UCLouvain

epl

École polytechnique de Louvain

Study of a multi-agent framework for the control of nitrogen removal

for the *Seine aval* wastewater treatment plant

Author: **Dimitri Somers**
Supervisors: **Denis Dochain, Ionela Prodan**
Reader: **Raphaël Jungers**

Academic year 2020–2021
Master [120] in Electro-mechanical Engineering

I would like to thank first Denis Dochain and Ionela Prodan, who supervised my master thesis, for their help and their availability throughout the year.

I also express my gratitude to my family, friends and all the people who have made my years of study at Louvain-la-Neuve a success.

Finally, I would like to thank Florine for her continued support.

Contents

List of Figures

Nomenclature

Biological constants

$\mu_{i,max}$	maximum specific growth rate	$1/h$
K_i	saturation constant associated to i for the modeling of a growth rate	mg/L or g/m^3
X	concentration of bacteria	$1250\,mg/L$ or g/m^3
$Y_{i/j}$	yield coefficient of i into j	g/g

Biological parameters

μ	specific growth rate	$1/h$

System's constants

V_1	volume of the nitrification unit	$30,000\,m^3$
V_2	volume of the post-denitrification unit	$30,000\,m^3$

System's parameters

Q_0	flowrate of water (bypass) from v_{in} to y	m^3/h
Q_1	flowrate of water from x_1 to y	m^3/h
Q_2	flowrate of water from x_2 to y	m^3/h
Q_{in}	inlet flowrate of water from the previous stages of SAV to v_{in}	m^3/h
Q_{seine}	flowrate of water from the Seine river to y	m^3/h
u_1	flowrate of water from v_{in} to x_1	m^3/h
u_2	flowrate of water from x_1 to x_2	m^3/h
u_3	flowrate of water from v_{in} to x_2	m^3/h
u_R	flowrate of water from x_2 to x_1	m^3/h

Other Symbols

e_i	normalised error of agent i i.e., $\frac{state_i - goal\ state_i}{goal\ state_i}$	/

Acronyms

EQS Environmental Quality Standards.

MAS Multi-Agent System.

SAV The Seine aval WWTP.

SIAAP Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne. Greater Paris Sanitation Authority.

UWWTD The Urban Waste Water Treatment Directivee 91/271/EEC.

WFD The Water Framework Directive 2000/60/EC.

WWTP Wastewater Treatment Plant.

Glossary

NH_4^+ ammonium. Sometimes noted in this document as NH4 or referred as ammonia although not technically correct.

NO_2^- nitrite. Often referred as NO2.

NO_3^- nitrate. Often referred as NO3.

Appendix the file attached to this thesis, `Somers_24721400_2021_Appendix.zip`.

Poissy measurement site of the Seine river after SAV.

receiving water the water receiving the WWTP discharge water. It is often a river.

Sartrouville measurement site of the Seine river before SAV.

Introduction

Wastewater treatment is essential, especially in densely populated areas. The water polluted by domestic or industrial usage cannot be returned untreated to the environment without consequences. Indeed, untreated or partially treated wastewater can lead to public health and environmental issues.

In Europe, the use of large Wastewater Treatment Plant (WWTP) and their regulation allow to tackle severe public health issues such as access to safe drinking water. Worldwide, it is estimated that 525 000 children under the age of five die each year from diarrhoea, mostly due to the drinking of untreated water [28]. As human excreta are implicated in the transmission of many diseases, good sanitation of the wastewater is primordial to tackle the spread of these diseases [9]. Nowadays WWTP can be used to monitor the Covid-19; indeed it is possible to detect a cluster of Covid-19 through the measurement of the viral load at the local WWTP [17]. On a lighter note, efficient wastewater treatment is needed to allow the swimming into city canals (that receive the water discharged by the WWTP) which is one of the objectives in Paris for the next Olympic Games in 2024 [27].

The environmental issues stem mainly from the high level of nutrients present in the wastewater. Indeed, nitrogen and phosphorus promote excessive growth of algae that chokes other life. This process, called eutrophication, quickly degrades the ecological state of the water body. The degradation includes a decreased biodiversity, toxicity effects, changes in fish species... and can lead to so-called dead zones where the level of dissolved oxygen is too low for fishes (among other species) to survive. The nitrogen removal from wastewater is costly and challenging, thus it is often the bottleneck to tackle the eutrophication.

Wastewater treatment in densely populated is a difficult task because the nitrogen concentrations in the wastewater are very high and the nitrogen removal process is poorly characterised. The process has been tuned over the years, in regular conditions, in order to meet the concentration limits fixed by the legislation. But the control of the nitrogen removal is in general based on heuristics and cannot deal with high variations in the input. And high input variations are not a rare thing in wastewater treatment. Indeed the flow rates of the wastewater can dramatically change: they can triple. Moreover the concentrations inside the inlet flow rate can also easily vary from simple to double. Wide variations at the input of the WWTP can possibly happen at every heavy rainfall. And they are strengthened in the dry season when the regular flow rates are lower.

The thesis aims to tackle the problem of nitrogen removal when the input is varying outside of the regular operation area. In particular, the thesis objective is to propose a control strategy to reduce the excess of nitrogen concentrations, in regards to the European legislation, during heavy rainfall.For this, the Multi-Agent System (MAS) framework is investigated on the The Seine aval WWTP (SAV), operated by the Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne (SIAAP i.e. the Greater Paris Sanitation Authority in English). Seine aval is the largest wastewater treatment plant in Europe, treating 70% of the wastewater from the Paris agglomeration (treatment capacity of 2.3 million cubic meters per day).

Multi-Agent System (MAS) is a growing field of research which studies the achievement of a global objective via the interactions of local agents. The agents and their interactions form a network that can be studied with graph theory. The MAS field has its own control approaches. For instance the *consensus algorithms* where all the agents converge to a single state although they interact only with their neighbour.

The MAS framework has already been used in many areas such as the modeling of the flocking behaviour of birds, the optimisation of bus schedules or for sensor fusion. In this thesis, the nitrogen removal section of SAV is modeled into a network of agents (where an agent represents a unit of the nitrogen removal system for instance) and their interactions (the communication of data such as the state of the water).

The master thesis is organised as follow. In the first section the state of the art for wastewater treatment is presented, including the modeling of nitrogen removal and the review of the legislation regarding wastewater treatment. Seine aval is described and parameters useful for the thesis are extracted from the literature. The multi-agent systems tools and notations used in the thesis are also reviewed. In the second section the case study is defined. The implementation of the simulation in MATLAB is then detailed. In the third section the heavy rainfall scenario is characterised and used in the simulation. A benchmark is proposed to evaluate the performance of various control approaches. In particular, the impact of various configurations of the interactions between the agents is discussed. In the last section the design of the MAS chosen in the thesis is discussed for future work.

1 State of the art

In this section, we give some necessary context about wastewater treatment and introduce basic tools of the multi-agent framework. In subsection 1.1, we present the mathematical model used for the nitrogen removal and describe the regulatory background. In subsection 1.2, the Seine aval Wastewater Treatment Plant (WWTP) specificity and previous work around nitrogen removal for SAV are summarised. The characterisation of useful parameters for the rest of the thesis is also given. Finally, subsection 1.3 outlines preliminaries for Multi-Agent System (MAS) such as graph theory. The proportional control law used further in the thesis is then presented.

1.1 Wastewater treatment

The wastewater is defined as water and solids that flow to a WWTP. The water is then discharged (hopefully once safely treated) in receiving water (typically a river), or reused locally. For Paris, wastewater refers to the used water from its citizens and industries and to the rainwater since it historically has a unitary network for all the water flows. In general, the treatment of wastewater comes in different size and shapes, depending on criterion such as the volume of water to be treated, its quality, the vulnerability of the receiving water... We focus here on large plants used for densely populated agglomerations. They have in general four main steps: a pre-treatment serves to remove large solids, oil and grease. A primary treatment allows suspended solids to settle at the bottom of the tank, called a clarifier. A secondary treatment partially removes organic material (carbon), nitrogen and phosphorus. For this it can use diverse methods, typically activated sludge and more recently membrane reactors. A secondary clarifier is also usually required. A tertiary treatment is more advanced and not always required. It is targeted to some materials or pollutants to meet specific needs. Wastewater treatment also produces sludge that is recycled from the secondary clarifier to the secondary treatment units but that has to be removed to avoid its accumulation in the plant and treated appropriately.

In regular conditions wastewater treatment performs well, as expected. However, the quality of the discharged water often dramatically falls during heavy rains. Indeed, although ammonium level decrease in the inflow, the amount of water coming at the WWTP may be larger than its treatment capacity. This means first that the treatment is degraded, as the WWTP has less time to perform the biological treatment of the water for instance.Besides, some inflow must sometimes bypass the WWTP and is thus directly released into the receiving water. Heavy rains can also lead to nitrate leaching in rural area, thus dramatically increasing the nitrate concentrations in the wastewater. The issue can be worsened by the receiving water conditions. For example the Seine river in Paris naturally produces some nitrite, and this production is increased when more ammonium is discharged in it. Also the flow of discharged water can sometimes be in the same order of magnitude as the flow of the river, amplifying the problem since there is less dilution of the discharged water [34]. This can appear when the river flow is low and the WWTP is in a highly impermeable area such as dense urban spaces. The water is then streaming quickly into the WWTP while the peak flow of the river is reached several (dozen of) hours later. The lag time depends on the basin characteristics and of the rain event. This lag time is often characterised in a storm hydrograph [31].

We now focus on the biological process known as *activated sludge*, and more specifically the biological nitrogen removal.

1.1.1 Activated sludge for nitrogen removal

The activated sludge corresponds to a selected suspended bacterial biomass, leading to the removal of specific organic matter. This process is known and applied since a century. The reference to describe it mathematically are the Activated Sludge Models (ASM), starting with ASM1 in 1987 [18]. They have shown good practical results and are widely used in WWTP [39].

The nitrogen removal allows to transform the ammonium (NH_4^+, typically coming from urea) into nitrogen gas (N_2). It is based on two different reactions : nitrification and denitrification.

1.1.1.1 Nitrification

The simplified reaction comes in two steps : nitritation (Equation 1) (NH_4^+ into NO_2^- by the *Nitrosomonas* bacteria) and nitratation (Equation 2) (NO_2^- into NO_3^- by the *Nitrobacter* bacteria)

$$2\,NH_4^+ + 3\,O_2 + 2\,H_2O \longrightarrow 2\,NO_2^- + 4\,H_3O^+ \tag{1}$$
$$2\,NO_2^- + O_2 \longrightarrow 2\,NO_3^- \tag{2}$$

Taken from [2], the standard dynamical model presented below in Equation 3 is in the form of a general biotechnological process [4] with no outflow in gaseous form (we are only interested in the dynamics of ammonium, nitrite and nitrate) and no transport of the biomass (variations in microbial biomass concentrations are negligible compared to other compounds). It neglects pH and temperature (i.e. we assume a good control of these parameters). Also, only one equivalent tank is considered for the treatment unit.

$$\frac{\mathrm{d}NH_4}{\mathrm{d}t} = \frac{Q_{in}}{V}(NH_{4,in} - NH_4) - \mu_1 X_1 \tag{3}$$
$$\frac{\mathrm{d}NO_2}{\mathrm{d}t} = \frac{Q_{in}}{V}(NO_{2,in} - NO_2) + \frac{\mu_1 X_1}{Y_{NH_4/NO_2}} - \mu_2 X_2 \tag{4}$$
$$\frac{\mathrm{d}NO_3}{\mathrm{d}t} = \frac{Q_{in}}{V}(NO_{3,in} - NO_3) + \frac{\mu_2 X_2}{Y_{NO_2/NO_3}} \tag{5}$$

With Q_{in} the inflow in $[m^3/h]$, V the volume of the tank in $[m^3]$. X_1, X_2, μ_1 and μ_2 are the concentration of *Nitrosomonas* and *Nitrobacter* bacteria in $[g/m^3]$ and their specific growth rates, respectively. Y_{NH_4/NO_2} and Y_{NO_2/NO_3} are the yield coefficient of ammonium into nitrite, and nitrite into nitrate, respectively. The specific growth rates are modeled by double Monod[1]equations as in Equation 6. Note that the dissolved oxygen influence will be omitted in the context of this document.

$$\begin{aligned}\mu_1 &= \mu_{1,max}\frac{NH_4}{K_{NH_4} + NH_4}\frac{O_2}{K_{O_2,1} + O_2}\\ \mu_2 &= \mu_{2,max}\frac{NO_2}{K_{NO_2} + NO_2}\frac{O_2}{K_{O_2,2} + O_2}\end{aligned} \tag{6}$$

[1]A empirical model for the growth rate, commonly used in ASM. See [4].

With $\mu_{1,max}$ and $\mu_{2,max}$ the maximum specific growth rates $[1/h]$ of each reaction. K_i are the saturation constants for i, in $[g/m^3]$.

In nominal conditions, the nitrification can be further approximated as a single reaction transforming ammonium into nitrate. However the formation of nitrous oxide (N_2O) is of rising concern due to his powerful greenhouse effect making for the majority of the equivalent carbon footprint of a WWTP [14].

1.1.1.2 Denitrification

Denitrification can be used before or after the nitrification process (this is denoted by using *pre-* or *post-* prefix), with a possible recirculation of the water between the units. This process reduces nitrate (NO_3^-) to dinitrogen (N_2) thanks to heterotrophic bacteria. When the reactant supply is not carefully managed, more nitrite is produced. This is often the case because the quantities of reactants used are often poorly managed [35].

The simplified reaction is : $5\,CH_2O + 4\,NO_3^- + 4\,H_3O^+ \longrightarrow 2\,N_2 + 5\,CO_2 + 11\,H_2O$

As for the nitrification, the denitrification equations come from [2]. The standard dynamical model given in Equation 7 is in the form of a general biotechnological process [4] with no outflow in gaseous form (we are only interested in the dynamics of ammonium, nitrite and nitrate) and no transport of the biomass (variations in microbial biomass concentrations are negligible compared to other compounds). It assumes a depleted medium in oxygen and sufficient supply of carbon and phosphorus (i.e. we assume a good control of these parameters). Also only one equivalent tank is considered for the treatment unit.

$$\frac{\mathrm{d}NH_4}{\mathrm{d}t} = \frac{Q_{in}}{V}\left(NH_{4,in} - NH_4\right) \tag{7}$$

$$\frac{\mathrm{d}NO_2}{\mathrm{d}t} = \frac{Q_{in}}{V}\left(NO_{2,in} - NO_2\right) + \frac{\mu_1 X}{Y_{NO_3/NO_2}} - \mu_2 X \tag{8}$$

$$\frac{\mathrm{d}NO_3}{\mathrm{d}t} = \frac{Q_{in}}{V}\left(NO_{3,in} - NO_3\right) - \mu_1 X \tag{9}$$

With Q_{in} the inflow in $[m^3/h]$, V the volume of the tank in $[m^3]$. X, μ_1 and μ_2 are the concentration of the bacteria in $[g/m^3]$ and the specific growth rates for nitrate and nitrite, respectively. Y_{NO_3/NO_2} is the yield coefficient of nitrate into nitrite. The specific growth rates are modeled by double Monod equations as in Equation 10. Note that the Chemical Oxygen Demand[2](COD) influence will be omitted in the context of this document.

$$\begin{aligned} \mu_1 &= \mu_{1,max}\frac{NO_3}{K_{NO_3}+NO_3}\frac{COD}{K_{COD,1}+COD} \\ \mu_2 &= \mu_{2,max}\frac{NO_2}{K_{NO_2}+NO_2}\frac{COD}{K_{COD,2}+COD} \end{aligned} \tag{10}$$

With $\mu_{1,max}$ and $\mu_{2,max}$ the maximum specific growth rates $[1/h]$ of each reaction. K_i are the saturation constants for i, in $[g/m^3]$.

Since denitrification is an anaerobic process whereas the nitrification needs an oxygen supply, the two reactions are often placed in different units. In some WWTP's, they are done successively in the same tank by alternating anaerobic and aerobic conditions.

[2] That is the mass of oxygen consumed by a volume of a solution. It relates to the evolution of organic material in wastewater.

1.1.2 Water framework directive

Wastewater treatment is an important public health and environmental issue. In Europe, the concern is nowadays focused on the environmental impact of the discharge of WWTP's into receiving water.

Early European water legislation began with setting binding quality target for the drinking water in late seventies. It also included quality objective on fish water, bathing waters and groundwater through the Dangerous Substances Directive. The second wave of European water legislation has been led by two important directives in 1991 : the Urban Waste Water Treatment Directive (UWWTD) and the Nitrates Directive. The first is directly addressed to WWTP, giving minimal standards on the quality of discharged waters. It requires a proper treatment of wastewater to reduce both risks to human health and damages to freshwater. Indeed, nitrogen and phosphorus promote excessive growth of algae that chokes other life (eutrophication). The second directive targets the water pollution by nitrates from agriculture. Both have been revised through the years, and they are still in force today [12] [13]. The requirements are some concentration limits and a minimum percentage of reduction of the incoming nitrogen and phosphorus material at the discharge. Also, the global volume of untreated water cannot exceed 5% of the urban wastewater volume in dry conditions. They are eventually met by SAV and are reported on the SIAAP website.

In 2000, The Water Framework Directive 2000/60/EC (WFD) has been introduced. From several fragmented directives, the goal was to give a global framework for a good management of the water in general. Water management in now based on river basins, and aims at achieving *good status* for all the water. For WWTP's, this materialise into new (more stringent) limits, but they apply on the receiving water instead of the discharged water. That is, at measuring stations downstream of the plant. The Environmental Quality Standards (EQS) for nitrogen material are given in Table 1. They are transposed in a French order in 2010 [21]. When referring to EQS in the following sections, these *good status* values are considered.

	high	good	moderate	poor
NH_4^+ $[mgNH_4^+/L]$	0.1	0.5	2	5
NO_2^- $[mgNO_2^-/L]$	0.1	0.3	0.5	1
NO_3^- $[mgNO_3^-/L]$	10	50	/	/

Table 1: Environmental Quality Standards (EQS) for nitrogen material from the WFD. A bar (/) means there is no agreement on the upper limit. The *good* limits (in green) are the ones to achieve

Note that these values can be exceeded locally at the discharge site, in a determined *mixing zone*. The status is calculated from the 90th percentile over two years.

Since then, some local orders have been issued (see [36]). Note that the European UWWTD is currently is process of revision : it is open to public consultation until 21 July 2021 and the Commission adoption is planned for the first quarter of 2022.

1.2 Seine aval

The SAV Wastewater Treatment Plant is managed by the SIAAP. It is by far the largest of the six WWTP's of the network covering four French departments of Paris. It treats 70% of the wastewater from the Paris agglomeration, or a population equivalent of 5 millions. Its nominal inlet flowrate is 1.5 million cubic meter per day and its treatment capacity is 2.3 million $[m^3/d]$ (26.6 $[m^3/s]$). It can accommodate larger flowrates, up to 70 $[m^3/s]$, but some water will not be treated [36]. SAV is in an upgrading process since 2013 (and scheduled until 2022) in order to meet regulatory requirements such as the The Water Framework Directive 2000/60/EC.

The treatment of the ammonia has been significantly improved in 2007 with the introduction of the nitrification and post-denitrification units. However the nitrite concentrations in the Seine river rose at the same time [35]. Since then the size of the operating units have been increased (2011) and a pre-denitrification unit has been implemented (2017) to reduce the nitrite levels. A membrane bioreactor line has also been added in parallel to the tree existing ones. This new line, based on a different technology, can take up to one fifth of the incoming water in nominal conditions and shows better performance [37].

1.2.1 Description

The nitrogen removal as of 2013 is presented in Figure 1. We can see the supply of oxygen (through air) and methanol, for the nitrification unit and the post-denitrification unit respectively. In case of rainfall (*Temps de pluie* in the schema), the water can bypass the two units and goes directly to the discharge (*Rejet* in French). The global scheme with the pre-treatment and the removal of carbon and phosphorus is given in Appendix A (in French).

