Hybrid Modelling and Receding Horizon Control of Combined Sewer Networks

Bernat Joseph i Duran

Institut de Robòtica i Informàtica Industrial

Universitat Politècnica de Catalunya

Consejo Superior de Investigaciones Científicas

A thesis co-advised by:

Gabriela Cembrano Gennari, Carlos Ocampo Martínez

A thesis submitted for the degree of

Doctor of Philosophy

Barcelona, 2014

Universitat Politècnica de Catalunya

Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial

Programa de doctorat:
Automàtica, Robòtica i Visió

Aquesta tesi ha estat realitzada a:

Institut de Robòtica i Informàtica Industrial, CSIC-UPC

Directors de la tesi:
Gabriela Cembrano Gennari
Carlos Ocampo Martínez

Tribunal de la tesi:
Xavier Litrico (LyRE, Lyonnaise des Eaux)
Peter-Jules van Overloop (Delft University of Technology)
Vicenç Puig Cayuela (Universitat Politècnica de Catalunya)
Roberto Griñó Cubero (Universitat Politècnica de Catalunya)
Josep Maria Mirats Tur (CETaqua)

With equal passion I have sought knowledge.
I have wished to understand the hearts of men.
I have wished to know why the stars shine.
And I have tried to apprehend the Pythagorean power
by which number holds sway about the flux.
A little of this, but not much, I have achieved.

Bertrand Russell

All models are wrong, but some are useful.

George E. P. Box

Agraïments

En primer lloc, vull donar les gràcies al meus tutors, Carlos Ocampo i Gabriela Cembrano, per la seva orientació al llarg d'aquests anys. També per haver-me deixat sempre tenir llibertat de moviments i per haver-me permès decidir les coses de manera conjunta i acordada, la qual cosa ha contribuït que aquest treball s'hagi convertit en una experiència molt més personal i motivadora.

En segon lloc, vull donar les gràcies a la meva família: els meus pares, la meva germana i els meus avis. M'agrada pensar que jo tinc una mica de cada un de vosaltres, així que aquesta tesi, d'alguna manera, també és vostra.

Evidentment, també vull donar les gràcies a tots el meus amics de Canet de Mar i de la Facultat de Matemàtiques i Estadística (FME): perquè els bons moments que he passat amb vosaltres, en moltes ocasions m'han donat energies per continuar treballant.

Finalment, vull donar les gràcies a tot el personal de l'Institut de Robòtica i Informàtica Industrial (IRI) pel bon tracte rebut durant aquests anys, i a en Michael N. Jung i en Sebastian Sager, del Interdisciplinary Center for Scientific Computing (IWR) de Heidelberg, per haver-me acollit en el seu departament durant uns mesos i per tot el que he après treballant amb ells.

Resum

Les xarxes de clavegueram combinades transporten conjuntament aigües residuals i aigües pluvials. En absència de pluges, tota l'aigua és conduïda cap a plantes de tractament on és degudament tractada abans de ser retornada als cossos aquàtics adjacents. En canvi, durant episodis de pluja intensa, la capacitat de la xarxa pot esdevenir insuficient donant lloc a inundacions en zones urbanes i abocaments d'aigua no tractada als medis receptors. Per tal de mitigar aquests efectes, les xarxes de clavegueram combinades acostumen a disposar de dipòsits de retenció i elements de redistribució del cabal, regulats amb la finalitat d'aprofitar al màxim la capacitat de la xarxa. En les últimes dècades s'han desenvolupat tècniques de control automàtic per a la regulació d'aquests elements d'emmagatzematge i redistribució, essent el control a temps real, global i predictiu basat en models la tècnica considerada més eficient, donat que és capaç de tenir en compte mesures instantànies del sistema i prediccions d'intensitat de pluja.

En aquesta tesi, es proposa una metodologia completa per al desenvolupament d'un controlador a temps real, global i predictiu basat en model per minimitzar els efectes contaminants en xarxes de clavegueram combinades. El model físic que descriu els fluxos en canals oberts es basa en un sistema d'equacions en derivades parcials que s'ha de resoldre numèricament. Com que en una estrategia de control predictiu a temps real les equacions del model s'han de resoldre moltes vegades per avaluar els efectes de diferents accions de control, el temps necessari per resoldre les equacions limita l'ús del model físic a xarxes petites i amb topologies simples. Per tant, és una pràctica habitual utilitzar models simplificats orientats a control per al control a temps real.

La primera part de la tesi es centra en el desenvolupament, calibratge i validació d'un model simplificat orientat a control del moviment de l'aigua en xarxes de clavegueram

combinades, tenint en compte tres característiques principals: la precisió, la facilitat de calibratge i la velocitat computacional. El model presentat descriu el cabal a través dels elements i estructures hidràuliques més comunes en xarxes de clavegueram combinades, algunes de les quals requereixen l'ús de funcions definides a trossos.

Una vegada les equacions del model han estat presentades, es desenvolupen procediments per al calibratge de tots els paràmetres del model. La metodologia de modelat i calibratge és aleshores aplicada a un cas d'estudi corresponent a una xarxa de clavegueram real i es presenten resultats de validació. Finalment, es duu a terme una anàlisi de sensitivitat respecte als paràmetres més rellevants del model i respecte a la intensitat dels escenaris de pluja considerats.

La segona part de la tesi està dedicada al control òptim basat en el model. En primer lloc, les equacions definides a trossos del model són reformulades per obtenir una expressió del sistema en termes d'un conjunt d'equacions i desigualtats lineals incloent variables contínues i binàries. Usant aquesta expressió general es presenta un procediment basat en matrius per a la formulació de problemes de Control Òptim i Estimació d'Estat.

Mitjançant una implementació de la xarxa del cas d'estudi en un simulador comercial de xarxes de clavegueram que resol les equacions del model físic complet com a realitat virtual, s'avalua el controlador basat en model descrit anteriorment. Resolent iterativament problemes d'Estimació d'Estat i de Control Òptim i utilitzant el simulador per obtenir mesures de la xarxa, se simula una estratègia de control amb horitzó lliscant. La inclusió de problemes d'Estimació d'Estat en llaç de control permet la simulació del controlador amb *output feedback*, tenint en compte que el nombre de mesures disponibles en una xarxa de clavegueram és limitat. Finalment, es discuteixen els resultats obtinguts en aquestes simulacions corresponents a diferents escenaris de disponibilitat de mesures.

Paraules clau: modelat de xarxes de clavegueram, control a temps real de xarxes de clavegueram, control amb horitzó lliscant, estimació amb horitzó mòbil.

Abstract

Combined sewer networks carry wastewater and storm water together. During normal operation all the water is delivered to wastewater treatment plants, where it is treated before being released to surrounding natural water bodies. However, during heavy rain events, the network capacity may become insufficient leading to untreated water discharges to the receiving environments. To mitigate these undesired effects, combined sewer networks are usually provided with detention tanks and flow redirection elements, managed to fully take advantage of the network capacity. In the last few decades automatic control techniques for the regulation of these storage and redirection elements have been developed, with real-time, global, model-based predictive ones being widely regarded as the most efficient ones due to their capacity to take advantage of instantaneous network measurements and rain intensity forecasts.

In this thesis a complete methodology to develop a real-time, global, model-based predictive controller to minimize pollution effects in combined sewer networks is proposed. The physically-based model for open-channel flow is based on a set of partial differential equations, which must be solved numerically. Since in a real-time predictive control strategy the model equations must be solved many times to evaluate the effect of different control actions, the time needed to solve the equations limits the use of the physically-based model to small network instances with simple topologies. Therefore, it is a common practice to use simplified control-oriented models for real-time control.

The first part of the thesis is focused on the development, calibration and validation of a simplified control-oriented model for water transport in combined sewer networks, taking into account three main features: accuracy, calibration ease and computational speed. The proposed model describes the flows through the most common elements and hydraulic structures present in combined sewer networks, some of which requiring

the use of piecewise equations.

Once the model equations are presented, calibration procedures to compute all the model parameters are developed. The modelling and calibration methodology is then applied to a real case study and validation results are provided. Finally, sensitivity analysis is conducted with respect to both the most relevant model parameters and the intensity of the considered rain scenarios.

The second part of the thesis is devoted to model-based optimal control. First, the piecewise equations of the model are reformulated to obtain a general expression of the system by means of a set of linear equations and inequalities including continuous and binary variables. Using this general expression, matrix-based procedures for the formulation of Optimal Control Problems and State Estimation Problems are presented.

Using an implementation of the case study network in a commercial sewer network simulator solving the complete physically-based model equations as virtual reality, the proposed model-based controller is evaluated. By iteratively solving State Estimation Problems and Optimal Control Problems and using the simulator to provide network measurements, a Receding Horizon Control strategy is simulated. The inclusion of State Estimation Problems in the control loop allows to perform output feedback control simulations taking into account that in a sewer network the number of available measurements is limited. Finally, a discussion of the results obtained with these simulations corresponding to different measurement availability scenarios is provided.

Keywords: sewer network modelling, real-time control of sewer networks, receding horizon control, moving horizon estimation.

Contents

List of Figures

List of Tables

List of Acronyms

CSO Combined Sewer Overflow

CSP Constraint Satisfaction Problem

FSM Full State Measurement

HLD Hybrid Linear Delayed

MHE Moving Horizon Estimation

MHEF Moving Horizon Estimation with Flow measurements

MHEL Moving Horizon Estimation with Level measurements

MHEC Moving Horizon Estimation with level and Collector flow measurements

MHEC2 Moving Horizon Estimation with level and 2 Collector flow measurements

MILP Mixed Integer Linear Programming

MLD Mixed Logical Dynamic

MPC Model Predictive Control

ODE Ordinary Differential Equation

OCP Optimal Control Problem

PDE Partial Differential Equation

RHC Receding Horizon Control

RTC Real-Time Control

SEP State Estimation Problem

VT Virtual Tank

WWTP Wastewater Treatment Plant

Part I

Preliminaries

Chapter 1

Introduction

Combined sewer networks are present in many large cities all over the world. These networks carry both wastewater and storm water together. During low to moderate rain events, this water is carried to Wastewater Treatment Plants (WWTPs), where it is treated before being released to the receiving environment (usually a river or the sea). However, during heavy-rain events both the network and WWTP capacities can be easily overloaded, causing urban surface flooding as well as untreated water discharges to the environment, known as Combined Sewer Overflows (CSOs). In most cities with combined sewer networks, the current available infrastructures were designed and build in different phases as a response to the city expansion with little or no planning regarding the potential future growth leading, in some cases, to underdimensioned networks. These planning deficiencies are expected to become even more relevant in the years to come according to some predictions stating that 80% of the world population will live in urban areas by the year 2050 [Price, 2000].

The effects of CSO discharges include visible matter, infectious (pathogenic) microorganisms, oxygen-demanding materials, suspended solids, nutrients and toxicants (e.g., heavy metals, pesticides and petroleum hydrocarbons) [Field et al., 2004]. Aside from the evident ecological and public health problems derived from these effects, water pollution in urban areas has also a direct socioeconomic impact. The aesthetic value of urban open spaces surrounding rivers, beaches and lakes makes these spaces especially suitable for recreational use turning them into potential sources for economic activity and citizen welfare.

To avoid unwanted CSO discharges, detention tanks are usually built along combined sewer networks to store the water and wastewater during the peak rain intensity periods and later release it at lower flow rates suitable for WWTPs. Since these infrastructures are clearly expensive and difficult to locate in urban areas, its efficient operation has become a topic of major interest. In addition to increase the network storage capacity by means of the construction of detention tanks, it is also evident that the homogeneous distribution of the water along the network to fully take advantage of its volumetric capacity plays an important role in avoiding CSOs. This can be achieved by the proper management of flow redirection elements such as gates, weirs or pumps.

Real-Time Control (RTC) techniques offer a solution to the regulation of sewer network infrastructure that takes advantage of nowadays scientific knowledge, technological means and computer power. While conventional regulation approaches consist of simple rules (or man-made decisions) based on local measurements of the network status or even by means of static regulation elements, global RTC takes advantage of centralized knowledge of the whole network measurements and rainfall forecasts and, by means of predictions of the future network status corresponding to different management actions, select the most convenient option every few minutes using the last available information. These features correspond to a *global, real time, model-based optimal control* strategy and is widely regarded as the best control option for sewer network regulation [Ocampo-Martínez et al., 2013, Pleau et al., 2010, Puig et al., 2009a, Schütze et al., 2004, Vezzaro and Grum, 2014].

Although the results appear to be promising, the development and implementation of such an RTC solution must take into account a number of issues to perform as expected. From the control algorithm point of view, these issues can be mainly classified as modelling issues and control issues, which are the focus of this thesis. A complete real implementation should also take into account budget for equipment and instrumentation, operator training, maintenance, etc.

The basic element needed to set up a global, real-time, model-based optimal controller is a suitable model to describe the system dynamics. Three main features are expected from such a model: it must provide suitable approximations of the system dynamics, it must be easily calibrated and validated using real or artificially generated

data and it must be possible to pose and solve Optimal Control Problems (OCPs) and State Estimation Problems (SEPs) based on the model in real-time.

As discussed later in this thesis, a system model to be used in a control strategy that takes advantage of measurements need not be a very complex and accurate model of the system dynamics. In fact, complex physically-based models, especially for large-scale systems such as a sewer network, are not commonly used in RTC due to the extended computational times required to be evaluated. Even if a single evaluation of the model can be executed in a few seconds, in an optimization-based predictive strategy the model is evaluated hundreds or thousands of times at each iteration to take into account all possible present and future control actions for a given time window. To overcome this difficulty, simplified control-oriented models are developed providing an acceptable trade off between accuracy and computational burden. In a simplified model, some elements of the system dynamics are omitted and its influence is concentrated in the form of model parameters. These parameters must be calibrated using system data, preferably from a restricted range of operation points. It often occurs that, while parameters in a physically-based model have physical meaning and can be measured or calibrated without much difficulty, parameters in simplified models do not. Therefore, when developing a model for RTC, a third feature, in addition to accuracy and computational time, plays an important role: the ease of calibration. A common practice is to calibrate a complex physically-based model by means of its physical properties and real measurement data and use it to generate further data to calibrate the simplified model for RTC.

Important control issues in the development and implementation of real-time controllers for sewer networks are dependent on the control strategy. It can be said that, if a model fulfilling the three requirements mentioned above is available, chances are that any control strategy based on the model will outperform any local or static control strategy. From a general point of view, the most relevant issue, aside from computational time, is measurement availability. Sewer networks are usually large-scale systems for which measurements are only available at specific locations. Therefore, to apply most control techniques, including model-based and optimization-based ones, estimation techniques must be used to reconstruct the system state. Notice that in this

context the corresponding estimation algorithms must also be executed in real-time, adding another reason to select a computationally light model.

Finally, it is worth noting that measurements of the system hydraulics can be taken in form of flow measurements and water level measurements. Level sensors are the most common available ones to provide measurements in sewer networks since they are cheaper, more reliable and require less maintenance than flow meters. In spite of this, most simplified hydraulic control models are based on flow variables requiring flow-level transformations to be carried to run the control algorithms. The relation of these two variables can become quite complex in certain flow situations and is not always unique.

Recent works on combined sewer network regulation show a number of different modelling and optimization-based control approaches depending on the characteristics of the specific case studies. Large-scale networks are usually modelled using topological aggregations with only some flows between catchments actually appearing in the model. However, topological simplification is a difficult task due to the complex interconnections among the network elements. Moreover, once the topological aggregation has been performed, these approaches rely heavily on on-line calibration techniques due to the lack of physical meaning of some parameters and the fact that the same models usually describe both the (hydrologic) rainfall-runoff and the (hydraulic) water transport processes.

On the other hand, smaller network instances can be modelled with more detailed and complex models. In this case the obtained models are non-linear, which leads to difficult optimization problems that must be solved using time-consuming and computationally demanding algorithms that do not guarantee global optimality or convergence within the time available for real-time control.

A last feature present within both large- and small-scale sewer network modelling is the explicit modelling of weir and overflow variables. In both cases the inclusion of these variables requires the use of piecewise equations that make the model non-differentiable. Therefore the resulting optimization problems cannot be solved by standard derivative-based procedures and Mixed Integer or Global Optimization algorithms, requiring longer computational times, must be used.

In this thesis, all the above mentioned issues regarding the development and implementation of an RTC strategy have been addressed: modelling, calibration, validation,

optimal receding-horizon control and moving horizon state estimation. The developed methodology is aimed at general instances of large-scale networks with the objective of overcoming topological aggregation and calibration difficulties while still providing suitable computational times. The presented model is based on modelling individual network elements, thus avoiding topological aggregation, simplifying the calibration procedure and leading to separate rainfall-runoff and hydraulic models. Linear equations for flow through sewer pipes taking into account flow delay and attenuation are used together with piecewise linear equations for weir flows and overflows resulting into Mixed Integer Linear Programming (MILP) OCPs and SEPs that can be efficiently solved within the available time.

All the developed methodologies have been applied to a modelled real case study where several scenarios of measurement availability have been considered. All the proposed techniques have been tested by means of simulations using a detailed physically-based model of the case study network as virtual reality.

1.1 Thesis Objectives

The objective of the thesis is the development of a complete methodology for modelling and control of combined sewer networks to minimize the impact of pollution in presence of heavy rain events by influencing the hydraulics of the network. This methodology is aimed to be applicable to a wide range of network instances sharing common elements such as detention tanks, weirs, gates, collectors and overflow points. Two aspects have been especially taken into account: the computational times needed for model evaluation and, consequently, OCPs and SEPs, and the ease to set up a model for a given network instance taking into account topological simplification and parameter calibration. In addition, procedures to validate the model accuracy and assess the control performance, including state estimation, have also been developed.

The above described general objective is detailed in the following list of specific objectives:

1. Modelling

(a) To develop a hybrid linear delayed model for combined sewer networks including the equations for each individual element: sewer pipes, tanks, weirs, overflow points and collectors.

(b) To develop calibration procedures to determine the model parameters for each element, based on the minimization of the error between the model-generated flow approximations and those computed by a complete physically-based model simulator.

(c) To validate the model accuracy according to quantitative indices, based on the error between the model-generated flow approximations and those computed by a complete physically-based model simulator.

2. Control

(a) To reformulate the model equations introducing binary variables so that the corresponding model-based OCPs and SEPs can be handled by standard optimization solvers.

(b) To develop a closed-loop simulation algorithm to evaluate the performance of a model-based Receding Horizon Control (RHC) strategy using a detailed physically-based model description of a network as virtual reality.

(c) To assess the controller performance according to several real-data scenarios of measurement availability including flow measurements and water level measurements.

1.2 Outline of the Thesis

The thesis is structured in four main parts: Preliminaries, Modelling, Control and Concluding Remarks. In the following, a brief summary of the contents of each chapter is given.

Preliminaries

Chapter 2: Background

Before starting with the description of the modelling and control techniques proposed in this thesis, a review of the most common modelling and control techniques in sewer

networks management is provided in this chapter. First, hydraulic models for water motion in open channels are presented, including both physically-based modelling and simplified control-oriented modelling. Secondly, an outline of other modelling techniques for water quality used in the so called *integrated modelling* is also given. Finally, an overview of common algorithms used for sewer network management and operational control are discussed with special emphasis on receding-horizon optimization-based strategies.

Large-scale sewer networks are usually modelled by means of aggregate models that lead to some difficulties to determine the control model topology and to calibrate the model parameters. On the other hand, detailed complex models are not suitable for large-scale networks due to extended computational times. The model presented in this thesis uses a hybrid linear framework to model individual network elements while providing suitable computational times for real-time control of large-scale sewer networks.

Modelling

Chapter 3: Sewer Network Modelling

In this chapter, the equations of two sewer network models are given: the *virtual tank* model and the *hybrid linear delayed* model. The *hybrid linear delayed* model is the novel model proposed in this thesis and has been developed as an extension of the *virtual tank* model including new elements and an improvement of some other elements and general features.

After the two models are presented, calibration procedures to compute all the model parameters for the *hybrid linear delayed* model are described. These procedures are based on minimization of model error with respect to data generated by a physically-based model, although they could also be applied using real measurement data, if available.

The *hybrid linear delayed* is based on modelling each network element individually, thus avoiding the need of topological aggregation, simplifying the calibration process and leading to accurate approximations. The basic flow equations are linear equations taking into account the flow delay and attenuation. Piecewise linear equations have

been used to model overflows and weirs. Since the number of switching equations is limited to the number of weirs and overflow points the model can be fastly evaluated, making it suitable for large-scale networks.

Chapter 4: Case Study and Model Validation

To validate the modelling and calibration procedures presented in Chapter 3, a case study corresponding to a real network is presented: the Riera Blanca sewer network. A complete description of this network has been made available by the company responsible of the network regulation in the form of an implementation in a commercial physically-based model simulator. This network includes all elements considered in the *hybrid linear delayed* and urban overflows and CSOs have been shown to occur during episodes of heavy rain. Aside from the network description, historical data corresponding to real rain events of different intensities and durations has been provided by the company. After applying a mild topological simplification, the model is calibrated using data generated by a physically-based model corresponding to four real rain events.

Validation results are provided in the form of error indices comparing the network flows generated by the *hybrid linear delayed* model and the physically-based model. Finally, sensitivity analysis is applied to show the impact of model parameters on the predicted flows and the impact of the intensity of the rain events on the model accuracy.

Control

Chapter 5: Model Reformulation and Control Problems

To use the *hybrid linear delayed* model for control purposes, model-based OCPs and SEPs must be formulated. However, the model equations as presented in Chapter 3 are not suitable for the formulation of such problems. Therefore, a reformulation of the model in the form of a set of linear equations and inequalities including binary variables is described, based on the so-called Mixed Logical Dynamic systems approach.

Using the general expression obtained from the model reformulation, three optimization problems can be simply formulated in a form that can be efficiently handled by standard MILP solvers. First, a Constraint Satisfaction Problem is presented to be used for model simulation. Secondly, an OCPs is formulated to compute optimal

control actions with respect to an objective function describing the required system performance. Finally, a SEPs to reconstruct the system state from a set of measurements is described.

Chapter 6: Receding Horizon Control

This chapter presents the results obtained by applying the *hybrid linear delayed* model of Chapter 3, together with the OCPs and SEPs of Chapter 5 to the case study network of Chapter 4. To evaluate the performance of the proposed controller in real-time control operation taking into account measurements, a physically-based model of the case study network is used as virtual reality. A closed-loop simulation algorithm to simulate a RHC with Moving Horizon Estimation (MHE) strategy by means of solving a series of OCPs and SEPs and performing physically-based model simulations is described, including some implementation issues. After describing the management objectives for the case study network and presenting the corresponding objective function for the OCPs, performance results are presented and discussed. These results include simulations corresponding to several measurement scenarios including full flow measurements, limited flow measurements and limited water level measurements. Finally, the computational times needed to solve the optimization problems involved in the simulations are presented to evaluate the suitability of the proposed controller for real-time operation.

Aside from the specific performance results for the case study network, the results presented in this chapter are aimed to show that, by means of the closed-loop simulation algorithm, the proposed modelling and control methodologies are not only useful for control purposes but also as a tool for infrastructure and instrumentation planning.

Concluding Remarks

Chapter 7: Conclusions and Future Work

In this chapter, the fulfillment of the objectives proposed in Chapter 1 is evaluated and a summary of the main contributions of the thesis is given. In addition, future research lines to further improve the presented results are outlined.

Appendices

The appendices section is mostly devoted to display extended results of the RHC simulations described in Chapter 6, including flow approximations, collector volume approximations, SEPs solutions and performance of the local gate controllers with respect to the OCPs set-points.

Chapter 2

Background

2.1 Introduction

Regulation of combined sewer networks, as many other engineering problems, is a problem that requires putting together knowledge coming from different fields. Physical principles are applied to derive the model equations for water motion along open channels and, in the so called *integrated modelling* approach, chemical and biochemical principles are applied to derive the model equations for water quality processes. These model equations turn out to be too complex to be solved in real time for most control applications. In response to this problem, simplified models have been developed from different physical, mathematical and engineering approaches: simplified physically based-equations, conceptual models and data-based black-box models. Depending on the type of model equations, different control engineering techniques can be used, the most common being the optimization-based ones. Optimization problems can be solved on-line, to compute control actions based on the last available measurements and disturbance forecasts, or off-line, to develop rule-based control algorithms. Therefore, considerations regarding numerical methods for optimization problems depending on the nature of the model equations (derivative based/derivative-free, linear, nonlinear, mixed integer) must be taken into account.

This chapter provides an introduction and literature review of the main topics involved in the contributions of this thesis:

- Physically-based modelling of open channel flow

- Control-oriented hydraulic models
- Integrated sewer network modelling
- Real-time control of sewer networks
- Receding horizon optimization-based control

It also aims to put the problem in context and provide arguments to justify the choice of the techniques used.

2.2 Hydrology and Hydraulics

The physical processes describing the water motion in a sewer system can be mainly classified in two fields: hydrology and hydraulics. Hydrology focuses on the study of the distribution and properties of water on the atmosphere and surface of the Earth. Regarding sewer network systems, the most important hydrologic process involved is the rainfall-runoff process. Rainfall-runoff models aim to compute net flows of a rain catchment towards a receiving system such as a river, the sea or, in this case, a sewer network, as a function of the rain intensity. This is a very complex process that depends on the spatial and temporal distribution of the rain and on the catchment geometry and materials, and is usually modelled by means of conceptual models. In a first step, losses with respect to the measured rain data due to surface wetting, infiltration, evaporation and surface storage are subtracted. Then, the net rain inflow is transported along the catchment surface until it enters to the sewer network. This transport process is often modelled by means of a reservoir or a series of reservoirs emptying in a linear or nonlinear way or by means of a unit hydrograph (c.f. Schütze et al. [2002]). Nowadays, these models are well established and described in the literature and widely used in the sewer network modelling and control community [Rauch et al., 2002].

On the other hand, hydraulics focus on the mechanic properties of liquids moving through, usually man-made, conduits and hydraulic structures such as weirs, gates, dams, etc. This is the most relevant modelling part from the control point of view, since it is the part of the process which can actually be controlled.

Since this work is focused on control-oriented modelling and control of sewer systems, it deals with the hydraulic part of the water transport along the sewer system. The specific hydrologic rainfall-runoff model used for model calibration, validation and control is described in Section 3.3.7, but no further attention is given to hydrological modelling. An early review of rainfall-runoff models for sewer systems can be found in Previdi et al. [1999]. Sewer network simulation software is usually equipped with such models as reported in MOUSE [2007b]. Finally, an exhaustive review and discussion of current and future trends can be found in Beven [2011].

2.3 Physically-based Model of Open-Channel Flow

2.3.1 Flow Routing Model

The physical model for water motion in sewer networks is based on the 1-Dimensional Saint-Venant equations [de Saint-Venant, 1871] with constant channel cross-sectional area [Marinaki and Papageorgiou, 2005, Ocampo-Martínez, 2011, Rauch et al., 2002, Schütze et al., 2002]. These equations are hyperbolic nonlinear Partial Differential Equations (PDEs) relating the flow and water level in an open channel/sewer pipe. The following assumptions are made in the derivation of the Saint-Venant equations [Litrico and Fromion, 2009]:

- The flow is one-dimensional: the velocity is uniform over the cross-section and the water level across the section is horizontal.
- The streamline curvature is small and vertical accelerations are negligible, hence the pressure is hydrostatic.
- The effect of boundary friction and turbulence can be accounted for through resistance laws analogous to those used for steady-state flow.
- The average channel bed slope is small so that the cosine of the angle with the horizontal may be replaced by one.
- The variation of the channel width is small.

With these assumptions, the Saint-Venant equations read

$$\frac{\partial A(x,t)}{\partial t} + \frac{\partial Q(x,t)}{\partial x} = 0, \tag{2.1}$$

$$\frac{\partial Q(x,t)}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q(x,t)^2}{A(x,t)}\right) + g\,A(x,t)\frac{\partial Y(x,t)}{\partial x} = g\,A(x,t)\Big(S_b(x) - S_f(x,t)\Big), \tag{2.2}$$

where:

- x is the longitudinal coordinate [m],
- t the time [s],
- $Q(x,t)$ the flow $\left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$,
- $A(x,t)$ the cross-sectional area of the flow [m^2],
- $Y(x,t)$ the water level [m],
- $S_b(x)$ the bed slope [dimensionless],
- $S_f(x,t)$ the friction slope [dimensionless], approximated by the Manning formula [Chaudhry, 2008, Chow, 1959, Litrico and Fromion, 2009]:

  $$S_f = \frac{Q(x,t)^2 n^2}{A(x,t)^2 R_h(x,t)^{4/3}},$$

  where n is the Manning coefficient $\left[\mathrm{s\,m}^{-1/3}\right]$ (depending on the channel physical properties) and $R_h(x,t)$ the hydraulic radius [m], defined as

  $$R_h(x,t) = \frac{A(x,t)}{P(x,t)},$$

  where $P(x,t)$ is the wetted perimeter [m],
- g the gravitational acceleration $\left[\frac{\mathrm{m}}{\mathrm{s}^2}\right]$.

Equation (2.1) is called the *continuity* or *mass conservation equation* and equation (2.2) is called the *momentum equation*. A derivation of the equations can be found in several books on fluid dynamics or open-channel hydraulics [Chaudhry, 2008, Chow, 1959, Litrico and Fromion, 2009]. In the presented form, there are three independent

variables, namely $q(x,t)$, $h(x,t)$ and $a(x,t)$. It is common, however, to use an area-level relation of the form

$$A(x,t) = f(Y(x,t)),$$

to eliminate one of both variables. Another variable that is commonly used in some formulations of the Saint-Venant equations is the flow velocity $V(x,t)$ $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$, which can be used to replace the flow $Q(x,t)$ by means of the following relation:

$$Q(x,t) = A(x,t)\,V(x,t).$$

In the following, it is assumed that flows $Q(x,t)$ and water levels $Y(x,t)$ are the variables of choice to express the equations.

For a complete sewer network model, the Saint-Venant equations are applied to each sewer pipe and coupled by means of internal and boundary conditions defined at joints, sewer pipe geometry changes and hydraulic structures. These internal and boundary conditions imply that the dynamics of the network need to be solved as a single system, not for each sewer pipe, so that the problem becomes computationally very demanding for big networks with complex topologies.

2.3.2 Flow Classification

2.3.2.1 Temporal and Spacial Variation

A solution of the Saint-Venant equations is said to be *uniform* if it does not change along the spatial coordinate x and *steady* if it does not change along time.

Steady solutions are of special importance since the linearization of the Saint-Venant equations around steady-state solutions are the basis for the development of some simplified models and the corresponding controllers. To identify steady solutions, the terms in the Saint-Venant equations involving temporal partial derivatives are removed letting the flow, height and area variables depend only on the spatial coordinate x.

2.3.2.2 Flow Regime

Another way to classify the solutions of the Saint-Venant equations takes into account the flow velocity $V(x,t)$ and the celerity $C(x,t)$ $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$: the velocity of a wave traveling

along the fluid surface,

$$C(x,t) = \sqrt{g\,\frac{A(x,t)}{W(x,t)}},$$

where $W(x,t)$ is the flow top width [m]. Now, depending on the relation between $V(x,t)$ and $C(x,t)$, the flow is said to be:

- Subcritical, if $V(x,t) < C(x,t)$,
- Critical, if $V(x,t) = C(x,t)$,
- Supercritical, if $V(x,t) > C(x,t)$.

This is an important property since it has a direct relation with the kind of boundary conditions needed to solve the Saint-Venant equations.

2.3.3 Initial, Boundary and Internal Conditions

As usual, to solve any system of PDEs, initial and boundary conditions must be provided. Initial conditions for the Saint-Venant system of equations are of the form

$$Q(x,0) = Q_0,\ Y(x,0) = Y_0,\ x \in [0,L],$$

where L is the open-channel length.

Boundary conditions must be given in terms of the following variables:

$$Q(0,t),\ Y(0,t),\ t \in [0,t_f],$$
$$Q(L,t),\ Y(L,t),\ t \in [0,t_f],$$

where t_f is the final time. In order for the problem to be well-defined, a boundary condition must be given at each end of the channel, except for the supercritical flow case, for which two upstream conditions are needed Chaudhry [2008], Litrico and Fromion [2009].

The following choices are common both in the sewer networks and irrigation canals fields:

$$Q(0,t) = Q_{in}(t),\ t \in [0,t_f], \quad \text{or} \quad Y(0,t) = Y_{in}(t),\ t \in [0,t_f],$$
$$Q(L,t) = f\big(Y(L,t), p(t)\big),\ t \in [0,t_f].$$

where $Y_{in}(t)$ ($Q_{in}(t)$) is the upstream water level (inflow) and f is a function relating the flow and water level at the downstream end of the channel that usually depends on the hydraulic structure present at that point (with structure parameters $p(t)$). Mathematical expressions corresponding to common structures are reported in Section 2.3.4.

Finally, internal conditions are used to link channels with different geometries or interconnections of several channels as part of a branching network. If no hydraulic structure is present, internal conditions impose water level continuity and mass balance at the interconnection junction.

2.3.4 Hydraulic Structure Models

The classical physical models for hydraulic structures can be found in a number of textbooks on hydrodynamics [Chaudhry, 2008, Chow, 1959] and control of hydrosystems [Litrico and Fromion, 2009, Malaterre and Baume, 1998] and relate the flow and water levels up- and downstream of the structure. The most common structures present in sewer networks are weirs and gates. The flow over a weir according to the physical description reads [Litrico and Fromion, 2009]:

- Free flow:

$$Q_w(t) = C_f \, L_w \, \sqrt{2g} \left(Y_1(t) - Y_w(t)\right)^{3/2},$$

- Submerged flow:

$$Q_w(t) = C_s \, L_w \, Y_2(t) \, \sqrt{2g \left(Y_1(t) - Y_2(t)\right)},$$

where Q_w $\left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$ is the weir flow, Y_w is the weir sill elevation [m], Y_1 and Y_2 [m] are the water levels up- and downstream of the weir, L_w is the weir length [m] and C_f and C_s are dimensionless discharge coefficients. Overflows in junctions of a sewer network are also modelled using the weir formulas.

Similarly, the gate flow formulas read [Litrico and Fromion, 2009]:

- Free flow:

$$Q_g(t) = K_f \, L_g \, Y_g(t) \, \sqrt{2g \, Y_1(t)},$$

- Submerged flow:

$$Q_g(t) = K_s\, L_g\, Y_g(t)\, \sqrt{2g\left(Y_1(t) - Y_2(t)\right)},$$

where Q_g $\left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$ is the gate flow, Y_g is the gate opening [m], Y_1 and Y_2 [m] are the water levels up- and downstream of the gate, L_g is the gate length [m] and K_f and K_s are dimensionless discharge coefficients.

2.3.5 Numerical Methods

The Saint-Venant equations lack an explicit solution for general channel geometry. Therefore, numerical methods are used to solve the equations. Several methods appear in the literature, each exploiting different aspects of the theory of hyperbolic PDEs and its associated numerical methods, with the most widespread being the *method of characteristics* [Abbott, 1966, Chanson, 2004, Chaudhry, 2008, Litrico and Fromion, 2009] and *finite differences methods* [Akan, 2006, Chaudhry, 2008, Chow, 1959, Litrico and Fromion, 2009]. Finite differences models can be classified into *explicit* and *implicit*. In explicit methods, the variables at a given time instant are computed by means of an explicit expression in terms of variables at previous times, whose values are already known while in implicit methods unknown variables at different time instants are related by means of an equation that must be solved using root-finding algorithms. Although explicit methods are easier to implement, they have a major drawback since a stability condition (Courant-Friedrichs-Levy condition [Akan, 2006, Chaudhry, 2008, Litrico and Fromion, 2009]) in terms of the discretization steps must hold:

$$\Delta t \leq \frac{\Delta x}{|V \pm C|},$$

where Δt and Δx are respectively the temporal and spatial discretization steps, V the flow velocity and C the flow celerity.

2.3.6 Preissmann Slot

Although combined sewer networks are designed to work as open-channel networks, peak rain inflows can eventually cause the sewer pipes to fill completely turning the flow into a pressurized one. To properly approximate the transition between open-channel and pressurized flow, in some sewer network simulation tools a fictitious slot

(the *Preissmann Slot*) is added to the top of each sewer pipe. This way, the closed conduit is approximated by an open channel and by properly selecting the slot width the pressurized flow is suitably approximated [MOUSE, 2007a]. The water levels above the actual sewer pipe height, filling the Preissmann Slot, must be interpreted as piezometric heads.

2.4 Control-oriented Hydraulic Models

The Saint-Venant equations are widely used for simulation purposes but they may become unsuitable for real-time optimal or predictive control algorithms. When applied to medium-size to large-scale networks, the number of variables involved in the discretization of the equations and their nonlinear nature produce big and hard root-finding problems that take too long to be solved. Therefore, a number of simplified control-oriented models have been developed in the literature, which allow to compute control actions within the available times. In the following, these models have been classified as:

- Models based on simplification of the Saint-Venant equations
- Models based on discretization of the Saint-Venant equations
- Models based on linearization of the Saint-Venant equations
- Conceptual models
- Model based on identification (*black-box*)

These simplified flow models do not apply just to the sewer network control field, but also to the irrigation canals control one. Therefore, models and references from both areas are discussed in the following.

2.4.1 Models Based on Simplification

Assumptions on the flow characteristics can lead to the elimination of some terms in the momentum equation (2.2). The resulting models omit some phenomena but might be suitable to approximate the flow properties if the assumptions hold. In all cases the

mass conservation equation (2.1) is kept. Table 2.1 shows the phenomena described by the two simplifications described below: the *diffusive wave equation* and the *kinematic wave equation*. In this context, the full Saint-Venant system is known as the *dynamic wave equations*.

Table 2.1: Phenomena described by the Saint-Venant equations and their simplifications, as in Schütze et al. [2002].

	Kinematic Wave	Diffusive Wave	Dynamic Wave
Backwater effects	×	✓	✓
Attenuation of flood waves	×	✓	✓
Flow acceleration	×	×	✓

Notice that these models can still be regarded as physically-based since they are special cases of the Saint-Venant equations under some assumptions on the flow characteristics.

Diffusive Wave Equation

The *diffusive wave equation* is obtained by removing the local and convective acceleration terms $\frac{\partial Q(x,t)}{\partial t} + \frac{\partial}{\partial x}\left[\frac{Q^2(x,t)}{A(x,t)}\right]$ from equation(2.2), leading to

$$\frac{\partial Y(x,t)}{\partial t} = S_b - S_f(x,t).$$

Combining this expression with equation (2.1) and writing the water level in terms of the absolute water level $z(x,t)$ [m] (with respect to horizontal datum) the following single equation is obtained according to Litrico and Georges [1999]:

$$\frac{\partial Q}{\partial t} + C(Q,z,x)\frac{\partial Q}{\partial x} - D(Q,z,x)\frac{\partial^2 Q}{\partial x^2} = 0.$$

This equation, called the diffusive wave equation, can be in turn linearized around a steady solution, leading to the *Hayami model*, consisting of a linear partial differential equation with constant coefficients. The Hayami model can be used to derive an inflow-outflow transfer function, and by approximating this function by rational transfer functions, controllers can be designed [Litrico and Georges, 1999, Litrico and Pomet, 2003, Litrico et al., 2010].

Kinematic Wave Equation

For steady uniform flow, all the left-hand-side terms of equation (2.2) can be removed leading to the the *kinematic wave equation*:

$$S_f = S_b.$$

Using the Manning formula, the last expression reduces to

$$S_b = \frac{Q^2 n^2}{A^2 R_h^{4/3}}.$$

Here, all terms are constant and using a level-area relation, a flow-level expression can be obtained.

2.4.2 Models Based on Discretization of the Saint-Venant Equations

By applying discretization schemes to the Saint-Venant equations, the kinematic wave equation or the diffusive wave equation, nonlinear discrete-time models are obtained.

In Duchesne et al. [2001], discretizations of the kinematic wave and diffusive wave equations are used together with a discretization of the momentum equation (2.1) for flow modelling including the transition between open-channel and pressurized flow. The presented model switches between the kinematic and diffusive wave models according to an iterative trial-and-error decision algorithm. Similarly, in Schwanenberg et al. [2010], discretized versions of the kinematic wave and momentum equation (2.1) are used for the design of an optimization-based predictive controller.

