in optimal control schemes in various recent applications, such as Multiple Model Predictive Control (MMPC) for open water systems [van Overloop et al., 2008], or equivalent approaches applied on storehouses [Doeswijk et al., 2006], and on energy systems [Lukasse et al., 2006; Zavala et al., 2009]. MMPC, and these equivalent methods, find the optimal control series using the entire ensemble in a stochastic optimization, and are apparently more robust than deterministic MPC. However MMPC still solves an open-loop optimal control problem; the control problem is set as if the entire control sequence is to be applied. Open-loop scheme does not consider that, in the next decisions instants, it will be possible to adjust the control according to the ongoing scenario using the last up-to-date information. In fact, future observations will reveal which ensemble members can be excluded and which ones can still happen, hence the controller will be able to adapt its control strategy accordingly. In this sense, MMPC, even if stochastic, is non-adaptive [Bertsekas, 1976].

MMPC, not considering that new information will be available in the next future, overestimates the uncertainty, and results in an over-conservative control policy. In addition, there is a larger risk of getting infeasible solutions in the presence of state constraints. Closed-loop control schemes add robustness, but are more computation-

ally intensive. In our approach we limit the closed-loop response to the most relevant uncertainty. The algorithm we propose, called Tree-Based Model Predictive Control (TB-MPC), sets a closed loop optimal control problem limited to EF uncertainty. TB-MPC embeds the ensemble data in a Multistage Stochastic Programming (MSP) problem. MSP is the proper problem formulation when decisions are interdependent and some of them can be taken after uncertainty is resolved [Birge and Louveaux, 1997].

MSP has been the subject of many studies [Dupačová et al., 2003; Birge and Louveaux, 1997], even specific to MPC and control of dynamic systems, [Muñoz de la Peña et al., 2005], or to water management [Watkins and Wei, 2004; Watkins Jr et al., 2000; Pallottino et al., 2005; Escudero, 2000], yet its major practical drawback is its intractability. In MSP in fact, every time some uncertainty is resolved, the system dimension is multiplied by the number of possible outcomes of the uncertain variable. Assuming a system multiplication at every time step, the number of variables increases exponentially with the horizon length. For its prohibitive complexity, Mayne et al. [2000] consider MSP for control of dynamic systems as "conceptual rather than practical" method. TB-MPC, instead, sets up a MSP problem limiting the number of "branches" in the tree, still capturing the major uncertainties.

3.2 Methodology

In MPC scheme, at each control instant t_0, the open-loop optimal control problem (3.1) is solved in real time.

$$\min_{\{u_t\}_{t=1}^{h}} \sum_{t=1}^{h-1} g_t(x_t, u_t, d_t) + g_h(x_h, u_h, d_h) \tag{3.1a}$$

$$x_t = f_t(x_{t-1}, u_t, d_t) \tag{3.1b}$$

$$c(x_t, u_t) \leq 0 \tag{3.1c}$$

Where t, the time index, goes from 1 to the final time step of the control horizon, h; g_t is the time step cost function, g_h the final penalty that sums up all the future costs beyond the control horizon, x_t, u_t and d_t, vectors of appropriate dimension, are respectively the state, control, and disturbance at time t, f is the process model, and $c(\cdot)$ the inequality constraints present in the system. The initial conditions (x_0) and the series of deterministic disturbances ($d_t\ \forall t$) are given.

MPC solves Problem (3.1) at each control time step. In MPC, the argument of the optimization problem is the control series $\{u_t\}_{t=1}^{h}$, that is a single control strategy.

Instead, in TB-MPC, the argument of the optimization problem is a "control tree". A control tree is a larger space than the control series. Using a tree, the optimal control problem in TB-MPC takes explicitly into account that the control can change along the control horizon dependently on the occurring ensemble member, finding a different optimal control strategy for each "branch".

In TB-MPC, ensemble data are used to generate a tree structure, then the tree and the ensemble are used in the optimization algorithm. Tree structure generation and optimal control are two separate problems. We introduce the tree in (3.2.1), then, in (3.2.2), we show how to use the tree and the ensemble in TB-MPC.

3.2.1 From ensemble to tree

An EF consists of different trajectories of the possible future outputs of a meteorological or hydrological model. EF typically refers to meteorological predictions, where trajectories are generated running the model n times, under different initial conditions or different numerical representations of the atmosphere, accounting for the major sources of forecast uncertainty [Gneiting and Raftery, 2005]. However, we extend the definition of EF to include also hydrological forecasts, that are generally called Ensemble Streamflow Prediction (ESP) [Cloke and Pappenberger, 2009; Day, 1985; Schaake and Larson, 1998; Faber and Stedinger, 2001]. In our terminology, ESP is just a subclass of EF.

Generally, the ensemble trajectories have small differences at the initial stage of the forecast, and tend to diverge in time. In the EF representation of uncertainty, randomness is among trajectories, but not within them. The ensemble trajectories are $d_{t,z}$, where $t \in \{1, 2, ..., h\}$ is the time index, and z is the ensemble member index, outcome of the stochastic variable $z \in Z = \{1, 2, \ldots, n\}$. The ensemble dimension n is generally a number between 5 and 100. Each ensemble member z has an assigned probability $p(z) \geq 0$, such that $\sum_{z=1}^{n} p(z) = 1$. Members are generally equiprobable. In this case, z's distribution is uniform, and $p(z) = 1/n, \forall z$.

While an ensemble is simply a collection of possible trajectories, a tree specifies when these trajectories diverge from each other. In a tree, at the initial stage $t = 1$, all trajectories are possible. At any *branching point* the sample space of possible ensemble members splits in two subsets. A branching point is placed where trajectories diverge, indicating the moment when uncertainty on the occurring sub-branch is going to be resolved, or in other words, the moment in which it will be clear which sub-branch is going to occur. From this instant onwards, these subsets are mutually exclusive, i.e. the occurrence of one excludes the occurrence of the other.

The most compact way to define a tree structure is assigning a *parent* $P(\cdot)$ and a *branching point* $B(\cdot)$ to each ensemble member, [Raso et al., 2012b], i.e.:

$$P(z), B(z)\ \forall z \tag{3.2}$$

$B(z)$ and $P(z)$ define respectively the time on the forecast horizon and the member from which member z branches out.

Generating a proper tree structure is a problem separate and prior to the optimal control problem. Stive [2011] shows that the tree structure has a strong influence on the control performance, stressing the importance of the tree generation problem. Existing methods to build up a tree from complete scenarios include: [Šutienė et al., 2010], presenting an empirical method to bundle scenarios based on the k-means technique, [Growe-Kuska et al., 2003], showing an algorithm based on transportation metrics in which the problem is formulated as a minimal mass transportation problem, and [Raso et al., 2012b], introducing an approach that takes into account the available information along the forecasting horizon.

The tree structure influences the control performances. A correct tree structure should neither overestimate nor underestimate the capacity to resolve uncertainty along the control horizon. In the extreme case, a tree is made of a fan of individual scenarios, as in [Dupačová et al., 2003]. Such a tree structure implies that uncertainty is completely resolved already at the second time step. This, in general, overestimates the future capacity of resolving uncertainty. At the other opposite extreme, a tree without any out-branching, equivalent to the MMPC control problem, assumes no expected uncertainty resolution, hence overestimating the total uncertainty. A proper tree structure is likely to lie between these two extremes.

The *scenario tree nodal partition matrix*, $\mathbf{M}(t, z)$ [Dupačová et al., 2003]. is a different way to represent the tree structure. Although less concise than (3.2), $\mathbf{M}$ provides the necessary labeling scheme to set up the MSP problem. $\mathbf{M}$ can be obtained from (3.2), and viceversa, as described in Equation (3.3).

$$\mathbf{M}(t, i) = \mathbf{M}(t, j) \text{ when } t < B(i), P(i) = j;\ i, j = 1, \ldots, n \tag{3.3}$$

The matrix $\mathbf{M}$ has dimension $h \times n$. As explanatory example, the same simple tree (having $h = 5$ and $n = 3$) is shown in Figure 3.2.1 and defined in (3.4), using both $\mathbf{M}$ and (3.2).

$$\mathbf{M} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 6 & 7 & 8 \\ 1 & 2 & 6 & 7 & 9 \end{bmatrix}^T \longleftrightarrow \begin{array}{rcccc} \text{z} & 1 & 2 & 3 \\ \hline \text{P(z)} & 1 & 1 & 2 \\ \text{B(z)} & 1 & 3 & 5 \end{array} \tag{3.4}$$

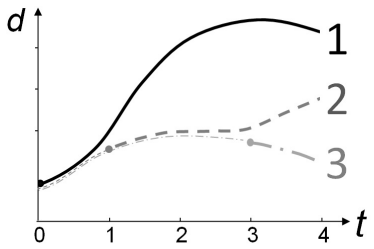

Figure 3.1: Graphical representation of the tree defined in (3.4)

3.2.2 From tree to optimal control

Consider an ensemble member and its parent. TB-MPC finds two different control strategies from the branching point onwards. Controls after the branching points can adapt to the occurring ensemble member, and improve. In dynamic systems, each state x_t depends on all the controls up to t. Controls u_t at different time steps t are thus interdependent. For this reason, controls before the branching point also improve, including the first one $\{u_t\}_{t=1}$ that, coherently with the receding horizon approach, is eventually applied to the real system.

At the decision instant t_0, if the controller knew the occurring ensemble member z, it could set the optimal control conditional to this information, i.e. find the optimal control strategy u_t using $\{d_{t,z}\}_{t=1}^{h}$ as disturbances. But, at instant t_0, z is a stochastic variable, outcome of Z, and the best that the controller can do is to weigh the effects of $d_{t,z}$ by their probability $p(z)$. Along the control horizon, however, direct or indirect observations of d_t will become available, giving information about which z is actually occurring [Raso et al., 2012b]. The feedback on which occurring scenario is occurring can be used to adapt the control strategy to it. From the instant that the controller will be able to distinguish z as the occurring scenario, it will find a control strategy optimal to z only. These moments along the control horizon are marked by the branching points.

The branching point $B(z)$ identifies when z is distinguishable from its parent $P(z)$. For $t < B(z)$, both z and $P(z)$ are possible outcome of Z, thus the control strategy from $t = 1$ to $B(z)$ is found by weighting the effects of $\{d_{t,z}\}_{t=1}^{B(z)}$ by $p(z)$. After the

branching point, from $t = B(z)$ to the end of the control horizon h, the controller is free to take different control strategies for z and $P(z)$.

The key idea of TB-MPC is to consider a different control strategy for each ensemble member (or set of them), to be applied from the moment that this ensemble member (or set) branches out. This is expressed in Equation (3.5), in which for different ensemble members, i and j, up to their branching points, the controls must be the same.

$$u_{t,i} = u_{t,j} \text{ when } \begin{cases} P(i) = j \\ t < B(i) \end{cases} \quad i, j = 1, \ldots, n \tag{3.5}$$

Equation (3.5) translates the non-anticipativity condition, saying that controls should not depend on the outcome of stochastic variables that have not been extracted yet [Birge and Louveaux, 1997].

The non-anticipative condition, as expressed in Equation (3.5), is implemented in TB-MPC by enlarging the control set, using $\mathbf{M}$. The optimal control problem for TB-MPC is defined in (3.6).

$$\min_{u_{\mathbf{M}(t,z)}} \sum_{z=1}^{n} p(z) \left[\sum_{t=1}^{h-1} g_t(x_{t,z}, u_{\mathbf{M}(t,z)}, d_{t,z}) + g_h(x_{h,z}, u_{h,z}, d_{h,z}) \right] \tag{3.6a}$$

$$x_{t,z} = f_t(x_{t-1,z}, u_{\mathbf{M}(t,z)}, d_{t,z}) \tag{3.6b}$$

$$c(x_{t,z}, u_{\mathbf{M}(t,z)}) \leq 0 \tag{3.6c}$$

where $u_{\mathbf{M}(z,t)}$ is the control space, a matrix of dimension $m \times N_u$. m is defined as the largest value of $\mathbf{M}$, and N_u is the control dimension. In the example tree of (3.4) and when $N_u = 1$, $\mathbf{M}$ is a 9×1 vector.

Matrix $\mathbf{M}(t, z)$ provides the "address" (i.e. the row) in u_M, given t and z. Before a branching point, for different z's, $\mathbf{M}$ gives the same address, thus the same control value, which is the non-anticipativity condition already expressed in Equation (3.5). Consider, as an example, a member z and its parent $P(z)$. For $t < B(z)$, when it is not know which member will happen, the same control must find an optimal compromise between the effects of z and $P(z)$, weighted by their probability, $p(z)$ and $p(P(z))$. After the branching point, Matrix $\mathbf{M}$ returns a different address for different z's, thus different control values. In this way, $\mathbf{M}$ takes into account the future capacity to change the control strategy depending on the occurring sub-branch.

The first control is to be implemented to the real system, thus it must be unique. Imposing the condition that branching points lay only from the next time step onwards ($B(z) > 1\ \forall z$) guarantees their uniqueness. On $\mathbf{M}$, this condition implies that on the first columns all rows have the same value.

$$\mathbf{M}(i,t) = \mathbf{M}(j,t)\ \forall i,j \in Z; t = 1 \tag{3.7}$$

The future possibility of changing control strategy along the control horizon implies a larger control space; the degrees of freedom in TB-MPC problem is determined by the tree structure, but $u_{\mathbf{M}}$'s dimension is never smaller than in deterministic MPC, being $m \geq h$. The larger dimension of the optimization problem influences the computational complexity of TB-MPC. It is difficult to draw general conclusions on the computational burden, which depends on a large number of elements. Nonetheless, it has been shown that, for the interior point method, the complexity has a polynomial growth with respect to the number of variables [Forsgren et al., 2002], which is a rather benign condition. Moreover, to speed up the solving time, TB-MPC can be easily parallelized, as shown in [Maestre et al., 2012a,b]. Some more empirical data about the computational time are presented in section (3.3.3).

3.3 Application

TB-MPC is applied for the operation of Salto Grande, a hydropower reservoir on the Uruguay River, at the border between Uruguay and Argentina. The reservoir extends for 800 km^2 and the storage volume is about $5.5 \times 10^9 m^3$. The catchment area is approximately 240.000 km^2 and covers parts of Argentina, Brazil and Uruguay. The average inflow is about 4700 m^3/s and the installed capacity is 1890 MW.

Section (3.3.1) describes the control problem. Section (3.3.2) shows how to generate the ensemble and the tree. Section (3.3.3) describes the control method against which TB-MPC is compared and the results.

3.3.1 Controlled system and objectives

The controlled reservoir is modeled by a (discrete time) mass balance equation

$$x_t = x_{t-1} + \Delta t \cdot \left(d_t^{in} - u_t^{turb} - u_t^{spill} \right) \tag{3.8}$$

Where the state x is the reservoir volume, the disturbance d^{in} is the hydrological inflow, and u^{turb} and u^{spill} are the controlled releases from the turbines and the

spillways. The controlled system is made of one state, one disturbances and two controls. The control time step Δt is six hours. The inflow to the reservoir is the disturbance obtained from the hydrological model and described in section (3.3.2). Evaporation and other flows are neglected.