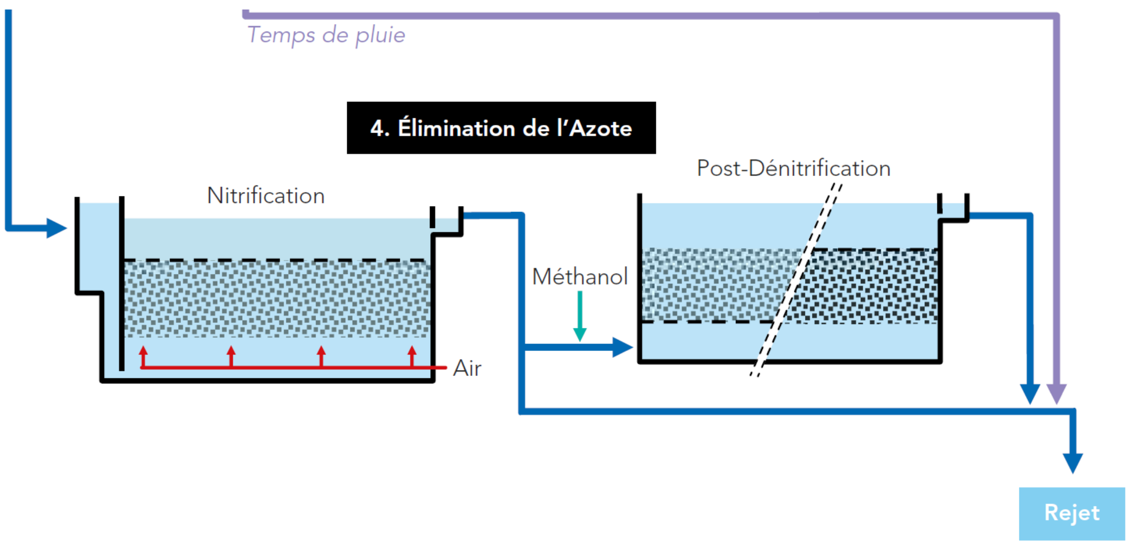


Figure 1: Nitrogen removal stage of SAV, before 2017 (without the pre-denitrification unit)

1.2.2 Previous work

The research program MOCOPEE (MOdélisation, Contrôle et Optimisation des Procédés d'Epuration des Eaux) has been launched in 2014. It aims to be a long-lasting exchange space between scientists, stakeholders and industries in the field of urban wastewater treatment. The programme deals with issues relating to metrology applied to water treatment, the modeling of the operation of treatment processes, their control, and the emergence of innovative concepts in the field of wastewater treatment [24]. This program produced many of the results used in this thesis.

Among the results, there is first the thesis on detailed modelisation of SAV [41], at the heart of the case study in section 2. There is also the study of the nitrite issue in the Paris agglomeration [34] [35], for other intermediate products of the nitrification and better sensing techniques [23] [14]. Finally, there are the works about the management of the different WWTP's as an integrated network [33] [5] [2], that include a simple model of SAV [32].

1.2.3 Parameters description

The parameters of the dynamical model for the ammonium, nitrite and nitrate in the nitrification and denitrification units (Equation 2 to Equation 10) are taken from [2]. They are summarised in Table 2 and Table 3.

Parameter	Value	Unit
$\mu_{1,max}$	$6.96\,10^{-2}$	$gNH_4/gX/h$
$\mu_{2,max}$	$3.67\,10^{-2}$	$gNO_2/gX/h$
K_{NH_4}	0.01	gNH_4/m^3
K_{NO_2}	0.01	gNO_2/m^3
$K_{O_{2,1}}$	$1.50\,10^{-4}$	gO_2/m^3
$K_{O_{2,2}}$	$4.42\,10^{-13}$	gO_2/m^3
Y_{NH_4/NO_2}	6.87	gNH_4/gNO_2
Y_{NO_2/NO_3}	0.14	gNO_2/gNO_3

Table 2: Calibrated parameters for the nitrification model, from [2]. The saturation constants for oxygen are negligible (supplied in excess by the aeration of the unit)

Parameter	Value	Unit
$\mu_{1,max}$	24.04	$gNO_3/gX/h$
$\mu_{2,max}$	$1.29\,10^{-2}$	$gNO_2/gX/h$
K_{NO_3}	0.10	gNO_3/m^3
K_{NO_2}	$4.42\,10^{-8}$	gNO_2/m^3
K_{COD_1}	70.29	$gCOD/m^3$
K_{COD_2}	$1.00\,10^{-3}$	$gCOD/m^3$
Y_{NO_3/NO_2}	27.57	gNO_3/gNO_2

Table 3: Calibrated parameters for the post-denitrification model, from [2]

The dataset for SAV measurements and measurements in the Seine river before and after the plant (at *Sartrouville* and *Poissy* respectively) is the same as in the paper modeling SAV [32]. It is given in the Appendix file (`Somers_24721400_2021_Appendix.zip`) attached to this thesis.

1.3 Multi-agent systems

The study of Multi-Agent System (MAS) can be traced back to more than 50 years ago. In particular, the work of C.W. Reynolds in 1987 shed the light on MAS with the simulation of a very large number of *boids* (bird-oid, or a bird like object) that could mimic the flocking behaviour of birds [30]. However, MAS emerged as a separate research field only in the last fifteen years, with notably the influential article from Olfati-Saber dating back to 2004 [26]. MAS are theorized and applied in many fields, from biology or psychology to computer science and provide insights leading to new developments in different disciplines [10].

A MAS can be defined as *a system composed of two or more* ***agents*** *which work together in an environment to achieve a* ***global task/objective*** *via* ***local interactions*** [10]. An agent can be a very different object depending on the context. However, it generally has some autonomy, is able to achieve simple computation, can communicate with other agents and perceive its environment. An agent can be a vehicule, a sensor, a computer process, a factory... The global task/objective may not be known by each agent and is beyond the capacity of a single agent. The interactions between agents are local in the sense that each agent can only use the information from agents in its reach (typically, in spatial proximity). A MAS can be *centralized* (one agent has access to all the other agents), *decentralized* (the system is split into several centralized communities), or *distributed* (there is no centralized agent). A distributed MAS provides robustness and scalability. Robustness because the system is less sensitive to the failure of an agent. Scalability because we can massively increase the number of agents without increasing as much the computing and control cost of the agents.

Some applications of MAS are found in the exploration and mapping of an unknown environment by a swarm of robots, optimisation of public bus schedules, sensor fusion, smart grids control, machine learning, simulation of social networks, task assignment for multi-core microprocessors, district heating optimisation, study of the emergence of collective intelligence...

There are still many research issues in this highly multidisciplinary field. Some examples are time delays, quantization from digital sensors and noise in general, dynamic connectivity, topology optimisation, non-linear dynamics, learning control and in general the transposition of concepts from classical control which remain partial [10]. Even the weighting of first-order average consensus for cooperative control is not clearly settled in general [22].

1.3.1 Graph theory

The agents of a MAS form a network. The topology of this network is usually described through the graph theory. It allows a mathematical characterisation of how the information is exchanged between agents [20].

A **graph** $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is composed of a set of vertex $\mathcal{V}$ and a set of edges $\mathcal{E}$. The vertices represent the agents while the edges describe the interconnection between them. The neighbourhood $\mathcal{N}$ of an agent is the set of agent from which it can receive information i.e., the neighbourhood of agent i $\mathcal{N}_i \subseteq \mathcal{V}$ is defined as $j \in \mathcal{V} \mid (i,j) \in \mathcal{E}$.

A graph is **undirected** if $\forall i, j \in \mathcal{V}, (i,j) \in \mathcal{E} \Rightarrow (j,i) \in \mathcal{E}$. Otherwise, the graph is said to be **directed** i.e. the communication between two agents can be unidirectional. Note that many properties apply to undirected graph only. Directed graphs need more specific tools.

A graph is **connected** if there exists a path between each pair of vertices. A directed graph is strongly connected if and only if there exists a path *that follows the direction of the edges* between each pair of vertices.

The **adjacency matrix** $\mathcal{A}$ is a square matrix representing the communication network: $A \in \mathbb{R}^{nxn}$ for n agents, with $a_{ij} = 1$ if $(j, i) \in \mathcal{E}$ and 0 elsewhere. For undirected graphs, this matrix is symmetric. In a *weighted* adjacency matrix, the unitary gain of each edge is replaced by its appropriate weight. A directed graph is said to be balanced iff $\sum_j a_{ij} = \sum_j a_{ji} \,\forall i$ (an undirected graph is always balanced).

The **degree matrix** $\mathcal{D}$ is a diagonal matrix with the degree of each agent. An agent degree is the size of its neighbourhood. For a directed graph, a distinction has to be made between in-degree (number of directed edges coming into the agent) and out-degree (number of directed edges coming out of the agent). In any case, the (in-)degree d^{in} of an agent i is defined as the sum of the elements of the ith row of the adjacency matrix : $d_i^{in} = \sum_{j=1}^{n} a_{ij}$.

The **graph Laplacian matrix** $\mathcal{L}$ is an important matrix in the graph theory. It is defined as

$$\mathcal{L} = \mathcal{D} - \mathcal{A}$$

For undirected graph, it is positive semi-definite. Its eigenvalues are real and can be ordered as

$$\lambda_1(\mathcal{L}) = 0 \leq \lambda_2(\mathcal{L}) \leq ... \leq \lambda_n(\mathcal{L})$$

For directed graph, this Laplacian can still be used, but its analysis can be misleading. A symmetric Laplacian $\hat{\mathcal{L}}$ based on the *mirror graph* $\hat{\mathcal{G}}$ is proposed in [26] : $\hat{\mathcal{L}} = \hat{\mathcal{D}} - \hat{\mathcal{A}}$ with $\hat{\mathcal{A}}$ computed from an augmented set of edges, $\hat{\mathcal{E}}$ which is the union of $\mathcal{E}$ and *flipped* $\mathcal{E}$ (*flipped* means that all the directions of the edges have been reversed). This results in an undirected graph, so the usual properties can be observed.

The second smallest eigenvalue λ_2 is known as the **algebraic connectivity** or the Fiedler value. A graph is (strongly) connected iff $\lambda_2(\mathcal{L}) > 0$. The algebraic connectivity is used to reflect how well connected the graph is, and how it is linked to desirable control properties [8] [20].

Example. The graph $\mathcal{G}$ in Figure 2 is directed, connected (but not strongly connected) and not balanced. The associated matrices and eigenvalues are :

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathcal{D} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathcal{L} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -2 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad eig(\mathcal{L}) = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 3 \end{pmatrix}$$

Note that the algebraic connectivity is zero since $\mathcal{G}$ is not strongly connected. The graph $\hat{\mathcal{G}}$ in Figure 2 is connected and undirected. The associated matrices and eigenvalues are :

$$\hat{\mathcal{A}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \hat{\mathcal{D}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \hat{\mathcal{L}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -2 & 0 \\ 0 & -2 & 3 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \quad eig(\hat{\mathcal{L}}) = \begin{pmatrix} 0 \\ 0.76 \\ 2 \\ 5.24 \end{pmatrix}$$

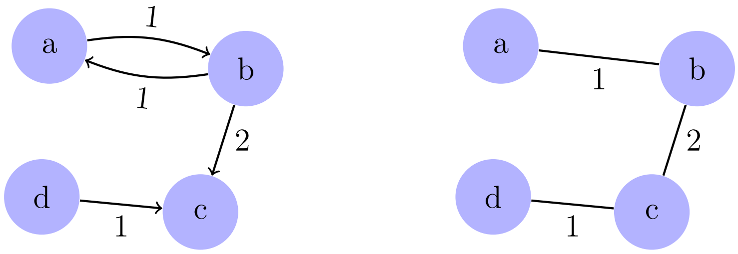


Figure 2: Example of a directed graph $\mathcal{G}$ with 4 vertices and 4 edges (left) and its mirror graph $\hat{\mathcal{G}}$ with 3 undirected edges (right)

1.3.2 Consensus problem

One of the fundamentals in MAS is the consensus algorithm. It specifies local interactions between agents to achieve an agreement on their state value. The consensus algorithm in its simplest form is :

$$u_i = \sum_{j=1}^{n} a_{ij}(x_j - x_i) \tag{11}$$

Where u_i is the control input of the ith agent, x_i its state, n the number of agents and a_{ij} the i,j entry of the adjacency matrix. Note that a_{ij} in non zero only if the agent j is a neighbour of agent i : agent $j \in \mathcal{N}_i$. Let $u = [u_1, ..., u_n]$ and $x = [x_1, ..., x_n]$. Equation 11 can be rewritten as

$$u = -\mathcal{L}x \tag{12}$$

where $\mathcal{L}$ is the Laplacian matrix. We can add a positive proportional control gain K_p. For a single-integrator MAS ($\dot{x}_i = u_i$), the closed loop system is

$$\dot{x} = -K_p\mathcal{L}x$$

where a consensus is reached asymptotically.

In MAS, many of the control approaches studied to this day stems from this consensus algorithm (11). An overview of the control algorithm possibilities is presented in [10]. In this thesis, a formation control algorithm is used. Let $e_i = (x_i - x_i^*)$ the error on the state of agent i from its desired state x_i^*. The desired formation is then denoted by $x^* = [x_1^*, ..., x_n^*]$. A simple feedback control is then

$$u_i = -e_i \tag{13}$$

The collaboration between agents is not strictly necessary to maintain each agent to its desired state. However the introduction of an agent collaboration term into Equation 13 can improve the convergence speed of the system. Again, we can add a positive proportional control gain K_p :

$$u_i = -K_p\left(e_i + \sum_{j=1}^{n} a_{ij}(e_i - e_j)\right) \tag{14}$$

The Equation 14 is known as formation control with absolute information.

In [10], other formation controls with relative information, distance information and rigid formation control are presented. In [1], consensus, flocking and formation control are illustrated trough different MATLAB simulations. In [6], distributed control is applied to the control problem of microgrids.

1.3.3 MAS for WWTP

There is no established set of criteria to determine when an agent approach should be used for a project [7]. A agent-based approach can be useful :

- When adaptability to a changing or uncertain environment is needed
- When agents are a natural description of the problem
- When the control or the expertise availability is distributed
- When high modularity is required
- When it is not possible to specify a case-by-case behaviour
- When the overall system properties should be maintained e.g. through self-healing.

In the context of WWWTP's, the multi-agent framework can be considered for the network of different WWTP's within a river basin or for a single WWTP considered as a network.In the first case, the goal is to exploit the communication among different dynamical agents such as WWTP's, upstream measurement stations and industrial players in order to manage the water quality. This has been studied in Spain [40] and in China [25]. The MAS at the WWTP scale is not much studied yet. A decentralised architecture is presented in [3], but it is focused on software interfaces. A recent paper (April 2021) used multi-agent deep reinforcement learning on a WWTP model [11]. Two agents are optimising the dissolved oxygen and the the chemical dosage under different reward functions. Artificial intelligence is making its way to WWTP [29].

1.4 Conclusion

MAS is an exciting field with different uses in many disciplines. In this thesis, the focus is on the topology of the agent network i.e., how the communications between local agents of the nitrogen removal section impact the quality of the discharged water. Indeed difficulties in the measurement of important parameters in the wastewater denitrification unit alongside poor mathematical models lead to an inconsistent open-loop control of the unit in spite of its ecological impact [34]. A simple feedback control was only proposed recently, relying on nitrite measure at the output of the unit.

Using the MAS framework on the nitrogen removal section of SAV, the thesis has two goals. First, to lay the foundations of a tool to better characterise the impact of the flows between the different units of the WWTP and the impact of the information availability of the flows concentration. Second, to control these flows in a way to conform to the latest water quality norms, even in case of heavy rain conditions.

2 Case study for regular days

In this section, the MAS for the nitrogen removal of SAV is presented. The scheme is given alongside the hypotheses and the constraints.Then initial values and the equilibrium point are specified. The MATLAB implementation is presented in subsection 2.4, after which the results are given for regular days.

2.1 Case study

The nitrogen removal system of SAV is the MAS. In the thesis, the pre-denitrification unit is not considered. This corresponds to the state of SAV between 2007 and 2017. The system is represented in Figure 3. Each agent has 3 state variables : its concentration in NH_4, NO_2 and NO_3 (in this order, e.g. NH_4 will always be referred as the first state variable). x_1 is the nitrification unit and x_2 is the (post-)denitrification unit. They are the two agents with inner dynamics, described in Equation 17-22. The other agents either serve as input values/perturbations (v_{in} and v_{seine}) or as the output values (y). v_{in} represent the water state after the pre-denitrification unit. The concentrations are in $[g/m^3]$ (equal to $[mg/L]$, which is the usual unit in the domain). The flow rates u_i (controllable) and Q_i (non-controllable directly but u_i-dependent) are in $[m^3/h]$ although their value will often be given in $[m^3/s]$ for better intuition. Each agent controls the flow rates u_i going towards themselves e.g. x_1 controls u_1 and u_R. Since we work at constant volumes in the units x_1 and x_2, the non-controllable flow rates are given as :

$$\begin{aligned}
Q_0 &= Q_{in} - u_1 - u_3 \\
Q_1 &= u_1 + u_R - u_2 \\
Q_2 &= u_2 + u_3 - u_R \\
Q_{out} &= Q_{in} + Q_0 + Q_1 + Q_2
\end{aligned}$$

We assume that all the chemical reactions happen inside the units and that they are perfectly regulated. That is, we will not take into account other parameters than the three concentrations of interest (e.g. oxygen demand or reactant supply). Although the flow rate will often be presented here in $[m^3/s]$, the time unit used in the equations is the hour, to keep up with the kinetics of the reactions.

The objective is to obtain output concentrations below the legal constraint (see Table 1) by using a multi-agent framework. The regulation will apply on the different flow rates, without using integrative or derivative action (only a proportional action).

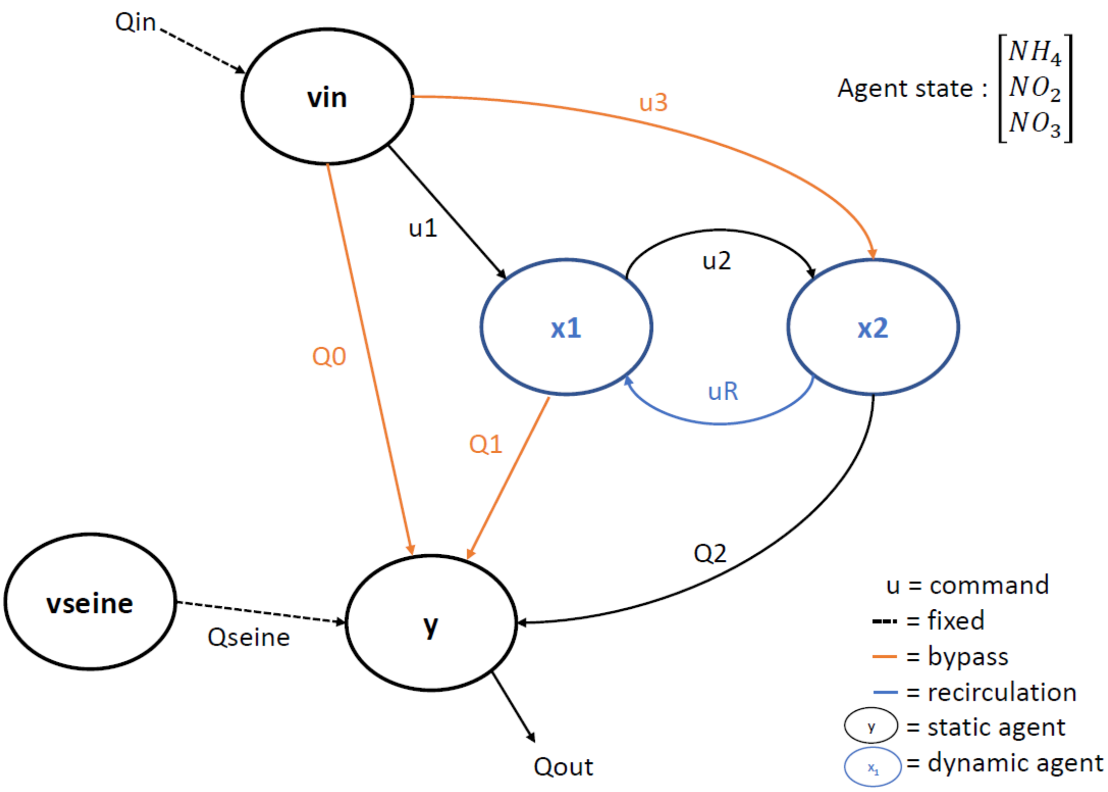


Figure 3: The system

The general equations for the two dynamical agents x_1 and x_2, and for the output y are :

$$\underline{\dot{x}}_i = A_i(u)\underline{x} + D_i(u)\underline{v} - K_i\underline{\phi}(x) \tag{15}$$

$$\underline{y} = \frac{1}{Q_{out}}\left(Q_1(u)\underline{x}_1 + Q_2(u)\underline{x}_2 + Q_0\underline{v}_{in} + Q_{seine}\underline{v}_{seine}\right) \tag{16}$$

Where $\underline{x}_i = [x_{11}\, x_{12}\, x_{13}]^T$, $\underline{y} = [y_1\, y_2\, y_3]^T$, $\underline{v}_{in} = [v_{in,1}\, v_{in,2}\, v_{in,3}]^T$,$\underline{v}_{seine} = [v_{seine,1}\, v_{seine,2}\, v_{seine,3}]^T$, $\underline{x} = [x_1\, x_2]^T$ and $\underline{v} = [v_{in}\, v_{seine}]^T$. Note that x_{ij} represents the jth concentration of the unit x_i according to this order : [NH4 NO2 NO3]. $A_i(u)$ is the 3x3 matrix associated to x_i, $D_i(u)$ is the 3x6 matrix associated to $\underline{v}$, K_i is the yield coefficients 3x4 matrix associated to x_i and $\phi(x)$ a 4x1 vector with the growth rates.