Finally, Xu [2013], Xu et al. [2011, 2012] propose a water level model based on a discretization of the full Saint-Venant equations. A linear time-varying model is obtained by performing a simulation with the complete Saint-Venant model and substituting most variables in the discretization by those resulting of the simulation, leaving only water levels and decision variables free. This model is supposed to be updated on line for control purposes, by performing subsequent simulations with the complete model.

2.4.3 Models Based on Linearization of the Saint-Venant Equations

As in the linearization of an Ordinary Differential Equation (ODE) around an equilibrium point, the Saint-Venant equations can be linearized around a steady solution to

describe the variations of the flow and water level with respect to this solution. The linearized Sain-Venant equations provide the theoretical basis for several simplified models which are widely used in the irrigation systems community where regulation of water levels near a steady state is a common problem [Baume et al., 1998, Litrico and Fromion, 2009, Malaterre and Baume, 1998, Schuurmans et al., 1995, 1999]. However, these models are not commonly used in sewer systems since the variations of the flows during intense rain events are too big to be properly described by the linearized system.

Steady solutions are those solutions which are constant in time, and can be identified by imposing

$$\frac{\partial Q}{\partial t} = 0, \quad \frac{\partial Y}{\partial t} = 0.$$

By substituting these equalities in equations (2.1) and (2.2), the equations for a stationary solution $(Q_0(x), Y_0(x))$ read

$$\begin{aligned} \frac{\mathrm{d}Q_0(x)}{\mathrm{d}x} &= 0, \\ \frac{\mathrm{d}Y_0(x)}{\mathrm{d}x} &= \frac{S_b - S_{f0}(x)}{1 - F_0(x)^2}, \end{aligned}$$

where $F_0 = \frac{V_0}{C_0}$ is the Froude number [dimensionless] and A_0 and C_0 are the flow cross-sectional area and celerity corresponding to the steady flow solution, respectively.

The linearized equations are those fulfilled by the deviations $(q(x,t), y(x,t))$ of a general solution $(Q(x,t), Y(x,t))$ from a given steady solution $(Q_0(x), Y_0(x))$,

$$Q(x,t) = Q_0(x) + q(x,t), \quad Y(x,t) = Y_0(x) + y(x,t).$$

The derivation of the linearized equations requires some tedious calculations that are out of the scope of this work [Litrico and Fromion, 2009]. After some manipulations, the equations can be written in the following compact form:

$$\frac{\partial \xi}{\partial t} + \mathbf{A}(x)\frac{\partial \xi}{\partial x} + \mathbf{B}(x)\,\xi = 0,$$

where $\mathbf{A}(x)$ and $\mathbf{B}(x)$ are obtained from the steady solutions and from physical parameters of the channel and where

$$\xi(x,t) = (a(x,t), q(x,t))^\top$$

with $a(x,t)$ the deviation of the solution cross-sectional area $A(x,t)$ from the steady solution one $A_0(x)$. $a(x,t)$ can be expressed as $a(x,t) = W_0(x)\, y(x,t)$, where $W_0(x)$ is the flow top width corresponding to the steady solution.

Applying the Laplace transform (denoted $\mathcal{L}\{f(t)\} \equiv \hat{f}(s)$) to the previous equation, a linear system of ODEs depending on the Laplace variable $s \in \mathbb{C}$ is obtained:

$$\frac{\partial \hat{\xi}(x,s)}{\partial x} = \mathcal{A}(x,s)\hat{\xi}(x,s) + \mathcal{B}(x,s)\xi(x,0), \tag{2.3}$$

where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are obtained from $\mathbf{A}$ and $\mathbf{B}$.

Equation (2.3) cannot be solved analytically in the general case. However, in the frictionless horizontal flow case and the uniform flow case, matrices $\mathbf{A}(x)$ and $\mathbf{B}(x)$ are constant and explicit solutions can be obtained. In the general case an approximation by splitting the channel into subchannels with uniform flow is given in Litrico and Fromion [2004a, 2009].

Once equation (2.3) is solved, exactly or approximately, an expression for $\hat{\xi}(x,s)$ in terms of $\hat{\xi}(x,0)$ is obtained. By means of further algebraic manipulation, and substituting $x = L$, with L the length of the channel [m], irrational transfer functions of the following form are obtained:

$$\begin{pmatrix} \hat{y}(0,s) \\ \hat{y}(L,s) \end{pmatrix} = \begin{pmatrix} p_{11}(s) & p_{12}(s) \\ p_{21}(s) & p_{22}(s) \end{pmatrix} \begin{pmatrix} \hat{q}(0,s) \\ \hat{q}(L,s) \end{pmatrix}.$$

These irrational transfer functions can, in turn, be approximated by rational ones (plus delay) for fast simulation or controller design purposes. The most well-known approaches are the Integrator Delay (ID) model [Litrico and Fromion, 2004b, Schuurmans et al., 1999], approximating functions $p_{ij}(s)$ by transfer functions of the form

$$\begin{aligned} p_{11}(s) &\approx \frac{1}{A_u s}, \quad & p_{12}(s) &\approx -\frac{e^{-\tau_u s}}{A_u s}, \\ p_{21}(s) &\approx \frac{e^{-\tau_d s}}{A_d s}, \quad & p_{22}(s) &\approx -\frac{1}{A_d s}, \end{aligned}$$

and the Integrator Delay Zero (IDZ) model [Litrico and Fromion, 2009], with approximations of the form

$$\begin{aligned} p_{11}(s) &\approx \frac{1}{A_u s} + \tilde{b}_u, & p_{12}(s) &\approx -\left(\frac{1}{A_u s} + b_u\right) e^{-\tau_u s}, \\ p_{21}(s) &\approx \left(\frac{1}{A_d s} + b_d\right) e^{-\tau_d s}, & p_{22}(s) &\approx -\left(\frac{1}{A_d s} + \tilde{b}_d\right). \end{aligned}$$

Since these expressions correspond to a frequency description of the system, classic control techniques are used together with this modeling approach. However, optimization-based predictive controllers based on the state space reformulation have also been proposed in van Overloop [2006], van Overloop et al. [2005, 2008], where by means of a multimodel approach several disturbance scenarios are taken into account into a single objective function for robustness.

2.4.4 Conceptual Models

Conceptual models arise from simple mathematical descriptions of the most relevant qualitative features of the flow routing phenomenon: transport delay and flow attenuation. These models provide suitable approximations of the variation of the flow along sewer pipes/open channels and are very appealing from the computational point of view since they can be evaluated faster than the physically-based models (even the simplified ones) and are especially suitable for optimization-based control. Notice that conceptual model equations need not have direct mathematical relation with the original Saint-Venant equations.

In the following, two different kinds of conceptual models will be described: conceptual hydraulic models and aggregated models. The former describe the flow along an open channel by means of simplified equations that usually take into account flow delay and attenuation. The latter go a step further in the conceptualization process and model entire network catchments mixing hydrologic and hydraulic phenomena. The catchments are usually defined to have a unique outflow which is computed taking into account rain inflows (the hydrologic part) as well as inflows from other catchments.

Discrete time translation and delay

This model is the base of multiple works by Marinaki and Papageorgiou [1998, 2001, 2003, 2005]. The main equation is a discrete-time relation between the total inflow Q_{in} and the outflow Q_{out} of a sewer pipe:

$$Q_{out}(k+1) = \left(1 - \frac{\Delta t}{\tau}\right) Q_{out}(k) + \frac{\Delta t}{\tau} Q_{in}(k),$$

where Δt is the sampling time and τ a calibration parameter to be estimated experimentally.

The total inflow Q_{in} is computed as the sum of n inflows, $Q_{in}^j,\ j = 1, \ldots, n$, from different sewer pipes that are regarded as having different delay effects,

$$Q_{in}(k) = \sum_{j=1}^{n} Q_{in}^j(k - \kappa_j)$$

where κ_j is the delay of inflow Q_{in}^j.

Muskingum Model

The Muskingum model takes its name from the Muskingum river, where it was first applied. It is based on modelling the channel as a storage element with the additional assumption of the volume being a linear combination of the in- and outflow of the following form [Akan, 2006, Chaudhry, 2008]:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= Q_{in}(t) - Q_{out}(t), \\ V(t) &= KXQ_{in}(t) + K(1-X)Q_{out}(t), \end{aligned}$$

where $V(t)$ is the stored volume, $Q_{in}(t)$ the inflow, $Q_{out}(t)$ the outflow and K and X two parameters to be estimated, as discussed below.

The following discretization scheme is then applied to the previous expressions:

$$\begin{aligned} \frac{V(t+1) - V(t)}{\Delta t} &= \frac{Q_{in}(t+1) + Q_{in}(t)}{2} - \frac{Q_{out}(t+1) + Q_{out}(t)}{2}, \\ V(t) &= KXQ_{in}(t) + K(1-X)Q_{out}(t). \end{aligned}$$

Now, for a given known inflow $Q_{in}(t)$, the corresponding outflow can be computed as

$$Q_{out}(t+1) = c_1 Q_{out}(t) + c_2 Q_{in}(t+1) + c_3 Q_{in}(t), \tag{2.4}$$

with

$$c_1 = \frac{2K(1-X) - \Delta t}{2K(1-X) + \Delta t},\ c_2 = \frac{-2KX + \Delta t}{2K(1-X) + \Delta t},\ c_3 = \frac{2KX + \Delta t}{2K(1-X) + \Delta t}.$$

To calibrate parameters K and X, and therefore c_1, c_2 and c_3, two procedures are commonly described in the literature. The first one, corresponding to the original Muskingum method, consists in using real inflow and outflow measurements together with some curve fitting strategy. The second one is known as the Muskingum-Cunge

model, and is based on using the continuity equation (2.1) together with the kinematic wave equation. After a discretization scheme is applied to the resulting expression, an equation with the same structure as (2.4) is obtained, with coefficients $c_1,\ c_2$ and c_3 being functions of the channel physical parameters [Chaudhry, 2008].

Tanks in series

The *tanks in series model*, also known as the *Nash cascade model*, the *Kalinin-Myliukov-Nash cascade model* or the *unit hydrograph model*, consists in representing the channel as a series of interconnected tanks, each one emptying towards the next at a rate depending on its contained volume. Although it was first developed as a conceptual hydraulic model [Nash, 1957], the tanks in series model has also been used as an aggregate model in several works and control-oriented modelling tools [Meirlaen, 2002, Solvi, 2006, Vanrolleghem et al., 2005].

The volume contained in the tank is described using the following mass balance equation:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = Q_{in}(t) - Q_{out}(t),$$

where Q_{in} and Q_{out} are the inflow and outflow of the tank. In the original linear case, the outflow is computed as a fraction of the contained volume:

$$Q_{out}(t) = k\,V(t).$$

Now, a series of n interconnected tanks is used to compute the outflow of a sewer pipe in terms of its inflow. Each tank has a volume $V_i(t),\ i = 1, \ldots, n$, and an outflow $Q^i_{out}(t) = k\,V_i(t),\ i = 1, \ldots, n$. The inflow to each tank is equal to the outflow from the previous one with the exception of the first tank, whose inflow is equal to the inflow to the sewer pipe being modelled. The outflow of the sewer pipe is then approximated by

the outflow of the last tank:

$$\begin{aligned}
\frac{\mathrm{d}V_1}{\mathrm{d}t} &= Q_{in}(t) - kV_1(t), \\
\frac{\mathrm{d}V_2}{\mathrm{d}t} &= kV_1(t) - kV_2(t), \\
&\vdots \\
\frac{\mathrm{d}V_n}{\mathrm{d}t} &= kV_{n-1}(t) - kV_n(t), \\
Q_{out}(t) &= kV_n(t).
\end{aligned}$$

Discretizing the system and assuming the inflow to be constant between two consecutive discretization time instants, the system can be solved analytically to obtain a closed expression of the outflow in terms of the inflow, avoiding the need to use all the intermediate volume and flow variables. A similar procedure can be carried out by linearly interpolating the inflow between two consecutive discretization time instants [Szilagyi, 2003, Szöllösi-Nagy, 1982].

Parameters k and n are related to the flow attenuation and delay represented by the model and must be calibrated from real data. A generalization for fractional values of n is described in Szilagyi [2005] based on using a different value of k for the last tank.

Another generalization of this model consists in assuming a nonlinear relation between the volume and outflow of each tank. In Meirlaen [2002], Meirlaen et al. [2001], for example, the relation is assumed of the form

$$Q_{out}(t) = \alpha(h(t) - \beta)^{\gamma},$$

with $h(t)$ the tank water level and α, β and γ constant parameters to be calibrated from real or simulation data. It is also assumed that the relation between the water level $h(t)$ and the cross-sectional area $A(t)$ is quadratic and that each tank represents a segment of the channel of length Δx, thus obtaining the following relations

$$\begin{aligned}
A(t) &= a\,h(t)^2 + b\,h(t), \\
V(t) &= \Delta x\,A(t)
\end{aligned}$$

where, a and b are constant parameters related to the catchment shape. By combining the previous equations with the tank mass balance equations, a system of ODEs in

terms of the tank water levels $h_i,\ i = 1, \ldots, n$, is obtained. Notice that the resulting system is nonlinear and therefore must be solved by means of numerical methods, since no analytical solution exists.

Virtual tank model

The *Virtual Tank* model is an aggregate model for large networks with complex topologies, where modelling individual sewer pipes may result in big systems of equations. To reduce the number of variables, a topological simplification of the network topology can be performed by dividing the network into catchments. Each catchment can then be modelled as a tank (a *virtual tank*) taking into account its total inflow and computing its outflow by means of a linear function of the tank volume, thus omitting all the catchment internal dynamics. In this approach, therefore, the flow in individual sewer pipes is not modelled.

The mathematical expressions are analogous to the tanks in series model. Two main differences can be highlighted between the two models. First, in the tanks in series model the tanks are placed in a row to model a single sewer pipe while in the virtual tank model the interconnections among the tanks are provided by the network topological simplification. Second, the tanks in series model is an hydraulic transport model while the virtual tank model must also take into account an hydrological phenomenon: the inflow due to the rainfall-runoff on the total catchment area.

A detailed description of this model is provided in Section 3.2.

2.4.5 Models Based on Identification

Identification approaches are a very general technique that can be used to obtain a model for a given input-output process from real data [Ljung, 1999]. Once the structure of the model is chosen, the involved parameters are obtained as the ones that best fit input-output data by means of either trial and error or optimization procedures. The most well-known approach is to assume a linear relation between the input $u(t)$ (inflow) and output $y(t)$ (outflow) of the system of the form:

$$y(t) + a_1 y(t-1) + \ldots + a_n y(t-n) = b_1 u(t-1) + \ldots + b_m u(t-m).$$

Different values of m and n are usually tested to finally select the most appropriate in terms of the balance between number of variables and accuracy.

These techniques applied to sewer network processes can be found in Eurén and Weyer [2007], Weyer [2001], Pleau et al. [2005] and Puig et al. [2009b].

2.5 Integrated Modelling

In the urban drainage systems community, the name *integrated modelling* refers to models that take into account sewer networks, WWTPs and receiving water bodies as a whole system. In integrated models not only water quantity is modelled, but also water quality, by means of the concentration of several dissolved and suspended chemical substances.

Integrated models are build by coupling sub-models describing the behaviour of different processes. Common sub-models of an integrated model include [Rauch et al., 2002]:

- Rainfall-runoff model
- Hydraulic model
- Water Quality model
  - Pollutant accumulation
  - Pollutant wash-off
  - Pollutant transport
  - Pollutant process
- Wastewater treatment plant
  - Flow propagation and mixing
  - Unit process
    - Clarifiers
    - Activated sludge
    - Biofilms

    - Anaerobic digestion

- Receiving water bodies
    - River Flow
    - Pollutant transport in rivers
    - Biochemical processes

These sub-models can either be executed one after the other or as a single whole. For example, once the hydraulic model has been run and flows and volumes over time are known, they can be used as inputs to run the pollutant, WWTP and receiving water bodies models. In other network instances, however, it might be necessary to run the hydraulic model coupled with the pollutant model and WWTP model if interactions between the WWTP and the sewer network require to do so.

The chemical and biochemical processes describing water quality involve complex physically-based models that can be suitable for simulation but are too computationally demanding for control purposes. Therefore, as with the hydraulic models case described in Section 2.4, simplified models are used for this purpose. Examples of control-oriented studies based on integrated models can be found in Solvi [2006], Meirlaen [2002], Meirlaen et al. [2001, 2002], Vanrolleghem et al. [2005] (based on the integrated model simulator WEST), Butler and Schütze [2005], Lau et al. [2002], Schütze et al. [2002] (based on the integrated model simulator SYNOPSIS), Rauch and Harremöes [1999a,b] (based on SAMBA/MOUSE simulator for the hydraulic sub-model) and Fu et al. [2008, 2009]. Both WEST and SYNOPSIS run the hydraulic module KOSIM, simulating flow in catchments by means of the tanks in series model and flow between catchments by translation.

WWTP and river water quality models have become standard in all integrated models after the work of research groups of the International Water Association (IWA). The current state-of-the-art models for WWTPs and river water quality are respectively the the Activated Sludge Model (ASM1, 2 & 3, Henze [2000]) and River Water Quality Model (RWQM1, Reichert [2001]). Since the focus of this thesis is the development of a control oriented model for the hydraulic sub-model of a sewer network, the description of these quality models is out of the scope.

A recent survey on integrated modelling including a historical overview, model classification, exhaustive literature review and future trends can be found in Bach et al. [2014]. Results of RTC implementations for several real case studies are detailed in Benedetti et al. [2013].

2.5.1 Emission- vs. Immission-Based Control

Much discussion has taken place in the sewer management community regarding whether the management objectives should be defined in terms of water quantity, thus requiring only hydraulic models (*emission*-based control), or quality, thus requiring integrated models (*immission*-based control). It has been shown that minimization of untreated water overflows may not correlate with the optimization of certain water quality indicators [Butler and Schütze, 2005, Lau et al., 2002, Rauch and Harremöes, 1999a] and, therefore, it is suggested that to meet the water quality standards imposed by legislation, quality models should be used [Meirlaen, 2002, Schütze et al., 2002, Solvi, 2006, Vanrolleghem et al., 2005]. On the other hand, others have pointed that quantity can still be used to define control objectives provided several considerations regarding the storage time in storm water tanks and capacity and characteristics of the WWTP are taken into account [Lau et al., 2002].

Integrated real-time control of sewer systems faces a main drawback regarding monitoring. Campisano et al. [2013] notice that water quality sensors still need to be improved since the harsh environmental conditions found in combined sewer pipes leads to serious maintenance problems and point at this fact as the cause for most real applications still defining management objectives in terms of water quantity (overflows, flood, CSOs) while quality-based control techniques have been proven mainly by simulation. The scarcity of water quality data for calibration purposes, robustness problems and extended computation times of integrated modelling have also been identified as reasons for emission-based control to be the most extended approach in real applications [Duchesne et al., 2004].

The model developed in this thesis describes the hydraulic part of the network and therefore the proposed controllers are of emission-based type. However, notice that, as the name suggest, integrated models are formed by coupling independent models. Therefore, due to the modular nature of integrated modelling, the model presented in

this thesis could be used as the hydraulic sub-model oriented integrated model. Notice also that, as developed in subsequent chapters in this thesis, one of the guiding lines in the development of the proposed model is the computational speed, which is one of the most appealing features of a model to be used in an integrated approach.

2.6 Real-Time Control

RTC is the common name for a number of techniques for the dynamic regulation of processes in response to measurements to meet operational objectives. RTC systems are structured in different levels depending on the abstraction of the decisions taken [Schütze et al., 2002]:

- Management level: definition of the management objectives and control strategy.
- System level: computation of set-points for local controllers, according to the objectives defined in the management level.
- Actuator level: local controllers regulate the actuators to achieve the set-points computed in the system level.

The actuator level is common in all RTC techniques and is composed by *control loops* consisting of *sensors* that monitor some system variables, *actuators* that can modify some system variables and *controllers* that adjust the actuators so that the controlled variables achieve the desired *set-point* values. A telemetry system must also be present to transfer data between these elements. In the case of RTC of sewer networks, common sensors and actuators include:

- Sensors
  - Level meters
  - Flow meters
  - Velocity meters
- Actuators
  - Storage elements

    - Detention tanks
    - Collectors with in-line detention capacity
  - Redirection elements
    - Gates
    - Weirs
    - Pumps
    - Valves
  - WWTPs

Further details on the characteristics of each type of sensor and actuator can be found in Campisano et al. [2013], where equipment for water quality control is also discussed.

The control algorithm in the system level of a RTC strategy is its fundamental and distinctive part. Given the management goals and the measurements of the current state of the network, in the system level set-points for all network actuators are computed. Notice that, assuming that the controllers in the actuator level are properly tuned and can achieve the desired set-points with suitable accuracy, the actual regulation of the network is decided by the control algorithm. Control algorithms can be classified according to the following relevant features [USEPA, 2006]:

- *Manual/Supervisory/Automatic*: depending on the degree of automation.
- *Local/Global*: local, if the set-points are computed at the actuator location depending only on local measurements or global, if decisions are taken at a central control station using measurements from the whole system.
- *Predictive/Reactive*: depending on whether forecasts are used to predict the future system state in addition to current and past measurements.
- *Model-based*: if a mathematical model of the process is executed on-line to compute set-points. Since a mathematical model is needed to compute predictions of the system, predictive controllers are model-based controllers.

- *Rule-based/Optimization-based*: depending on whether decisions are taken according to a pre-computed set of rules or according to the minimization/maximization of some quantitative measure of the system performance evaluated on-line (using a mathematical model).

In the following, the control algorithms in the system level of real-time controllers are discussed.

Unlike other control problems, the objective of the regulation of a sewer network is not to track a desired trajectory or set-point but to minimize the effects of an external disturbance, the rain inflows, over a system that, under normal operation, would not require any action. Moreover, the problem shows the following important features:

- Network structure.
- Presence of delayed phenomena.
- Uneven distribution of the disturbances both temporally and spatially.

These features suggest that a *local control* strategy, where actuators are only provided with monitoring information at their location, might not be a good option. On the contrary, a *global control* approach that takes into account the state of the whole network and is able to predict its future behaviour appears to be clearly more suitable.

The third feature in the list also suggests that a single strategy might not be suitable for all possible disturbances. To overcome this problem, predictions of the disturbance variable can be used to anticipate the future network state. However, these predictions are valid only within a short time window and must be updated constantly. According to these characteristics of the problem, it is clear that a *predictive controller* computing actions in *real time* taking advantage of the most recent measurements and rain forecasts would be the best option.

2.6.1 RTC of Combined Sewer Networks

At the light of the previous discussion, control strategies for the regulation of sewer networks in presence of intense rain episodes are usually *global, predictive, real-time control* strategies [Schütze et al., 2004]. Since the objective of the management is

to minimize the effect of the disturbance on the system optimization-based control appears to be the natural solution, and is, in fact, the most common. However, several rule-based solutions have also been proposed, especially in the integrated modelling community [Meirlaen, 2002, Meirlaen et al., 2001, 2002, Solvi, 2006, Vanrolleghem et al., 2005].

A comparison of the cost-efficiency of RTC implementations with respect to static solutions based on infrastructure development (separation of wastewater and stormwater sewer systems, increasing sewer pipes capacity, construction of detention tanks) for CSO reduction in combined sewer networks can be found in Beeneken et al. [2013], Dirckx et al. [2011a,b].

2.6.1.1 Optimization-based Controllers for Sewer Networks

Optimization-based control is a type of model-based control that uses the mathematical model of the system to formulate a *constrained optimization* problem (called the Optimal Control Problem), the solution of which provides the optimal control actions that minimize[1] a quantitative measure of the system performance. The equations of the system dynamics appear as constraints of the optimization problem and the manipulated variables (the actuators set-points) are left free, usually between some physical constraints. Each configuration of the manipulated variables propagates through the system equations in the constraints of the optimization problem to result in a unique sequence for all the system variables. These sequences can be used to measure the system performance through an *objective function*. The task of the optimization algorithm (*solver*) consists in systematically evaluating possible configurations of the manipulated variables to come up with the one that results in a minimum value of the objective function. A very appealing property of optimization-based control is that physical and operational constraints on the system variables can be easily taken into account, since they are naturally handled by optimization algorithms.

[1]Performance indicators for a system can either indicate "how good" or "how bad" it is performing. Depending on the case, the performance indicator should be minimized or maximized. In the following it will be assumed that the performance index is to be minimized. Notice that the minimum of a function $f(x)$ equals minus the maximum of the function $-f(x)$, thus any maximization problem can be turned into a minimization one.

Depending on the type of model and optimization algorithm used, two main paradigms for RTC of sewer networks appear in the literature. In the first one, authors develop their own model, formulate the corresponding optimization problem and use derivative-based algorithms to solve it. In the second one, the model is provided by some fast simulation software, with the consequence that no direct access to the equations is possible and the resulting optimization problem is non explicit and must be solved by derivative-free solvers.

In the first group, the work of Gelormino and Ricker [1994] is one of the first references: it uses a reservoir-based aggregated model of the network to formulate and solve a quadratic optimization problem. This approach was further developed in Ballester Rodés et al. [1998] in the so called *virtual tank model*. The virtual tank model was later improved by adding overflow variables and additional hydraulic structures (this version is presented in Section 3.2) and is the basis of the works in Cembrano et al. [2004], Ocampo-Martínez [2011], Ocampo-Martínez and Puig [2010], Ocampo-Martínez et al. [2013], Puig et al. [2009a] in which the resulting OCP are solved using quadratic programming or gradient-based nonlinear programming methods. A similar catchment-based approach to minimise CSOs is developed in Vezzaro and Grum [2012, 2014] and later used in an integrated modelling approach to include water quality optimization in Vezzaro et al. [2014]. In these works, although no external software is used and the model equations are available, global optimization is used to simply deal with non-linearities and piecewise functions.

In the work of Marinaki and Papageorgiou [1998, 2001, 2003, 2005] a discrete-time translation and delay model with nonlinear piecewise functions for weir flow and overflows is solved by an ad-hoc *iterative feasible direction* algorithm. A similar approach is presented in Fradet et al. [2010], Pleau et al. [2005, 2010]. On the other hand, in Duchesne et al. [2001, 2003, 2004] a nonlinear model based on switching between discretizations of the kinematic wave and diffusive wave equations is used together with a gradient method.

The use of available commercial or research software is very common among the integrated modelling community. Since the different sub-models involved in the integrated modelling require high level of expertise in each topic (hydrology, hydraulics, chemistry, biochemistry), it is easier to use sub-models developed independently and

plug them together by means of interfacing software. Aggregated models are common in this context to speed-up computation times, even for small or mid-scale networks. Once the whole system can be evaluated, it can be used together with global derivative-free solvers that by means of massive evaluation of the model come up with an optimal solution. To this end, the model evaluation must be performed very fast, therefore simplified models are common in this area, although complex models can be used too if the network dimensions are small. A popular choice for global derivative-free optimization solvers used in this context are Genetic Algorithms [Butler and Schütze, 2005, Lau et al., 2002, Rauch and Harremöes, 1999a,b, Schütze et al., 2002] although random search algorithms and other evolutionary algorithms have been reported [Schütze et al., 2002].

2.6.2 Receding Horizon Control

RHC also known as Model Predictive Control (MPC) or model-based predictive control is an optimization-based control strategy consisting in solving on-line a series of finite horizon model-based optimal control problems based on the most recent available measurements of the system and, eventually, the most recent disturbance forecasts. At a given time step, expressions of the future states of the system in terms of the initial conditions and the future control actions are used as constraints of an optimisation problem whose objective is to minimize a performance index.

In the following, an outline the whole RHC algorithm and the structure of the OCPs solved in each RTC iteration is given, based on the algorithm description in Camacho and Bordons [2004], Maciejowski [2002], Rawlings and Mayne [2009]. Further details and modifications to adapt the algorithm to specific properties of the model presented in this work will be presented in Sections 5.5 and 5.6 and in Chapter 6. To proceed with the algorithm description, let the following vector variables be defined:

- State variables x: these variables describe the dynamic evolution of the system. These evolution is affected by the exogenous disturbances and the manipulated variables as described by means of the dynamic equations of the system

$$x(t+1) = f(x(t), u(t), w(t)),$$

In the case of a sewer network, these variables describe flows through sewer pipes and volumes in reservoirs and collectors.

- Manipulated variables u: also known as input variables, these are the variables through which it is possible to influence the system behaviour. The objective of the optimization problem is to compute their value. In the case of a sewer network, these variables describe flows through gates, flows over weirs and flows through pumps and valves.

- Output variables y: these are the variables that are measured during the system evolution and are used for feedback control. They are expressed as a function of the state variables and the manipulated variables:

$$y(t) = g(x(t), u(t)).$$

  Measurement noise can also be added to the output equation, though it will not be considered in this thesis. In the case of a sewer network, these variables usually correspond to some of the water levels or flows in sewer pipes and volumes in tanks.

- Exogenous disturbances w: these variables describe external effects of the environment over the system. It is not possible to modify their dynamics. In fact, they are included into the optimization problem as fixed values, provided by means of forecast techniques. In the case of a sewer network, these variables describe rainfall-runoff inflows to the network.

Notice that, in the above description of the system variables and dynamics, a discrete time approach has been used, with discrete time variables t describing the variable values at $t \cdot \Delta t$ seconds after the start of the modelled event, where Δt is the discretization time step.

In addition to the above relations, physical and operational constraints can be taken into account in the OCP formulation:

$$x^{min} \leq x \leq x^{max},$$
$$y^{min} \leq y \leq y^{max},$$
$$u^{min} \leq u \leq u^{max}.$$

These inequalities applied to vector variables indicate element-wise relations.

By imposing the dynamic and output equations and the constraints over a *prediction horizon* of H future time steps, the OCP at time step t, $P(t)$, can be written as:

$$
\begin{array}{lll}
\min & J(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{w}), & \\
\text{s.t.} & x(t) = \hat{x}(t), & \\
& x(k+1) = f(x(k), u(k), w(k)), & k = t \dots t+H-1, \\
& y(k) = g(x(k), u(k)), & k = t \dots t+H, \\
& x^{min} \leq x(k) \leq x^{max}, & k = t \dots t+H, \\
& y^{min} \leq y(k) \leq y^{max}, & k = t \dots t+H, \\
& u^{min} \leq u(k) \leq u^{max}, & k = t \dots t+H.
\end{array}
\qquad (P(t))
$$

where $\hat{x}(t)$ is the measured or estimated state at time step t and

$$
\begin{aligned}
\mathbf{x}(t) &= (x(t)^\top, x(t+1)^\top, \dots, x(t+H)^\top)^\top, \\
\mathbf{y}(t) &= (y(t)^\top, y(t+1)^\top, \dots, y(t+H)^\top)^\top, \\
\mathbf{u}(t) &= (u(t)^\top, u(t+1)^\top, \dots, u(t+H)^\top)^\top, \\
\mathbf{w}(t) &= (w(t)^\top, w(t+1)^\top, \dots, w(t+H)^\top)^\top.
\end{aligned}
$$

If the system is properly defined, the state and output evolution is uniquely determined by the sequences of manipulated variables and disturbances and by the initial condition, so in fact

$$J(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t), \mathbf{w}(t)) = \tilde{J}(\hat{x}(t), \mathbf{u}(t), \mathbf{w}(t)).$$

This relation cannot be obtained in an explicit form in the general case and it is implicitly implied by the constraints. In the linear case (that is, when f and g are linear functions), however, it is possible to express all the relations in terms of the manipulated variables, disturbances and initial conditions to obtain an OCP with a reduced number of variables [Maciejowski, 2002].

Now, the RHC algorithm works as follows. At time step t:

- Solve OCP $P(t)$,
- Apply to the system the control actions corresponding to the first time step of the solution sequence $u(t)$,
- Let the system evolve one time step ahead in time (i.e., Δt seconds),

- Measure/estimate the current system state $\hat{x}(t+1)$ and compute new disturbance forecastings $\mathbf{w}(t+1)$,
- Set $t := t+1$ and repeat the procedure.

Notice that, in subsequent problems the prediction time window moves forward in time: the first one covers from t to $t+H$, the second one from $t+1$ to $t+H+1$ and so on: hence the name *receding horizon control*.

When the RHC algorithm is used as the system level controller of an overall management strategy, the values of the manipulated variables resulting as the solution of each OCP are used as set-points for the local controllers at the actuator level. Therefore, in this context, it is also possible that measures of the manipulated variables $\hat{u}(t)$ are performed to check whether the local controllers managed to reach the set-points and take into account this information to formulate the next OCPs.

2.7 Summary

In this chapter, an overview of the most common techniques used for hydraulic modelling and control of open channel flow systems have been presented with an emphasis on those studies devoted to sewer network control. Depending on the specific problem characteristics and objectives different combinations of models and control techniques can be used. To take advantage of measurements and rain forecasts, predictive model-based Receding Horizon Control is the most widespread approach for sewer network regulation.

For pure hydraulic models with quantity-based objectives, more accurate models can be used and the corresponding optimization problems can be solved by means of efficient derivative-based solvers. Studies have shown that quantity based control may turn into slightly suboptimal pollution results [Butler and Schütze, 2005, Lau et al., 2002, Rauch and Harremöes, 1999a], but the implementation of such controllers is the most widespread due to ease of implementation and model and sensor reliability. On the other hand, integrated models taking into account quality variables lead to complex models that produce non-explicit optimization problems that must be solved by global derivative-free algorithms. Although theoretically, optimal management of pollution

indices can be achieved with integrated control, this approach is still not commonly adopted in real applications mainly due to instrumentation problems and mistrust in the complex models involved, though much research effort is currently being put to overcome this difficulties.

In this thesis, a *hybrid linear delayed model* has been chosen as the modelling framework for the sewer network description since it efficiently deals with three main aspects of the problem. Firstly, the presence of *delays* in the model is a common element in any water transportation model. Secondly, the *hybrid* approach allows to model the presence of overflows in the network, which only occur when a given flow is above a threshold value, thus according to a logical condition. Finally, the *linear* framework is especially suited to guarantee acceptable computation times, since sewer networks usually have a high number of variables, making the systems belong to the *large-scale* class. Such a model leads to MILP optimization problems which can be efficiently solved with appropriate solvers. The model has been developed to be used in quantity based control, thus taking only into account the system hydraulics. However, since computational speed has been one of the main factors taken into account in selecting the model features, it could be used as the hydraulic module of an integrated model for quality-based control.

Part II

Modelling

Chapter 3

Sewer Network Modelling

This chapter is partially based on:

- *B. Joseph-Duran, M. Jung, C. Ocampo-Martinez, S. Sager, and G. Cembrano. Minimization of sewage network overflow.* Water Resources Management*, 28(1): 41–63, 2014a.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid modeling and receding horizon control of sewer networks.* Water Resources Research*, 2014d.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid control-oriented modeling of combined sewer networks: Barcelona case study.* Hydroinformatics Conference*, 2014c. New York, USA.*

3.1 Introduction

In this chapter, a novel model for sewer networks is presented: the Hybrid Linear Delayed (HLD) model. The modelling principles and techniques of the HLD model are based in the so called Virtual Tank (VT) model. The VT model is a conceptual model that has been used as a control-oriented model for the Barcelona sewer network since the late 90s, both in industry [Ballester Rodés et al., 1998] and academia [Cembrano et al., 2004, Joseph-Duran et al., 2014a, Ocampo-Martínez, 2011, Puig et al., 2009a]. The VT model is also the control model behind the development of the sewer network control tool CORAL (Spanish for Optimal Control of Sewer Networks) [Figueras et al., 2002, Puig et al., 2009a]. The main objective of the VT model, based on an early work on the topic by Gelormino and Ricker [1994], is to be computationally suitable for computation of control actions in real time for large-scale networks, inspired by the

characteristics of the Barcelona sewer network, involving 1450 km of sewer pipes within an area of 98 km^2 [Ballester Rodés et al., 1998].

From the control point of view, the VT model is designed to be used as a system level global controller. The values corresponding to gate flow variables computed by the OCPs based on the model are used as set-points for local PID controllers at the network gates that regulate the gate position as a function of the measured outflow or water level, depending on the available sensors.

As outlined in the Section 2.4, the VT model is based on a conceptual approximation of the water distribution and transport along the network by means of dividing the network into catchments. By means of a tank model, the storage time and wave attenuation inside the catchment is taken into account. Notice that this is a highly conceptual model, where not only the transport phenomenon is conceptualized, but also the problem topology, by means of the relations and interactions between elements.

The HLD model developed in this work aims to cope with several drawbacks of the VT model. Although control experiences to minimize flooding and CSOs based on the VT model have proven successful in several studies [Cembrano et al., 2004, Puig et al., 2009a] by evaluating the performance of the model-based controller against a physically-based model simulator, it has been noted that the division of the network into catchments to be modelled as virtual tanks and the calibration of some of the parameters associated to these catchments is not straightforward. In the mentioned studies, a proper topology and parameter calibration (also using on-line calibration procedures) was possible thanks to the involvement of the company responsible of the sewer network management, who provided data and expert knowledge of the network.

In the HLD model a conceptual model for the hydraulics of each particular sewer pipe is presented, thus avoiding the need of strong topological simplifications. The model still makes use of the sub-models for hydraulic structures already presented in the latest versions of the virtual tank model [Ocampo-Martínez, 2011], such as tanks, weirs and overflow points, though some parameters have been added to improve the model accuracy. This approach allows for simpler implementation and parameter calibration than the VT model approach and provides a more detailed description of the network dynamics.

The HLD model must not be seen as a complete paradigm shift with respect to the VT model, but as an extension. Although in this thesis emphasis is made on the hydraulic part of the presented model (using an already established rainfall-runoff model for the hydrologic part), the HLD model is fully compatible with the virtual tank model. Therefore it would be possible to model parts of a network with virtual tanks and other parts using the hydraulic description of the HLD model, thus adapting the modelling to the needs of the specific case study and data availability.

With the aim to outline the modelling principles and techniques that have led to the HLD model developed in this thesis, the VT model is first presented.

3.2 Virtual Tank Model

In the *virtual tank* model approach, sewer networks consist of several elements, which are described in the following. For water storage, there are *real* water retention *tanks* built by the network operator and so-called *virtual tanks*, each of which representing a set of sewage collectors for a specific zone of the city. According to Ocampo-Martínez [2011], a virtual tank can be defined as follows:

> "At any given time, let the *virtual tank* be a storage element that represents the total volume of sewage inside the sewer mains associated with a determined sub-catchment of a given sewer network. The sewage volume is computed via the mass balance of the stored volume, the inflows and the outflows related to the sewage mains, and the equivalent inflow associated with rainwater."

Then, there are *sewer pipes* to connect the different tanks and virtual tanks, which can be partly controlled with pumps and valves. In some sewer pipes, there are *redirection gates* to manipulate and redirect the flow. Other sewer pipes are connected by simple junctions. Both these structures are treated as tanks with zero maximum volume and where all inflow is directly forwarded as outflow. These sewer networks can be displayed as directed graphs [Joseph-Duran et al., 2014a]. A conceptual example of such a network is displayed in Figure 3.1.

