POLITECNICO DI MILANO
Scuola di Ingegneria Civile, Ambientale e Territoriale
Corso di Laurea in Ingegneria per l'Ambiente ed il Territorio

Optimal and centralized reservoir management for drought and flood protection on the upper Seine-Aube river system

Supervisor:
Prof. Ing. Rodolfo SONCINI-SESSA

Internship tutor:
Dott. Ing. Luciano RASO

Master Courses Thesis by:
Mattia CHIAVICO
Mat. 801766

AA. 2014-2015

Contents

Index of figures

Index of tables

Abstract

Seine river region is an extremely important logistic and economic junction for France and Europe. The hydraulic protection of the river relies on four controlled reservoirs, managed by basin authority *EPTB Seine-Grands Lacs*. Presently, reservoirs operation is not centrally coordinated, and release rules are based on empirical rule curves. In this study, we analyze how a centralized release policy can face flood and drought risks, optimizing water system efficiency.

The optimal and centralized decisional problem is solved by Stochastic Dual Dynamic Programming (SDDP) method, which provides the optimal policy minimizing the totality of environmental impact indicators. SDDP allows us to include into the system: 1) the hydrological discharge, specifically a stochastic semi-distributed auto-regressive model, 2) the hydraulic transfer model, represented by a linear lag and route model, 3) the diversions and 4) the water stocks in reservoirs. The novelty of this study lies on the combination of reservoir and hydraulic models in SDDP for flood and drought protection problems.

The study case covers the Seine basin until the confluence with Aube River: this system includes two reservoirs, the city of Troyes, and the Nuclear power plant of Nogent-Sur-Seine. The conflict between the interests of flood protection, drought protection, water use, life on lakes and ecology leads to analyze the environmental system in a Multi-Objective perspective.

Keywords: Water management, Stochastic Dual Dynamic Programming, flood and drought risk, statistical hydrology, optimal control.

Estratto

Il bacino della Senna costituisce un nodo logistico ed economico estremamente importante per la Francia e l'Europa. La protezione idraulica del fiume è affidata a quattro serbatoi artificiali, gestiti dall'autorità di bacino *EPTB Seine Grands Lacs*. Ad oggi, le operazioni nei serbatoi non sono controllate centralmente, e le regole di rilascio sono determinate sulla base di curve di riempimento empiriche. In questo studio si analizza come una politica di rilascio centralizzata possa far fronte al rischio di piene e siccità, ottimizzando l'efficienza del sistema.

La strategia risolutiva adottata per risolvere il problema decisionale ottimo e centralizzato è data dalla *programmazione dinamica stocastica duale* (SDDP), che permette di ottenere una politica che minimizzi l'insieme degli indicatori di impatto ambientale. La SDDP permette di descrivere il sistema in termini di: 1) afflussi idrologici, dati da un modello auto-regressivo semi distribuito, 2) trasferimento idraulico, rappresentato da un modello lineare di ritardo ed attenuazione (*lag and route*), 3) diversioni e 4) stoccaggio nei serbatoi. La novità di questo studio è legata alla combinazione fra gestione dei serbatoi e trasferimento idraulico in SDDP per problemi di piene e siccità.

Il caso di studio copre il bacino della Senna sino alla confluenza con l'Aube: questo sistema include due serbatoi controllati, la città di Troyes e la centrale Nucleare di Nogent-sur-Seine. Il conflitto tra gli interessi di protezione da piene , di uso dell'acqua, di ecologia fluviale e di qualità della vita lacustre, inserisce il problema in una prospettiva di analisi a Molti Obiettivi.

Preface

This Master Course Thesis resumes the work I carried out during my internship at IRSTEA Montpellier, UMR G-EAU. I feel particularly fortunate to have had this opportunity: the world of research is fascinating for its richness in culture and novelties; moreover, the diversity on water management competences that I found within people at IRSTEA, belonging to mathematics, political sciences, sociology, geography, economics, hydraulics and much more, had been really inspiring for me and for my knowledge on this theme. By the way, I may suggest to the one who adopts integration principle for his project to integrate himself first by immersing in other disciplines: sometimes it is not as simple as it might be.

I am really grateful to all people who made this experience possible: thanks to my internship tutor, Luciano Raso, for his inspiring teachings, for his optimism, and also for his hospitality and the bicycle; thanks to my family and Charlotte, who always gave me support for my decisions. Above all, thanks for always being encouraging in difficult moments.

I owe a big thank to all people who collaborated for this project. In particular, thanks to: David Dorchies, for providing essential data from ClimAware project and for his continuous support; Pierre-Olivier Malaterre, Jean-Claude Bader and Jean-Cristophe Pouget for their close cooperation; Andrea Ficchì for providing information on the project and for his suggestions on the world of scientific research.

I would like to thank also all friends researchers and students of G-Eau for the pleasant time spent together during my internship.

List of abbreviations

ACF: Auto Correlation Function;

AR(*p*): Auto-Regressive model of order *p*;

CAR(*p*) : Conditioned Auto-Regressive model of order *p*;

cdf: Cumulate Distribution Function;

G-EAU : fr. *Gestion de l'Eau, Acteurs, Usages* ;

EPTB : fr. *Etabilissement Public Territorial de Bassin* ;

IRSTEA: fr*. Institut de Recherche en Science et Technologies pour l'Environnement et l'Agriculture* ;

IWRM: Integrate Water Resources Management;

LLR: Linear Lag and Route hydraulic model;

PAR(*p*) : Periodic Auto-Regressive model of order *p*;

PIP: Participatory and Integrated Planning procedure;

R^2: *r-square*, coefficient of determination for linear regressions;

RC: Rule Curve;

RRV: Reliability, Resilience and Vulnerability estimators;

SDDP : Stochastic Dual Dynamic Programming ;

SDP : Stochastic Dynamic Programming ;

SGL: Seine Grands Lacs;

SF: System Failure;

UMR: fr. *Unité Mixte de Recherche*.

List of symbols

$\circ$: Element-wise multiplication (Hadamard product) for matrices;

$\oslash$: Element-wise division for matrices;

$\times$: Matrix product;

diag(M): Returns a squared matrix with the diagonal elements of M;

vec(M): Returns a column vector with the diagonal elements of M;

sum(M): Returns a column vector with the sum of rows of M;

I: Identity matrix;

Introduction

Project goal definition

The present work deals with optimal control theory applied to water resources management on a real study-case. As a matter of facts, control theory represents an interesting manner to study natural resources, because it is directly linked to the impacts (*i.e.* the costs) generated by their exploitation. Thus, it can be seen as a transversal study between the resources, their uses and their relationship with the environment. By analyzing this relationship, and measuring it, we can derive an optimal management policy based on the composition of all the particular interests related to the totality of the system. This integration paradigm is adopted in this study in order to respond to the limits of empirical and de-centralized management of water systems. The solving technique adopted here is Stochastic Dual Dynamic Programming: this algorithm provides a management policy centralized and optimized on the base of stakeholders interests. The study case relative to this report is the upper Seine-Aube river system, characterized by floods and drought risk on urban, industrial and agricultural milieus. The novelty of this study lies on the application of SDDP for flood and drought protection, put into action by a combination of reservoir management and hydraulic transfer modeling.

Therefore, the main research questions of this study are the followings:

- How does centralized SDDP optimize the behavior of the system?
- Does SDDP perform better than current management?

Research context

Stochastic Dual Dynamic Programming method has been introduced by (Pereira, 1989) and analyzed in (Shapiro A. , 2010) and (de Matos, 2012) . In the field of natural resources management, the main application of this technique has been the hydrothermal scheduling, sometimes coupled with irrigation (Tilmant, 2007). Our purpose is to analyze SDDP efficiency for flood and drought protection, which means using a combination of stock and hydraulic propagation state variables.

For what concern the study-case, it has also already been analyzed with different purposes. The main reference about control issues in Seine River region is(Dorchies, 2013) in the context of ClimAware project, where the main goal is to evaluate climate change impacts on reservoirs management. In (Ficchì, 2015) the resolution strategy of Model Predictive Control (MPC) and Tree-Based Model Predictive Control (TB-MPC) has been employed for the same project. In our analysis we will focus on a part of Seine river system (the upper Seine-Aube River system) trying to derive an optimal off-line policy for Seine and Aube reservoirs.

Problem solving procedure

Following the Participatory and Integrated Planning procedure, PIP (Soncini-Sessa, 2007), the first phase of the problem identification for water resources planning and management is *Reconnaissance*, which defines the study-case context. Then, once a solution strategy is defined, objectives are defined in terms of *step-costs* (or *indicators*). A model may be required as step cost functions' argument: for the case of floods and droughts protection, the system is characterized by uncertainty, considered as a disturbance, which influences, through an hydrological model, a combination of reservoir management operations and hydraulic transfer. At the end of the modeling phase, the policy is designed and impacts can be analyzed in order to respond to the initial research questions. Figure 1 provides a scheme of the problem solving procedure adopted in this study:

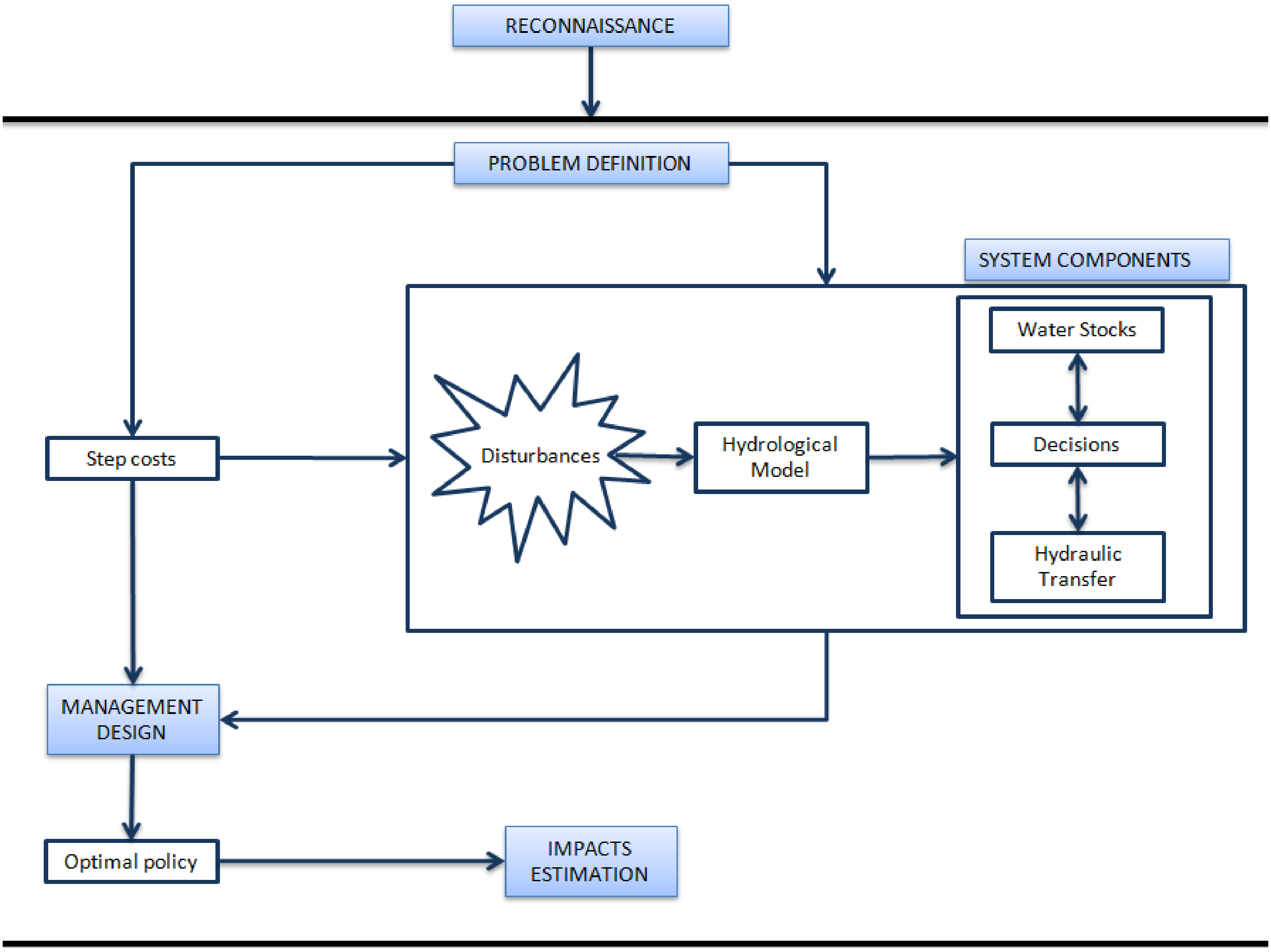


Figure 1: Problem solving procedure adopted in this study.

Organization of this report

Chapter 1. *Reconnaissance*. This is the problem characterization phase, which aims to shed a light on the environmental, sociologic, economic and normative context. A significant part of this chapter, in particular sections *Current management* and *High flow and low flow thresholds*, is taken from (Ficchi', 2013), where more detailed information about the study-case can be found.

Chapter 2. *Solution Strategy*. This chapter provides a formulation of the methodology of our analysis and the presentation of the theory of SDDP algorithm.

Chapter 3. *Components model*. This chapter describes the mathematical models for all components of the system, including hydrological inputs, reservoirs and water courses.

Chapter 4. *Management Design*. This chapter is dedicated to the optimization design and analysis.

Chapter 5. *Impacts estimation*. In this chapter results of optimization are discussed in terms of impacts.

Chapter 6. *Conclusions and perspectives*. In this chapter final conclusions ad answers to initial research questions are shown.

Models and optimal management are implemented using MATLAB software. The optimization tool used for linear programming is CPLEX.

1 Reconnaissance

In this chapter the case study will be presented and contextualized. The Reconnaissance phase defines: the spatial and temporal boundaries of the system, the principal stakeholders involved, the normative and planning context and, finally, the data available and to be collected,. Then, reservoirs and their actual management rules will be presented, highlighting the limitations of the current management strategy faced in the past.

1.1 Spatial boundaries: Upper Seine-Aube river basin

The study case covers Seine River's basin until the confluence with Aube River. The region is characterized by a western European oceanic climate affected by the North Atlantic Current (Köppen climate classification: *Cfb*). The flow regime is characterized by low flows in summer and high flows in winter. For exhaustive description of the physical, meteorological and hydrological characteristics of the Seine river basin, the reader may refer to (Ducharne, 2007).

As showed in Figure 2: Seine river basin [Wikipedia]: in evidence the upper Seine-Aube river basin, Seine and Aube rivers have their source in the same mountain system (Plateau des Langres) and they flow from south-East to north-West until their confluence, where they change direction to South-West. Because of the gentle slope of Seine Valley, the two rivers have numerous meanders, which imply a slow water runoff.

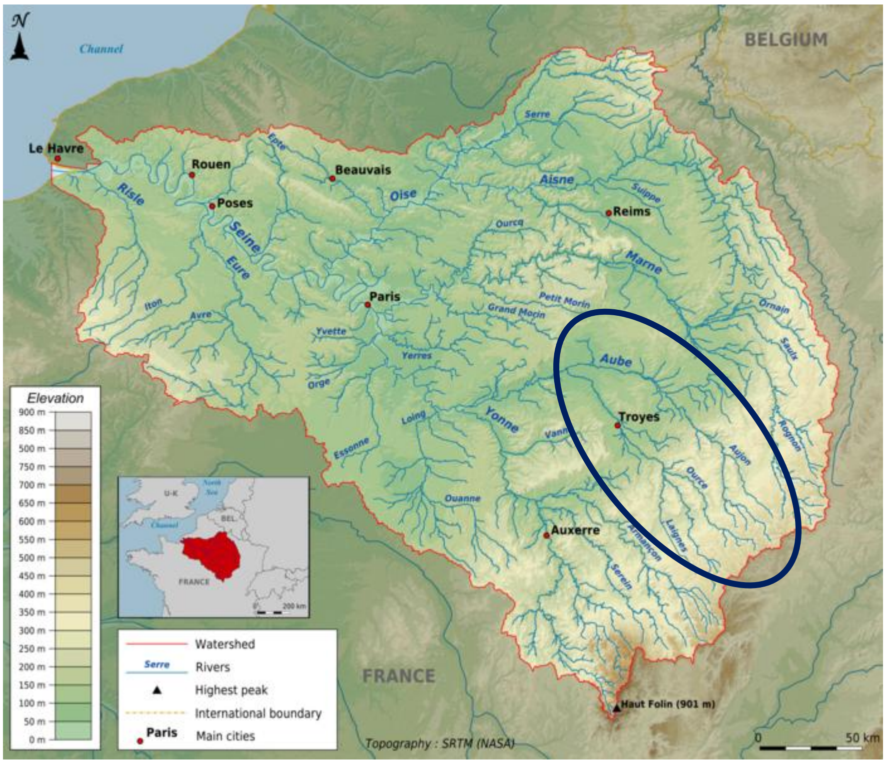


Figure 2: Seine river basin [Wikipedia]: in evidence the upper Seine-Aube river basin

From a hydrological point of view, Seine and Aube upper basins are strictly correlated, because of similar climatic and geo-morphologic conditions. As well, Figure 3 shows that in the northern part of the system both rivers are influenced by lateral hydrological contributions and rainfalls, which contributes to differentiate their flows.

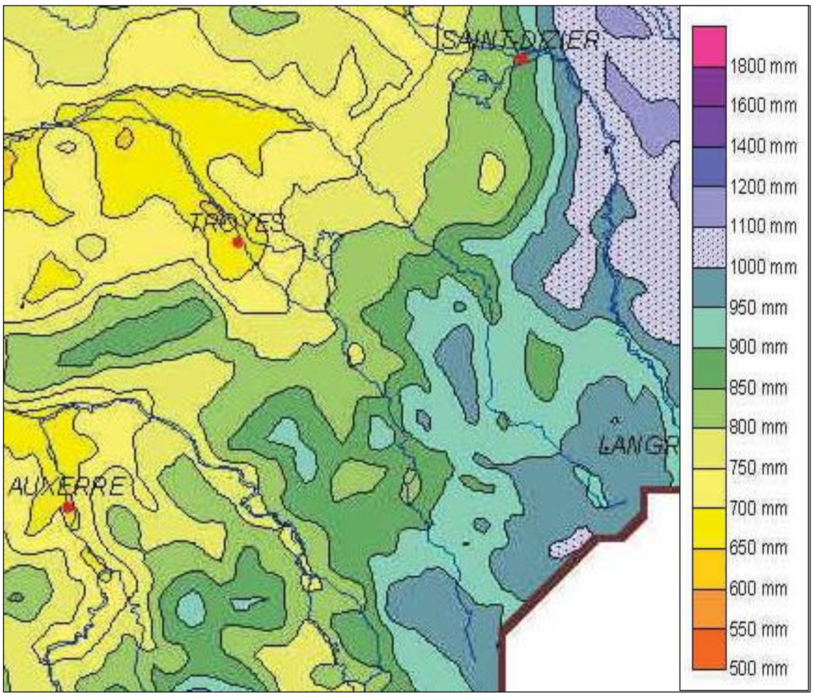


Figure 3: Normal annual rainfall on Seine-Aube upper basin

Until 1966, Aube and Seine watercourses were not controlled. Then, on memory of 1910, 1924 and 1955 flood events, Seine Reservoir was built. Aube Reservoir construction dates to 1990, eight years after 1982 flood.