For the nitrification unit (x_1) it gives :

$$\dot{x}_{11} = \frac{u_1}{V_1}(v_{in,1} - x_{11}) + \frac{u_R}{V_1}(x_{21} - x_{11}) - k_{11}\left(\mu_{1,max}X\frac{x_{11}}{K_{11} + x_{11}}\right) \tag{17}$$

$$\dot{x}_{12} = \frac{u_1}{V_1}(v_{in,2} - x_{12}) + \frac{u_R}{V_1}(x_{22} - x_{12}) + k_{12}\left(\frac{\mu_{1,max}X}{Y_1}\frac{x_{11}}{K_{11} + x_{11}}\right) - k_{22}\left(\mu_{2,max}X\frac{x_{12}}{K_{12} + x_{12}}\right) \tag{18}$$

$$\dot{x}_{13} = \frac{u_1}{V_1}(v_{in,3} - x_{13}) + \frac{u_R}{V_1}(x_{23} - x_{13}) + k_{23}\left(\frac{\mu_{2,max}X}{Y_2}\frac{x_{12}}{K_{12} + x_{12}}\right) \tag{19}$$

These are Equation 3 with the system parameters plugged-in and additional kinetic coefficients. k_{ij} are coefficients (related to μ_i and concentration j), $\mu_{i,max}$ and K_{1j} the maximum specific growth rates $[1/h]$ and the saturation constant for each specific growth rate, modeled with the Monod equation, respectively. V_1 is the volume of the tank unit $[m^3]$, X the biomass concentration $[g/m^3]$ and Y the yield coefficient. The numerical values of the constants are given in Table 2, at the exceptions of the k_{ij} coefficient introduced here to ease the finding of equilibrium values.

For the denitrification unit (x_2) it gives :

$$\dot{x}_{21} = \frac{u_3}{V_2}(v_{in,1} - x_{21}) + \frac{u_2}{V_2}(x_{11} - x_{21}) \tag{20}$$

$$\dot{x}_{22} = \frac{u_3}{V_2}(v_{in,2} - x_{22}) + \frac{u_2}{V_2}(x_{12} - x_{22}) + k_{a2}\,(\frac{\mu_{a,max}X}{Y_a}\frac{x_{23}}{K_{23} + x_{23}}) - k_{b2}\,(\mu_{b,max}X\frac{x_{22}}{K_{22} + x_{22}}) \tag{21}$$

$$\dot{x}_{23} = \frac{u_3}{V_2}(v_{in,3} - x_{23}) + \frac{u_2}{V_2}(x_{13} - x_{23}) - k_{a3}\,(\mu_{a,max}X\frac{x_{23}}{K_{23} + x_{23}}) \tag{22}$$

These are Equation 7 with the system parameters plugged-in and additional kinetic coefficients. The numerical values of the constants are given in Table 3.

The volume is set to $V_1 = V_2 = 30,000\,[m^3]$. This is approximated from the total volume of SAV [32] and measurements on Google Maps. The biomass concentration is hard to measure. Here we set it to $1250[mg/L]$ for all SAV units, based on the biomass concentration used in [32]. The maximal instantaneous inlet flow rate of SAV is $70[m^3/s]$. For the nitrification and denitrification units, the flow rates are limited to about 90% and 70% of the treatment capacity of SAV (that is 23.9 and $18.4[m^3/s]$ respectively).

2.2 Initial concentration values

We have concentration values for the nitrification, pre- and post-denitrification units for May 2018 in ([41], Annex 4), given in Table 4. There are also typical concentrations derived from the concentrations measured by the SIAAP during 2009-2014 at the input and output of SAV and in the Seine river before (Sartrouville) and after (Poissy) the WWTP. The extraction of these values from the dataset is presented in subsection B.1.

	NH_4 $[mg/L]$	NO_2 $[mg/L]$	NO_3 $[mg/L]$	flow $[m^3/s]$
pre-denitrification unit (v_{in})	19.7	0.3	2.2	/
nitrification unit (x_1)	1.2	0.3	19.2	/
post-denitrification unit (x_2)	0.7	1.0	1.7	/
SAV_{input}	44	/	/	16.67
SAV_{output}	2.62	4.45	21.96	/
Sartrouville	0.09	0.08	20.04	350
Poissy	0.22	0.34	20.88	/

Table 4: Concentrations and flow in different places for typical good conditions

2.3 Equilibrium values

With the values from Table 4 and Equation 17 to 22, equilibrium values for the simulation for regular days can be obtained i.e. $\dot{x}_ij = 0$. Firstly, it is easy to see in Equation 20 that there is no equilibrium for the ammonium concentration in x_2 unless both inlet flow rates

u_3 and u_2 are zero (since both $v_{in,1} - x_{21}$ and $x_{11} - x_{21}$ are positive values 4). However, we can assume that at equilibrium u_3 is zero, and the ammonium concentrations in x_1 and x_2 are equal. Thus relaxing u_2.

Secondly, the values in Table 4 need to be scaled so that v_{in} numbers meet SAV_{input} and the mix of x_1 and x_2 meets SAV_{output} (Q_0 is set to 0 at equilibrium). This is done by following the chemical reactions from v_{in} to x_1 to x_2 (e.g. around 6% of $NH_{4,in}$ is left in x_1 and the rest is transformed into nitrate).

Thirdly, since the ammonia concentration remains constant in x_1 and x_2, we set it to SAV_{output}, which also happens to be around 6% of SAV_{input} concentration. If we put nitrite and nitrates concentration of SAV_{input} to 0 (they are unknown), SAV_{output} concentrations cannot reasonably be achieved. We arbitrarily set them to 2 $[mg/L]$.

The resulting concentrations are in Table 5. These are the ones used and referred further as the (initial) equilibrium concentration values. With these, the equilibrium flow rates can be found to match SAV_{output} nitrite and nitrate concentrations. At the exception of u_R which can take an arbitrary value between 0 and Q_1. Once u_R is set, for example at $0.6 \cdot Q_2$, the k_{ij} coefficients introduced in equations 17 to 22 can be found at equilibrium. The equilibrium system is represented with its flow rates in Figure 4, Table 7 and the coefficients are given in Table 6. The concentrations found for y are consistent with the Poissy concentrations in Table 4.

	NH_4 $[mg/L]$	NO_2 $[mg/L]$	NO_3 $[mg/L]$
v_{in}	44	2	2
x_1	2.62	2	43.36
x_2	2.62	6.34	4.34

Table 5: Equilibrium concentration values

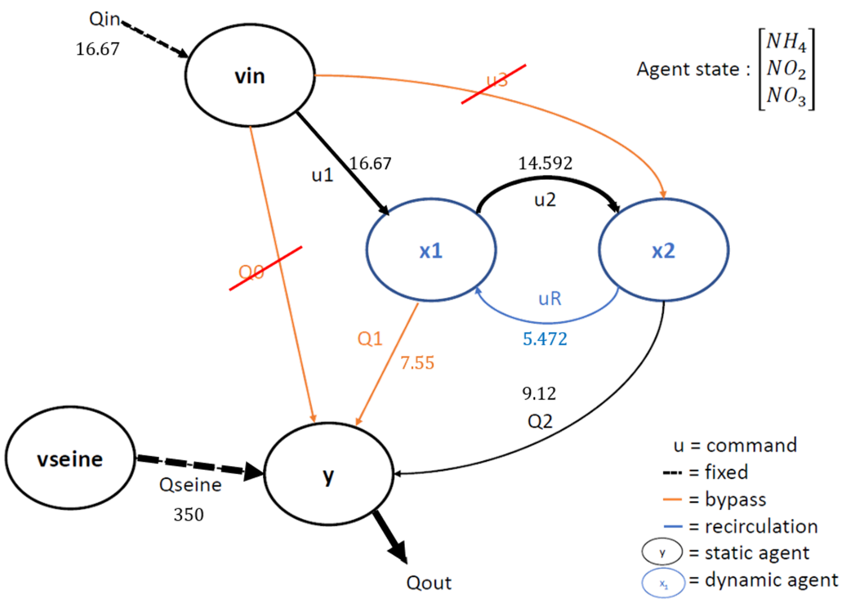


Figure 4: The system at equilibrium. Larger black arrows indicate the main flow paths. Flow rates crossed in red are set to zero.

k_{11}	k_{12}	k_{22}	k_{23}	k_{a2}	k_{b2}	k_{a3}
0.955	0.086	0.086	0.333	0.007	0.007	0.002

Table 6: Coefficient values in equations 17 to 22 (lower than 1)

u_1	u_2	u_3	u_R	Q_{in}	Q_0	Q_1	Q_2	Q_{seine}
16.67	14.592	0	5.472	16.67	0	7.55	9.12	350

Table 7: Flow values at equilibrium, in $[m^3/s]$

2.4 Matlab implementation

A flowchart of the main MATLAB files used for the simulation is given in Figure 5. Each block represent a file, with the name in purple. The code is given in the Appendix attached to the thesis (`Somers_24721400_2021_Appendix.zip`).

- `SAV.m` is the main file. There is one loop to compute the evolution of the system (Equation 15). Every minute, the control values u are updated based on the last state of the system.
- `init_SAV.m` defines all the initial values for the simulation parameters. These are the values from Table 4 to 7.
- `Adj_SAV.m` defines the adjacency matrix of the system, possibly weighted with the current flows values.
- `u_SAV.m` is the controller function.
- `constrain_SAV.m` applies the physical constraints given in subsection 2.1.
- `my_dyn.m` uses ode45 to solve Equation 15
- `A_SAV.m` defines the matrices $A(u)$ and $D(u)$ from Equation 15 with the current flows values
- `Kphi_SAV.m` defines the vector $K\phi(x)$ from Equation 15 with the current states values. The biological constants are also defined here.

Finally, the simulation displays figures such as the state of each agent, the control values (before and after being constrained) and the state deviation of each agent from its equilibrium values.

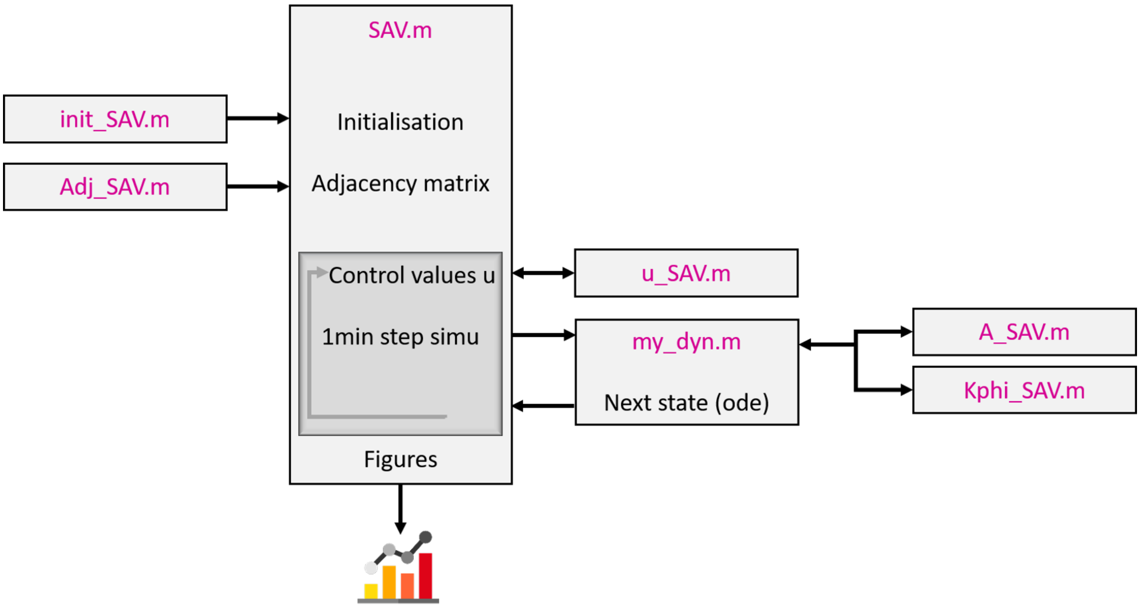


Figure 5: Simulation code flowchart

There are also different MATLAB scripts in the Appendix attached to the thesis. They are summarised in Appendix C.

2.5 Simulation for typical days

The objective is to show satisfying behaviour of the simulation. Different stages of regulation are explored in the following section, but the values are not properly tuned as this effort will be reserved for the final simulation of interest : for heavy rain conditions (section 3).

2.5.1 Open loop simulation

Using the equilibrium values given in Table 5, we first simulate the system with constant flows. A small disturbance is added in the input concentration (+5%) at time t=20h. The simulation for the concentrations in x_2 (Figure 6) shows firstly that the initial values are indeed at equilibrium, and secondly that it can reach another equilibrium state in few hours. x_1 holds similar dynamics.

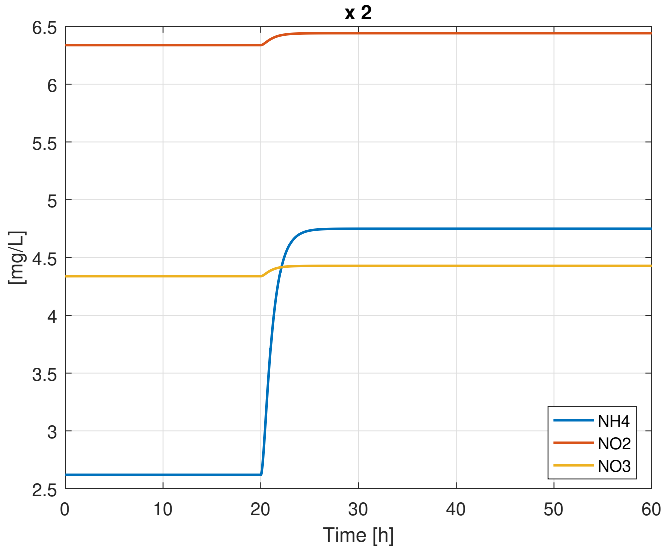


Figure 6: x_2 concentrations for constant flows and input disturbance at times t=20h

2.5.2 Proportional control

A simple proportional regulation is then introduced: $r_i = -k \cdot e_i$, with k a constant and e_i the error (see Equation 13). It is decided to use directly the error on the concentrations of the receiving unit to regulate the flows. That is : $u_{i,(t+1)} = u_{i,(t)} - k \cdot e_{x_a}$, with x_a the appropriate unit for the flow u_i. The error is computed from the normalised error between the unit current concentrations and its goal concentrations, the goal being the initial equilibrium concentrations. The error is a scalar, obtained by a weighted sum of these values. For this section, the weights are set to $[1\,1\,1]/3$ for the NH4, NO2 and NO3 respectively. So we have :

$$\begin{aligned} u_1 &= u_1 - k \cdot e_{x_1} \\ u_2 &= u_2 - k \cdot e_{x_2} \\ u_3 &= u_3 - k \cdot e_{x_2} \\ u_R &= u_R - k \cdot e_{x_1} \end{aligned} \tag{23}$$

The flow rates must settle first for the concentration values to reach a steady state. An instantaneous application of the final flow rates (very high gain k) is not desirable. If the disturbance is small, the dynamics are typical as displayed in Figure 7. Ten hours after the disturbance, $e_{x_1} \approx 0$ while $e_{x_2} \approx 0.21$ which can be indirectly observed as u_1 and u_R quickly stabilise while u_2 is still decreasing. u_3 is not affected as it is always negative here, and thus cropped to zero by the constrains.

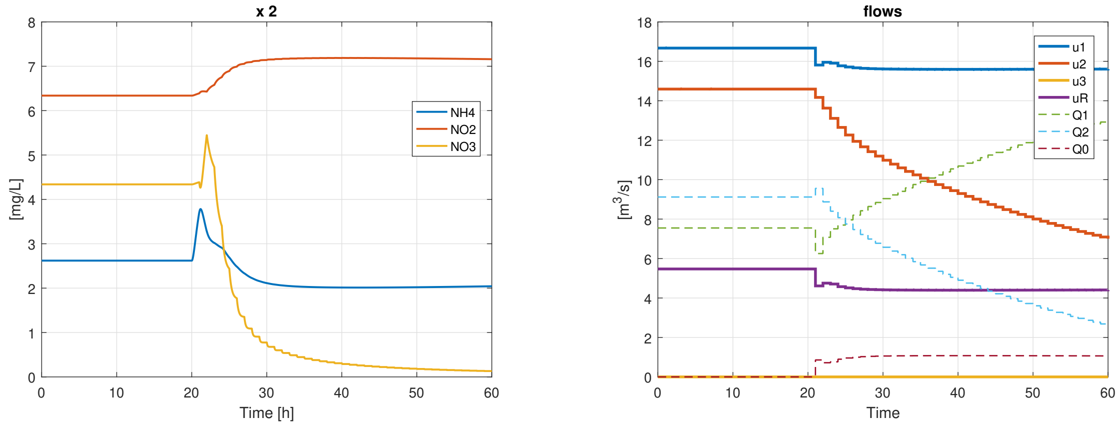


Figure 7: x_2 concentrations with $k = 36,000$, for an input disturbance of 5% at time t=20h. Ripples are caused by the discrete update of u_i. Note that the NO3 low value is due to u_2 being low (little flow rate from x_1 to x_2)

2.5.3 Proportional control with collaboration between the agents

The formation control with collaboration is adapted from the literature (Equation 14). That is :

$$u_{i,(t+1)} = u_{i,(t)} - k \cdot \left(e_{x_a} + \sum_{j=1}^{n} a_{aj}(e_{x_a} - e_{x_j}) \right) \tag{24}$$

with n the number of agents and a_{aj} the adjacency matrix's elements. The index $a = 1, 2$ refers to the dynamic agents. Here we suppose that all the agents know the state of the others, i.e. $a_{aj} = 1 \, \forall \, a \neq j$. Note that the flow rates u_1 and u_R will thus have the same control here, as well as the flow rates u_2 and u_3, since they are respectively the input flow rates of x_1 and x_2.

If only x_1 and x_2 are taken into account for the collaboration term, we get :

$$\begin{aligned}
u_1 &= u_1 - k \cdot e_{x_1} - k_u \cdot u_R(e_{x_1} - e_{x_2}) \\
u_2 &= u_2 - k \cdot e_{x_2} - k_u \cdot u_2(e_{x_2} - e_{x_1}) \\
u_3 &= u_3 - k \cdot e_{x_2} - k_u \cdot u_2(e_{x_2} - e_{x_1}) \\
u_R &= u_R - k \cdot e_{x_1} - k_u \cdot u_R(e_{x_1} - e_{x_2})
\end{aligned}$$

Where the weight of the adjacency matrix is k_u times the flow connecting the two agents. With $k = 25,200$ and $k_u = 0.3$, the dynamics are very similar to the simple proportional control. When k_u is increased, oscillations appears during the transient phase (each agent responds strongly with the variations of the other one).

It gets interesting when collaboration with the output agent y is added. Its error is $e_y = \frac{y - EQS}{EQS}$ with EQS being the Environmental Quality Standard (the concentrations limits) from Table 1. This term is negative since in typical conditions, we have $y = [0.20 \, 0.28 \, 20.13] < EQS = [0.5 \, 0.3 \, 50]$.

The simulation, with the collaboration with all agents and a unique gain k, without disturbance, is shown in Figure 8. This simulation is with older conditions such as concentration weights set to $[1 \, 10 \, 1]/12$, but the results are still interesting to show. Since e_y is non zero, another equilibrium is deemed more successful even without disturbance.

Now the flows are going only through the nitrification unit. Indeed, in typical conditions, the concentrations in x_1 are good enough to respect the EQS, and all the water can be treated by this unit alone. Oscillations appears on the flows around t=25h (but nor the weights nor the gains are tuned). In fact, this is due to u_3 crossing the 0 threshold : at this point all the other flow rates are already constrained to have $Q_{in} = Q_0 + Q_1 + Q_2$. So when u_3 increases, the other flow rates have to decrease. The ammonium concentration in x_2 rise during these oscillations when $u_3 > 0$ carries high concentrations from the input (remember that this unit does not process any ammonium, see Equation 20). As the command is not bounded by the constraints, the constraint application after the computation of the command leads for example to u_R set to zero most of the time. At t=30h, the error e_{x_2} oscillates close to zero while e_{x_1} oscillates around -0.45. This is due to the deviation of NO2 concentration and the fact that a negative e_{x_1} drives the decrease of u_2 and u_3, and thus indirectly closes the connection with x_2 which is the nitrite source. Only one configuration is presented here, different weights distribution lead sometimes to different scenario (especially when u_2 meets u_R value). So there is room to play with the regulation, to adapt it to our specific case.