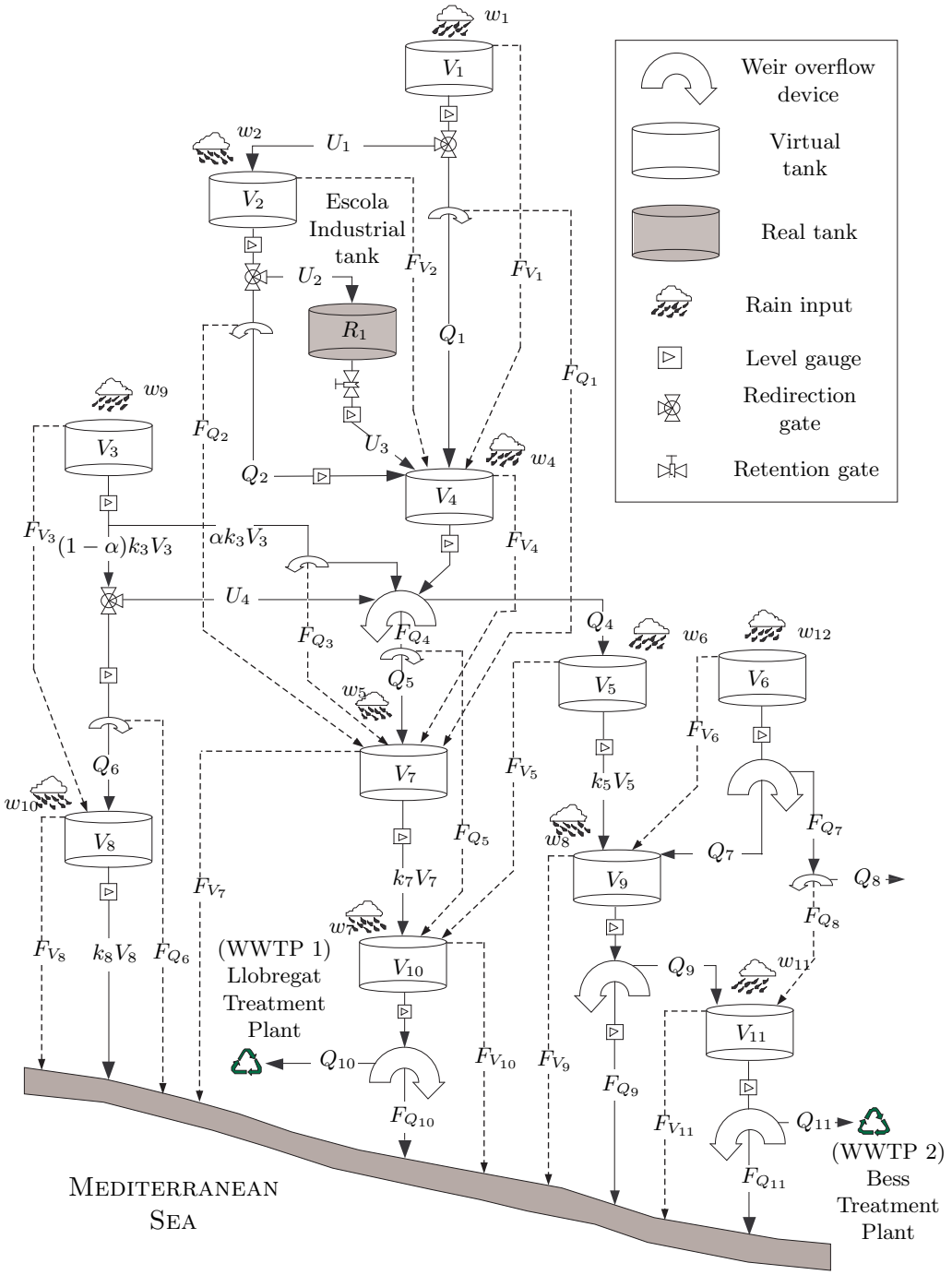

Figure 3.1: Diagram of the virtual tank model of a part of the Barcelona sewer network including 11 virtual tanks and 1 real tank. Taken from Ocampo-Martínez and Puig [2010].

The system states are the volumes $R(t)$ [m^3] in each real tank and the volumes $V(t)$ in each virtual tank [m^3]. The uncontrolled (gravity-driven) intercommunicating flows between tanks are denoted $Q(t)$ [m^3/s] and the controlled ones $U(t)$ [m^3/s]. The discrete time variable t accounts for the number of time steps of duration Δt [s] since the start of the modelled event. For each tank of the system, there is only one regular outflow – controllable or not. The outflows that completely exit the system are the desired flows to WWTPs and undesired CSOs into the receiving environment.

All the controlled flows are limited due to physical constraints such as sewer pipe sizes and pump capacities. It is assumed that the volumes of the real detention tanks are limited as well and they cannot overflow since they are often placed underground and not connected to the surface. Their inflows are hence always controlled to prevent overflow, which physically could happen. In real applications, an overflow emergency mechanism is always present in case there is a malfunction in the controlled devices, but this special situation is not taken into account in the VT model, as is usual in control-oriented models.

3.2.1 Real Tank Equations in the Virtual Tank Model

The volume $R(t)$ stored in a real tank is modelled with a forward Euler discretization of the mass balance equations with time step Δt:

$$R(t+1) = R(t) + \Delta t \left(U_{in}(t) - U_{out}(t)\right),$$

where $U_{in}(t)$ and $U_{out}(t)$ are controlled the in- and out-flows of the tank.

3.2.2 Virtual Tank Equations in the Virtual Tank Model

For the volume $V(t)$ in a virtual tank, the same mass balance equations as for real tanks is applied, with several considerations regarding the inflows $Q_{in}(t)$ and outflows $Q_{out}(t)$

$$V(t+1) = V(t) + \Delta t \left(Q_{in}(t) + w(t) - Q_{out}(t)\right) \tag{3.1}$$

Here, $Q_{in}(t)$ accounts for the total network inflow to the catchment modelled by the virtual tank (possibly the sum of several controlled or uncontrolled flows from different catchments) and $w(t)$ for the external rain inflow, obtained by multiplying the rain

intensity $I\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$ by the catchment area $A\left[\mathrm{m}^2\right]$ and scaling with a dimensionless ground absorption coefficient ϕ (calibrated on-line) to account for infiltration losses [Puig et al., 2009a]:

$$w(t) = \phi\, A\, I(t). \tag{3.2}$$

This is the hydrologic rainfall-runoff sub-model of the VT model, that is, the model that computes the net inflow to the network from rain intensity data or measurements.

Finally, the outflow $Q_{out}(t)$ of a virtual tank is computed as proportional part of the volume stored in it, i.e.,

$$Q_{out}(t) = k\, V(t),$$

where parameter k is obtained from historical sensor data, or to be calibrated online in a real-time control approach. Figure 3.2 shows how the flow delay and attenuation are modelled by means of parameter k.

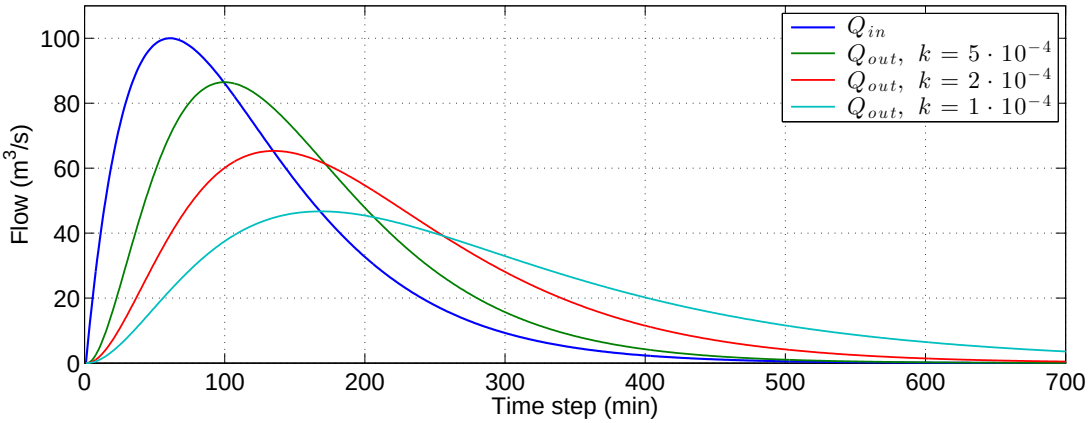

Figure 3.2: Inflow to a virtual tank (assuming null rain inflow) and corresponding outflow for several values of parameter k.

3.2.3 Mass Balance Equations in the Virtual Tank Model

Flows coming from virtual tanks or controlled devices can join and split at junctions where a mass balance equation must hold:

$$Q_{in}(t) = Q_{out}(t), \tag{3.3}$$

where $Q_{in}(t)$ is the total inflow to the junction and $Q_{out}(t)$ the total outflow. These total inflows and outflows can be the sum of several controlled and uncontrolled flows.

If there are several outflows present at the junction, each one takes a proportional part of the total inflow. The proportionality factors for each outflow can be calibrated off-line, using historical data or data generated by a physically-based model, or on-line, using the last available measurements.

3.2.4 Overflow Equations in the Virtual Tank Model

During normal operation, the network can easily transport the sewage towards the treatment plant. However, in the presence of heavy rain, it may happen that there exists no viable flow path for the network to process all the incoming water. In these scenarios, overflow happens and flow paths appear, which were not present before and depend on the system state and inputs. These overflow paths are represented as dash-dotted lines in Figure 3.1.

In a virtual tank, overflow happens if the volume obtained by the volume equation (3.1) aims to exceed the maximum capacity V_{max}. Similarly, in a junction, overflow happens if the total inflow exceeds the maximum total outflow capacity. In both cases, all the excess volume is considered as overflow.

To take into account virtual tank overflows $F_V(t)$, the mass conservation equation (3.1) of the corresponding virtual tank must be changed to

$$V(t+1) = V(t) + \Delta t \left(Q_{in}(t) + w(t) - Q_{out}(t) - F_V(t)\right). \tag{3.4}$$

Then, the overflow variable $F_V(t)$ can be modelled with logical decisions:

$$\begin{aligned} \text{IF} \quad & V_{max} \leq V(t) + \Delta t \left(Q_{in}(t) - Q_{out}(t)\right) \\ \text{THEN} \quad & F_V(t) = \frac{1}{\Delta t}\left(V(t) + \Delta t \left(Q_{in}(t) - Q_{out}(t)\right) - V_{max}\right) \\ \text{ELSE} \quad & F_V(t) = 0, \end{aligned} \tag{3.5}$$

where the THEN expression, together with equation (3.4), also sets $V(t+1) = V_{max}$. Virtual tank overflows can also be modelled directly with the maximum function

$$F_V(t) = \frac{1}{\Delta t} \max\left\{0, V(t) + \Delta t \left(Q_{in}(t) - Q_{out}(t)\right) - V_{max}\right\}.$$

Then, the system equations are piecewise affine equations and thus nonlinear and non-differentiable. Therefore, each system directly containing these constraints becomes

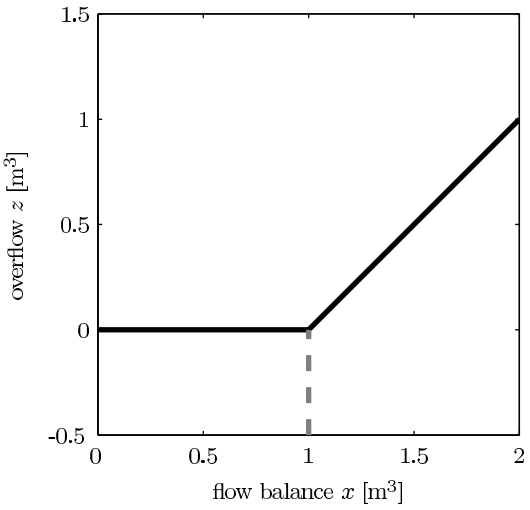

Figure 3.3: The overflow function $F_V(x) = \max\{0, x - 1\}$ with maximum capacity $V_{max} = 1$ and water volume change x. It is nonlinear, nonconvex and nondifferentiable in the kink at $x = 1$.

not only nonlinear but also nonconvex and nondifferentiable. One such constraint is displayed in Figure 3.3.

Analogous to the modelling of virtual tank overflows $F_V(t)$, overflows in junctions $F_Q(t)$ are modelled by considering a joint as a tank with zero capacity, and hence constant volume of 0. The mass balance equation (3.3) is modified to take into account the overflow contribution,

$$Q_{in}(t) = Q_{out}(t) + F_Q(t),$$

and the overflow variables are defined by means of logical statements so that overflows occur when the total inflow to the joint $Q_{in}(t)$ exceeds the maximum outflow Q_{max}:

$$\begin{aligned} \text{IF} \quad & Q_{in}(t) \geq Q_{max} \\ \text{THEN} \quad & F_Q(t) = Q_{in}(t) - Q_{max} \\ \text{ELSE} \quad & F_Q(t) = 0. \end{aligned}$$

Again, the THEN expression together with the mass balance equation on junctions sets $Q_{out}(t) = Q_{max}$. As in the virtual tank overflow case, the overflow in a junction can be modelled using the maximum function

$$F_Q(t) = \max\{0, Q_{in}(t) - Q_{max}\}.$$

Both the virtual tank overflows $F_V(t)$ and the junction overflows $F_Q(t)$ can be redirected as inflows to other virtual tanks, i.e., to other sewer catchments, or to the receiving environment in the form of CSOs. Only in the latter case, the overflow volume leaves the network permanently.

In Joseph-Duran et al. [2014a], a more detailed description of the virtual tank model together with several reformulations to obtain optimal control problems are described. The piecewise equations for overflows can be reformulated by means of nonlinear smoothing of the nondifferentiable equations or through the introduction of integer variables (by means of the Mixed Logical Dynamical Systems approach or the Generalized Disjunctive Programming approach). The paper also presents an *ad hoc constraint branching* algorithm to directly optimize the system of piecewise affine equations obtained through the formulation presented in this section. Results comparing the solutions and computational times obtained with each approach are also provided.

3.3 Hybrid Linear Delayed Model

The Hybrid Linear Delayed model has been developed as an extension of the VT model, to overcome some difficulties related to the model construction when applied to specific case studies. These problems are not related to the performance of the controllers based on the model, but to the simplification of the network topology and the estimation of some parameters. The division of a network into catchments is a difficult task if the network is highly interconnected. Notice that, since each virtual tank has only one outgoing flow, network parts with complex interconnections must be necessarily aggregated and modelled as a single catchment. This can lead to big catchments for which the rain inflow computed by the hydrologic model in equation (3.2) might not be accurate due to inhomogeneous distribution of the rain intensity along the catchment area. The area of the catchment and the ground absorption coefficient in equation (3.2) can also be difficult to estimate for big catchments since it might not be clear which is the geographical extension on which rain has a relevant influence to the network inflow and since the physical characteristics of a big area might not be homogeneous. On the other hand, deliberately omitting interconnections among catchments in order to obtain smaller areas can lead to inaccurate approximations of the volume contained in the catchment, and therefore inaccurate outflow approximations. Finally, even when

the network division into catchments modelled by virtual tanks is available, the definition of overflow paths among these catchments is, again, not evident. Depending on the rain inflow distribution, different overflows with different intensities might be happening at different points within each catchment that should be redirected to different destinations, according to the geometric properties of the catchment surface near the overflow point. Since in the VT model, overflows are defined for entire catchments, with overflow paths being static, the definition of these paths can be difficult, especially for big catchments.

To overcome these difficulties, the HLD model directly describes the flows and volumes in each particular network element: flows through sewer pipes and through hydraulic structures (weirs and gates), overflows in junctions, runoff flows returning to the network after overflow events and the volumes in detention tanks and collectors (big sewer pipes with in-line storage capacity). Therefore, the need for topological aggregation is avoided.

Moreover, the hydrological rainfall-runoff model computing the rain inflow to the network is completely separated from the hydraulic part of the network and could, therefore, be substituted by any other similar model. This is not the case for the VT model, in which the rainfall-runoff computation is directly linked to the area and characteristics of the catchment described by the virtual tank (c.f. equations (3.1) and (3.2)), which is, in turn, an element conceptually describing hydraulic phenomena such as flow attenuation and delay. The selected hydrological rainfall-runoff sub-model for this study computes the inflow to some network junctions through which rain inflow can enter the network. Therefore, the area of rain influence to these junctions is smaller than in the VT model and can be better approximated. Moreover, the presence of additional parameters in the rainfall-runoff model allows for additional flexibility of the model to fit measurements during parameter calibration.

Analogous to the rain inflows, overflows in the HLD model are defined to happen at the junction level. A new feature of the model is that, since the area of influence of an overflowing junction is composed only of its surroundings, overflow volume is defined to enter at the same junctions where it went out, once the overflow event has finished. A simple modification could be added to allow overflow paths to nearby junctions, but this situation will not be considered in the following description for the sake simplicity.

Summarizing, the HLD model provides the following advantages with respect to the VT model, especially regarding the model setup:

- The control model network topology is obtained directly from the real network.
- Straightforward off-line parameter calibration can be performed using real data or data generated by a physically-based model simulator.
- Switching phenomena (weir flows, overflows and flows re-entering the network after overflow) are described with better accuracy since they are modelled at the sewer pipe/junction level instead of at a catchment level.
- The hydrological rainfall-runoff sub-model is independent of the hydraulic sub-model.
- Further insight on the network dynamics is obtained, since flow approximations in all network parts are provided.

Although the HLD model has been developed based on the data and characteristics of a specific case study (described in Section 4.1), the modelling approach is to model each element independently so that the model can be used in a wide range of network instances. The considered sub-models for different network elements/phenomena are described in this section. The main model variables are listed in Table 3.1. Further auxiliary variables will be defined in Section 5.2 to describe the switching nature of some flows.

All the parameters involved in the equations of the following model description can be calibrated using data generated with a simulator based on the full partial differential equation (PDE) physically-based model as described in detail in Section 3.4.

3.3.1 Flow Equations in the Hybrid Linear Delayed Model

The flow model describes the main features of the water transport along a sewer network: the mass balance in junctions and the flow delay and attenuation along sewer pipes. To this end, as shown in Figure 3.4, for each sewer pipe, two flows are considered: the upstream flow $q^{in}(t)$ (or inflow) and the downstream flow $q^{out}(t)$ (or outflow).

Table 3.1: Notation for the variables of the system.

Description	Symbol	Units	Indexing
Flow entering sewer pipes	$q_i^{in}(t)$	m^3/s	$i = 1 \dots n_q$
Flow leaving sewer pipes	$q_i^{out}(t)$	m^3/s	$i = 1 \dots n_q$
Volume in tanks	$v_i(t)$	m^3	$i = 1 \dots n_v$
Flow under gates	$g_i(t)$	m^3/s	$i = 1 \dots n_g$
Flow over weirs	$w_i(t)$	m^3/s	$i = 1 \dots n_w$
Overflows	$f_i(t)$	m^3/s	$i = 1 \dots n_f$
Overflow volume	$v_f^i(t)$	m^3	$i = 1 \dots n_f$
Flood runoff flow	$q_f^i(t)$	m^3/s	$i = 1 \dots n_f$
Collector volume	$v_c^i(t)$	m^3	$i = 1, \dots N$
Collector flow	$q_c^i(t)$	m^3/s	$i = 1, \dots N$
Collector overflow	$f_c(t)$	m^3/s	-
Rainfall-runoff inflow	$r_i(t)$	m^3/s	$i = 1 \dots n_r$

Mass balance equations

The mass balance equations describe how, at each network junction, the total inflow must equal the total outflow. For each sewer pipe $i = 1 \dots n_q$, the total inflow is computed as the sum of all inflows at the junction where it is connected. Therefore, the flow upstream of each sewer pipe is now defined as a fraction $\alpha_i \in (0, 1]$ of the total inflow coming from other sewer pipes, weirs, gates or rain inflows plus the effects of overflow $f(t)$ and flooding runoff $q_f(t)$ (described in the following sections):

$$q_i^{in}(t) = \lambda_i \left(\sum_{j=1}^{n_q} a_{ij}^q q_j^{out}(t) + \sum_{j=1}^{n_w} a_{ij}^w w_j(t) + \sum_{j=1}^{n_g} a_{ij}^g g_j(t) + \sum_{j=1}^{n_c} a_{ij}^r r_j(t) + \sum_{j=1}^{n_f} a_{ij}^f f_j(t) + \sum_{j=1}^{n_c} a_{ij}^{q_f} q_{f_j}(t) \right). \tag{3.6}$$

See Table 3.1 for a description of the variables related to the different flows involved in this equation. Coefficients $a_{ij}^q,\ a_{ij}^r,\ a_{ij}^{q_f} \in \{0, 1\}$, $a_{ij}^f \in \{0, -1\}$ and $a_{ij}^w,\ a_{ij}^g \in \{0, 1, -1\}$ indicate which elements are interconnected; therefore they contain the topological information of the network. The value of $\lambda_i \in (0, 1]$ describes whether sewer pipe i is the only outgoing sewer pipe from the junction where it is connected ($\lambda_i = 1$) or there

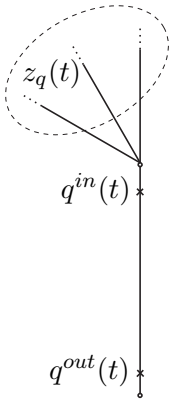

Figure 3.4: Flow model diagram.

are other outgoing sewer pipes ($\lambda_i < 1$). In the latter case, the λ parameters of all the outgoing sewer pipes of a junction should add up to 1 for mass conservation.

To simplify the notation in the description of the sub-models for different elements in the rest of this chapter, it is convenient to define also here the net inflow to a junction as:

$$\begin{aligned} z_{q_i}(t) = &\sum_{j=1}^{n_q} a_{ij}^q q_j^{out}(t) + \sum_{j=1}^{n_w} \max\{0, a_{ij}^w\}\, w_j(t) + \\ &\sum_{j=1}^{n_g} \max\{0, a_{ij}^g\}\, g_j(t) + \sum_{j=1}^{n_c} a_{ij}^r r_j(t), \end{aligned} \tag{3.7}$$

where, now, all the coefficients are either 0 or 1. Notice that, for sewer pipes i and j connected to the same upstream junction, $z_{q_i}(t) = z_{q_j}(t)$.

Flow equations

To account for transport delays and flow attenuation, the flow downstream of each sewer pipe is computed as a convex combination of the upstream flows at two consecutive previous time steps. Hence, for each sewer pipe $i = 1 \dots n_q$,

$$q_i^{out}(t) = a_i\, q_i^{in}(t - t_i) + (1 - a_i)\, q_i^{in}(t - t_i - 1), \tag{3.8}$$

with $a_i \in (0, 1]$. This model has been chosen because the delay in sewer pipes may not be a multiple of the sampling time unless the latter is chosen to be very small, which would lead to a high number of variables in the problem to cover reasonable

simulation or optimization time windows. By means of a convex combination of flows at two consecutive time steps, delays of any magnitude can be suitably approximated.

The flow attenuation introduced by this equation can be regarded as a consequence of the discretization, as it vanishes as the sampling time approaches zero. However, taking into account the usual time steps used in sewer network control (1 to 5 minutes), this attenuation has a noticeable positive effect in the model accuracy, especially as the flow travels through many sewer pipes and the attenuation phenomena accumulate.

It is also worth noticing the fact that coefficients a_i and $1-a_i$ add to 1 implies that the model is mass conservative in the sense that the total flow entering the sewer pipe equals the total flow leaving it.

3.3.2 Tank Equations in the Hybrid Linear Delayed Model

To match the discrete-time equations of the flow transport, the volumes $v(t)$ in the network tanks are described by the following discretization of the volume equation, with sampling time Δt:

$$v(t) = v(t-1) + \Delta t\big(g_{in}(t-1) - g_{out}(t-1)\big). \tag{3.9}$$

For ease of notation, it is assumed that both the inflow $g_{in}(t)$ and outflow $g_{out}(t)$ of the tanks are controlled by gates. However, the model could be easily extended to consider inflows and outflows from sewer pipes or weirs.

3.3.3 Weir Equations in the Hybrid Linear Delayed Model

Weirs are flow-regulation hydraulic structures that divert part of the inflow of a sewer pipe to a secondary sewer pipe, called a spillway, when a certain water level is reached. Moveable weirs have many applications in rivers and irrigation channels, where they are used to regulate the flow and water levels, to activate the in-line detention capacity of the channel or as safety elements to redirect part of the flow outside the channel. In the present model, movable weir flows are considered as controlled variables in the same way as gate flows: to run the model for simulation, weir overflow threshold values have to be provided as inputs while for optimal control purposes weir flows are left as free variables to be computed by the optimization problem and later be used as set-points

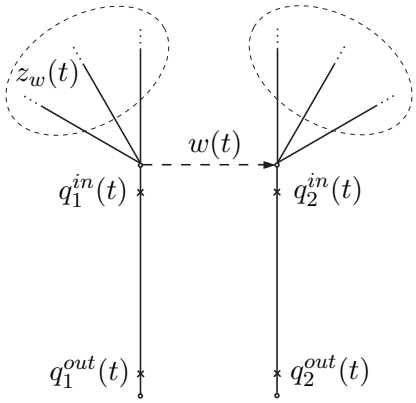

Figure 3.5: Weir model diagram.

for local PID controllers. Therefore, in the following only weirs with fixed position will be modelled.

As described in Section 2.3.4, in physically-based models, the flow over a weir is, in fact, a function of the square root of the difference between the water level and the weir crest level. Since the proposed model does not include water levels, an approximation for fixed weirs in terms of flow is used to keep the model structure.

From simulation data, a flow value q_w^{max} is determined as the maximum inflow to the junction that does not produce any flow through the spillway. Thus, while the inflow is below this threshold, the flow over a weir $w(t)$ is zero. For inflow values higher than q_w^{max}, $w(t)$ is defined as a fraction $0 < a_w < 1$ of the difference between the inflow and q_w^{max}. Mathematically, for a junction with total inflow $z_w(t)$ (see Figure 3.5), the flow over a weir $w(t)$ can be computed as

$$w(t) = \max\{0, a_w(z_w(t) - q_w^{max})\}. \tag{3.10}$$

As mentioned earlier, the weir flow does not actually depend on flow values but on water level, so it can be observed in data obtained from a complete physically-based model simulator that flow values at the main sewer pipe can reach values greater than q_w^{max}. This fact is suitably approximated by the introduction of parameter $a_w \in (0, 1]$. The value of parameter a_w is also determined from simulation data as a the one that minimizes the model prediction error, as detailed in Section 3.4. The introduction of

this parameter is a new feature with respect to previous similar works like Ocampo-Martínez [2011], Ocampo-Martínez and Puig [2010], Ocampo-Martínez et al. [2007] and has proven to improve the predictions considerably.

Notice that for each weir in the network, the inflow $z_{w_i},\ i \in \{1, 2, \ldots, n_w\}$, equals the total inflow z_{q_j}, for some $j \in \{1, 2, \ldots, n_q\}$.

3.3.4 Overflow Equations in the Hybrid Linear Delayed Model

Overflows are defined at junctions in a way that is completely analogous to the weir flow, i.e.,

$$f(t) = \max\{0, a_f(z_f(t) - q_f^{max})\}, \tag{3.11}$$

where $f(t)$ is the overflow, $z_f(t)$ the total inflow to the junction and q_f^{max} the inflow value at which overflow starts. Again, $a_f \in (0, 1]$ is introduced to better approximate the fact that the outflow can be greater than q_f^{max}.

Although overflows could be defined in every network junction, it is enough to define them only at those prone to suffer from overflow events. This junctions can be easily determined from data generated by a physically-based model simulator. Avoiding the definition of overflow variables at those junctions where overflows are very unlikely to occur improves the model computational speed since it is strongly dependent on the amount of switching equations, especially in the optimal control case.

Notice that, as in the weir inflow definition, for each overflow junction defined in the network, the inflow $z_{f_i},\ i \in \{1, 2, \ldots, n_f\}$, equals the total inflow z_{q_j}, for some $j \in \{1, 2, \ldots, n_q\}$.

3.3.5 Flood Runoff Equations in the Hybrid Linear Delayed Model

A novel feature of the proposed model consists in keeping track of the volume that goes out of the network through overflows to let it return to the network when the overflow event has finished. A similar model based on water levels is implemented in the physically-based model simulator used for calibration, validation and control throughout this thesis. Therefore, this model is developed as a flow-based approximation of the one implemented in the physically-based model simulator.

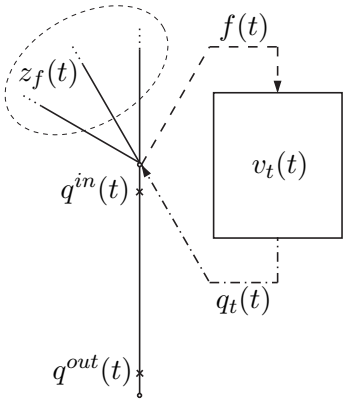

Figure 3.6: Overflow and flood runoff diagram. The proposed overflow model keeps track of the overflow volume and lets it return to the network when the overflow event has finished.

As shown in Figure 3.6, to keep track of the volume flowing out of the network through overflows, for each overflow variable $f(t)$ a volume variable $v_f(t)$ is defined, which acts like a tank that stores all the overflow volume, i.e.,

$$v_f(t) = v_f(t-1) + \Delta t\big(f(t-1) - q_f(t-1)\big), \tag{3.12}$$

where $q_f(t)$ is the emptying flow defined as

$$q_f(t) = \min\left\{\max\left\{0, b_f\left(q_f^{max} - z_f(t)\right)\right\}, \frac{v_f(t)}{\Delta t}\right\}, \tag{3.13}$$

with $b_f \in (0, 1]$. To understand the meaning of equation (3.13), notice from the definition of the overflow variable (3.11) that

$$f(t) > 0 \implies q_f(t) = 0,$$

therefore, the tank does not start emptying until the overflow event has finished. On the other hand, (3.11) also implies that

$$f(t) = 0 \implies q_f(t) = \min\left\{b_f\left(q_f^{max} - z_f(t)\right), \frac{v_f(t)}{\Delta t}\right\}.$$

This means that the tank can never provide more flow than that which would empty it in a single time step (i.e., $v_f(t)/\Delta t$). If there is enough volume available (i.e., $\frac{v_f(t)}{\Delta t} > b_f(q_f^{max} - z_f(t))$), the tank empties at a rate b_f proportional to the difference between the overflow threshold q_f^{max} and the inflow $z_f(t)$. Similar to parameters a_w and a_f of the weir and overflow model, parameter b_f is introduced for calibration purposes.

3.3.6 Collector Equations in the Hybrid Linear Delayed Model

Collectors are big sewer pipes with an in-line detention capacity of the same order as a tank. For optimal control purposes, collectors with a downstream gate controlling their outflow are better modelled as one or more tanks. This modelling allows to keep track of the volume $v_c(t)$ contained in the collector coming from upstream pipes $q(t)$ to better decide the amount available to be released through the downstream gate $g(t)$. Three different models have been developed and compared in this study: a single tank model and two multiple tank models, one only accounting for delay and the other also accounting for the water distribution along the collector. In all cases a manipulated gate $g(t)$ is assumed to be placed at the downstream end of the collector and an overflow variable $f_c(t)$ is added at its upstream end to model possible flooding. This feature will be of special interest when using the model for optimal control purposes to avoid infeasibilities in case the collector becomes overloaded. As will be described in Chapter 6, this situation is to be avoided by means of a strong penalization of variable $f_c(t)$ in the objective function of the optimal control problem.

Single Tank

The equation for the volume contained in the collector using a one-tank model is analogous to the one used previously for the volume contained in a tank,

$$v_c(t) = v_c(t-1) + \Delta t\big(q(t-1) - g(t-1) - f_c(t-1)\big), \tag{3.14}$$

where $f_c(t)$ is the overflow variable, defined in as the VT model overflows in virtual tanks,

$$f_c(t) = \frac{1}{\Delta t}\max\Big\{0\,,\, v_c(t-1) + \Delta t\big(q(t-1) - g(t-1)\big) - v_c^{max}\Big\}, \tag{3.15}$$

i.e., variable $f_c(t)$ equals the the part of the inflow $q(t)$ that does not fit in the collector, which has the physical limitation $v_c(t) \leq v_c^{max}$.

Single Tank Plus Delay

The second model consists in adding a delay to the inflow to the tank, thus making the volume available to be released through the downstream gate only some time steps

after it has entered the collector. An easy way to implement this is to represent the collector as a series of N tanks, each one adding a one time step delay to the flow, with only the last one acting as a storage element (see Figure 3.7). In the following, the sub-index c in the tank volumes will be dropped for notational simplicity. The tank equations are the same as in the tank model with correspondingly modified in- and outflows:

$$\begin{aligned} v_c^1(t) &= v_c^1(t-1) + \Delta t\big(q_{in}(t-1) - q_c^1(t-1) - f_c(t-1)\big), \\ v_c^i(t) &= v_c^i(t-1) + \Delta t\big(q_c^{i-1}(t-1) - q_c^i(t-1)\big),\ i = 2, \dots, N-1, \\ v_c^N(t) &= v_c^N(t-1) + \Delta t\big(q_c^{N-1}(t-1) - g(t-1)\big). \end{aligned} \tag{3.16}$$

To obtain the desired one time step delay effect, the communicating flows $q_i(t)$ between the tanks are defined as

$$q_c^i(t) = \frac{v_c^i(t)}{\Delta t},\ i = 1, \dots, N-1. \tag{3.17}$$

This means that each tank completely empties towards the next one every time step except the last one, which is controlled by a gate flow $g(t)$. In this case, the overflow variable is defined as

$$f_c(t) = \frac{1}{\Delta t} \max\Big\{0, \sum_{i=1}^{N} v_c^i(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\Big\}, \tag{3.18}$$

where v_c^{max} is the total collector volume.

N Tanks

The last model also consists of a series of N tanks for which equations (3.16) hold with the only difference being the definition of the interconnecting flows q_c^i. In this case each of the tanks is defined to have the same maximum capacity

$$v_N^{max} = \frac{v_c^{max}}{N},$$

where v_c^{max} is the total collector volume. The last tank is again controlled by a gate, acting as a decision variable (c.f. Figure 3.7). This can cause the last tank to become full. If this happens the second downstream tank starts filling. The same procedure applies to the other tanks on until the first one. If the first tank becomes full, any additional inflow is regarded as overflow.

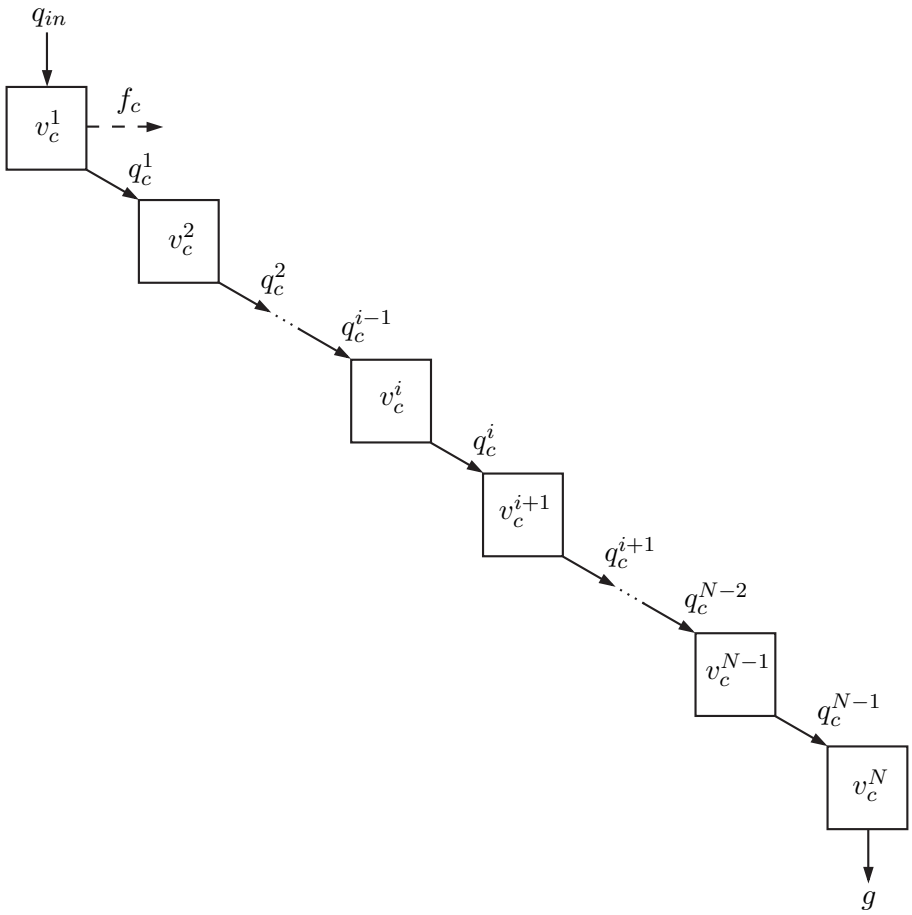

Figure 3.7: Model diagram and variables for the Single Tank Plus Delay and the N Tanks models.

The equations for the tanks and the overflow variable are the same as for the previous model. However, in this case the flows communicating the tanks are defined as follows:

$$q_c^i(t) = \min\left\{\frac{v_c^i(t)}{\Delta t}, \frac{v_N^{max} - v_c^{i+1}(t)}{\Delta t} + q_c^{i+1}(t)\right\},\ i = 1, \ldots, N-2, \tag{3.19}$$

with the modified expression for the last tank, accounting for the controlled outflow

$$q_c^{N-1}(t) = \min\left\{\frac{v_c^{N-1}(t)}{\Delta t}, \frac{v_N^{max} - v_c^N(t)}{\Delta t} + g(t)\right\}. \tag{3.20}$$

Notice that with these expressions, even when several tanks are full, there is still flow through the collector: all the full tanks provide the next one a flow equal to the downstream gate flow $g(t)$.

In this case, the overflow variable takes only into account the first tank

$$f_c(t) = \frac{1}{\Delta t} \max\Big\{0, v_c^1(t-1) + \Delta t\big(q_{in}(t-1) - q_1(t-1)\big) - v_N^{max}\Big\}. \tag{3.21}$$

3.3.7 Rainfall-Runoff Equations in the Hybrid Linear Delayed Model

Rainfall-runoff models are conceptual models that describe the relationship between rainfall intensity and runoff flow. These models were originally developed to compute the net inflow from rain catchments to rivers, but more recently they have been used to compute inflow to sewer networks.

The rainfall-runoff model used in this thesis is the one provided by the physically-based model simulator used for calibration, validation and control simulations so that both the control model and the physically-based model have the same rain inflows. The rainfall-runoff model, called *Kinematic Wave Model B* [MOUSE, 2007b], is based on modelling the rain catchments as *nonlinear reservoirs*, that is, reservoirs emptying at a rate which depends nonlinearly on the water volume contained therein, i.e.,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} V(t) &= A\, I_{\mathit{eff}}(t) - r(t), \\ r(t) &= \frac{W\, S^{1/2}}{n\, A^{5/3}}\, V(t)^{5/3}, \end{aligned}$$

where $r(t)$ is the runoff flow $[\frac{\mathrm{m}^3}{\mathrm{s}}]$ entering the network, $V(t)$ the volume $[\mathrm{m}^3]$ stored in the catchment/reservoir, $I_{\mathit{eff}}(t)$ the effective rain intensity $[\frac{\mathrm{m}}{\mathrm{s}}]$ (after subtracting losses due to several hydrological phenomena), A the catchment area $[\mathrm{m}^2]$, W the catchment width [m], S the catchment slope [dimensionless] and n the Manning coefficient $[\mathrm{s\,m}^{-1/3}]$. The second equation is obtained from an approximation of the Manning formula for large rectangular channels [Litrico and Fromion, 2009].

Notice that this model is a conceptual model and, therefore, some of its parameters, such as the catchment slope or width, do not represent the real physical properties of the catchment, which are not homogeneous. In order to provide a suitable approximation of the rainfall-runoff relation, these parameters must be calibrated from real data for each particular catchment. For further details on rainfall-runoff modelling and applications see Beven [2011].

Each network catchment is connected to a network junction where the computed inflow is attached.

3.4 Model Calibration

As mentioned above, it is assumed that a simulator based on the complete physically-based model is available for calibration purposes. As proposed in the literature (e.g. in Meirlaen et al. [2001], Solvi [2006], Vanrolleghem et al. [2005]), using a detailed model, which has previously been calibrated with real data, larger amounts of virtual data covering a wider range of situations can be generated to calibrate the control model. Through simulation of several rain events, the physically-based model should provide comprehensive information regarding flows and water levels in junctions, sewer pipes, weirs and gates. A description of the specific physically-based model simulator used for this work is detailed in Section 4.1. Nonetheless, it is important to remark that the calibration procedures presented in this section are based only in flow, water-level and rain intensity data, so that the process would be the same if many real field measurements were available.

Since the objective of this thesis is the description of a control-oriented model, its calibration, validation and use for RTC, the calibration procedures described in this section focus on the ability of the control model to aprehend the hydraulic behaviour explained by the physically-based model. Therfore, calibration procedures for the physically-based model and the rainfall-runoff model are out of the scope of this work. In fact, it is a common practice for the companies/institutions responsible of operating sewer networks to use such model simulators, which have been calibrated using real measurement data and expert knowledge. Therefore, it is reasonable to assume that, prior to the implementation of any RTC technique, a physically-based model for the hydraulics and a rainfall-runoff model for the hydrology are available and properly calibrated.