The water level can be obtained from the volume using the storage curve $A(\cdot)$ as in Equation (3.9).

$$h_t^{forebay} = A\left(x_t\right) \tag{3.9}$$

The control u_t^{turb} and u_t^{spill} are constrained according to inequalities (3.10).

$$0 \leq u_t^{turb} \leq u_t^{turb,\max} \tag{3.10a}$$
$$0 \leq u_t^{spill} \leq u_t^{spill,\max} \tag{3.10b}$$

where $u_t^{turb,\max}$ and $u_t^{spill,\max}$ are the maximum release through the turbines and the spillways. These values depend on the water level h according to their rating curves R as in Equations (3.11).

$$u_t^{turb,\max} = R_1\left(h_t^{forebay} - h_t^{tailwater}\right) \tag{3.11a}$$
$$u_t^{spill,\max} = R_2\left(h_t^{forebay}\right) \tag{3.11b}$$
$$h_{t+1}^{tailwater} = R_3^{-1}\left(u_t^{out}\right) \tag{3.11c}$$

$h_t^{tailwater}$ is the tailwater elevation, downstream of the reservoir. Equations R's are static, i.e. they do not take into account unsteady conditions. Nonetheless, the given accuracy is sufficient for the purposes of this study.

The objective of the reservoir management, as expressed by the reservoir manager, can be phrased as:

> Maximizing energy production and avoiding upstream and downstream floods.

In addition, a minimum flow must be maintained for environmental requirements.

Table 3.3.1 shows the objectives included in the optimization problem. Referring to the points of Table 3.3.1, points 1 and 3 penalize the exceedings of the upstream flood threshold level (35 m) or emptying of the reservoir (water level under 30 m). Points 2 and 4 are intended to maximize the energy production by maximizing the water level in the reservoir and minimizing the spillage. This indirect way of formulating the energy production maximization has the advantage of offering closed loop stability [Bemporad and Morari, 1999] without requiring a cost-to-go function $g_h(\cdot)$, which is difficult to define. Points 5 and 6 maintain the downstream flow

i	Sub-objective	Coefficient $[w_i]$	Formula
1	Flood upstream	100	$\left[\max\left(h_t^{salto} - 35.0, 0\right)\right]^2$
2	Energy production	1	$\left[\min\left(h_t^{salto} - 35, 0\right)\right]^2$
3	Resource deployment	5	$\left[\min\left(h_t^{salto} - 30.0, 0\right)\right]^2$
4	Energy production	5x10E-4	$\left[u_t^{spill}\right]^2$
5	Flood downstream	10E-4	$\left[\max\left(u_t^{tot} - 20.000, 0\right)\right]^2$
6	Environment	10E-4	$\left[\min\left(u_t^{tot} - 500, 0\right)\right]^2$
7	Wear and tear of structures	5x10E-7	$\left[u_t^{spill} - u_{t-1}^{spill}\right]^2$
8	Wear and tear of structures	5x10E-7	$\left[u_t^{turb} - u_{t-1}^{turb}\right]^2$

Table 3.1: Sub-objectives that make up the total cost function

between the minimum environmental requirements (500 m^3/s) and the flooding threshold level (20,000 m^3/s). Points 7 and 8 smoothen the control trajectories to avoid continuous and rapid changes. All these sub-objectives i, considered for the entire control horizon h, are weighted by their weight w_i and summed up in the total cost function J to be minimized.

$$J = \sum_{t=1}^{h} g_{tot,t} = \sum_{t=1}^{h} \sum_{i=1}^{m} w_i \cdot g_{i,t} \tag{3.12}$$

Both the forecasting and the control horizon are 15 days. The control horizon is relatively short because the reservoir, compared to the inflow, is relatively small. This also implies little controllability of the system, as shown in the results.A new weather forecast is produced every 6 hours, which is also the time step of the hydrological model, the reservoir model, and the control. Therefore, optimal control problem uses a new inflow EF and tree at each time step. The next section describe how to generate them.

3.3.2 Ensemble and Tree Generation

To produce the streamflow EF, A meteorological forecasting system provides the inputs to the hydrological model, whose output is the inflow to the reservoir.

The Numerical Weather Predictions (NWP), both in its deterministic and ensemble version, is produced by the US National Center for Environmental Prediction (NCEP). The weather EF is made of 20 equiprobable ensemble members. The forecasting horizon is 15 days and the frequency is 6 hours. The weather EF used as input

for the hydrological model is made of $\{P_{t,z,k}, T_{t,z,k}\}$, where $t = [0, 1, \ldots, 60] \times 6h$, $z = 1, \ldots, 20$, and $k = \{$Sao Joacquim, Joacaba, Santa Rosa, Sao Borja, Uruguaiana$\}$. P is the precipitation, T is the 2-m temperature, t is the time index, z the ensemble member index and k the location index.

The hydrological model uses the weather EF as input and produces a streamflow EF $\{d_{t,z}\}$, which is the inflow to the reservoir. Description of the hydrological model is in the appendix.

Streamflow data assimilation is realized using an AR output correction, as described in Schwanenberg et al. [2011b]. This module updates the hydrological model output at the three locations where observed discharge is available: Irai, Sao Borja, and Salto Grande. A Bayesian Model Averaging (BMA) method, as described in Raftery et al. [2005], using no bias correction and equiprobable ensemble members, is used to post-process the original ensemble and include the hydrological uncertainty in the EF.

The BMA enlarge the EF in order to include the hydrological uncertainty. However, in this enlarged ensemble, many members that are similar to each other increase the computational burden without adding much information. To reduce their number, members can be effectively aggregated using the scenario reduction technique described in Dupačová et al. [2003]. This method, solving iteratively a minimum mass transportation problem, finds the two closest ensemble members, deletes one and assigns its probability to the other. This reduce the computation burden without loosing much in terms of performances, as shown in Stive [2011] for a similar system. The streamflow EF $d^{in}_{t,z}$ is reduced to 15 members, and used in the optimal control problem.

The tree structure is generated using the information-based approach described in Raso et al. [2012b] by setting the threshold probability to $p^* = 0.95$. The present inflow to the reservoir is the only available observation considered. The observational uncertainty is considered independent on previous values and normally distributed, with standard deviation 300 m^3/s, quantified using an simplified approach derived from Di Baldassarre and Montanari [2009].

3.3.3 Results

The main objective, energy production, is realized by keeping a high water level without spilling water. The turbine capacity is limited. Therefore, before flood events, the controller must lower the water level to create storage for the approaching peak that, otherwise, cannot flow entirely trough the turbines. A higher water level in-

creases the risk of spillage. Performances are linked to the ability to find an optimal trade-off between these two needs.

The closed loop horizon, for which we have EF data, extends from 26 February 2010 to 13 June 2011. In this period, there are just 5 events that require anticipation. This is a rather short horizon, affecting the representativeness of performance statistics used to compare the methods.

Objective of this TB-MPC application is to show its potential advantage. The system on which it is tested, made of a single reservoir, was selected for its simplicity of interpreting and communicating the results. Even if, in this specific system, the control problem could be solved with other off-line methods, similar on-line optimal control methods offer a more sensible benchmark.

The control methods against which TB-MPC is tested are: i) MPC, ii) MMPC [van Overloop, 2006], and iii) MPC under Perfect Forecasts (PF). MPC, in its original form as defined in Equations (3.1), neglects uncertainty completely. Deterministic inflow forecasts to the reservoir are obtained from the hydrological model, run with the deterministic weather forecast provided by NCEP. MMPC uses the same ensemble in a simple stochastic programming scheme. PF serves as reference, being an upper boundary of the achievable performances. PF conditions are emulated by running the MPC using, as forecasted disturbance, the observed inflow $\{d^{in}_{t,obs}\}^{k+h-1}_{t=k}$ in the optimization horizon window, where k is the closed loop simulation time index.

Results are summarized in Table 3.3.3, which shows that TB-MPC performs better than both MPC and MMPC. This suggests that EF contains valuable information for optimal control of water systems. Secondly, enlarging the control space by a tree is significantly preferable to simple stochastic programming. The latter result is meaningful because both TB-MPC and MMPC use EF, thus TB-MPC makes better use of the same information. When analyzed by sub-objective, TB-MPC scores better than MMPC on both the main objectives, i.e. energy production and flood upstream. All the other sub-objectives are not taken into account because, in the closed loop results, sub-objectives 3, 5, and 6 do not play any role and sub-objectives 7 and 8 are negligible.

Figure 3.2 shows the closed-loop water levels for the considered controllers. A good controller must correctly lower the water level in order to store the peaks. PF serves as reference point, showing the best possible trajectory. The emerging patterns are: i) even under PF conditions, some upstream flood still happens. ii) MPC tends to underestimate the losses due to peak events, not lowering the water level enough before them. iii) Along the entire horizon, MMPC tends to stay below the target level. This is due to its conservative behavior, forcing the system to stay safely un-

Controller	J_{rel} [-]	Spillage $\sum_t Q_t^{spill}$ [10e8 $m^3/year$]	Flood upstream $\sum_t g_1(h_t)$ [10e3]
PF	0%	6.7	9.0
TB-MPC	**37%**	**12.2**	**19.6**
MMPC	64%	13.1	25.2
MPC	100%	12.9	31.0

Table 3.2: Results summary for the different control methods. J_{rel} is J calculated for the closed loop simulation, standardized between zero (best case, PF) and one (worst case, MPC)

der the upstream flood threshold. Nonetheless, this precaution does not help before peak events. When it is really necessary to be conservative, MMPC does not anticipate enough. iv) Before expected peaks, TB-MPC is always the most precautionary controller. This behavior follows from the open-loop results of MMPC and TB-MPC.

Figures 3.3 and 3.4 give a better insight on how MMPC and TB-MPC work, also explaining why, before a peak event, MMPC does not lower the water level as TB-MPC does. MMPC, not considering the feedback in the open-loop optimization, finds a single control strategy, then the predicted open-loop water levels spread out along the open-loop horizon (Figure 3.4, bottom panel). For some members, water levels go below 30 m, activating sub-objective 3. To avoid possible costs in case of occurrence of a very low inflow, the controller limits the lowering of the water level. On the other hand, TB-MPC, correctly considering the possibility of changing control strategy, does not show this behavior (Figure 3.3, bottom panel). For each ensemble member, TB-MPC finds an control branch that brings the water level back to the target.

To give an insight into the computational complexity, MPC optimal control problem has $h \times N_u$ variables. In this test case $h = 61$ and $N_u = 2$ variables, and the optimal solution is found in about 2-4 seconds. In TB-MPC, the optimization argument has $m \times N_u$, where m depends on the tree structure, which changes at each time steps. For this test case and for a tree made of 15 branches, m is about 400-500, and the optimal solution is found in 15-22 seconds.

The advantage of TB-MPC over MMPC, in terms of avoided spillage, was 8.3×10^7 $m^3/year$, that if turbined at average conditions would produce approximatively $5.5 \times 10^3 MWh$, which monetized at the present energy price of 200 $US\$/MWh$ gives a revenue of 1.1×10^6 $US\$/year$.

3.4 Conclusion and Discussion

This chapter presented the Tree Based Model Predictive Control (TB-MPC), a method for control of water systems when the disturbance forecast is represented in form of ensemble. TB-MPC proved to be effective, and results from the test case indicates the advantages of TB-MPC over other existing on-line optimal control methods, including MMPC, that uses the same ensemble information.

Adaptivity of TB-MPC is limited to the preserved ensemble members which, even when properly selected, are a subset of all the possible disturbance trajectories. TB-MPC is not adaptive to deviations from the EF, or to jumps from one trajectory to another. The ensemble must correctly represent the relevant uncertainties affecting the controlled system. Processing of the ensemble forecasting is a required precondition, prior to its application in TB-MPC. A proper EF is a necessary condition for the effectiveness of the method.

TB-MPC is an on-line optimal control method. At each control instant, TB-MPC solves a closed loop optimal control problem where the feedback is limited to EF uncertainty. TB-MPC solves in fact a Multistage Stochastic Programming (MSP) problem. MSP is performed using a tree. The tree, enlarging the control space, makes the control strategy adaptive, i.e. capable to consider a different control strategy for each ensemble member (or set of them), to be applied from the moment that this ensemble member (or set) diverges from the others. This has an advantage in terms of control performances at the cost of a higher computational complexity, yet without reaching the obstructive complexity of some other methods, such as DP. Future works will explore how TB-MPC performs when scaled up to larger systems, as in [Ficchì et al., 2013], where the system is made of 55 states and 8 controls.

TB-MPC was tested on Salto Grande reservoir, a system consisting of a single reservoir. TB-MPC, however, not suffering the curse of dimensionality, can easily be applied to the control of large scale systems with a high degree of interconnectivity. EF are in fact particularly attractive to control systems consisting of multiple reservoirs, because EF captures the correlation among streamflows in multiple, neighboring basins [Faber and Stedinger, 2001].

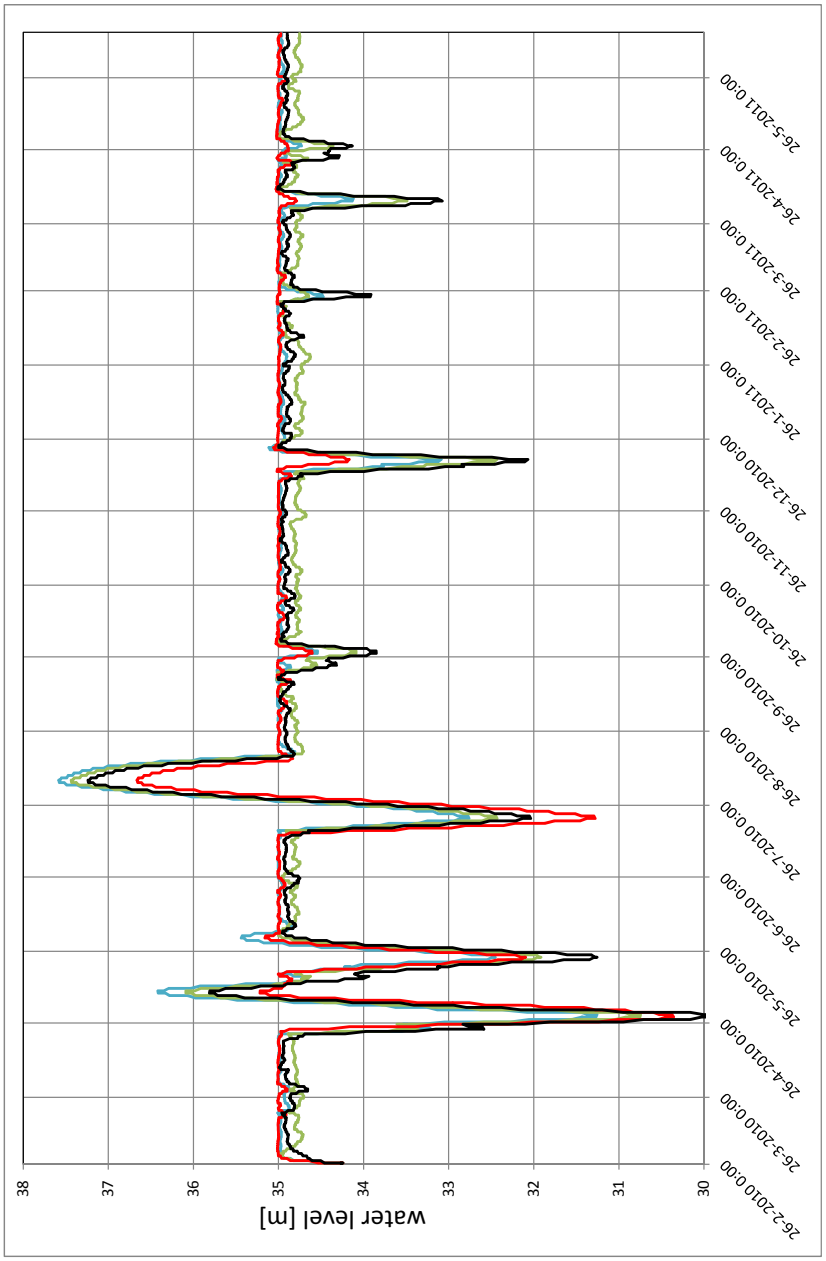

Figure 3.2: Water levels for the closed-loop horizon of TB-MPC (black line) MMPC (green line), MPC (blue line), and PF (red line)