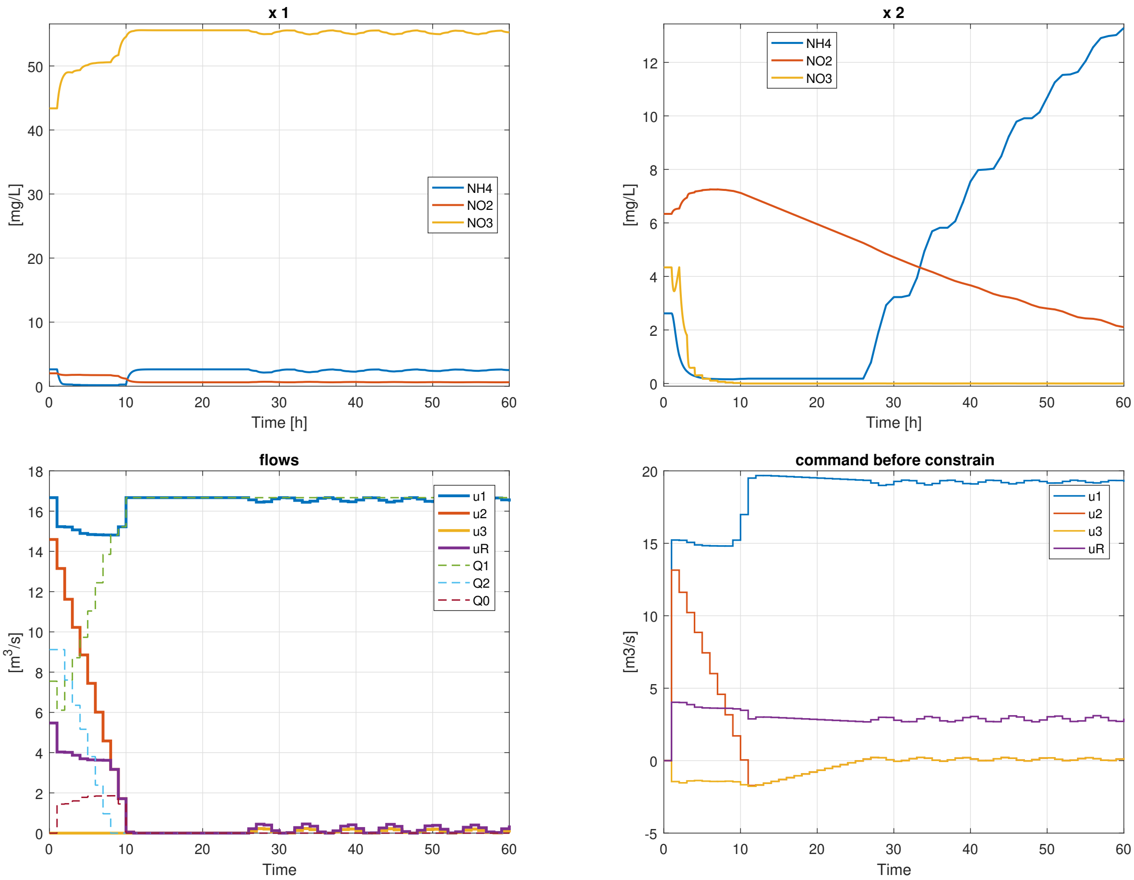


Figure 8: x_1, x_2 concentrations and flows with k=500 without disturbance. Note that the path through x_2 is largely ignored

A more recent simulation (done with the code provided in Appendix), is displayed in Figure 9. The concentration weights are normalized again (so k has been increased to 6000), and some numerical errors (arising from the numerical difference for the regulation) have been taken care of. The previous example is thus out of date, but it allows to show the effect of the constraints so we kept it.

Here all the flow rates pass through u_1 and Q_2 (Q_1 goes to zero), with some recirculation between the two units. It shows better performances : it converges from the initial state in 15h. At t=40h, the initial disturbance (+5%) is added, and a new equilibrium is reach few hours after. At t=50h, $e_{x_1} \approx -0.31$ while $e_{x_2} \approx 0.09$. e_y moves from -0.12 at t=0 to -0.10 at t=50.

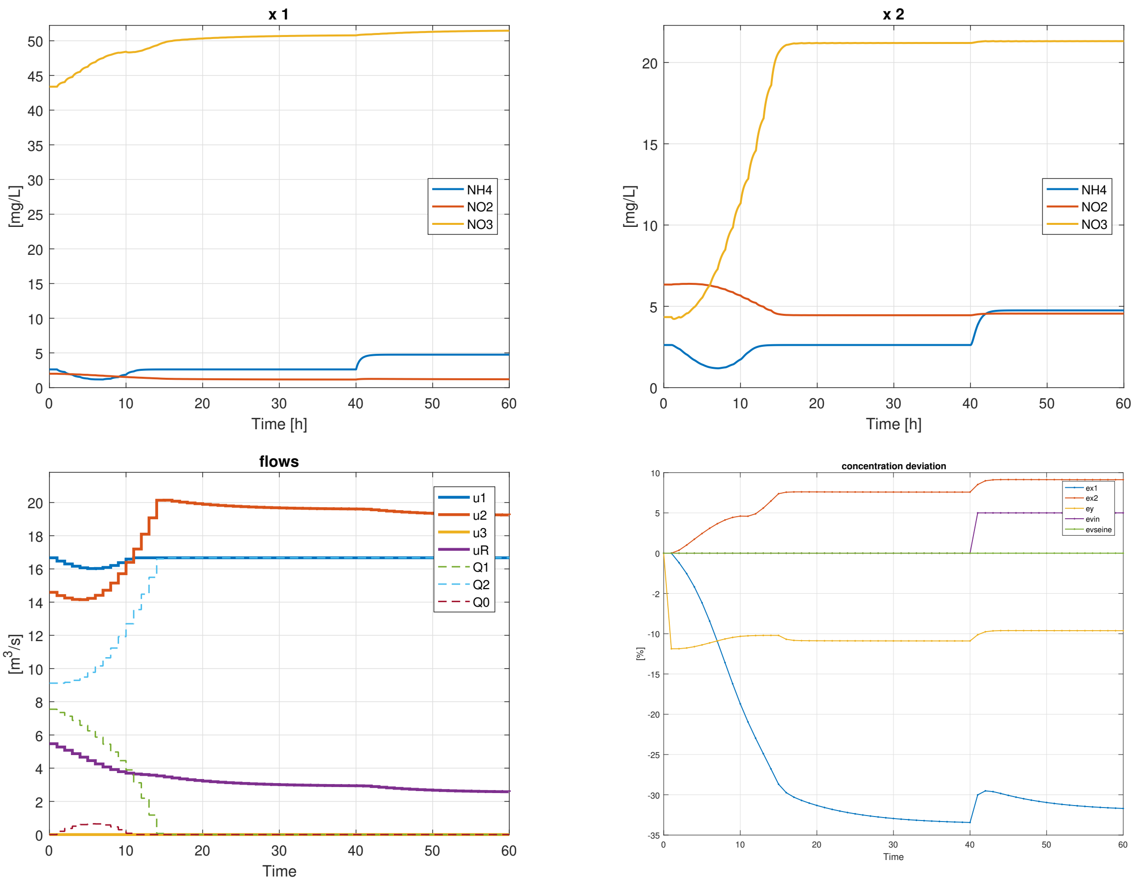


Figure 9: x_1, x_2 concentrations, flows and normalized errors, with k=6000. Input disturbance of 5% at t=40h

2.6 Discussion

The simulation developed in this section has a coherent behaviour, and it will allow the study of several types of control schemes for heavy rain conditions where the WWTP has difficulties to stay below the EQS. This is the object of the next section (section 3).

The simulation captures the behaviour of the nitrification, post-denitrification unit, and the interactions with water flow rates. The local control of the units (e.g. air supply) is considered as an independent problem and it is thus not investigated here. Since high level of nitrite can also occur when the supply of the reactant for the post-denitrification unit is poorly controlled, the model cannot capture all the dynamics related to the nitrogen. Also the value in Table 5 to 7 are mostly extrapolated from the literature available, and may deviate from the reality.

The physical constraints influence the dynamics, as discussed in subsubsection 2.5.3 and illustrated in Figure 8. These constraints are applied after the computation of the control input, but they could be included into the control itself (e.g. by mapping the control input to a sigmoid function).

The error on the output agent, e_y, is negative when the output concentrations are below the EQS. And it is not in the interest of the WWTP operator to do more than the norms require. Indeed finer treatment of the water cost money as more reactant is needed. In consequence the error on the output concentrations is often rectified to zero so that it is only taken into account when the norms are exceeded (i.e. when $e_y > 0$) [33]. This is implemented in the next section.

3 Case study for heavy rainfall scenario

In this section, a heavy rainfall scenario is characterised for the simulation. Disturbances to apply on the system are designed, and a benchmark is given to observe the performance of different control approaches. The impact of various communications between agents is then assessed and discussed.

The MAS is similar to the previous section.

3.1 Heavy rainfall conditions

Concentration levels in the Seine river and in the discharged water from SAV are derived from the same data set (see Appendix), but considering only some days with high inlet flow rate at SAV. The values are given in Table 8, following the methodology in annex B.2. The extrapolation of data for the state of x_1 and x_2 is questionable. However the concentration values will be simulated in subsubsection 3.1.1. The nitrite concentration at the input of SAV is set to $1\,[mg/L]$ to follow similar dilution as the ammonium. The nitrate concentration has been increased to 6 $[mg/L]$ because of nitrate run-off during heavy rainfall, although the lack of available data makes this value relatively uncertain.

	NH4 $[mg/L]$	NO2 $[mg/L]$	NO3 $[mg/L]$	Q $[m3/s]$
$SAVinput$	17.34	/	/	35
$SAVoutput$	7.56	2.34	19.00	35
Sartrouville	0.12	0.02	20.80	/
Poissy	0.28	0.47	20.57	766

Table 8: Values for heavy rainfall conditions

The initial values for the flow rates are similar to the ones for regular days, but scaled in function of Q_{in} e.g. $Q_1 = 0.45 \cdot Q_{in}$. The initial flow rates are often at their maximal value since Q_{in} (35 $[m^3/s]$) is larger than the treatment capacity of SAV (26.6 $[m^3/s]$). As a result Q_0 is non-zero. u_3 has been set arbitrarily to $1\,[m^3/s]$ to better see the dynamics of the different control approaches (otherwise it would often remain at zero).

3.1.1 Equilibrium values

The simulation is used to find equilibrium values for heavy rainfall conditions. The idea is to start from the equilibrium for regular days (subsection 2.4), then slowly adapt the flow rates Q_{in}, Q_{seine} and their concentration to the values of Table 8, and let the system adapt to this in open loop. Figure 10 shows the slow evolution of the flow rates (during 400 hours). Note the saturation of u_1 and u_2 (in blue and orange, respectively) and the rise of Q_0 when they are both saturated.

The values for x_1, x_2 and SAV_{output} are then collected at the end. These values are given in Table 9. Note that the concentrations of SAV_{output} are not imposed on the simulation as it is not an input variable. However the two SAV_{output} values can be compared: the simulated values are about 20% lower. This is not a big concern since the reference values are the means of a few days with different conditions. Indeed the mean of a set of values is not always a good depiction of this set of values e.g. when the values are either around

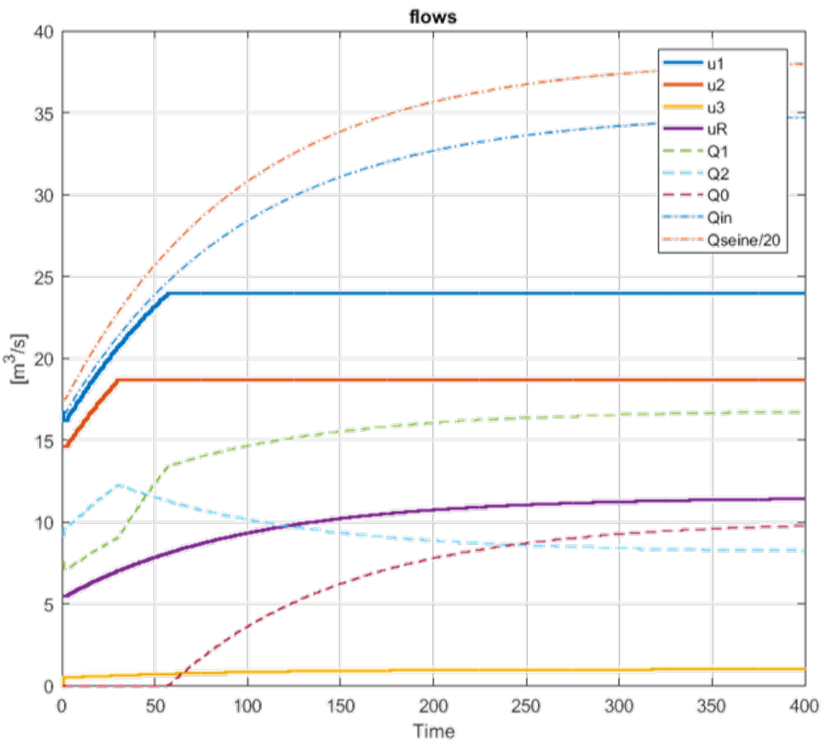


Figure 10: Slow evolution of the flows from the regular days equilibrium to the heavy rainfall scenario

-1 or +1, the mean 0 is not representative. The same holds for the set of averaged output values. Thus the input values from Table 8 do not imply the output values, and they only serve as guidance.

The values for the state of the units can be compared with the steady states for regular days: the ammonium level decreases in the units, which can be explained by the low ammonium concentration in the inlet flow rate. Note that the nitrate content in x_1 is greater than that of v_{in}. This is due to the initial concentration in x_1 (43.36 $[mg/L]$).

	NH_4 $[mg/L]$	NO_2 $[mg/L]$	NO_3 $[mg/L]$
v_{in}	17.34	1	6
x_1	0.02	1.32	29.88
x_2	0.90	4.35	1.22
SAV_{output}	5.20	1.94	16.36

Table 9: Equilibrium concentration values

The equilibrium system is represented with its flow rates in Figure 11 and in Table 10.

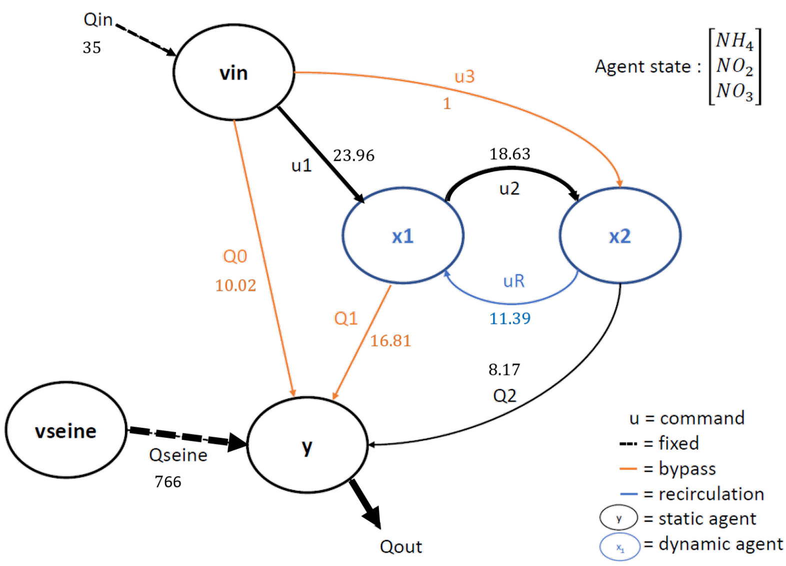


Figure 11: The system at the heavy rainfall equilibrium. Larger black arrows indicate the main flow paths. Flows crossed in red are set to zero

u_1	u_2	u_3	u_R	Q_{in}	Q_0	Q_1	Q_2	Q_{seine}
23.96	18.63	1	11.39	35	10.02	16.81	8.17	766

Table 10: Flow rates values at equilibrium, in $[m^3/s]$

3.1.2 Disturbance

The disturbance is represented in Figure 12 and designed as follows:

t=5h : step for Q_{in} (t_1)

t=10h: step for Q_{seine} (t_2)

t=15h: step back to the initial conditions (t_3)

t=25h: step for ammonium concentration in Q_{in} : $+10[mg/L]$ (t_4)

t=35h: step back to the initial conditions (t_5)

The delay between the disturbance in Q_{in} and Q_{seine} is used to model the lag time (subsection 1.1). It is important to model this lag time: without it, the poorly treated water would simply be diluted in the large flow rate of the river. It is indeed when the flow rate of discharged water of poor quality is large while the river flow rate is still regular that overshoot of the EQS happens. The step back to the initial conditions, at t=15h, is there to confirm the good behaviour of the controller and to observe the settling of the system to the reference conditions. The step at t=25h is to test the limits of the control approaches and observe their stability for a harsh disturbance, when the EQS cannot be respected. A ramp function could be used instead of a step. Indeed it is more natural not to have an instantaneous inlet flow rate. But a ramp is difficult to characterise and varies with the rain event. A step function can also be seen as a worst-case scenario compared to a ramp function. This is consistent with the objective of the thesis: to study control approaches for extreme events.

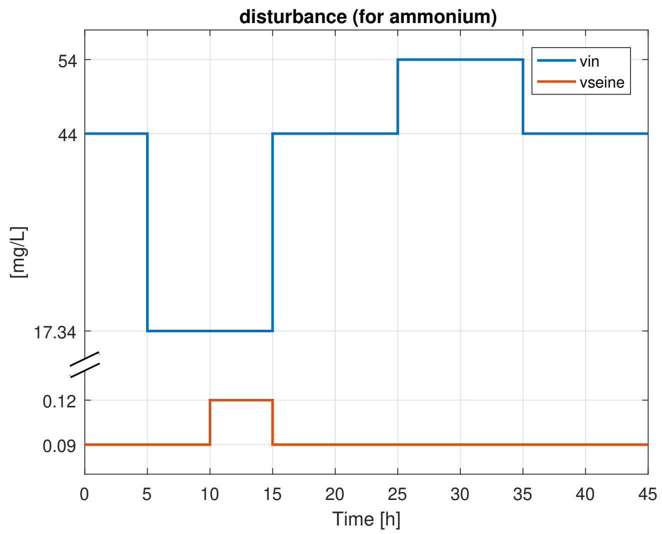


Figure 12: Ammonium disturbance applied to the system at equilibrium for regular days. The disturbance also concerns nitrite, nitrate and flow rates, not displayed here for clarity. Not to scale

3.1.3 Benchmark

The different control approaches used in the next section will be evaluated via the transient time and final state values. Five indicators are used for each step:

- the time above EQS (the sum of the time above EQS for each output concentration): $\#EQS$
- the settling time, within a 5% range: t_s
- the presence of oscillation (none, small, medium, large, unstable): oscillation
- the deviation from EQS, at steady state (the one that deviates most among the output concentrations): $\tilde{EQS}$
- the deviation of x_1 and x_2 from the equilibrium values, at steady state: $\tilde{x}_1$ and $\tilde{x}_2$

The first indicator is the most important, particularly during t_1. Since it is also easily quantifiable, it will be the one used for automatic ranking of large set of simulations in paragraph 3.2.3.3.

3.2 Simulations

The benchmark is applied to different control approaches, with the open loop as the reference case.

3.2.1 Open loop simulation

The controllable flow rates u_i are not kept constant, otherwise the additional inlet flow rate only goes trough Q_0 without affecting the WWTP system. Instead the same logic as in subsubsection 3.1.1 is used: the flow rates are proportional to Q_{in}.

The simulation for the output agent y and the flow rates are displayed in Figure 13 and the benchmark is given in Table 11. t_1 is the key time period. The ammonium limit is exceeded during all five hours because of the bypassed water, and the nitrite limit is slightly exceeded for 7 minutes. When the river flow rate increases, the ammonium concentration drops due to the dilution. During t_4 all the water is treated but the ammonium level is too high to handle for the nitrification unit. There is no oscillation. The steady error of x_1 and x_2 are comparable.

The big overshoot at the output at t=5h is caused by the steepness of the perturbation: the adaptation of the flow rates to the inlet flow rate variation is not instantaneous and thus large amounts of water are bypassed for a moment. Besides note that the flow rates are set to follow the input step instead of ramping up more slowly, which is not realistic. In the next sections, appropriate control gains limit the increase of the flow rates.