1.2 Stakeholders

Seine river basin represents important socio-economic stakes in France, especially because of Paris urban area. The presence of cities and industries is the cause of high water demand and vulnerability to floods. Thus, a large set of local stakes is involved in this problem:

- Urban areas, i.e. cities of Troyes, Romilly-sur-Seine, Arcis-sur-Aube, etc.;
- Industries, i.e. a nuclear power plant in Nogent-sur-Seine;
- Users and providers of navigation service on the Seine River;
- Hydroelectric production plants (which is a minor goal and will not be considered in this study);
- Water users for irrigation and supply;
- Rivers ecology;
- Anthropic and natural areas by lakes;

Finally, we need to consider also Paris metropolitan region, which is not a local stakeholder, but is directly influenced by decision took in the upper Seine-Aube River system. Its interest can be identified in not having worsening between current and optimized management for both floods and droughts.

All these stakeholders can be grouped in three main sectors, shown in Table 1: Sectors, stakeholders and optimization criteria.

Sectors	Main stakeholders	Operational Criteria
Urban areas, activities and industries by Seine and Aube rivers	Arcis Sur Aube, Troyes, Mery Sur Seine, Nogent Sur Seine (which represents also downstream riparian region until Paris), Nuclear plant.	Minimize flood risk
Navigation, irrigation, drinking and industry water supply, ecology	Aube and Seine watercourse after derivations, Arcis Sur Aube, Troyes, Mery Sur Seine, Nogent Sur Seine (which represents also downstream riparian region until Paris), Nuclear plant.	Minimize drought risk and respect environmental flow

Lakes Life Quality	Seine and Aube Lake tourism and life; water quality	Possibly have full lake.

Table 1: Sectors, stakeholders and optimization criteria

The nuclear power plant protection, for its high relevance in the system, requires a description on what are the risks associated with it. The plant uses Seine river water for its cooling system. Thus, an extreme drought can potentially damage the functioning of the power plant for the switching-off risk caused by the lack of cooling water flow; damages may be caused also by floods for the risk of water intakes obstruction. Nogent-sur-Seine, for its critical challenges such as the protection of both Parisian metropolitan region and nuclear power plant, is the highest-priority station. The high-density urban areas predominantly along Seine River are the second priority; the city of Troyes is the main critical point, with Romilly-sur-Seine and Mery-sur-Seine. Aube River, being less urbanized, has an inferior priority to Seine River and its protection is restricted to Arcis-sur-Aube. The last level of priority is given to lakes life quality, principally because our purpose is to give an answer to flood and drought protection, which perceived cost is greater than having empty lakes.

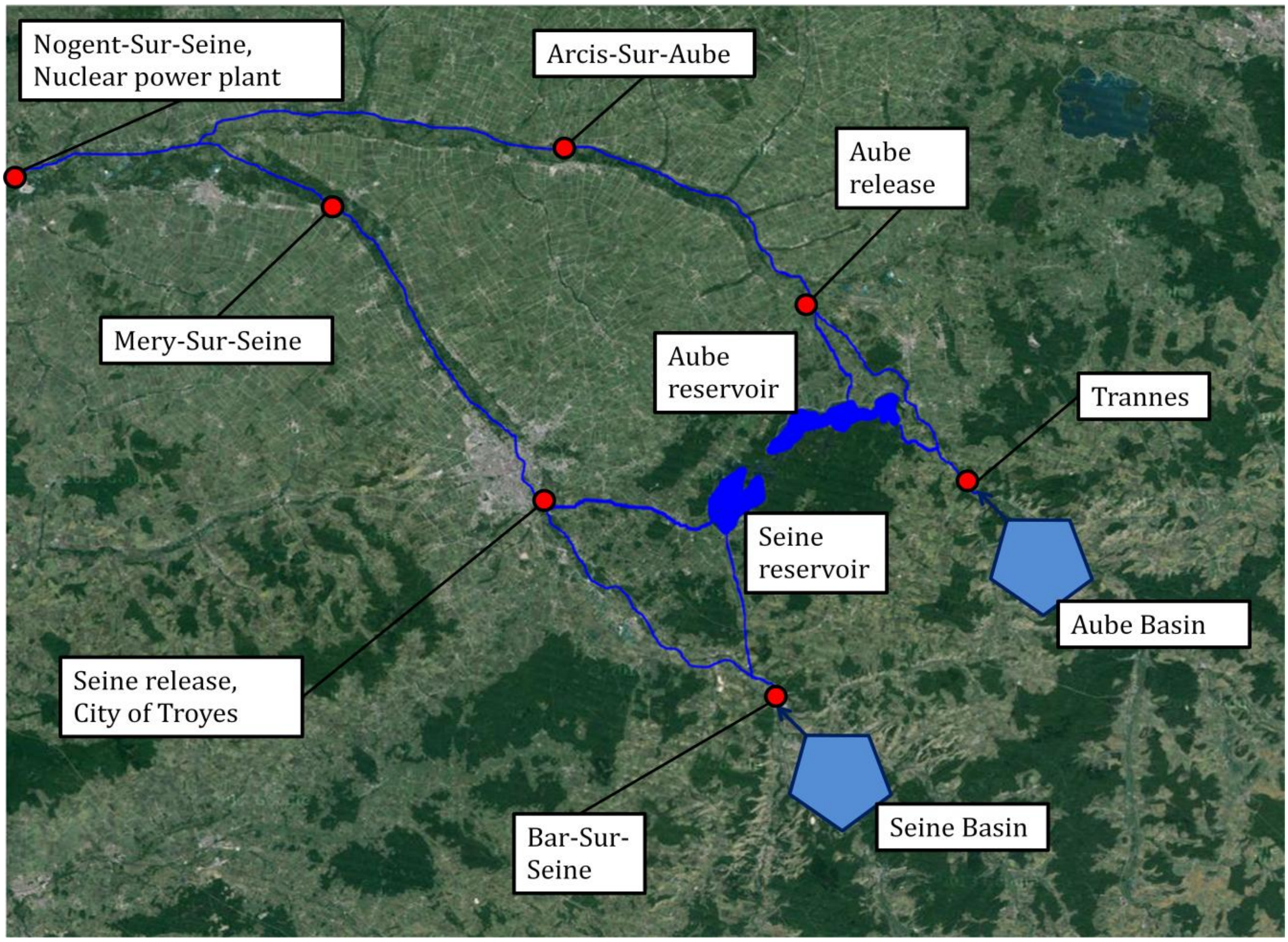


Figure 4: Upper Seine-Aube river system (in evidence Seine and Aube Rivers and Reservoirs) and gauging stations (red dots)

1.3 Data survey and temporal boundaries

Most data are provided by works of (Dorchies, 2013),(Ficchi', 2013) and (Dehay, 2012).

In the context of ClimAware project, Seine Grands Lacs provided data and information on the four reservoirs management:

- Current management rules, expressed as objective RCs (Ch. 1.3);
- Thresholds for environmental and reference flows (paragraph 1.4.2);
- Reservoirs and inlet/outlet channels capacity.

As a result of SGL data analysis for ClimAware project, IRSTEA Montpellier provided:

- A data-base of hydrological insertions at every control and gauging station for the 1961-2009 period. This 48-years temporal domain is used for calibration both of hydrological parameters (paragraph 3.1.2) and cost-to-go function boundaries (paragraph 2.2.2.1).
- A simulation of the current system management during the 1961-1991 period. This 30-years simulation is used to evaluate Impacts Estimation (Ch. 5). This can be defined as the real temporal boundary of the project, even if data calibration is related to a longer time series.
- Parameters of hydraulic transfer model along watercourses, which are used in this study in order to have the same system model than the one used for ClimAware project.

NB: Thresholds values for the flows of the river system at different control stations are used by SGL for performance assessment of the management.

1.4 Current management

1.4.1 Objective Rule Curves

Current reservoirs management is decentralized following some *Objective Rule Curves* (RCs), empirically constructed off-line for each reservoir and respecting some constraints downstream the connections of the lakes. The RCs are designed to store water during the low-flows season (from beginning of July to end of October), and filling the reservoirs from autumn (November) to spring-summer (June). The target volume to reach at the end of the filling season is set in order to have enough water to satisfy a minimal flow threshold in the river for all the releasing season. The slopes of the curve in the filling period are calculated on the basis of the average inflows to the reservoirs. The releasing period can be prolonged if the river flow immediately downstream the dams is below the second low-flow threshold (paragraph 1.4.2) until the 1st of January at most. In this condition of prolonged or delayed droughts the releasing period ends when the flows begin to be higher than this threshold. The rule curves of all the lakes and the maximum and minimum volumes are provided in Annex I: Rule Curves for Seine and Aube Lakes.

The filling trajectory is calculated taking into account the historical floods of the 20th century so that there is enough space (storing capacity) available for these floods control. There were two flood events of the 20th century that current management rules could not be able to control: those of June 1982 and August 1910. Besides 1982 and 1910 events, there are some other critical events (1970, 1988, 1999 and 2001) for which the capacity available in the reservoir is just enough to control the floods.

1.4.2 Environmental and reference flows

In addition to following the rule curves, the current management rules has to respect some legal constraints on the river flows, if possible. The constraints are the following:

- ***Reference flow (or Retention flood level)***: is the threshold which indicates the *occurrence of a flood* on the sub-basin *downstream the dam*. In the current reservoirs management, if downstream the inlet and outlet channels the flow overcomes this value, *the excess flow is stored in the reservoir*. The value of the retention flood level depends on the season, because the flood areas depend on the adjacent territory agricultural uses.
- ***Environmental flow***: is a minimum flow to assure life, movement and reproduction of all species in the river. This is a *legal obligation* (Article L432-5 of the French Environmental code). In case of off-river reservoirs this threshold defines the minimum flow to let in the river downstream of the inlet channel, while for reservoirs directly on the river it indicates the minimum flow to be discharged.

1.5 High flow and low flow thresholds

For assessing the management of the reservoir, the flow is monitored at several strategic stations downstream of the lakes. On each of these gauging stations, flow thresholds define critical low and high flows. For low-flows, thresholds are defined from local decree corresponding to restrictions on the water uses (Arrêté cadre sécheresse , 2012):

- *Vigilance threshold*: at this first threshold any restriction of uses is defined but the river is extremely sensitive to pollutions; this threshold has been used to derive rule curves;
- *Alert threshold*: at which 30% restriction of uses;
- *Reinforced alert threshold*: 50 % restriction of uses;
- *Crisis threshold*: all uses are prohibited except a minimum use for drinking water.

For high flows, the thresholds correspond to three critical levels, respectively*: limit of flooding, flooding in regular area, exceptional flooding*. The monitoring stations and their thresholds are presented in the following table.

Stations	**Drought thresholds [m³/s]**	**Flood thresholds [m³/s]**
Arcis-sur-Aube (09)	Vigilance: 6.3 **Alert: 5** Reinforced Alert: 4 **Crisis: 3.5**	Limit: 110 **Regular: 260** **Exceptional: 400**
Mery-sur-Seine (12)	Vigilance: 7.3 **Alert: 5** Reinforced Alert: 4 **Crisis: 3.5**	Limit: 140 **Regular: 170** **Exceptional: 400**
Nogent-sur-Seine (13)	Vigilance: 25 **Alert: 20** Reinforced Alert: 17 **Crisis: 16**	Limit: 180 **Regular: 280** **Exceptional: 420**

Table 2: Thresholds for monitoring stations; values in bold are the ones used for optimization

Following the IWRM paradigm, a good optimal controller should be chosen on the base of stakeholder's interests, thus on the minimization of *real damages*. Therefore, we chose not to consider a cost for trespassing *vigilance* and *limit of flooding* thresholds: their mere *informative* importance does not imply either direct or indirect damages. From Table 2: Thresholds for monitoring stations we can also observe that *reinforced alert* and *crisis* thresholds are close, therefore, they can be merged into a unique value.

2 Solution Strategy

In this chapter the choice of Stochastic Dual Dynamic Programming (SDDP) as solving algorithm will be justified. Then, the main features of SDDP will be described, also in order to focus on the best suitable model to represent the whole system behavior.

2.1 How to control hydrological risk?

Floods and droughts are rare but dangerous events; moreover, even if they are knowable in some river sections, they can evolve in unpredictable manners along the watercourse, because they are the superposition of numerous stochastic events. Our primary goal is to find a reservoirs management capable of minimize hydrological risk for all the stations. This main objective leads us to solve an optimal control problem, i.e. to minimize an indicator, which should quantify risk. We can instantiate risk in different ways: if we treat a flood or a drought event as a system failure (SF), we should be able to derive some indicators that may respond to our inquietudes about them, as, for instance: *how many times does the system fail?*, *how long can a failure take?, what are the damages associated to failures?,* etc. Whatever the perceived indicators, one can represent them as a set of multiple objectives to be minimized, generically identifiable as *costs*. Let g_t be a chosen cost (also called *step-cost*) associated to an event (flood or drought) for a particular river station at stage t, u_t be a decision variable (for example to withdraw water or not), and x_t be the state of the system (for instance reservoirs volume) at *t.* The expected cost depends on the hydrological flow disturbance, which is given by a climatic scenario, called ε_{t+1}(subscript *t+1* means that the flow begins in *t* and terminates in *t+1*, so it is not known at stage *t*). Therefore, for a pure management problem, the indicator's value over a horizon *h* is expressed as it follows:

$$i = \sum_{t=0}^{h-1} g_t(x_t, u_t, \varepsilon_{t+1}) + g_h(x_h)$$

where g_h, called *penalty*, is the cost associated to the final stage[1] of the problem.

The relation above means that the indicator's value varies in function of decisions u_t. Therefore, at each stage of the sequential decisional problem, a trade-off must be found between the current protection from failures or the ensuring for future expected events, which can be more dangerous (*costly*) than the present ones. The minimization of both current cost and expected future cost is

[1] For periodic systems, as our own, $g_h$can be considered equal to zero if h is sufficiently larger than the period T and to the overall system resilience, which can be estimated as the time for lakes to completely discharge (around 75 days for study-case system).

represented by the recursive Bellman equation, which is the center of Stochastic Dynamic Programming (SDP). In its standard formulation, thus for *risk neutral* operations[2], it is represented as it follows:

(2-1)

$$F_t^*(x_t) = \min_{u_t}\left\{ \underset{\varepsilon_{t+1}}{E} \left[g_t(x_t, u_t, \varepsilon_{t+1}) + F_{t+1}^*(x_{t+1})\right]\right\}$$

where $F_t^*(x_t)$, called cost-to-go function, is the optimal value at stage t. Therefore, the optimal policy is given by the following:

$$m_t^*(x_t) = \arg\min_{u_t}\left\{ \underset{\varepsilon_{t+1}}{E} \left[g_t(x_t, u_t, \varepsilon_{t+1}) + F_{t+1}^*(x_{t+1})\right]\right\}$$

which is a closed-loop control scheme for an off-line reservoirs management.

We can observe that solving problem ((2-1) requires g_t being separable with respect of time. Moreover, state transition is required to be a Markov process, i.e. all information has to be conditional to the further one-stage state. Unfortunately, the solution of Bellman equation is limited to small-scale problems, because the computational effort increases exponentially with the number of states x_t in the system. Various strategies have been developed to cope with the course of dimensionality posed by dynamic programming: some particularly efficient ones have already been analyzed in (Tilmant, 2007).

2.2 Stochastic Dual Dynamic Programming (SDDP) Algorithm

2.2.1 Main features

SDDP Algorithm is an approximation of the Stochastic Dynamic Programming method that provides a closed loop and off-line control scheme. It can be used for large-scale systems at large state-space, but it presents two important limitations:

- The overall system model has to be linear.
- The cost-function must be convex.
- SDDP cost-to-go function $\tilde{F}_t$ is an approximation (more specifically a lower bound) to the true function F_t^*.

[2] The choice between neutrality or adversity to risk for the specific operational research problem is discussed at the end of paragraph 2.2.2.1

About the first limitation, in some cases linearity does not pose problems, in others non-linearities can be represented by piece-wise linear functions.

2.2.2 SDDP Technique

SDDP cost-to-go function is approximated by a convex, piece-wise linear function. The approximated cost-to go is built up by an iterative procedure composed by a backward and forward phase. The backward phase aims at constructing the piecewise linear approximation of the cost-to-go functions at each stage *t* by Bender's cuts. Then, for every cut added, a forward phase produces a series of state trajectories. At the end, L linear cuts at the most are added to the optimization problem as lower boundaries of the cost-to-go functions and the solution converge to the real cost-to-go value. So, every adding-cut iteration has a backward and a forward phase.

2.2.2.1 The backward phase

Vector x_t is defined as the ensemble of all state variables, decisions and step-costs. Beginning from the final stage *t=H*, where x_{t-1} is given by historic data and cost-to-go is zero, a number N of hydrologic input scenarios is launched. At time t, for every *j*-th scenario the following problem is solved:

(2-1)

$$\min_{x_t} \pi^T x_t + F_{t+1}^*$$

$$subject\ to: \begin{cases} A^{eq} x_t = \left(B_t^{eq} x_{t-1} + c_t^{eq}\right) \circ \Xi_t^j \\ x_t \in \mathcal{X} \end{cases}$$

Then, for every scenario results of optimization J_t^j and dual solutions λ_t^j are calculated and stored. After scenarios extraction, two elements can be derived:

1) the cost of the passage at stage t, called J_t^*, can be estimated by averaging the results of optimization between the scenarios:

$$J_t^* = \frac{1}{N}\left[\sum_{j=1}^{N} J_t^j\right]$$

2) because of properties of dual variables, we have the following relationship:

$$\frac{\partial J_t^j}{\partial x_{t-1}} = B_t^{eq^T} \lambda_t^j$$

therefore, the sub-gradient of J_t^* at x_{t-1}, called Θ_t^* is calculated as the average of $B_t^{eq^T} \lambda_t^j$ along the different scenarios:

$$\Theta_t^* = \frac{1}{N}\left[\sum_{j=1}^{N} B_t^{eq^T} \lambda_t^j\right]$$

The information on J_t^* and dJ_t^* permits to construct a linear cutting plane, given by the relation:

$\ell_{t-1} := J_t^* + \Theta_t^{*T} \cdot (x_{t-1} - \overline{x}_{t-1})$

which gives the constraint:

$F_t^* \geq \ell_{t-1}$

consequently, the components of the linear cut are added to the problem, and $\overline{x}_{t-1}$ is updated as an optimal solution of the problem.

N.B.: the choice of *averaging* costs of different scenarios means to use Laplace criterion for a Monte Carlo simulation, which implies operator's *neutrality to risk*. Risk adverse SDDP has already been discussed in (Shapiro A. W., 2013), but it will not be implemented as long as this analysis aims to discern how standard SDDP method works on contexts of risk.

2.2.2.2 The forward phase

To backward phase follows an additive optimization that uses the information on the cuts previously created. It goes from the first to the final stage, and, as for the backward phase, x_{t-1} is given by historic data. Thus, at every stage t the following problem is solved:

$$\min_{x_t} \pi^T x_t + F_{t+1}^*$$

$$\textit{subject to}: \begin{cases} A^{eq} x_t = \left(B_t^{eq} x_{t-1} + c_t^{eq}\right) \circ \Xi_t^j \\ F_{t+1}^* \geq \ell(x_t) \\ x_t \in \mathcal{X} \end{cases}$$

Finally, the total cost of the forward phase, called $\overline{J_1}$ can be compared with J_1^* for convergence analysis (subscript 1 means that total costs are given by cost-to-go value at first stage). If convergence is not attained, another backward phase is launched and another linear cut for every stage is added.

2.2.2.3 Computational effort

SDDP algorithm reduces the problem of the "curse of dimensionality"; anyway, solving time linearly increases with the number of active constraints, as, for instance, state transitions and linear cuts. Therefore, it is necessary sometimes to find a trade-off between the quality of system definition (number of state transitions) and the efficiency of its resolution (number of linear cuts).