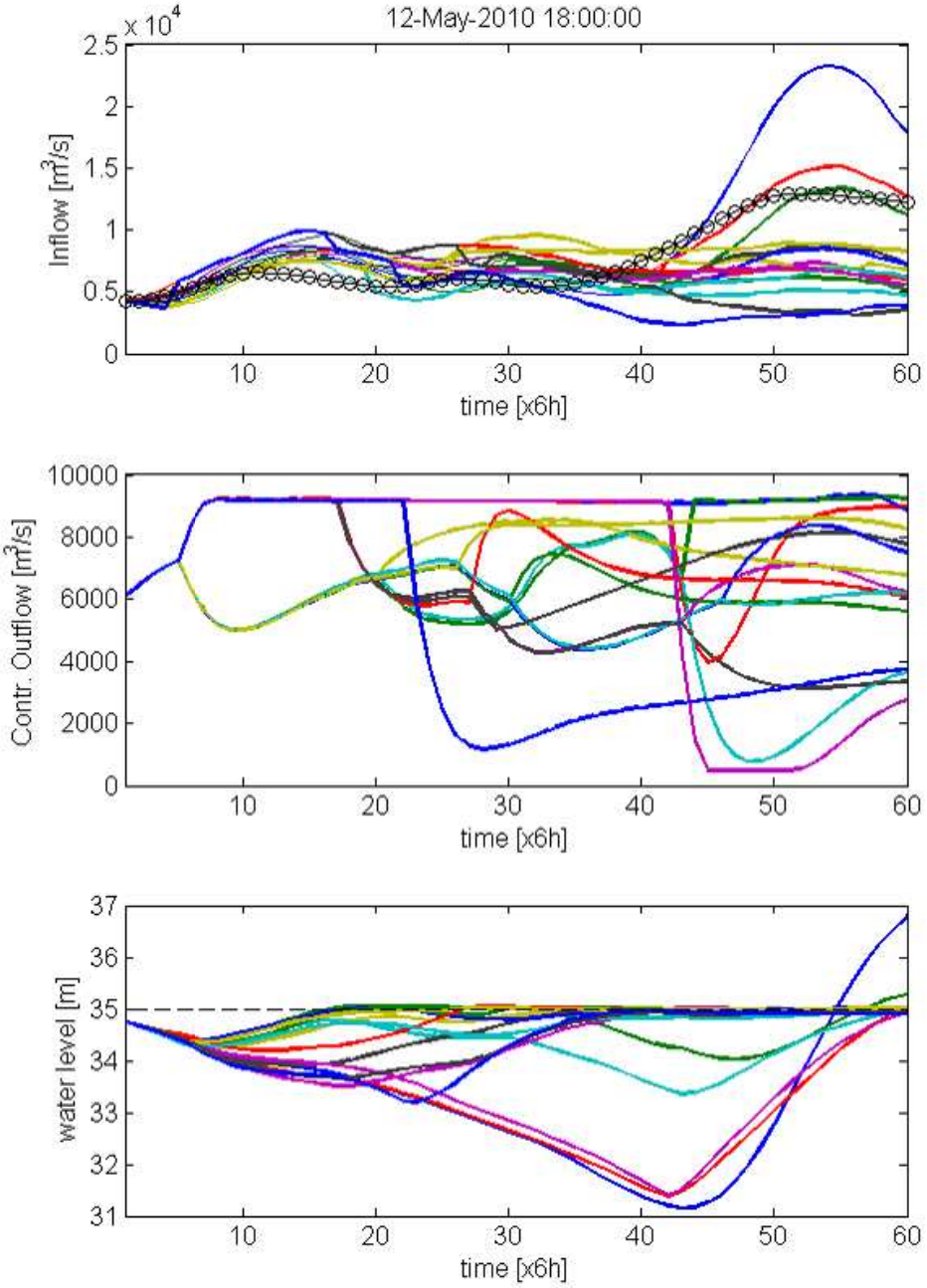

Figure 3.3: Open loop optimization results for the TB-MPC controller at 12 May 2010, 18:00. Top panel: disturbance forecast, EF of the inflow to the reservoir, and actual future inflow (circled line). Central Panel: optimal control tree, discharge from the reservoir. Bottom panel: open loop output, water levels

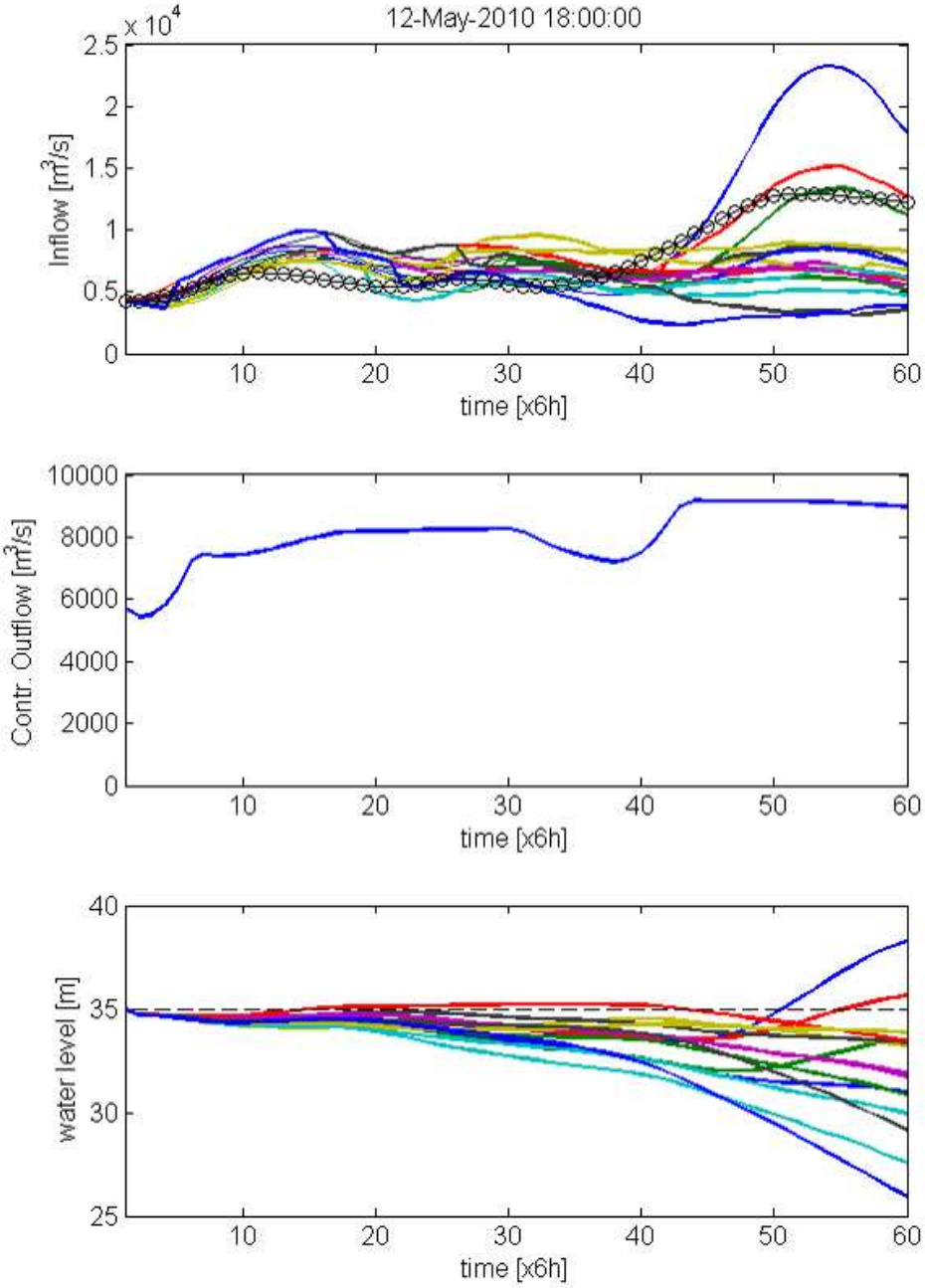

Figure 3.4: Open loop optimization results for the MMPC controller at 12 May 2010, 18:00. Top panel: disturbance forecast, EF of the inflow to the reservoir, and actual future inflow (circled line). Central Panel: optimal control, discharge from the reservoir. Bottom panel: open loop output, water levels

Chapter 4

Tree Structure Generation from Ensemble Data

This chapter presents a new methodology to generate a tree from an ensemble. The reason to generate a tree is to use the ensemble data in multistage stochastic programming, as seen in Chapter 3. A correct tree structure is of critical importance because it strongly affects the performance of the optimization. A tree, in contrast to an ensemble, specifies when its trajectories diverge from each other.

A tree can be generated from the ensemble data by aggregating trajectories over time until the difference between them becomes such that they can no longer be assumed to be similar, at such a point the tree branches.

The proposed method models the information flow: it takes into account which observations will become available, at which moment, and their level of uncertainty, i.e. their probability distributions (pdf). No conditions are imposed on those distributions.

The method is well suited to trajectories that are close to each other at the beginning of the forecasting horizon, and spread out going on in time, as ensemble forecasts typically are.

4.1 Introduction

When disturbance forecasts are available, this information can be integrated in the control strategy and the water management becomes *proactive*, Zavala et al. [2009]: before the realization of the disturbance, the controller sets the system to a state which is optimal to accommodate the expected disturbance. An example would be lowering the water level of a reservoir before an expected storm event in order to avoid floods. This generally enhances the reliability of the system.

However, weather is difficult to forecast and meteorological models can be wrong. Especially when using only one deterministic forecast, the control is more vulnerable

to forecast uncertainty, running the risk of taking action to counteract a predicted event that will not occur.

Ensemble Forecasting recognizes this uncertainty producing a large number of possible future trajectories. These are close to each other at the early stages of the forecasting horizon and spread out in time. Ensemble Forecasts can be used for control by setting up stochastic programming. This makes the control less vulnerable to the forecast uncertainty, especially in the presence of nonlinearity in the cost function or in the system to be controlled.

Ensembles have been directly used in stochastic programming, but this overestimates the uncertainty, as it does not consider its expected future resolution. In fact the proper problem formulation is Multistage Stochastic Programming (MSP). A tree serves the scope of embedding ensemble data in MSP, specifying the moments when some uncertainties are resolved.

Generating a tree from ensemble data is both difficult and of a critical importance. This can be done by bundling ensemble members at their initial stages, until the point where they are similar to each other. This is a different problem from that of other studies regarding tree generation, like those by Latorre et al. [2007] or Šutienė et al. [2010], in which the starting point is a much larger set of scenarios (or the conditional probability to generate them), then the tree is created by reducing their number, but the tree structure is given. On the contrary, when building up a tree from an ensemble, defining the tree structure is the core of the problem, and all the trajectories must be used.

Growe-Kuska et al. [2003] present a method to bundle complete scenarios based on a transportation metrics (Kantorovic distance), in which the problem is formulated as a minimal mass transportation problem. However, when this method is extended to tree generation, the aggregation rule must be slightly modified. In Heitsch and Römisch [2009] the authors recognize that this method is not appropriate to derive a tree structure from complete trajectories.

Defining the tree structure is part of the modelling process. It is recognized as a crucial problem in applications by Dupačová et al. [2003]. Defourny et al. [2011] already noticed that this field is still in a state where the scope of existing methods is not well defined, and the algorithmic description of the methods is incomplete, especially concerning the branching structure of the trees, despite its importance. Stive [2011] has shown in fact how the tree structure strongly affects the control performances.

We propose here a new methodology to produce a tree structure from ensemble data.

This approach models the information flow to the controller. This implies the explicit definition of the observations available in the future and their degree of uncertainty.

The paper is organized as follows. The next two sections introduce the ensemble and the tree as uncertainty models, the difference between them, and the way to use a tree in multistage stochastic programming. The subsequent methodology section is dedicated to the detailed description of the new approach for the tree structure generation. The test case section presents an application to the discharge forecast of the Salzach river. Implications, advantages, and limitations of the new method are discussed in the conclusion.

4.2 Stochastic programming using ensemble forecasts

An Ensemble Forecast (EF) is a prediction composed of representative trajectories of the possible future states of the system. The trajectories have generally small differences at the initial stage of the forecast, then they tend to diverge because of the chaotic nature of the underlying model. The most important Ensemble Prediction Systems (EPS) for short-medium term EF are based at the European Center for Medium-Range Weather Forecast (ECMWF) and the U.S. National Centers for Environmental Prediction (NCEP). They use global numerical weather prediction models.

The EF method offers the advantage of simulating uncertainty directly from a complex, physically based model. The trajectories produced are physically consistent and no derivation of a simplified stochastic model is required.

An Ensemble emulates the contemporary presence of a deterministic and a stochastic component assuming uncertainty among its members, but determinism within them. The ensemble trajectory is $y_t|z$, where t is the time instant, with $t \in T : \{1, 2, \ldots, H\}$, H being the length of the forecast horizon. Each trajectory comprises d different variables (for example ground temperature, total precipitation, wind speed, etc). z, representing the ensemble member, is a random variable; its sample space Ω is composed of N values, where N is the number of ensemble members, generally a number between 5 and 100.

$$z \in \Omega : \{1, 2, \ldots, N\} \tag{4.1}$$

Each ensemble member z has an assigned probability $p(z) \geq 0$, such that $\sum_{z=1}^{N} p(z) = 1$. Generally members are equiprobable, in this case the distribution of z is discrete uniform and $p(z) = 1/N \ \forall z$.

Ensemble Forecasts (EF) have often been used for optimal control van Overloop et al. [2008]; Doeswijk et al. [2006]; Lukasse et al. [2006]; Zavala et al. [2009]).

In the cited approaches, however, uncertainty is considered the same along the entire forecasting horizon. Implicitly, one assumes that there will be no future information to update the control. In reality, uncertainty differs if decisions are at an early or late stage of the forecast horizon: New information will become available on time, revealing which ensemble member is going to happen. When this uncertainty is resolved, the controller will change its control strategy, adapting it to the occurring ensemble member. Earlier decisions, depending on later ones, also improve when considering the future possibility to adapt.

4.3 Multistage stochastic programming using trees

In recursive decision making, when decisions are interdependent and some of them can be taken after uncertainty reduction, Multistage Stochastic Programming (MSP) (Birge and Louveaux [1997]) is the proper problem formulation.

In practical MSP applications the number of scenarios is limited, both because of calculation time and, as in ensembles, because they are originally few. In this case a tree serves the scope of setting those scenarios in a MSP setting. A tree, in fact, locates uncertainty resolution at specific moments along the forecasting horizon.