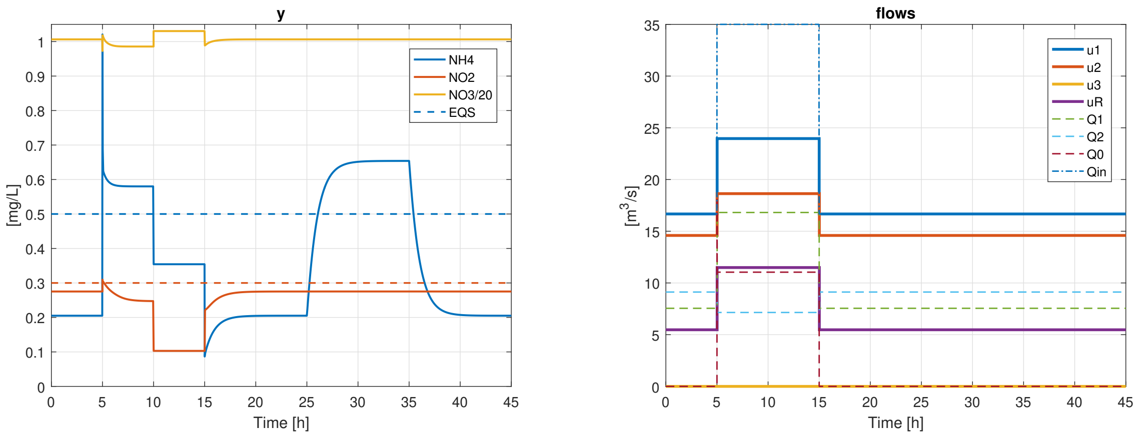


Figure 13: Open loop simulation. The evolution of the output agent y (left) and the flows (right)

OL	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	307	1	0	535	26	869
$t_s\,[h]$	/	1.6	/	2.8	2.9	2.9	
oscillation	/	/	/	/	/	/	
$\tilde{EQS}\,[\%]$	0	16	0	0	31	0	
$\tilde{x}_1\,[\%]$	0	-64	-64	0	126	0	
$\tilde{x}_2\,[\%]$	0	-67	-67	0	126	0	

Table 11: Benchmark for the open loop simulation

3.2.2 Proportional control

The control approach is the same as in subsubsection 2.5.2. Three different gain types are tested: general gains (g *general*) and gains tuned around each equilibrium (g *regular* and g *rain*, respectively)

For the general gains case, the equation is the same as in Equation 23, but different gains k_1 and k_2 are chosen for the inlet flow rate (u_1 and u_3) and the flow rates between the two units (u_2 and u_R), respectively. k_2 is fixed to half the value of k_1 to prioritise the control of the inlet flow rate over the circulation between the two units.

The simulation for the output agent y and the flow rates is displayed in Figure 14 and the benchmark is given in Table 12. There are two main observations: concentration can be driven below the EQS in t_1, and the control become unstable in t_4 as the flow rates crash to zero (all the inlet flow rate is thus bypassed). The system oscillates during t_3. The steady state error of x_2 is remarkably low, due to the use of u_3. Indeed the error e_{x_2} is the normalised sum of its tree concentration deviation, so when the nitrite and nitrate levels decrease, an increase in ammonium through u_3 can compensate for these losses.

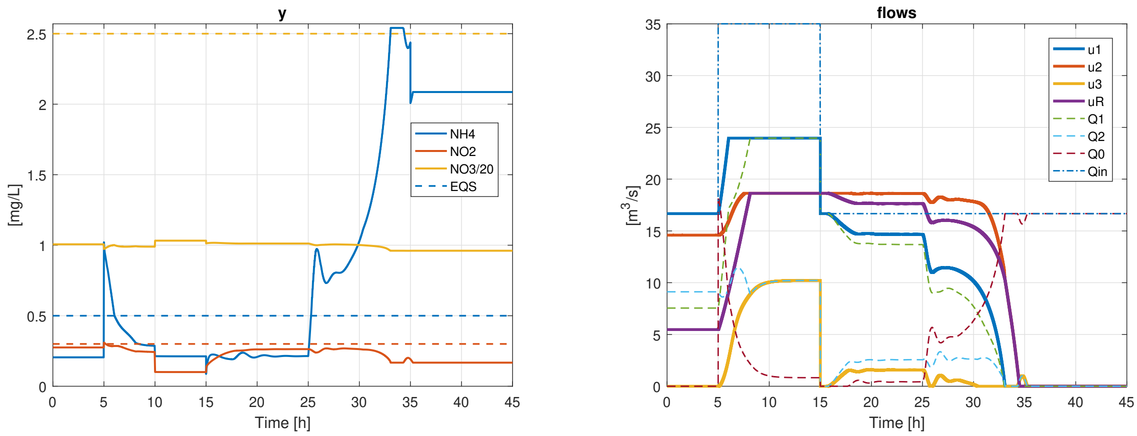


Figure 14: Proportional control with general gains, $k_1 = 1000$ and $k_2 = 500$. The evolution of the output agent y (left) and the flows (right)

g general	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	72	0	0	583	600	1255
$t_s\,[h]$	/	3.4	/	5.0	8.0	2.0	
oscillation	/	small	/	large	*unstable*	/	
$\tilde{EQS}\,[\%]$	0	0	0	386	317	0	
$\tilde{x}_1\,[\%]$	0	-58	-58	0	114	130	
$\tilde{x}_2\,[\%]$	0	-1	0	0	13	13	

Table 12: Benchmark for the proportional control with general gains (g general)

3.2.2.1 Parameters tuning

The system (Equation 15) is linearized around each equilibrium. Then the optimal matrix gain is computed with the MATLAB function `lqr()`[3]. Equation 23 becomes :

$$u = u - K \cdot E \tag{25}$$

Where $u = [u_1\, u_2\, u_3\, u_R]^T$, K is a 4x6 gain matrix and E is a 6x1 vector with the error of x_1 and x_2 relative to their equilibrium state (not normalised). The gain matrices (rounded to 1 decimal place) are K_{reg} and K_{rain} for the initial and the heavy rainfall equilibrium, respectively.

$$K_{reg} = \begin{pmatrix} 88.5 & -5.1 & -46.2 & 0 & -0.2 & -0 \\ 0 & 0 & 0 & 2.3 & -11.3 & 99.3 \\ 0 & -0 & 0 & 99.3 & -10.7 & -3.5 \\ -44.9 & 15.0 & -88.0 & 0 & 0.2 & -0 \end{pmatrix}$$

[3]See `Somers_24721400_2021_Appendix.zip: simulation/lin_sys.m`

$$K_{rain} = \begin{pmatrix} 41.7 & -76.8 & -61.3 & -5.6 & -33.0 & -3.7 \\ -0 & -0 & 0 & -12.9 & -8.6 & 98.6 \\ 0.1 & -0.1 & 0 & 97.6 & -18.0 & 11.2 \\ -17.6 & 60.1 & -78.9 & 3.5 & 20.3 & 2.2 \end{pmatrix}$$

In general, the obvious relations are confirmed: u_1 is important for the ammonium level in x_1 and dilute the nitrate inside the unit. u_2 is mainly related to the nitrate into x_2 and u_3 to the ammonium. u_R dilute the ammonium and the nitrate inside x_1 and increase its nitrite concentration. The flow rates have little impact on the unit they leave. Compared to K_{reg}, K_{rain} emphasizes the presence of high nitritelevels for heavy rainfall i.e. the values in the second and the fifth columns, corresponding to the nitrite concentration in x_1 and x_2, significantly increase. It also highlights the increased coupling of the flow rates: the limit between the first three columns (x_1) and the others (x_2) becomes less distinct e.g. the first row of K_{reg} is clearly cut between x_1 and x_2, the latter having only near zero coefficients.

As expected, the controller with these gains perform well when the disturbance is small. However the two equilibrium differ too much (i.e. there are large non-linearities in the system) for larger disturbance (as it is the case in this thesis). The control approach has good performance only when the gains and their associated state equilibrium (used to compute E) are both dynamically updated i.e. they are changed at t=5h and t=15h. This case is shown in Figure 16. Since the set point for heavy rainfall is case dependent and poorly characterised at the time of the event, this approach is limited. The eigenvalues of the linearized system have been computed with the two gain matrices. When the gain matrix of the other equilibrium is used for the linearized system, either its stability is reduced or eigenvalues become complex. Figure 15 shows the case where K_{rain} is used and the reference state equilibrium is the one for regular days. This control completely fails to stay below the EQS, even when the flow rate of the river is increased (t=10h).

The optimised control (Figure 16) adequately follows the set points given. Note that the equilibrium values for heavy rainfall computed in subsubsection 3.1.1 are above EQS when the Seine river is still at its initial state.

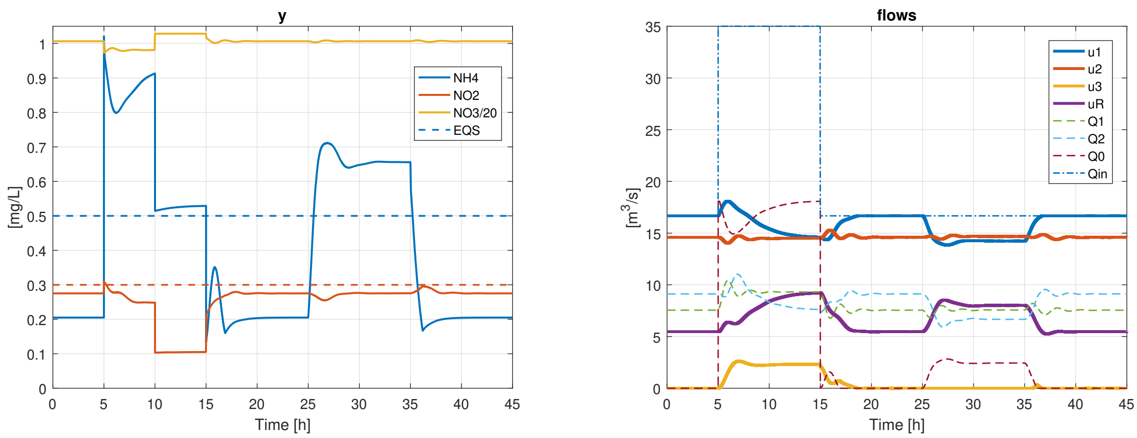


Figure 15: Evolution of the output agent y (left) and the flows (right) with K_{rain} and the initial state equilibrium.

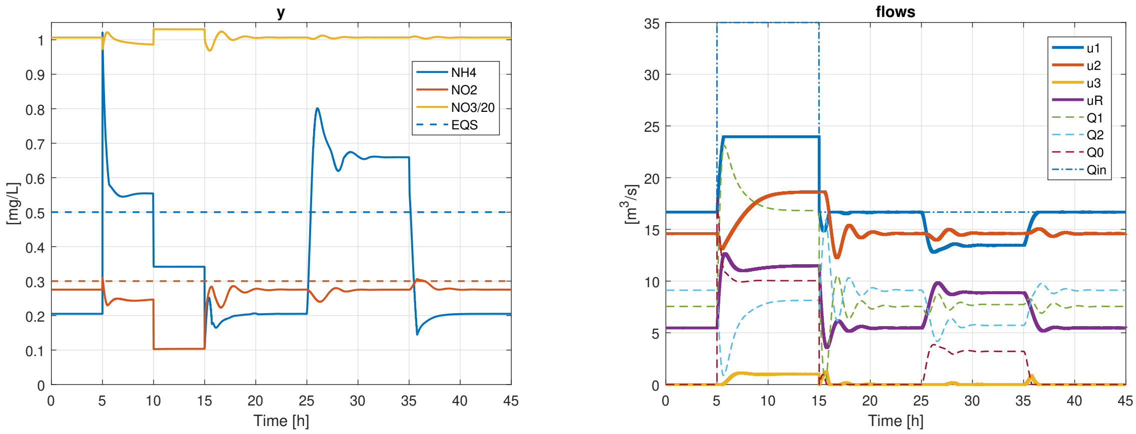


Figure 16: Evolution of the output agent y (left) and the flows (right) with optimised gains and set points dynamically adapted to the conditions

3.2.3 Proportional control with collaboration between the agents

The Equation 26 is studied in this section, and more specifically the impact of various communication links, represented by the term a_{aj}. It is equivalent to Equation 24, with the separation of a unique proportional gain into different gains k_1, k_2 and k. When k is zero, the equation becomes the simple proportional control studied in subsubsection 3.2.2.

$$
\begin{aligned}
u_1 &= u_1 - k_1 \cdot e_{x_1} - k \cdot \sum_{j=1}^{n} a_{1j}(e_{x_1} - e_{x_j}) \\
u_2 &= u_2 - k_2 \cdot e_{x_2} - k \cdot \sum_{j=1}^{n} a_{2j}(e_{x_2} - e_{x_j}) \\
u_3 &= u_3 - k_2 \cdot e_{x_2} - k \cdot \sum_{j=1}^{n} a_{2j}(e_{x_2} - e_{x_j}) \\
u_R &= u_R - k_1 \cdot e_{x_1} - k \cdot \sum_{j=1}^{n} a_{1j}(e_{x_1} - e_{x_j})
\end{aligned}
\tag{26}
$$

3.2.3.1 Full communication

The adjacency matrix $\mathcal{A}$ is simply a 5x5 matrix fulled with ones, except on the diagonal. So every agent is connected to the others. The gains are $k_1 = 300$, $k_2 = 150$ and $k = 300$. They are reduced compared to subsubsection 3.2.2 to compensate the influence added by the collaboration term i.e. if the gains are kept unchanged the overall proportional gain would be larger.

The simulation for the output agent y and the flow rates is displayed in Figure 17 and the benchmark is given in Table 13. The performance regarding EQS is the best so far (50 at t_1). Compared to the simple proportional control, the settling time is better in general but has similar settling time for the first disturbance at t=5h. The oscillations are not significantly different, except at t_4. Indeed the communication between the agent allows to prevent the crash of the flow rates observed during t_4 for the simple proportional control. Note that although the output concentration are almost identical at t=0h end at the end (t=45h), the distribution of the flow rates has significantly changed. In particular

u_3 is non zero and the recirculation flow rate u_R is in the same order as Q_{in}. However the resulting distribution alone cannot be exploited by the WWTP operator since other parameters such as the amount of reactant are not taken into account.

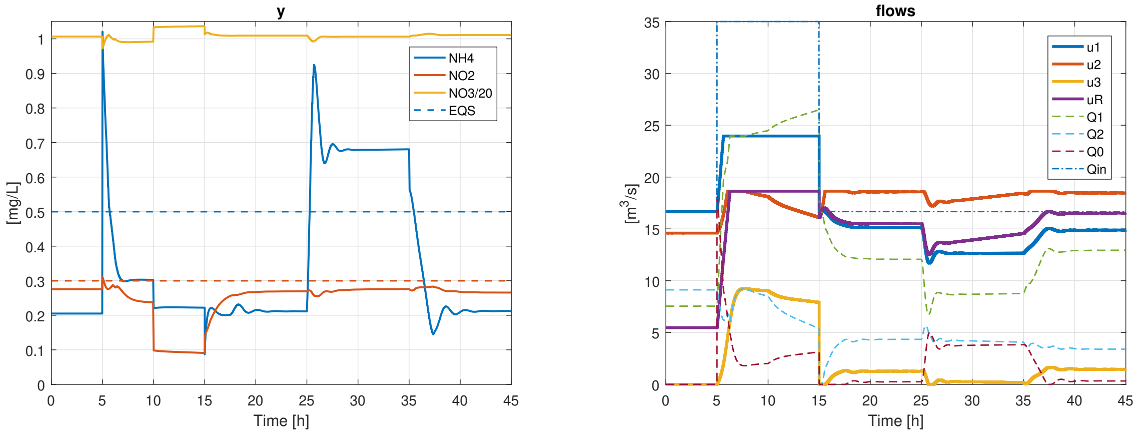


Figure 17: Formation control with full communication, $k_1 = 300$ and $k = 300$. The evolution of the output agent y (left) and the flows (right)

collab full	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	50	0	0	583	29	662
$t_s\,[h]$	/	3.4	/	3.8	2.0	3.0	
oscillation	/	small	/	large	large	medium	
$\tilde{EQS}\,[\%]$	0	0	0	36	0	0	
$\tilde{x}_1\,[\%]$	0	-58	-58	0	5	0	
$\tilde{x}_2\,[\%]$	0	-5	-8	0	5	0	

Table 13: Benchmark for the formation control with full communication

3.2.3.2 Communication based on physical flows

In this section, the adjacency matrix $\mathcal{A}$ reflects the physical connections e.g. the flow rate u_2 from the agent x_1 into the agent x_2. That is:

$$\mathcal{A} = \begin{array}{c} \\ x_1 \\ x_2 \\ v_{in} \\ v_{seine} \\ y \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & u_R & u_1 & 0 & 0 \\ u_2 & 0 & u_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ Q_1 & Q_2 & Q_0 & Q_{seine} & 0 \end{pmatrix} \end{array}$$

The in-degree of v_{in} and v_{seine} are zero, as the out-degree of y. The gain k_1 is still 300, but k is set to 0.0104 since the flow rates are in the order of tens of thousands $[m^3/h]$. The simulation for the output agent y and the flow rates is displayed in Figure 18 and the benchmark is given in Table 14.

Since the flow rates and the errors of the agents are variable, the overall proportional gain (which is a combination of k_1, k_2, k and a_{aj}) differs from the one with full communication (where a_{aj} is 1 $\forall a \neq j$). So the difference in performance could be due to a difference

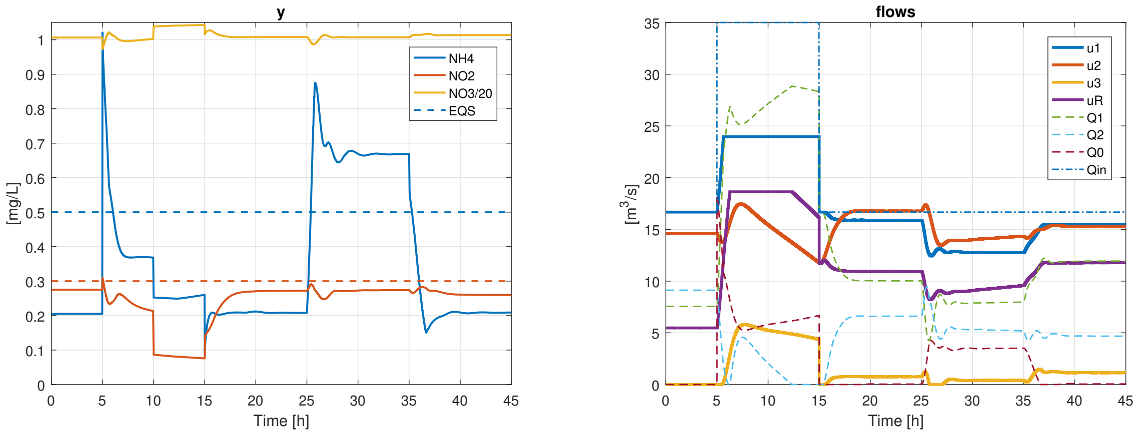


Figure 18: Formation control with communication based on physical flows, $k_1 = 300$ and $k = 0.0104$. The evolution of the output agent y (left) and the flows (right)

collab flows	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	67	0	0	577	18	662
$t_s\,[h]$	/	4.2	/	3.3	3.2	2.5	
oscillation	/	medium	/	small	large	medium	
$\tilde{EQS}\,[\%]$	0	0	0	34	0	0	
$\tilde{x}_1\,[\%]$	0	-56	-57	0	4	0	
$\tilde{x}_2\,[\%]$	0	-23	-21	0	3	0	

Table 14: Benchmark for the formation control with communication based on physical flows

in the proportional gain rather than the change in the communication weights. To this end, k is calibrated so that the saturation of the flow rate u_1 is reached at the same time. Tuning k to have the same $\#EQS$ at t_1 is possible, but the dynamics are no longer similar (it has more oscillations as a result of a higher gain).

If the mirror adjacency matrix $\hat{\mathcal{A}}$ is used, the communication with the output agent y is added. The gain k is reduced to 0.0068. There is no mention to the case where there is bidirectional directed communication between two agents but with different weight (as u_2 and u_R) in the literature. It is decided to keep u_1 and u_R apart here (another decision could be to form an undirected edge with an unique gain $u_2 + u_R$, but this lose the discrimination between the flows).

Again, the simulation for the output agent y and the flow rates is displayed in Figure 19 while the benchmark is given in Table 15. The results are very similar to the control based on physical flows (without the communication with the output agent y), but with worse settling time in general. Indeed when the output concentrations are below the EQS, e_y is zero so the reduced gain k increases the settling time during t_3 and t_5.

While the scenario with full communication and a unitary gain has clearly the best performance ($\#EQS = 50$ for t_1), the introduction of weights based on the physical connections has little effect. Weights set to the inverse of the flow rates have been investigated but it is worse, as expected. Indeed it is reasonable to think that the larger the flow rates are, the larger the control inputs should be.