2.3 Step-costs definition

2.3.1 Floods, droughts and ecology

Step-cost functions for floods, droughts and ecology are defined dealing with risk assessment and SDDP constraints. Thus, these functions have to be convex, linear and separable in time. A good indicator is the one associated to *vulnerability*(Kjeldsen, 2004), under which costs are directly proportional to thresholds trespassing volume. Since we do not have any information about real costs, parameters of floods and droughts costs are derived as it follows:

	Droughts cost [$/m³]	Floods cost [$/m³]
Regular Events	1	10
Exceptional Events	100	1000

Table 3: Droughts and floods step costs for regular or exceptional events

Thus, step-costs for floods and droughts are defined as it follows:

$$g_t^{i,F} = w_{i,j}^F \left(q_t^i - \bar{q}_t^{F\ i,j}\right)^+$$

$$g_t^{i,D} = w_{i,j}^D \left(\bar{q}_t^{D\ i,j} - q_t^i\right)^+$$

Where:

$g_t^{i,F}$ is step-cost relative to a flood event for station *i* at time *t*;

$w_{i,j}^F$ is flood events cost relative to trespassing threshold *j* at station *i*;

q_t^i is station *i* flow at time *t*.

$\bar{q}_t^{F\ i,j}$ is flood threshold *j* for station *i* at time *t*.

The same symbols have been used for droughts step-costs by replacing letter *D* to letter *F*.

The following graph shows the step-cost function for Nogent-sur-Seine station:

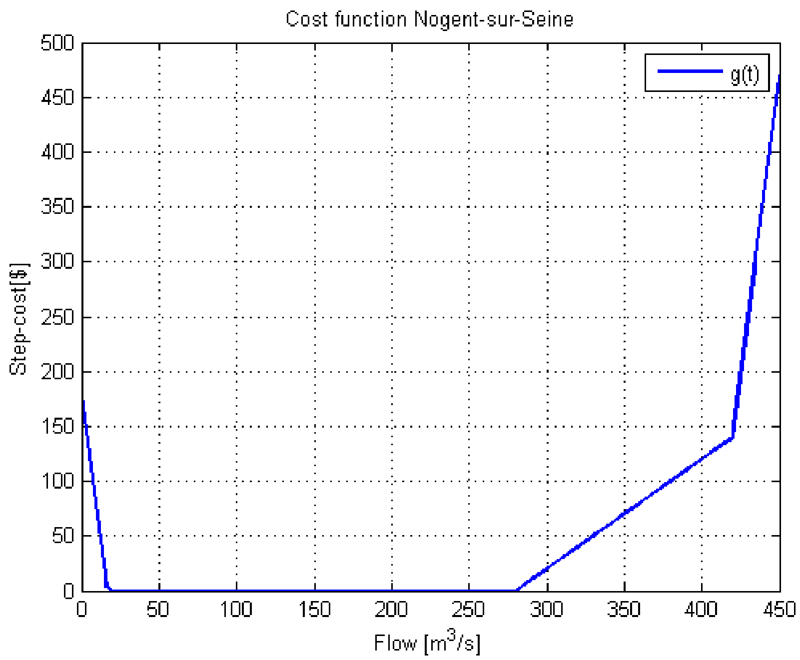


Figure 5 : Step-cost function for Nogent-sur-Seine station

In this study, ecology is supposed to be protected by respecting normative on environmental flows (paragraph 1.4.2). As a matter of fact, environmental flow can be treated as a so-called *soft constraint*, which differs from *hard constraints* because is not given by physical (so unbreakable) limits, but by legislative ones. Therefore, soft constraints can be violated by paying an extremely high price, which should discourage their exceeding. Thus, in this analysis we consider the same cost function for both environmental flow violations and as exceptional drought events.

	Environmental Flow [m^3/s]	**Violation cost [$\$/m^3$]**
Aube River	2	100
Seine River	3	100

2.3.2 Lakes life quality

The indicator for lakes life quality has been chosen on the supposition that lakes operators aims to fill the reservoir until the top of the rule curve. Thus, under and above this volume there is a cost on lakes life quality: in fact, a too empty lake is sensitive to pollution and ecological carrying capacity diminution, with effects also on tourism and fishing; while an extremely filled lake can cause damages or discomforts to its operators.

3 Components model

This chapter is dedicated to the modeling of system components. It begins with the hydrological disturbances model, then it shows the linear model used for water stocks in reservoirs and finally it describes how hydraulic transfer has been derived.

3.1 Hydrological Scenarios Model

3.1.1 Model definition

3.1.1.1 Multivariate Periodic Auto Regressive models

The purpose of this study is not to give predictions on hydrological inflows, but to generate scenarios that should represent possible evolutions of the system. Auto-Regressive models are a reliable method to represent this kind of information.

Auto-Regressive models are classified as statistical empirical models. They are particularly interesting in describing Markov processes because they assume that a time-series output $y_t$is given by a linear combination of its precedent values y_{t-i}, as it follows:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t$$

where p is the auto-regression order of the model and ε_t is the error between model and time series. The model is validated if the error is estimated being a white noise, which means that it is no longer algorithmically compressible. In practice this situation is given when correlation of ε_t in time is close to zero, therefore, in other words, when all the possible information on the correlation of y_t in time has been used. At this point, ε_t can be assumed as a random variable with a given probability distribution. In the most simple case ε_t is Gaussian, so: $\varepsilon_t \sim \mathcal{N}(0, \sigma)$, σ being noise standard deviation. This is called an *Auto-Regressive model of order p*, and is indicated with the acronym AR(p).

For natural time series, the correlation of y_t in time can be strictly different across seasons. That's why ϕ parameter is preferred to be time-variant with yearly periodicity. So, Periodic AR(p) model (PAR(p)) is given by:

$$y_t = \sum_{i=1}^{p} \phi_{\tau,i} y_{t-i} + \varepsilon_t$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma_\tau)$$

with $\tau \in [1, \dots T]$

As described in chapter 1.1, the correlation between Seine ad Aube river systems is not negligible. Therefore, it is necessary to integrate also spatial correlation to PAR models, that can be given by using a multivariate distribution of noise, as, for instance: $\varepsilon_t \sim \mathcal{N}(0, \Sigma_\tau)$, being Σ_τ noise covariance matrix. However, spatial correlation on noise is not sufficient for describing the study case time series, where the two rivers have a strong correlation in the entire hydrological dynamic. In fact, if noises exhibit some correlations due to the proximity of basins, some correlation is expected also for y.[3] The following paragraph provides a description of a possible strategy to add further spatial information to multivariate PAR(p) models.

3.1.1.2 Conditioned Periodic Auto Regressive model

High correlation on catchments may be well explained if we consider that, together with noise, also time series data have an influence on other basin's inflows. Physically this correlation does not imply any casual relation, nonetheless, in the context of information theory, it gives an additive piece of information to the overall time series description[4].

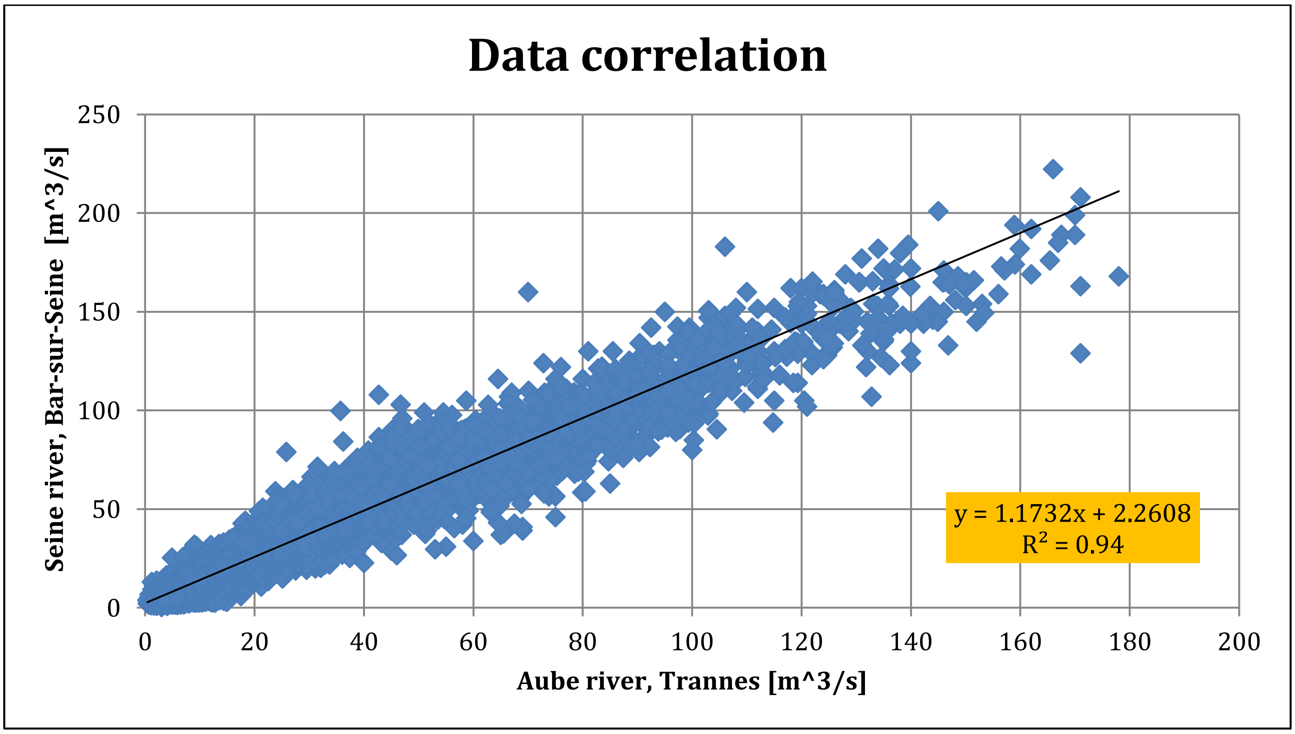


Figure 6 : Correlation between Seine and Aube rivers' flows at the first gauging stations, Bar-sur-Seine and Trannes.

[3] AR models with correlation on noises and not on time series are usually called "*Contemporaneous*", (Bartolini, 1988).

[4] *"The aim of an empirical model is to describe the relationship between the inputs and the outputs of a system without describing its internal* (physical) *processes"*. (Soncini-Sessa, 2007)

Given two correlated hydrological catchments, as the ones in Figure 6, the first, called y_t^M, is the main basin and the second, called y_t^C, is influenced (or conditioned) by the first. The behavior of the main catchment is obtained by a standard multivariate PAR(p) model, while the behavior of the conditioned one can be seen as a random difference β_t from the regression line between the two observed flows. The difference can be represented by an Auto Regressive model, to represent its temporal auto-correlation. Therefore, being $\varepsilon_t \sim \mathcal{N}(0, \Sigma_\tau)\ with\ \tau = 1, \dots, T$, the main basin model is represented as it follows:

(3-1)

$$y_t^M = \sum_{i=1}^{pM} \phi_{\tau,i}^M y_{t-i}^M + \varepsilon_t^M$$

While the conditioned basin model is given by:

(3-2)

$$y_t^C = \alpha y_t^M + \beta_t$$

where α is the linear regression coefficient between observed y_t^C and y_t^M. β_t, obtained by the relation $y_t^C - \alpha y_t^M$, is also given by a multivariate PAR(p) model, as it follows:

(3-3)

$$\beta_t = \sum_{i=1}^{pC} \phi_{\tau,i}^C \beta_{t-i} + \varepsilon_t^C$$

Rearranging equations (3-2) and (3-3), and using the information on equation (3-1), we obtain the following relation for conditioned catchments:

$$y_t^C = \underbrace{\alpha \sum_{i=1}^{\max(pM,pC)} \left(\phi_{\tau,i}^M - \phi_{\tau,i}^C\right) y_{t-i}^M}_{Main\ basin\ autocorrelation\ influence} + \underbrace{\sum_{i=1}^{pC} \phi_{\tau,i}^C y_{t-i}^C}_{Auto\ Regression} + \underbrace{\alpha \varepsilon_t^M + \varepsilon_t^C}_{Disturbance}$$

This is a particular version of lower-triangular periodic Auto-Regressive models, which are described in(Salas, 1985). This model will be called CAR(p) (Conditioned periodic Auto-Regressive model of order p). It can be expressed in the more general form:

(3-4)

$$y_t^j = \sum_{i=1}^{p} \left[\varphi_{\tau,i}^{j|M}\ y_{t-i}^M + \phi_{\tau,i}^j y_{t-i}^j\right] + \varepsilon_t^j$$

Where:

y_t^j is the time series j at time t (main or conditioned);

p is the maximum auto-regression order for the considered basins;

$\varphi_{\tau,i}^{j|M} = \alpha\left(\phi_{\tau,i}^M - \phi_{\tau,i}^j\right)$ is main basin influence parameter. If y_t^j is itself a main basin, α is zero and this parameter have no influence;

$\phi_{\tau,i}^j$ is the Auto-Regression parameter of basin j;

$\mathcal{E}_t^j = \alpha\varepsilon_t^M + \varepsilon_t^j$ is the noise of basin j at time t. If y_t^j is itself a main basin, α is zero and $\mathcal{E}_t^j$ is simply ε_t^j;

3.1.1.3 Multiplicative formulation

For ensuring solution positivity and enhance correct parameters calibration, y are logarithmic time series of inflows with zero-average:

$$y_t^j = \log\left(a_t^j\right) - \frac{\left(\sum_{\tau=1}^{Nyears} \log\left(a_{t\in\tau}^j\right)\right)}{Nyears}$$

where a_t^j is the flow contribution (from French: *"apport"*) of basin *j* at time *t*.

Hydrological flows are so given by the following relation:

(3-5)

$$a_t^j = \gamma_\tau^j \cdot \mathrm{e}^{y_t^j}$$

where γ_τ^j is time series geometrical mean.

We can derive the direct hydrological state transfer by applying equation (3-5) on (3-4):

$$a_t^j = \kappa_\tau^j \cdot \prod_i^p \left[\left(a_{t-i}^M\right)^{\varphi_{\tau,i}^{j|M}} \cdot \left(a_{t-i}^j\right)^{\phi_{\tau,i}^j}\right] \cdot \xi_t^j$$

Where:

$$\kappa_\tau^j = \frac{\gamma_\tau^j}{\prod_i^p \left(\gamma_{\tau-i}^M\right)^{\varphi_{\tau,i}^{j|M}} \cdot \left(\gamma_{\tau-i}^j\right)^{\phi_{\tau,i}^j}}$$

$$\xi_t^j \sim \log\mathcal{N}(\mathrm{M}_\tau, \Sigma_\tau)$$

3.1.1.4 Multiplicative model linearization

Multiplicative model utilization involves the problem of non-linearity of hydrological state transfer. Model linearization is obtained here via first order Taylor's expansion:

$$\tilde{f}(x_1, \dots, x_n) = f(\overline{x_1}, \dots, \overline{x_n}) + \frac{\partial f(\overline{x_1}, \dots, \overline{x_n})}{\partial x_1}(x_1 - \overline{x_1}) + \dots + \frac{\partial f(\overline{x_1}, \dots, \overline{x_n})}{\partial x_n}(x_n - \overline{x_n})$$

By centering Taylor's expansion on the climatic geometric average[5], the final linear multiplicative CAR(p) model becomes:

(3-6)

$$\tilde{a}_t^j = \sum_i^p \left(\rho_{\tau,i}^{j|M} a_{t-i}^M + \rho_{\tau,i}^j a_{t-i}^j\right)\xi_t^j + \psi_\tau^j \xi_t^j$$

[5] The choice of centering Taylor's expansion on geometric (and not arithmetic) average is given by the fact that the multiplicative model of a_t^j is itself centered on γ_τ^j.

Where:

$$\rho_{\tau,i}^{j|M} = \gamma_\tau^j \frac{\varphi_{\tau,i}^{j|M}}{\gamma_{\tau-i}^M}$$

$$\rho_{\tau,i}^{j} = \gamma_\tau^j \frac{\phi_{\tau,i}^{j}}{\gamma_{\tau-i}^j}$$

$$\psi_\tau^j = \gamma_\tau^j \left(1 - \sum_i^p \left(\varphi_{\tau,i}^{j|M} + \phi_{\tau,i}^j\right)\right)$$

Some final considerations about linear multiplicative CAR models:

These models do not ensure solution's positivity. However, if their order (given by parameter p) is adequately low, the probability of having negative values of $\tilde{a}_t^j$ remains sufficiently close to zero.

Linearization via first order Taylor expansion may overestimate errors if compared to non-linear model. This difference can be minimized by including overestimation in the overall error of the model, represented by random noise ξ_t. The calibration of ξ_t is described in paragraph 3.1.2.

3.1.1.5 Linear CAR(p) model: Matrix representation[6]

$$\mathbf{x}_t^{Hydro} = \sum_i^p \left(\boldsymbol{\rho}_{\tau,i} \times \mathbf{x}_{t-i}^{Hydro}\right) \circ \boldsymbol{\xi}_t + \boldsymbol{\Psi}_\tau \circ \boldsymbol{\xi}_t \qquad (3\text{-}7)$$

Where:

$$\mathrm{x}_t^{Hydro} = \begin{bmatrix} a_t^1 \\ \vdots \\ a_t^n \end{bmatrix}$$

$$\xi_t = \begin{bmatrix} \xi_t^1 \\ \vdots \\ \xi_t^n \end{bmatrix}$$

$$\rho_{\tau,i} = \Upsilon_{\tau,i} \circ \left(\Gamma_\tau \oslash \Gamma_{\tau-i}^{\mathrm{T}}\right)$$

$$\Upsilon_{\tau,i} = \mathcal{A} \circ \left(\Phi_{\tau,i}^{\mathrm{T}} - \Phi_{\tau,i}\right) + \mathrm{diag}(\Phi_{\tau,i})$$

$$\Psi_\tau = \sum_i^p \left[\mathrm{vec}(\mathrm{I}) - \mathrm{sum}\left(\Upsilon_{\tau,i}\right)\right] \circ \mathrm{vec}(\Gamma_\tau)$$

$$\Phi_{\tau,i} = \begin{bmatrix} \phi_{\tau,i}^1 & \cdots & \phi_{\tau,i}^1 \\ \vdots & \ddots & \vdots \\ \phi_{\tau,i}^n & \cdots & \phi_{\tau,i}^n \end{bmatrix} \quad \Gamma_\tau = \begin{bmatrix} \gamma_\tau^1 & \cdots & \gamma_\tau^1 \\ \vdots & \ddots & \vdots \\ \gamma_\tau^n & \cdots & \gamma_\tau^n \end{bmatrix} \quad \mathcal{A} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}$$

3.1.2 Model calibration

CAR(p) linear model requires the calibration of four sets of parameters:

1. Linear regression parameters for spatial correlation: α_{jM} ;
2. Auto-Regression parameters: $\phi_{\tau,i}^j$;

[6] Operators notation used for matrix representation is described in *List of symbols.*

3. Hydrologic time series geometric mean values: γ_τ^j;
4. Statistic properties of log normal noise distribution, given by: μ_τ and Σ_τ.