In the following, flows in sewer pipes, weirs and gates and the rainfall-runoff inflows provided by the physically-based model simulator will be denoted with hats: $\hat{q}_i^{in}(t)$, $\hat{q}_i^{out}(t)$, $\hat{w}_i(t)$, $\hat{g}_i(t)$ and $\hat{r}_i(t)$, with $t = 1, \ldots, t_s$ and t_s the duration of the simulated event. From this data, the inflows to each sewer pipe $\hat{z}_{q_i}(t)$ (and therefore to each weir

$\hat{z}_{w_i}(t)$ and gate $\hat{z}_{f_i}(t)$) can also be obtained as in equation (3.7):

$$\hat{z}^{in}_{q_i}(t) = \sum_{j=1}^{n_q} a^q_{ij} \hat{q}^{out}_j(t) + \sum_{j=1}^{n_w} \max\{0, a^w_{ij}\}\, \hat{w}_j(t) + \sum_{j=1}^{n_g} \max\{0, a^g_{ij}\}\, \hat{g}_j(t) + \sum_{j=1}^{n_c} a^r_{ij} \hat{r}_j(t).$$

The overall calibration strategy involves comparing the flows computed by the control model with the flows computed by the physically-based model simulator. This comparison is performed for different values of the parameters (by trial and error or, indirectly, by optimization methods) and the values that provide the best approximation are chosen.

3.4.1 Flow Model Parameters

Mass Balance Equations Parameters

Since all $a^\bullet_{ij}$ coefficients in equation (3.6) are such that $a^\bullet_{ij} \in \{0, 1, -1\}$ depending on the network topology, λ_i are the only parameters to be calibrated in those equations. Moreover, parameters λ_i need to be calibrated only for sewer pipes whose upstream junction has more than one outflow since otherwise $\lambda_i = 1$.

If N outflows are considered, let $i_1, \ldots, i_N$, be the indices of the outflowing sewer pipes. Denote $\hat{q}(t)$ the total outflow to the junction (the sum of all inflows), i.e.,

$$\hat{q}(t) = \sum_{k=1}^{N} \hat{q}^{in}_{i_k}(t).$$

Then, each parameter λ_{i_k}, $k = 1, \ldots, N$, is computed as

$$\lambda_{i_k} = \arg \min_{\tilde{\lambda}_{i_k} \in (0,1)} \sum_{t=1}^{t_s} \left(\tilde{\lambda}_{i_k}\, \hat{q}(t) - \hat{q}^{in}_{i_k}(t) \right)^2.$$

This problem has an explicit solution. Denoting

$$\hat{\mathbf{q}} = (\hat{q}(1), \ldots, \hat{q}(t_s))^\top$$
$$\hat{\mathbf{q}}^{in}_{i_k} = \left(\hat{q}^{in}_{i_k}(1), \ldots, \hat{q}^{in}_{i_k}(t_s)\right)^\top,\ i_k = 1, \ldots, N,$$

coefficients λ_{i_k} are computed as

$$\lambda_{i_k} = \frac{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}_{i_k}^{in}}{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}}.$$

Using this expression is easy to show that the sum of all coefficients for a junction add to one, as desired for mass balance:

$$\sum_{k=1}^{N} \lambda_{i_k} = \sum_{k=1}^{N} \frac{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}_{i_k}^{in}}{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}} = \frac{\hat{\mathbf{q}}^\top \sum_{k=1}^{N} \hat{\mathbf{q}}_{i_k}^{in}}{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}} = \frac{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}}{\hat{\mathbf{q}}^\top \hat{\mathbf{q}}} = 1.$$

Flow Equation Parameters

Delays t_i and attenuation parameters a_i are computed by minimizing the difference of the left- and right-hand sides of (3.8) when using data computed by the physically-based model simulator, thus leaving the parameters as the only free variables, i.e.,

$$(t_i, a_i) = \arg \min_{\substack{\tilde{t}_i \in \mathbb{Z}^+ \\ \tilde{a}_i \in (0,1]}} \sum_{t=1}^{t_s} \left(\hat{q}_i^{out}(t) - \tilde{a}_i \, \hat{q}_i^{in}(t - \tilde{t}_i) - (1 - \tilde{a}_i) \, \hat{q}_i^{in}(t - \tilde{t}_i - 1) \right)^2.$$

In this case, since no explicit solution is available, parameters are computed by evaluating all possible combinations of values of a_i in a partition of interval $(0, 1]$ and $t_i \in \{0, 1, \ldots, \tilde{T}\}$, where $\tilde{T}$ is a rough upper bound on the maximum network delay determined beforehand from observation of simulation data. If the results of this procedure for a given $\tilde{T}$ are such that $t_i = \tilde{T}$, for some $i \in \{1, \ldots, n_q\}$, the upper bound $\tilde{T}$ has to be increased until it is not the case anymore.

3.4.2 Weir Model Parameters

According to equation (3.10), two parameters are to be determined for the weir flow equation: the maximum inflow at the junction before water starts to flow through the spillway q_w^{max} and the weir parameter a_w. The maximum inflow q_w^{max} is defined as the inflow at the time instant when the flow over the weir starts,

$$q_w^{max} = \hat{z}_w(t_w),$$

with

$$t_w = \min\{\, t \,|\, \hat{w}(t) > 0\}.$$

The weir parameter is computed so that the maximum weir flow obtained with (3.10) using simulator data equals the maximum provided by the simulator, that is,

$$a_w = \frac{\max \hat{w}(t)}{\max\{\hat{z}_w(t) - q_w^{max}\}}.$$

3.4.3 Overflow and Flooding Runoff Model Parameters

Although the expressions for the overflow and weir flow are completely analogous, the calibration procedure for the involved parameters is slightly different because in some cases physically-base model simulators do not provide explicit overflow variables. For example, the physically-based model simulator used in this thesis (see Section 4.1 for details) simulates overflow using only water levels: by means of the Preissmann slot and extending the junction geometry above the ground level with a conic shape that simulates the flooded area. When the junction water level is above the ground level, thus filling the conic extension, overflow is taking place. However, it does not provide any quantitative measure (at least no one that can be accessed from the output files) of either the overflow or the spilled volume. Therefore, calibration must be carried out by using only the inflows and outflows of the overflowing junction. Another useful variable for overflow parameter calibration that is not used anywhere else in the model, but is always provided by simulators, is the junction water level. This variable is used de determine when the overflow event starts as follows:

$$t_f = \min\{\, t \,|\, \hat{h}(t) > h_0\},$$

where h_0 is the ground level at the junction. Thus, it is being considered that overflow starts when the water level in the junction surpasses its ground level. The maximum inflow before overflow starts is, therefore, defined as the inflow at the starting of the overflow event, i.e.,

$$q_f^{max} = \hat{z}_f(t_f).$$

Defining the following parameter-dependent functions,

$$\begin{aligned}
\hat{f}(t, a_f) &= \max\{0, a_f(\hat{z}_f(t) - q_f^{max})\}, \\
\hat{q}_t(t, b_f) &= \min\left\{\max\left\{0, b_f\left(q_f^{max} - \hat{z}_f(t)\right)\right\}, \frac{\hat{v}_t(t)}{\Delta t}\right\}, \\
\hat{v}_t(t, a_f, b_f) &= \hat{v}_t(t-1, a_f, b_f) + \Delta t\big(\hat{f}(t-1, a_f) - \hat{q}_t(t-1, b_f)\big),
\end{aligned}$$

the overflow parameters are obtained as

$$(a_f, b_f) = \arg \min_{\substack{\tilde{a}_f \in (0,1] \\ \tilde{b}_f \in (0,1]}} \sum_{t=1}^{t_s} \Big(\hat{q}_i^{in}(t) - \lambda_i \big(\hat{z}_f(t) + \hat{f}(t, \tilde{a}_f) - \hat{q}_t(t, \tilde{b}_f) \big) \Big)^2.$$

Notice that functions $\hat{f}(t, a_f)$, $\hat{q}_t(t, b_f)$ and $\hat{v}_t(t, a_f, b_f)$ are only defined here for ease of notation and to make apparent the resemblance of the minimization problem with equation (3.6). Since no explicit solution is available for this optimization problem, these parameters are computed by trying all combinations of different values of a_f and b_f in a grid within $(0, 1] \times (0, 1]$. Notice that this computation is only carried out off line.

3.4.4 Collector Model

For the collector model only one parameter needs to be calibrated: the number of tanks. Since every tank adds a one time step delay, a first estimate of the number of tanks can be obtained by comparing the inflow and outflow of the collector according to the data provided by the physically-based model. However, notice that the delay obtained by this procedure is strongly affected by the rain intensity and the downstream gate opening. Therefore, it is recommended that, to use the model for control purposes, closed loop-simulations with different number of tanks are performed for different rain scenarios to decide which is the best value. These results applied to a specific case study are provided in Chapter 6.

Regarding the maximum volumetric capacity of the collector, notice that this parameter can be obtained from the geometry of the collector. Since collector overflowing is a critical and dangerous situation and the collector models included in the HLD model are only approximate, it is recommended to use these models with the measured maximum volume decreased by a safety factor for control purposes. Again, to assess the suitability of the security factor, closed-loop simulations are recommended.

3.5 Summary

In this chapter, a novel control-oriented sewer network model has been presented: the Hybrid Linear Delayed (HLD) model. The equations for each element have been de-

scribed in detail, including three different sub-models for collectors. The HLD model has been developed as a natural extension of the Virtual Tank model, to overcome some difficulties that appear in the topological simplification and parameter calibration when applying this model to specific case studies. By modelling individual elements of the network including sewer pipes, weirs, gates, tanks and overflow points, the HLD model topology can be directly obtained from the real network one and the corresponding model parameters have direct physical meaning.

The *hybrid linear delayed* modelling framework, naturally extends the usual linear systems approach by including time delays and switching equations. These features allow to properly describe the transport delays in sewer pipes and the logics behind the weir flow and overflow phenomena. On the other hand, the underlying linear nature of the transport and mass balance equations allows for fast and accurate computation, suitable for large scale systems, as discussed in the next chapters.

After the model description, calibration procedures for the computation of all the model parameters have been developed. The direct correspondence of the HLD model variables and the physically-based model variables allows for simple procedures based on the minimization of the error between HLD model predictions and simulation data generated by a physically-based model, although real measurements could also be used, if available.

The whole modelling approach has been developed with the aim to be a systematic methodology that could be readily applied to any network including the considered elements without the need of additional management experience or expert knowledge. All the steps involved in the modelling can be automated by extracting the network topology from a physically-based model simulator description. Similarly, calibration procedures can be readily implemented by performing physically-based model simulations and solving the corresponding optimization problems.

In the following chapter, the proposed modelling and calibration methodologies will be applied to a real case study and validation will be presented.

Chapter 4

Case Study and Model Validation

This chapter is partially based on:

- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid modeling and receding horizon control of sewer networks.* Water Resources Research*, 2014d.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Receding horizon control of hybrid linear delayed systems: Application to sewer networks.* IEEE Conference on Decision and Control*, 2013a. Firenze, Italy.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. A control-oriented hybrid modelling approach for sewer networks: Barcelona case study.* IWA Conference on Instrumentation, Automation and Control*, 2013b. Narbonne, France.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Output-feedback control of sewer networks thorugh moving horizon estimation.* IEEE Conference on Decision and Control*, 2014b. Los Angeles, USA.*

4.1 Case Study Description

To calibrate and validate the model and later apply it in a control context, a specific network has been studied: the Riera Blanca network. This network is a part of the Barcelona city sewer network that spans an area of approximately 26 km^2. Full information about this network has been provided by CLABSA (Clavegueram de Barcelona S.A.), the company responsible of its management, by means of a highly detailed implementation in the sewer network physically-based model simulator MOUSE [MOUSE, 2007c], including three-dimensional coordinates of sewer pipes and junctions, cross-sectional geometries and materials of sewer pipes, tank geometries and gate charac-

teristics. The implementation also includes the definition of rainfall-runoff catchments related to real rain gauges, modelled with the rainfall-runoff model described in Section 3.3.7. As mentioned in Section 3.4, all the parameters for both the hydraulic and the rainfall-runoff model have been calibrated and provided by CLABSA.

By means of a simple program that analyses the text files containing the network information of the MOUSE implementation, lists of all the network sewer pipes (links) and junctions (nodes) can be automatically obtained. Using these lists, an automatic simplification of the network topology has been performed in order to avoid the use of redundant variables: as shown in Figures 4.1 and 4.2, only junctions involving more than two connected sewer pipes have been considered, with the exception of those that are also defined to have a rainfall-runoff inflow attached (shown in grey in the figures). As a result of this simplification process, new lists of links and nodes are obtained, which are used to define coefficients $a_{ij}^{\bullet}$ in equations (3.7) and (3.6). Also some simplifications regarding the geometry of the detention tanks have been carried out, since the presence of several bodies in the tanks with intercommunicating flows thorough gates and weirs has not yet been taken into account. After this processes, the number of each element in control model is as follows:

$$\begin{aligned} n_v &= 2 \text{ tanks}, \\ n_q &= 145 \text{ sewer pipes}, \\ n_w &= 3 \text{ weirs}, \\ n_f &= 11 \text{ overflows}, \\ n_g &= 10 \text{ gates}, \\ n_c &= 1 \text{ collector}, \\ n_r &= 68 \text{ rain inflows}. \end{aligned}$$

The sampling time has been chosen of $\Delta t = 1\,\text{min}$ with a maximum delay in sewer pipes of $T = 6\,\text{min}$.

Out of the ten controllable gates, two are used to redirect part of the flow from the upstrem part of the network toward the two detention tanks, six are used to regulate in- and outflow of the two detention tanks, one to regulate the collector outflow and one to redirect the flow to the WWTP.

As shown in Figure 4.3, through gates g_1 and g_2, part of the uncontrolled flow entering at the upstrem part of the network can be diverted to the two main network branches at its middle part. Folowing each of these branches, detention tanks with

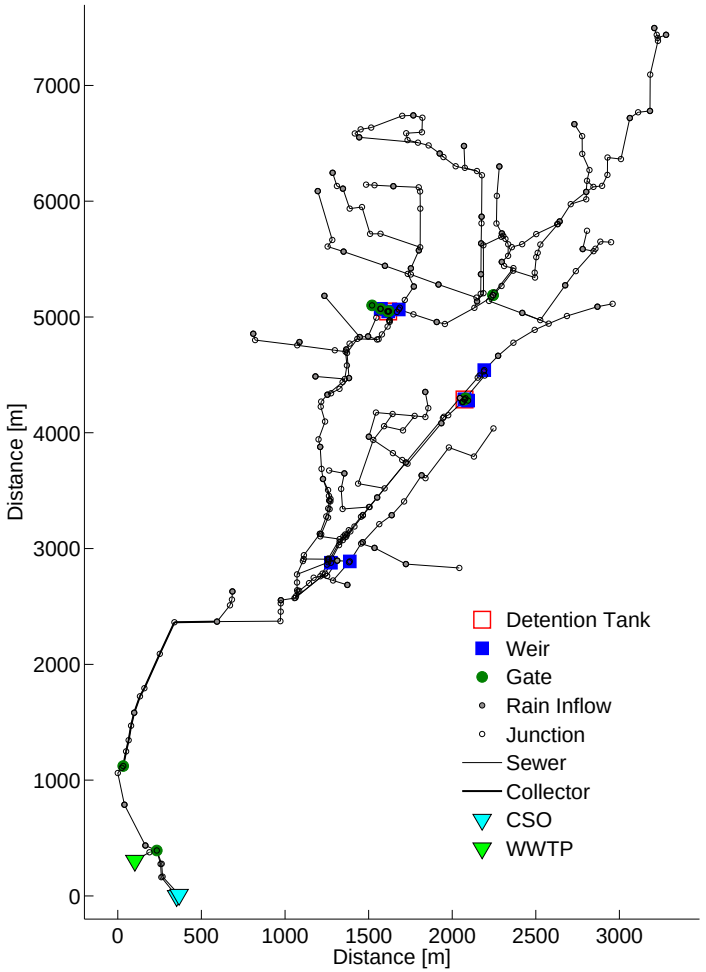

Figure 4.1: Diagram of Riera Blanca sewer network as implemented in MOUSE.

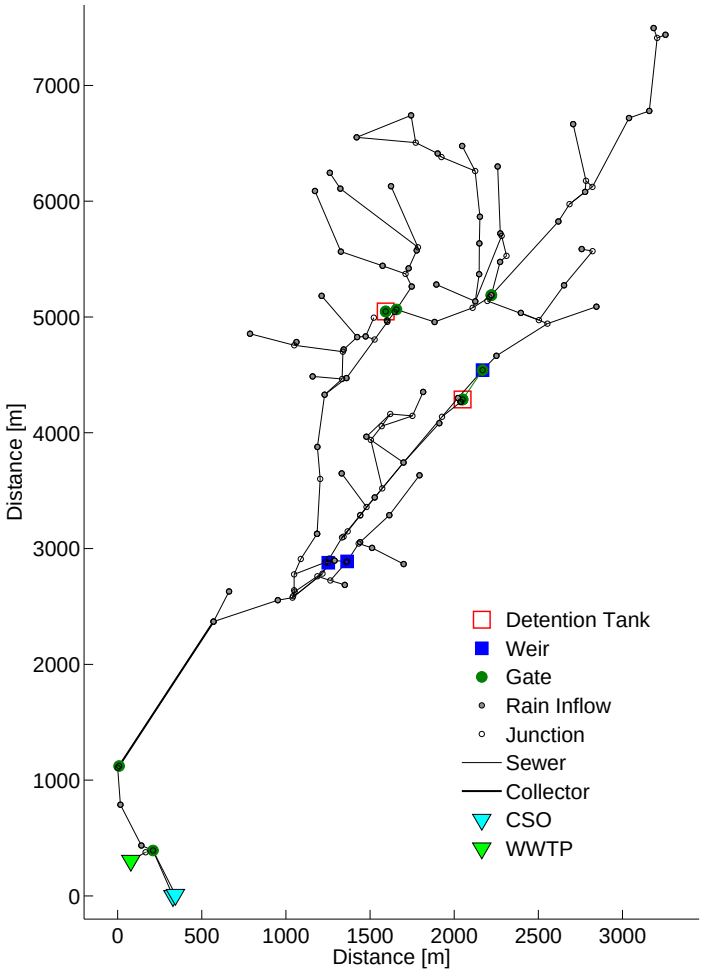

Figure 4.2: Diagram of Riera Blanca sewer network after simplification.

maximum capacities of $v_1^{max} = 102524\,\mathrm{m}^3$ and $v_2^{max} = 54918\,\mathrm{m}^3$ can be used for flow detention through inflow gates g_8 and g_9 and outflow gates g_4 and g_5 and g_3 and g_6, respectively. Figures 4.5 and 4.6 show the detail of the network topology near the detention tanks.

The Riera Blanca sewer network converges at its downstream end to a single sewer pipe, denoted as q_{139}, controlled by a gate at its downstream end, denoted g_7. Figure 4.4 shows a detail of the downstream end of the network including sewer pipe q_{139} and gate g_7. Sewer pipe q_{139} is a big sewer pipe of over 1.5 km long, with very small slope and with a total volume of $64490\,\mathrm{m}^3$. Table 4.1 shows the physical characteristics of the different sections of sewer pipe q_{139}. Due to these features, this sewer pipe has been modelled using the collector models of Section 3.3.6. Finally, as also shown in Figure 4.4, all the outflow through gate g_7 is either routed to the WWTP, which has a maximum inflow capacity of 2 m^3/s, through gate g_{10} or is released to the Mediterranean sea as CSO.

Table 4.1: Physical characteristics of the Riera Blanca collector.

	Slope	Length [m]	Width [m]	Height [m]	Area [m^2]	Volume [m^3]
Section 1	$3 \cdot 10^{-4}$	254.8	13.5	3.5	47.25	11759
Section 2	$3 \cdot 10^{-4}$	286.9	13.5	3.2	43.2	12395
Section 3	$1 \cdot 10^{-4}$	310.5	13.5	3	40.5	12576
Section 4	$1 \cdot 10^{-4}$	74.4	13.5	3	40.5	3011
Section 5	$7 \cdot 10^{-4}$	145.8	13.5	3	40.5	5906
Section 6	$4 \cdot 10^{-4}$	114.2	13.5	3	40.5	4619
Section 7	$3 \cdot 10^{-4}$	125.6	13.5	3	40.5	5087
Section 8	$2 \cdot 10^{-4}$	98.4	13.5	3	40.5	3985
Section 9	$3 \cdot 10^{-4}$	127.2	13.5	3	40.5	5152
Total		1537.7				64490

For the model calibration and for the validation results shown in this section, real pluviometer data provided by CLABSA corresponding to four real-rain events from years 2002, 2006 and 2011 has been used. The total rain inflow for each event is shown in Table 4.2 together with its duration. Figure 4.7 shows the total rain inflow to the network as computed by the rainfall-runoff model described in Section 3.3.7 (i.e., the sum of all 68 inflows as a function of time for the studied rain events).

Using these inflows as input data for the physically-based model, the four rain events have been simulated with fixed position for the network gates to generate the data sets

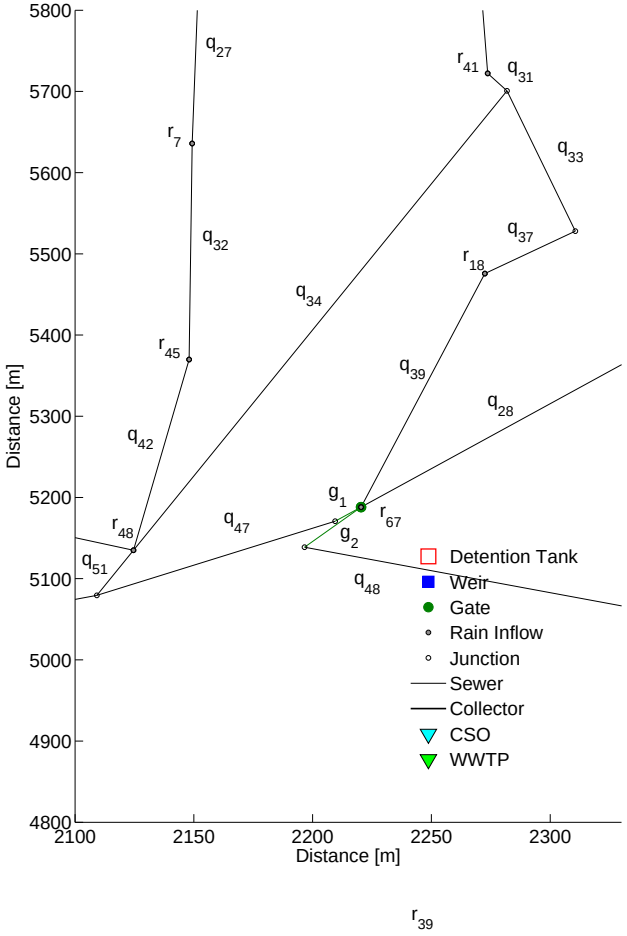

Figure 4.3: Redirection gates g_1 and g_2.

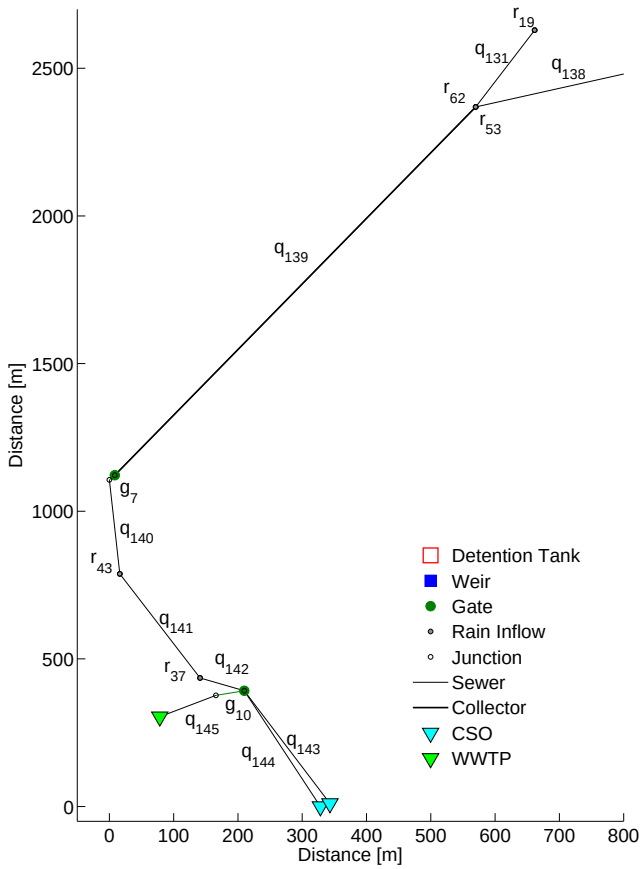

Figure 4.4: Detail of the downstream part of the Riera Blanca sewer network.

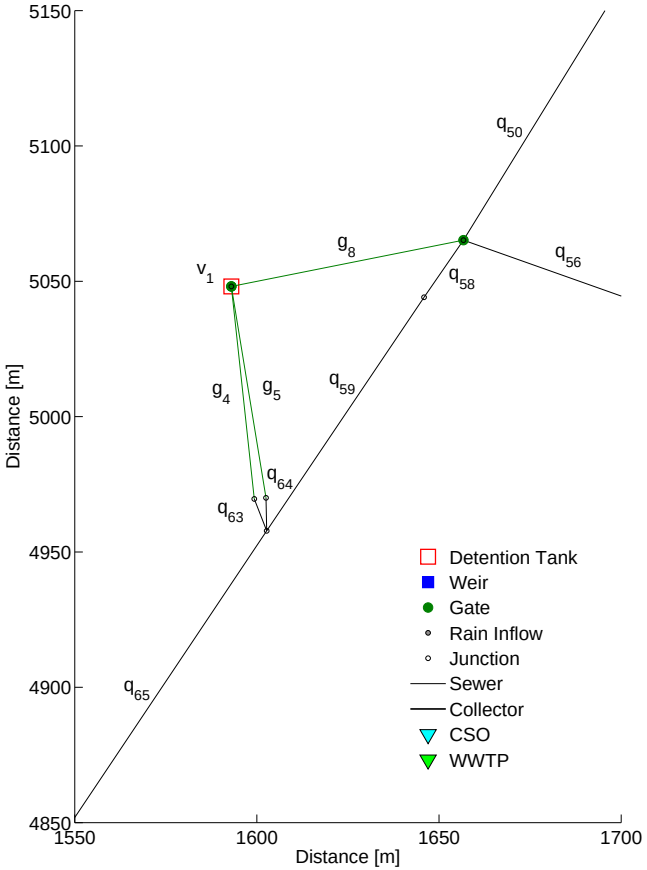

Figure 4.5: Detail of the Riera Blanca sewer network in the surroundings of detention tank v_1.

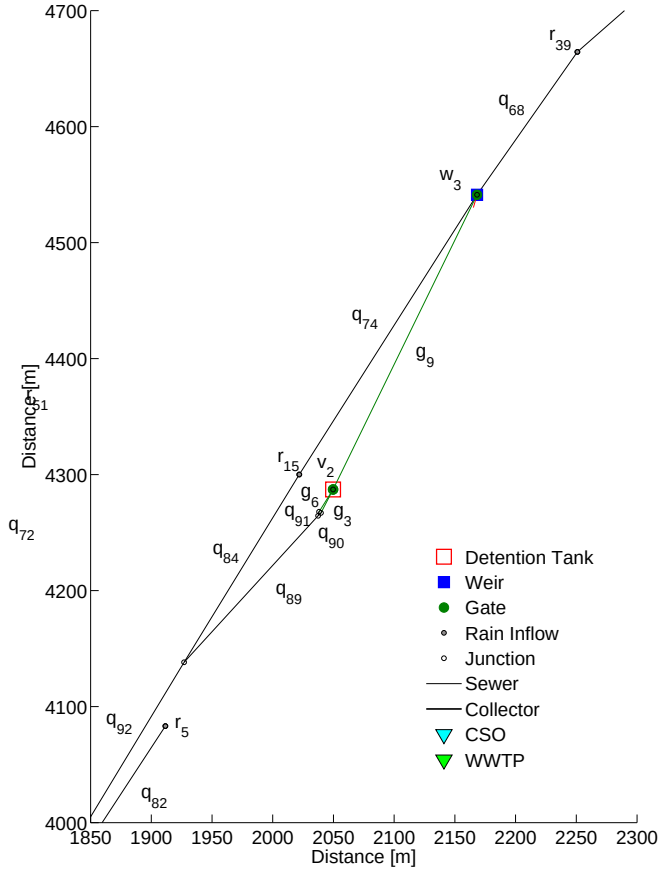

Figure 4.6: Detail of the Riera Blanca sewer network in the surroundings of detention tank v_2.

Table 4.2: Total rain inflow and duration of the studied rain events.

Episode	Total Inflow [m^3]	Duration [min]
17-09-2002	140958.34	529
09-10-2002	554135.48	606
15-08-2006	115489.84	397
30-07-2011	169875.10	339

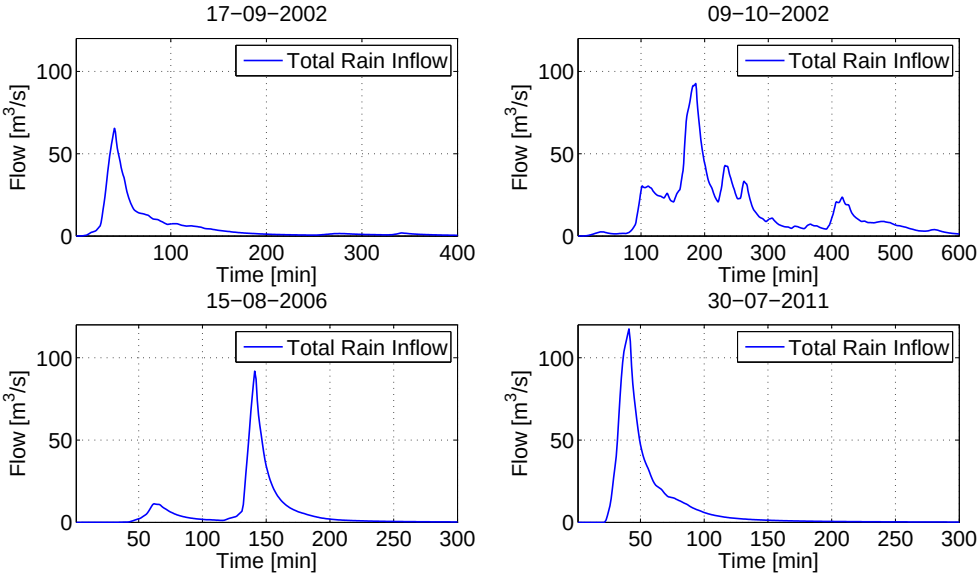

Figure 4.7: Total rain inflow to the network for the four studied rain events.

used for calibration and validation. The calibration procedure has been applied to each scenario and a final parameter set has been obtained by trial and error as a weighted average of the individual scenario parameters.

4.2 Model Validation

To validate the model, the flow values of all sewer pipes in the network as computed by the HLD model are compared with those provided by the physically-based model. For each sewer pipe, define $\bar{e}_i,\ i = 1 \dots n_q$, as the mean value of the accumulated absolute error over the simulation duration, expressed as the total number of time steps t_s, i.e.,

$$\bar{e}_i = \frac{1}{t_s} \sum_{t=1}^{t_s} \left| q_i^{in}(t) - \hat{q}_i^{in}(t) \right| \quad \left[\frac{\mathrm{m}^3}{\mathrm{s}} \right].$$

To obtain a description of the overall model accuracy, the mean and the maximum of these errors are used as indicators, expressed as:

$$E_1 = \frac{1}{n_q}\sum_{i=1}^{n_q} \bar{e}_i \ \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right], \quad E_2 = \max_i \bar{e}_i \ \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right].$$

Table 4.3 shows the error values for the different rain events. The maximum error E_2 is achieved in all cases at the big sewer pipe at the downstream end of the network (q_{139}, Figure 4.4). The nonlinear effects of open-channel flow are especially relevant for big sewer pipes with low slope. These effects are increased by the presence of a gate at the downstream end of the sewer pipe since, unless the gate is completely open, water accumulates causing changes in the flow and the total sewer pipe delay. The approximations of the inflow to the collector as computed by the control model and the physically-based model are shown in Figure 4.8.

Table 4.3: Model Error.

Episode	$E_1 \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$	$E_2 \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$	t_s
17-09-2002	0.075	0.853	529
09-10-2002	0.115	1.183	606
15-08-2006	0.108	1.321	397
30-07-2011	0.117	1.468	339

Figures 4.9 and 4.10 show respectively the flows at an overflowing node and over a weir as computed by the presented control model and by physically-based model simulator for the most intense rain event (09-10-2002).

4.3 Sensitivity Analysis

The use of simplified control-oriented models implies that some aspects of the system dynamics are omitted. To compensate this fact, parameters are included into the model to be calibrated and better approximate the system behaviour. In the present problem, the values of the parameters are highly dependent on the characteristics of the exogenous disturbance: the rainfall-runoff flow entering the network, which is determined by the rain intensity. Depending on the rain intensity, the flow velocity through the network pipes changes, resulting into variable transport delays and affecting the flow-level

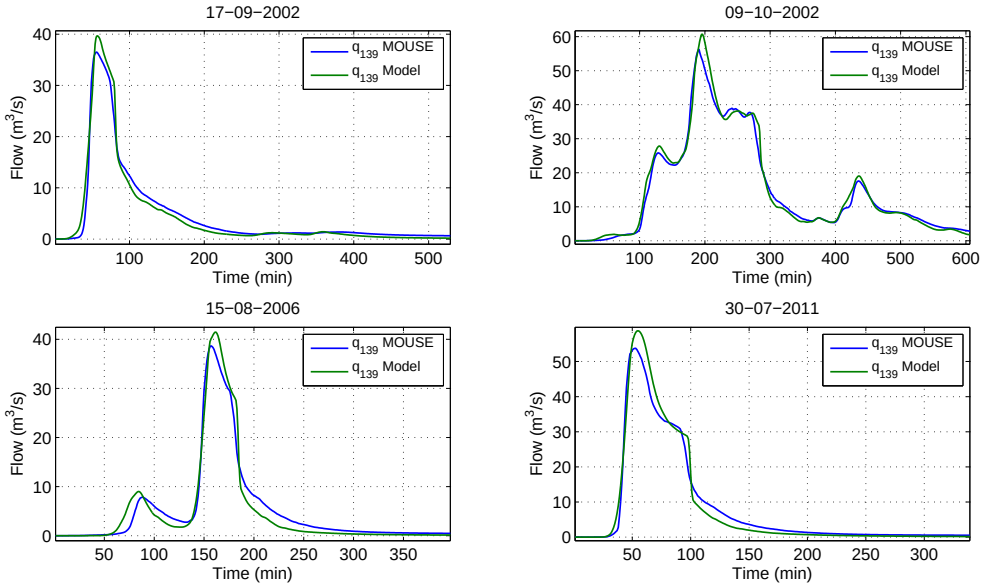

Figure 4.8: Flow at sewer pipe q_{139}, at the downstream part of the network as computed by the presented control model and by MOUSE simulator. The maximum error E_2 occurs in this sewer pipe for the four simulated rain events.

relationships. These phenomena are not explicitly taken into account by the model and should be reflected by means of the parameter values, obtained by the calibration procedures. Therefore, in order to obtain a suitable set of parameters, the studied rain events used for calibration and control should be of similar intensities. In the following, a discussion of the model performance against variations of the rain intensity is presented.

The selected rain events for calibration, validation and closed loop control simulation used in this thesis have different profiles and peak values (see Figure 4.7) but result in peak flows and velocities of the same order. Therefore, as shown in Table 4.4, they produce similar parameters.

Figures 4.11 and 4.12, show the variations on the predictions of weir flow and outflow at an overflowing node for slightly different values of some model parameters. It can be noticed that small variations in these parameters already turn into bad approximations at several time instants. Moreover, due to the network structure, these errors would be propagated and accumulated to all the following downstream sewer pipes.

To evaluate the model sensitivity against the variation of the rain intensity, simu-

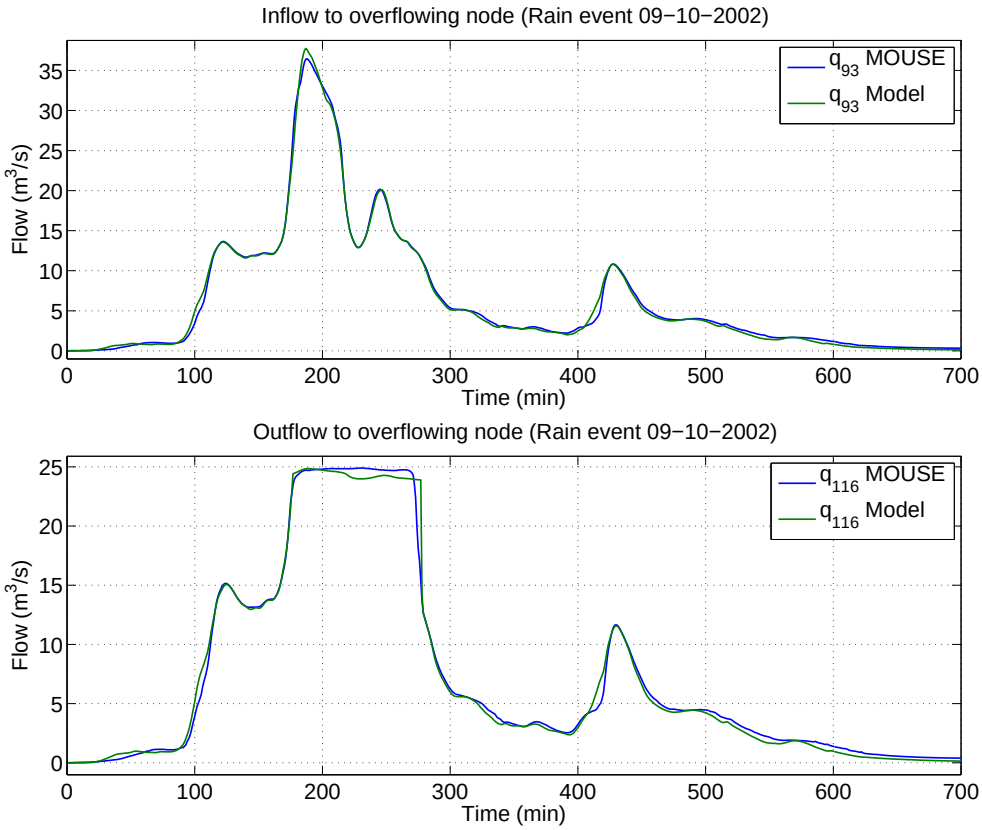

Figure 4.9: Inflow and outflow at an overflowing node as computed by the presented control model and by MOUSE simulator.

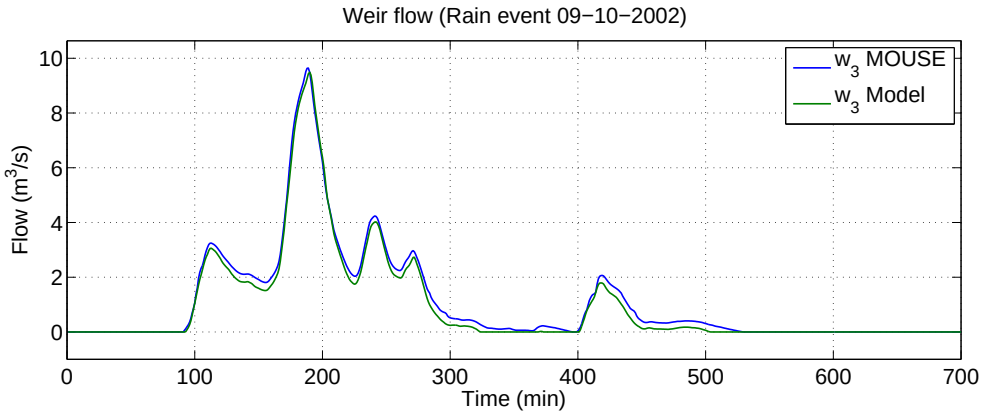

Figure 4.10: Weir flow as computed by the presented control model and by MOUSE simulator.

Table 4.4: Values of selected model parameters obtained by the calibration procedure for the different rain events.