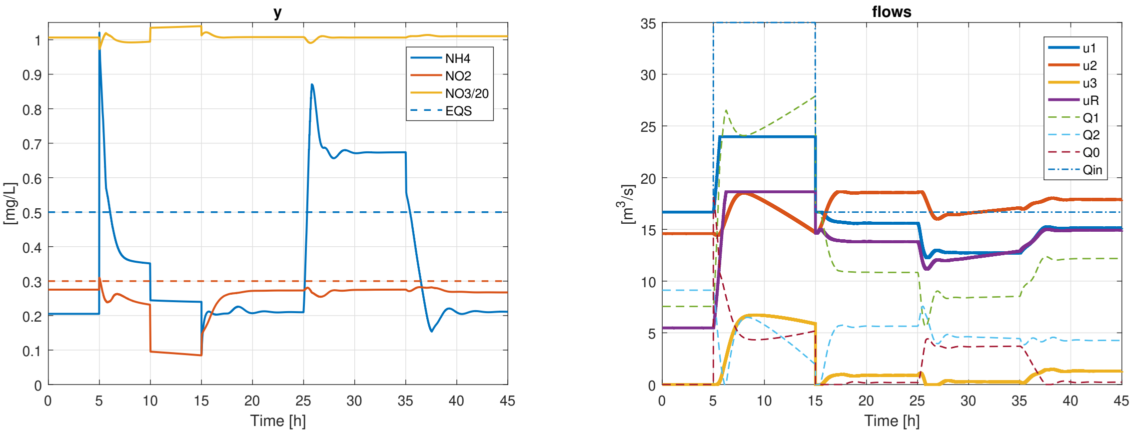


Figure 19: Formation control with communication based on mirror physical flows, $k_1 = 300$ and $k = 0.0068$. The evolution of the output agent y (left) and the flows (right)

collab mirror flows	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	68	0	0	579	26	673
$t_s\,[h]$	/	4.2	/	3.7	3.2	3.9	
oscillation	/	medium	/	medium	large	medium	
$\tilde{EQS}\,[\%]$	0	0	0	35	0	0	
$\tilde{x}_1\,[\%]$	0	-56	-56	0	5	0	
$\tilde{x}_2\,[\%]$	0	-18	-17	0	4	0	

Table 15: Benchmark for the formation control with communication based on mirror physical flows

In general there are noticeable differences in the flow rates evolution and distribution among the different adjacency matrices tested, but this highlights the many possibilities in the control of the system rather than providing informative data.

Finally, the unitary gain equivalent to the adjacency matrix for physical connections has been tested. This is:

$$\mathcal{A} = \begin{array}{r} \\ x_1 \\ x_2 \\ v_{in} \\ v_{seine} \\ y \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

Compared to the case with full communication, the performance is very similar thanks to the tuning of the gain k. However the settling times are larger during t_1 and t_4 (when $e_y > 0$) since there is no communication with the agent y. The dynamics of the output agent y are also slightly degraded on the nitrite, which comes closer to the EQS (but stays below it).

Overall the performance is better with unitary gains ($\#EQS \approx 50$ for t_1) than with the weighted adjacency matrix ($\#EQS \approx 68$ for t_1). So the weighted approach will be discarded in the next section.

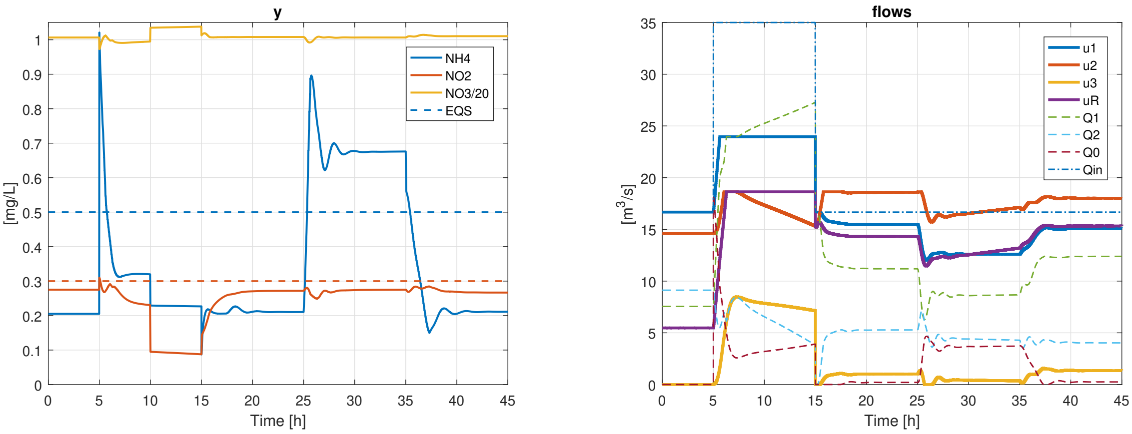


Figure 20: Formation control with unitary communications based on physical connections, $k_1 = 300$ and $k = 615$. The evolution of the output agent y (left) and the flows (right)

collab flows	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	49	0	0	579	26	654
$t_s\,[h]$	/	4.1	/	3.9	4.5	3.0	
oscillation	/	medium	/	medium	large	medium	
$\tilde{EQS}\,[\%]$	0	0	0	35	0	0	
$\tilde{x}_1\,[\%]$	0	-57	-57	0	5	0	
$\tilde{x}_2\,[\%]$	0	-10	-10	0	5	0	

Table 16: Benchmark for the formation control with unitary communications based on physical connections

3.2.3.3 Limited communication between agents

In this section, the impact of the communication between agents is studied for a wide range of communication configurations. To this end, a script runs the simulation with all the possible adjacency matrices and store their $\#EQS$ metric for t_1, t_5 and for the total time. Communication are not weighted (either set to 1 or 0). This gives $2^8 = 256$[4]different realisations since there are 8 communications studied (4 to x_1 and 4 to x_2).

The global results are displayed in Figure 21, where each of the 256 realisations is measured with three EQS metrics ($\#EQS$ for t_1 in blue, for t_5 in red and for the total in yellow) and with the algebraic connectivity, shown in the form of $1 - \lambda_2$ to better see the correlation with $\#EQS$ (see Appendix D for the computation of the algebraic connectivity metric). The same data is presented in the two graphs, but they are sorted differently: in the ascending order for the $\#EQS$ metric for t_1 on the left, and for the algebraic connectivity on the right. Note the total $\#EQS$ in yellow can be lower than for t_5 because it has been scaled down for clarity. Similarly the algebraic metric has been scaled up.

There are two main observations. First the algebraic correlates with the performance measured through $\#EQS$. But for the global trend only, as there is an important case to case variation (see Figure 21). Also there is higher variance in the performance when

[4]Intuitively, there are 8 bits (one bit for each communication) and every realisation from 00000000 to 11111111 is tested. Combination can also be used to find the number: $\sum_{k=0}^{n} nCk = 2^n$, with $n = 8$.

the algebraic connectivity is low, and this variance reduces as the algebraic connectivity increases. Indeed when there is only one communication, performance can vary considerably depending on the communication in use. Second there exist different thresholds in the performance (see the two blue plateaus in the rightmost part of Figure 21 or the big red plateau). This suggests that a minimal set of communication is necessary to achieve significantly better performance than without communication at all (i.e. better performance than with the simple proportional control). The larger plateau gap seen for t_5 and for the total (in red and yellow, respectively) is due to the instability of the control at the end of the simulation, as shown in Figure 14. Two different areas can be identified too. At around two thirds of the realisations, the $\#EQS$ metrics start to grow faster (the frontiers of the areas vary slightly with the $\#EQS$ metric chosen). The results presented here are for the $\#EQS$ metric for t_1 since it is the time window for which we are the most interested in (i.e. heavy rain transition).

The first (rightmost) plateau surprisingly does not contain the case with no communication at all. Indeed communication should increase the performance of the system so the worst case should be the case without any communication. This suggests that the communications present in this set should have negative weights. All these communications only concern x_2. In particular the communication from v_{in} and y are the most represented. The communication from v_{seine} is present in half the cases.

The second plateau again concerns only communications with x_2, and it contains the case with no communication at all. This time the communication from x_1 is the most represented and communication from v_{seine} is present in half the cases. It is harder to conclude on these weights except that they are of a lesser importance.

The break around two thirds of the realisations includes the communication to x_1, but only from y or x_2 (at the exception of cases with only one communication link). This suggests that these are the least important concerning x_1. There are also the cases with only one communication link to x_1. They are, in decreasing performance order : v_{in} (98), y (102), v_{seine} (113), x_2 (141). Otherwise they are only communications to x_2.

The best five cases scores 49 ($\mathcal{A}_1$ and $\mathcal{A}_2$), 50 ($\mathcal{A}_3$) or 51. They are:

$$\mathcal{A}_1 = \begin{array}{c} \\ x_1 \\ x_2 \end{array}\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array} \qquad \mathcal{A}_2 = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

$$\mathcal{A}_3 = \begin{array}{c} \\ x_1 \\ x_2 \end{array}\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

$$\mathcal{A}_4 = \begin{array}{c} \\ x_1 \\ x_2 \end{array}\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \qquad \mathcal{A}_5 = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

$\mathcal{A}_3$ is the case with full communication studied in paragraph 3.2.3.1. As expected from the previous observations, all the communication toward x_1 are active, while the communication to x_2 from y or x_1 are not always active. Note that the communication from v_{in} to x_2 is always present even if it was previously identified as one of the worst communication link. This highlights the need of considering the whole network instead of analysing the communication in an individual way.

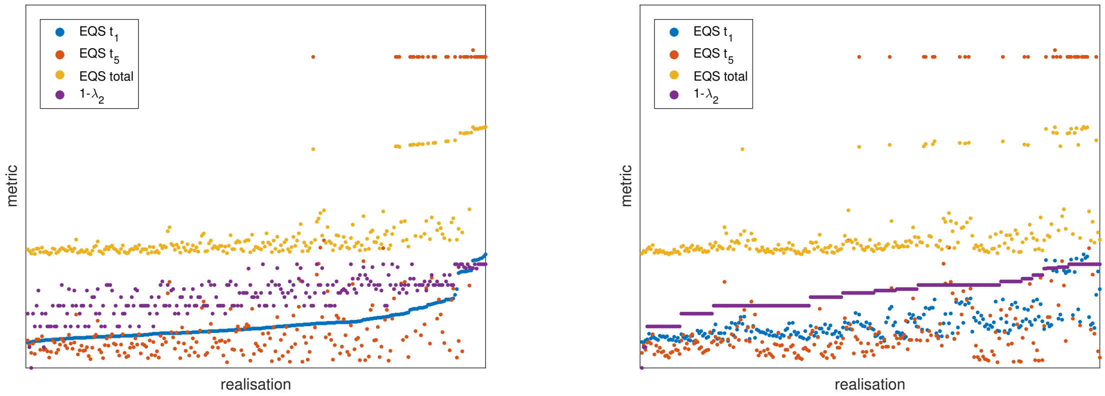


Figure 21: Performance of the simulation for all the possible adjacency matrices. Sorted following the $\#EQS$ metric for t_1 (left) and the algebraic connectivity (right)

3.2.3.4 Communication on a specific concentration

Another possibility for the adjacency matrix is to use it at full scale. That is developing the state of each agent to obtain a 15x15 matrix. It allows to chose the weight for each agent and for each concentration.

The main point here is only to test if the performance of the controller remains satisfying when only the ammonium concentrations are measured. Indeed the nitrite sensors are often either nonexistent or poorly calibrated and the nitrate sensors have not been investigated much due to the nitrate concentrations being well under the norms. The case used is the one with full communication ($\mathcal{A}_3$).

The simulation for the output agent y and the flow rates are displayed in Figure 22 and the benchmark is given in Table 17. The settling times of the other concentrations (nitrite and nitrate) are very large, if they settle at all before the next step. The control on x_2 is clearly poorer than usual (x_2 does not process the ammonium at all). The global performance regarding the EQS are degraded. However it recovers better than usual at t_5 (where the disturbance only concerns the ammonium).

This highlights that the simulation can take into account the absence of (reliable) sensors in the WWTP. It also shows that the weights of the different concentrations could be tuned (in a centralised manner, i.e. by the operator of the WWTP) to adapt the network to particular conditions e.g. in order to recover from an extreme event.

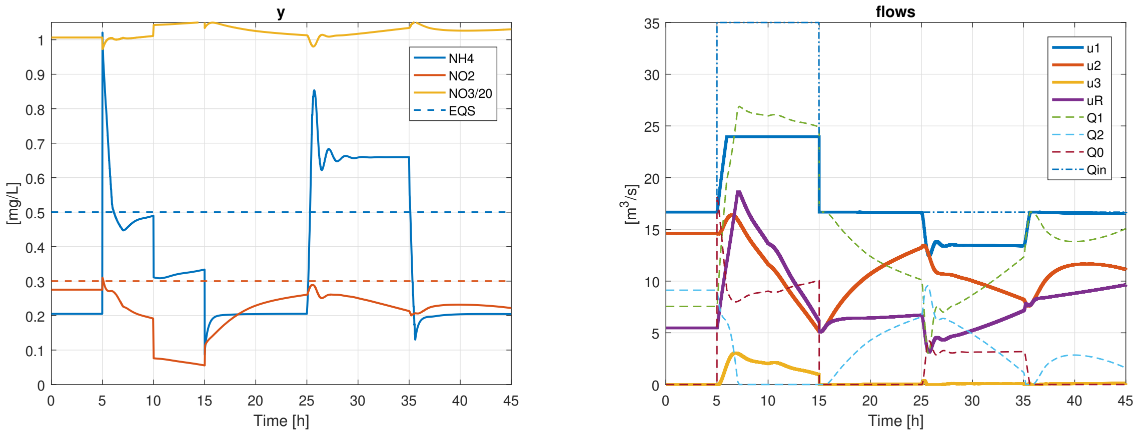


Figure 22: Formation control with full communication but only with the ammonium concentration, $k_1 = 300$ and $k = 300$. The evolution of the output agent y (left) and the flows (right)

collab flows	t_0	t_1	t_2	t_3	t_4	t_5	total
$\#EQS\,[min]$	0	73	0	0	581	8	662
$t_s\,[h]$	/	2.7*	/	1.4*	2.3*	1.0*	
oscillation	/	medium	/	/	large	small	
$\tilde{EQS}\,[\%]$	0	0	0	32	0	0	
$\tilde{x}_1\,[\%]$	0	-55	-61	3	18	7	
$\tilde{x}_2\,[\%]$	0	-45	-44	-24	-15	-25	

Table 17: Benchmark for the formation control with full communication but only with the ammonium concentration. * settling time of the ammonium concentration

3.2.4 Agent-based control

It is interesting to compare the performance in the previous sections, built on the multi-agent framework, to a agent-based control approach built on the knowledge of the specific problem.

First the controllable flow rates of each agent are studied from the perspective of each agent e.g. the nitrification unit sees an ammonium source (v_{in}) and a nitrite source (x_2). If the ammonium level rises too much, the flow rate from v_{in} should be lowered. These insights are shown in Figure 23. The agent with the disturbance could also modify the weights of each concentration.

This leads to the following control approach:

$$\begin{aligned}
u_1 &= u_1 \quad -k\cdot e_{x_1,1} - k\cdot e_{x_1,3} \qquad +k\cdot e_{v_{in},1} + k\cdot e_{y,1} - k\cdot e_{y,3} \\
u_2 &= u_2 \qquad - k\cdot e_{x_2,2} - k\cdot e_{x_2,3} \qquad + k\cdot e_{y,3} - k\cdot e_{y,2} \\
u3 &= u_3 \qquad - k\cdot e_{x_2,1} \qquad +k\cdot e_{v_{in},3} + k\cdot e_{v_{in},2} + k\cdot e_{y,1} - k\cdot e_{y,2} \\
u_R &= u_R \qquad -k\cdot e_{x_1,2} \qquad + k\cdot e_{y,2} - k\cdot e_{y,3}
\end{aligned} \tag{27}$$

Note that some gains are negative. The graphs are given in Figure 24. k is tuned with the saturation time of u_1, see paragraph 3.2.3.2. The results are comparable to the one

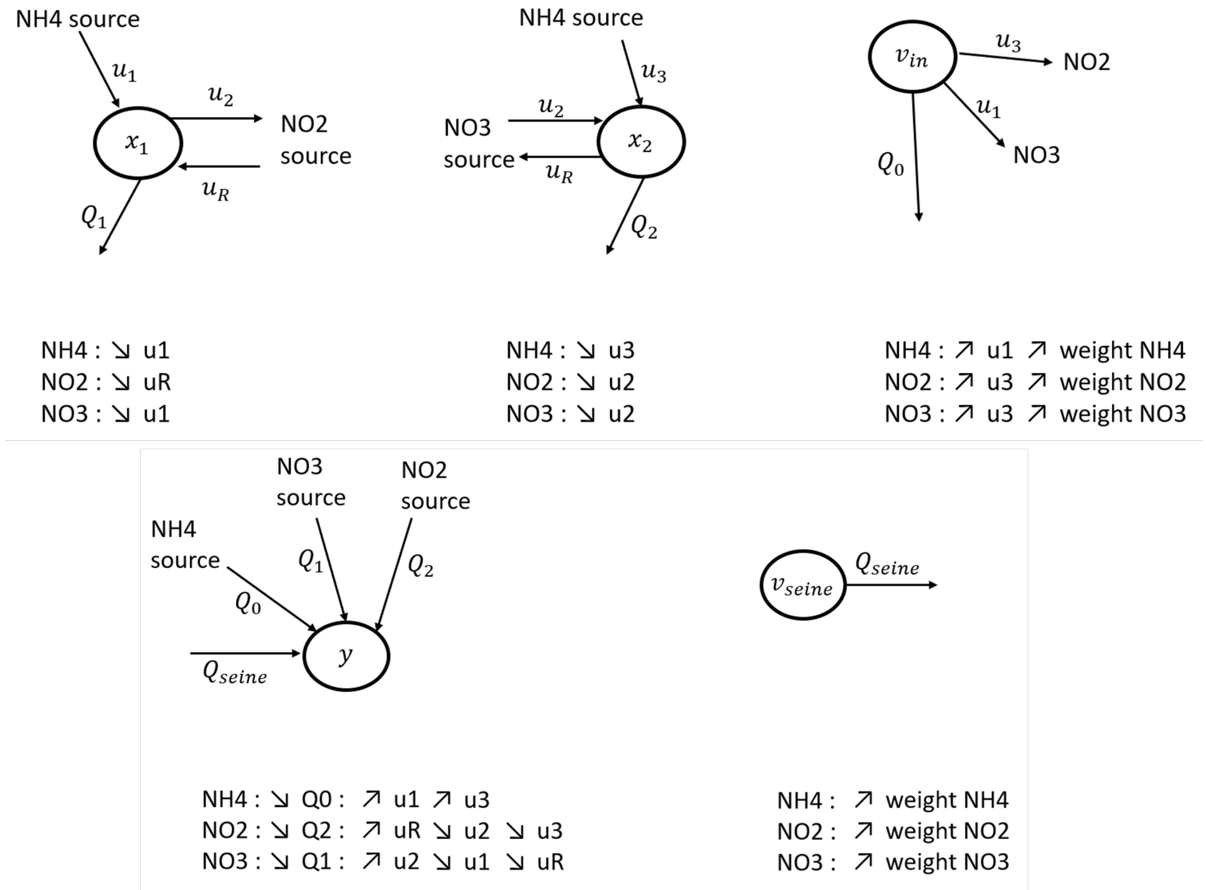


Figure 23: Evolution of the controllable flow rates from the perspective of each agent

achieve with the multi-agent framework. The $\#EQS$ metric is poorer due to the increase in u_3 and Q_2, leading to more nitrite at the output. The gains must be adapted by trial and error, but could become inadequate for other disturbances.

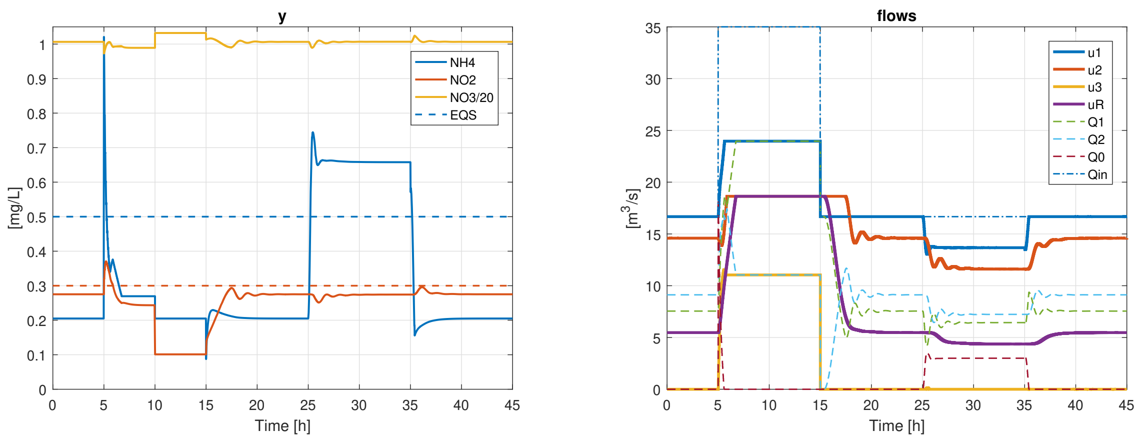


Figure 24: Agent-based control, $k = 1200$. The evolution of the output agent y (left) and the flows (right)

3.3 Discussion

This chapter highlights the general link between the number of communications and the performance of the system. It also shows that the addition of collaboration between agents can stabilise the simple proportional control. Furthermore, the communications can easily be adapted to a specific scenario e.g. if the nitrite sensor of a unit is down, this specific communication can be set to 0.