The first two sets of parameters have been obtained using the mean-squared method on logarithmic time series, while the third is easy to derive directly from data. The fourth set of parameters has been calculated on the difference between linear model and data, in order to include linearization correction in an adjusted distribution of noise (which generally has a mean value inferior to unity, to avoid overestimation). Therefore, being e_t^j the overall error of multiplicative linear CAR(p) model, ξ_t^i is represented as it follows:

$$\xi_t^j \sim \log \mathcal{N}\left(E\left[e_{t\in\tau}^j\right], \Sigma\left[e_{t\in\tau}^j\right]\right)$$

With:

$$e_t^j = \frac{a_t^{j,observed}}{a_t^{j,linear\ model}}$$

3.1.3 Model validation

3.1.3.1 Application on study case

3.1.3.1.1 Choosing between main and conditioned basins

The first element to consider in the application of CAR(p) model for a real study-case is the choice on what basins should be *conditioning* (main basins) or *conditioned* (influenced by main basins). This issue can be decomposed in two separate instances:

- Which basins have an interrelation such that we can state that they have mutual influence;
- Which of them should be the main basin;

About the first instance, linear correlation coefficients can be employed in order to observe mutual influence between basins. One measure of goodness of fit is the *coefficient of determination*, or R^2. This statistic indicates how closely values obtained from fitting a regression line match the dependent variable the line is intended to represent. Statisticians often define R^2 using the residual variance from a fitted model:

$$R^2 = 1 - SS_{resid} / SS_{total}$$

SS_{resid} is the sum of the squared residuals from the regression. SS_{total} is the sum of the squared differences from the mean of the dependent variable (*total sum of squares*). An R^2 greater than 0.8 usually confirms that two basins are well correlated.

For what concerns the second instance, once a group of mutually dependent basins is found, the definition of main basin is relative to volume incidence: the basin with greatest mean flow is the one that should influence the others.

The following table resumes these parameters for study-case gauging stations:

R^2	Trannes	Lassicourt	Arcis	Bar	Mery	Nogent	**Mean flow [m^3/s]**
Trannes	1	**0.9304**	0.4874	**0.9445**	0.6608	0.3669	19.2717
Lassicourt	**0.9304**	1	0.4544	**0.8853**	0.6357	0.3476	7.1738
Arcis	0.4874	0.4544	1	0.5324	**0.8997**	**0.8741**	5.5064
Bar	**0.9445**	**0.8853**	0.5324	1	0.7004	0.4133	24.9575
Mery	0.6608	0.6357	**0.8997**	0.7004	1	**0.8081**	4.4120
Nogent	0.3669	0.3476	**0.8741**	0.4133	**0.8081**	1	6.1879

Table 4 : Parameters used for the determination of main and conditioned basins: R^2 indicates which basins are best correlated, while the higher mean flow indicates which basins are suitable to be mean basins. Coefficients in bold are the ones which express a good estimated correlation

From precedent table we can observe that the system is composed by two groups of spatially correlated hydrological basins, the first, measured at Trannes, Lassicourt and Bar-sur-Seine stations, the three highest-flow catchments, and the second composed by the catchments measured at Arcis-sur-Aube, Mery and Nogent-sur-Seine stations. For the first group, Bar-sur-Seine has been chosen as main basin because of its preponderant mean flow, while, for the second group, Arcis-sur-Aube was preferred to Nogent-sur-Seine for its better correlation with other basins.

Table 4 shows a great dissimilarity on the inter-correlation between the two groups of basins. Such a difference may be explained by the fact that time series on second catchments' group are not provided by data, but derived from ClimAware project's hydrological model.

3.1.3.1.2 Catchments at reservoirs inlets and outlets

In the context of ClimAware project, four intermediary basins have been artificially created to consider hydrological contributions at the inlet and at the outlet of Seine and Aube reservoirs. Inflows have been calculated as a fraction of the closest downstream station flow, with respect to the relative catchment area. Therefore, the ones relative to Aube Reservoir are directly proportional to Arcis-sur-Aube time series, while the ones relative to Seine Reservoir are directly proportional to Mery-sur-Seine time series. Respectively, they have been called: Aube In, Aube Out, Seine In, Seine Out.

In figure below we can finally observe catchments localization and main features. Figure 7: Catchments spatial localization and main features

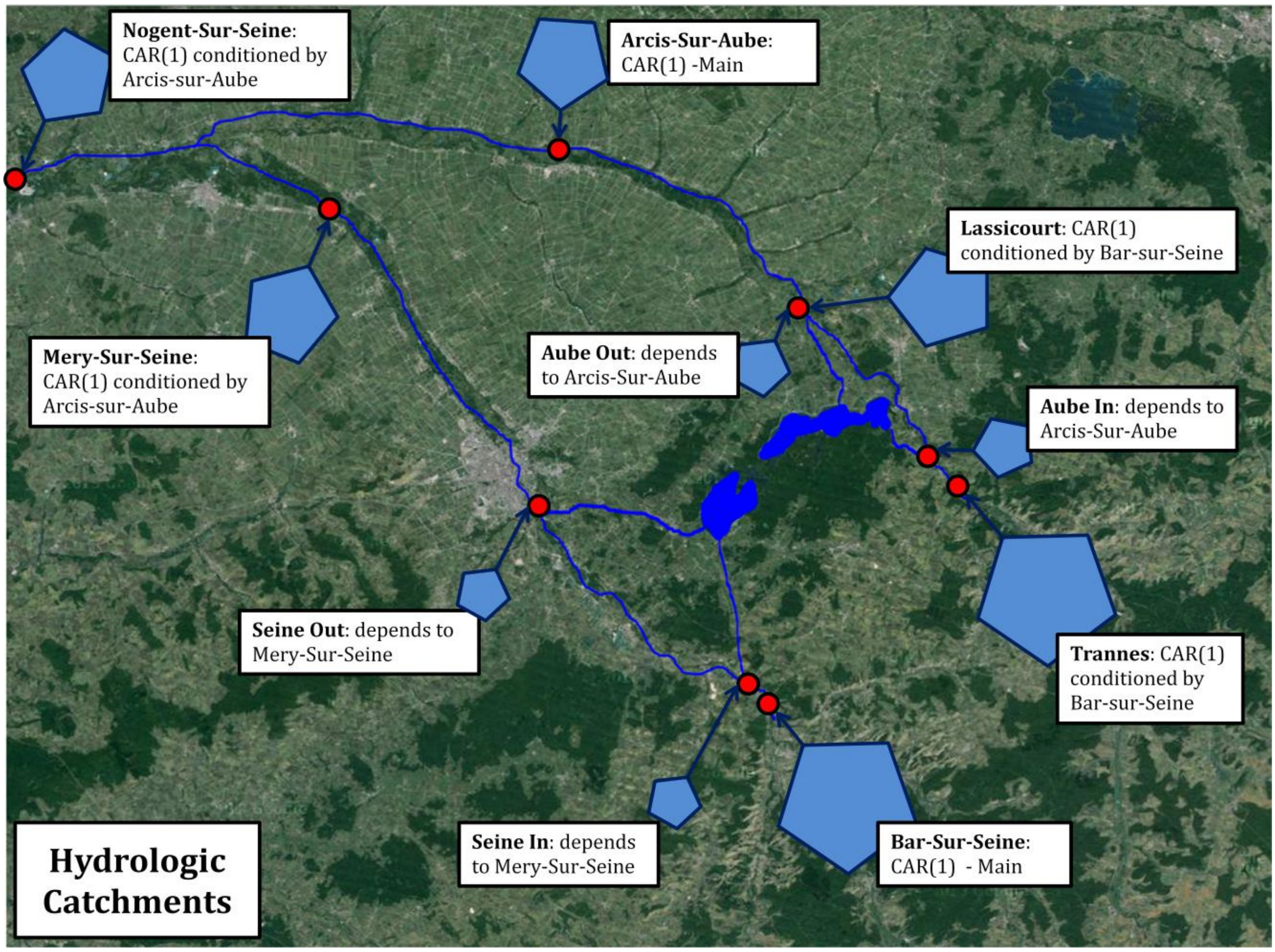


Figure 7: Catchments spatial localization and main features.

Hydrological state transfer equations for the study case are provided in Annex III: State transfer equations.

3.1.3.2 Linear CAR(p) model testing

Test the multivariate linear CAR(p) model:

1. Model's error should be white, thus uncorrelated in time.
2. Observed model error should be distributed as pdf of ξ_t.
3. Artificially generated scenarios must well represent observed time series in terms of some hydrological statistics.
4. Good representation of inter-correlation among catchments.

The following paragraphs provide a description of these four aspects.

3.1.3.2.1 Error auto-correlation

The Auto-Correlation Function (ACF) measures the signal auto-correlation at different lags. Theoretically, error autocorrelation should stay inside an interval close to zero for every lag value. If it decreases with exponential rate, it means that CAR(p) model's order should be incremented; if it doesn't decrease to the confidence interval, a CAR(p) model will never have white noise. In this case, auto-correlation converges to values which are often outside confidence interval. Figure below shows ACF for the hydrological inflows of the system.

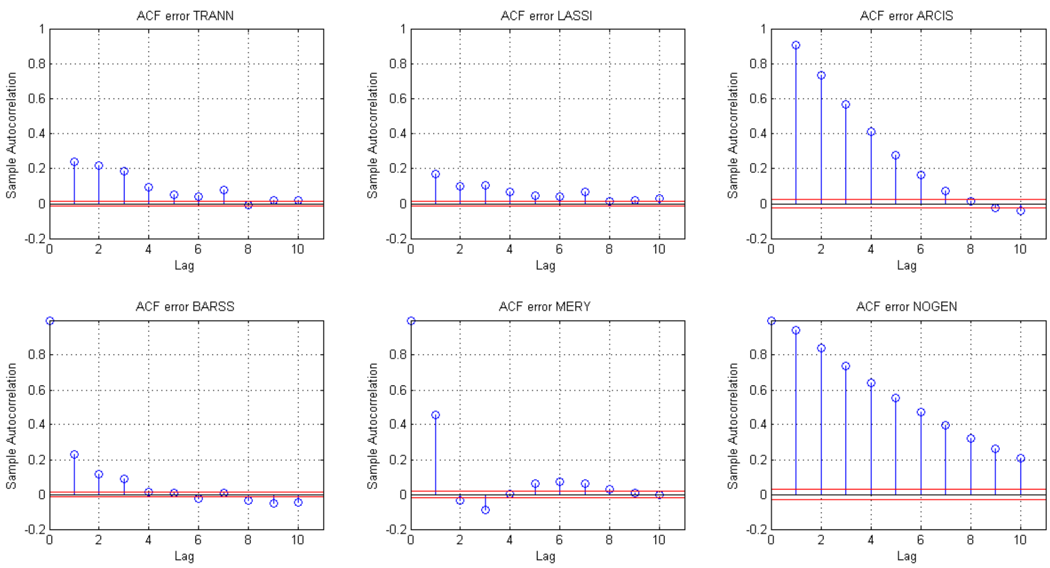


Figure 8: Auto-Correlation Function on errors.

From Figure 8, it is clear that error is not a completely white process: while lag-1 autocorrelation for the first group of basins (Trannes, Lassicourt and Bar-sur-Seine) never exceeds 0.3, the one relative to the others is very high. We think that this may be due to the differences on data survey explained in paragraph 3.1.3.1.1, in fact hydrological model provided by ClimAware project suffers of a high slow term error autocorrelation: this can explain why autocorrelation on Arcis and Nogent decreases very slowly. In this case a better description of the hydrological process may be given by an auto regression order increment. However, being Mery, Arcis and Nogent all weak contributions to the overall inflow, we preferred not to treat them with additional regressors. Moreover, to have a perfect description of overall hydrologic autocorrelation, a lot of state variables should be added, with negative consequences on SDDP computational effort. Therefore, we judged auto regression order of the model sufficient for the purpose of this study.

3.1.3.2.2 Error distribution convergence

Observed disturbances should match with errors distribution (which is log-normal). Therefore, error time series' logarithm can be compared to a normal distribution to observe the coherence of noise distribution law. For this purpose, normal probability plots (called *Normplots*) have been employed:

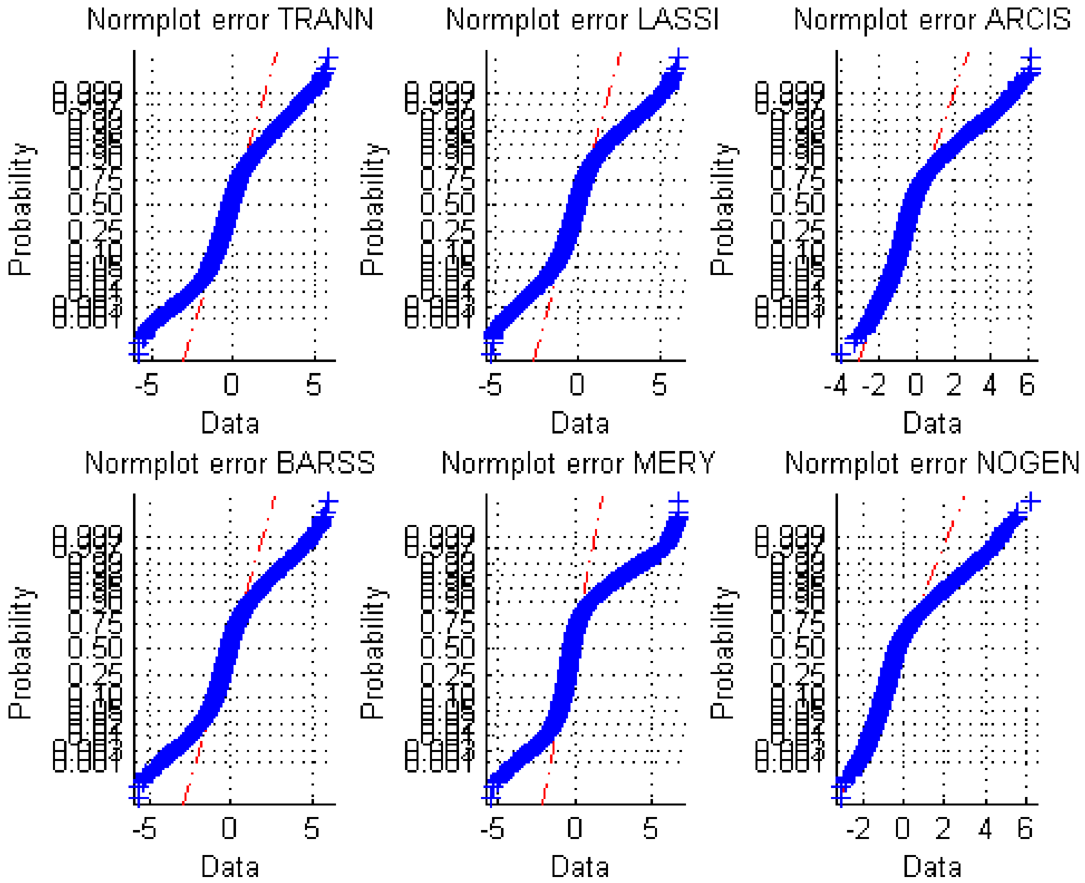


Figure 9: Normal probability plots on standard logarithmic error distributions. Time series distribution (in blue) is compared to a Gaussian cdf (red).

We can observe that logarithmic errors do not follow exactely a normal cdf; this is the case of most natural time series. For the study-case, the behavior close to average is good, but extreme values probability (mostly for high flows) is generally underestimated by a standard Gaussian function. It is interesting to observe that errors distribution directly depends to errors autocorrelation degree: for poorly auto correlated basins, such as Arcis and Nogent, error cdf has a particular pattern which follows Gaussian behavior for low flows. Thus, a correct estimation for the real distribution should be taken on white (or "*no longer algorithmically compressible*") errors, with, for instance a generalized log-normal function. This could be an interesting improvement on noise scenarios generation.

3.1.3.2.3 Time series statistics convergence

The overall model statistics should converge to data for a large simulation horizon. For our case, the first important information concerns the respect of inflows' mean value. Additive information is given by time series simulated standard deviation: when larger than observed data, the model explores a wider ensemble of daily scenarios than data, and vice versa.

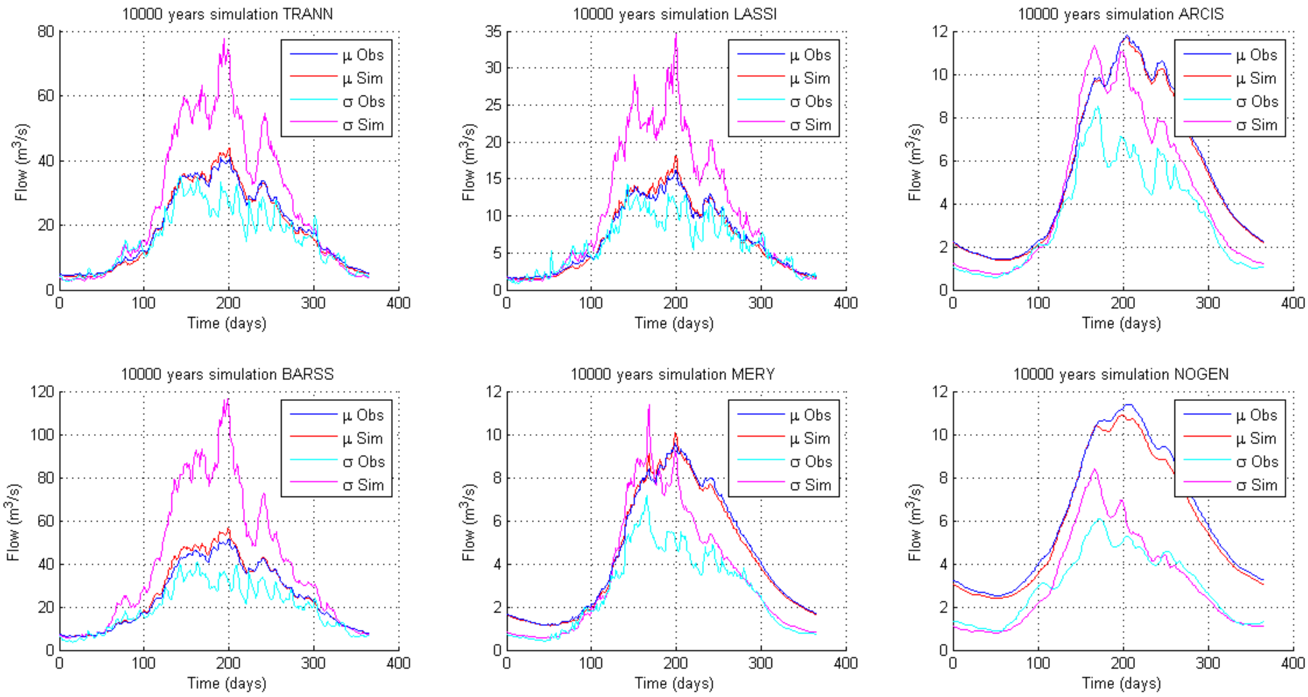


Figure 10: Seasonal mean and standard deviation for observed and simulated time series for a 10 000 years simulation

Figure above shows that mean flow value is respected, while simulated standard deviation is greater than the observed one. This may be due to the low autocorrelation order of the model, which is responsible of the generation of "low-memory" scenarios, where informative deficit on autocorrelation is propagated along the simulation. Thus, high standard deviation may be caused by the super position of low-memory errors for a prolonged sequence of stages. Anyway, this phenomenon will not occur in the optimization phase, because SDDP scenarios extraction is stage-wise, with initial condition always given by data.

3.1.3.2.4 Spatial inter-correlation

The analysis of correlation between basins is carried out by simple comparison on dispersion diagram between data and a 48-years simulation (as many years as dataset). Results are showed in the image below.