Episode	α_{33}	a_{w_3}	$q_{w_3}^{max}$	a_{f_1}	$q_{f_1}^{max}$	b_{f_1}	a_{138}	t_{138}
15-08-2006	0.52	0.82	0.67	1.00	24.5	0.90	0.63	7
17-09-2002	0.53	0.83	0.75	0.97	24.0	0.85	0.71	8
09-10-2002	0.53	0.81	0.58	0.96	24.5	1.00	0.21	7
30-07-2014	0.52	0.84	1.19	1.00	24.5	0.97	0.87	7

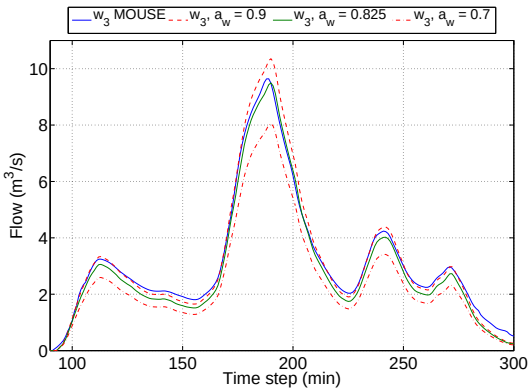

Figure 4.11: Detail of the approximation of a weir flow for different values of the parameter a_w. The solid green line shows the flow values as computed using the parameters obtained from the calibration process.

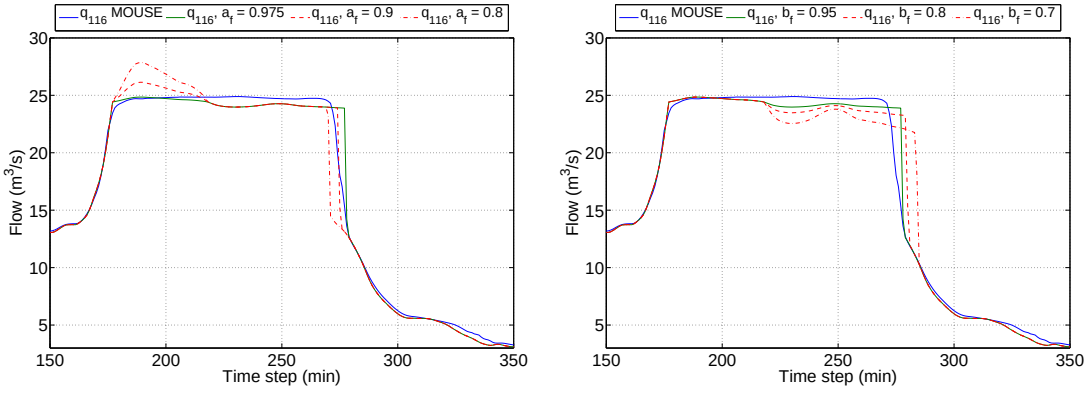

Figure 4.12: Detail of the approximation of outflow to an overflowing node for different values of parameters a_f and b_f. The solid green line shows the flow values as computed using the parameters obtained from the calibration process.

Table 4.5: Error indices for a design rain event with several increasing factors.

Episode	$E_1 \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$	$E_2 \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$	t_s
DRE 1	0.073	0.838	480
DRE 1.5	0.096	1.110	480
DRE 2	0.138	1.486	480
DRE 2.5	0.214	2.048	480

lations have been performed using a *design rain event* (DRE). Design rain events are artificially generated rain profiles used for simulation purposes. The specific procedures are developed to meet standard intensities and durations for a given climate and are out of the scope of this work. Table 4.5 shows the error indices introduced in Section 4.2 for a DRE that has been scaled with factors 1, 1.5, 2 and 2.5. As expected, due to the variation of the parameters in different rain intensity scenarios, for a fixed parameter set the model accuracy decreases as the rain intensity increases. Figure 4.13 shows the model approximations and errors with respect to physically-based model simulator data for sewer pipe q_{139}, located at the downstream end of the network, where model error from all previous elements accumulates. Looking at these approximation results with further detail, it can be noticed that the parts of the simulation events where the error takes its greatest values occur at two specific points. First, at the peak flow instants which are related to the overflow and weir flow threshold parameters q_w^{max} and q_f^{max}. These values are higher for intense rain events where flows reach higher velocities and flow values for a given water level are also higher. Secondly, at the end of the rain event, when a sudden decrease of the flow value occurs. Again, for high intensity events the higher flow velocities lead to shorter delays. In presence of sudden flow changes the delay accuracy is of capital importance in order to properly approximate the flows. In both situations, recalibration of the model parameters using rain events of suitable intensities would lead to improved-accuracy approximations.

Notice also that when the model is used for RTC it is expected that network measurements provided every few minutes are used as initial values for the model, correcting partially the approximation errors. All the simulations and plots shown in this section have been performed without any measurement update.

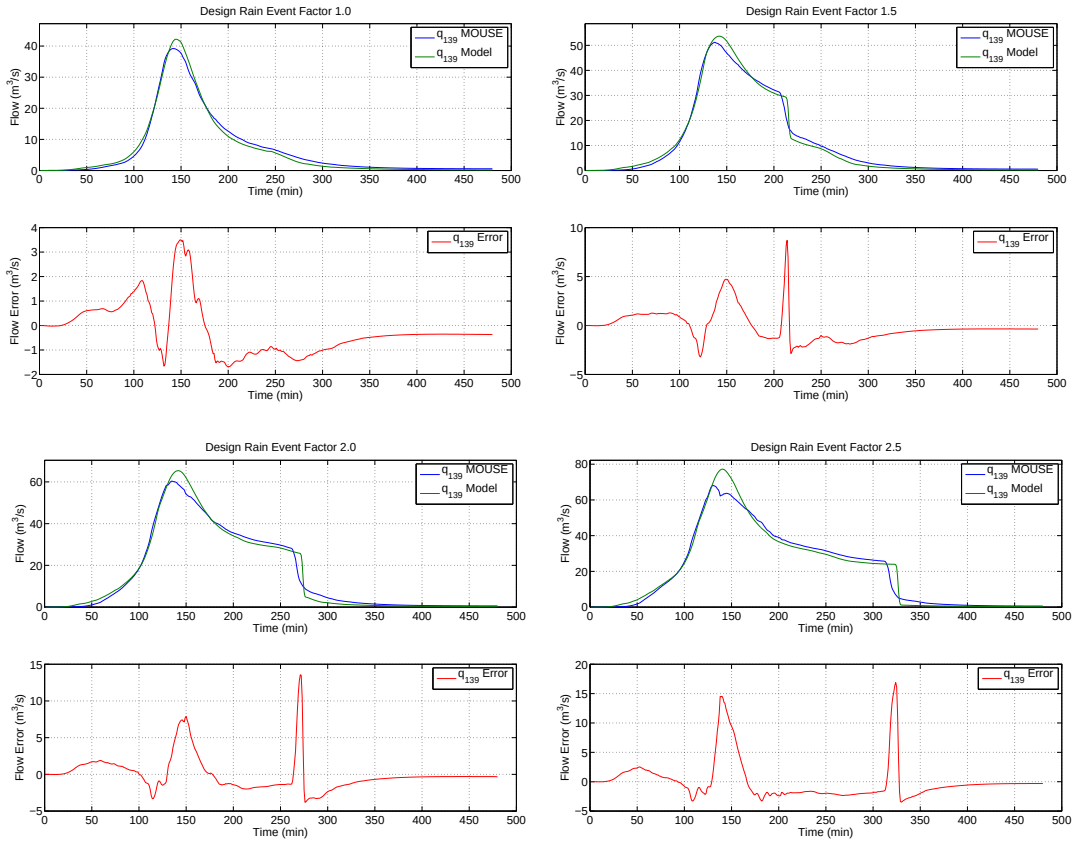

Figure 4.13: Flow at sewer pipe q_{139} (as computed by the presented control model and by MOUSE simulator) and approximation error for a design rain event with different increasing factors.

4.4 Summary

In this chapter, the case study of the Riera Blanca Sewer network has been presented. The modelling and calibration techniques described in Chapter 3 have been applied to this case study using data generated by an implementation of this network in a complete physically-based model simulator. To assess the model accuracy in predicting the network flows, a comparison between flows computed by the HLD model and a physically-based model simulator has been performed and summarized by means of two accuracy indices. The values of the accuracy indices for four real rain events show that the model provides accurate approximations of the network flows.

In addition to the validation results, sensitivity analysis has been performed regarding the variation of the network parameters as a function of the rain intensity and its effect on the model predictions. First, the variation of the flows computed by the model when modifying individual the values of some of the model parameters has been tested to better understand their impact on the flow predictions. Secondly, the model accuracy in predicting the network flows for design rain events of increasing intensity has been assessed. It has been shown that when the rain intensities increase with respect to the calibration events, weir and overflow threshold parameters show a special sensitivity in the predictions of the peak flows resulting in considerable local errors while the overall accuracy at other time instants remains acceptable. Still, it can be concluded that, for control purposes, the model is sufficiently accurate, especially taking into account that predictions for RTC cover only short time windows and benefit from constant updates from real measurements.

Part III

Control

Chapter 5

Model Reformulation and Control Problems

This chapter is partially based on:

- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid modeling and receding horizon control of sewer networks.* Water Resources Research*, 2014d.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Receding horizon control of hybrid linear delayed systems: Application to sewer networks.* IEEE Conference on Decision and Control*, 2013a. Firenze, Italy.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. A control-oriented hybrid modelling approach for sewer networks: Barcelona case study.* IWA Conference on Instrumentation, Automation and Control*, 2013b. Narbonne, France.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Output-feedback control of sewer networks thorugh moving horizon estimation.* IEEE Conference on Decision and Control*, 2014b. Los Angeles, USA.*

5.1 Introduction

In this chapter, an Optimal Control Problem (OCP) and a SEP associated with the model presented in Chapter 3 will be formulated. To formulate these problems, it is useful to work with a compact expression of the model equations. Therefore, previous to the OCP and SEP formulation, a matrix form reformulation of the model is developed. Notice, however, that some of the model equations presented in Section 3.3 include piecewise functions, in the form of maximum and minimum functions, which cannot

be directly written as linear expressions to fit a matrix formulation. To overcome this difficulty, binary variables may be introduced, whose value describe which of the branches of the maximum and minimum functions is chosen.

The systematic reformulation of piecewise linear equations into linear ones involving binary variables is a well-known technique in the control-oriented modelling field, leading to the so-called Mixed Logical Dynamic (MLD) systems: a modelling framework for systems described by "interdependent physical laws, logic rules and operating constraints" [Bemporad and Morari, 1999]. In the MLD framework, constrained linear systems are extended by allowing the presence of binary variables, which can have a direct physical meaning (that is, binary/integer states, inputs or outputs) or can arise in the description of logical conditions on the system variables. These variables describe the different *modes* of the system. Under some assumptions, MLD systems have been shown to be equivalent to other *hybrid systems* modelling formats including *linear complementarity systems*, *extended linear complementarity systems*, *piecewise affine systems*, and *max-min-plus-scaling systems* [Heemels et al., 2001].

The system description involving binary variables is usually performed in two steps: first, logical conditions are turned into linear inequalities, whose fulfilment defines the value of the binary variables, and secondly, piecewise linear functions are defined by means of linear combinations of products between binary variables and continuous variables. As an example, consider the case of the maximum function

$$F(x) = \max\{0, f(x)\}, \tag{5.1}$$

where f is a linear function, which will be extensively used in the formulation of the overflow and weir flow equations. Notice that since

$$\max\{f_1(x), f_2(x)\} = \max\{0, f_2(x) - f_1(x)\} + f_1(x),$$

only the case of the maximum between a function and zero needs to be covered.

According to the two-step procedure mentioned above, first, the logical condition behind the equation needs to be identified and reformulated. To this end let equation (5.1) be rewritten as

$$F(x) = \max\{0, f(x)\} = \begin{cases} f(x), & \text{if } f(x) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

The condition that imposes which of the two branches of the maximum function is chosen is whether $f(x) \geq 0$ or $f(x) < 0$. Therefore, let binary variable $\delta(x) \in \{0, 1\}$ be defined by means of the following equivalence relation:

$$[\delta(x) = 1] \leftrightarrow [f(x) \geq 0].$$

According to the MLD formulation, this condition is equivalent to the fulfilment of the following set of linear inequalities:

$$\begin{aligned} f(x) &\geq m\,(1 - \delta(x))\,, \\ f(x) &\leq (M + \varepsilon)\,\delta(x) - \varepsilon, \end{aligned} \tag{5.3}$$

where

$$\begin{aligned} m &:= \min_{x \in \mathcal{D}} f(x), \\ M &:= \max_{x \in \mathcal{D}} f(x), \end{aligned}$$

and ε is a small tolerance parameter, added for computational reasons, beyond which the constraint is considered to be violated. Assuming f to be a linear function and variables x to belong to a bounded domain $\mathcal{D}$, m and M can be computed or at least under- and overestimated, respectively, which is enough for the equivalence between the logic statement and the set of inequalities to hold. As usual with these so-called *Big-M-formulations*, the reformulation works better computationally the smaller the entries of M and the larger the entries of m, as long as they remain valid bounds [Williams, 1999].

Now that variable $\delta(x)$ describing the logic behind function $F(x)$ has been defined, it can be rewritten as

$$F(x) = \delta(x)\, f(x). \tag{5.4}$$

Surprisingly enough, products of binary and continuous variables can be naturally handled by the MLD formulation, since they are, again, equivalent to the fulfilment of a set of linear inequalities. According to the MLD formulation, the definition of $F(x)$ as given in (5.4) can be performed by imposing the following inequalities:

$$\begin{aligned} F(x) &\geq m\,\delta(x), \\ F(x) &\leq M\,\delta(x), \\ F(x) &\leq f(x) - m(1 - \delta(x)), \\ F(x) &\geq f(x) - M(1 - \delta(x)), \end{aligned} \tag{5.5}$$

where m and M are the same as defined above. The idea of defining these products is, in fact, one of the most powerful features of the MLD formulation.

Similarly, equation

$$G(x) = \min\{0, g(x)\}$$

is equivalent to

$$\begin{aligned} & g(x) \leq M(1 - \delta(x)), \\ & g(x) \geq (m - \varepsilon)\delta(x) + \varepsilon, \\ & G(x) \geq m\,\delta(x), \\ & G(x) \leq M\,\delta(x), \\ & G(x) \leq g(x) - m(1 - \delta(x)), \\ & G(x) \geq g(x) - M(1 - \delta(x)), \end{aligned} \tag{5.6}$$

where

$$[\delta(x) = 1] \leftrightarrow [g(x) \leq 0].$$

Now, variables F or G can be used elsewhere in the model, provided inequalities (5.3) and (5.5) or (5.6), respectively, are also imposed. Notice that to evaluate a model involving MLD reformulations not only dynamic equations must be evaluated but also the fulfilment of a set of inequalities must be solved. This kind of problems belong to the category of the so-called Constraint Satisfaction Problems (CSPs). Section 5.4 shows how to formulate and solve a CSP to use the HLD model for simulation purposes.

After applying the MLD reformulation rules to all the system logical decisions, the whole system can be written in the following general format [Bemporad and Morari, 1999]:

$$\begin{aligned} & x(t+1) = A_t\,x(t) + B_{1t}\,u(t) + B_{2t}\,\delta(t) + B_{3t}\,z(t), \\ & y(t) = C_t\,x(t) + D_{1t}\,u(t) + D_{2t}\,\delta(t) + D_{3t}\,z(t), \\ & E_{2t}\,\delta(t) + E_{3t}\,z(t) \leq E_{4t}\,x(t) + E_{5t}, \end{aligned} \tag{MLD}$$

where x is the vector of the system states, y the vector of the system outputs, u the vector of the system controlled variables and z and δ are vectors of continuous and binary auxiliary variables, respectively, defined in the reformulation of switching equations. A_t, B_{it}, C_t, D_{it} and E_{it} are, in the general case, time-dependent matrices of suitable dimensions describing the system dynamics and MLD inequalities. This format has been proven suitable for the formulation of OCPs and SEPs, and for the study of

several aspects of the systems such as equivalence with other modelling approaches, stability, observability, etc. However, due to the presence of delays, the sewer network model has been formulated using a slightly more general expression. As will be shown in the following, this expression appears to be the most natural form to express the equations and MLD inequalities of the sewer network model and turns out to be suitable for the simple formulation of the OCPs and SEPs.

A discussion of other reformulation techniques for optimal control for the maximum and minimum functions can be found in Joseph-Duran et al. [2014a]. There, aside from the MLD approach, reformulations of the VT model based on nonlinear smoothing, generalized disjunctive programming (GDP, Grossmann and Ruiz [2012]) and an ad-hoc algorithm based on relaxing the piece-wise linear constraints are described and compared taking into account computational times and implementation issues.

5.2 Sewer Network Model Equations and MLD Reformulation

In this section, the matrix reformulation of the model equations described in Section 3.3 will be developed for the case of a general network with an arbitrary number of elements of each type, except for the collector case. Since for each collector the number of tanks involved in a general network modelling would change, for simplicity of notation only one collector will be considered. For the elements involving piecewise linear functions, the MLD reformulation will be described, while for the others just the matrix expressions will be given. The notation used for vectors collecting the model variables is detailed in Table 5.1.

5.2.1 Flow model

Equations (3.6) can be easily rearrangend into matrix form by defining matrices A_Q, A_W, A_G, A_C, A_F, A_T containing coefficients $a_{ij}^{\bullet}$ and diagonal matrix Λ containing coefficients λ_i, i.e.,

$$\begin{aligned} Q_{in}(t) = \Lambda \Big(& A_Q Q_{out}(t) + A_W W(t) + A_G G(t) + \\ & A_R R(t) + A_F F(t) + A_{Q_f} Q_f(t) \Big), \end{aligned} \tag{5.7}$$

Table 5.1: Notation for vector variables of the system.

Description	Symbol	Vector	Dimensions
Flow entering sewer pipes	$q_i^{in}(t)$	$Q_{in}(t)$	n_q
Flow leaving sewer pipes	$q_i^{out}(t)$	$Q_{out}(t)$	n_q
Volume in tanks	$v_i(t)$	$V(t)$	n_v
Flow through gates	$g_i(t)$	$G(t)$	n_g
Flow over weirs	$w_i(t)$	$W(t)$	n_w
Overflows	$f_i(t)$	$F(t)$	n_f
Overflow volume	$v_f^i(t)$	$V_f(t)$	n_f
Flooding runoff flow	$q_f^i(t)$	$Q_f(t)$	n_f
Collector volume	$v_c^j(t)$	$V_c(t)$	N
Collector flow	$q_c^j(t)$	$Q_c(t)$	$N-1$
Collector overflow	$f_c(t)$	$f_c(t)$	1
Rainfall-runoff inflow	$r_i(t)$	$R(t)$	n_r

Similarly, equations (3.8) are rewritten as

$$Q_{out}(t) = A_0 Q_{in}(t) + \sum_{i=1}^{T} A_i Q_{in}(t-i), \tag{5.8}$$

where $A_i,\ i = 1 \dots T$, are diagonal matrices containing the a_i or $1 - a_i$ coefficients for each sewer pipe, and where

$$T = \max_{i=1,\dots,n_q} t_i + 1$$

In order to reduce the number of variables in the model, equations (5.7) and (5.8) can be combined. By substituting Q_{out} in (5.7) for its expression in (5.8) and solving for Q_{in}, the following expression is obtained:

$$\boxed{Q_{in}(t) = \sum_{i=1}^{T} \tilde{A}_i Q_{in}(t-i) + \tilde{A}_W W(t) + \tilde{A}_G G(t) + \tilde{A}_R R(t) + \tilde{A}_F F(t) + \tilde{A}_{Q_f} Q_f(t),} \tag{5.9}$$

with

$$\begin{aligned}
\tilde{A}_i &= (I - \Lambda\, A_Q\, A_0)^{-1} \Lambda\, A_Q\, A_i,\ i = 1 \dots T,\\
\tilde{A}_W &= (I - \Lambda\, A_Q\, A_0)^{-1} \Lambda\, A_W,\\
\tilde{A}_G &= (I - \Lambda\, A_Q\, A_0)^{-1} \Lambda\, A_G,\\
\tilde{A}_R &= (I - \Lambda\, A_Q\, A_0)^{-1} \Lambda\, A_R,\\
\tilde{A}_F &= (I - \Lambda\, A_Q\, A_0)^{-1} \Lambda\, A_F,\\
\tilde{A}_{Q_f} &= (I - \Lambda\, A_Q A_0)^{-1} \Lambda\, A_{Q_f}.
\end{aligned}$$

Since most of the variables in a sewer network model correspond to flows through sewer pipes, this reduction allows to obtain a considerably smaller model so that the control and estimation problems based on it may be solved faster. Notice also that, although equation (5.9) may suggest that the inflow in a sewer pipe depends on its past values this is not the case. After this simplification, the inflow to a sewer pipe depends on the inflow to its upstream sewer pipes (see equations (3.6) and (3.8)) not on the sewer pipe itself. This is reflected by the fact that the elements in the diagonal of matrices $\tilde{A}_i,\ i = 1, \dots, T$, are null.

5.2.2 Tank model

The tank model follows the discrete-time volume equation

$$v(t) = v(t-1) + \Delta t\big(g_{in}(t-1) - g_{out}(t-1)\big),$$

where g_{in} and g_{out} are the total net flow into the tank (inflow) and out of the tank (outflow). These flows are controlled flows through gates g_i. Hence, the matrix expression for this part of the model becomes

$$V(t) = V(t-1) + \Delta t B_{in}\, G(t-1) - \Delta t B_{out}\, G(t-1),$$

where B_{in} and B_{out} are 0-1 matrices selecting the suitable gate variables. Finally, defining $B_G = B_{in} - B_{out}$, yields

$$\boxed{V(t) = V(t-1) + \Delta t B_G\, G(t-1).} \tag{5.10}$$

5.2.3 Weir model

The weir flow equation (3.10) involves the maximum function. Therefore, the MLD reformulation to obtain an equivalent set of linear equalities and inequalities will be applied. According to the procedure described in Section 5.1, equation (3.10) can be rewritten as

$$w(t) = \begin{cases} a_w\,(z_w(t) - q_w^{max}) & , \text{ if } a_w\,(z_w(t) - q_w^{max}) \geq 0 \\ 0 & , \text{ otherwise,} \end{cases}$$

or, equivalently,

$$w(t) = a_w\,\delta_w(t)\,(z_w(t) - q_w^{max}), \tag{5.11}$$

where

$$\delta_w(t) = \begin{cases} 1, & \text{if } a_w\,(z_w(t) - q_w^{max}) \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{5.12}$$

Expressions (5.11) and (5.12) are analogous to (5.4) and (5.2) in Section 5.1, with $f(x) = a_w\,\delta_w(t)\,(z_w(t) - q_w^{max})$. Therefore, they are equivalent to the following set of inequalities:

$$\begin{aligned} a_w\,(z_w(t) - q_w^{max}) &\geq m_w(1 - \delta_w(t)), \\ a_w\,(z_w(t) - q_w^{max}) &\leq M_w\delta_w(t) + \varepsilon(\delta_w(t) - 1), \\ w(t) &\leq M_w\delta_w(t), \\ w(t) &\geq m_w\delta_w(t), \\ w(t) &\leq a_w\,(z_w(t) - q_w^{max}) - m_w(1 - \delta_w(t)), \\ w(t) &\geq a_w\,(z_w(t) - q_w^{max}) - M_w(1 - \delta_w(t)), \end{aligned}$$

as in (5.3) and (5.5), and where

$$\begin{aligned} m_w &= \min a_w\,(z_w(t) - q_w^{max}), \\ M_w &= \max a_w\,(z_w(t) - q_w^{max}). \end{aligned}$$

In matrix form, the previous inequalities can be written as:

$$\begin{pmatrix} -a_w \\ a_w \\ 0 \\ 0 \\ -a_w \\ a_w \end{pmatrix} z_w(t) + \begin{pmatrix} -m_w \\ -M_w - \varepsilon \\ -M_w \\ m_w \\ -m_w \\ M_w \end{pmatrix} \delta_w(t) + \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} w(t) + \begin{pmatrix} m_w + a_w\,q_w^{max} \\ \varepsilon - a_w\,q_w^{max} \\ 0 \\ 0 \\ m_w + a_w\,q_w^{max} \\ -M_w - a_w\,q_w^{max} \end{pmatrix} \leq 0.$$

To come up with the matrix equations for all the n_w weirs in the network, one instance of the above-referred set of inequalities must be written for each weir, with its specific set of parameters. Therefore, define

$$p_j^w = \begin{pmatrix} -a_{w_j} \\ a_{w_j} \\ 0 \\ 0 \\ -a_{w_j} \\ a_{w_j} \end{pmatrix}, \quad q_j^w = \begin{pmatrix} -m_{w_j} \\ -M_{w_j} - \varepsilon \\ -M_{w_j} \\ m_{w_j} \\ -m_{w_j} \\ M_{w_j} \end{pmatrix}, \quad r^w = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad s_j^w = \begin{pmatrix} m_{w_j} + a_{w_j}\, q_{w_j}^{max} \\ \varepsilon - a_{w_j}\, q_{w_j}^{max} \\ 0 \\ 0 \\ m_{w_j} + a_{w_j}\, q_{w_j}^{max} \\ -M_{w_j} - a_{w_j}\, q_{w_j}^{max} \end{pmatrix},$$

where the subindex $j = 1, \ldots, n_w$, indicates each one of the network weirs. By arranging these matrices in a block-diagonal structure, all the inequalities can be written together as

$$\begin{aligned} &\underbrace{\begin{pmatrix} p_1^w & & & \\ & p_2^w & & \\ & & \ddots & \\ & & & p_{nw}^w \end{pmatrix}}_{E_{Z_W}} Z_W(t) + \underbrace{\begin{pmatrix} q_1^w & & & \\ & q_2^w & & \\ & & \ddots & \\ & & & q_{nw}^w \end{pmatrix}}_{E_{\Delta_W}} \Delta_W(t) + \\ &\underbrace{\begin{pmatrix} r^w & & & \\ & r^w & & \\ & & \ddots & \\ & & & r^w \end{pmatrix}}_{E_W} W(t) + \underbrace{\begin{pmatrix} s_1^w \\ s_2^w \\ \vdots \\ s_{n_w}^w \end{pmatrix}}_{E_w} \leq 0, \end{aligned} \tag{5.13}$$

where

$$\begin{aligned} W(t) &= \Big(w_1(t), \ldots, w_{n_w}(t)\Big)^\top, \\ Z_W(t) &= \Big(z_{w_1}(t), \ldots, z_{w_{n_w}}(t)\Big)^\top, \\ \Delta_W(t) &= \Big(\delta_{w_1}(t), \ldots, \delta_{w_{n_w}}(t)\Big)^\top. \end{aligned}$$

Now, the set of inequalities defining the weir flow variables W and Δ_W has the form

$$\boxed{E_{Z_W} Z_W(t) + E_{\Delta_W} \Delta_W(t) + E_W W(t) + E_w \leq 0.} \tag{5.14}$$

5.2.4 Overflow model

The overflow variables are defined in the same way as the flow over weirs, i.e.,

$$f(t) = \max\{0, a_f(z_f(t) - q_f^{max})\} = a_f\, \delta_f(t)\, (z_f(t) - q_f^{max}), \tag{5.15}$$

with

$$\delta_f(t) = \begin{cases} 1, & \text{if } a_f(z_f(t) - q_f^{max}) \geq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{5.16}$$

and with $z_f(t)$ being the flow entering the junction where the overflow is considered to potentially occur.

Since the overflow model is analogous to the weir one, the procedure is not repeated again here. Defining matrices E_{Z_F}, E_{Δ_F}, E_F and E_f analogous to E_{Z_W}, E_{Δ_W}, E_W and E_{C_W}, the overflow variables are defined by means of the following matrix inequalities:

$$\boxed{E_{Z_F}\, Z_F(t) + E_{\Delta_F}\, \Delta_F(t) + E_F F(t) + E_f \leq 0,} \tag{5.17}$$

where,

$$\begin{aligned} F(t) &= \Big(f_1(t), \ldots, f_{n_f}(t)\Big)^\top, \\ Z_F(t) &= \Big(z_{f_1}(t), \ldots, z_{f_{n_f}}(t)\Big)^\top, \\ \Delta_F(t) &= \Big(\delta_{f_1}(t), \ldots, \delta_{f_{n_f}}(t)\Big)^\top. \end{aligned}$$

5.2.5 Flood runoff model

For the flood runoff flow model two sub-models must to be reformulated. First, the volume for the fictional tank collecting all the flooding volume and secondly the flow runoff returning to the network after the overflow event. The former, does not involve any piecewise linear expression; therefore, the volume equation for each overflowing junction (3.12) just needs writing in matrix form as

$$\boxed{V_f(t) = V_f(t-1) + \Delta t\big(F(t-1) - Q_f(t-1)\big).} \tag{5.18}$$

On the other hand, the flood runoff equation (3.13),

$$q_f(t) = \min\left\{\max\left\{0, b_f\left(q_f^{max} - z_f(t)\right)\right\}, \frac{v_f(t)}{\Delta t}\right\}, \tag{5.19}$$

involving a maximum function within a minimum one may suggest that several binary variables will be needed to reformulate it by means of the MLD precedures. However, making use of the already defined variables $f(t)$, only one such variable will be needed for each flood runoff. First, recall the definition of $f(t)$ as

$$f(t) = \max\{0, a_f(z_f(t) - q_f^{max})\} = a_f\, \max\{0, z_f(t) - q_f^{max}\}.$$

Making use of the identity $\max\{0, a-b\} - \max\{0, b-a\} = a - b$, yields

$$\max\{0, q_f^{max} - z_f(t)\} - \max\left\{0, z_f(t) - q_f^{max}\right\} = q_f^{max} - z_f(t),$$

and the maximum function in (5.19) can be replaced by

$$\max\{0, q_f^{max} - z_f(t)\} = q_f^{max} - z_f(t) + \frac{f(t)}{a_f}.$$

In a last step, another binary variable can be avoided by forcing one of the two arguments of minimum function to be zero, i.e.,

$$\begin{aligned} q_f(t) &= \min\left\{b_f\left(q_f^{max} - z_f(t) + \frac{f(t)}{a_f}\right), \frac{v_f(t)}{\Delta t}\right\} \\ &= \min\left\{b_f\left(q_f^{max} - z_f(t) + \frac{f(t)}{a_f}\right) - \frac{v_f(t)}{\Delta t}, 0\right\} + \frac{v_f(t)}{\Delta t}. \end{aligned}$$

Finally,

$$q_f(t) = q_{aux}(t) + \frac{v_f(t)}{\Delta t},$$

with

$$q_{aux}(t) = \delta_q(t)\, f_{aux}(t),$$

and

$$f_{aux}(t) = b_f\left(q_f^{max} - z_f(t) + \frac{f(t)}{a_f}\right) - \frac{v_f(t)}{\Delta t},$$

$$\delta_q(t) = \begin{cases} 1 & , \text{ if } f_{aux}(t) \leq 0, \\ 0 & , \text{ otherwise.} \end{cases}$$

Now, the MLD formulation of the previous expressions becomes

$$\begin{aligned} & f_{aux}(t) \leq M_{q_f}(1 - \delta_q(t)), \\ & f_{aux}(t) \geq m_{q_f}\delta_q(t) + \varepsilon\,(1 - \delta_q(t)), \\ & q_{aux}(t) \leq M_{q_f}\delta_q(t), \\ & q_{aux}(t) \geq m_{q_f}\delta_q(t), \\ & q_{aux}(t) \leq f_{aux}(t) - m_{q_f}(1 - \delta_q(t)), \\ & q_{aux}(t) \geq f_{aux}(t) - M_{q_f}(1 - \delta_q(t)), \end{aligned}$$

where the set of inequalities for the minimum function reformulation (5.6) has been used with

$$m_{q_f} = \min b_f \left(q_f^{max} - z_f(t) + \frac{f(t)}{a_f} \right) - \frac{v_f(t)}{\Delta t} = -b_f \, z_f^{max} - \frac{v_f^{max}}{\Delta t},$$
$$M_{q_f} = \max b_f \left(q_f^{max} - z_f(t) + \frac{f(t)}{a_f} \right) - \frac{v_f(t)}{\Delta t} = b_f \, q_f^{max} + \frac{b_f}{a_f} f^{max}.$$

Expanding the terms in f_{aux} and regrouping yields

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} q_{aux}(t) + \begin{pmatrix} -b_f \\ b_f \\ 0 \\ 0 \\ b_f \\ -b_f \end{pmatrix} z_f(t) + \begin{pmatrix} b_f/a_f \\ -b_f/a_f \\ 0 \\ 0 \\ -b_f/a_f \\ b_f/a_f \end{pmatrix} f(t) +$$
$$\frac{1}{\Delta t} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} v_f(t) + \begin{pmatrix} M_{q_f} \\ m_{q_f} - \varepsilon \\ -M_{q_f} \\ m_{q_f} \\ -m_{q_f} \\ M_{q_f} \end{pmatrix} \delta_q(t) + \begin{pmatrix} b_f q_f^{max} - M_{q_f} \\ -b_f q_f^{max} + \varepsilon \\ 0 \\ 0 \\ -b_f q_f^{max} + m_{q_f} \\ b_f q_f^{max} - M_{q_f} \end{pmatrix} \leq 0.$$

Defining vectors collecting the variables corresponding to all the overflow points in the network

$$V_f(t) = \left(v_f^1(t), \ldots, v_f^{n_f}(t) \right)^\top,$$
$$Q_f(t) = \left(q_f^1(t), \ldots, q_f^{n_f}(t) \right)^\top,$$
$$\Delta_f(t) = \left(\delta_{q_f}^1(t), \ldots, \delta_{q_f}^{n_f}(t) \right)^\top,$$
$$Q_{aux}(t) = \left(q_{aux}^1(t), \ldots, q_{aux}^{n_f}(t) \right)^\top,$$

and repeating the set of inequalities in a block-diagonal way analogous to (5.13) with the corresponding parameters a_{f_j}, b_{f_j}, $q_{f_j}^{max}$, $m_{q_f}^j$ and $M_{q_f}^j$, the matrix expression for the set of inequalities defining the flood runoff variables $Q_f(t)$ for the entire network is obtained as

$$E_{Q_f} Q_{aux}(t) + E_{Z_T} Z_F(t) + E_{F_f} F(t) + E_{V_f} V_f(t) + E_{\Delta_f} \Delta_f + E_{C_T} \leq 0.$$

Finally, in order to reduce the number of variables to be used in the problem, variables $Q_{aux}(t)$ are substituted with their expression in terms of $Q_f(t)$ and $V_f(t)$,

$$Q_{aux}(t) = Q_f(t) - \frac{1}{\Delta t} V_f(t),$$

to get

$$\boxed{E_{Q_f} Q_f(t) + E_{Z_T} Z_F(t) + E_{F_f} F(t) + \left(E_{V_f} - \frac{1}{\Delta t} E_{Q_f} \right) V_f(t) + E_{\Delta_f} \Delta_f + E_{q_f} \leq 0.} \tag{5.20}$$

5.2.6 Collector Model

As mentioned before, only one collector will be considered in the MLD reformulation. This does not imply a loss of generality, since by defining the necessary variables and adding copies of the equations described below in a block-diagonal manner (as in (5.13)), any number of collectors can be described. However, considering one single collector simplifies the matrix notation considerably, avoiding the need to define many matrices and vectors including extra indices and subindices.

Before developing the MLD formulation for the specific equations of each of the three proposed submodels, notice that in all cases the tank volume equations (3.14) and (3.16) can be written in matrix form as follows:

$$\boxed{V_c(t) = V_c(t-1) + K_{Q_c} Q_c(t-1) + K_Q Q_{out}(t-1) + K_G G(t-1) + K_{F_c} F_c(t-1),} \tag{5.21}$$

where

$$Q_c(t) = \left(q_c^1(t), \ldots, q_c^{N-1}(t) \right)^\top,$$
$$V_c(t) = \left(v_c^1(t), \ldots, v_c^N(t) \right)^\top,$$

and

$$K_{Q_c} = \Delta t \begin{pmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{pmatrix}, \; K_{in} = \Delta t \begin{pmatrix} 0 & \ldots & 1 & \ldots & 0 \\ 0 & \ldots & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{pmatrix},$$

$$K_g = \Delta t \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & -1 & \dots & 0 \end{pmatrix}, \; K_f = \Delta t \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Here, K_{in} and K_G have only one nonzero element in the position corresponding to the inflow sewer pipe and outflow gate, respectively.

In the case of one Single Tank model, $N = 1$, and therefore neither variable Q_c nor matrix K_Q are defined. In this case matrices K_{in} and K_G become row vectors with length n_q and n_c respectively with a single nonzero entry each one, indicating the positions of the index of the inflowing sewer pipe and outflowing gate, respectively, and $K_{f_c} = -\Delta t$.

Since the volume equations can be written in the same way for the three sub-models, in the following, only the overflow equations and, when needed, the communicating flows among tanks will be reformulated.

Single Tank

The single tank overflow equation (3.15) can be rewritten as

$$f_c(t) = \delta_{f_c}(t) \frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right),$$

with

$$\delta_{f_c}(t) = \begin{cases} 1, & \text{if } \frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right) \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

In this case inequalities (5.3) and (5.5) become

$$\begin{aligned}
&\frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right) \geq m_c \left(1 - \delta_{f_c}(t) \right), \\
&\frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right) \leq M_c \, \delta_{f_c}(t) + \varepsilon \left(1 - \delta_c(t) \right), \\
&f_c(t) \leq M_c \, \delta_{f_c}(t), \\
&f_c(t) \geq m_c \, \delta_{f_c}(t), \\
&f_c(t) \leq \frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right) - m_c \left(1 - \delta_{f_c}(t) \right), \\
&f_c(t) \geq \frac{1}{\Delta t} \left(v_c(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_c^{max} \right) - M_c \left(1 - \delta_{f_c}(t) \right),
\end{aligned}$$

where

$$m_c = \min \frac{1}{\Delta t}\big(v_c(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\big) = -g^{max} - \frac{v_c^{max}}{\Delta t},$$
$$M_c = \max \frac{1}{\Delta t}\big(v_c(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\big) = q_{in}^{max}.$$

In matrix form:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} f_c(t) + \frac{1}{\Delta t}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} v_c(t-1) + \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} q_{in}(t-1) +$$
$$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} g(t-1) + \begin{pmatrix} -m_c \\ -M_c - \varepsilon \\ -M_c \\ m_c \\ -m_c \\ M_c \end{pmatrix} \delta_{f_c}(t) + \begin{pmatrix} m_c + \frac{v_c^{max}}{\Delta t} \\ \varepsilon - \frac{v_c^{max}}{\Delta t} \\ 0 \\ 0 \\ m_c + \frac{v_c^{max}}{\Delta t} \\ -M_c - \frac{v_c^{max}}{\Delta t} \end{pmatrix} \leq 0.$$

The previous set of inequalities can be written in matrix form in terms of the flow and gate vectors Q_{in} and G respectively, as:

$$\boxed{E_{f_c}\, f_c(t) + E_{v_c}\, v_c(t-1) + E_{Q_{in}}\, Q_{in}(t-1) + E_{G_c}\, G(t-1) + E_{\delta_{f_c}}\, \delta_{f_c}(t) + E_c \leq 0,}$$

with $E_{Q_{in}}$ and E_{G_c} appropriately taking into account the specific flow and gate variables descibing the in- and outflow of the collector.

Single Tank Plus Delay

Regarding this formulation, in addition to the overflow variable (3.18), MLD reformulation for the piecewise linear equations (3.17) defining the interconnection flows has to be performed.