In a tree, up to the first branching point, all the trajectories are possible. From that point on, and at any other branching point, the entire set of possible scenarios splits in two subsets. A branching point represents the instant when uncertainty over which subset will occur is resolved. The occurred subset can be thought as determined by the result of a coin tossed at that instant. From this instant on, these subsets are mutually exclusive. Before a branching point, the series of control is optimal on average for all the considered trajectories. From the branching point onwards, two optimal control strategies (one per each branch) can be specified. The tree bifurcates repeatedly until the subsets are all made of a single scenario.

Compared to an ensemble, a tree is a different uncertainty model. A tree, specifying the moment when trajectories diverge, contains more information than an ensemble, which is merely a set of possible trajectories. Therefore, an ensemble cannot be used directly in MSP, but a tree structure must be generated from it first.

Dupačová et al. [2003] define a tree using a *scenario tree nodal partition matrix*. We do it here in the most compact way. A tree structure is in fact completely defined by assigning a *parent* $P(\cdot)$ and a *branching point* $B(\cdot)$ to each ensemble member. The

parent function $P(z)$ is defined from Ω to Ω and it indicates from which member the member z branches out. By convention, the parent power zero of a member is the same member, $P^0(z) = z$, and parent power n is the n-times recursive evaluation of the function, for example $P^2(z) = P(P(z))$. The branching point function $B(z)$ is defined from Ω to T and it indicates the time on the horizon when the member z branches out.

One ensemble member, called "root", is the one from which all the others branch out. It is a special member because it begins at $t = 1$ and it does not branch out from any other member, thus it has no parent. The convention used is that the root's parent is the root. In a tree there is only a single root member, i.e. $\exists! z | P(z) = z, B(z) = 1$, and the branching point of a member always follows the branching point of its parent, i.e. $B(P(z)) < B(z)$.

The formal definition of a tree is given in (4.2). The sample space of a tree Ξ depends on the time t and the considered member ξ.

$$\mathbf{\Xi}(\xi, t) = z | \bigcup_{k=0}^{n-1} \Big[P^k(z) = S, B(P(z)) > t \Big] \tag{4.2a}$$

$$S = P^j(\xi) | \min_j B(P^j(\xi)) < t \tag{4.2b}$$

Equation (4.2b) defines the *stem* member S. This is the member from which all other members in the considered sample space $\Xi(\xi, t)$ will branch out. Equation (4.2a) indicates the ensemble members (and all their future "offsprings") branching out from the stem member after t. The union stops at $n - 1$ because, in a tree of n members, that is the largest number of possible "generations".

Figure 4.1 presents an illustrative example. It shows the trajectories of a simplified ensemble, having only four members.

Table 4.1: Parent and branching point function for each member z of the ensemble in Figure 4.1

z	1	2	3	4
$P(z)$	1	1	1	3
$B(z)$	1	27	17	29

Table 4.1 contains all information needed to define the relative tree. In the picture, the branching points are highlighted using a dot and the part of the trajectory when a member has branched out using a slightly bolder line. For every sub-branch, the

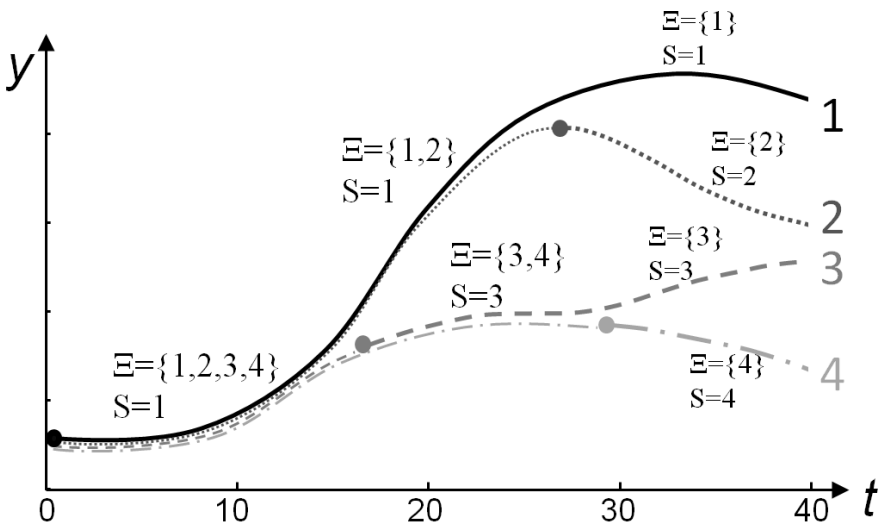

Figure 4.1: Illustrative ensemble. Sample space and stem member for every subbranch are specified.

sample space and the stem member are shown. These can be derived from Table 4.1 using Equation (4.2).

The MSP problem, defined using a tree is then:

$$\min_{u_t^z} \sum_{z=1}^{N} p(z) J(y_t^z, u_t^z) \tag{4.3a}$$

$$u_t^i = u_t^j \text{ when } \begin{cases} P(i) = j \\ B(i) < t \end{cases} \quad i, j = 1, \ldots, N \tag{4.3b}$$

$$\text{other constraints} \tag{4.3c}$$

Where $J(\cdot)$ is a generic cost function depending on the disturbances trajectories y_t^z and the control $u_t^z \in R^h \times R^N$, in the case that there is a single controllable value. Equation (4.3b) ensures the *non-anticipativity* (see Birge and Louveaux [1997]) of the control strategies.

4.4 The Information Flow Modelling approach: methodology

The premise of the Information Flow Modelling approach is that a decision strategy optimal for a specific ensemble member (or a set of them) is applied as soon as that

ensemble member (or that set) is considered certain to happen. The key idea is to identify this moment and to place the proper branching point there.

A branching point is in fact a pivotal element of a tree. It represents the instant when it is certain if its relative ensemble member (and its "offspring") has occurred or not. Along the horizon, new information will become available and this is expected to give, up to that moment, sufficient evidence to determine which of the two branches is the one actually happening. The approach we propose models how available information changes the ensemble member's probability taking explicitly into account which observations are available, when, and with what degree of certainty. Notice that observation is intended here in a broader sense than its usual one, following Weijs [2011]. Typically observation is understood as the sensing of the present state of a system using devices, but we include also the acquaintance of future states using models. Consequently, the definition of observational uncertainty is also different from the one typically used in hydrology and water resources (as by Götzinger and Bardossy [2008] for example), encompassing all the uncertainty sources. This will be shown and clarified further in the application.

At the forecasting instant, the ensemble member's probability is the a priori one. Going on in time, new information enters the system. Information at time t is composed of observations available at that instant I_t and their likelihood of coming from an ensemble member, $f_{I_t}(I_t|z)$. This function is in fact the observational uncertainty, i.e. the probability distribution of the observations.

When information is available, the ensemble member's probability will change in time, increasing if the collected observations are more likely to have been produced by this member, decreasing otherwise. Bayes' theorem describes how new information changes the conditional probability of an event (Jaynes and Bretthorst [2003]). Bayes' rule can be used to calculate the ensemble member's probability conditional to new observations. When the observation is continuous and one-dimensional, the form of the Bayes' rule is:

$$p_{t+1}(z|I_t) = \frac{f_{I_t}(I_t|z)p_t(z)}{\sum_{j=1}^{N} f_{I_t}(I_t|z_j)p_t(z_j)} \tag{4.4}$$

Equation (4.4) describes how new observations update the ensemble member probabilities. Here, $p_t(z)$ is the a priori ensemble member's probability, i.e. before making the observation, the left part of the equation $p_{t+1}(z|I_t)$ is the conditional probability of z given I_t, i.e. when I_t has been observed.

In Equation (4.4), the observations come only from the ensemble trajectories. The implicit assumption is that no deviations from the ensemble trajectories or jumps

from a trajectory to another one are allowed: the ensemble covers all the possible outcomes. This is clearly an approximation: the set of all possible trajectories is much larger than the ensemble (made of relatively few members). Nevertheless defining the possible trajectories by producing a sufficient number of them is a separate problem on which we do not have any influence.

Equation (4.4) can not be used directly for our purpose, because the tree has to be generated at the forecasting instant, when no observations have been made yet. Nevertheless, Bayes' rule can be used "on average" to calculate the expected evolution of the ensemble member probabilities. If one assumes that member z_k happens, observations I_t will be extracted from the distribution $f_{I_t}(I_t|z_k)$, and the new expected ensemble member's probability is

$$\underset{I_t|z_k}{E}\left[p_{t+1}(z|(I_t|z_k))\right] = \int_{-\infty}^{+\infty} p_{t+1}(z|I_t) \cdot f_{I_t}(I_t|z_k) dI_t \tag{4.5}$$

where $p_{t+1}(z|I_t)$ is Bayes' rule, the right side of Equation (4.4) written as a function of I_t only, with $p_t(z)$ given.

A possible interpretation of Equation (4.5) can be given following the decomposition of information in meaning and surprise of Applebaum [1996]. The observation I_t can be interpreted as the surprising component, and its meaning is given by its relation with the ensemble members, $p(z|I_t)$. Surprise is the quantity of information conveyed by the observation, increasing when the probability of that event decreases. Less expected observations convey much information. The function $p(z|I_t)$ represents the meaning, attaching significance to the new (more or less surprising) observation, indicating from which ensemble member it is likely to come, which is what we are interested in. Equation (4.5) represents thus the meaningfulness of the surprise that is expected at time t.

Equation (4.5) can be applied at every time step over the forecasting horizon to evaluate the expected evolution of an ensemble member's probability. Given the observation available if member z_k happened, it is possible to predict the probability dynamic of any z.

$$p_t(z) = p(z) \qquad \text{for } t = 1 \tag{4.6a}$$

$$p_{t+1}(z) = \phi(p_t(z), f_{I_t}(I_t|z), z_k) \qquad \text{for } t = 1, \ldots, H-1 \tag{4.6b}$$

Equation (4.6a) states that, at the first time step, the probability is the a priori one, $p(z)$, given from the ensemble. From that point on, probability evolves in time. In

Equation (4.6b), ϕ is a generic dependence function. It represents Equation (4.5) written in a form that clearly presents its inputs. Going on in time over the forecasting horizon and as long as more information becomes available, $p_t(z_k)$ approaches one and $p_t(z \neq z_k)$ approach zero.

Knowledge of the probability dynamics can be used to determine the instant in which an ensemble member is expected to be distinguishable from another. Information on distinguishability is contained in the *distinguishability matrix* D. Each element of this matrix, $D_{k,j}$, is the earliest moment at which, considering only two members at a time, z_k and $z_j (k \neq j)$, it is sufficiently certain that member z_k is the occurring one, given that the available observations I_t up to that moment come from this member. Sufficient certainty means that the probability of that event is larger than a predefined confidence level p^*, chosen sufficiently large (say equal to 0.95).

The procedure to build up $D(p^*)$ is described below. For every $D_{k,j}$ the sample space is reduced to the two considered members. The reduced sample space is thus $\Omega^r = \{z_k, z_j\}$ and its probability is $p_t(\Omega^r) = p_t(z_k) + p_t(z_j)$, being the events mutually exclusive. The probability of scenario z_k, considering the reduced sample space, is

$$p_t^r(z_k) = \frac{p_t(z_k)}{p_t(\Omega^r)} = \frac{p_t(z_k)}{p_t(z_k) + p_t(z_j)} \tag{4.7}$$

Then, for $\forall z_j \neq z_k$

$$D_{k,j}(p^*) = t | min_t \left[p_t^r(z_k) > p^* \right] \tag{4.8}$$

The distinguishability matrix is not defined on the diagonal. The matrix D depends, among other things, on the a priori member's probability. The matrix is a function of the confidence level p^*. A larger p^* requires more evidence from the observations, shifting ahead in time the instant when two members are distinguishable, corresponding to larger values in the D matrix. It can happen that two trajectories are so close to each other that, even at the end of the horizon, when all the available observation will have been collected, one ensemble member is not distinguishable with sufficient certainty from the other. In this specific case, $D_{k,j}$ will have an infinitive value.

The D matrix contains sufficient information to build up a tree, indicating, for each ensemble member, the most similar member and the instant they are sufficiently distinguishable. We describe now how to generate a tree from a distinguishability matrix. Starting from the complete D matrix:

1. Find the maximum value of D and the member corresponding to its row and column, called "row" and "column" member
2. The $\max D$ value and the column member are assigned to the row member as branching point and parent respectively
3. Matrix D is reduced, removing row and column of "row member" (its row and its column)
4. The procedure from (1) to (3) is repeated until a single member remains
5. The last member is the root, its parent is itself and its branching point is 1

Notice that a tree does not use the entire information contained in D. Choosing the maximum value implies considering only the case when a member is distinguishable from all the others. Fuzzy situations, when a member is distinct from some, but not all members, are ignored. This approximation obviates the need for a more complicated object than a tree, in practice a more general graph, with no limits on the number of "parents". Nevertheless, the objective of this study is to build up a tree, that it is considered a sufficient approximation, then the implicit assumption is that those fuzzy situations are of minor importance. The limits of the tree as uncertainty model will be shown in the application too.

4.5 Application

The method is tested on ensemble data for the Salzach river at Saltzburg. The basis of this ensemble is a meteorologic ensemble from ZAMG (Austrian meteorological office, Vienna) based on the INCA system, Haiden et al. [2010], with 50 ensemble members. The discharge at Salzburg is computed using COSERO (see Stanzel et al. [2008] and references therein), a continuous, semi-distributed rainfall-runoff model developed at the University of Natural Resources and Life Sciences (BOKU) of Vienna. The ensemble used was produced the 26th May 2011 at 8.00 am. The time step is 15 minutes and the forecast horizon is 3 days. The ensemble is reduced from 50 to 10 members using the scenario reduction algorithm of Growe-Kuska et al. [2003].

The approach requires the definition of the observable variables at every time step and their level of uncertainty. In river discharge forecasts, the observable variable is the water level, from which it is possible to derive the discharge using a rating curve. Di Baldassarre and Montanari [2009] show that this function is affected by uncertainty. We will follow here the method Di Baldassarre and Montanari [2009] proposed to analyze and quantify the uncertainty of this relation, extending their results to our test case.

Observations of water levels are assumed without uncertainty. The total error ϵ in the rating curve is normal and heteroscedastic, depending on the water level. It is composed of i) uncertainty in river discharge measurements and ii) rating curve uncertainty. The first component ϵ_1 is the error in the measurements of the rating curve.The second component ϵ_2 can be decomposed in: $\epsilon_{2,1}$, error in the rating curve interpolation-extrapolation, $\epsilon_{2,2}$, error due to the presence of unsteady flow conditions, and $\epsilon_{2,3}$, error due to the seasonal changes of the river roughness. Those errors are assumed to be all independent from each other and normally distributed, having zero mean and a quantified variance (see Appendix), thus

$$\epsilon = \epsilon_1 + \epsilon_{2,1} + \epsilon_{2,2} + \epsilon_{2,3} \tag{4.9}$$

and the distribution of ϵ is $N \sim (0, \sum_i \sigma_i^2)$.