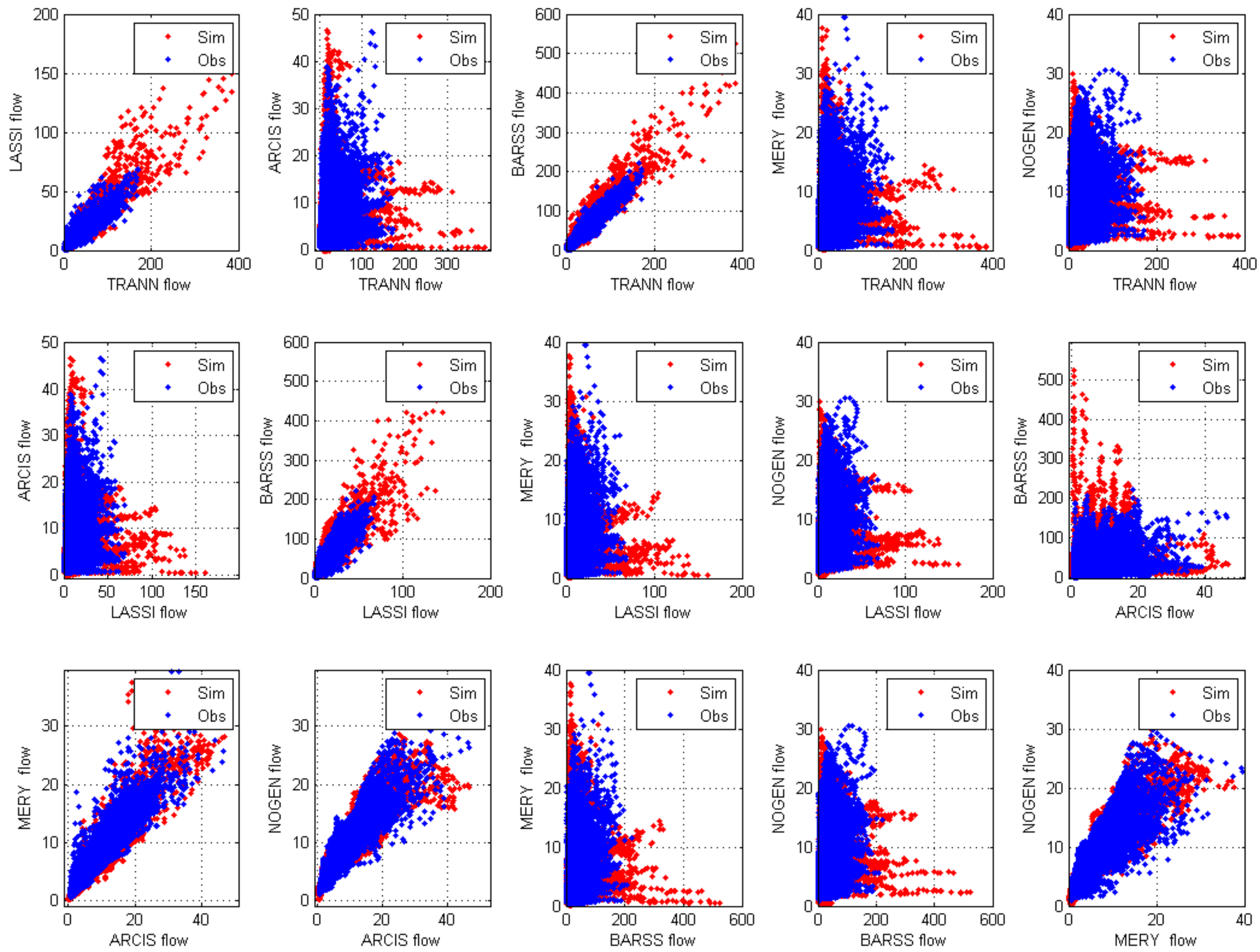


Figure 11: Data and simulations spatial correlation over a 48 years simulation (as many as dataset). Blue dots are observed flows [m^3/s] and red dots are simulated flows [m^3/s].

Dispersion diagrams show that CAR(*p*) model gives a good interpretation of basins mutual influence.

3.1.3.2.5 Final considerations on model testing

CAR(*p*) model tested in this study is able to extract a stage-wise scenario for all considered inflows by using a whole information ensemble on disturbances, temporal auto-correlation and spatial inter-correlation. Improvements on the model performance may be given by acting in two directions:

- Have a better representation on temporal autocorrelation, *i.e.*, compress information on disturbances.
- Calibrate a more specific probability distribution law for disturbances.

Nonetheless, thanks to its overall good performances, this model is the most suitable for the study purpose, also because it permits to relieve computational effort given by excessive complexity.

3.2 Water stocks model

Water stocks in reservoirs can be represented using a linear mass-balance model by observing the following hypothesis:

- Lakes' inflows and releases are known at time *t*;
- Releases are not sensitive to storage-discharge relation;
- Lakes are cylindrical.

The first hypothesis is verified for our system because lakes' inflows and releases are supposed to be completely controlled. For what concerns the second hypothesis, releases do not depend on storage-discharge relation because they have been designed to provide always the exact quantity of water needed. The third hypothesis is verified too, because, as long as evaporation is neglected, lakes' fillings are linear.

The final model for reservoirs is given by the following:

$$s_t^i = s_{t-1}^i + \Delta t \cdot u_t^i - \Delta t \cdot r_t^i \qquad (3\text{-}8)$$

$subject\ to$:

$$s_{min}^i \leq s_t^i \leq s_{min}^i \qquad (3\text{-}9)$$

$$0 \leq r_t^i \leq r_{max}^i \qquad (3\text{-}10)$$

Where:

s_t^i : is the volume of reservoir *i* at time *t*; [m^3]

u_t^i: is reservoir's *i* controlled inflow at time *t*; [m^3/s]

r_t^i: is reservoir's *i* controlled discharge at time *t*; [m^3/s]

$\Delta t = 86400 \left[\frac{seconds}{day}\right]$;

s_{min}^i: is reservoir's *i* minimum volume; [m^3]

s_{max}^i: is reservoir's *i* maximum volume; [m^3]

r_{max}^i: is release *i* maximum capacity. [m^3/s]

Water stocks state transfer equations for the study case are provided in Annex III: State transfer equations.

3.3 Hydraulic model

3.3.1 Model definition

The model employed for representing hydraulic transfer along natural watercourses is LLR (Linear Lag and Route) model, the same as the one used in ClimAware project, and was already calibrated in that context. Using the same hydraulic model because this gives the possibility to compare our model results with the ones of ClimAware project. The LLR model is defined as it follows:

(3-11)

$$q_t^d = \sum_{i}^{n\ branches} (1 - m_i) q_{t-1}^d + m_i \left[d_i q_{t-1-n_i}^i + (1 - d_i) q_{t-n_i}^i \right]$$

Where:

- q_t^d [m3/s] is downstream flow at time t;
- $n\ branches$ is the total number of upstream branches;
- q_t^i [m3/s] is upstream branch *i* flow at time t [m3/s];
- m_i [#] is the attenuation coefficient of upstream branch *i* flow: $m_i \in [0,1]$;
- n_i [days] is delay integer part for upstream branch *i* flow: $n_i \in [0, \infty]$;
- d_i [days] is delay decimal part for upstream branch *i* flow: $d_i \in [0,1]$;

Hydraulic transfer is conservative, therefore all mass losses are implicitly considered on hydrologic information, which is supposed to provide net inflows.

Hydraulic state transfer equations for the study case are provided in Annex III: State transfer equations.

3.3.2 Withdrawals

3.3.2.1 Reservoirs controlled intakes

Withdrawals for reservoirs filling are subject to physical constraints, related to their capacity, and normative constraints too, related to ecological flow (see paragraph 1.4.2). The following scheme provides an example of controlled withdrawal to reservoirs (expressed by the variable *u(t)*):

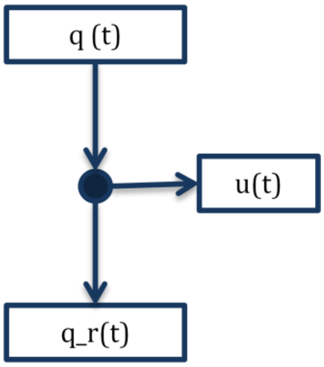


Figure 12: Scheme for exemplifying withdrawals for reservoirs: q(t) is the total inflow, u(t) is reservoir's withdrawal, q_r(t) is environmental flow.

Reservoirs inlets capacity constrictions are expressed as *hard* constraints by the following:

$$0 \leq u(t) \leq \bar{u} \qquad (3\text{-}12)$$

Besides, one more hard constraint must be added to ensure non-negative flows:

$$q\_r(t) \geq 0 \qquad (3\text{-}13)$$

Ecological flows are expressed as *soft* constraints, thus, their violation gives a cost, as explained in paragraph 2.3.1.

3.3.2.2 Withdrawals for civil use

Along the rivers there are two withdrawals that have not been implicitly considered in hydrological information:

- Arcis-sur-Aube: withdrawals for industries and irrigation purposes;
- Nogent-sur-Seine: withdrawals for industries, supply and irrigation purposes.

However, water demand at these points is very small. Therefore, these withdrawals have been taken into account in the hydraulic model (thus as hard constraints) because they can always be ensured.

4 Management Design

This chapter provides the system setting for optimization with SDDP algorithm. Thus, all system components, such as hydrological model, stocks model and hydraulic model will be incorporated in a unique state transition relation, and additive constraints will be added to the overall setting of the problem. Then, step-costs will be defined in terms of objectives and supplementary constraints, and relative weights will be chosen. Once given the overall architecture of the problem, SDDP parameters will be planned for optimization. Finally, optimization performance will be analyzed.

4.1 Optimal Control problem

Optimal control problem is defined as a finite horizon, yearly-periodic and daily time step management problem under uncertainty and multi-objective (also called partial rationality) conditions. Uncertainty is filtered using Laplace criterion, and partial rationality is faced via weighting method:

$\min_{u_t}\{\mathrm{E}_{\xi_t}[\pi \cdot g_t + F_t^*]\}$		(4-1)
$x_t = f(x_{t-1}, u_{t-1}, \xi_t)$	$t = 1, \dots, H$	(4-2)
$g_t \geq h(x_t, u_t)$	$t = 1, \dots, H$	(4-3)
$x_t \in \mathcal{X}$	$t = 1, \dots, H$	(4-4)
$u_t \in \mathcal{U}$	$t = 1, \dots, H$	(4-5)
$\xi_t \sim \log \mathcal{N}(\mu_\tau, \Sigma_\tau)$	$t = 1, \dots, H$	$\tau = 1, \dots, T$

where *f* and *h* are linear functions. Equation (4-6) (objective function) and (4-7) are discussed in paragraph 4.1.2; relations (4-8) and (4-9) are discussed in paragraph 4.1.1.4. This problem is the adaptation of problem (2-1) to the present study-case.

4.1.1 State transition

Overall state transition is given by the linear relation:

$$A^{eq} x_t = \left(B_t^{eq} x_{t-1} + c_t^{eq}\right) \circ \Xi_t$$

It can be obtained by the composition of system components state transitions. They can be resumed in three relations, one for hydrologic inputs, one for hydraulic transfers and one for stocks:

4.1.1.1 Hydrologic inputs

Following relation (3-7) for CAR(1) hydrological model (which is the case of all our catchments), it is easy to derive the relative state transition:

$$A^{eq\_Hydro} x_t^{Hydro} = \left(B_t^{eq\_Hydro} x_{t-1}^{Hydro} + c_t^{eq\_Hydro}\right) \circ \Xi_t^{Hydro}$$

Where A^{eq_Hydro} is a squared diagonal identity matrix, $B_t^{eq\_Hydro}$ corresponds toρ_τ, $c_t^{eq\_Hydro}$ is Ψ_τ and Ξ_t^{Hydro} is ξ_t.

4.1.1.2 Hydraulic transfers

Besides hydraulic variables, water transfer depends to hydrologic inputs and decisions too. Therefore, A^{eq_Hydra} and $B_t^{eq\_Hydra}$ have dimensions ($N_{hydra} \times N_{hydro}+N_{hydra}+N_{Dec}$), while vector $c_t^{eq\_Hydra}$ have dimensions ($N_{hydra} \times 1$). Thus, relation (3-11) is expressed by the following:

$$A^{eq\_Hydra} x_t^{Hydro\_Hydra\_Dec} = B_t^{eq\_Hydra} x_{t-1}^{Hydro\_Hydra\_Dec} + c_t^{eq\_Hydra}$$

Where A^{eq_Hydra} contains parameters relative to hydraulic variables and hydraulic lag variables at time *t*, while B^{eq_Hydra} contains parameters relative to hydraulic and hydraulic lag variables at time *t-1*. Finally, $c_t^{eq\_Hydra}$ contains information on eventual withdrawals for civil uses explained in paragraph 3.3.2.2.

4.1.1.3 Water stocks in reservoirs

Reservoirs are influenced by decisions. That's why, A^{eq_Res} and $B_t^{eq\_Res}$ matrices have dimensions ($N_{Res} \times N_{Res}+N_{Dec}$), while vector $c_t^{eq\_Res}$ have dimensions ($N_{Res} \times 1$) Thus, mass-balance relation (3-8) is given by:

$$A^{eq\_Res} x_t^{Res\_Dec} = B_t^{eq\_Res} x_{t-1}^{Res\_Dec}$$

Where A^{eq_Res} contains coefficients for inflow, release and volume at time *t*, while $B_t^{eq\_Res}$ contains unit coefficients for reservoirs in *t-1*.

4.1.1.4 Inequality constraints for state transition

Constraints on reservoirs' and channels' capacity, as well as reservoirs' intakes withdrawals, given by relations (3-9), (3-10), (3-12) and (3-13), are inside dominions (4-4) and (4-5), that in standard programming languages are represented by lower and upper boundaries vectors:

$$lb \le x_t \le ub$$

4.1.2 Objective function

4.1.2.1 Representing piece-wise linear step-costs

Step-costs defined in paragraph 2.3 are piece-wise linear functions which can be employed in linear programming by adding an inequality constraint for every function's linear "piece". For instance, the minimization of step-costs given by the relation $g_t = w(q_t - \bar{q}_t)^+$ can be treated as it follows:

$$\min_{u_t}\{g_t\}$$
$$\text{s.t.}\; g_t \geq w(q_t - \bar{q}_t)$$
$$g_t \geq 0$$

It is easy to derive that a sum of piece-wise linear step-costs for the same station $g_t = \sum_i^N w_i\left(q_t - \bar{q}_t^i\right)^+$ can be obtained by adding N constraints of the type: $g_t \geq w_i\left(q_t - \bar{q}_t^i\right)$ to the previous problem. All these constraints are contained in relation (4-10) of the overall problem.

The number of variables g is so given by the number of objective points, that are:

- Three stations for flood and drought protection (Arcis-sur-Aube, Mery-sur-Seine and Nogent-sur-Seine)
- Two stations for ecology protection (Aube and Seine Lake Inlets)
- Two points for lakes life quality protection (Aube and Seine Lakes).

that gives 7 step-cost variables.

4.1.2.2 Weighting step-costs

Step costs are weighted on the base of priority, which have been discussed in *Reconnaissance* phase (paragraph 1.2). Thus, priority order has been defined as it follows:

I. Nogent-sur-Seine, for its strategic importance in protecting Paris and the nuclear power plant;
II. Aube and Seine Lake Inlets, for being constrained by law on ecology protection;
III. Mery-sur-Seine, for representing the interest of the city of Troyes;
IV. Arcis-sur-Aube;
V. Aube and Seine Lakes, for minor priority on lakes life quality protection.

However, weights do not depend to priority only; they are assigned also according to practical considerations on real system behavior: high flows' thresholds trespassing are rare for Mery-sur-Seine and Arcis-sur-Aube stations, while they occur easily for Nogent-sur-Seine station. This happens because flood thresholds for Nogent are strictly inferior to the sum of Mery-sur-Seine and Arcis-sur-Aube flood thresholds. The same situation is given for droughts: droughts thresholds for Nogent are strictly greater than the sum of Mery-sur-Seine and Arcis-sur-Aube droughts thresholds. Thus, Nogent's weight should take into account its major sensibility to floods and droughts, besides the fact that it is the first priority station.

System behavior is very sensitive to lakes life quality weights definition, because of their direct influence on reservoirs state. Thus, they are strictly inferior to other weights, but not negligible. Table below shows the combination of weights chosen for this study:

Stations	**Priority**	**Weight [%]**	**Objective**
Aube Inlet	II	10	Ecology protection
Arcis-sur-Aube	IV	1	Flood and drought protection
Seine Inlet	II	10	Ecology protection
Mery-sur-Seine	III	5	Flood and drought protection (also for Troyes)
Nogent-sur-Seine	I	100	Flood and drought protection (also for Paris and nuclear power plant)
Aube Reservoir	V	10^{-3}	Lakes life quality
Seine Reservoir	V	10^{-3}	Lakes life quality

Finally, cost-to-go function, being practically considered as a step-cost to be minimized, needs the attribution of a weight. We chose to assign the value 0.99 to avoid influence of too far events in the future[7]. Therefore, future total cost is discounted.

4.1.3 Overall System Architecture

Optimal control problem defined at the beginning of paragraph 4.1 is finally represented in the following generalized form:

$$\min_{x_t} \pi^T x_t$$

$$subject\ to: \begin{cases} A^{eq} x_t = \left(B_t^{eq} x_{t-1} + c_t^{eq}\right) \circ \Xi_t^j \\ A^{ineq} x_t \geq b_t^{ineq} \\ x_t \in \mathcal{X} \end{cases}$$

Where:

x_t is a vector containing all system variables at time t: from state variables, to decisions, until step-costs and cost-to-go;

π is weights vector;

[7] For periodic systems, the influence of events at distance T to the present one is weighted $(\pi)^T$. Thus for system with T=365, a value π =0.99 ensures next year's influence sufficiently close to zero.

A^{eq} and B_t^{eq} are matrices composed by as many rows as state transitions and as many columns as the dimension of x_t;

Ξ_t^j is disturbances vector

A^{ineq} is a matrix composed by as many rows as step-costs inequalities and as many columns as the dimension of x_t;

$\mathcal{X}$ is the definition ensemble for states, decisions and step-costs.

This problem shows an evident correspondence to problem (2-1). SDDP algorithm will then provide a collection of linear cuts, which parameters will be added to A^{ineq} and b^{ineq} matrices.

4.2 Optimization setting

SDDP algorithm requires a set of initialization parameters:

- Initial value of x_t, called x_0;
- Maximum number of cuts and number of scenario extractions;
- Optimization horizon.

For the hydrological and hydraulic part, x_0 has been initialized on the average value of historic observations, while reservoirs part has been initialized to the objective rule curve of the current management; other components of x_0 (step-costs and cost-to-go) have been initialized to zero.

The total number of linear cuts has been fixed to 5000, with 1000 adding-cut iterations. Such a number of linear cuts is due to the rarity of floods/droughts events.

Optimization horizon is defined in order to have a good representation of cost-to-go function at different stages. As a matter of fact, being the final stage cost-to-go equal to zero, if chosen horizon is too short, decisions will not be affected by long-term effects. On the other hand, it is unnecessary to have too long optimization horizon, because for periodic systems cost-to-go is periodic too, thus it would not give any improvements, and computational effort would increase. Therefore, optimization horizon has been chosen for a two-years period: this means that cost-to-go effects are relative to a lapse of one year at least, which corresponds to system periodicity.

5 Impacts estimation

The optimal policy found with SDDP algorithm is tested in simulation. In the first section of this chapter, a set of impacts estimation indicators is provided on the base of initial research questions. In the second section, indicators are compared to current management. The third section provides a critical analysis on SDDP optimization final results.