Four points of discussion come out of this section:

- the communication between the agents
- the impact of the communications and the tuning of their weights
- the optimisation of the gain in the context of a nonlinear application
- a limitation on the results

In general the communication with the nitrification unit is more impactful than with the denitrification unit. Some specific communications seem to play a small role. Some of them even lead to a worse performance than without any communication, but they are not always the one discarded in the best solutions (see $\mathcal{A}_1$ and $\mathcal{A}_2$) e.g. the communication from v_{in} to x_2. Replacing the missing communication(s) in these best solutions with a negative one (i.e. putting -1 instead of 0) degraded the performance ($\#EQS \approx 55$ for t_1). This highlights the complex behaviours that can emerge from the collaboration of multiple agents. The impact of a specific communication is not the same when it is isolated or when it is part of a network.

Although the physical intuition is usually used to tune the communication weights, it has shown poor performance for the simulation of the nitrogen removal of SAV. Indeed using the flow rates as the communication weights lead to poorer performance than with unitary gains. In fact the weighting process for MAS is not well characterised. Weighting strategies to achieve average consensus of a simple integrator are still the topic of recent publications [22].

Nonlinear systems are often handled via different setpoints where the system is linearized. Thus the well established toolbox for linear systems can be used around these setpoints. So optimised gains can be computed (see K_{typ} and K_{rain}). However these gains are interesting only in a certain range around the setpoint. If they are used in the collaboration control, it will inevitably needs to track the setpoint which is not the objective (else we would use linear control).

A agent-based control using knowledge of the specific application has been tested. But in the end it also relies on setpoints as the input knowledge is reasoned on the basis of different known scenarios.It thus also misses the issue for the transition between each scenario.

Finally, although physical constraints are respected, the control of the flow rates is idealised. For example it does not take into account the consumption of reactant of the nitrification and denitrification units, which could possibly lead to unacceptable costs for the operator of the WWTP. Thus the results cannot be transposed as it is into guidelines for SAV, they should be examined by the operator first.

4 Discussion

In this section, different multi-agent design approaches for SAV are discussed. The physical and regulatory constraints are then examined. Finally, the perspectives and the path for future work are presented in subsection 4.3.

4.1 Different designs

The system in Figure 25 pictures the design used in this thesis.

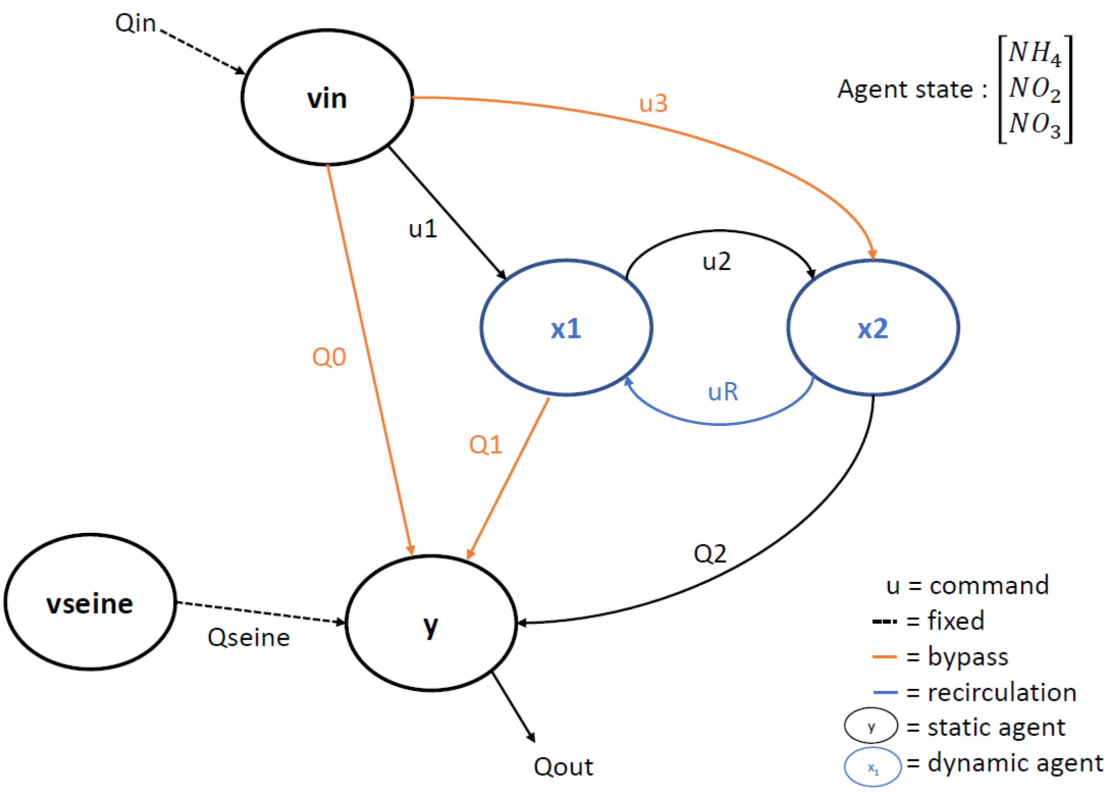


Figure 25: The system

The agents have been defined as the units. This is the most straightforward design choice. Another choice is the number of controllable flow rates. When four of them are fixed, all the other flow rates are determined due to the constant volume hypothesis (i.e. the volumes of water stay constant in the units). So for simplicity only four controllable flow rates have been chosen. They have been assigned to the two dynamical agents (the nitrification and denitrification units) as the other agents are static i.e. there is no (significant) microbial activity.

A different design could be to spread the four controllable flow rates between the agents so that they each have one (except for v_{seine} which physically is not controllable). Another design could stem from the problem-specific agent-based control given in subsubsection 3.2.4. The agents thus control all their flow rates, and the MAS is used to reach a consensus on the different flow rates. The consensus values would then need to respect the physical constraints (see subsection 4.2).

4.1.1 Concentration-based MAS

Different design choices can also be made for the agents themselves. One interesting possibility is to chose the concentrations as the agent. This choice is less straightforward and it is harder to predict how it will perform.

The idea is the following: each agent (ammonium, nitrite and nitrate) knows their possible flow paths through SAV, and each want to choose their best path i.e. to be below the EQS at the exit. The three different paths then need to be merged into one single path. A consensus algorithm would need to be used so that the three agents can agree on a single path.

However, even when the concentrations are decoupled, finding their best flow path is not trivial. Also the trade-off that need to be made by each agent to reach the consensus would probably be the main variable to tune, and it is not clear if a consensus law will be able to grasp the coupling between the concentrations (but maybe it does not need to).

4.2 Constraints

The constraints on the flow rates come from the flow rate conservation (the volume in the units is constant) and from the maximal capacity of the pipes. In both cases, the constraints are applied strictly in the simulation: when the maximal capacity of the pipe is reached, the controlled flow rates are saturated. When a flow rate increases, another must decrease immediately. While a strict application of the constraints is desirable in the first case to meet the real life capabilities of the system, some slack could be given to the hypothesis of constant volumes in the units. The variations in the volumes should be limited though.

In the previous section, a MAS design where all the flow rates must reach a consensus is proposed. Relaxing the constant volume hypothesis could be of particular importance for this consensus (unless the consensus is computed separately of the control of the flow rates).

4.2.1 Saturation due to constraints

A simple simulation is given in Figure 26 to illustrate a constrained consensus of three agents (x_1, x_2 and x_3) on the values for five flow rates ($u = [u_1\, u_3\, Q_0\, Q_1\, Q_2]$, the ones from the system in Figure 25). For each flow rate, the classical consensus is applied. But in order to constrain the final values of the flow rates, some extra terms are added: first $\sum_{inlet} - \sum_{outlet}$ so that the flow rate at the inlet of the WWTP ($u_1 + u_3 + Q_0$) is equal to the outlet ($Q_0 + Q_1 + Q_2$); then $\sum_{inlet} - Q_{in}$ so that the inlet flow rate Q_{in} is set properly. The consensus successfully produced flow rates values that do not require to be rectified by external constraints. The constraints are reflected on all the agents and the system will not be disturbed by an external application of the constraints

This gives the following equation based on the consensus algorithm notation (Equation 11), where the constraints are explicitly taken into account:

$$x_i = x_i - k\left(\sum_{j=1}^{n} a_{ij}(x_i - x_j) + (\sum_{inlet} x_i - \sum_{outlet} x_i) + (\sum_{inlet} x_i - Q_{in})\right)$$

The values used in Figure 26 are $k = 0.1$, $n = 3$, $x_1 = [2\,8\,9\,11\,6]$, $x_2 = [0\,10\,5\,10\,5]$, $x_3 = [0\,7\,9\,7\,9]$. $\mathcal{A}$ is the full communication adjacency matrix. The value for a flow rate is in the same color for each of the three agents. The agents find an agreement for each flow rate while respecting the constraints (inlet flow rate = outlet flow rate = 16, in red).

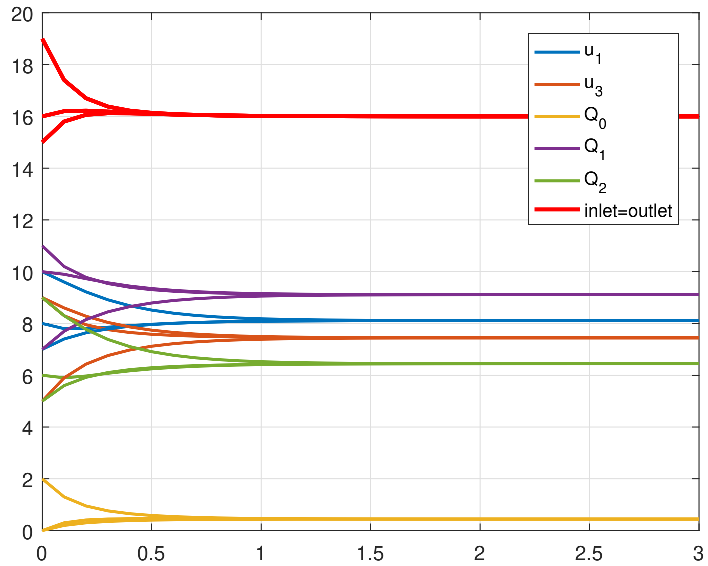


Figure 26: Simulation of a simple set of agents with constrained values

The simulation explicitly takes the constraints into the control law. There are thus taken into account in a continuous way rather than being applied externally. This allows among other things to prevent instabilities as the constraints no longer act as external disturbances. It is not really a concern when the system is stable and with a proportional control. However a controller with an integral action absolutely needs to take the constraints into account or to have good anti-windup.

The EQS constraint could be explicitly taken into account too.

4.3 Perspectives

The thesis is a first step into the study of the MAS framework for the control of the nitrogen removal of SAV. Four main directions can be taken in future work.

First, different MAS designs could be explored, as discussed in subsection 4.1. The designs could be ranked based on their ease of implementation or their robustness to losses in communication for instance.

Second, the MAS should be extended, at least to add the pre-denitrification unit introduced in 2017 in SAV. With the addition of this unit into the simulation and more detailed data from the WWTP (e.g. hourly measures of the concentrations for a regular day on for a heavy rainfall event), the output concentrations could be compared with the real ones. If this produces interesting results in the eye of SIAAP, a MAS for the entire plant could be investigated.

Another direction for future work is the addition of an integral action. Indeed a proportional control cannot address the steady state error in all cases. Since the controlled values can be saturated, an anti-windup mechanism should be implemented alongside the integral action. Because the steady state values are controlled, the reference steady state selected for each agent should be chosen carefully to obtain the best results.

Finally, the optimisation of the gains and the robustness of the control action regarding communication losses or uncertainties could be investigated. The agents could also communicate other information such as a signal to increase or decrease the weights on some concentrations i.e. use gain scheduling.

The question of the amount of reactant used for the units should also be addressed, to assess if it should be implemented in the model or if it can be decoupled from the control of the flows (by restraining the magnitude of the concentration variations inside the units for instance). In all cases, a closer contact with the SIAAP and their expertise in the wastewater treatment could increase the relevance of this work and provide guidance on the directions that the operator of the WWTP would want to explore.

Final conclusions

The aim of the master thesis was to study a multi-agent control strategy to tackle the problem of nitrogen removal for the Seine aval wastewater treatment plant during heavy rainfall. The nitrogen removal system studied is formed by the nitrification unit and denitrification unit modeled in [41].

To start, the state of the art for wastewater treatment has been covered, more particularly the nitrogen removal part, and the associated European legislation have been reviewed. The WWTP used for the thesis, Seine aval, has been presented and the parameters needed for the work have been extracted from the literature. Relevant tools from the multi-agent framework such as the graph theory and the consensus problem have been presented.

Then the multi-agent system chosen to model the nitrogen removal of Seine aval has been designed. It is characterised for the conditions of regular days based on data from the WWTP operator (SIAAP). It has been transposed to a MATLAB simulation and the various control approaches have been presented and evaluated in these conditions.

The main part of the thesis focused on the disturbances caused by heavy rainfall. An adequate disturbance and a benchmark to evaluate the performance of the control approaches have been designed. Then the various control approaches tested in regular conditions have been simulated with the disturbance. The study of the different interactions between each agent and the impact on the control performance has been conducted and discussed.

The outcomes of the thesis are as follows:

First the effective exploration of the application of the multi-agent framework to a wastewater treatment plant.

Second the development of a simulation tool to analyse different multi-agent control strategies for the nitrogen removal, and to explore the impact of the communication between the agents.

Third the characterisation of the importance of some specific communications. Mainly, the nitrification unit is more sensitive to the communications so the efforts to control the nitrogen removal should be placed on this unit. Also the communication between each agent is not necessary because a configuration with the 6 adequate communications has a similar performance as the configuration with full communication (i.e. 8 communications).

Finally a discussion for future work has been done. Four directions have been identified: the exploration of different MAS designs, the extension of the current MAS, the addition of an integral action to the controller, and the use of adaptive gains.

Bibliography

[1] Anokina Shalimoon. "Cooperative Control of Multi-Agent Systems: Consensus, Flocking and Formation Control". EN. MA thesis. California state university, Northridge, 2018.

[2] Ayoub Ben Ayed. *Mid-term evaluation report Analysis and Control of interconnected systems*. EN. Tech. rep. Oct. 2020.

[3] Juan Baeza et al. "A distributed control system based on agent architecture for wastewater treatment". In: *Computer-Aided Civil and Infrastructure Engineering* 17.2 (2002), pp. 93–103.

[4] G Bastin and Denis Dochain. *On-line estimation and adaptive control of bioreactors*. EN. Elsevier, 1990.

[5] Ayoub Ben Ayed. "Real-time optimisation of a wastewater treatment network". EN. MA thesis. UCLouvain, 2018. URL: http://hdl.handle.net/2078.1/thesis:14881.

[6] Sif Eddine Benahmed. "Distributed Cooperative Control for DC Microgrids". 2021LORR0056. PhD thesis. 2021. URL: http://www.theses.fr/2021LORR0056/document.

[7] Paul Bogg, Ghassan Beydoun, and Graham Low. "When to use a multi-agent system?" EN. In: *Intelligent agents and multi-agents systems*. Hanoi: Springer, 2008. ISBN: 3540286535.

[8] Ming Cao and Chai Wah Wu. "Topology design for fast convergence of network consensus algorithms". In: *Proceedings - IEEE International Symposium on Circuits and Systems* (2007), pp. 1029–1032. ISSN: 02714310. DOI: 10.1109/iscas.2007.378145.

[9] Richard Carr. "Excreta-related infections and the role of sanitation in the control of transmission". In: *Water Quality: Guidelines, Standards and Health*. 2001. Chap. 5, pp. 89–113. ISBN: 1 900222 28 0.

[10] Fei Chen and Wei Ren. "On the control of multi-agent systems: A survey". In: *Foundations and Trends in Systems and Control* 6.4 (2019), pp. 1–164. ISSN: 23256826. DOI: 10.1561/260-000-0019.

[11] Kehua Chen et al. "Optimal control towards sustainable wastewater treatment plants based on multi-agent reinforcement learning". In: *Chemosphere* 279 (2021), p. 130498. ISSN: 0045-6535. DOI: https://doi.org/10.1016/j.chemosphere.2021.130498. URL: https://www.sciencedirect.com/science/article/pii/S0045653521009681.

[12] European Commission. *Introduction to the EU Water Framework Directive*. [Online; accessed 27-May-2021]. 2021. URL: https://ec.europa.eu/environment/water/water-framework/info/intro_en.htm.

[13] European Commission. *Urban waste water: improvement in collection and treatment across the EU helps reducing pollution in the environment*. [Online; accessed 27-May-2021]. 2020. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1563.

[14] Justine Fiat. "Analyse and modelling of N2O emissions from full-scale tertiary nitrifying BAFs - public PhD defence". EN. PhD thesis. INSA Toulouse, IRSTEA, 2019.

[15] Freepik. 2021. URL: https://www.flaticon.com/.

[16] Dina Irofti. "An anticipatory protocol to reach fast consensus in multi-agent systems". In: *Automatica* 113 (2020), p. 108776. ISSN: 00051098. DOI: 10.1016/j.automatica.2019.108776. URL: https://doi.org/10.1016/j.automatica.2019.108776.

[17] Raphaël Janssens, Hadrien Maloux, and Marie Lesenfants. "National surveillance of SARS-COV-2 in wastewater". 2021.

[18] Ulf Jeppsson and G Olsson. "A General Description of the IAWQ Activated Sludge Model No. 1". EN. In: *Department of Industrial Electrical Engineering and Automation* 1 (1996), p. 444.

[19] Rick Katz. "Statistics of extremes in hydrology". EN. In: *Advances in Water Resources* 25 (2002), pp. 1287–1304.

[20] Mehran Mesbahi and Magnus Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. STU - Student edition. Princeton University Press, 2010. ISBN: 9780691140612. URL: http://www.jstor.org/stable/j.ctt1287k9b.

[21] Ministerial Order. *Arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement*. FR. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000021865356/. 2010.

[22] Furugh Mirali, Antonio Mendez Gonzalez, and Herbert Werner. "First-Order Average Consensus for Cooperative Control Problems Using Novel Weighting Strategies". EN. In: *IFAC-PapersOnLine* 50.1 (2017), pp. 14302–14307. ISSN: 24058963. DOI: 10.1016/j.ifacol.2017.08.1865. URL: https://doi.org/10.1016/j.ifacol.2017.08.1865.

[23] MOCOPEE. "Colloque annuel ASTEE". FR. In: June 2016, pp. 1–12.

[24] MOCOPEE. *Mocopee presentation*. [Online; accessed 27-May-2021]. 2021. URL: http://www.mocopee.com/.

[25] Jianjun Ni et al. "A multiagent dynamic assessment approach for water quality based on improved Q-learning algorithm". In: *Mathematical Problems in Engineering* 2013 (2013). ISSN: 1024123X. DOI: 10.1155/2013/812032.

[26] Reza Olfati-Saber and Richard M. Murray. "Consensus problems in networks of agents with switching topology and time-delays". In: *IEEE Transactions on Automatic Control* 49.9 (2004), pp. 1520–1533. ISSN: 00189286. DOI: 10.1109/TAC.2004.834113.

[27] Noemie Olive. *Canal pool shows Paris could have Seine swimming for 2024 Olympics, mayor says*. Ed. by Reuters. [Online; accessed 11-June-2021]. 2017. URL: https://www.reuters.com/article/us-olympics-2024-paris-swimming-idUSKBN1A323M.

[28] World Health Organization. *Diarrhoeal disease*. [Online; accessed 11-June-2021]. 2017. URL: https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease.

[29] Par Patrick Philipon. "Intelligence artificielle et supervision : l 'enjeu de demain ..." FR. In: *L'Eau, l'Industrie, les Nuisances (EIN)* 431 (Apr. 2020).