The system observation, specifically the water level h, happens every 15 minutes, thus $I_t = h_t$ for $t = 1, 2, \ldots, H$. Notice that the not direct observability of the forecast variable is not a limitations, having a relation in probability between the observable variable (water level) and the variable of interest (discharge). Therefore

$$f_{q_t}(q_t|z) = f_{q_t}(q_t|h_t) \cdot f_{h_t}(h_t|z) \tag{4.10}$$

Where $f_{q_t}(q_t|h_t)$ is the probabilistic rating curve and $f_{h_t}(h_t|z)$ is a deterministic relation (having assumed no uncertainty in the water level observation). The confidence level is chosen sufficiently high ($p^* = 0.95$).

Table 4.2 shows the distinguishability matrix. Its asymmetry is also due to the different a priori probability of the ensemble members, being the ensemble made of 10 members obtained by reduction of the original one by deleting trajectories and aggregating probabilities.

Table 4.3 shows the tree structure.

Infinity as branching value means that that ensemble member never branches out from its parent. This happens when the trajectories of the considered member and its parent are so close that all the future observations along the horizon are not sufficient to distinguish them with the required certainty.

Figure 4.2 shows the trajectories of the original ensemble (thin lines) and that of the tree produced from it (bold lines). The numbers label the ensemble members at their branching point. The most different trajectories bifurcate first. Going on in time, also the more similar ones are distinguishable and branch out. The yellow-greenish trajectory (ensemble member 6) has an interesting behavior: despite its

Table 4.2: Distingushibility matrix obtained from the ensemble of Figure 4.2.

$z_k \backslash z_j$	1	2	3	4	5	6	7	8	9	10
1	NaN	55.5	27.5	26.25	24.25	35.5	34	27.25	27	27.75
2	Inf	NaN	28.75	27.5	27.25	44.5	35.75	28.5	27.75	39.25
3	41	41.25	NaN	35.75	59.5	36.5	59	66.5	45.5	49
4	29.25	29.5	31.75	NaN	33	28	32.25	34.5	41.75	30.25
5	20.25	24	20.75	21.75	NaN	18	28.75	34.75	32.5	25.75
6	44.5	45.25	49.25	30.25	43.75	NaN	49.5	32.75	30.5	56.25
7	Inf	56.75	48.75	32.25	51	30.25	NaN	42.75	35.5	45.25
8	33.75	35	62.5	37.5	51.75	27.25	53.75	NaN	Inf	35.5
9	27.5	28	31	44	38.75	26	32.25	46.5	NaN	28.75
10	44.25	45.25	54.75	31.25	48.75	54.75	53.5	36.5	32.25	NaN

Table 4.3: Tree structure defined by Parent and Branching Point for each ensemble member z, obtained from the D matrix (Table 4.2). The branching point values are expressed in hour

z	1	2	3	4	5	6	7	8	9	10
$P(z)$	5	1	5	5	5	10	1	9	4	5
$B(z)$	24 15’	Inf	59 30’	33	1	56 15’	Inf	Inf	44	48 45’

deviation from other ensemble members (for example from the light-blue trajectory, member 4) already at some earlier stages, it branches out relatively late. This happens because the algorithm that build up a tree from the D matrix considers only the highest value. Thus, an ensemble member branches out only when it becomes univocally recognizable from any other one, including the most similar one, which in this example is the red trajectory (member 10). Nevertheless, this limit is a characteristic of the tree as uncertainty model and not of the specific tree generation algorithm, as explained in section (4.4).

In Figure 4.3 two ensemble members are separated. Plot A shows their trajectories, plot B shows the probability of z_k, assuming that that is the happening ensemble member. On the first part of the horizon, the trajectories are very close, the observational pdf's $f_{q_t}(q_t|z)$ overlap, thus observations here are not informative. When the trajectories diverge, observations are less ambiguous and the confidence that z_k is occurring increases. After some time, from hour 40 onwards, it is already clear that the trajectories have diverged, thus the additional observations do not add much more information. The branching point is placed when this confidence is sufficiently high. The graph below shows also to what extent the branching point is postponed

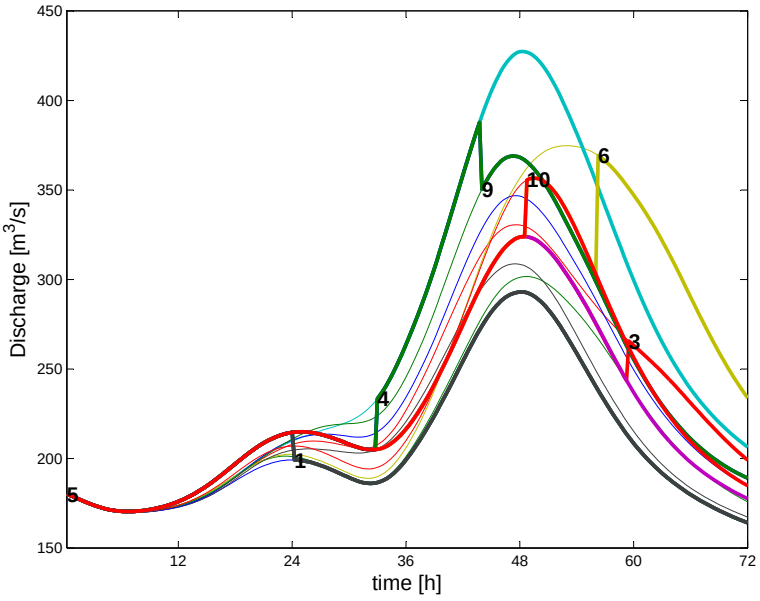

Figure 4.2: The reduced ensemble at Saltzburg, 26th May 2011 at 8.00 am (thin line), and the tree generated from it (bold line), defined in Table 4.3.

when the confidence level p^* increases.

If the control is applied every 6 hours, then observations with a higher frequency cannot be used until the new control is chosen. Therefore at every control time step (every 6 hours) the last 24 past observations are available (one observation every 15 minutes for six hours). The available information is in this case $I_t = \{h_\tau\}_{\tau=t-23}^{t}$

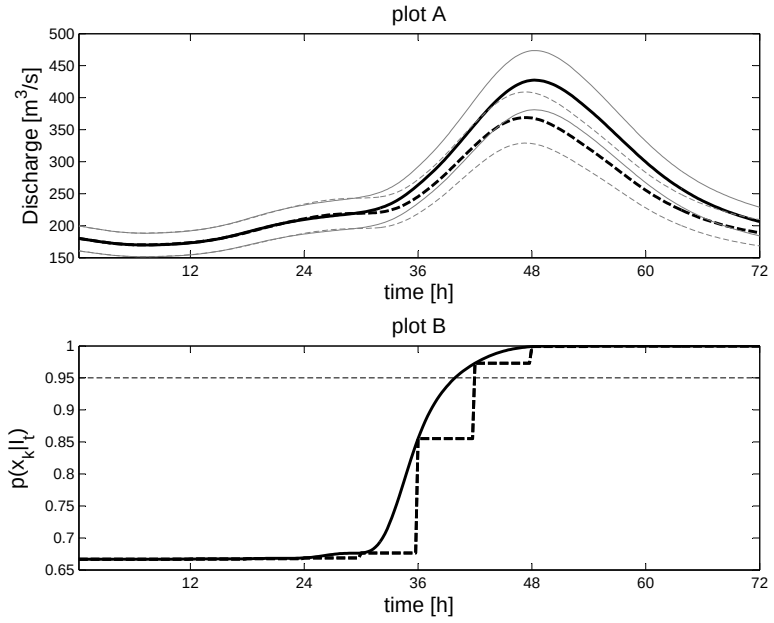

Figure 4.3: Plot A: Trajectories of two ensemble members (bold lines) and the $\pm\sigma$ bounds of their observational uncertainty (thin line). Plot B: the evolution of the reduced probability dynamic for z_k on time (Equation 4.6) when $I_t = \{y_t\}$ $t = 1, 2, 3, \ldots, H$ (continuous line) and $I_t = \{y_\tau\}_{\tau=t-23}^{t}$ $t = 24, 48, 72, \ldots, H$ (dashed line).

for $t = 24, 48, 72, \ldots, H$. I_t is a 24-dimensions variable, assuming independence of the errors in time, then $I \sim \mathcal{N}(0, \Sigma)$, and $\Sigma = I \cdot \sigma_\epsilon$ (I is the identity matrix). The dotted line in Figure 4.3, plot B shows the probability dynamic for this case. The

trend has steps corresponding to the moments when information enters the system and a new control can be applied. The information flow has a coarser discretization and the branching points are postponed.

Until now, we have considered the information gained observing the water level at every decision time step. Nevertheless, the river stage is not the only available observation. Considering the broader definition of observation that we gave, the controller disposes also of a forecasting system that every hour provides an estimation of the future discharge values. This information source should be modeled too.

The information flow including the future forecasts comprises $I_t = \{q_t\}_t^H$ with $t = 4, 8, 12, 16, \ldots, H$. The forecast information will not be tested here because of the difficulty to integrate a H-dimensional variable (in this ensemble H is 288). In case, it would be necessary to define an uncertainty model of the ensemble forecasting as well.

4.6 Conclusions

This paper presents a new methodology to create a tree from ensemble data, called Information Flow Modelling. The methodology models how information becomes available to the system controller in time. Information consists of the observations and their uncertainty. Observational uncertainty $f_I(I|z)$ is also the link between the probability of observing I to the ensemble member z.

Just as in the information theory framework described in Weijs et al. [2010], observations are seen here as a communication process in which uncertainty about the outcome of a random event is reduced by delivering an informative message to the user, in this case the controller. Information Flow Modelling approach requires the explicit definition of what the controller can observe and its level of certainty. These requirements give to the method a formal character and guarantee a strong theoretical basis.

The ease of use has been shown by applying the method to a test case. Differently from other methods, information on the conditional probability in time is not required. All one needs to define is the observational uncertainty $f_{I_t}(I_t|z)$ and the confidence level p^*. The first is more difficult to determine, although in the test case it was already defined. Hydrologists have recently put more and more effort into associating a distribution characterization to a forecast, see for example Schaefli et al. [2007], Pianosi and Raso [2012], Montanari and Brath [2004], Solomatine and Shrestha [2009] and Schoups and Vrugt [2010]. The confidence level is a threshold

value, representing the certainty level required for considering a ensemble member as the one actually happening. Typical values are 0.95 or 0.99. Its value is rather arbitrary, as any other hypothesis testing threshold, and its influence on the control performances should be tested.

Another important property of this approach is its generality. The function $f_{I_t}(I_t|z)$ can be any type of distribution, for example, a discrete one (rain, no rain) or a deterministic relation. In the application shown it is a normal distribution, but this is not a generally required condition.

Further research will investigate the influence that the tree structure has on the control performance when trees are generated by the Information Flow Modelling approach, applying it to the real time control of dynamic systems.

Chapter 5

Conclusions and recommendations

5.1 Contribution of this thesis

In the context of operational water management, controlling the hydraulic structures by deterministic optimization offers many advantages, reviewed in Chapter 1. A deterministic optimization can harmoniously manage complex systems even under stringent constraints. It requires a model of the physical process and an objective function.

Uncertainty affects the performance, for unexpected events can drive the system off from the optimal conditions. Stochastic optimization explicitly takes into account the presence of uncertainty and reduces the performance loss due to it. Uncertainty can be present in any element of the control problem: the system model, the state-output relation, the objective function, the initial state estimation, and future disturbances. In water systems, the controlled components are generally hydraulic structures (canals, reservoirs, or a combination of them), which are well known, when compared to the hydrological or meteorological components, where the uncertainty is much larger. Meteorological and hydrological systems are both upstream of the controlled system, thus are not affected by the control. They take part in the control problem forming the disturbances that can be partially predicted. Even if disturbances can be forecasted, a large uncertainty still remains. This thesis focused on the forecast uncertainties relevant for the real time control of water systems.

Errors in the hydrological model can, under some conditions, be much larger than under regular conditions. I showed this in Chapter 2, where I also showed a simple and effective dynamic modeling of the error standard deviation, based on a regression using the absolute error. The approach has been called Dynamic Uncertainty Modelling By Regression on Absolute Errors (DUMBRAE). The standard deviation model can be built up using the time series of residual errors only, independently of the hydrological model. The necessary assumptions on the error distribution may

necessitate some appropriate transformation when the assumptions are not satisfied. The method proved to be effective under any indicator, including maximum likelihood. DUMBRAE outperformed all other methods to which it was compared and showed to be able to effectively reproduce the heteroscedasticity of the residual errors, both under a fast response catchment, as in the Lake Maggiore test case, and a slower process dominated by snowmelt, as in the Rhone River.

In Chapter 3 and 4, I have shown a way to use ensemble forecasts within the MPC framework by building up a tree and using it to implement multistage stochastic programming. In order to set an optimal control problem, the analyst typically models the physical system, defining how the water flows from one component to another. Part of my work has shown the importance of modelling a different flow, the information flow, made of observations (in a broad sense) that the controller is able to make. The information flow modeling is part of the problem setting. The future uncertainty resolution affects the present decisions; this is counterintuitive, but taking it into account affects (positively) the control performances.

Chapter 4 showed a method to build up a tree structure from a set of complete scenarios by quantifying the expected level of certainty along the control horizon. This is an advantage when compared with other existing methods that are empirically based or adapted from other problems. The approach has been called *Information Flow Modelling*, because the tree is built up following how and when information becomes available to the system controller. Information consists of the observations and their uncertainty, therefore this method requires the explicit definition of what the controller can observe, when, and its level of certainty. Further research should investigate the influence that the tree structure has on the control performance when trees are generated by the Information Flow Modelling approach, applying it to the real time control of dynamic systems.

In Chapter 3, the tree is then used in the proposed control algorithm called *Tree-Based Model Predictive Control* (TB-MPC). This control method considers the feedback mechanism of MPC. It does not search for an optimal control time series, but for many control strategies, one for each branch of the tree. The tree is in fact used to enlarge the control space in order to set up multistage stochastic programming. This makes the control adaptive. Even if, eventually, only the first control is applied to the system, considering the future capacity to change strategy improves the closed loop results. TB-MPC has been applied for the operation of Salto Grande reservoir, at the border between Uruguay and Argentina, which is a simple system made of a single reservoir. TB-MPC, not suffering of the curse of dimensionality, shows its advantages especially on large, hardly constrained, strongly connected systems.

In this thesis, I used the tree for real time control. The method is general and can be extended to any type of scenario analysis. For example, consider climate change scenarios. If climate change trees are produced, we would have available, apart from the scenarios themselves, the information about when it will be known (with sufficient certainty) that each scenario is actually the happening one.

5.2 Recommendations for Further Research

As any good research, my work raised more questions that it answered. Some of them are presented here.

5.2.1 Balance between model complexity and uncertainty

A practical question, central for any analyst modelling a controlled system under uncertainty, is how much detail to reserve for physics or uncertainty modelling. Computer power is in fact limited and an optimal compromise between the two needs to be found.

In the last decades, computer power has grown following Moore's law. This exponential growth gives the idea that computation capacity will, at some point, overcome all the limitations that we now face. Computer power allocation seems just a temporary problem that will be wiped out in some future. However, we do not believe in eternal exponential growth. Moore's law is just an empirical extrapolation of past data. There are no reasons to think that it will grow infinitely. Actually, it is more sensible to think that, at some point, for some physical constraint, it will saturate. Moreover, if we also look at the recent history of environmental modeling, we will see that the system boundaries have grown with time too [Washington et al., 2009]. The bonanza of computer power has been used for more integration. The trend is, and very likely will continue to be, towards more integrated models, and not, as someone stated, towards more spatial and temporal model accuracy. There will always be a limited computer capacity then, and always the need for a trade-off in its allocation.