5.1 Risk assessment

5.1.1 Probability, Duration, Intensity

Operational step-costs, defined in chapter 2.3, have been derived on the base of SDDP requirements, such as convexity, linearity and separability. On the other side, impacts estimation indicators are not subject to any constraints, and can be freely defined on the base of stakeholders' real interests. Literature on risk assessments for water resources systems is wide and several indicators (or combinations of them) have been proposed to cope with this issue. In last decades, since the publications by (Hashimoto, 1982) and (Fiering, 1982), it is common to evaluate risk sustainability of a scenario on the base of Reliability, Resilience and Vulnerability estimators (RRV). However, this set is not rigorously defined in literature and different arrangements of RRV indicators are proposed. In next paragraphs we derive a combination of RRV estimators inspired by (Kjeldsen, 2004) and based on the principle of *communicative transparency*: I think that one important level of the integration paradigm should be among *information*: elaborate estimators are mastered principally by hydrologists, while they should be easily understood by the totality of stakeholders.

5.1.1.1 Event Probability

The first indicator is linked to reliability formulation. It is generally adopted in the estimation of SF impacts on activities connected to water. It is defined as it follows:

$$\mathbb{P}_E = \frac{1}{N \cdot T} \cdot \sum_{j=1}^{M} d_j^E \qquad \left[\frac{Failure\ Days}{Total\ Days}\right]$$

Where:

E indicates the SF event, such for instance flood or drought event;

N is the simulation horizon [years];

T is system period, practically 365 [days/year];

d_j^E is the duration of j-th event of type E [days];

M is the total number of events over the simulation horizon.

$\mathbb{P}_E$ estimator gives the Sistem Failure (SF) probability over the simulation horizon. This is a good indicator to be graphically represented for comparison, but its practical value, being a probability value, is not easily grasped by the reader. Therefore, we chose to define a more transparent declination of this estimator for having a better representation of its output value. As a matter of fact, event probability can be represented also in terms of *Return Period*, as it follows:

$$\mathbb{T}_E = \frac{(\mathbb{P}_E)^{-1}}{T} \qquad [years]$$

Thus, for instance, a graphical comparison of two events characterized by $\mathbb{P}_1 = 0.0001$ and $\mathbb{P}_2 = 0.0002$ gives the information that "event 2 occurs with double probability than event 1", while a comparison on return periods gives that "event 1 returns every $\mathbb{T}_1 = 27{,}4$ years, while event 2 returns every $\mathbb{T}_2 = 13{,}7$ years". Furthermore, one can also derive, for instance, that events comporting a threshold trespassing probability $\mathbb{P} \le \mathbb{P}_{thr}$, have return period $\mathbb{T} \ge \mathbb{T}_{thr}$.

5.1.1.2 Event Expected Duration

The second indicator is directly related to by resilience formulation. Its interest is about granting the system to cope with external stresses and disturbances given by SF events. It is defined as it follows:

$$\mathbb{D}_E = \frac{1}{M} \cdot \sum_{j=1}^{M} d_j^E \qquad \left[\frac{Failure\ Days}{Total\ Failure\ Events}\right]$$

$\mathbb{D}_E$ estimator gives the expected duration of a SF over the simulation horizon.

5.1.1.3 Event Expected Intensity

The last indicator corresponds to vulnerability standard formulation. It is generally considered as the most important because it gives an idea of physical and ecological potential damages implied in a SF. It is defined as it follows:

$$\mathbb{V}_E = \frac{1}{M} \cdot \sum_{j=1}^{M} v_j^E \qquad \left[\frac{Trespassing\ Flow}{Total\ Failure\ Events}\right]$$

Where v_j^E is a threshold trespassing flow for the j-th event of type E.

$\mathbb{V}_E$ estimator gives the expected trespassing flow over the simulation horizon. Thus, it can be interpreted as the expected intensity of a flood or drought event.

5.1.1.4 Further considerations on risk assessment indicators

Probability, duration and intensity estimators are calculated as average values for SFs. It could be more interesting to analyze also other statistics than average, as, for instance, the maximum value or the p–th fractile in the empirical cdf. Anyway, some of these statistics (as maximum value) are less robust than average value because they diverge by augmenting the simulation horizon, while others (as p–th fractile) are more difficult to communicate.

A way to manage this inconvenient can be given by using the same estimators on different types of events: for instance exceptional floods indicators can be added in order to complement information on flood indicators.

5.1.2 Impacts estimation setting

Impacts estimation analysis aim is to display a confrontation between different alternatives based on indicators application for stakeholders criteria. This can be provided by a table, called "*Table of Impacts*", which, in our case, is divided in "Impacts concerned with risk", and "Impacts not concerned with risk". The following paragraphs provide a description of these features. (NB: The table of impacts for the study case is available in Annex IV: Table of Impacts for 1961-1991 period.

5.1.2.1 Impacts concerned with risk

Monitoring stations being linked with risk assessment estimators are the ones concerned with flood and drought protection. Thus, every monitoring station for flood and drought protection has a set of three indicators ($\mathbb{P}_E$, $\mathbb{D}_E$, $\mathbb{V}_E$) for four events (floods, exceptional floods, droughts, exceptional droughts).

5.1.2.2 Impacts for Lakes Life Quality

This study provides also information on lakes life quality indicators, which are not concerned with risk assessment. Indicators proposed for lakes life quality are the followings:

- Lakes mean volume, given by relation:

$$\mu_{LLQ}^{j} = \frac{1}{H}\sum_{t=1}^{H} s_t^j \qquad [Mm^3]$$

- Lakes mean variation, given by relation:

$$\sigma_{LLQ}^{j} = \frac{1}{H}\sum_{t=1}^{H} \left| s_t^j - \mu_{LLQ}^{j} \right| \qquad [Mm^3]$$

The objective is to maximize lakes volume and minimize lakes variation.

5.2 Comparison between current and optimal management

The comparison between current and optimal management has been realized on 01/08/1961 – 31/07/1991 period. The table of impacts for the study case, available in Annex IV: Table of Impacts for 1961-1991 period, provides a confrontation between three alternatives:

- Natural system: impacts are calculated on the naturalized system, so as it should be without reservoirs influence.
- Current management: impacts are calculated on current policy simulation.
- SDDP Optimization: impacts are calculated on the simulation of centrally optimized management for a given weights combination.

In the following paragraphs a more detailed confrontation is provided by visual interpretation of time series and indicators.

5.2.1 Performance on floods and droughts protection

Results obtained from a multi-objective control problem are strongly sensitive to the choice of objective function weights. The following results are given by a combination of costs which gives priority to flood protection (see paragraph 2.3.1, *Floods, droughts and ecology*) and a combination of weights which gives priority to Nogent-sur-Seine station (see paragraph 4.1.2.2, *Weighting step-costs*).

Figure 13 shows a confrontation between current and SDDP management for Nogent-sur-Seine:

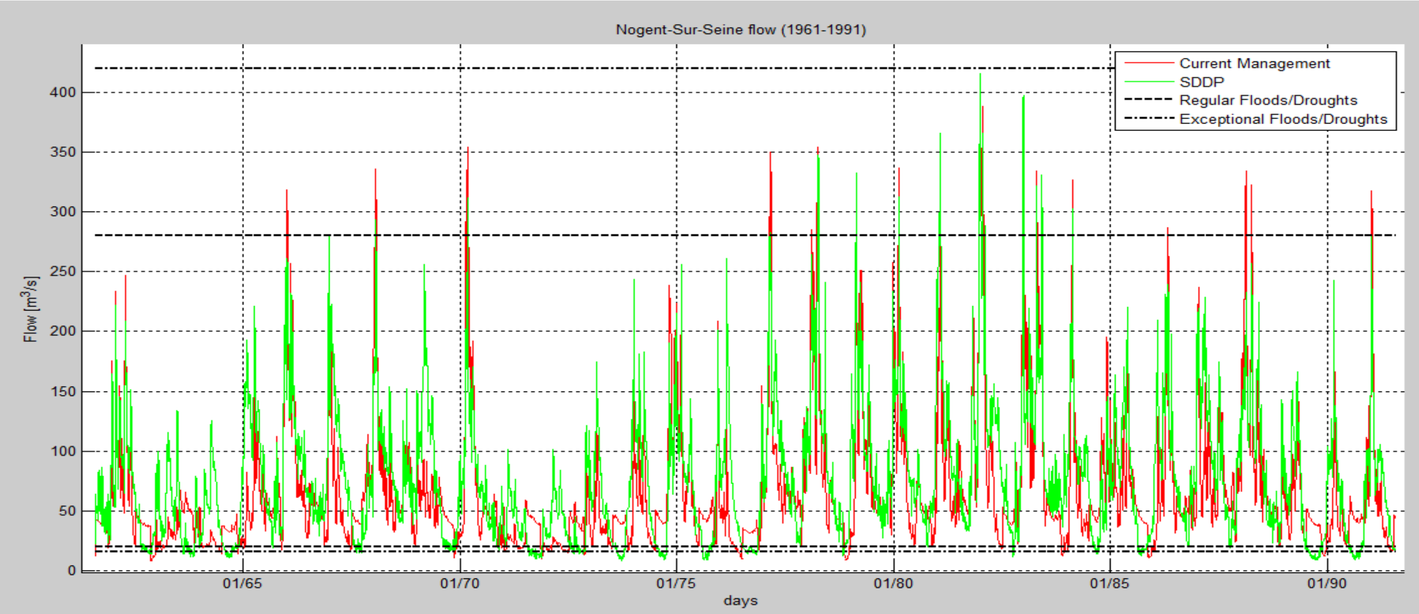


Figure 13: Seine River at Nogent-sur-Seine for current and SDDP management. In the proposed case, SDDP gives an overall better protection for floods, while current management seems to be more reliable for droughts.

We can observe that floods are generally better contained by SDDP technique than current management. Even still, peaks for 1982 and 1983 flood events are higher for optimized policy. This may be caused by the fact that SDDP algorithm is risk neutral, consequently it finds preferable to minimize these floods by acting on their duration rather than their peak. Figure below shows the detail of 1982 flood event:

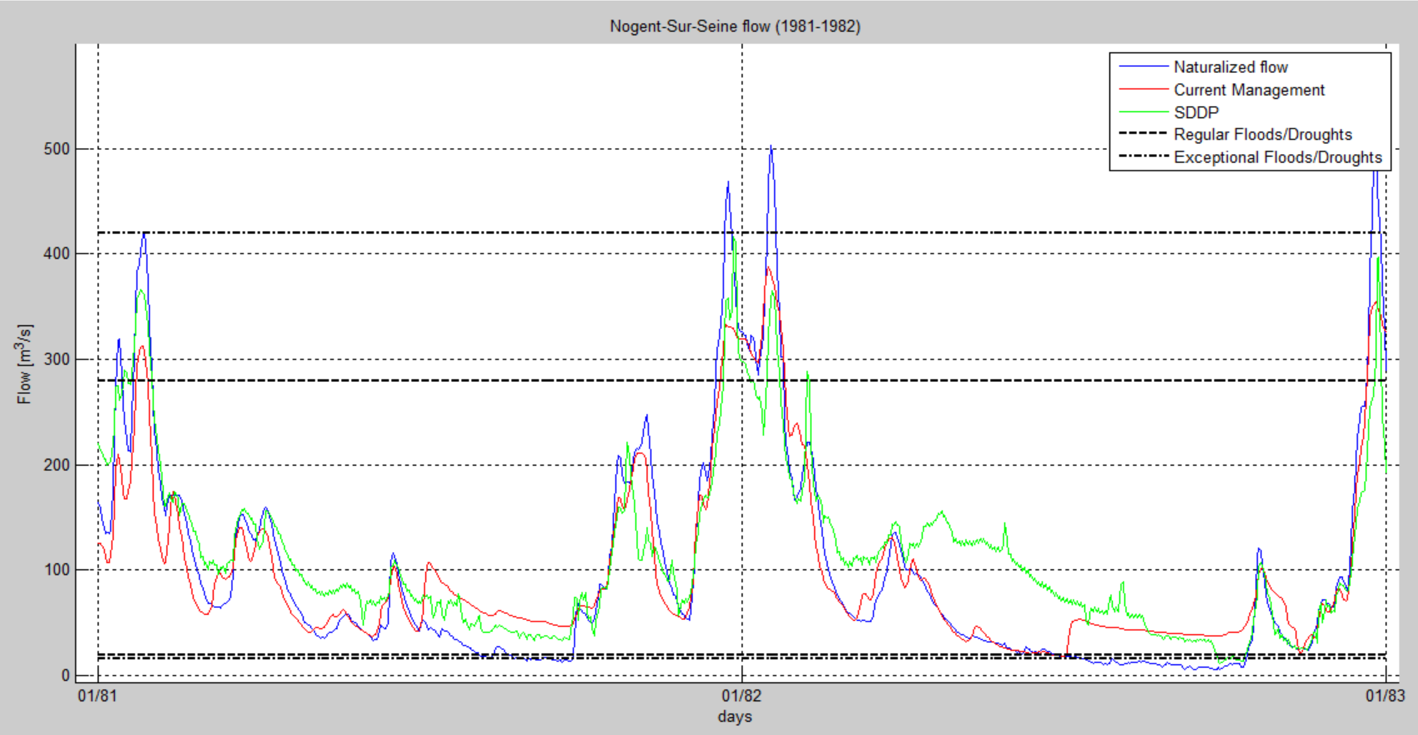


Figure 14: Seine River at Nogent-sur-Seine station, focus on 1982 flood event. The flood is contained by both current and SDDP management (exceptional flooding threshold is not attained). Flood duration for SDDP policy is lower than current management.

A better comparison between SDDP performances and current management may be given by plotting flows discharge-duration curve. This may be obtained by ranging time series in descendent order and assigning a cumulative frequency to events, here given by the following:

$$\forall t \in 1, \dots, H: \qquad fr(t) = \frac{t - 0.5}{H}$$

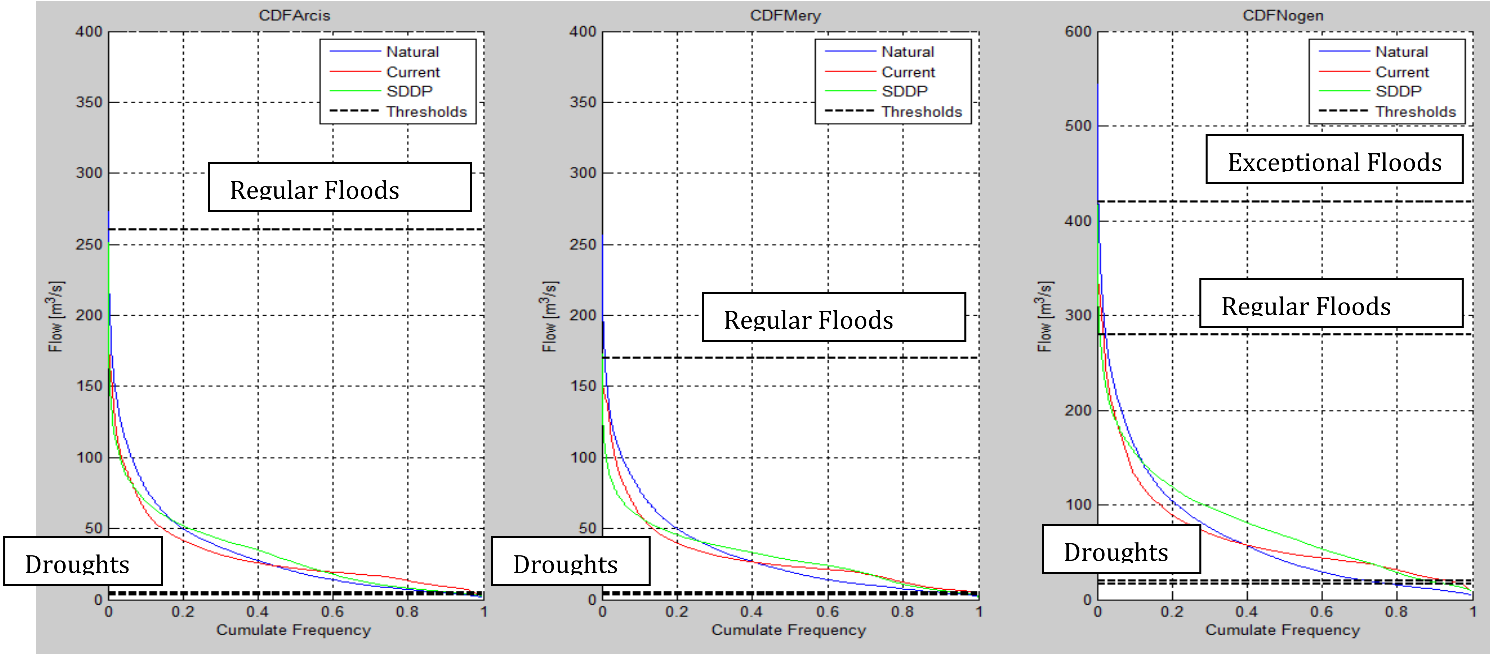


Figure 15: Flows discharge-duration curves for the three stations. It can be observed that Nogent-sur-Seine is the most critical station.

Next figures focus on discharge-duration curves at drought and flood thresholds:

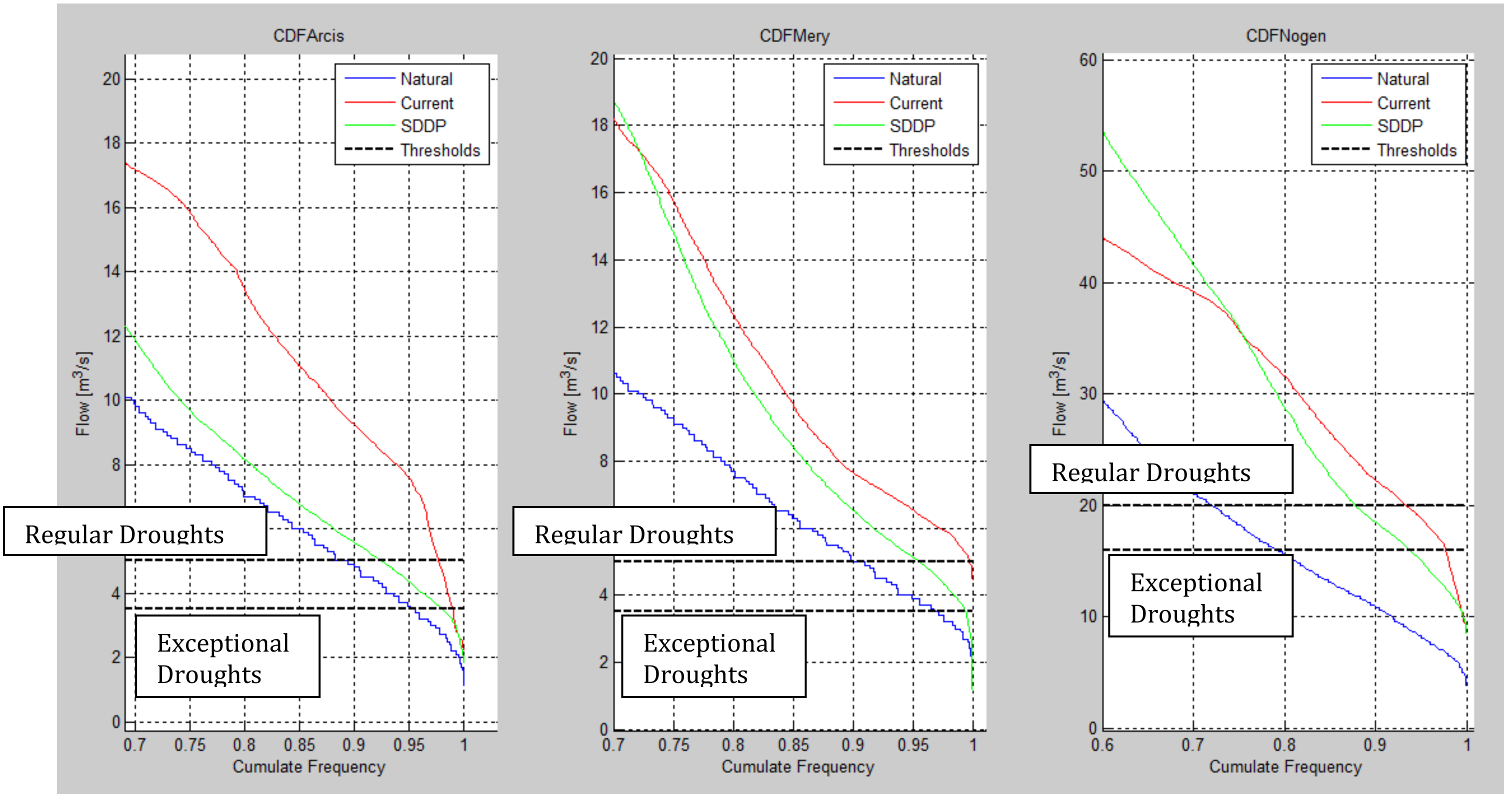


Figure 16: Flows discharge-duration curves for the three stations, focus on droughts. By giving priority to floods, SDDP algorithm minimizes droughts, but performs globally worse than current management.