Collector Overflow in the Single Tank Plus Delay Model

In the case of the Single Tank Plus Delay model, the overflow equation (3.18) can be rewritten as

$$f_c(t) = \delta_{f_c}(t)\frac{1}{\Delta t}\left(\sum_{i=1}^{N} v_c^i(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\right),$$

with

$$\delta_{f_c}(t) = \begin{cases} 1, & \text{if } \frac{1}{\Delta t}\left(\sum_{i=1}^{N} v_c^i(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\right) \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

By comparing this expression with the Single Tank model one, it can be noticed that the only difference is in the volume terms. Therefore, the MLD reformulation is analogous to the Single Tank model one with a modified volume term, i.e.,

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} f_c(t) + \frac{1}{\Delta t} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \sum_{i=1}^{N} v_c^i(t-1) + \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} q_{in}(t-1) +$$

$$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} g(t-1) + \begin{pmatrix} -m_c \\ -M_c - \varepsilon \\ -M_c \\ m_c \\ -m_c \\ M_c \end{pmatrix} \delta_{f_c}(t) + \begin{pmatrix} m_c + \frac{v_c^{max}}{\Delta t} \\ \varepsilon - \frac{v_c^{max}}{\Delta t} \\ 0 \\ 0 \\ m_c + \frac{v_c^{max}}{\Delta t} \\ -M_c - \frac{v_c^{max}}{\Delta t} \end{pmatrix} \leq 0,$$

with

$$m_c = \min \frac{1}{\Delta t}\left(\sum_{i=1}^{N} v_c^i(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\right) = -g^{max} - \frac{v_c^{max}}{\Delta t},$$

$$M_c = \max \frac{1}{\Delta t}\left(\sum_{i=1}^{N} v_c^i(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_c^{max}\right) = q_{in}^{max}.$$

As in the Single Tank case, the previous set of inequalities can be written in matrix form in terms of the flow and gate vectors Q_{in} and G respectively, as:

$$\boxed{E_{f_c}\, f_c(t) + E_{V_c}\, V_c(t-1) + E_{Q_{in}}\, Q_{in}(t-1) + E_{G_c}\, G(t-1) + E_{\delta_{f_c}}\, \delta_{f_c}(t) + E_c \leq 0,}$$

with $E_{Q_{in}}$ and E_{G_c} appropriately taking into account the specific flow and gate variables descibing the in- and outflow of the collector.

Interconnection Flows in the Single Tank Plus Delay Model

The intercommunicating flows among tanks in the Single Tank plus Delay model follow the simple linear equation

$$q_c^i(t) = \frac{v_c^i(t)}{\Delta t},$$

which can be readily rewritten in matrix form as

$$Q_c(t) = K_{V_c} V_c(t),$$

with

$$K_{V_c} = \frac{1}{\Delta t} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & 0 \end{pmatrix}.$$

N Tanks

In the N Tanks model, in addition to the overflow variable (3.21), MLD reformulation for the piecewise linear equations (3.19) and (3.20) defining the interconnection flows has to be performed.

Collector Overflow in the N Tanks Model

The overflow formula for the N Tanks model (3.21) is rewritten as

$$f_c(t) = \delta_{f_c}(t) \frac{1}{\Delta t} \left(v_c^1(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_N^{max} \right),$$

with

$$\delta_{f_c}(t) = \begin{cases} 1, & \text{if } \frac{1}{\Delta t} \left(v_c^1(t-1) + \Delta t \big(q_{in}(t-1) - g(t-1) \big) - v_N^{max} \right) \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Again, this expression analogous to that of the Single Tank model with the difference that, in the N Tanks model, overflow occurs when the upstream tank is full. Therefore, the MLD reformulation is the same using the first interconnection flow q_1 as outflow of the tank instead of the gate flow g. The corresponding set of linear inequalities of the

MLD reformulation reads:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} f_c(t) + \frac{1}{\Delta t} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} v_c^1(t-1) + \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} q_{in}(t-1) +$$

$$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} q_1(t-1) + \begin{pmatrix} -m_c \\ -M_c - \varepsilon \\ -M_c \\ m_c \\ -m_c \\ M_c \end{pmatrix} \delta_{f_c}(t) + \begin{pmatrix} m_c + \frac{v_N^{max}}{\Delta t} \\ \varepsilon - \frac{v_N^{max}}{\Delta t} \\ 0 \\ 0 \\ m_c + \frac{v_N^{max}}{\Delta t} \\ -M_c - \frac{v_N^{max}}{\Delta t} \end{pmatrix} \leq 0,$$

where

$$m_c = \min v_c^1(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_N^{max} = -\Delta t\, g^{max} - v_N^{max},$$
$$M_c = \max v_c^1(t-1) + \Delta t\big(q_{in}(t-1) - g(t-1)\big) - v_N^{max} = \Delta t\, q_{in}^{max}.$$

Finally, in matrix form:

$$\boxed{E_{f_c}\, f_c(t) + E_{V_c}\, V_c(t-1) + E_{Q_c}\, Q_c(t-1) + E_{Q_{in}}\, Q_{in}(t-1) + E_{\delta_{f_c}}\, \delta_{f_c}(t) + E_c \leq 0.} \tag{5.22}$$

Interconnection Flows in the N Tanks Model

Two flow expressions were presented in Section 3.3 for the interconnection flows of the N Tank model: equations (3.19) for the first $N-2$ flows and equation (3.20) for the last one, which is related to the controlled gate flow of the last tank. Since the treatment for both cases is exactly the same, changing the outflow term q_c^{i+1} for g, only the MLD reformulation for equations (3.19) is described in the following.

First, the minimum expression for $q_c^i(t)$ in equation (3.19) is rearranged to obtain an expression of the form $\min\{0, x(t)\}$, which helps reduce the number of binary variables to be defined, i.e.,

$$\begin{aligned} q_c^i(t) &= \min\left\{ \frac{v_c^i(t)}{\Delta t}, \frac{v_N^{max} - v_c^{i+1}(t)}{\Delta t} + q_c^{i+1}(t) \right\} \\ &= \frac{1}{\Delta t} \min\left\{0, v_N^{max} - v_c^{i+1}(t) - v_c^i(t) + \Delta t\, q_c^{i+1}(t)\right\} + \frac{v_i(t)}{\Delta t} \\ &= q_i^{aux}(t) + \frac{v_i(t)}{\Delta t}, \end{aligned}$$

with

$$q_i^{aux}(t) = \delta_c^i(t)\, f_i^{aux}(t),$$

and

$$\delta_c^i(t) = \begin{cases} 1, & \text{if } v_N^{max} - v_{i+1}(t) - v_i(t) + \Delta t\, q_{i+1}(t) \leq 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_i^{aux}(t) = \frac{1}{\Delta t}\left(v_N^{max} - v_c^{i+1}(t) - v_c^i(t) + \Delta t\, q_c^{i+1}(t)\right).$$

Now, according to the MLD formalism, the previous expressions are equivalent to the following set of linear inequalities:

$$\begin{aligned}
&f_i^{aux}(t) \leq M_{q_c}(1 - \delta_c^i(t)), \\
&f_i^{aux}(t) \geq m_{q_c}\, \delta_c^i(t) + \varepsilon\,(1 - \delta_c^i(t)), \\
&q_i^{aux}(t) \leq M_{q_c}\, \delta_c^i(t), \\
&q_i^{aux}(t) \geq m_{q_c}\, \delta_c^i(t), \\
&q_i^{aux}(t) \leq f_i^{aux}(t) - m_{q_c}(1 - \delta_c^i(t)), \\
&q_i^{aux}(t) \geq f_i^{aux}(t) - M_{q_c}(1 - \delta_c^i(t)),
\end{aligned}$$

with

$$\begin{aligned}
m_{q_c} &= \min\ f_i^{aux}(t) = -\frac{v_N^{max}}{\Delta t}, \\
M_{q_c} &= \max\ f_i^{aux}(t) = \frac{v_N^{max}}{\Delta t} + q_i^{max}.
\end{aligned}$$

The previous expressions can be rewritten in matrix form as

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} q_i^{aux}(t) + \frac{1}{\Delta t}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} v_c^{i+1}(t) + \frac{1}{\Delta t}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} v_c^i(t)$$
$$+ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} q_c^{i+1}(t) + \begin{pmatrix} M_{q_c} \\ m_{q_c} - \varepsilon \\ -M_{q_c} \\ m_{q_c} \\ -m_{q_c} \\ M_{q_c} \end{pmatrix} \delta_c^i(t) + \begin{pmatrix} \frac{v_N^{max}}{\Delta t} - M_{q_c} \\ \varepsilon - \frac{v_N^{max}}{\Delta t} \\ 0 \\ 0 \\ m_{q_c} - \frac{v_N^{max}}{\Delta t} \\ \frac{v_N^{max}}{\Delta t} - M_{q_c} \end{pmatrix} \leq 0.$$

Now, using expression $q_i^{aux}(t) = q_c^i(t) - \frac{v_c^i(t)}{\Delta t}$, the previous inequalities are obtained in terms of the desired variables $q_c^i(t)$ as

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} q_c^i(t) + \frac{1}{\Delta t} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} v_c^{i+1}(t) + \frac{1}{\Delta t} \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} v_c^i(t)$$
$$+ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} q_c^{i+1}(t) + \begin{pmatrix} M_{q_c} \\ m_{q_c} - \varepsilon \\ -M_{q_c} \\ m_{q_c} \\ -m_{q_c} \\ M_{q_c} \end{pmatrix} \delta_c^i(t) + \begin{pmatrix} \frac{v_N^{max}}{\Delta t} - M_{q_c} \\ \varepsilon - \frac{v_N^{max}}{\Delta t} \\ 0 \\ 0 \\ m_{q_c} - \frac{v_N^{max}}{\Delta t} \\ \frac{v_N^{max}}{\Delta t} - M_{q_c} \end{pmatrix} \leq 0.$$

By repeating the inequalities in a block-diagonal matrix structure, the following expression for each of the $N-1$ flows is obtained:

$$\boxed{H_{Q_c} Q_c(t) + H_{V_c} V(t) + H_{\Delta_c} \Delta_c(t) + H_G G(t) + H_c \leq 0.} \tag{5.23}$$

Filling Order Property in the N Tanks Model

The first simulation tests carried using the N Tanks Model showed that, due to the addition of further binary variables describing the interconnection flows, the computational times for the optimization problems increased beyond the acceptable ones for RTC. However, by imposing additional constraints on the binary variables, these times can be significantly reduced to achieve the suitable ones (for a comparison of the required computational times with the different models see Section 6.7). In order that these constraints do not alter the system dynamics, they must be obtained from the analytic expressions defining the variables.

In this section, it will be proven that, since the tanks in the N Tanks Model get full one after the other in the upstream direction, the binary variables describing the interconnection flows fulfill the following inequalities:

$$\delta_c^1(t) \leq \delta_c^2(t) \leq \ldots \leq \delta_c^{N-2}(t) \leq \delta_c^{N-1}(t). \tag{5.24}$$

Notice that the fulfilment of the previous set of inequalities reduces the possible configurations of $\left(\delta_c^1(t), \ldots, \delta_c^{N-1}(t)\right)^\top$ from 2^{N-1} to N.

Notice also that since $\delta_c^i(t) \in \{0,1\}$, the previous set of inequalities are equivalent to

$$\delta_c^i(t) = 1 \implies \delta_c^{i+1}(t) = 1, \quad i = 1, \ldots, N-2.$$

To show this, it is easier to prove the equivalent statement

$$\delta_c^{i+1}(t) = 0 \implies \delta_c^i(t) = 0 \quad i = 1, \ldots, N-2.$$

Suppose that $\delta_c^{i+1}(t) = 0$. From the definition, it follows that

$$\delta_c^{i+1}(t) = 0 \iff q_{i+1}(t) = \frac{v_{i+1}(t)}{\Delta t}.$$

Therefore,

$$\begin{aligned} q_i(t) &= \min\left\{\frac{v_i(t)}{\Delta t}, \frac{v_{max} - v_{i+1}(t)}{\Delta t} + q_{i+1}(t)\right\} \\ &= \min\left\{\frac{v_i(t)}{\Delta t}, \frac{v_{max} - v_{i+1}(t)}{\Delta t} + \frac{v_{i+1}(t)}{\Delta t}\right\} \\ &= \min\left\{\frac{v_i(t)}{\Delta t}, \frac{v_{max}}{\Delta t}\right\} = \frac{v_i(t)}{\Delta t}. \end{aligned}$$

And, again, from the definition

$$q_i(t) = \frac{v_i(t)}{\Delta t} \iff \delta_c^i(t) = 0.$$

Finally, inequalities in (5.24) are written in matrix form as

$$K_{\Delta_c}\, \Delta_c(t) \leq 0, \tag{5.25}$$

with

$$K_{\Delta_c} = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{pmatrix},$$

and

$$\Delta_c(t) = \left(\delta_c^1(t), \delta_c^2(t), \ldots, \delta_c^{N-1}(t)\right)^\top.$$

5.2.7 Inflow to Nodes

In the weir, overflow and flood runoff MLD reformulations (5.14), (5.17) and (5.20), variables describing the inflows to each weir $Z_W(t)$ and overflow junction $Z_F(t)$ have been used in order to keep the notation clear and compact. In order to reduce the number of model variables, it is possible to avoid explicitly including these inflow variables in the model by expressing them in terms of the already defined flows of each element type.

To this end, let vector $Z_Q(t)$ collect the inflows to each sewer pipe as defined in equation (3.7):

$$Z_Q(t) = (z_{q_1}(t), \ldots, z_{q_{n_q}}(t))^\top.$$

Now, $Z_Q(t)$ can be expressed in matrix form as

$$\begin{aligned} Z_Q(t) &= A_Q Q_{out}(t) + A_W^+ W(t) + A_G^+ G(t) + A_C C(t) \\ &= \sum_{i=0}^{T} A_Q A_i Q_{in}(t-i) + A_W^+ W(t) + A_G^+ G(t) + A_R R(t), \end{aligned}$$

where A_W^+ and A_G^+ collect only the positive terms in A_W and A_G, respectively.

The components of $Z(t)$ corresponding to inflows to nodes connected to a weir can be selected using a matrix S_W defined as follows:

$$(S_W)_{ij} = \begin{cases} 1, \text{ if } w_i \text{ is connected upstream to the same junction as } q_j \\ 0, \text{ otherwise.} \end{cases}$$

A matrix S_F is defined in the same way, to select the components of $Z(t)$ corresponding to links connected upstream to a junction where overflow is considered to be possible.

Now, using S_W and S_F inflows $Z_W(t)$ and $Z_F(t)$ can be defined as

$$\begin{aligned} Z_W(t) &= S_W Z(t), \\ Z_F(t) &= S_F Z(t). \end{aligned}$$

Remark 5.1. *Some additional parameters have been defined for the computation of the MLD bounds $m_\bullet$ and $M_\bullet$. As mentioned in Section 3.4, the maximum volume for a collector v_c^{max} can be computed from its geometry. On the other hand, for the maximum overflow value f^{max} and the maximum volume of the fictional tank collecting the overflow v_f^{max}, only an upper bound needs to be defined, which can be obtained from simulations or by a trial and error procedure. Similarly, the maximum inflow to a*

collector, weir or overflowing junction (q_{in}^{max}, z_w^{max}, z_f^{max}, respectively) can be chosen as an upper bound value from simulation data. In general, the MLD bounds need not be exact bounds and underestimation of parameters $m_\bullet$ and overestimation of parameters $M_\bullet$ are enough for the formulation to remain valid. However, the tighter the bounds, the more efficiently the computations can be performed when using the model to solve either OCPs or SEPs.

5.3 General Expression

For the formulation of model-based control problems such as OCPs and SEPs, it is simpler to work with general system expressions. To this end, the abstract type of system obtained after the matrix reformulation of the network equations will be used in the following. In addition, the procedures developed in the following section apply to any system which can be written in the same way, and not only to the sewer network model case.

The considered model is formed by the following set of equalities and inequalities relating the system variables at different time-steps:

$$\begin{aligned} \sum_{i=0}^{T} M_i\, X(t-i) &= m(t), \\ \sum_{i=0}^{T} N_i\, X(t-i) &\leq n(t), \end{aligned} \tag{5.26}$$

where $t \in \mathbb{Z}$ is the discrete-time variable and

$$X(t-i) = (x_1(t-i), \ldots, x_n(t-i))^\top,\ i = 0, \ldots, T,$$

with $x_j(t-i) \in \mathbb{R}$ for a subset of indices $j \in \mathcal{C} \subset \{1, \ldots, n\}$ and $x_j(t-i) \in \{0, 1\}$ for a subset of indices $j \in \mathcal{B} \subset \{1, \ldots, n\}$. Index sets $\mathcal{C}$ and $\mathcal{B}$ are such that $\mathcal{C} \cap \mathcal{B} = \varnothing$ and $\mathcal{C} \cup \mathcal{B} = \{1, \ldots, n\}$. M_i, $i = 0, \ldots, T$, and N_i, $i = 0, \ldots, T$, where T is the maximum system delay, are matrices of appropiate dimensions. Discrete variables can have a direct physical meaning but more commonly arise in the formulation of piecewise equations of the model.

Vectors $X(t-i),\ i = 0, \ldots, T$, include all system variables, making no distinction whether they are either state variables or controlled variables. The influence of any disturbance variable at any time step is included in vectors $m(t)$ and $n(t)$.

In Appendix A, the definition of the vector of unknowns X for the sewer network problem is given together with the precise form of matrices $M_i,\ N_i,\ i = 1, \ldots, T$, in terms of the matrices defined in Section 5.2.

5.4 Simulation through Constraint Satisfaction Problems

For simulation purposes, the general expression presented in the previous section can be used to compute the value of $X(t)$, given the values of $X(t-i),\ i = 1, \ldots, T$. Notice that the presence of inequalities and binary variables implies that the computation of $X(t)$ cannot be carried out by simple evaluation of the equality constraints. The problem of finding the values of $X(t)$ that fulfill both the system equalities and inequalities is called a Constraint Satisfaction Problem (CSP). A general description of such problems can be found in Jaulin et al. [2001]. From a practical point of view, CSPs can be solved by means of optimization problems. Constrained optimization algorithms usually solve a CSP in order to find an initial feasible candidate to the solution before starting to iterate towards the optimal one. Taking advantage of this property, CSPs can be solved by any constrained optimization solver by imposing the equations and inequalities of the problem as constraints of an optimization problem minimizing a constant cost function. For system (5.26), such a problem can be written as

$$\begin{aligned} &\min_{X(t)} \ 0 \\ &\text{s.t. } M_0 X(t) = -\sum_{i=1}^{T} M_i \hat{X}(t-i) + m(t), \\ &\qquad N_0 X(t) \leq -\sum_{i=1}^{T} N_i \hat{X}(t-i) + n(t), \\ &\qquad U(t) = \hat{U}(t) \end{aligned} \qquad \text{CSP}(t)$$

where $\hat{U}(t)$ are the known values of the system controlled variables at the current time step and $\hat{X}(t-i),\ i = 1, \ldots, T$, the known values of all the system variables at previous time steps. Notice that for simulation purposes, the values of the controlled variables at every time step must be known beforehand, since they affect the way the system will evolve. This is, in contrast with the OCP presented in Section 5.5, aimed to determine the values of the controlled variables that lead to an optimal performance of the system with respect to several management objectives.

By solving consecutive problems CSP(t) for $t = 1, 2, \ldots$ with initial conditions $\hat{X}(t-i),\ i = 1, \ldots, T$, updated using the solutions of the previous problems, simulations of any length can be performed. Only the initial conditions for the first CSP describing the system state before the start of the simulation event need to be provided as external data inputs.

This simulation method is the one that has been used for the validation results provided in Section 4.2.

Remark 5.2. *For simulations of the sewer network model using this method, the values of the gate flows used as controlled variable inputs are those provided by the physically-based model. Using these values can cause infeasibility problems in the mass balance equations of junctions for which the only outflows are gate flows. These infeasibilities arise from the fact that the inflows to the junction computed by the model need not add exactly to the gate outflows given by the physically-based model. Therefore, these mass balance equations are not included in the CSPs.*

5.5 Optimal Control Problem Formulation

To formulate the OCP associated to the control model (5.26), first the model is extended to include the network equations and MDL inequalities at several time instants ahead in the future as follows:

$$\sum_{i=0}^{T} M_i X(t-i+k) = m(t+k),$$
$$\sum_{i=0}^{T} N_i X(t-i+k) \leq n(t+k),$$
$$k = 1, \ldots, H,$$

where H is called the *prediction horizon*. At time instant t, it is assumed that all the network variables at the previous and current time instants are known, i.e.,

$$\mathcal{X}_0(t) = (\hat{X}(t)^\top, \ldots, \hat{X}(t-T+1)^\top)^\top.$$

Predictions of the rain inflows to the network are also assumed to be available to compute the independent terms $m(t+k)$ and $n(t+k)$, $k = 1, \ldots, H$.

To express the OCP in a compact matrix form, define the following vector collecting all the system unknowns:

$$\mathcal{X}(t) = (X(t+H)^\top, \ldots, X(t+1)^\top)^\top,$$

and the following block-matrices:

$$\mathcal{M}_1 = \left.\begin{pmatrix} M_0 & M_1 & \dots & \dots & M_T & & \\ & \ddots & & & & \ddots & \\ & & M_0 & M_1 & \dots & \dots & M_T \\ & & & M_0 & \dots & \dots & M_{T-1} \\ & & & & \ddots & & \vdots \\ & & & & & M_0 & M_1 \\ & & & & & & M_0 \end{pmatrix}\right\} \begin{matrix} H \\ \text{blocks} \end{matrix}$$

$$\mathcal{M}_2 = -\left.\begin{pmatrix} & & & & & & \\ & & & & & & \\ & & & & & & \\ M_T & & & & & & \\ M_{T-1} & M_T & & & & & \\ \vdots & \vdots & \ddots & & & & \\ M_2 & M_3 & \dots & M_T & & & \\ M_1 & M_2 & \dots & M_{T-1} & M_T & & \end{pmatrix}\right\} \begin{matrix} H \\ \text{blocks} \end{matrix},$$

$$\mathcal{M}_3(t) = (m(t+H)^\top, \dots, m(t+1)^\top)^\top,$$

with similar expressions for $\mathcal{N}_1$, $\mathcal{N}_2$ and $\mathcal{N}_3$. For the construction of these matrices it has been assumed that $H > T$, that is, the prediction horizon is greater than the largest delay in the system. Although this is not a necessary condition for the formulation of the OCP, it is a common assumption that allows the OCP to evaluate the performance of the system taking into account all the effects of the network dynamics.

Now, the OCP can be stated as

$$\begin{aligned} \min_{\mathcal{X}(t)} \ & J(\mathcal{X}(t)) = c^\top \mathcal{X}(t), \\ \text{s.t.} \quad & \mathcal{M}_1\, \mathcal{X}(t) = \mathcal{M}_2\, \mathcal{X}_0(t) + \mathcal{M}_3(t), \\ & \mathcal{N}_1\, \mathcal{X}(t) \leq \mathcal{N}_2\, \mathcal{X}_0(t) + \mathcal{N}_3(t), \\ & A_{eq}\, \mathcal{X}(t) = b_{eq}(t), \\ & A_{ineq}\, \mathcal{X}(t) \leq b_{ineq}(t). \end{aligned} \qquad \text{OCP}(t)$$

Details of the form of the cost function $J(\mathcal{X}(t))$ used to quantify the management objectives of the case study network are given in Section 6.3. In this case, a linear function of the problem variables has been used and the corresponding optimization problem becomes a MILP. More generally, a quadratic function of the form

$$J(\mathcal{X}(t)) = \mathcal{X}(t)^\top \Phi \mathcal{X}(t) + \phi^\top \mathcal{X}(t)$$

can be considered, where Φ and ϕ are respectively a square matrix and a vector of the dimensions of $\mathcal{X}$. In this case, the corresponding optimization problem would become

a *Mixed Integer Quadratic Programming* problem (MIQP). Both the MILP and the MIQP versions of the OCP can be efficiently solved by using appropriate optimization software. Details on the problem size and computational results for the case study are provided in Section 6.7.

5.5.1 Additional Constraints

Additional constraints in (OCP(t)) of the form

$$A_{eq}\,\mathcal{X}(t) = b_{eq}(t)$$
$$A_{ineq}\,\mathcal{X}(t) \leq b_{ineq}(t)$$

are added to the OCP to take into account:

1. Bounds on variables:
$$\mathcal{X}_{min} \leq \mathcal{X}(t) \leq \mathcal{X}_{max}.$$

2. Bounds on the variation of the gate flows for smooth control actions:
$$\Delta g_{min} \leq \Delta g(t) \leq \Delta g_{max}.$$

3. Mass balance in junctions with outflowing gates:
$$z_g(t) = \sum_{k=1}^{n} g_{i_k}(t),$$
where $z_g(t)$ is the total inflow to the junction where gates are connected.

4. Filling order property (5.25), for the N Tanks collector model.

5.6 State Estimation Problem Formulation

State estimation problems aim to reconstruct the full system state out of a few output measurements. To this end, the difference between the system measurements and the outputs generated by the estimator model is minimized along finite past horizon by means of an optimization problem. Due to plant-model mismatch, generally, no sequence of model-generated outputs can meet exactly the measured values. To take

into account this fact, two formulations of the SEP are usually applied: the first one involves adding slack/noise variables to each system dynamic and output equation and minimizing these variables while forcing the model outputs to be equal to the measured ones [Bemporad et al., 2000, Ferrari-Trecate et al., 2002, Rao et al., 2001]. The second formulation is a direct minimization of the norm of the difference between the model-generated outputs and the measured ones. In this case, model outputs and the measured ones are not assumed to be equal [Breckpot et al., 2010, Busch et al., 2013, Michalska and Mayne, 1995]. While the former approach appears to be more appealing for theoretical purposes, the latter results into smaller optimization problems if the number of measured variables is smaller than the number of dynamic equations. Since this is likely to be the case in a sewer network, where usually only a few measurements are available, the second approach is the one that will be used in the following.

Notice that the well-known Kalman Filter (KF) technique is not commonly used in the context of hybrid systems. For hybrid systems, the efficient explicit recursive solution of the KF optimal estimate is not available. Therefore, a mixed integer SEP would be required to be solved at each control iteration. Since the KF is a *full information estimator* (that is, all the available previous measurements are used), the estimation horizon of each problem would be greater. With mixed-integer optimization, these extended estimation horizons would soon result into problems not solvable in real time. The MHE technique provides a simple solution to this problem by solving SEPs with a fixed-size horizon at each time step.

The SEP proposed in this thesis is analogous to the OCP one but the system dynamics and inequality constraints are enforced for the past states rather than for the future ones, i.e.,

$$\begin{aligned} &\sum_{i=0}^{T} M_i X_{\mathcal{O}}(t-i+k) = m(t+k), \\ &\sum_{i=0}^{T} N_i X_{\mathcal{O}}(t-i+k) \leq n(t+k), \\ &\qquad k = -H_{\mathcal{O}} + T + 1, \ldots, 0, \end{aligned} \tag{5.27}$$

where $H_{\mathcal{O}}$ is the number of past instant measured variables that will be used in the

problem. The vector of unkonwn variables for the SEP is then defined as

$$\mathcal{X}_{\mathcal{O}}(t) = (X_{\mathcal{O}}(t)^\top, \ldots, X_{\mathcal{O}}(t - H_{\mathcal{O}} + 1)^\top)^\top.$$

To express the constraints in matrix form, the following matrices are defined:

$$\mathcal{M}_1^{\mathcal{O}} = \left(\begin{array}{ccccccc} M_0 & M_1 & \ldots & M_{T-1} & M_T & & \\ & \ddots & \ddots & & \ddots & \ddots & \\ & & M_0 & M_1 & \ldots & M_{T-1} & M_T \end{array} \right) \Bigg\} \begin{array}{l} H_{\mathcal{O}} - T \\ \text{blocks} \end{array},$$

$$\mathcal{M}_2^{\mathcal{O}}(t) = (m_{\mathcal{O}}(t)^\top, \ldots, m_{\mathcal{O}}(t - H_{\mathcal{O}} + 1)^\top)^\top,$$

with analogous expressions for $\mathcal{N}_1^{\mathcal{O}}$ and $\mathcal{N}_2^{\mathcal{O}}$.

To formulate the SEP it is important to distinguish, among the system variables $X(t)$, the *output* variables $Y(t) \in \mathbb{R}^{n_y}$, which are the variables whose value is measured at every time step, and the *input* variables $U(t) \in \mathbb{R}^{n_u}$, whose value is known for being the system controlled variables. In presence of a system level controller providing the values of controlled variables as set-points for local controllers, these variables must also be measured, since the values reached by the local controllers may eventually be different than the set-points.

Since both the output and the controlled variables are a subset of the system variables, they can be obtained by means of 0-1 projection vectors π_Y and π_U as

$$\begin{aligned} Y(t) &= \pi_Y \, X(t), \\ U(t) &= \pi_U \, X(t). \end{aligned}$$

Similarly, the extended vectors to the estimation horizon $H_{\mathcal{O}}$ are obtained as

$$\begin{aligned} \mathcal{Y}(t) &= \Pi_{\mathcal{Y}} \mathcal{X}_{\mathcal{O}}(t) = (Y(t)^\top, \ldots, Y(t - H_{\mathcal{O}} + 1)^\top)^\top, \\ \mathcal{U}(t) &= \Pi_{\mathcal{U}} \mathcal{X}_{\mathcal{O}}(t) = (U(t)^\top, \ldots, U(t - H_{\mathcal{O}} + 1)^\top)^\top, \end{aligned} \tag{5.28}$$

with block diagonal matrices $\Pi_{\mathcal{Y}}$ and $\Pi_{\mathcal{U}}$ defined by

$$\Pi_{\mathcal{Y}} = \left(\begin{array}{ccc} \pi_Y & & \\ & \pi_Y & \\ & & \ddots \\ & & \pi_Y \end{array} \right) \Bigg\} \begin{array}{l} H_{\mathcal{O}} \\ \text{blocks} \end{array}, \quad \Pi_{\mathcal{U}} = \left(\begin{array}{ccc} \pi_U & & \\ & \pi_U & \\ & & \ddots \\ & & \pi_U \end{array} \right) \Bigg\} \begin{array}{l} H_{\mathcal{O}} \\ \text{blocks} \end{array}. \tag{5.29}$$

Now, the state estimation problem can be written as:

$$\begin{aligned}
\min_{\{X_{\mathcal{O}}(t),\,\varepsilon_{\mathcal{Y}},\,\varepsilon_{\mathcal{U}}\}} \quad & \mathbf{1}_{\mathcal{Y}}^{\top}\,\varepsilon_{\mathcal{Y}} + \mathbf{1}_{\mathcal{U}}^{\top}\,\varepsilon_{\mathcal{U}}, \\
\text{s.t.} \quad & M_1^{\mathcal{O}}(t)\,\mathcal{X}_{\mathcal{O}}(t) = M_2^{\mathcal{O}}(t), \\
& N_1^{\mathcal{O}}(t)\,\mathcal{X}_{\mathcal{O}}(t) \leq N_2^{\mathcal{O}}(t), \\
& -\varepsilon_{\mathcal{Y}} \leq \Pi_{\mathcal{Y}}\,\mathcal{X}_{\mathcal{O}}(t) - \hat{\mathcal{Y}}(t) \leq \varepsilon_{\mathcal{Y}}, \\
& -\varepsilon_{\mathcal{U}} \leq \Pi_{\mathcal{U}}\,\mathcal{X}_{\mathcal{O}}(t) - \hat{\mathcal{U}}(t) \leq \varepsilon_{\mathcal{U}}, \\
& A_{eq}\,\mathcal{X}_{\mathcal{O}}(t) = b_{eq}(t), \\
& A_{ineq}\,\mathcal{X}_{\mathcal{O}}(t) \leq b_{ineq}(t),
\end{aligned} \qquad \text{SEP}(t)$$

where $\hat{\mathcal{Y}}(t)$ and $\hat{\mathcal{U}}(t)$ are the measured values of the input and output variables, $\mathbf{1}_{\mathcal{Y}}$ and $\mathbf{1}_{\mathcal{U}}$ are vectors of ones of dimensions $H_{\mathcal{O}} \cdot n_y$ and $H_{\mathcal{O}} \cdot n_u$, respectively, and $\varepsilon_{\mathcal{Y}}$ and $\varepsilon_{\mathcal{U}}$ are auxiliary variables used to reformulate the minimization of the 1-norms $\|\Pi_{\mathcal{Y}}\,\mathcal{X}_{\mathcal{O}}(t) - \hat{\mathcal{Y}}\|_1$ and $\|\Pi_{\mathcal{U}}\,\mathcal{X}_{\mathcal{O}}(t) - \hat{\mathcal{U}}\|_1$ as a *mixed integer linear problem* (MILP) [Boyd and Vandenberghe, 2004]. Additional equalities $A_{eq}\,\mathcal{X}(t) = b_{eq}(t)$, and inequalities $A_{ineq}\,\mathcal{X}(t) \leq b_{ineq}(t)$, are the same as those commented in Section 5.5.1 for the OCP case.

Notice that, in (5.27), the system equations are only enforced for the last $H_{\mathcal{O}} - T$ time instants: $X_{\mathcal{O}}(t)$ to $X_{\mathcal{O}}(t - H_{\mathcal{O}} + T + 1)$. Therefore, $H_{\mathcal{O}} \geq 2\,T$ is assumed so that the system equations are enforced for the variables needed to be used as initial conditions for the OCP. The rest of variables at the first T time instants $t - H_{\mathcal{O}} + T, \ldots, t - H_{\mathcal{O}} + 1$, are left free. In this way, estimated inputs and outputs at these times will take exactly the same values as the measured ones as a result of the optimization contributing to a better estimation of the rest of variables.

Details on the problem size and computational results for the case study are provided in Section 6.7.

5.7 Summary

In this chapter, the sewer network model presented in Chapter 3 has been reformulated to be expressed in a compact matrix form for a network with an arbitrary number of elements of each type. To this end, the piecewise linear functions defining the weir flow, the overflows, the flood runoff and the collector flows have been substituted for

equivalent sets of linear inequalities involving newly defined binary variables, following the Mixed Linear Dynamic (MLD) systems approach. After coupling all the system equations and MLD inequalities and taking into account the presence of delays in the system, a final general expression for the model has been obtained, which could also be used to represent other systems with similar features.

Based on the general system expression, simple matrix-based procedures for the formulation of simulation and control-related problems have been developed. First, a Constraint Satisfaction Problem is formulated, which allows to compute the system variables at a given time step using past values and control inputs. By iteratively solving a series of such problems, simulations of any length can be performed. Secondly, by coupling the system equations and inequalities at several future time instants, an Optimal Control Problem has been formulated. Through the minimization of an objective function describing the system performance, optimal control actions over a finite optimization horizon can be computed. Finally, by imposing the system equations and inequalities at past time instances, a State Estimation Problem is presented. By minimizing the difference between measured variables and those computed by the model, the whole system state can be suitably approximated. All the obtained problems are mixed-integer linear problems that can be solved by means of specialized standard software avoiding the need to implement *ad hoc* optimization routines.

Chapter 6

Receding Horizon Control

This chapter is partially based on:

- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid modeling and receding horizon control of sewer networks.* Water Resources Research*, 2014d.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Receding horizon control of hybrid linear delayed systems: Application to sewer networks.* IEEE Conference on Decision and Control*, 2013a. Firenze, Italy.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. A control-oriented hybrid modelling approach for sewer networks: Barcelona case study.* IWA Conference on Instrumentation, Automation and Control*, 2013b. Narbonne, France.*
- *B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Output-feedback control of sewer networks thorugh moving horizon estimation.* IEEE Conference on Decision and Control*, 2014b. Los Angeles, USA.*

6.1 Receding Horizon Control and Moving Horizon Estimation

Receding Horizon Control (RHC) is a RTC strategy aimed to take full advantage of model-based control techniques using real-time measurements and disturbance forecasts. To this end, after solving a finite-time optimal control problem, only the part of the sequence of control actions obtained as a solution corresponding to the first time step is applied to the system. After letting the system respond to this action for the corresponding time step, feedback measurements are taken and a new OCP is formulated, using the latest available information, and solved to compute the control action

for the next time step and repeat the whole procedure again, along the simulation scenario. Depending on the available measurements, the initial conditions for each of the subsequent OCPs can be directly obtained or must be estimated. In the latter case, before solving each OCP, a SEP is solved to reconstruct the full-state initial condition necessary to formulate the OCP. This technique consisting in solving a SEP at each time step based on the last available measurements is known as Moving Horizon Estimation (MHE) and is regarded as the observer counterpart of the RHC strategy.

In some cases, the time step used in the control model to provide sufficient accuracy might not be adequate to be used in the RHC strategy as described above. This fact might be due to additional time required to gather system measurements from a SCADA system and formulate and solve the SEPs and OCPs or due to limitations in the actuators. In any case, the RHC/MHE strategy can still be applied by updating and solving the SEPs and OCPs every few time steps instead of every one. The number of time steps elapsed between updating and solving two consecutive SEPs and OCPs is called the *control interval*.

In the sewer network corresponding to the case study described in Chapter 4, to provide a suitable approximation of the flow delay in the sewer pipes, a sampling time of $\Delta t = 1\,\text{min}$ was chosen. Taking into account that gates can only move at limited speeds, this time step is not sufficient for the local controllers to achieve the gate flow set-points $G^*(t)$ obtained as the solution of the OCPs. Therefore, a control interval of five time steps (i.e., five minutes) was chosen and the set-points produced by the OCP are assumed to be constant for five minutes periods. To take this into account in the control model, a constraint forcing gate flows to remain constant along five time steps was added to the OCPs as

$$g(t + 5\,k) = \ldots = g(t + 5\,(k + 1) - 1),$$

for $k = 0, 1, \ldots, (H - 5)/5$. Notice that, if a sampling time of five minutes was chosen, this additional constraint would not be required. In that case, however, the prediction accuracy of the model would decrease. On the other hand, the computational time required to solve the SEPs and OCPs would also be decreased due to the reduction of the number of variables needed to cover the same prediction horizon. Since in the case-study network the computational times for a one minute time step are shown to

be suitable (c.f. Section 6.7), this sample time was kept in order to achieve a higher prediction accuracy.

Following the RHC/MHE strategy, four different scenarios depending on the available measurements have been considered and will be discussed in the following sections:

- State feedback RHC
  - Full flow measurements
- Output feedback RHC with MHE
  - Limited flow measurements of sewer pipes and collector
  - Limited water level measurements of sewer pipes and collector
  - Limited water level measurements of sewer pipes plus flow measurement at the collector inflow

6.2 Closed-Loop Simulation Algorithm

As mentioned in previous chapters, to test this RHC strategy, the commercial physically-based model simulator MOUSE (MOUSE [2007c]) has been used as a virtual reality providing what in a real case would correspond to flow or water level measurements. According to the RHC/MHE strategy described above, for the simulation of a closed-loop control event a series of SEPs, OCPs and physically-based model simulations (substituting the real evolution of the system) must be solved and executed.

From an implementation point of view, closed-loop simulations require a bidirectional communication between the physically-based simulator and the optimization module. The overall closed-loop algorithm is written as a MATLAB script, which solves the SEPs and OCPs and calls the simulator executable `mouse604.exe` through command-line as:

```
"path1/mouse604.exe" "path2/file.MPR" HD Run Close NoPrompt Hide
```

for the hydraulic simulation and

```
"path1/mouse604.exe" "path2/file.MPR" RD Run Close NoPrompt Hide
```

for the hydrologic simulation. The `.MPR` file is the MOUSE project file including the information of the simulated event such as initial and final simulation times, rain intensity time series files (`.bbf`), hydraulic and hydrologic models, network description file (`.UND`), hydrologic catchment description file (`.HGF`) and dry weather flow parameters file (`.DWF`). The additional parameters `Close NoPrompt Hide` force the application to close at the simulation end, not prompt confirmation messages and hide the simulation status window so that simulations can be executed in a loop.

The result of such simulations are binary encoded files, `.PRF` for the hydraulic case and `.CRF` for the hydrologic case. By means of extractor application `m11extra.exe` results in the form of text files can be generated. First, the extractor file is copied into the same directory as the result file and is executed over it as

```
m11extra.exe file.PRF
```

to generate a new text file `M11.OUT` with a list of the variables included in the result file. This file must be copied under the name `M11.IN` and edited to indicate which are the variables to be extracted by inserting a `1` at the beginning of the line with the variable name. Finally, the extractor is executed again indicating the name of the output text file

```
m11extra.exe file.PRF outputfile.txt
```

The same procedure is applied to the hydrologic results file `.CRF` to extract the rainfall-runoff inflows computed from rain intensity data. Since rain inflows are not affected by the control actions this step is not performed at every closed-loop simulation iteration but only once, before strating the simulation.

The final text files containing the simulation results can then be read by MATLAB and transformed into `.mat` files used to update the SEPs and OCPs initial conditions. After the OCP is solved, the constant set-point for the gate flows for the next five-minutes simulation are written into the Real-Time Control section of the network description file `.UND` to repeat the whole procedure again.