Uncertainty is very computational demanding. In the worst case, i.e. without some smart algorithm, treating uncertainty requires multiplying the problem complexity for the number of realization of any stochastic variable. At the same time, uncertainty is very significant for decision making.

The question on which uncertainties to include has not a clear solution yet and it is left to the analyst's subjective choice. Some general rule, guiding the analyst in the problem setting, would be a valuable tool. This rule should quantify the benefits, in

terms of performance improvement and generalized costs, brought by an extended problem, and then suggest what to include next and when to stop.

5.2.2 Optimal uncertainty reduction

In order to manage a system, a decision maker (DM) tries to make the best decision under uncertainty, having partial knowledge of the effects of his/her decision. However, instead of just accepting the present uncertainty, he/she can reduce it by getting more observations. Uncertainty is lack of information and can be reduced by new observations. Observation is intended here in a general sense, as a carrier of a piece of information. Possible observations are, for example, getting a forecast, building up a model, or sensing the system at some location for some variable.

Observations reduce uncertainty, but are costly. The DM should then reduce uncertainty and, at the same time, minimize the costs due to observation until an optimal level of *rational ignorance* Simon [1990] is reached. Uncertainty reduction is in fact not an end to itself, but aimed at taking better decisions; at some point, accepting uncertainty will be more effective than reducing it further. I called this the *Optimal Observation Problem* (OOP) [Raso et al., 2013], suggesting that, in addition to the optimal control problem, there is a secondary decision, linked to the first, concerning the most effective observation. OOP suggests what to observe next and if the next observation is worth its cost.

5.2.3 Defining the Objective Function

Optimal control is essentially selecting a control value given a process model and an objective function. Presently, a large part of the research focus on improving the model representing the physical process; the objective function is simply given. Nevertheless, the objective function determines the final results as much as the process model does.

In MPC, the loss function is often nothing more than a quadratic function [Camacho and Bordons, 2004]. But this is mostly a legacy of previous MPC applications, where the typical problem was trajectory tracking. The reason to stick to a quadratic function are mostly historical. Secondary, but not less important, the success of quadratic cost function is linked to the ease to solve a quadratic programming problem. However, now that we have better solving algorithms [van den Boom and Backx, 2003] a quadratic cost function is no more a necessity.

In water resources management, researchers often focus on the presence of multiple objectives. There is in fact a large research production using multi-objective

optimization, for which we have analytical instruments. The presence of multiple stakeholders with conflicting requirements from the same water system is a common situation, and a multi-objective analysis formalizes and offers a solution to this conflict. The output of a multi-objective optimization is generally a Pareto front, a graph that is supposed to inform the decision maker about the possible optimal trade-offs among the conflicting objectives. Even if multi-objective analysis is a valuable tool, multiplying the objectives adds complexity and should be limited to cases when the objectives are really irreducible, because of incommensurability or irreconcilable groups.

Nevertheless, even when the objective is unique, or aggregated, objective function definition is a complex task. It requires identifying the interests at stake in the system and reducing them to a mathematical object. Castelletti and Soncini-Sessa [2006] do this using a participative process, and Fishburn [1982] shows how to translate statements and preferences into a cardinal function. These are useful tools to define what I call the *subjective rationality*, which does not ensure the *objective rationality*.

As remarkable example, let us focus on risk aversion, intended as the nonlinearity of the loss function to negative effects [Hashimoto et al., 1982]. Imagine that risk aversion can be summarized in a single parameter; the analyst estimates it by getting the necessary information from the stakeholder. Even if the stakeholder is satisfied with the analyst output, this is a subjective assessment. It could be the case that the stakeholder is risking too little, or too much. Subjective rationality, got from the interaction with the stakeholder, does not necessarily provide objective rationality, represented by the "optimal" risk aversion. Objective rationality is here defined as the probability of surviving. This definition is teleologically based because entities are selected based on it. Those who did not have this objective, are already extinguished. Thus an objectively rational decision maximizes the probability of surviving.

Even if useful as reference point to reject the relativism of subjective rationality, this definition has little usefulness in practice. No model has surviving probability as output. Therefore the research question, on how to define a good objective function, is still an open one, bringing with it some related problems, first of all the *verifiability problem*.

The procedure to define the objective function should not be different from that of a physical model. Hypotheses are formulated, expressed in mathematical formulas, and eventually verified. In hydrology, for example, the way to verify the goodness of a model is typically by looking at the difference between the predicted and the observed values, trying to minimizing this distance. Objective functions should be verified as well, and if the quality criterion for a model is to make good predictions,

the quality criterion for an objective function is to make (objectively) good decisions. Presently, objective functions are not verified; the acceptance by the stakeholders is considered a sufficient criterion for its quality.

5.2.4 Closed Loop Stability

One of the main drawbacks of MPC is that the horizon finiteness introduces the issue of *closed loop stability* [Bemporad and Morari, 1999]. Closed loop stability ensures that the series of decisions obtained from the open loop optimization do not bring the system state away from the equilibrium in the closed loop iteration. It can be compared to the idea of sustainability, i.e. the present decision about the resource use does not lead to its over-exploitation.

The control theory community has developed many theorems dedicated to finding sufficient conditions for stability. In a finite horizon scheme, stability can be guaranteed by a cost function that is Lyapunov, or a terminal constraint (or penalty) on the state of the last time step. However, for reservoir management, the cost function is not always of the Lyapunov type, and obtaining a terminal penalty $g_h(x_h)$ is not trivial. The terminal penalty should summarize all the costs from h to infinity, which is the costs of having left the system in state x_h. A possible solution to this problem is using, a terminal cost function obtained from DP. However, if the DP problem can be formulated, MPC looses its usefulness.

An infinite horizon guarantees stability [Maciejowski, 2002]. Further research in this direction is needed, to apply MPC to systems with memories longer than control horizon, as reservoirs typically are. Therefore, if MPC for reservoir management wants to reach a mature state, exploring different forms of infinite horizon MPC is a priority.

5.2.5 Decision Support Systems (DSS) vs Automatic Control

Looking at the history of control [Bennett, 1996], automatic control has proved to be an effective solution in a variety of fields. For industrial application, experts initially thought that optimal control algorithms were aimed at suggesting the decision only. Nowadays the automation is pervasive, and for some systems close to complete. The question for water resources management researchers is whether the same trend will rise in water systems. Even if water seems to be different from other commodities [Savenije, 2002], the super-positivist hypothesis of a completely automatic system cannot be excluded.

My opinion is that the implementation of automatic control in water management encounters an essential difference with respect to the other classical applications. Gourbesville [2008] stresses the complexity of river management, thus the need for DSS. However, it is not so much the physical complexity that hampers the automatization (a powerplant is also complex), but the political complexity, i.e. the presence of conflicting interests, that can generate a resistance to delegation of responsibility to an algorithm. Moreover, water is a strategic resource[1], too important to bypass the presence of a human decision maker for its management. Nonetheless, his/her decision can anyway be supported by the application of control theory techniques within a DSS.

Process Oriented Algorithms for an Operational DSS

Experience has shown that a suggested optimal control can be meaningless for the decision maker, who will eventually discard it. Borowski and Hare [2007] documents the existence of a gap between the research community developing model based tools, which sees the innovation of such instruments and the potential improvement in performance of water systems, and the decision makers at an operative level, who do not understand the meaning of the tool output.

Researchers, instead, tend to neglect the importance of an established existing procedure developed to take decisions using all available information, and assume that introducing the tool will reshape straightforwardly the decision procedure. In applications where the tool radically changes the procedure, this approach can lead to the failure of its effective implementation. Process Oriented Algorithms propose an integrated design of the model based tool and the desired process. Its objective is to close the gap between the existing optimization algorithms and the decision making process, by analyzing the information requirements of the decision maker, as suggested by Janssen et al. [2009],

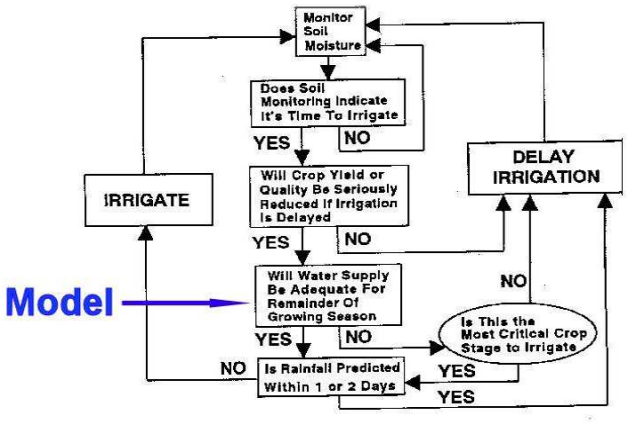

Figure 5.1: a simple possible scheme of the decision process and the role of the model within it.

1. For the same reason, its strategic nature, water cannot be reduced to a mere economic good

and shaping both the tool and the decision making process to adapt them to each other.

Appendices

Appendix A: Expected value of the absolute value of a normal variable

First it can be proved that if $X \sim N(0, \sigma^2)$ and $Y = |X|$, then

$$f_Y(y) = 2 \cdot f_X(x) \quad \forall x \in \mathbb{R}, y = |x|$$

where $f_Y(\cdot)$ and $f_X(\cdot)$ are the pdfs of X and Y respectively. In fact, from the definition of Y, it follows that $P[Y < y] = P[|X| < y] = P[-y < X < y]$ and thus

$$\int_0^y f_Y(\xi)d\xi = \int_{-y}^y f_X(\xi)d\xi$$

Since $f_X(\cdot)$ is an even function,

$$\int_0^y f_Y(\xi)d\xi = \int_0^y 2 \cdot f_X(\xi)d\xi$$

from which the thesis follows.

Therefore the expected value of Y is given by

$$E(Y) = \int_0^\infty y\, 2 f_X(y)\, dy = \frac{2}{\sqrt{2\pi}} \int_0^\infty \frac{y}{\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = \frac{2}{\sqrt{2\pi}}\sigma = k\sigma \qquad (5.1)$$

Appendix B: Periodic model of the standard deviation based on Fourier decomposition

A periodic model of the error standard deviation is

$$\sigma_t = g(t, \boldsymbol{\phi}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \sin\left(\frac{2\pi}{T} it\right) + \beta_i \cos\left(\frac{2\pi}{T} it\right) \qquad (5.2)$$

where T is the time period (for instance, if the modelling time step is 24 hours, $T = 365$ day) and t is the current time step counter. Identification of model (5.2) is

straightforward, in fact, for given order n of the Fourier decomposition, the coefficients $\boldsymbol{\phi} = |\alpha_0, \alpha_1, ..., \alpha_k, \beta_1, ..., \beta_k|$ can be estimated by least squares according to the well known formula

$$\boldsymbol{\phi} = \left(\sum_{t=1}^{N} \mathbf{u}_t \ \mathbf{u}_t' \right)^{-1} \sum_{t=1}^{N} \mathbf{u}_t \ y_t$$

where N is the number of data, the regressor vector $\mathbf{u}_t$ is given by

$$\mathbf{u}_t = \left[1, \ \sin\left(\frac{2\pi}{T} t\right), ..., \ \sin\left(\frac{2\pi}{T} n \ t\right), \ \cos\left(\frac{2\pi}{T} t\right), ..., \ \cos\left(\frac{2\pi}{T} n \ t\right) \right]$$

and the observation y_t is simply given by $k \ |e_t|$, according to equation (2.9).

Appendix C: Additional information to Chapter 3

5.2.6 Programming environment

The hydrological model, the reservoir model, and the optimization are programmed in RTC-Tools, a open source, free software developed by Deltares Foundation. EF's can be freely downloaded from the TIGGE portal.

5.2.7 Hydrological model

The hydrological model is a semi-distributed HBV model [Bergström et al., 1995] with a topology according to Figure 5.3, simulating both the rainfall-runoff process and the routing process. The routing process is modeled by a triangular unit hydrograph. The model is calibrated on one and half year of data, from December the 28th, 2009 to June the 28th 2011, using weather data from five weather stations and observed discharge from three hydrological stations (see Figure 5.2). The HBV model inputs are precipitation, temperature, monthly average temperature, and monthly average evaporation. The Nash index, as a measure of model performance, scores 0.74.

Appendix D: Observational uncertainty estimation

The variance of ϵ_1 is: $\sigma_{\epsilon_1}(t) = 0.027 \cdot q_t$

For the errors $\epsilon_{2,i}$, Di Baldassarre and Montanari [2009] give the 95% of the error falls in the α_i% of q_t, thus $\sigma_{\epsilon_{2,i}}(t) = \frac{\alpha_i \cdot q_t}{1.96}$, where 1.96 is the 97.5-quantiles of a standard normal distribution, that takes into account the 2.5% on both the left and the right tail of the distribution.

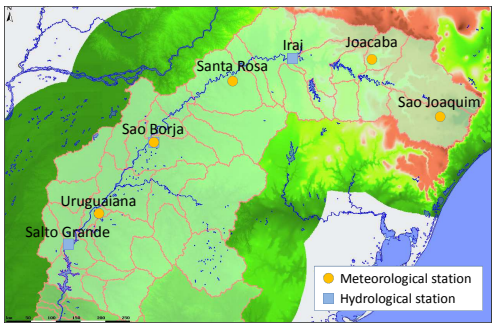

Figure 5.2: Catchment of the Uruguay River with location of the weather and hydrological stations

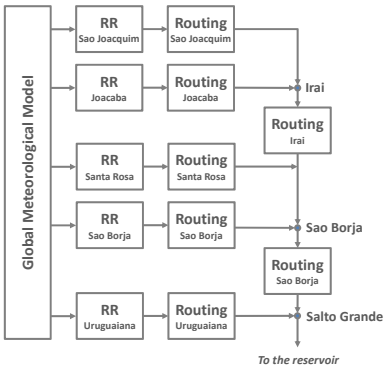

Figure 5.3: Diagram of the hydrological model

References

Applebaum, D., 1996. Probability and information: An integrated approach. Cambridge Univ Pr, Cambridge.

Astrom, K., 1995. Pid controllers: theory, design and tuning. Instrument Society of America.

Bates, B. C., Campbell, E. P., 2001. A markov chain monte carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resources Research 37 (4), 937–947.

Batina, I., 2004. Model predictive control for stochastic systems by randomized algorithms. Technische Universiteit Eindhoven.

Bellman, R. E., Dreyfus, S. E., 1966. Applied dynamic programming. Vol. 7962. Princeton University Press.

Bemporad, A., Morari, M., 1999. Robust model predictive control: A survey. Robustness in identification and control, 207–226.

Bennett, S., 1996. A brief history of automatic control. Control Systems, IEEE 16 (3), 17–25.

Bergström, S., Singh, V., et al., 1995. The HBV model. Computer models of watershed hydrology., 443–476.

Bertsekas, D., 1995. Dynamic programming and optimal control. Vol. 1. Athena Scientific Belmont, MA.

Bertsekas, D. P., 1976. Dynamic programming and stochastic control. Vol. 125. Academic Pr, Belmont, MA.

Birge, J. R., Louveaux, F., 1997. Introduction to stochastic programming. Springer Verlag.

Biswas, A. K., 2004. Integrated water resources management: a reassessment. Water international 29 (2), 248–256.