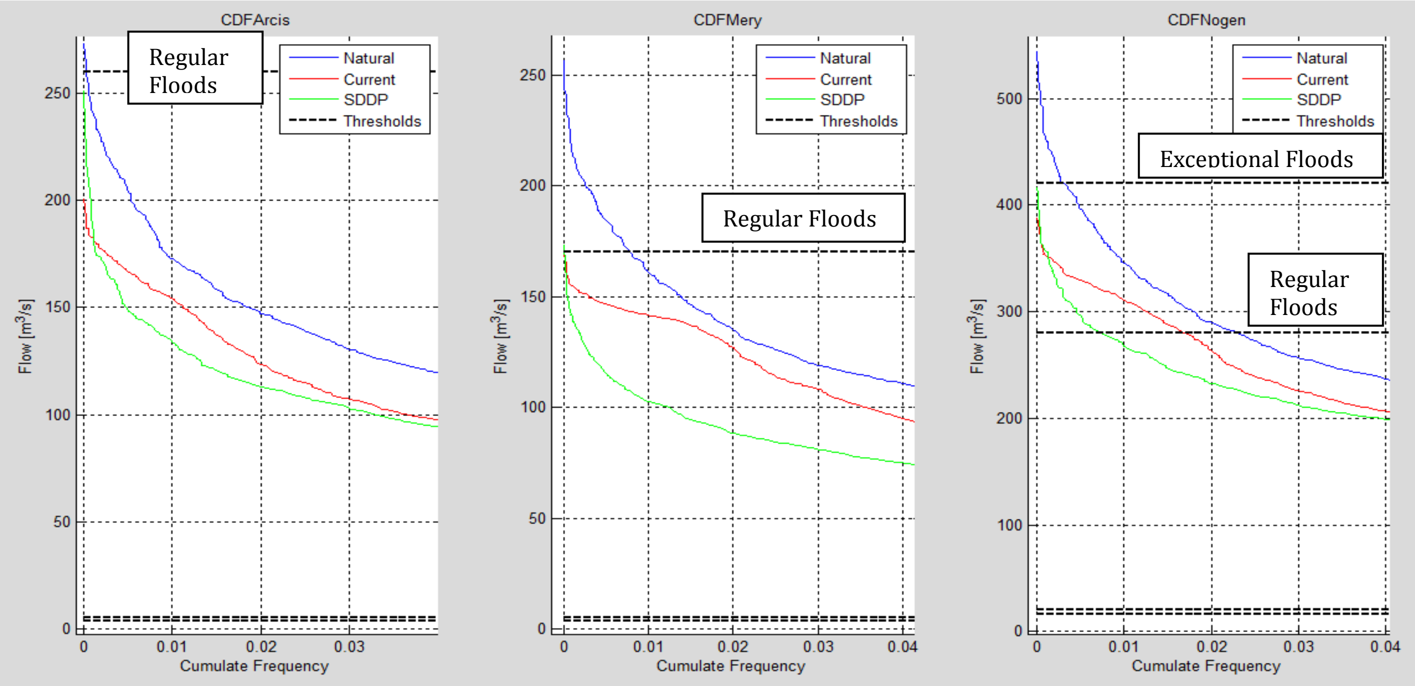


Figure 17: Flows discharge-duration curves for the three stations, focus on floods. Floods in Nogent-sur-Seine are minimized. Other stations do not have particular flood issues.

Graphics on discharge-duration curve convey very good information on intensity and probability for flood and drought events, which are related to system vulnerability and reliability. On the other hand, they do not provide any information on expected duration for floods and droughts, which is linked to system resilience. Furthermore, discharge-duration curve are interesting for the observation of system behavior, but do not provide a score to system performance that is better represented by risk assessment estimators.

Figures below show risk assessment indicators for natural system, current management and SDDP optimization for all main stations. Regular and exceptional floods estimators for Mery and Arcis have not been considered, as well as exceptional floods estimators for Nogent.

- Arcis-sur-Aube: regular and exceptional droughts

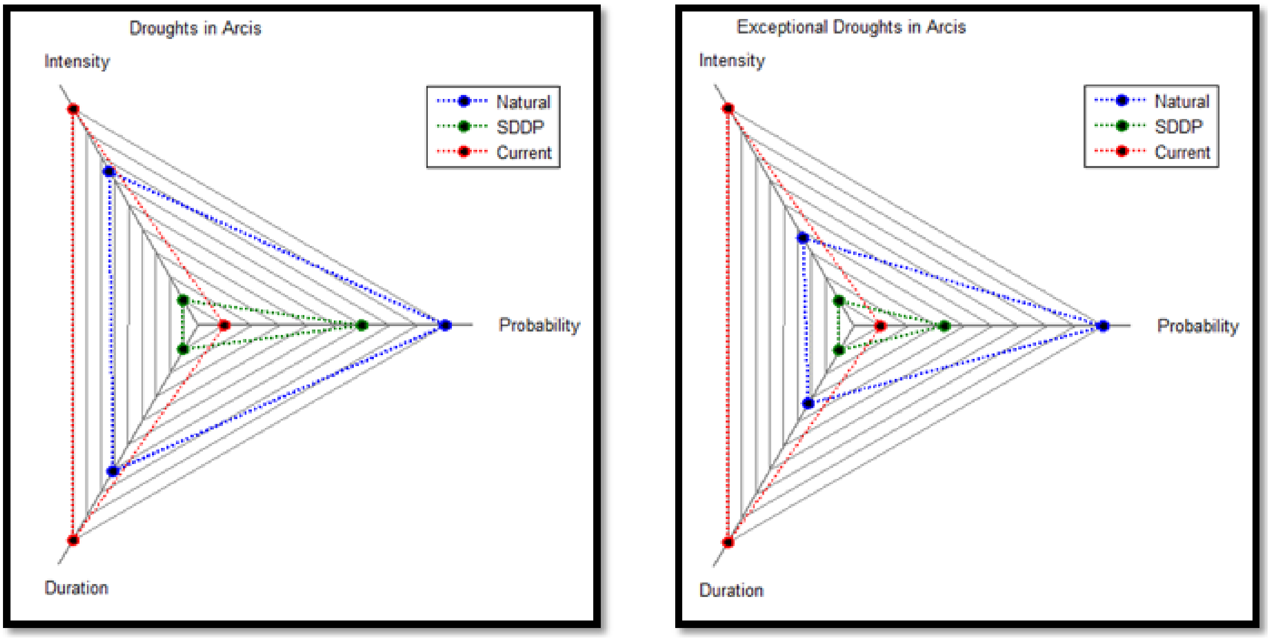


Figure 18: Estimators for droughts and exceptional droughts at Arcis-sur-Aube.

Droughts estimators for Arcis-sur-Aube show SDDP being better than current rules for intensity and duration indicators. This means that SDDP creates more drought events characterized by lower intensity than current management. We remark also that droughts expected duration and intensity is greater for current management than for naturalized flows.

- Mery-sur-Seine: regular and exceptional droughts

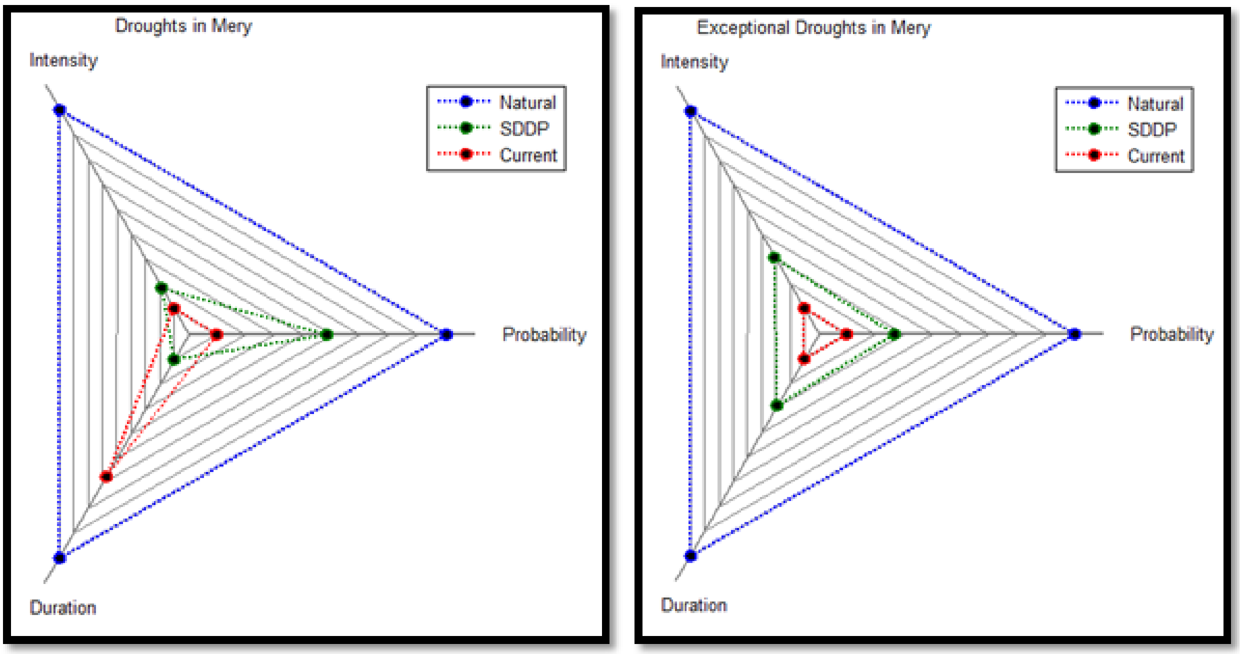


Figure 19: Estimators for droughts and exceptional droughts at Mery-sur-Seine.

Droughts estimators for Mery-sur-Seine show an overall better performance for current management, except to drought expected duration for regular events.

- Nogent-sur-Seine: regular floods; regular and exceptional droughts

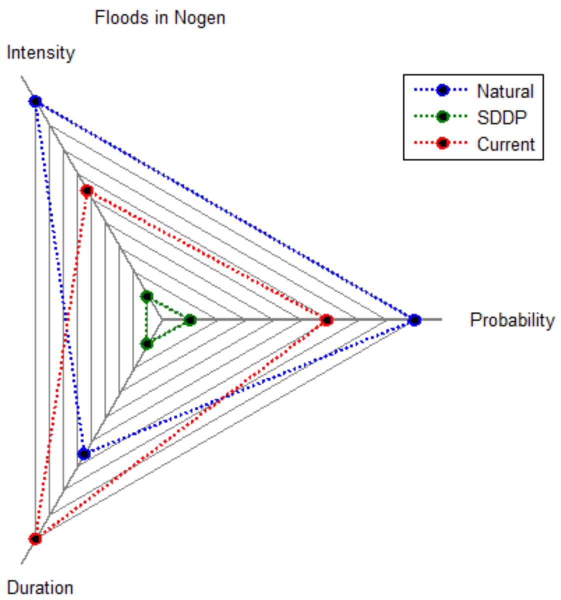
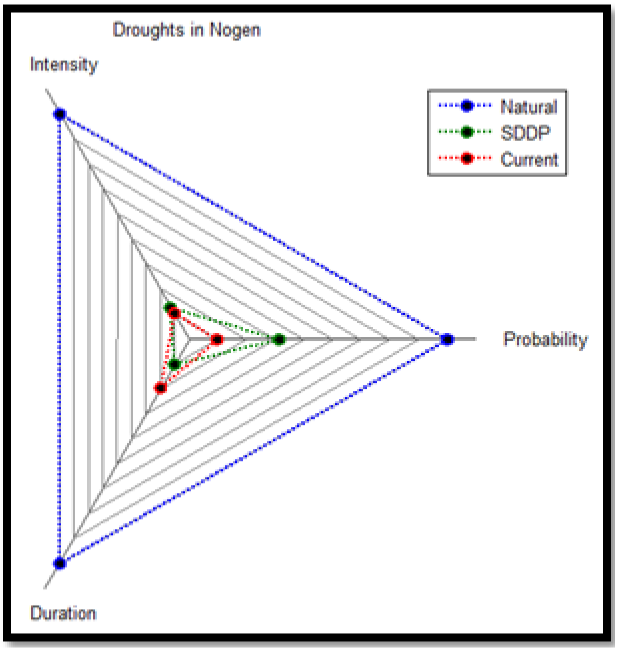


Figure 20: Estimators for floods at Nogent-sur-Seine.

Here, SDDP algorithm performs better for all estimators, confirming weights propensity to flood protection.



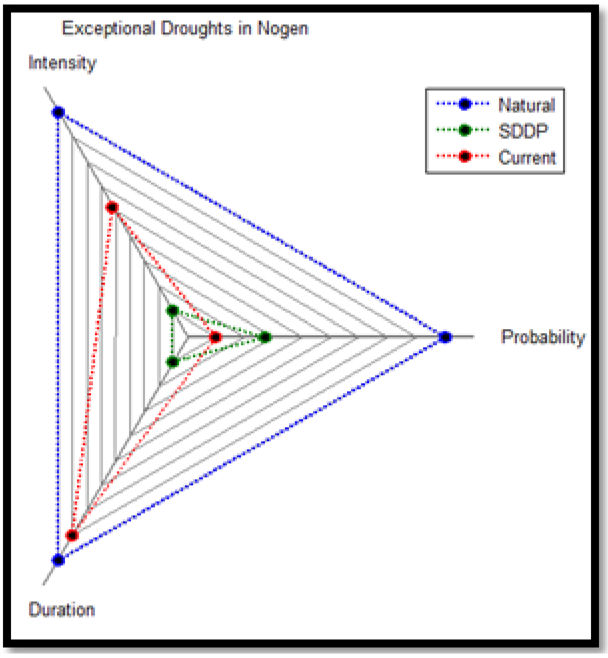


Figure 21: Estimators for droughts and exceptional droughts at Nogent-sur-Seine.

The global performance for drought protection at Nogent-sur-Seine station is similar to the one of Arcis-sur-Aube, with high droughts probability and low intensity and duration. However, for regular droughts SDDP performance is closer to the one given by current management.

5.2.2 Lakes life quality indexes

Lakes life quality indexes, resumed in Annex IV: Table of Impacts for 1961-1991 period, confirm the actual flood protection policy: they are more discharged on average, but on the other hand they are not affected by sensitive level variations.

6 Conclusions and perspectives

6.1 Final considerations and answers to initial research questions

SDDP solution strategy has been adopted in order to face flood and drought risk. The requirements of this algorithm, such as linearity and convexity, led to assume several hypothesis for modeling that were not always easy to validate: for instance, hydrological model's complexity is the result of a long calibration process. Moreover, a trade-off must found between the quality of model description and the total amount of computational effort, therefore, the choices on system definition depended also to the conflict between precision and dimension. Finally, issues on weighting objectives influence a lot the final solution, and weights have to be chosen in accurate combinations.

Answer to the initial research questions:

- How does centralized SDDP optimize the behavior of the system?

SDDP optimizes water system management by protecting main stations from flood and drought events. Moreover, it permits to avoid risk for all the ensemble of vulnerability, resilience and reliability indicators. Nonetheless, risk neutrality hypothesis conveys critical situations of flood and drought peaks which are not attained using the current policy.

Finally, the use of SDDP technique gives the solution for the current control problem, characterized by 20 state-variables, with a relatively small computational effort (from 10 hours of computational time, depending to the number of scenarios extractions). For that reason, SDDP represent an interesting tool for the evaluation of off-line water systems' management.

- Does SDDP perform better than current management?

For current weights combination, SDDP performs better in protecting the system from floods. Moreover, while current management avoids principally floods and droughts *probability*, SDDP optimizes the combination of *probability*, *expected intensity* and *expected duration* for floods and droughts. As a matter of fact, SDDP reduces the total costs as defined in the cost function, differently from present management that reduces droughts and floods locally for some objective points. Its reliability is thus given by reducing risk according to *stakeholders' interests*, and not to top-down planned objectives.

6.2 Future developments

There are three main directions to develop this project:

1. Improve SDDP algorithm by considering risk aversion (Shapiro A. W., 2013);
2. Refine hydrology, hydraulics and stocks models for the study case:
   - Statistic properties of disturbances can be developed in order to have a better representation of model scenarios;
   - Hydraulic propagation of flows may be affected by flood plains, which should be considered for the study case;
   - The hypothesis of perfect control on stocks' inflows can be relaxed by considering the effects of evaporation and lateral contributions.
3. Extend the project to the entire Seine River basin, adding Marne and Pannecière Reservoirs to the system and modeling the protection of Paris; moreover, a planning problem can be involved in the extended system by considering the actual project of a new flood plain, called "*la Bassée*":

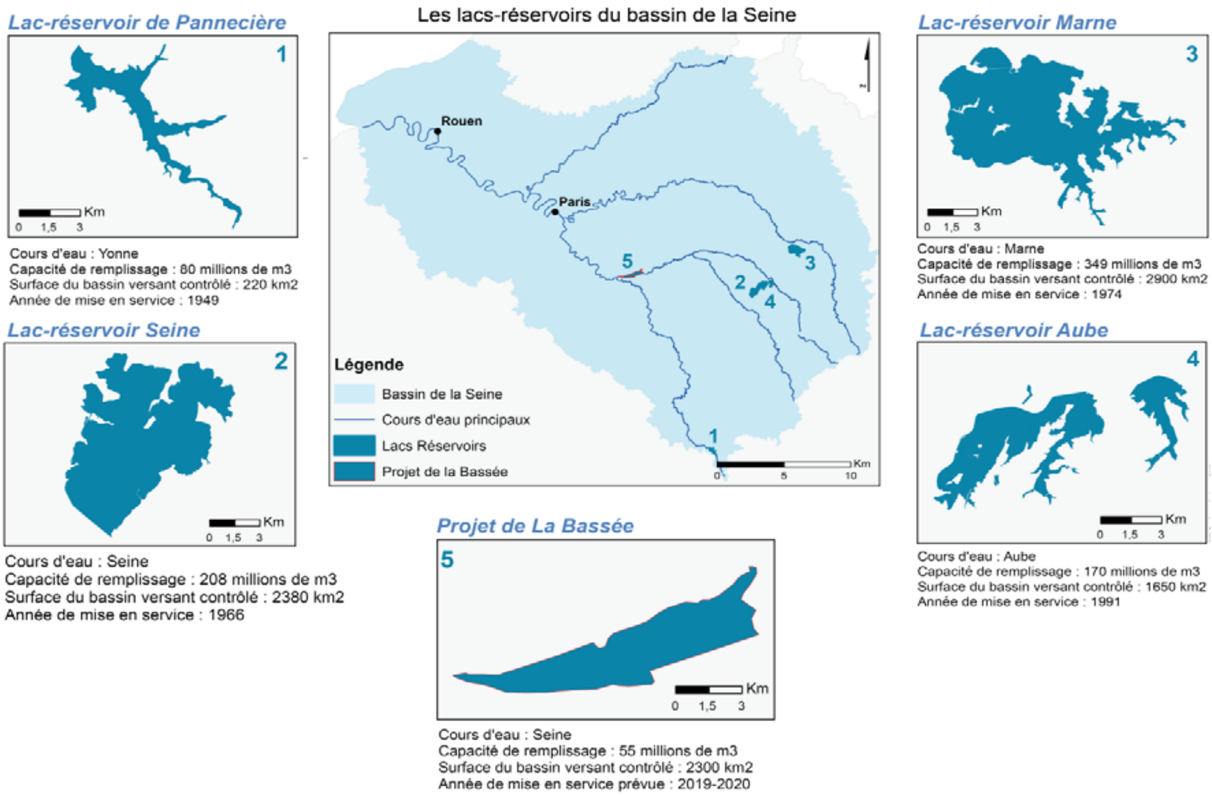


Figure 22: Main water stocks on Seine River basin: the fifth is the actual flood plain project for floods protection. (OECD, 2014)

Bibliography

Arrêté cadre sécheresse . (2012). *(N° 2012 094-0001)* .