6.3 Management Objectives and Cost Function

The RHC strategy has been applied to the four rain events that were used for the model calibration and validation to achieve the following control objectives:

1. Minimize overflows
2. Minimize CSO discharges
3. Maximize waste water treatment plant (WWTP) usage

These objectives are quantified in the following multi-objective cost function:

$$\begin{aligned} J(\mathcal{X}(t)) = &\gamma_{COF}\, J_{COF}(\mathcal{X}(t)) + \gamma_{OF}\, J_{OF}(\mathcal{X}(t)) + \\ &\gamma_{CSO}\, J_{CSO}(\mathcal{X}(t)) - \gamma_{WWTP}\, J_{WWTP}(\mathcal{X}(t)), \end{aligned} \tag{6.1}$$

where $J_{COF}(\mathcal{X}(t))$ is the overflow of collector q_{139}, $J_{OF}(\mathcal{X}(t))$ contains the sum of the rest of the overflow variables at junctions, $J_{CSO}(\mathcal{X}(t))$ contains the sum of flow variables corresponding to the sewer pipes connecting the network to the sea and $J_{WWTP}(\mathcal{X}(t))$ contains the sum of flow variables corresponding to the sewer pipes connecting the network to the WWTP.

The model is flexible enough to accommodate control objectives other than the ones proposed in objective function (6.1), e.g., minimizing of CSOs caused by weir flow, prioritizing CSO events at different points of the network, prioritizing the use of different WWTPs, etc. Moreover, using a quadratic objective function tracking objectives can also be defined such as tracking of flows to the WWTPs or tracking of levels in tanks. The choice of the terms and weights in (6.1) should reflect the order of importance of these individual control objectives. Moreover, the network topology plays an important role on the way objectives interact with one another and trial and error tests are necessary to correct the weights to compensate individual objective and global objective performances. One possible general procedure to determine objective function weights is to perform closed-loop simulations using single-goal objective functions. The performance results of these simulations provide a reference for each objective to be compared with the multiple-objective simulation results. Then, starting with a multi-objective function with weights of different orders of magnitude according to the priority of the objectives (normalization factors must also be included if both flows and volumes are involved), closed-loop simulation tests can be performed to assess whether the interactions among the different individual objectives cause global performance losses and then correct the weights accordingly.

The choice of weights in (6.1) for the topology of the case-study network has been performed as follows. Notice that the fulfilling of all the proposed control objectives for the case study benefits from low flow rates, which encourage the absence of overflows and contribute to not saturating collector q_{139} at the downstream end of the network, which, in turn, leads to avoid CSO discharges. Taking into account that the values of the weights are only relevant relative to one another, γ_{OF} and γ_{CSO} are set to 1. Overflows at the collector would be especially dangerous in case of high flow rates, therefore γ_{COF} has been set to 10, implying that the collector overflow is prevented possibly by means of allowing overflows to occur elsewhere upstream. Finally, the selected weight for the WWTP term is set to $\gamma_{WWTP} = 10^{-1}$. The negative sign of γ_{WWTP} in (6.1) is used to obtain maximization of this goal (while the others are minimized). Trial and error tests showed that negative weights in the WWTP term of the same order as the CSO ones lead to higher CSO results due to the fact that the negative and positive terms of the two objectives compensated [Joseph-Duran et al., 2013a].

6.4 State Feedback RHC

The first test to assess the performance of the proposed receding horizon controller is carried out assuming a rather improbable situation in which measurements of the network flows are available at all the network sewer pipes, gates and weirs. In this case, no SEP needs to be solved, since, using the measured flows, the rest of the model variables can be computed using the model equations. Although assuming full flow measurement is unrealistic, the results of this test will be useful as a reference to assess the performance of the RHC strategy when used together with the MHE technique. Moreover, using the results of this simulation, one of the collector models described in Section 3.3 will be selected to be used in the output feedback closed-loop simulations for sequel sections. Figure 6.1 shows a diagram of the closed-loop simulation algorithm for the state feedback configuration. The detailed procedure for a simulation event of t_s time steps is described in Algorithm 1.

The first closed-loop simulations of the proposed model-based controller were conducted using the single tank model for the collector, with a prediction horizon of $H = 30$ time steps (i.e., 30 minutes) and a safety factor for the collector volume of 0.7 [Joseph-Duran et al., 2013b, 2014d]. Table 6.1 shows the results obtained from those simulations

Table 6.1: RHC results and variations with respect to passive control.

Episode	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
17-09-2002	0.16 (-96%)	21.81 (-79%)	164.90 (64%)
09-10-2002	0.90 (-97%)	345.04 (-31%)	257.77 (108%)
15-08-2006	0.25 (-96%)	7.51 (-92%)	149.86 (75%)
30-07-2011	0.75 (-96%)	54.07 (-63%)	159.57 (83%)

and the variations in the objectives compared with the no control results obtained by simulating the rain events with gates set at fixed positions (*passive control*). For each objective J_{PC} in the passive control scenario and the corresponding one J_{RHC} in the RHC scenario, the percent variation has been computed as

$$\frac{J_{RHC} - J_{PC}}{J_{PC}} \cdot 100.$$

The actual network regulation is performed by expert operators and no data related to the real management of the network for the considered rain scenarios is available for comparison. Results show that appropriate management of the detention tanks at the upper part of the network can mitigate overflows almost completely (most overflow volume reported in Table 6.1 corresponds to overflow points upstream of any control action) by reducing the peak flows in the network sewer pipes. The volume stored in the tanks can be released later at adequate flow rates to maximize the use of the WWTP capacity. On the other hand, the use of the in-line capacity of sewer pipe q_{139} results in a reduction of the CSO volumes.

Due to the addition of extra delays in the collector models, simulations to compare the different approaches have been performed with an increased prediction horizon of $H = 40$. After analysing the results of the first simulations, the safety factor for the collector volume has been relaxed to 0.9. Table 6.2 shows the results of the closed-loop simulations for the different collector models, including different numbers of tanks (i.e. different delays). Notice that since the delay and the water distribution along the collector depends on the movements of the downstream gate, comparing the models by means of simple simulations, as in Section 4.2, does not provide much information on their predictive capabilities. Such simulations would only provide the volume distribution along the tanks, since the gate flow would be used as an input. Therefore, closed-loop simulations must be carried to assess the ability of the different submodels

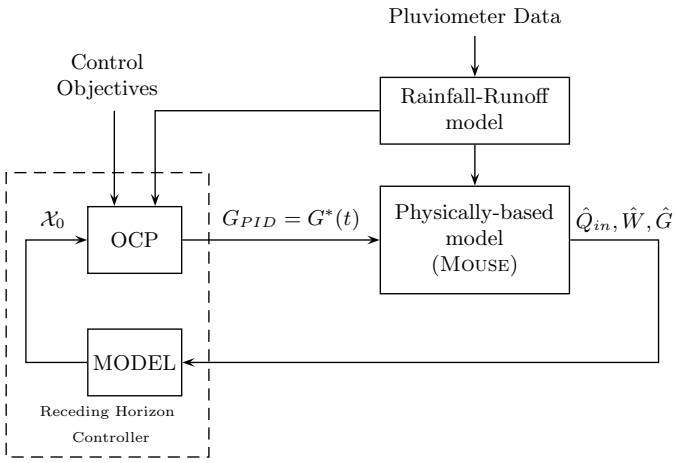

Figure 6.1: Closed-loop simulation algorithm diagram with full-flow measurements.

in computing the gate flow set-points and the effect of these set-points in the delay and volume distribution in future time steps.

Generally, it can be noticed that, although the delayed models usually perform better than the single tank one, none of the models performs clearly better than the others in this case. This is mainly due to the maximum volume constraint, present in all models. Although the delayed models take into account the time needed for the flow to reach the downstream gate, the volume constraint considers the volume along the whole collector, therefore taking into account the inflow volume as soon as it enters the collector: thus with no delay. Therefore, although for the three submodels the distribution of the volume along the collector is different, the total contained volume is similar and also is the computed outflow, leading to similar CSO results. It is also worth noting that the fact that the OCPs do not have a unique solution and that the solution at each time step is affected by all the previous ones (through the initial conditions), leads to different inflow curves to the collector for the different collector models. Thus, a detailed comparison of the controller performance according to the collector model is not possible based on the results in Table 6.2.

For the Moving Horizon Estimation simulations of the next sections, the Single Tank Model plus delay with a delay of N = 10 time steps (N tanks) has been chosen since it is the one providing the best results for the two most intense rain events (09-10-2002 and 30-07-2011) with marginal variations with respect to the other models for

Algorithm 1: Closed-Loop State Feedback RHC Algorithm

Input : $\mathcal{X}_0(0) = \left(\hat{X}(0)^\top, \ldots, \hat{X}(-T+1)^\top\right)^\top = \mathbf{0}$
begin
  Set $t := 0$
  while $t \leq t_s$ **do**
    Compute rainfall-runoff prediction $R_H(t) = \left(r(t+1)^\top, \ldots, r(t+H)^\top\right)^\top$
    Compute $\mathcal{M}_3(t)$, $\mathcal{N}_3(t), b_{eq}(t), b_{ineq}(t)$ from $\mathcal{X}_0(t)$, $R_H(t)$
    Solve OCP $\rightarrow \mathcal{X}^*(t) = \left(X^*(t+H)^\top, \ldots, X^*(t+1)^\top\right)^\top$
    Run MOUSE for simulation time $(t, t+5)$ with gate PID set-points $G_{PID} = G^*(t)$
    Extract MOUSE data from result files: $\hat{Q}_{in}(t+5)$,...,$\hat{Q}_{in}(t)$, $\hat{W}(t+5)$,...,$\hat{W}(5)$, $\hat{G}(t+5)$,...,$\hat{G}(t)$
    Compute $\hat{\mathcal{X}}(t+5) = \left(\hat{X}(t+5)^\top, \ldots, \hat{X}(1)^\top\right)^\top$, using the model
    Set $\mathcal{X}_0(t+5) := \left(\hat{X}(t+5)^\top, \ldots, \hat{X}(t+5-T+1)^\top\right)^\top$
    Set $t := t+5$
  end
end

the other two rain scenarios. The prediction horizon in all cases is kept to $H = 40$ time steps. In the following sections, the results corresponding to this scenario will be referred to as FSM (Full-State Measurement).

Table 6.2: Closed-loop simulation results of the diferent collector models with full flow measurements.

(a) 17-09-2002

Model	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
1T	0.15	10.21	161.80
1TD N= 5	0.16	7.76	151.99
1TD N=10	0.16	9.72	160.36
ND N= 5	0.16	4.16	162.73
ND N=10	0.16	2.52	163.45

(b) 09-10-2002

Model	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
1T	1.21	345.44	155.51
1TD N= 5	1.03	345.16	158.33
1TD N=10	1.01	342.09	158.27
ND N= 5	1.01	365.32	158.59
ND N=10	0.99	362.05	158.58

(c) 15-08-2006

Model	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
1T	0.25	5.39	135.41
1TD N= 5	0.25	5.72	111.04
1TD N=10	0.25	5.54	122.67
ND N= 5	0.25	5.92	136.19
ND N=10	0.25	4.72	135.97

(d) 30-07-2011

Model	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
1T	0.75	42.89	148.49
1TD N= 5	0.75	41.41	139.91
1TD N=10	0.75	39.92	155.87
ND N= 5	0.75	36.24	150.91
ND N=10	0.75	36.65	150.47

Remark 6.1. *The physically-based model simulator cannot perform simulations for completely empty sewer pipes due to numerical problems. Therefore, a minimum water level is always enforced at all the network sewer pipes. This, together with dry-weather flows, causes that, at the end of a closed-loop simulation event, the total volume that left the network (through CSOs or towards the WWTP) is greater than the total rain inflows reported in Table 4.2. This is especially noticeable for long events.*

6.5 Output Feedback RHC

Due to the large-scale nature of sewer networks, the most common situation is that measurements of the network variables are only available at certain points. Moreover, instrumentation for water level measurements is cheaper and more reliable than that aimed to measure flow rates. To take into account these facts, the model-based RHC/MHE strategy designed in this thesis has been applied to the case study network taking into account the available instrumentation. In fact, only level measurements through limnimeters are available in the Riera Blanca network. The measurement points are shown in Figure 6.2. Since the local PID controllers at the gates implemented in the physically-based model simulator need flow measurements to regulate

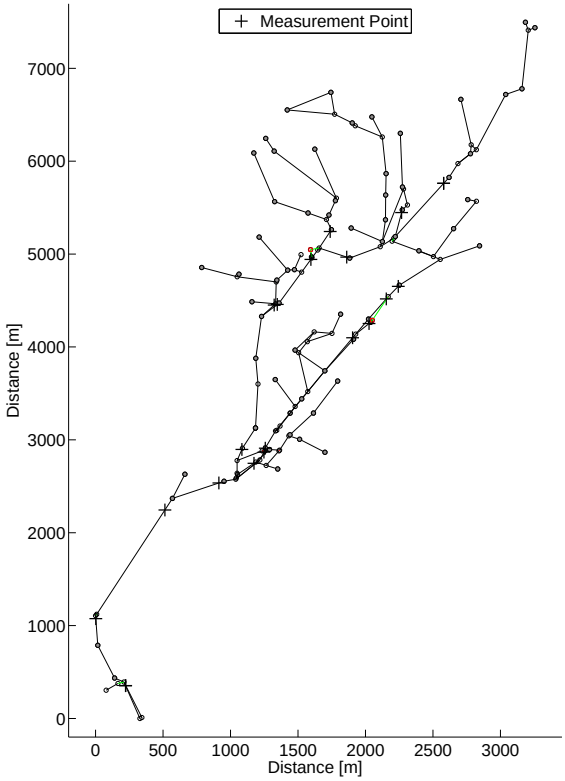

Figure 6.2: Measurement locations in the Riera Blanca sewer Network.

the gate position, it will be assumed in the following that flow measurements are always available at the gate outputs. Although this is not the case in reality, accurate approximations of the gate outflows can be obtained by means of the water levels up- and downstream and the gate position.

In the following, the performance of the model according to four configurations regarding the available measurements are compared and discussed. As a first test, flow measurements have been considered at the limnimeter locations (from now on, this scenario will be referred to as MHEF). Secondly, only water level measurements have been considered. In this scenario, flows at the measurement points have been reconstructed by means of flow-level polynomial approximations calibrated from data generated by the physically-based model (from now on, this scenario will be referred to as MHEL). Taking into account the results of the previous two scenarios, closed-loop simulations have been conducted assuming flow measurements at the collector inflow and water level measurements at the rest of the limnimeter locations. It is shown that

in thi measurement scenario a trade off between the results of the the MHEF and MHEL is obtained (from now on, this scenario will be referred to as MHEC). Finally, by adding a second flow measurement to the MHEC scenario, corresponding to the sewer pipe immediately upstream of the collector, the MHEF performance is recovered (from now on, this scenario will be referred to as MHEC2). Therefore it is concluded that by installing two flow-meters at the collector inflow and at its upstream sewer pipe, control performance could be improved.

6.5.1 Flow Measurements

To formulate the SEP in the case of flow measurements, it is necessary to define the measured and input variables as

$$Y(t) = \left(q_{i_1}(t), \ldots, q_{i_{n_y}}(t)\right),$$
$$U(t) = \left(g_1(t), \ldots, g_{n_g}(t)\right),$$

where i_j, $j = 1, \ldots, n_y$, are the indices of the sewer pipes for which a measurement is available and $n_y = 20$ is the number of measurements. As mentioned before, flows at the gate outputs are also measured. Then, matrices (5.28) and (5.29) can be readily constructed to formulate the SEP.

Algorithm 2: Closed-Loop Output Feedback RHC Algorithm with Flow Measurements

Input : $\mathcal{X}_0(0) = \left(\hat{X}(0)^\top, \ldots, \hat{X}(-T+1)^\top\right)^\top = \mathbf{0}$
begin
  Set $t := 0$
  while $t \leq t_s$ **do**
    Compute rainfall-runoff prediction $R_H(t) = \left(r(t+1)^\top, \ldots, r(t+H)^\top\right)^\top$
    Compute $\mathcal{M}_3(t)$, $\mathcal{N}_3(t), b_{eq}(t), b_{ineq}(t)$ from $\mathcal{X}_0(t)$, $R_H(t)$
    Solve OCP$(t) \rightarrow \mathcal{X}^*(t) = \left(X^*(t+H)^\top, \ldots, X^*(t+1)^\top\right)^\top$
    Run MOUSE for simulation time $(t, t+5)$ with gate PID set-points $G_{PID} = G^*(t)$
    Extract MOUSE data from result files: $\hat{G}(t+5)$,...,$\hat{G}(t)$, $\hat{Q}(t+5)$,...,$\hat{Q}(t)$
    Solve SEP$(t+5) \rightarrow \mathcal{X}^*_{\mathcal{O}}(t+5) = \left(X^*_{\mathcal{O}}(t+5)^\top, \ldots, X^*_{\mathcal{O}}(t+5-H_{\mathcal{O}}+1)^\top\right)^\top$
    Set $\mathcal{X}_0(t+5) := \left(X^*_{\mathcal{O}}(t+5)^\top, \ldots, X^*_{\mathcal{O}}(t+5-T+1)^\top\right)^\top$
    Set $t := t+5$
  end
end

Algorithm 1 is adapted to include the corresponding SEP to estimate the initial conditions of the OCPs from the measured variables $\hat{\mathcal{U}}$ and $\hat{\mathcal{Y}}$, as detailed in Algorithm 2 and shown in Figure 6.3. All SEPs have been formulated and solved with an estimation horizon of $H_{\mathcal{O}} = 15$ time steps.

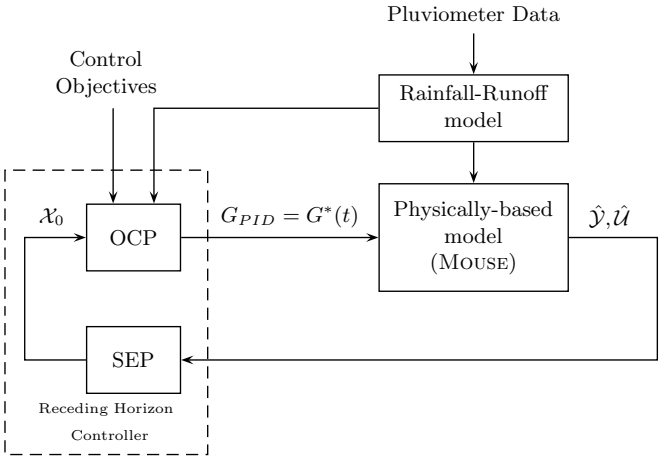

Figure 6.3: Closed-loop simulation algorithm diagram with available flow measurements.

Table 6.3 shows the closed-loop simulation results for the three network objectives and its variation with respect to the FSM scenario. It can be noticed that marginal variation of the considered objectives with respect to the FSM scenario is obtained, proving that the MHE strategy provides accurate estimates for the OCPs initial condition. Further discussion of this results and comparison with other measurement scenarios is provided in Section 6.6.

Table 6.3: RHC/MHEF results with flow measurements and variation with respect to full flow measurements (FSM, 1TD N=10 in Table 6.2).

Episode	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
17-09-2002	0.16 (0.00%)	4.06 (-55.86%)	107.43 (0.21%)
09-10-2002	1.08 (7.52%)	340.90 (-0.24%)	101.56 (0.28%)
15-08-2006	0.25 (0.00%)	5.26 (8.21%)	100.61 (-0.10%)
30-07-2011	0.75 (0.00%)	41.55 (5.51%)	108.18 (0.04%)

On the other hand, to assess the performance of the state estimation strategy in approximating the initial conditions for the OCPs, two error indices have been defined.

First, for each sewer pipe $i = 1, \ldots, n_q$, and each SEP solved $k = 1, \ldots, t_s/5$, the maximum error in the last T estimates (that is, the values used in the OCP updating) is computed as

$$e_i^{\mathcal{O}}(k) = \max_{\tau=5\,k-T+1,\ldots,5\,k} |\hat{q}_i(\tau) - q_i^{\mathcal{O}}(\tau)| \quad \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right].$$

And secondly, the previously defined maximum error is averaged over all the solved SEPs

$$\bar{e}_i^{\mathcal{O}} = \frac{5}{t_s} \sum_{k=1}^{t_s/5} e_i^{\mathcal{O}}(k) \quad \left[\frac{\mathrm{m}^3}{\mathrm{s}}\right].$$

These error indices provide a measure for the estimation accuracy for each network sewer pipe. The histogram in Figure 6.4 shows the frequency of the values of $\bar{e}_i^{\mathcal{O}}$ for the $n_q = 145$ sewer pipes when grouped in intervals of equal length in the Riera Blanca sewer network, for the MHEF scenario and for the four simulated rain events. As with the error indices defined for model validation in Section 4.2, the highest values occur for the estimation of the flow at the collector q_{139} and its immediate upstream sewer pipe q_{138}. In Figures 6.7, 6.8 and 6.9, plots of the flows obtained as the solution of several consecutive SEPs (including the one with the highest maximum error) and the corresponding flow values to be estimated are shown. The considered sewer pipes are a middle network sewer pipe, q_{92}, showing an accurate estimation and the collector sewer pipes q_{138} and q_{139}, showing higher deviations. Taking into account that flow values at these sewer pipes reach values of 30 to 50 m^3/s, an average maximum error of 2 m^3/s means that the approximations are sufficiently accurate, as also proven by the performance results when compared to the FSM scenario and by the controller flow and volume approximations in Figures 6.5 and 6.6. The oscillations in the flow computed by the physically-based model simulator for collector q_{139} in Figure 6.5 are due to the backwater effects caused by the movement of the downstream gate. The controller flows and volume in these figures, are built using the first five minutes of each OCP solution, i.e., the time during which the solution values for the gate set-points are applied. In the following, flows and volumes constructed in this way will be referred to as RHC solutions.

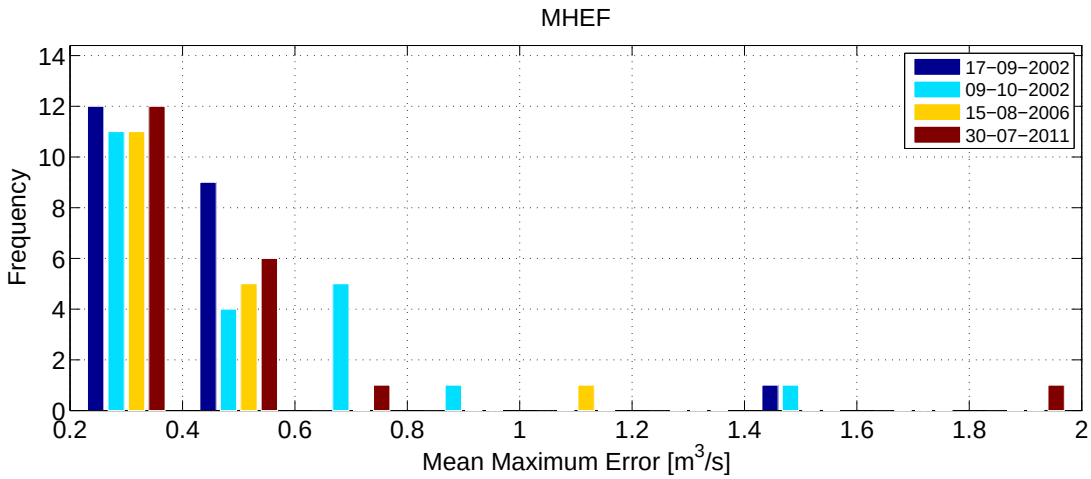

Figure 6.4: Frequency of the values of error indices $\bar{e}_i^{\mathcal{O}}$, $i = 1, \ldots, n_q$, for the MHEF scenario. Values of $\bar{e}_i^{\mathcal{O}}$ below the minimum shown in the histogram have frequencies of above 100 instances and have not been represented.

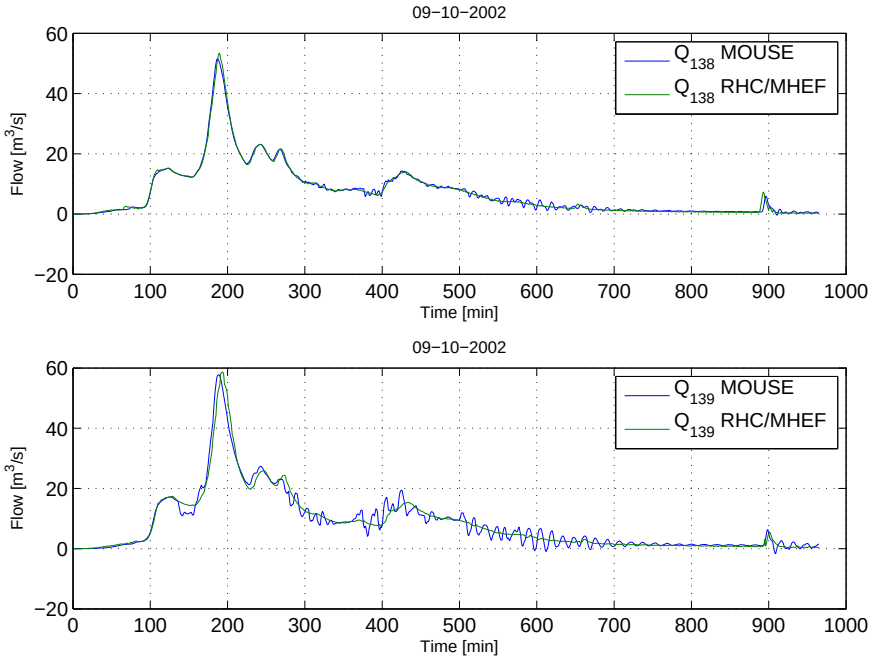

Figure 6.5: RHC prediction and MOUSE simulation values for the MHEF scenario for sewer pipes q_{138} and q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

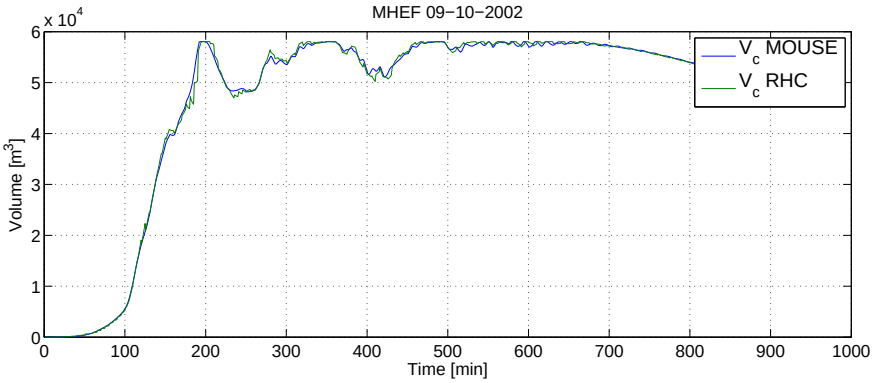

Figure 6.6: RHC prediction and MOUSE simulation values for the MHEF scenario for the volume contained in collector q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

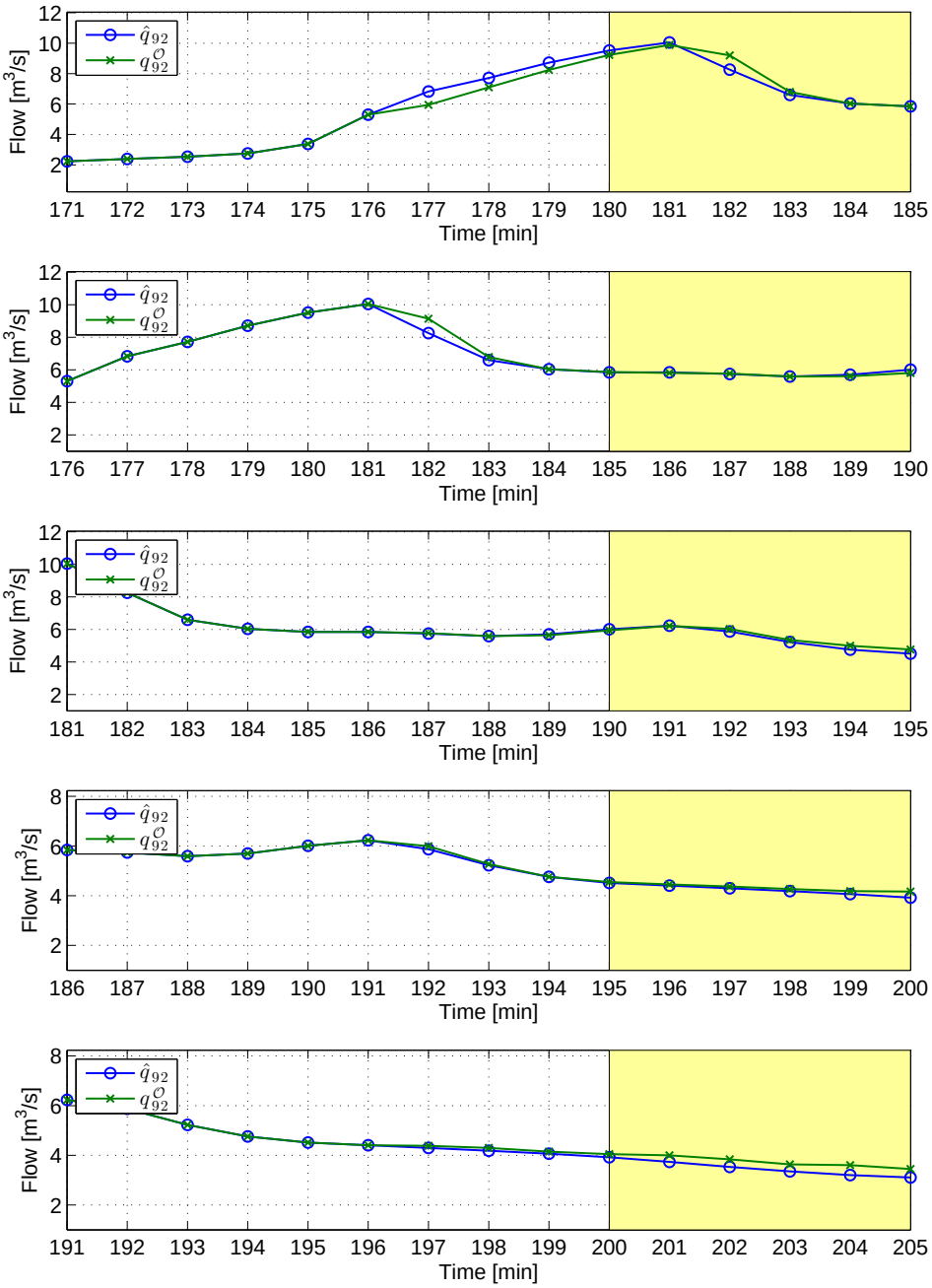

Figure 6.7: SEP solution corresponding to sewer pipe q_{92} for several consecutive problems in the MHEF scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

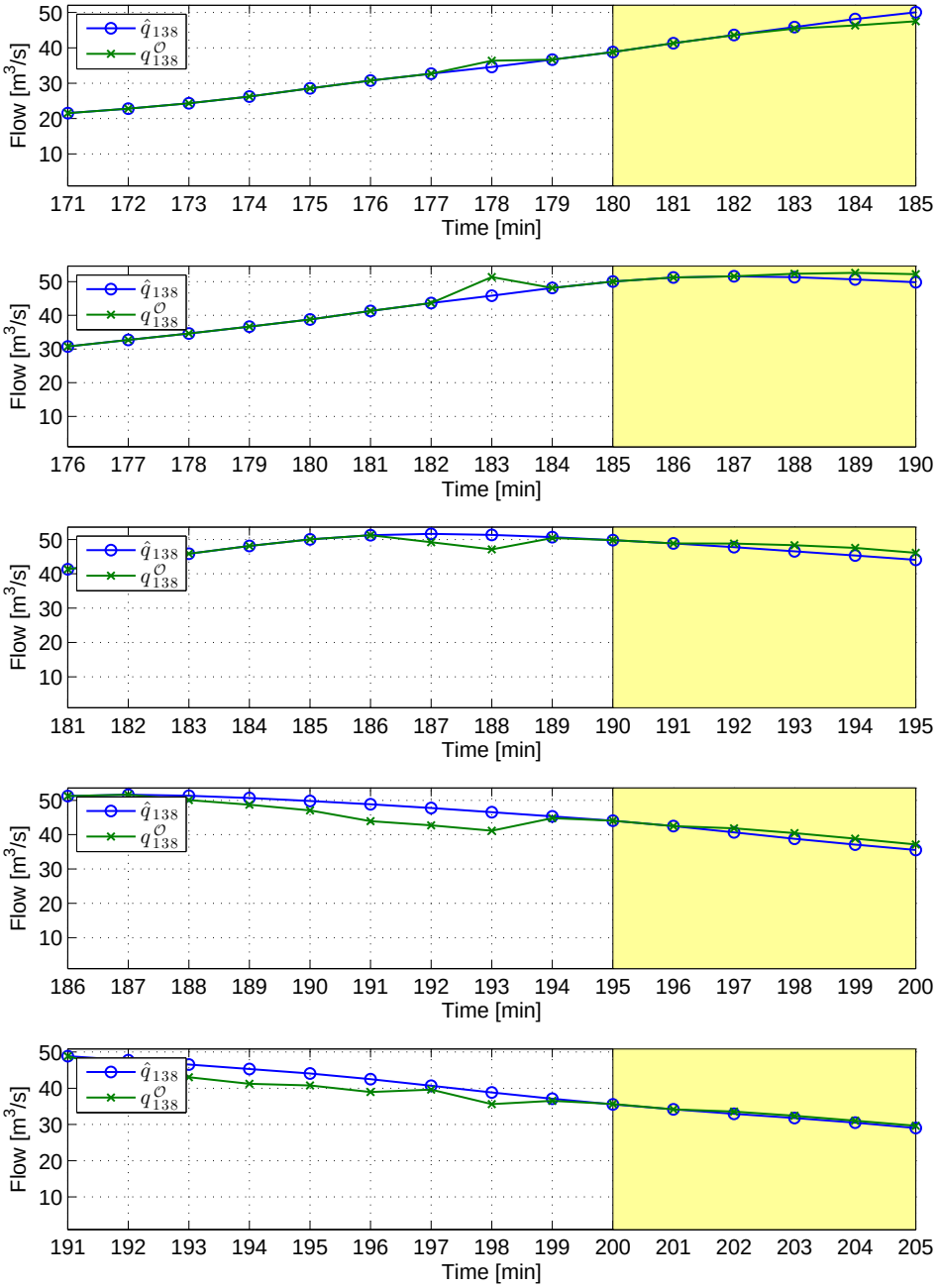

Figure 6.8: SEP solution corresponding to sewer pipe q_{138} for several consecutive problems in the MHEF scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

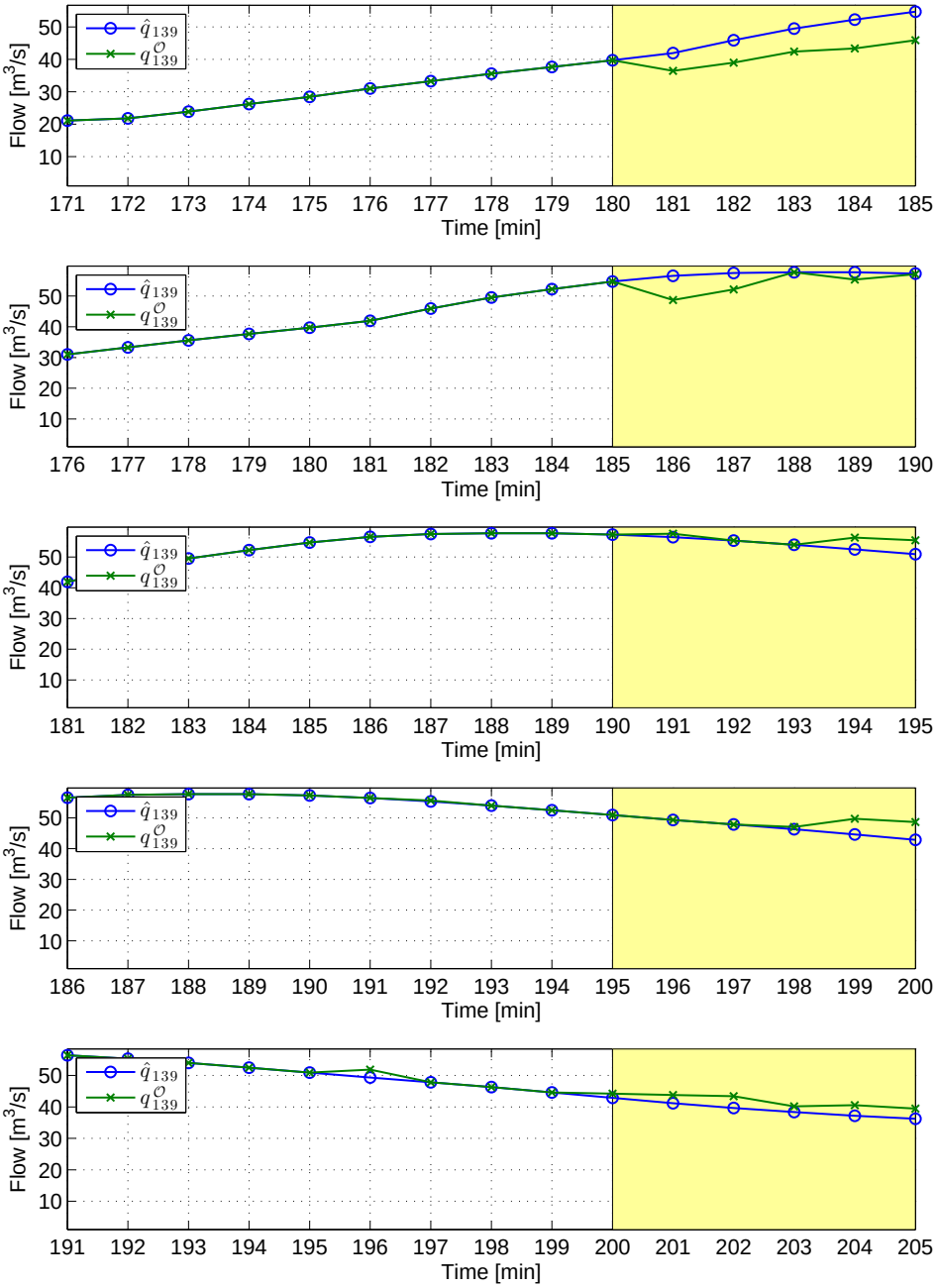

Figure 6.9: SEP solution corresponding to collector q_{139} for several consecutive problems in the MHEF scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

6.5.2 Water Level Measurements

To compute flow values from water level measurements, third-degree polynomial approximations for the flow-level relation have been used:

$$\hat{q}(t) = p_0\,\hat{h}(t)^3 + p_1\,\hat{h}(t)^2 + p_2\,\hat{h}(t) + p_3,$$

where $\hat{h}(t)$ is the measured water level and $\hat{q}(t)$ the flow approximation. Calibration of the polynomial coefficients $p_i,\ i = 0,\ 1,\ 2,\ 3$, has been performed by means of least squares fitting using data from the four available rain events. The resulting coefficients for each rain event have been averaged to obtain the final set of coefficients. The choice of the polynomial degree is based on trial and error tests, which showed that no improvement in the fitting is obtained using higher degrees.

More complex flow-level relations could be used in a similar way including, for example, the downstream gate position or other water level measurements at several locations along the collector. At this point, however, the objective of this thesis is to provide a suitable framework for output feedback control, leaving the details of flow-level relation modelling for future work.

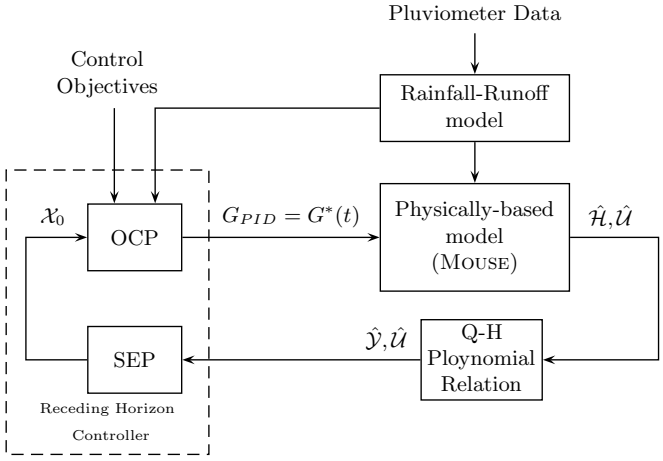

Figure 6.10: Closed-loop simulation algorithm diagram with available water level measurements, denoted $\hat{\mathcal{H}}$.