Biswas, A. K., 2008. Integrated water resources management: Is it working? Water resources development 24 (1), 5–22.

Borowski, I., Hare, M., 2007. Exploring the gap between water managers and researchers: difficulties of model-based tools to support practical water management. Water Resources Management 21 (7), 1049–1074.

Bowden, G. J., Dandy, G. C., Maier, H. R., 2005. Input determination for neural network models in water resources applications. part 1-background and methodology. Journal of Hydrology 301 (1), 75–92.

Box, G., Jenkins, G., Reinsel, G., 1970. Time series analysis. Holden-day San Francisco.

Camacho, E. F., Bordons, C., 2004. Model predictive control. Springer Verlag.

Cardwell, H. E., Cole, R. A., Cartwright, L. A., Martin, L. A., 2006. Integrated water resources management: Definitions and conceptual musings. Journal of contemporary water research & education 135 (1), 8–18.

Castelletti, A., de Rigo, D., Rizzoli, A. E., Soncini-Sessa, R., Weber, E., 2007. Neuro-dynamic programming for designing water reservoir network management policies. Control Engineering Practice 15 (8), 1031–1038.

Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2010. Tree-based reinforcement learning for optimal water reservoir operation. Water Resources Research 46 (9).

Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2008. Water reservoir control under economic, social and environmental constraints. Automatica 44 (6), 1595–1607.

Castelletti, A., Soncini-Sessa, R., 2006. A procedural approach to strengthening integration and participation in water resource planning. Environmental Modelling & Software 21 (10), 1455–1470.

Castelletti, A., Soncini-Sessa, R., 2007. Coupling real-time control and socio-economic issues in participatory river basin planning. Environmental Modelling & Software 22 (8), 1114–1128.

Cloke, H., Pappenberger, F., 2009. Ensemble flood forecasting: a review. Journal of Hydrology 375 (3), 613–626.

Day, G., 1985. Extended streamflow forecasting using nwsrfs. Journal of Water Resources Planning and Management 111 (2), 157–170.

De Nicolao, G., Magni, L., Scattolini, R., 1996. On the robustness of receding-horizon control with terminal constraints. Automatic Control, IEEE Transactions on 41 (3), 451–453.

Defourny, B., Ernst, D., Wehenkel, L., 2011. Multistage stochastic programming: A scenario tree based approach to planning under uncertainty. LE, Sucar, EF, Morales, and J., Hoey (Eds.), Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions. Hershey, Pennsylvania, USA: Information Science Publishing.

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrology and Earth System Sciences 13 (6), 913–921.

Doeswijk, T. G., Keesman, K. J., van Straten, G., 2006. Impact of weather forecast uncertainty in optimal climate control of storehouses. In: Proc. of 4th IFAC/CIGR workshop on Control Applications in Post-Harvest and Processing Technology (CAPPT 2006), Potsdam, Germany.

Draper, A. J., Jenkins, M. W., Kirby, K. W., Lund, J. R., Howitt, R. E., 2003. Economic-engineering optimization for california water management. Journal of water resources planning and management 129, 155.

Dupačová, J., Gröwe-Kuska, N., Römisch, W., 2003. Scenario reduction in stochastic programming. Mathematical programming 95 (3), 493–511.

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the Econometric Society, 987–1007.

Escudero, L., 2000. WARSYP: a robust modeling approach for water resources system planning under uncertainty. Annals of Operations Research 95 (1), 313–339.

Faber, B. A., Stedinger, J. R., 2001. Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. Journal of Hydrology 249 (1-4), 113–133.

Ficchì, A., Raso, L., Allemande, M., Dorchies, D., Malaterre, P., Pianosi, F., Van Overloop, P., 2013a. A centralized real-time controller for the reservoirś management on the seine river using ensemble weather forecasting. In: EGU General Assembly Conference Abstracts. Vol. 15. p. 10420.

Ficchì, A., Raso, L., Allemande, M., Dorchies, D., Malaterre, P., Pianosi, F., Van Overloop, P., Thirel, G., Ramos, H., July 2013b. Reservoir's management on the seine river using a centralized real-time controller and ensemble weather forecasting. In: IAHS-IAPSO-IASPEI joint assembly, Gothenburg, Sweden, Abstracts 2843596.

Fishburn, P., 1982. The foundations of expected utility. Theory & Decision Library.

Forsgren, A., Gill, P. E., Wright, M. H., 2002. Interior methods for nonlinear optimization. SIAM review 44 (4), 525–597.

Galelli, S., Gandolfi, C., Soncini-Sessa, R., Agostani, D., 2010. Building a metamodel of an irrigation district distributed-parameter model. Agricultural water management 97 (2), 187–200.

Gneiting, T., Raftery, A. E., 2005. Weather forecasting with ensemble methods. Science 310 (5746), 248.

Götzinger, J., Bardossy, A., 2008. Generic error model for calibration and uncertainty estimation of hydrological models. Water Resources Research 44 (12), W00B07.

Gourbesville, P., 2008. Integrated river basin management, ICT and DSS: challenges and needs. Physics and Chemistry of the Earth, Parts A/B/C 33 (5), 312–321.

Growe-Kuska, N., Heitsch, H., Romisch, W., 2003. Scenario reduction and scenario tree construction for power management problems. In: Power Tech Conference Proceedings, 2003 IEEE Bologna. Vol. 3. p. 7.

Gupta, H., Wagener, T., Liu, Y., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrological Processes 22, 3802–3813.

Haiden, T., Kann, A., Pistotnik, G., Stadlbacher, K., Wittmann, C., 2010. Integrated nowcasting through comprehensive analysis (INCA) system description. ZAMG report.

Harou, J., Pulido-Velazquez, M., Rosenberg, D., Medellín-Azuara, J., Lund, J., Howitt, R., 2009. Hydro-economic models: Concepts, design, applications, and future prospects. Journal of Hydrology 375 (3), 627–643.

Hashimoto, T., Loucks, D., Stedinger, J., 1982. Reliability, resiliency and vulnerability criteria for water resources systems. Water resources research 18 (1), 14–20.

Heitsch, H., Römisch, W., 2009. Scenario tree modeling for multistage stochastic programs. Mathematical Programming 118 (2), 371–406.

Henson, M., 1998. Nonlinear model predictive control: current status and future directions. Computers and Chemical Engineering 23 (2), 187–202.

Janssen, J. A. E. B., Hoekstra, A. Y., De Kok, J. L., Schielen, R. M. J., 2009. Delineating the model-stakeholder gap: framing perceptions to analyse the information requirement in river management. Water resources management 23 (7), 1423–1445.

Jaynes, E. T., Bretthorst, G. L., 2003. Probability theory: the logic of science. Cambridge Univ Pr.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press.

Kavetski, D., Kuczera, G., Franks, S. W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. theory. Water Resources Research 42 (3), W03407.

Koutsoyiannis, D., 2009. A random walk on water. Hydrology and Earth System Sciences Discussions 6, 6611–6658.

Kuczera, G., Parent, E., 1998. Monte carlo assessment of parameter uncertainty in conceptual catchment models: the metropolis algorithm. Journal of Hydrology 211 (1), 69–85.

Labadie, J. W., 2004. Optimal operation of multireservoir systems: State-of-the-art review. Journal of Water Resources Planning and Management 130 (2), 93–111.

Laio, F., Tamea, S., 2007. Verification tools for probabilistic forecast of continuous hydrologycal variables. Hydrology and Earth System Sciences 11, 1267–1277.

Latorre, J. M., Cerisola, S., Ramos, A., 2007. Clustering algorithms for scenario tree generation: Application to natural hydro inflows. European Journal of Operational Research 181 (3), 1339–1353.

Leutbecher, M., Palmer, T. N., 2008. Ensemble forecasting. Journal of Computational Physics 227 (7), 3515–3539.

Liu, Y., Gupta, H., Springer, E., Wagener, T., 2008. Linking science with environ-

mental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. Environmental Modelling & Software 23 (7), 846–858.

Löfberg, J., 2003. Minimax approaches to robust model predictive control. Univ.

Loucks, D. P., 1992. Water resource systems models: their role in planning. Journal of Water Resources Planning and Management 118 (3), 214–223.

Loucks, D. P., Kindler, J., Fedra, K., 1985. Interactive water resources modeling and model use: an overview. Water Resources Research 21 (2), 95–102.

Lukasse, L., van der Voort, A. J., de Kramer-Cuppen, J., 2006. Optimal climate control to anticipate future weather and energy tariffs. In: 4th IFAC Workshop on Control Applications in Post-harvest and Processing Technology, Bornimer Agrartechnische Berichte. Vol. 55. pp. 109–122.

Maciejowski, J., 2002. Predictive control: with constraints. Pearson education.

Maestre, J., Raso, L., van Overloop, P., De Schutter, B., 2012a. Distributed tree-based model predictive control on a drainage water system. Journal of Hydroinformatics.

Maestre, J., Raso, L., van Overloop, P., de Schutter, B., 2012b. Distributed tree-based model predictive control on an open water system. In: American Control Conference (ACC), 2012. IEEE, pp. 1985–1990.

Magni, L., Sepulchre, R., 1997. Stability margins of nonlinear receding-horizon control via inverse optimality. Systems & Control Letters 32 (4), 241–245.

Mayne, D. Q., Rawlings, J. B., Rao, C. V., Scokaert, P. O. M., 2000. Constrained model predictive control: Stability and optimality. AUTOMATICA-OXFORD- 36, 789–814.

Mitchell, B., 2004. Integrated water resources management: A reassessment by asit k. biswas. Water International 29 (3), 398–399.

Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resour. Res. 40.

Morari, M., et al., 1999. Model predictive control: past, present and future* 1. Computers & Chemical Engineering 23 (4-5), 667–682.

Muñoz de la Peña, D., 2005. Control predictivo para sistemas con incertidumbre. Universidad de Sevilla.

Muñoz de la Peña, D., Bemporad, A., Alamo, T., 2005. Stochastic programming applied to model predictive control. In: Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on. pp. 1361–1366.

Overloop, P. J., Negenborn, R. R., Schutter, B. D., Giesen, N. C., 2010. Predic-

tive control for national water flow optimization in the netherlands. Intelligent Infrastructures, 439–461.

Pallottino, S., Sechi, G., Zuddas, P., 2005. A DSS for water resources management under uncertainty by scenario analysis. Environmental Modelling & Software 20 (8), 1031–1042.

Pereira, M., Pinto, L. M., 1991. Multi-stage stochastic optimization applied to energy planning. Mathematical Programming 52 (1), 359–375.

Philbrick, C. R., Kitanidis, P. K., 1999. Limitations of deterministic optimization applied to reservoir operations. Journal of water resources planning and management 125 (3), 135–142.

Pianosi, F., Raso, L., 2008. Flood forecasting for heteroscedastic streamflow processes. In: Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea.

Pianosi, F., Raso, L., 2012. Dynamic modelling of predictive uncertainty by regression on absolute errors. Water Resources Research 48 (3).

Pianosi, F., Ravazzani, G., 2010. Assessing rainfall-runoff models for the management of lake verbano. Hydrological Processes 24, 3195–3205.

Pianosi, F., Soncini-Sessa, R., 2009. Real-time management of a multipurpose water reservoir with a heteroscedastic inflow model. Water Resources Research 45 (10), W10430.

Popper, K., 1959. The propensity interpretation of probability. The British journal for the philosophy of science 10 (37), 25–42.

Qin, S., Badgwell, T., 2003. A survey of industrial model predictive control technology. Control engineering practice 11 (7), 733–764.

Raftery, A. E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Using bayesian model averaging to calibrate forecast ensembles. Monthly Weather Review 133 (5), 1155–1174.

Raso, L., 2013. Socio-hydrology: conceptualising human-flood interactions. Hydrology and Earth System Sciences Discussions 10 (4), 2143–2146.

Raso, L., Schwanenberg, D., Schruff, T., Failache Gallo, N., 2012a. Tree based model predictive control for optimizing hydropower production and floods. In: 10th International Conference on Hydroinformatics, HIC 2012, Hamburg, Germany.

Raso, L., Schwanenberg, D., van Overloop, P., 2010. Decision under uncertainty, use of a flexible model predictive control on a drainage canal system. 9th International Conference on Hydroinformatics, 7th-10th September, Tianjin, China, HICA00114–00176.

Raso, L., Schwanenberg, D., van Overloop, P., van de Giesen, N., 2013a. Tree-based

model predictive control for short-term management of reservoir systems using ensemble forecasts. Submitted to Advances in Water Resources.

Raso, L., Schwanenberg, D., van Overloop, P. J., 2011. Tree-based model predictive control for operational management of open-water systems. In: EGU Geophysical Research Abstracts. Vol. 13. p. 8338.

Raso, L., van der Giesen, N., Schwanenberg, D., 2012b. Tree structure generation from ensemble forecasts for short-term reservoir optimization. In: AGU General Assembly 2012, San Francisco, 3rd-7th December 2012, Control ID 1487220. Vol. 15.

Raso, L., van Overloop, P., 2011. Model predictive control for operational water management: Optimal, integrated, anticipatory and robust to uncertainty. In: EGU General Assembly Conference Abstracts. Vol. 13. p. 11675.

Raso, L., van Overloop, P., Schwanenberg, D., 2012c. Optimal control of water systems under uncertainty using stochastic approximation algorithm. In: 10th International Conference on Hydroinformatics, HIC 2012, Hamburg, Germany.

Raso, L., Werner, M., Weijs, S., 2012d. Deciding under uncertainty vs. reducing uncertainty by observations: the optimal observation problem. In: AGU General Assembly 2012, San Francisco, 3rd-7th December 2012, Control ID 1495690. Vol. 15.

Raso, L., Werner, M., Weijs, S., 2013b. The optimal observation problem applied to a rating curve estimation including the "cost-to-wait". In: Geophysical Research Abstracts. Vol. 15. p. 5176.

Reed, P. M., Kasprzyk, J., 2009. Water resources management: The myth, the wicked, and the future. Journal of Water Resources Planning and Management 135, 411.

Reichert, P., Mieleitner, J., 2009. Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters. Water Resources Research 45 (10), W10402.

Reznicek, K., Cheng, T. C. E., 1991. Stochastic modelling of reservoir operations. European journal of operational research 50 (3), 235–248.

Romanowicz, R. J., Young, P., Beven, K. J., 2006. Data assimilation and adaptive forecasting of water levels in the river severn catchment, united kingdom. Water Resour. Res. 42.

Rotting, T., Gjelsvik, A., 1992. Stochastic dual dynamic programming for seasonal scheduling in the norwegian power system. Power Systems, IEEE Transactions on 7 (1), 273–279.

Sahinidis, N. V., 2004. Optimization under uncertainty: state-of-the-art and opportunities. Computers & Chemical Engineering 28 (6-7), 971–983.

Savenije, H., 2002. Why water is not an ordinary economic good, or why the girl is special. Physics and Chemistry of the Earth, Parts A/B/C 27 (11), 741–744.

Schaake, J., Larson, L., 1998. Ensemble streamflow prediction (esp): Progress and research needs. In: Special Symposium on Hydrology, Am. Meteorol., Soc., Boston, Mass, J19-J24.

Schaefli, B., Hingray, B., Niggli, M., Musy, A., 2005. A conceptual glacio-hydrological model for high mountainous catchments. Hydrology and Earth System Sciences 9 (1/2), 95–109.