[30] Craig W. Reynolds. "Flocks, Herds and Schools: A Distributed Behavioral Model". In: *SIGGRAPH Comput. Graph.* 21.4 (Aug. 1987), pp. 25–34. ISSN: 0097-8930. DOI: 10.1145/37402.37406. URL: https://doi.org/10.1145/37402.37406.

[31] *River discharge - hydrographs*. [Online; accessed 30-May-2021]. 2017. URL: https://8777928435989433391.weebly.com/river-discharge--hydrographs.html.

[32] C. E. Robles-Rodriguez et al. "A simple model of wastewater treatment plants for managing the quality of the Seine River". EN. In: *IFAC-PapersOnLine* 51.18 (Jan. 2018), pp. 880–885. ISSN: 24058963. URL: https://doi.org/10.1016/j.ifacol.2018.09.236.

[33] C. E. Robles-Rodriguez et al. "Management of an integrated network of wastewater treatment plants for improving water quality in a river basin". In: *IFAC* 52.1 (2019), pp. 358–363. ISSN: 24058963. DOI: 10.1016/j.ifacol.2019.06.088. URL: https://doi.org/10.1016/j.ifacol.2019.06.088.

[34] V Rocher et al. "Problématique environnementale du nitrite en agglomération parisienne - Origine et solutions techniques envisageables". FR. In: June 2017.

[35] Vincent Rocher et al. "La production de nitrites lors de la dénitrification des eaux usées par biofiltration - stratégie de contrôle et de réduction des concentrations résiduelles". FR. In: *Revue des sciences de l'eau* (2018). DOI: https://doi.org/10.7202/1047053ar.

[36] Christian Roux, Augé Anthony, and Alexandre Cussoneau. *Retreauspective, bilan 2018 du système d'assainissement*. FR. Tech. rep. SIAAP, 2019.

[37] SIAAP. *Fiches usines : Usine d'épuration Seine aval*. FR. Tech. rep. 2013. URL: https://www.siaap.fr/fileadmin/user_upload/Siaap/3_Equipements/Les_usines/Seine_aval/Fiche_usines_SAV_2013.pdf.

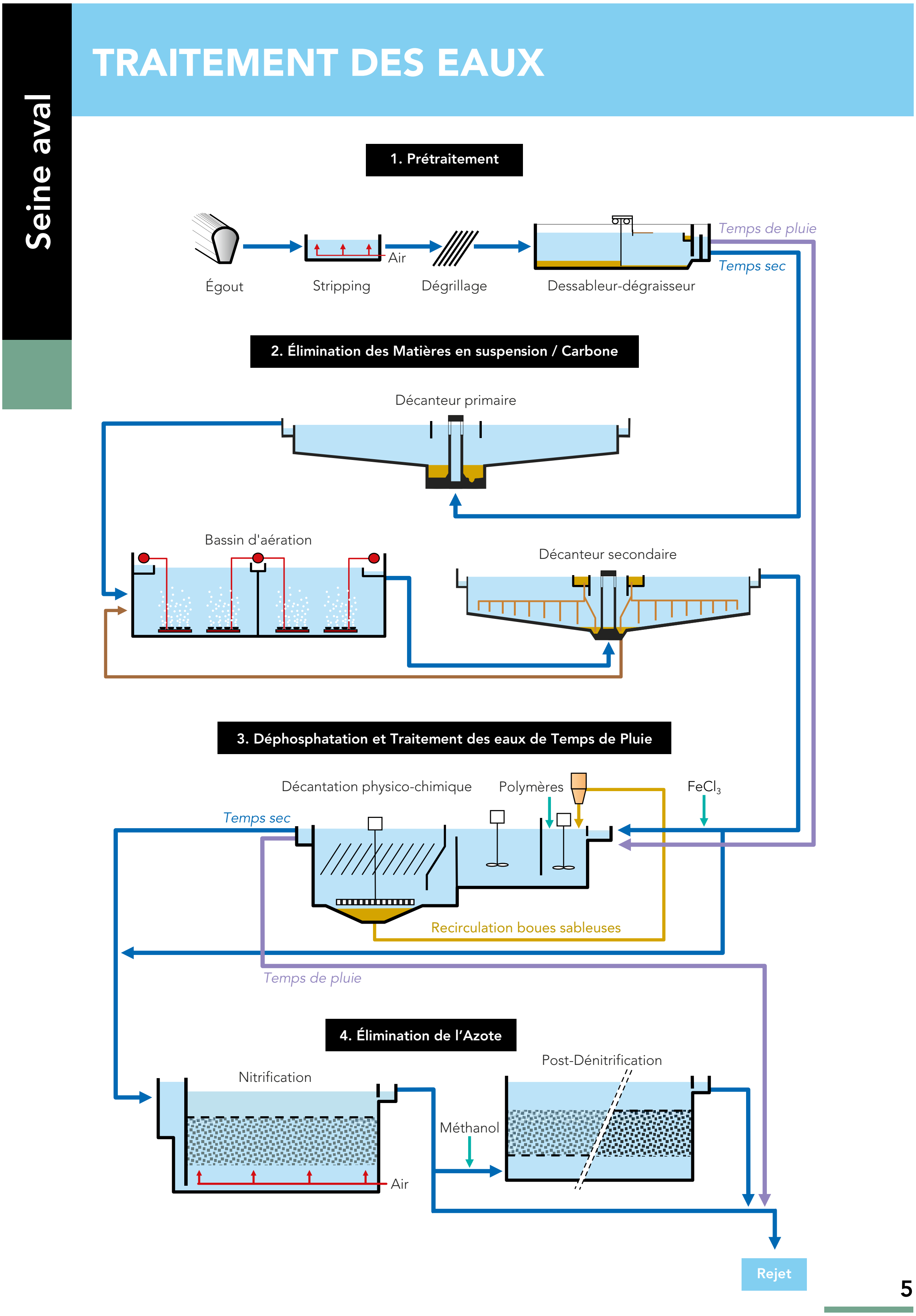
[38] SIAAP. *Seine aval - Présentation de l'usine*. FR. Tech. rep. 2017. URL: https://www.siaap.fr/fileadmin/user_upload/Siaap/6_Presse_et_publications/Espace_presse/dossiers_de_presse/2017_DP_SAV.pdf.

[39] M. C.M. Van Loosdrecht et al. "Twenty-five years of ASM1: Past, present and future of wastewater treatment modelling". EN. In: *Journal of Hydroinformatics* 17.5 (2015), pp. 697–718. ISSN: 14647141. DOI: 10.2166/hydro.2015.006.

[40] Marta Verdaguer, Narcis Clara, and Manel Poch. "Ant colony optimization-based method for managing industrial influents in wastewater systems". In: *AIChE journal* 58.10 (2012), pp. 3070–3079.

[41] Jialu ZHU. "Modélisation détaillée du fonctionnement de la filière complète de biofiltration de la station de traitement des eaux usées Seine Aval". FR. PhD thesis. 2020.

A SAV global treatment process



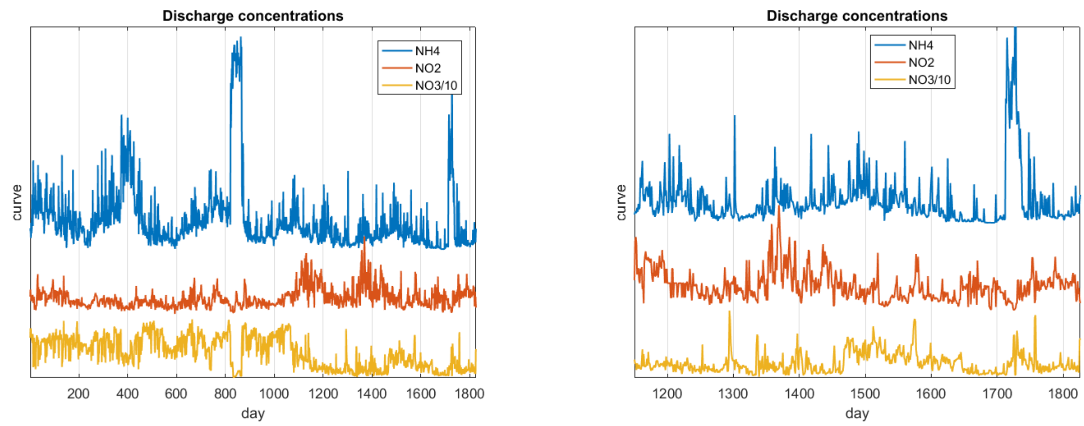
B Seine and SAV typical days

The data starts at January first, 2009. For SAV, we have daily measures for 5 years (until 1/1/2014). For the MeSeine measures (Poissy and Sartrouville) we have weekly data for almost 5 years (until 19/12/2013). The first years will not be taken into account, as the denitrification unit was upgraded in 2011 and more biofilters were introduced in early 2012.

We observe large fluctuations, as shown in Figure 27. We will take from now on only the data after the 1150th day (March 2012). The objective is to extract typical values for NH4, NO2 and NO3 concentrations in the Seine river before and after SAV (at Sartrouville and Poissy, respectively), and at the input and output of SAV. We want the typical values for 2 scenarios :

- normal (good) conditions
- heavy rain conditions (high flow rates)

We will use histogram to observe the most recurrent values. The GEV distribution (Generalized Extreme Value) is often a good fit, and is used for extreme events such as river discharge [19]. The data and the scripts used are available in Appendix 1 (`Somers_24721400_2021_Appendix1.zip`) attached to the thesis file.



(a) Whole dataset (b) From March 2012 (1150th day)

Figure 27: Evolution of the discharge concentrations from SAV. Not to scale.

B.1 Normal conditions

The histogram for Poissy is shown in Figure 28. From the GEV fit, we get the mean values : [NH4, NO2, NO3] = [0.22 0.34 20.88]. Note it still has nitrite concentrations above the regulation (0.3 $[mg/L]$).

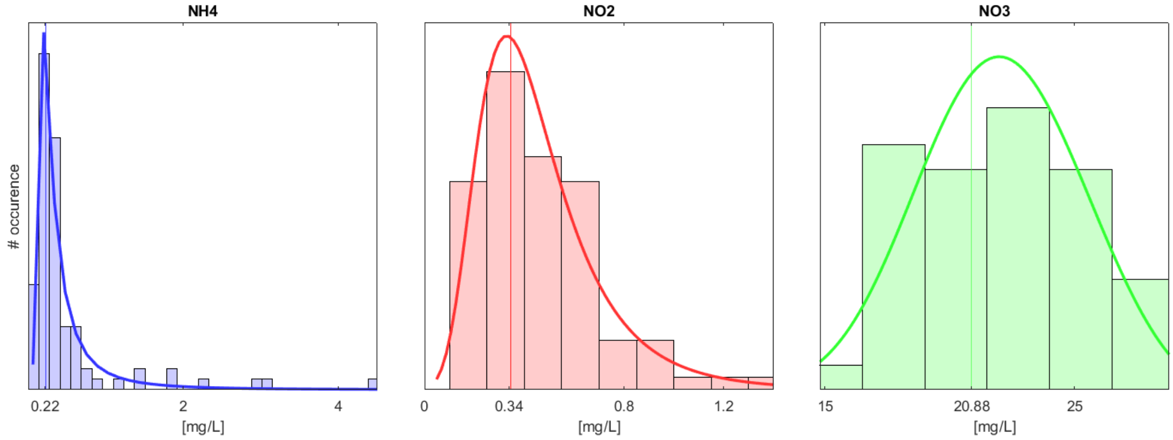


Figure 28: Poissy measures. Mean : NH4=0.22 , NO2=0.34 , NO3=20.88

For Sartrouville, with the same dates and method, we find [NH4, NO2, NO3] = [0.09 0.08 20.04] $[mg/L]$.

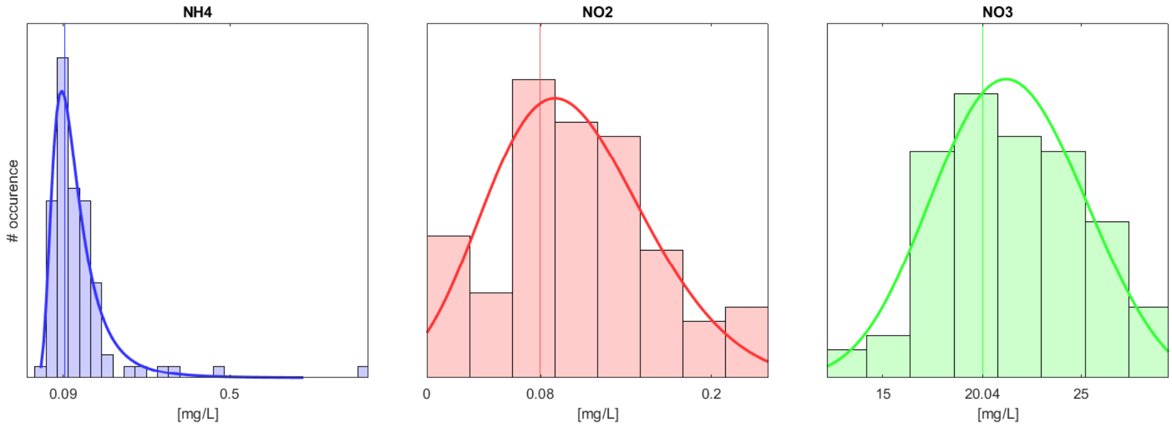


Figure 29: Sartrouville measures. Mean : NH4=0.09 , NO2=0.08 , NO3=20.04

For SAV, we have daily measures. The period remains the same (from day 1150 to the end). It gives [NH4, NO2, NO3] = [2.62 4.45 21.96] $[mg/L]$. Idem for the inlet flow rate $Q_{in} = 16.67$ $[m^3/s]$.

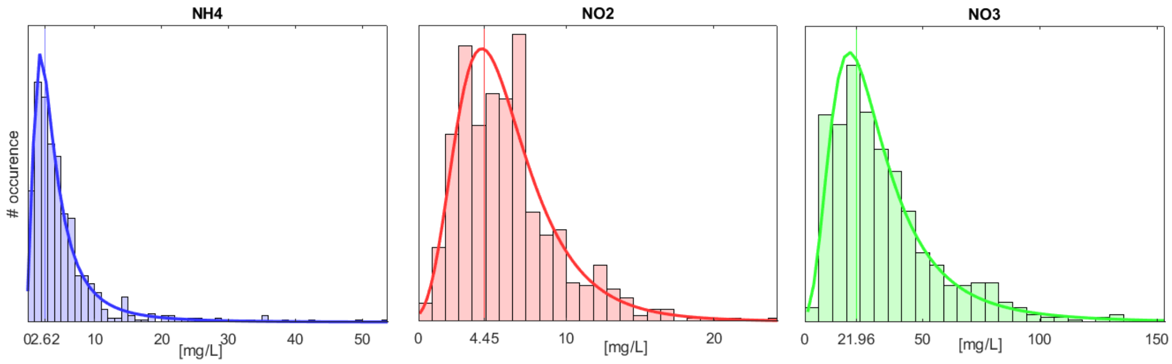


Figure 30: SAV output. Mean : NH4=2.62 , NO2=4.45 , NO3=21.96

For a Seine river with a flow rate equal to 350 $[m^3/s]$ and the concentrations of Sartrouville, it gives the following concentrations after SAV (perfectly mixed flows) : [0.20 0.28 20.13]. This is not so far from the Poissy measures (which are perturbed by the Oise river), and it is below the EQS (Environment Quality Standards) required for a *good status* of the Seine river ([0.5 0.3 50]).

B.2 Heavy rain conditions

We look here at the 3 peaks in the SAV inlet flow rate around day 1600 in Figure 31. They correspond to the end of May and June 2013, with May 20th having very high precipitations (nearly 40 $[m^3/s]$ coming in SAV) and May 30th very high Seine flow rate (close to 1000 $[m^3/s]$). The exact chosen days are 1601, 1602, 1611, 1631 and 1632. For Poissy and Sartrouville, we only have weekly measures of the concentrations so we took days 1603, 1610 and 1631.

The data set is small and rather heterogeneous. Each rain event has different parameters (intensity, duration, location,...) and depends on the previous days (drought, pollution, malfunction,...) among other things. Also, the averaging produces only an average set of conditions, not a "real" one. The values given in Table 18 are thus mostly indicatives.

	NH4 $[mg/L]$	NO2 $[mg/L]$	NO3 $[mg/L]$	Q $[m3/s]$
$SAV input$	17.34	/	/	35
$SAV output$	7.56	2.34	19.00	35
Sartrouville	0.12	0.02	20.80	/
Poissy	0.28	0.47	20.57	766

Table 18: Values for heavy rain conditions

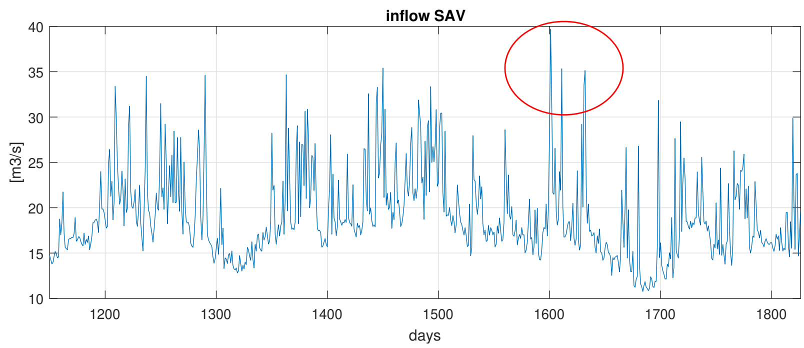


Figure 31: Measures at SAV for heavy rains

C MATLAB code

The various MATLAB scripts and functions in the Appendix attached to the thesis (`Somers_24721400_2021_Appendix1.zip`) are summarised here (except for the functions presented in subsection 2.4).

- `simulation/` contains the functions and scripts related to the simulation.

- simulation/coefficients_k.m computes the kinetic coefficients introduced in Equation 17-22.
- simulation/lin_sys.m computes the linearization of subsection 2.1 around the two equilibrium (regular days and heavy rain). It outputs the proportional gains for the simple controller in subsubsection 3.2.2.
- simulation/init_rain_SAV.m initialises the values for the heavy rain conditions (Table 10).
- simulation/my_dynrain.m works as my_dyn.m, but the flows Q_{in} and Q_{seine} are dynamic. It is only used to find the equilibrium values for heavy rain conditions (see subsubsection 3.1.1.
- simulation/lin_sys.m computes the optimal gains for the linearized system around each equilibrium.
- simulation/limit_con_script.m computes all the possible connections between the agents (i.e. all the possible adjacency matrices), executes the simulation for each and stores the metrics.
- adjacencies/ contains the results of the simulations done in simulation/limit_con_script.m.
- adjacencies/plots.m uses the metrics from limit_con_script.m, calculates the algebraic connectivity and displays them.
- constrained.m is used for the simple simulation in subsection 4.2.
- data/ contains the data set and the scripts used in Appendix B.
- data/SAV.m is the script used to analyse the data for the inputs and the outputs of Seine aval.
- data/Sartr.m is the script used to analyse the data for the Sartrouville measurements.
- data/Poissy.m is the script used to analyse the data for the Poissy measurements.
- data/flows.m is the script used to analyse the data for the flow rate of the Seine river at Sartrouville and at Poissy.
- data/data_use.xlsx is the data set used for the scripts above.
- data/DonnéesCR_DO_MNamont_STEP_2009_13v2.1.xlsx is the original data set.

D Algebraic connectivity metric

The only communications of interest in the analysis are the ones concerning x_1 and x_2 (the ones controlling the flow rates). To calculate a metric based on the algebraic connectivity, it is first necessary to fill the other lines of the adjacency matrix (the filled part is in blue in the example below). Then following the method in [26] for directed graph, the mirror Laplacian $\hat{\mathcal{L}}$ is built ($\hat{\mathcal{L}} = \frac{1}{2}(L + L^T)$). The algebraic connectivity is then simply the second smallest eigenvalue of $\hat{\mathcal{L}}$.

For example:

$$\mathcal{A} = \begin{array}{c c} & \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \qquad \hat{\mathcal{L}} = \frac{1}{2} \begin{array}{c c} & \begin{array}{ccccc} x_1 & x_2 & v_{in} & v_{seine} & y \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ v_{in} \\ v_{seine} \\ y \end{array} & \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & -1 \\ 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & -1 & 3 & -1 \\ 0 & -1 & -1 & -1 & 3 \end{pmatrix} \end{array}$$

$$eig(\hat{\mathcal{L}}) = \begin{pmatrix} 0 \\ 0.5 \\ 1 \\ 2 \\ 2 \end{pmatrix}$$

The algebraic connectivity is thus 0.5 in this example. The metric used in Figure 21 for each realisation i is $max(\lambda_2) - \lambda_{2,i}$, where $max(\lambda_2)$ is the maximal algebraic connectivity present in the data set.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**