Bartolini, P. S. (1988). Multivariate Periodic ARMA (1,1) Processes. *Water Resources Research* .

Bellmann, R. (1957). *Dynamic Programming.* Princeton, New Jersey: Princeton University.

de Matos, V. L. (2012). Improving the Performance of Stochastic Dual Dynamic Programming.

Dehay, F. (2012). *Etude de l'impact du changement climatique sur la gestion des lacs-réservoirs de la Seine.* Strasbourg, France: ENGEES.

Dorchies, D. T. (2013). Climate change impacts on multi-objective reservoir management: case study on the Seine River basin, France. *The International Journal of River Basin Management* .

Ducharne, A. B. (2007). Long term prospective of the Seine River system: confronting climatic and direct anthropogenic changes. *The Science of the total environment* .

Ficchi', A. D. (2013). *A centralized real-time controller for the reservoir's management on the Seine River using ensemble weather forecasting.* Master Thesis, Milano: Politecnico di Milano.

Ficchì, A. R.-O.-J.-A. (2015). Optimal operation of the multi-reservoir system in the Seine River basin using deterministic and ensemble forecasts. *Submitted to Journal of Water Resources Planning and Management* .

Fiering, M. B. (1982). Alternative indices of resilience. *Water Resources Research , 18* (1), 33-39.

Hashimoto, T. S. (1982). Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation. *Water Resources Research* .

Kjeldsen, T. R. (2004). Choice of reliability, resilience and vulnerability estimators for risk assessments of water resources systems. *Hydrological Sciences Journal* .

OECD. (2014). *Etude del'OCDE sur la gestion des risques d'inondation: la Seine en Ile-de-France 2014.*

Pereira, M. P. (1989). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming , 52*.

Salas, J. D. (1985). Approaches to multivariate modeling of water resources time series. *Water Resources Bulletin* .

Shapiro, A. (2010). Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research , 209*.

Shapiro, A. W. (2013). Risk neutral and risk averse Stochastic Dual Dynamic Programming method. *European Journal of Operational Research , 224*.

Soncini-Sessa, R. W. (2007). *Integrated and participatory water resources management-theory.* Amsterdam: Elsevier.

Tilmant, A. K. (2007). A stochastic approach to analyze trade-offs and risks associated with large-scale water resources systems. *Water Resources Research , 43* (W06425).

Annexes

Annex I: Rule Curves for Seine and Aube Lakes

- Aube Lake

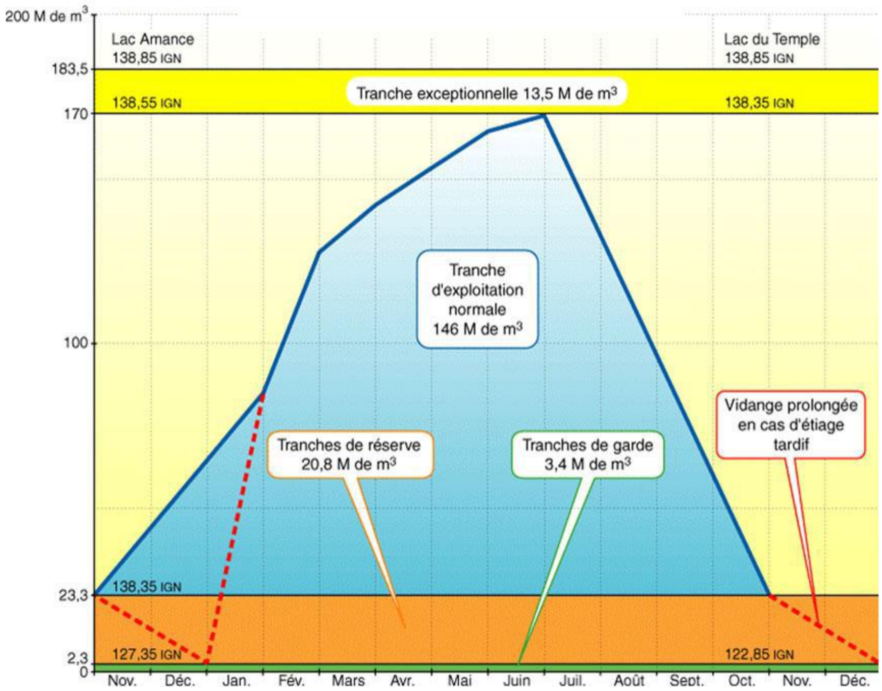


Figure 23: Aube Lake Filling Curve

- Seine Lake

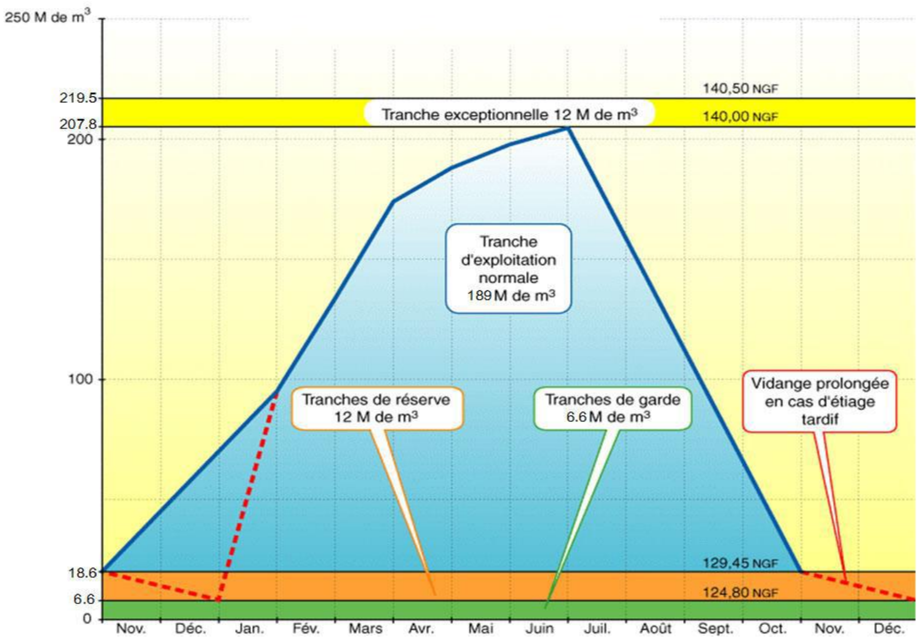


Figure 24: Seine Lake Filling Curve

Annex II: System control network and monitoring stations

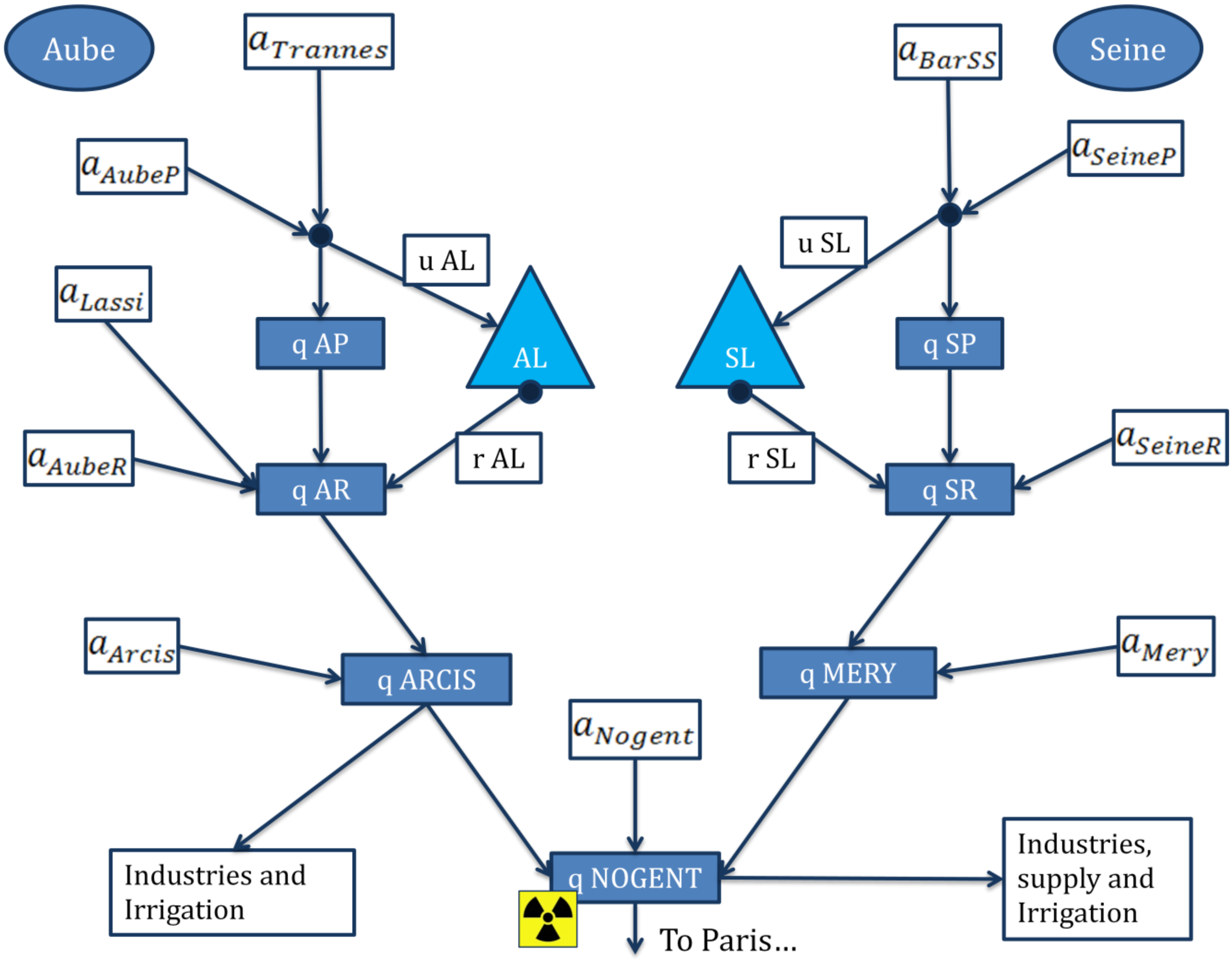


Figure 25: System control network

Symbol	Meaning	Interest
AP	Aube river after Aube Reservoir intake	Ensuring environmental flow for ecology
AL	Aube Lake	Ensuring lake's life quality
ARCIS	Aube river at Arcis-sur-Aube	Flood and drought protection
SP	Seine flow after Seine Reservoir intake	Ensuring environmental flow for ecology
SL	Seine Lake	Ensuring lake's life quality
MERY	Seine river at Mery-sur-Seine	Flood and drought protection
NOGENT	Seine river at Nogent-sur-Seine	Flood and drought protection

Table 5: System monitoring stations

Annex III: State transfer equations

- Hydrology

1. **Trannes Inflow**: Conditioned basin; CAR(1) model

$$a_t^{Tran} = \xi_t^{Tran}(\rho_\tau^{Tran} a_{t-1}^{Tran}) + \xi_t^{Tran}\left(\rho_\tau^{Tran|BSS} a_{t-1}^{BSS}\right) + \xi_t^{Tran}(\psi_\tau^{Tran})$$

2. **Lassicourt Inflow**: Conditioned basin; CAR(1) model

$$a_t^{Lassi} = \xi_t^{Lassi}\left(\rho_\tau^{Lassi|BSS} a_{t-1}^{BSS}\right) + \xi_t^{Lassi}\left(\rho_\tau^{Lassi} a_{t-1}^{Lassi}\right) + \xi_t^{Lassi}\left(\psi_\tau^{Lassi}\right)$$

3. **Arcis-sur-Aube Inflow**: Main basin ; CAR(1) model

$$a_t^{Arcis} = \xi_t^{Arcis}\left(\rho_\tau^{Arcis} a_{t-1}^{Arcis}\right) + \xi_t^{Arcis}\left(\psi_\tau^{Arcis}\right)$$

4. **Bar-sur-Seine Inflow**: Main basin CAR(1) model

$$a_t^{BSS} = \xi_t^{BSS}\left(\rho_\tau^{BSS} a_{t-1}^{BSS}\right) + \xi_t^{BSS}(\psi_\tau^{BSS})$$

5. **Mery-sur-Seine Inflow**: Conditioned basin ; CAR(1) model

$$a_t^{Mery} = \xi_t^{Mery}\left(\rho_\tau^{Mery} a_{t-1}^{Mery}\right) + \xi_t^{Mery}\left(\rho_\tau^{Mery|Arcis} a_{t-1}^{Arcis}\right) + \xi_t^{Mery}\left(\psi_\tau^{Mery}\right)$$

6. **Nogent-sur-Seine Inflow**: Conditioned basin CAR(1) model

$$a_t^{Nogent} = \xi_t^{Nogent}\left(\rho_\tau^{Nogent|Arcis} a_{t-1}^{Arcis}\right) + \xi_t^{Nogent}\left(\rho_\tau^{Nogent} a_{t-1}^{Nogent}\right) + \xi_t^{Nogent}\left(\psi_\tau^{Nogent}\right)$$

- Hydraulics (with enlarged states)

7. **Aube river after withdrawal for Aube Reservoir** :

$$q_t^{AP} = d_{AP|Tran} a_{t-1}^{Tran} + \left(1 - d_{AP|Tran}\right) a_t^{Tran} + \alpha^{AP|Arcis} a_t^{Arcis} - u_t^{AL}$$

8. **Confluence in Aube Release**:

$$q_t^{AR} = a_t^{Lassi} + \alpha^{AR|Arcis} a_t^{Arcis} + d_{AR|AP} q_{t-1}^{AP1} + \left(1 - d_{AR|AP}\right) q_t^{AP1} + r_t^{AL}$$

9. **Aube river at Arcis Sur Aube**:

$$q_t^{Arcis} = a_t^{Arcis} + d_{Arcis|AR} q_{t-1}^{AR} + \left(1 - d_{Arcis|AR}\right) q_t^{AR} - w_\tau^{Arcis}$$

10. **Seine River after withdrawal to Seine Reservoir** :

$$q_t^{SP} = d_{SP|BSS} a_{t-1}^{BSS} + (1 - d_{SP|BSS}) a_t^{BSS} + \alpha^{SP|Mery} a_t^{Mery} - u_t^{SL}$$

11. **Confluence in Seine Release**:

$$q_t^{SR} = \alpha^{SR|Mery} a_t^{Mery} + q_{t-1}^{SP} + r_t^{SL}$$

12. **Seine River at Mery Sur Seine**:

$$q_t^{Mery} = a_t^{Mery} + d_{Mery|SR} q_{t-1}^{SR1} + (1 - d_{Mery|SR}) q_t^{SR1}$$

13. **Seine and Aube Rivers confluence in Nogent Sur Seine**:

$$q_t^{Nog} = a_t^{Nog} + (1 - m_{Nog|Arc}) q_{t-1}^{Nog} + m_{Nog|Arc} \left[d_{Nog|Arc} q_{t-1}^{Arc2} + (1 - d_{Nog|Arc}) q_t^{Arc2} \right] + d_{Nog|Mery} q_{t-1}^{Mery1} + (1 - d_{Nog|Mery}) q_t^{Mery1} - w_\tau^{Nog}$$

- Hydraulic states enlargements:

14. $q_t^{AP1} = q_{t-1}^{AP}$

15. $q_t^{Arc1} = q_{t-1}^{Arc}$

16. $q_t^{Arc2} = q_{t-1}^{Arc1}$

17. $q_t^{SR1} = q_{t-1}^{SR}$

18. $q_t^{Mery1} = q_{t-1}^{Mery}$

- Reservoirs:

19. $s_t^{AL} = s_{t-1}^{AL} + u_t^{AL} - r_t^{AL}$ **Aube Lake**

20. $s_t^{SL} = s_{t-1}^{SL} + u_t^{SL} - r_t^{SL}$ **Seine Lake**

Annex IV: Table of Impacts for 1961-1991 period

Station	Criterion	Natural system (no management)			Current Management			SDDP Optimization		
		Return period	Expected Duration	Expected Intensity	Return period	Expected Duration	Expected Intensity	Return period	Expected Duration	Expected Intensity
ARCIS	Regular Floods	7,5 years	2 days	17,5 m³/s	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s
	Exceptional Floods	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s
	Regular Droughts	0,03 years	12,2 days	17,5 m³/s	0,12 years	16,9 days	24 m³/s	0,04 years	3,7 days	3,9 m³/s
	Exceptional Droughts	0,06 years	5,9 days	6,2 m³/s	0,27 years	16 days	13,9 m³/s	0,14 years	1,9 days	2,5 m³/s
MERY	Regular Floods	0,35 years	4,8 days	120,28 m³/s	>30 years	0 days	0 m³/s	30 years	1 day	3,1 m³/s
	Exceptional Floods	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s
	Regular Droughts	0,03 years	7,7 days	9, 5 m³/s	1,07 years	5,6 days	1,2 m³/s	0,06 years	2,5 days	2,1 m³/s
	Exceptional Droughts	0,09 years	4,8 days	4,1 m³/s	>30 years	0 days	0 m³/s	0,46 years	1,1 days	3,2 m³/s
NOGENT	Regular Floods	0,12 years	7,2 days	509,2 m³/s	0,16 years	9,2 days	352, 4 m³/s	0,37 years	4,5 days	168 m³/s
	Exceptional Floods	0,86 years	2,9 days	251,5 m³/s	>30 years	0 days	0 m³/s	>30 years	0 days	0 m³/s
	Regular Droughts	0,01 years	27, 4 days	202 m³/s	0,04 years	16,1 days	59,7 m³/s	0,02 years	14,6 days	64,3 m³/s
	Exceptional Droughts	0,01 years	19,3 days	104 m³/s	0,12 years	17,9 days	67,3 m³/s	0,04 years	8,2 days	26,7 m³/s

Table 6: Impacts concerned with risk. Cells in green represent best indicator values, cells in orange represent intermediary values, while cells in red represent indicator‘s worst performances.

Station	Criterion	Current Management		SDDP Optimization	
		Medium volume	Mean variation	Medium volume	Mean variation
AL	Lake's Life Quality	103,02 Mm^3	47,9 Mm^3	7,3 Mm^3	30,42 Mm^3
SL	Lake's Life Quality	114,92 Mm^3	62,43 Mm^3	7,12 Mm^3	29,71 Mm^3

Table 7: Impacts not concerned with risk. Cells in green represent indicator’s improvement, while cells in red represent indicator‘s worsening.