Schaefli, B., Talamba, D. B., Musy, A., 2007. Quantifying hydrological modeling errors through a mixture of normal distributions. Journal of Hydrology 332 (3-4), 303 – 315.

Schoups, G., Vrugt, J., 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-gaussian errors. Water Resources Research 46 (10).

Schwanenberg, D., Galelli, S., Sheret, I., Alliance, S., 2011a. Nonlinear model predictive control for heterogeneous process models in water resources. In: World Congress. Vol. 18. pp. 10565–10570.

Schwanenberg, D., van Breukelen, A., Hummel, S., 2011b. Data assimilation for supporting optimum control in large-scale river networks. In: Networking, Sensing and Control (ICNSC), 2011 IEEE International Conference on. pp. 98–103.

Simon, H., 1990. Reason in human affairs. Stanford University Press.

Solomatine, D. P., Shrestha, D. L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resour. Res. 45 (12).

Soncini-Sessa, R., Castelletti, A., Weber, E., 2007. Integrated and participatory water resources management. Elsevier Science.

Soncini-Sessa, R., Villa, L., Weber, E., Rizzoli, A., 1999. TwoLe: a software tool for planning and management of water reservoir networks. Hydrological sciences journal 44 (4), 619–631.

Sorooshian, S., Dracup, J. A., 1980. Stochastic parameter estimation procedures for hydrologie rainfall-runoff models: Correlated and heteroscedastic error cases. Water Resources Research 16 (2), 430–442.

Stanzel, P., Kahl, B., Haberl, U., Herrnegger, M., Nachtnebel, H., 2008. Continuous hydrological modelling in the context of real time flood forecasting in alpine danube tributary catchments. In: IOP Conference Series: Earth and Environmental Science. Vol. 4. p. 012005.

Stedinger, J. R., Sule, B. F., Loucks, D. P., 1984. Stochastic dynamic programming models for reservoir operation optimization. Water Resources Research 20 (11), 1499–1505.

Stive, P. M., 2011. Performance assessment of tree-based model predictive control. Master thesis, Delft University of Technology, Delft, The Netherlands.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S., Srikanthan, S., et al., 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using bayesian total error analysis. Water Resources Research 45, 22.

Tilmant, A., Pinte, D., Goor, Q., 2008. Assessing marginal water values in multipurpose multireservoir systems via stochastic programming. Water Resources Research 44 (12), W12431.

Trezos, T., Yeh, W. W.-G., 1987. Use of stochastic dynamic programming for reservoir management. Water Resources Research 23 (6), 983–996.

Tsoukias, A., 2008. From decision theory to decision aiding methodology. European Journal of Operational Research 187 (1), 138–161.

Turgeon, A., 1980. Optimal operation of multireservoir power systems with stochastic inflows. Water Resources Research 16 (2), 275–283.

Van de Water, H., Willems, J., 1981. The certainty equivalence property in stochastic control theory. Automatic Control, IEEE Transactions on 26 (5), 1080–1087.

van den Boom, T., Backx, T., 2003. Model predictive control. Lecture Notes for the Dutch Institute of Systems and Control, Winterterm 2004.

van Overloop, P., 2006. Model predictive control on open water systems. Ph.D. thesis.

van Overloop, P. J., Weijs, S., Dijkstra, S., 2008. Multiple model predictive control on a drainage canal system. Control Engineering Practice 16 (5), 531–540.

Šutienė, K., Makackas, D., Pranevičius, H., 2010. Multistage k-means clustering for scenario tree construction. Informatica 21 (1), 123–138.

Washington, W., Buja, L., Craig, A., 2009. The computational future for climate and earth system models: on the path to petaflop and beyond. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 367 (1890), 833–846.

Watkins, D. W., Wei, W., 2004. Scenario-tree framework for streamflow forecasting and water resources decision support. In: Proc., World Water & Environmental Resources Congress. ASCE Salt Lake City.

Watkins Jr, D., McKinney, D., Lasdon, L., Nielsen, S., Martin, Q., 2000. A scenario-based stochastic programming model for water supplies from the highland lakes. International Transactions in Operational Research 7 (3), 211–230.

Watkins Jr, D. W., McKinney, D. C., 1997. Finding robust solutions to water resources problems. Journal of Water Resources Planning and Management 123 (1), 49–58.

Weijs, S., van Leeuwen, E., van Overloop, P., van de Giesen, N., 2007. Effect of uncertainties on the real-time operation of a lowland water system in the netherlands. IAHS PUBLICATION 313, 463.

Weijs, S. V., 2011. Information theory for risk-based water system operation. Ph.D. thesis.

Weijs, S. V., Schoups, G., Van de Giesen, N., 2010. Why hydrological predictions should be evaluated using information theory. Hydrol. Earth Syst. Sci 14, 2545–2558.

Xu, M., Van Overloop, P., Van De Giesen, N., 2013. Model reduction in model predictive control of combined water quantity and quality in open channels. Environmental Modelling & Software 42, 72–87.

Xu, M., Van Overloop, P., Van De Giesen, N., Stelling, G., 2010. Real-time control of combined surface water quantity and quality: polder flushing.

Zavala, V. M., Constantinescu, E. M., Krause, T., Anitescu, M., 2009. On-line economic optimization of energy systems using weather forecast information. Journal of Process Control 19 (10), 1725–1736.

Zhao, T., Cai, X., Yang, D., 2011. Effect of streamflow forecast uncertainty on real-time reservoir operation. Advances in Water Resources 34 (4), 495–504.

Acknowledgements

I have the strong personal opinion that the self-made man is nothing but a myth. In my view, success is made of a combination of luck and choices. Luck comes in form of opportunities, then it is up to the individual to take them. For me, this opportunity was the access to education. Being aware that in other times and other part of the world the right to education is not so obvious, my first thank goes to everybody that participate and support social democracy, that grants this right. [2]

Thanks to this opportunity, I obtained a PhD, that is the highest level of education. This long path was not a solitary walk. If alone, I could not have made it. I was in fact supported by many people. I am glad to have here the opportunity to give the credit they deserve, and to thank them.

A good teacher can change your life. Therefore I would like to thank the teachers I met during my life, beginning with the first teachers, my parents. I am grateful to them, for having brought me up with love, teaching me the sense of responsibility, the family values, and the capacity to take my own decisions. Their presence gives me the inner strength to travel around the world, for they are always a safe harbour where I can come back anytime.

I am grateful to the the person that most influenced my education at university, Professor Rodolfo Soncini-Sessa. His vision for the application of optimization tools in the framework of integrated and participatory water resources management has been inspiring, and his work and method are still a reference point for me.

In the last four years in Delft, I had the opportunity to work with many people. I am glad to thank the people that made this time fruitful and enjoyable. Their friendliness gives me the feeling that this is just the beginning of a longer collaboration. The first person that I would like to mention is my promotor, Nick van der Giesen, that I consider close to be the ideal supervisor. Even if the time spent together was not much, he has often transmitted his experience with few words, and he has always

2. The Italian constitution recognizes education as a right, stating that "The capable and deserving, even without financial resources, have the right to attain the highest levels of education." (Art 34).

given me the freedom to pursue my own research. I am also extremely grateful to my daily supervisor, Peter Jules van Overloop. He has been always available for discussions or help. His unconventionality and informality create the nice conditions to have fun and working well at the same time, and his generosity is of great help. I am glad to have shared the room with you in these years. All the people with whom I shared the "control room" get also the credit for these pleasant years: Steven Weijs, Tian, Xu Min, and all the others. I have to spend some more words for Steven, for he helped me in the AXA Post-Doc research grant, that we both won. Your company was really a pleasure. It is nice to meet again at conferences or other events. I am also grateful to Dirk Schwanenberg, and all the people of the Operational Water Management Department at Deltares, for having funded my research and for the fruitful collaboration. Last person from Delft that I mention is the one that I met first, Jair Smit. He introduced me to the people in the department, and for this I am still grateful.

Even if the Water Resources Management group at TU Delft is a wonderful place to work, life would have been unbearable without friends. I shared relatively many experiences with relatively few friends, and I thank them. Therefore, thanks to all my friends in Holland, in particular to my two paranymphs: Lionello Lopes, who, even without high level of education, has a lot to teach, and Marika Mingardo, who, even if having studied so much, has a lot to learn.

There are still more people that I would like to thank, but it would take really too long. I have to focus on two remarkable people only. Thanks to Silvio Berlusconi, for having gave me the opportunity to talk about Italian politics with all the foreigners. Thanks to Rita Verdonk, for having obliged non-Europeans to follow the "inburgeringcursus", the course of Dutch language and culture, that I also followed. This was a great opportunity for me to learn the Dutch language and to get to know that immigration in The Netherlands is much more multi-colored and rich than who organized the course thinks.

The last teacher that I want to thank is the one that I hope to have next to me for the rest of my life, for I still have so many things to learn. Thanks for all the discussions, the love, the patience, the forgiveness, and the support that you give me, every day, especially the hardest ones.

About the author

Luciano Raso (LR) is a water resources management researcher.

LR was born in Reggio Calabria, on October the 1st, 1982. He attended there the *Liceo Scientifico Leonardo da Vinci*, where he met professor Maria D'Africa, to whom this thesis is dedicated. LR received his bachelor degree in environmental engineering at *Politecnico di Milano*, and the double master degree *cum laude* at *Politecnico di Milano* e *Politecnico di Torino*, still in environmental engineering. In the same period he spent one year as exchange student at the *Denmark Technical University (DTU)* of Lyngby, Denmark, and attended successfully the *Alta Scuola Politecnica*, getting a *Diploma with Merit*. Later, he won a scholarship for a master in Environmental Economics and Management at the *Bocconi University*, in Milan.

His main work experience outside university took place in Rome, first at *ENEL*, the largest Italian energy corporation, as stagier, and then at *Eurogroup Consulting*, a consulting company, as business analyst. In March 2009, LR returned to academia, joining the Water Resources Management group at the Delft University of Technology, chaired by Prof. Nick van de Giesen.

LR's core research interest is **decision making under uncertainty**, specifically for water systems. For this reason, LR is interested in Systems Dynamics and Control, Decision Support Systems (DSS), Environmental Economics and Policies, Hydrological Modelling, Stochastic Hydrology, Risk Analysis and Stochastic Programming. LR likes the interaction and collaboration with experts outside his field, for he believes that innovation lays in the intersection among disciplines.

During his PhD, LR was guest lecture at *Cornell University, ETH Zurich, Denmark Technical University, Universitat Politecnica de Valencia, Politecnico di Milano, Hogeschool Rotterdam,* and *Haagse Hogeschool*, assistant in the course "Probability & Statistics" at *TU Delft*, and supervisor of different master students. He attended and presented at many international conferences and workshops, reviewed different

papers for *Water Resources Research, Environmental Modelling & Software, Physics and Chemistry of the Earth, Hydrology Research, Journal of Hydroinformatics*, and *IFAC*.

LR won different scholarships and awards, among which, recently, the NCR days best oral presentation, and the AXA Post-Doc research fellowship (120k€). Thanks to the AXA grant, from October 2013, LR will be able to continue pursuing his research interests in the Water Management Group of *IRSTEA*, in Montpellier, France.

Luciano is enthusiastic and communicative, entrepreneurial, responsible, organized, willing to share and to work in a group. He speaks English, Dutch, and Italian. He is learning French and Spanish.

Publications

Peer-reviewed Journal Articles

Raso, L., Schwanenberg, D., van Overloop, P., van de Giesen, N., 2013a. Tree-based model predictive control for short-term management of reservoir systems using ensemble forecasts. Submitted to Advances in Water Resources.

Raso, L., van de Giesen, N., Stive, P., Schwanenberg, D., van Overloop, P., 2013a. Tree structure generation from ensemble forecasts for real time control. Hydrological Processes 27 (1), 75–82.

Pianosi, F., Raso, L., 2012. Dynamic modelling of predictive uncertainty by regression on absolute errors. Water Resources Research 48 (3).

Maestre, J., Raso, L., van Overloop, P., De Schutter, B., 2012a. Distributed tree-based model predictive control on a drainage water system. Journal of Hydroinformatics.

Conference Articles

Raso, L., Schwanenberg, D., Schruff, T., Failache Gallo, N., 2012. Tree based model predictive control for optimizing hydropower production and floods. In: 10th International Conference on Hydroinformatics, HIC 2012, Hamburg, Germany.

Raso, L., van Overloop, P., Schwanenberg, D., 2012c. Optimal control of water systems under uncertainty using stochastic approximation algorithm. In: 10th International Conference on Hydroinformatics, HIC 2012, Hamburg, Germany.

Maestre, J., Raso, L., van Overloop, P., de Schutter, B., 2012b. Distributed tree-based model predictive control on an open water system. In: American Control Conference (ACC), 2012. IEEE, pp. 1985–1990.

Raso, L., Schwanenberg, D., van Overloop, P., 2010. Decision under uncertainty, use of a flexible model predictive control on a drainage canal system. 9th International Conference on Hydroinformatics, 7th-10th September, Tianjin, China, HICA00114–00176.

Pianosi, F., Raso, L., 2008. Flood forecasting for heteroscedastic streamflow processes. In: Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea.

Conference Abstracts

Raso, L., Werner, M., Weijs, S., 2013. The optimal observation problem applied to a rating curve estimation including the "cost-to-wait". In: Geophysical Research Abstracts. Vol. 15. p. 5176.

Ficchì, A., Raso, L., Allemande, M., Dorchies, D., Malaterre, P., Pianosi, F., Van Overloop, P., 2013. A centralized real-time controller for the reservoirś management on the seine river using ensemble weather forecasting. In: EGU General Assembly Conference Abstracts. Vol. 15. p. 10420.

Ficchì, A., Raso, L., Allemande, M., Dorchies, D., Malaterre, P., Pianosi, F., Van Overloop, P., Thirel, G., Ramos, H., July 2013b. Reservoir's management on the seine river using a centralized real-time controller and ensemble weather forecasting. In: IAHS-IAPSO-IASPEI joint assembly, Gothenburg, Sweden, Abstracts 2843596.

Raso, L., van der Giesen, N., Schwanenberg, D., 2012b. Tree structure generation from ensemble forecasts for short-term reservoir optimization. In: AGU General Assembly 2012, San Francisco, 3rd-7th December 2012, Control ID 1487220. Vol. 15.

Raso, L., Werner, M., Weijs, S., 2012d. Deciding under uncertainty vs. reducing uncertainty by observations: the optimal observation problem. In: AGU General Assembly 2012, San Francisco, 3rd-7th December 2012, Control ID 1495690. Vol. 15.

Raso, L., Schwanenberg, D., van Overloop, P. J., 2011. Tree-based model predictive control for operational management of open-water systems. In: EGU Geophysical Research Abstracts. Vol. 13. p. 8338.

Raso, L., van Overloop, P., 2011. Model predictive control for operational water management: Optimal, integrated, anticipatory and robust to uncertainty. In: EGU General Assembly Conference Abstracts. Vol. 13. p. 11675.

Raso, L., Schwanenberg, D., van der Giesen, N., van Overloop, P. J., 2010. Tree-scenario based model predictive control. In: Geophysical Research Abstracts. Vol. 12. p. 3178.

Interactive Comments

Raso, L., 2013. Interactive Comment to: Socio-hydrology: conceptualising human-flood interactions, by Di Baldassarre et al., Hydrology and Earth System Sciences Discussions 10 (4), 2143–2146.