Once the flow variables have been recovered by using the flow-level approximations, the SEP and OCP are solved as in the flow measurements case, as shown in Figure 6.10

and detailed in Algorithm 3. All SEPs have been formulated and solved with an estimation horizon of $H_{\mathcal{O}} = 15$ time steps.

Polynomial approximations for flow-level relations are accurate when they are applied to sewer pipes that are not affected by backwater effects. However, for collector q_{139}, the presence of the downstream gate causes the flow-level relation to become not even one to one. This effect is even increased if the gate position changes: the loop present in the flow-level relation for calibration data with a fixed downstream gate becomes a much more complex curve in the case of a moving gate leading to poor polynomial approximations, as shown in Figures 6.11 and 6.12.

Algorithm 3: Closed-Loop Output Feedback RHC Algorithm with Water Level Measurements

Input : $\mathcal{X}_0(0) = \left(\hat{X}(0)^\top, \ldots, \hat{X}(-T+1)^\top\right)^\top = \mathbf{0}$
begin
  Set $t := 0$
  while $t \leq t_s$ **do**
    Compute rainfall-runoff prediction $R_H(t) = \left(r(t+1)^\top, \ldots, r(t+H)^\top\right)^\top$
    Compute $\mathcal{M}_3(t)$, $\mathcal{N}_3(t), b_{eq}(t), b_{ineq}(t)$ from $\mathcal{X}_0(t)$, $R_H(t)$
    Solve OCP$(t) \rightarrow \mathcal{X}^*(t) = \left(X^*(t+H)^\top, \ldots, X^*(t+1)^\top\right)^\top$
    Run MOUSE for simulation time $(t, t+5)$ with gate PID set-points $G_{PID} = G^*(t)$
    Extract MOUSE data from result files: $\hat{G}(t+5)$,…,$\hat{G}(t)$, $\hat{H}(t+5)$,…,$\hat{H}(t)$
    Compute flows from water level measurements: $\hat{\mathcal{Y}}(t) = P(\hat{\mathcal{H}}(t))$
    Solve SEP$(t+5) \rightarrow \mathcal{X}^*_{\mathcal{O}}(t+5) = \left(X^*_{\mathcal{O}}(t+5)^\top, \ldots, X^*_{\mathcal{O}}(t+5-H_{\mathcal{O}}+1)^\top\right)^\top$
    Set $\mathcal{X}_0(t+5) := \left(X^*_{\mathcal{O}}(t+5)^\top, \ldots, X^*_{\mathcal{O}}(t+5-T+1)^\top\right)^\top$
    Set $t := t+5$
  end
end

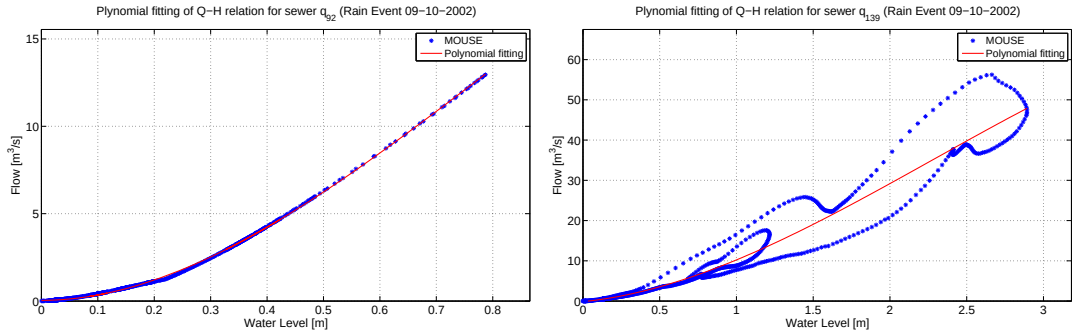

Figure 6.11: Calibration data and polynomial fitting for sewer pipes q_{92} and q_{139} for rain scenario 09-10-2002. Similar figures corresponding to the other calibration events can be found in Appendix C.

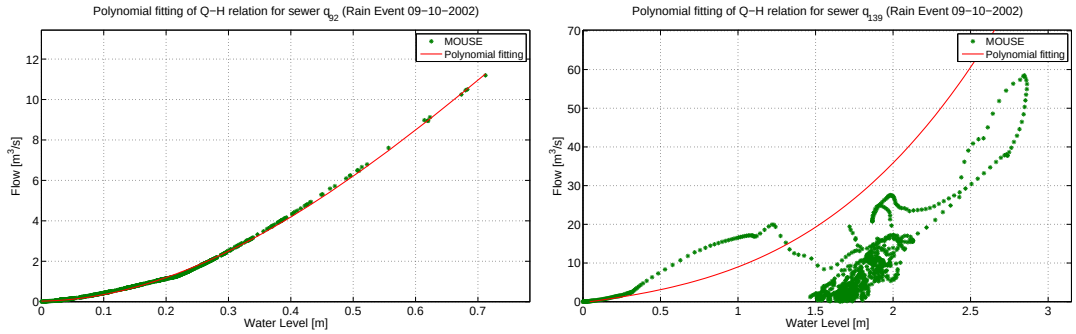

Figure 6.12: Closed-loop simulation data and polynomial approximation for sewer pipes q_{92} and q_{139} for rain scenario 09-10-2002. Similar figures corresponding to the other calibration events can be found in Appendix C.

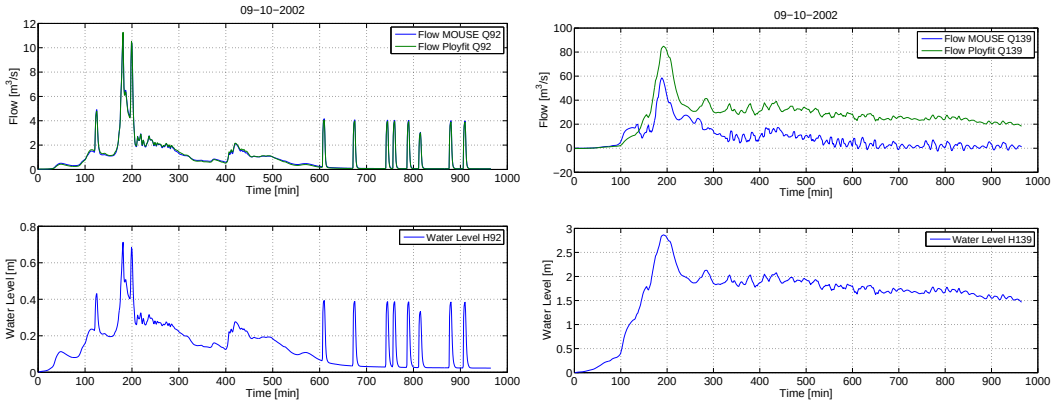

Figure 6.13: Flow estimation during closed-loop simulations for sewer pipes q_{92} and q_{139} for rain scenario 09-10-2002. Similar figures corresponding to the other calibration events can be found in Appendix C.

Figure 6.13 shows the water levels computed by the physically-based model during the closed-loop simulation corresponding to episode 09-10-2002 and the flow values obtained after applying the polynomial transformation for sewer pipe q_{92} and collector q_{139}. It can be noticed that the inflows to the collector are heavily overestimated by the flow level relations. This cause the SEPs to provide poor approximations of the system state, as shown in Figures 6.17–6.19 and by instances of mean maximum error above 10 m^3/s in the error histogram of Figure 6.14.

In Figure 6.15 it can be noticed that the RHC solution for the collector inflow q_{139} is partially corrected due to more accurate approximations in the upstream sewer pipe q_{138}. Each peak in the RHC flow for collector q_{139} corresponds to the first predicted flow in each of the OCPs. The first predicted flow in each OCP solution has a strong influence from the overestimated initial conditions provided by the polynomial approximations. In the following time steps the influence of the better approximated initial conditions and flows in sewer pipe q_{138} contribute to improving the prediction for q_{139} through the delayed transport equations. At the next OCP the initial condition is again overestimated causing another peak and giving the RHC solution an oscillating shape (recall that the RHC solution is built by using the first five time steps of each OCP solution). On the other hand, overestimations of the collector volume initial condition has an effect along the whole prediction horizon, as shown in the collector volume RHC solution in Figure 6.16. This overestimation in the collector volume leads to a conservative management of the collector downstream gate, since in the control model the collector is becoming full very fast and also must empty very fast to fulfill the maximum volume constraint. Therefore, as detailed in Table 6.4, higher CSO values are obtained when using flow-level relations to estimate the collector inflows. As mentioned before, further discussion of these results and comparison with the other measurement scenarios is provided in Section 6.6.

Table 6.4: RHC/MHEL results with water level measurements and variation with respect to full flow measurements (FSM, 1TD N=10 in Table 6.2).

Episode	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
17-09-2002	0.16 (0.00%)	32.61 (254.16%)	106.06 (-1.07%)
09-10-2002	1.01 (0.48%)	364.10 (6.55%)	100.81 (-0.45%)
15-08-2006	0.25 (0.00%)	11.04 (127.01%)	99.57 (-1.13%)
30-07-2011	0.75 (0.00%)	67.49 (71.38%)	107.27 (-0.80%)

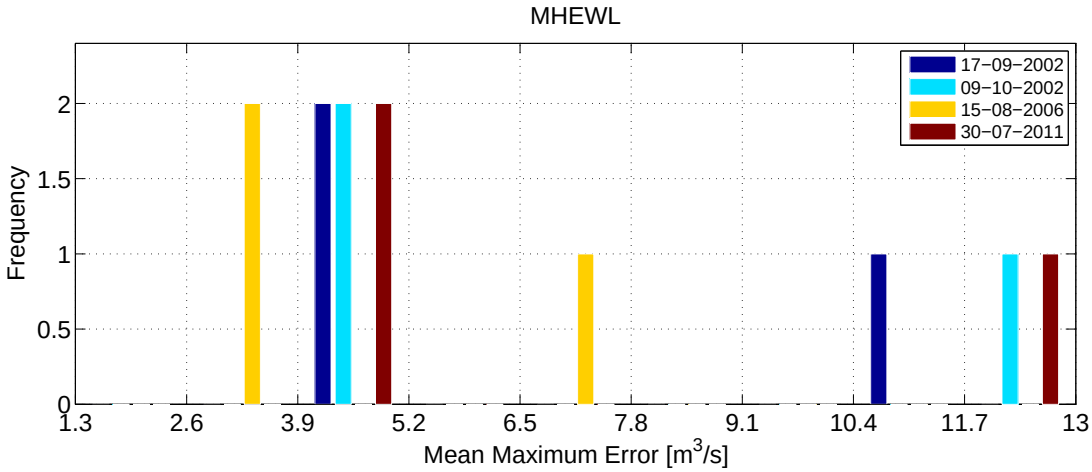

Figure 6.14: Frequency of the values of error indices $\bar{e}_i^{\mathcal{O}}$, $i = 1, \ldots, n_q$, for the MHEL scenario. Values of $\bar{e}_i^{\mathcal{O}}$ below the minimum shown in the histogram have frequencies of above 100 instances and have not been represented.

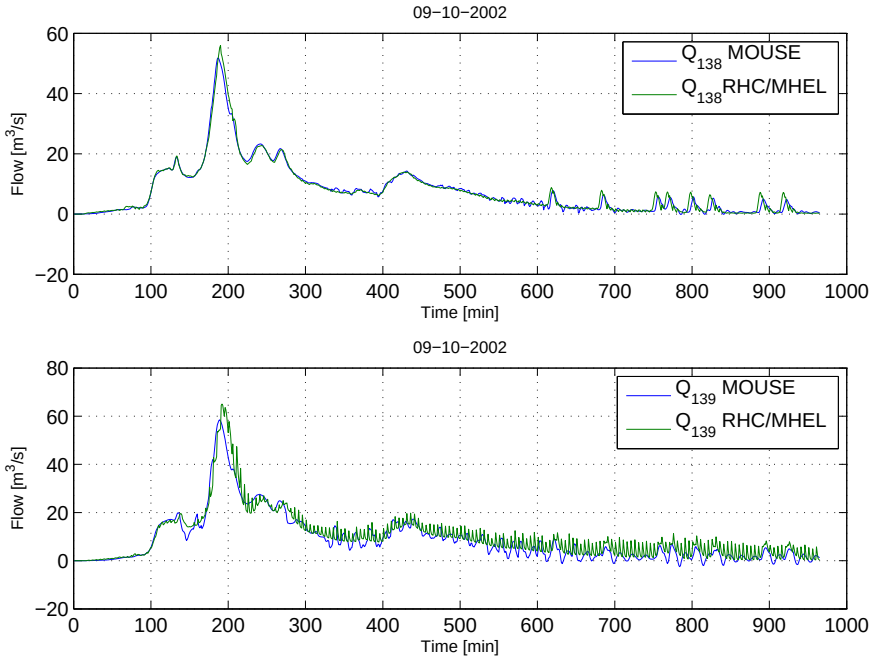

Figure 6.15: RHC prediction and MOUSE simulation values for the MHEL scenario for sewer pipes q_{138} and q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

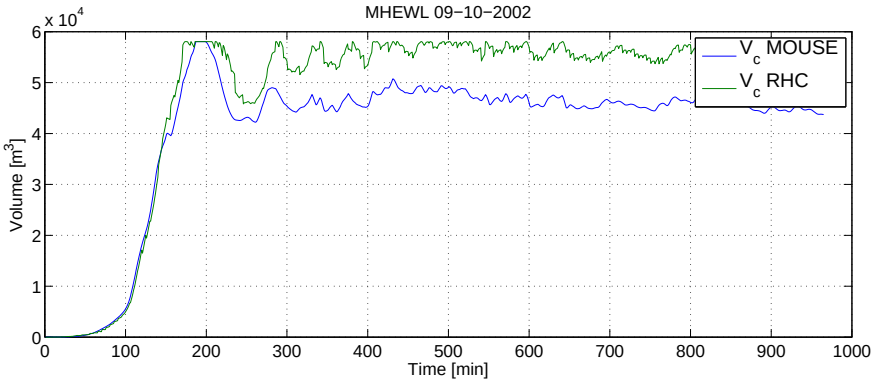

Figure 6.16: RHC prediction and MOUSE simulation values for the MHEL scenario for the volume contained in collector q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

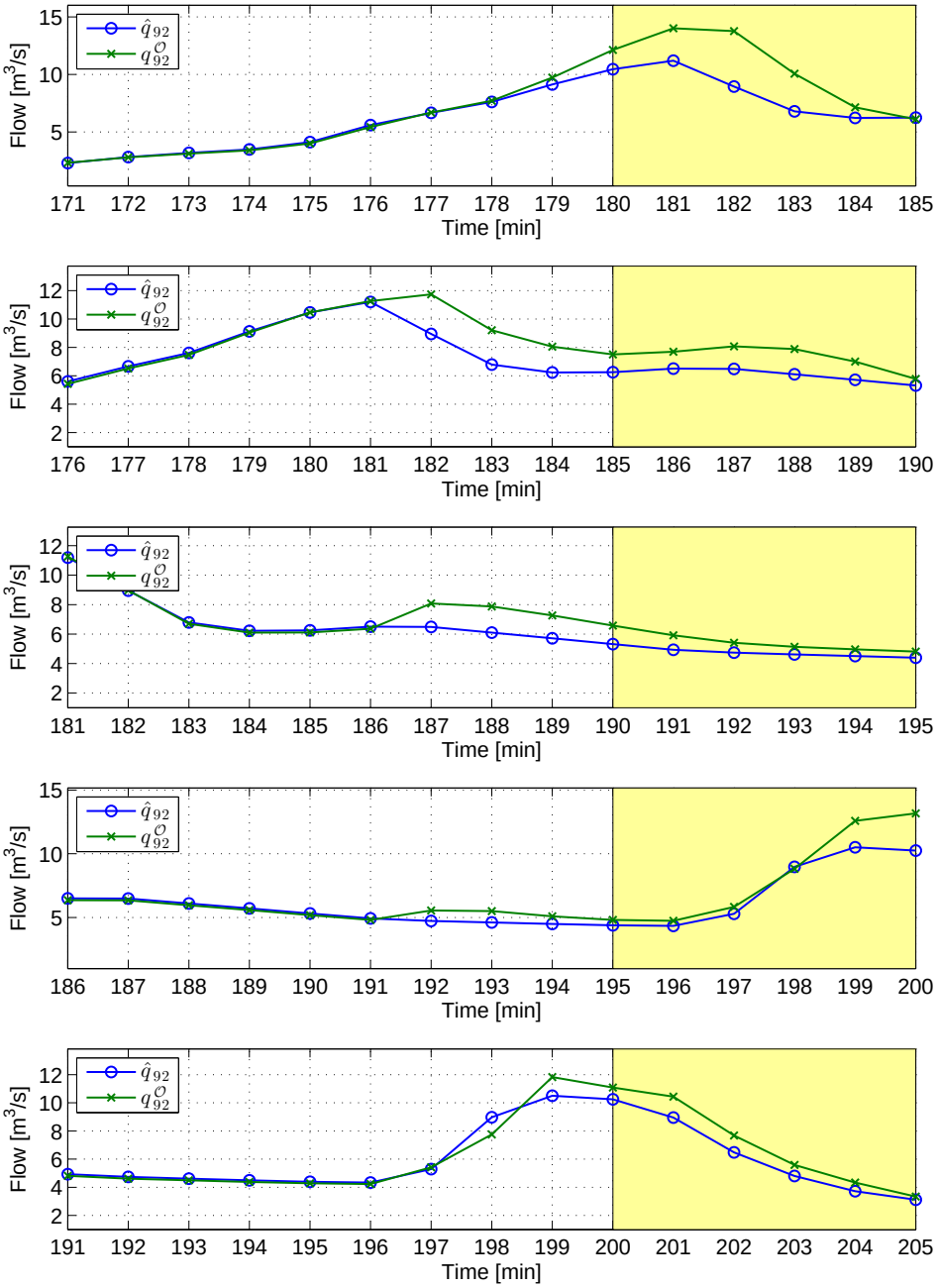

Figure 6.17: SEP solution corresponding to sewer pipe q_{92} for several consecutive problems in the MHEL scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

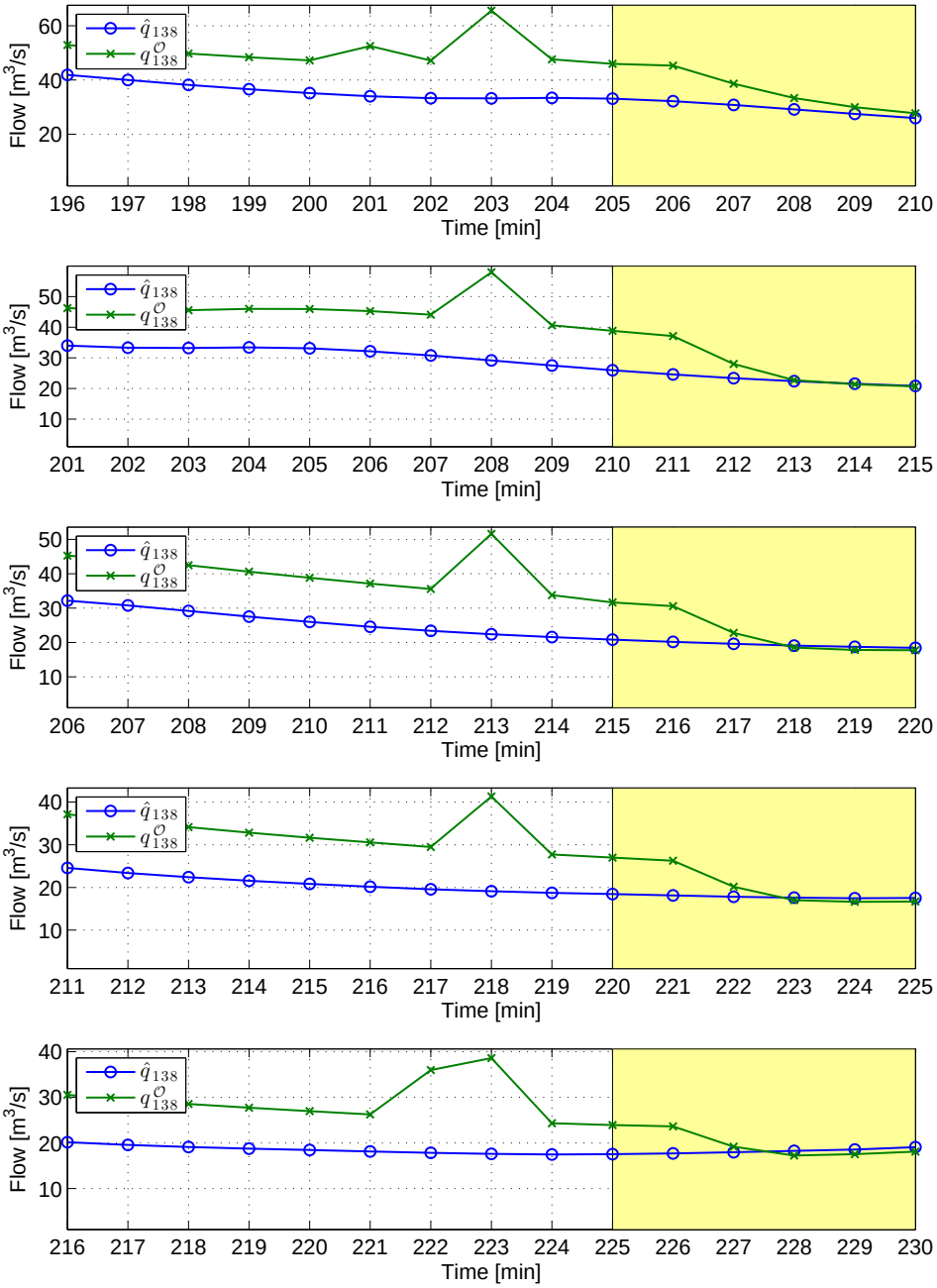

Figure 6.18: SEP solution corresponding to sewer pipe q_{138} for several consecutive problems in the MHEL scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

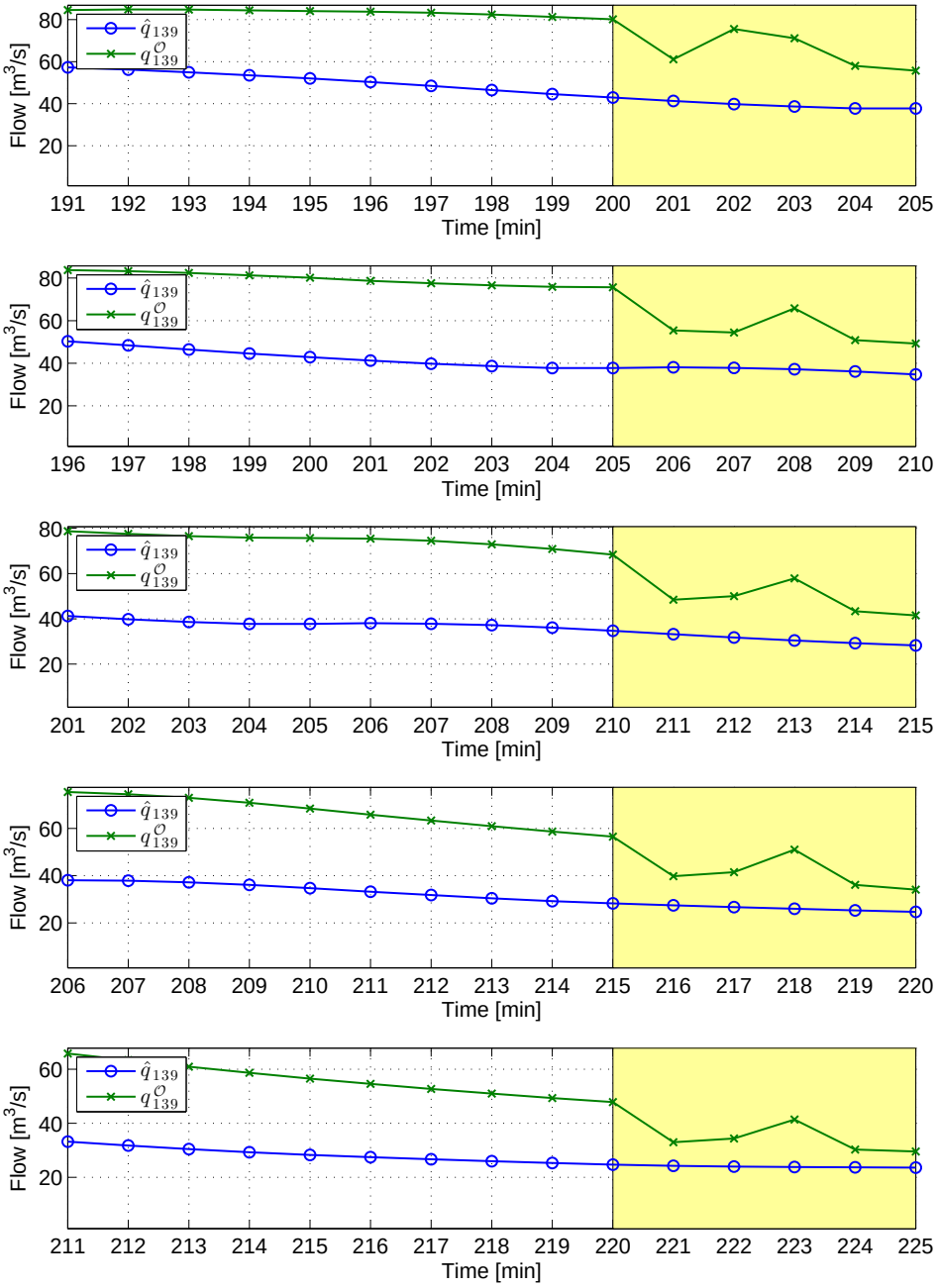

Figure 6.19: SEP solution corresponding to collector q_{139} for several consecutive problems in the MHEL scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

6.5.3 Water Level Plus Collector Inflow Measurements

To solve the SEP using both flow and water level measurements, one simply constructs the measured variables vector $\hat{\mathcal{Y}}$ by either using the direct flow measurements or the flows obtained from water level measurements and the flow-level relation described in the previous section, depending on the available measurement. Then, the SEP is solved as in the other cases, with an estimation horizon of $H_{\mathcal{O}} = 15$ time steps, and the whole closed-loop simulation algorithm continues in the same way.

As mentioned before, two cases for the mixed flow and level measurement scenario have been considered: first, only the inflow to the collector is measured and second, inflows to both the collector and its immediate upstream sewer pipe q_{138} are measured. In the former case, as shown in the histogram in Figure 6.20 and in Figures 6.24–6.26, it can be noticed that the collector inflow is estimated with more accuracy than in the MHEL scenario. However, these estimations still suffer from the influence of the poor approximations provided by the flow-level relation when backwater also affects the upstream sewer pipe q_{138}, since the flow at this sewer pipe is routed downstream by the model affecting the flow at the collector. As a consequence, although the overestimation of the controller predictions of the collector inflow and volume is reduced with respect to the MHEL scenario, it is still noticeable, as shown in the RHC solutions in Figures 6.22 and 6.23. Therefore, in comparison with the two previous approaches, measuring the collector inflow provides an improvement in the CSO objective with respect to the MHEL case but results are still far from the ideal MHEF case, as detailed in Table 6.5.

Table 6.5: RHC/MHEC results with water level and collector inflow measurements and variation with respect to full flow measurements (FSM, 1TD N=10 in Table 6.2).

Episode	Overflow [$\times 10^3 m^3$]	CSO [$\times 10^3 m^3$]	WWTP [$\times 10^3 m^3$]
17-09-2002	0.16 (0.00%)	17.06 (85.30%)	106.02 (-1.11%)
09-10-2002	1.03 (2.26%)	354.83 (3.83%)	100.90 (-0.37%)
15-08-2006	0.25 (0.00%)	6.16 (26.71%)	99.58 (-1.12%)
30-07-2011	0.75 (0.00%)	56.36 (43.11%)	107.14 (-0.93%)

The negative influence of the poor flow-level approximations at sewer pipe q_{138} is completely removed by measuring flow at this point. It can be seen in Table 6.6, that

the results obtained in this case virtually recover the performance obtained by the MHEF scenario. The error histogram in Figure 6.21 shows that the maximum errors in this case are half than in the single collector flow measurement situation. Figures 6.27 and 6.28 show that the controller predictions of the collector inflow and volume are as accurate as in the MHEF case. Plots of the convergence of the estimator to the desired values are provided in Figures 6.29, 6.30 and 6.31.

Table 6.6: RHC/MHEC2 results with water level and two collector inflow measurements and variation with respect to full-flow measurements (FSM, 1TD N=10 in Table 6.2).

Episode	Overflow $[\times 10^3 \mathrm{m}^3]$	CSO $[\times 10^3 \mathrm{m}^3]$	WWTP $[\times 10^3 \mathrm{m}^3]$
17-09-2002	0.16 (0.00%)	3.78 (-58.95%)	106.46 (-0.69%)
09-10-2002	1.01 (2.26%)	333.63 (-2.37%)	101.12 (-0.15%)
15-08-2006	0.25 (0.00%)	5.74 (17.93%)	99.91 (-0.79%)
30-07-2011	0.75 (0.00%)	40.13 (1.90%)	107.56 (-0.54%)

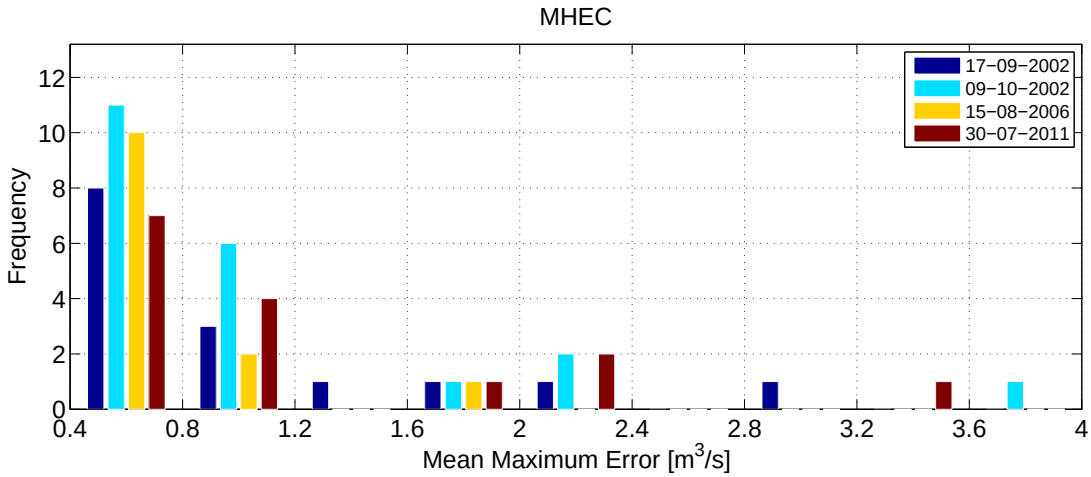

Figure 6.20: Frequency of the values of error indices $\bar{e}_i^{\mathcal{O}}$, $i = 1, \ldots, n_q$, for the MHEC scenario. Values of $\bar{e}_i^{\mathcal{O}}$ below the minimum shown in the histogram have frequencies of above 100 instances and have not been represented.

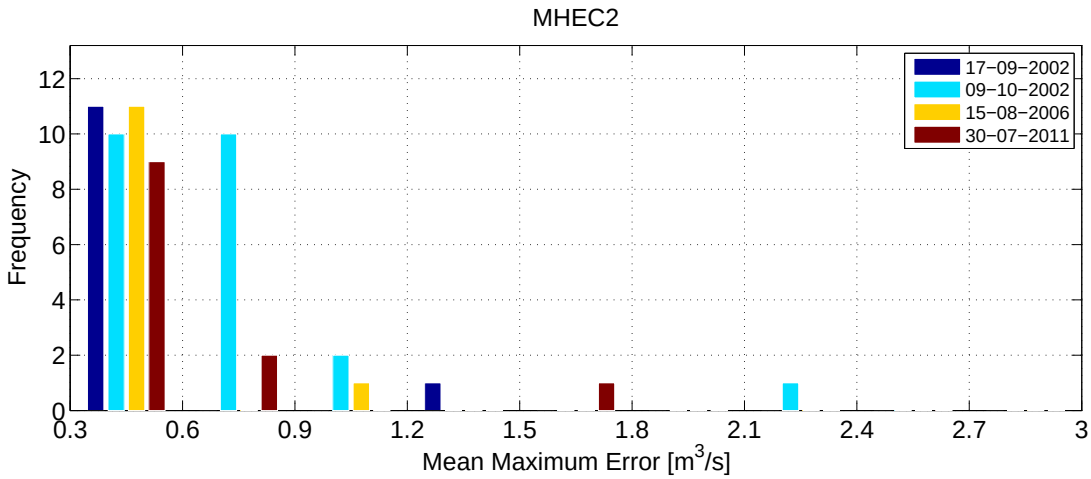

Figure 6.21: Frequency of the values of error indices $\bar{e}_i^{\mathcal{O}},\ i = 1, \ldots, n_q$, for the MHEC2 scenario. Values of $\bar{e}_i^{\mathcal{O}}$ below the minimum shown in the histogram have frequencies of above 100 instances and have not been represented.

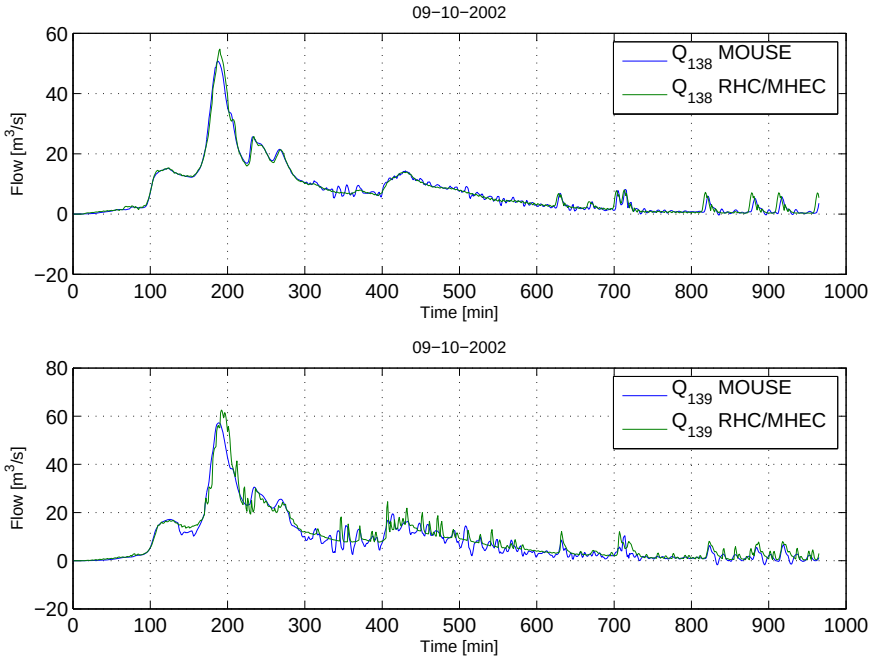

Figure 6.22: RHC prediction and MOUSE simulation values for the MHEC scenario for sewer pipes q_{138} and q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

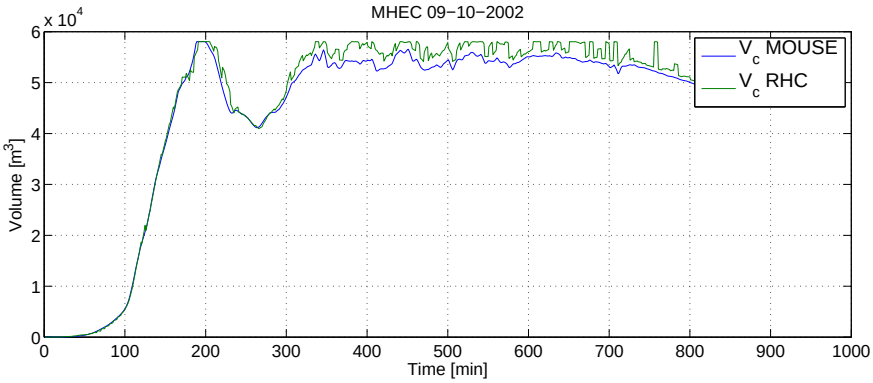

Figure 6.23: RHC prediction and MOUSE simulation values for the MHEC scenario for the volume contained in collector q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

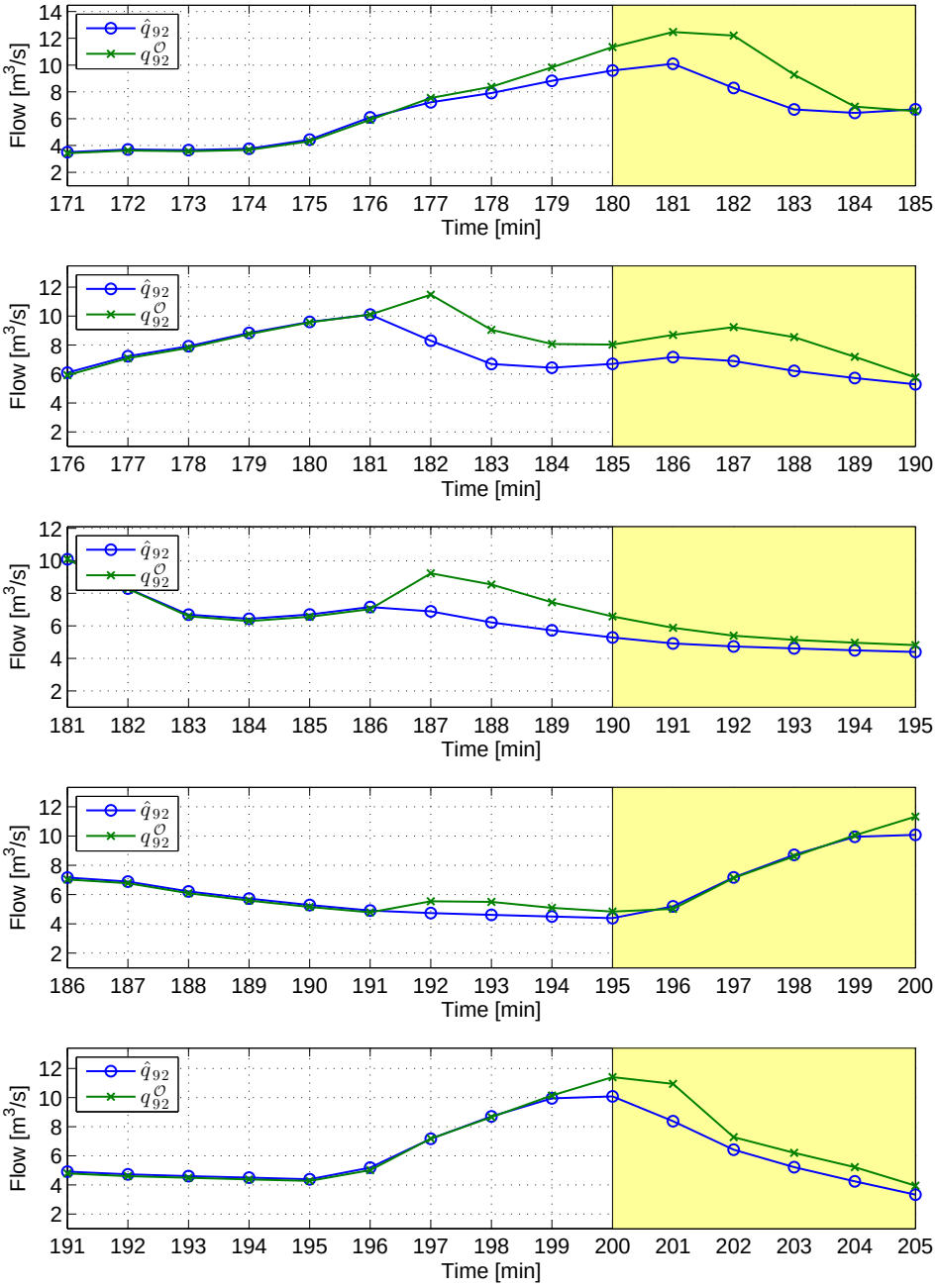

Figure 6.24: SEP solution corresponding to sewer pipe q_{92} for several consecutive problems in the MHEC scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

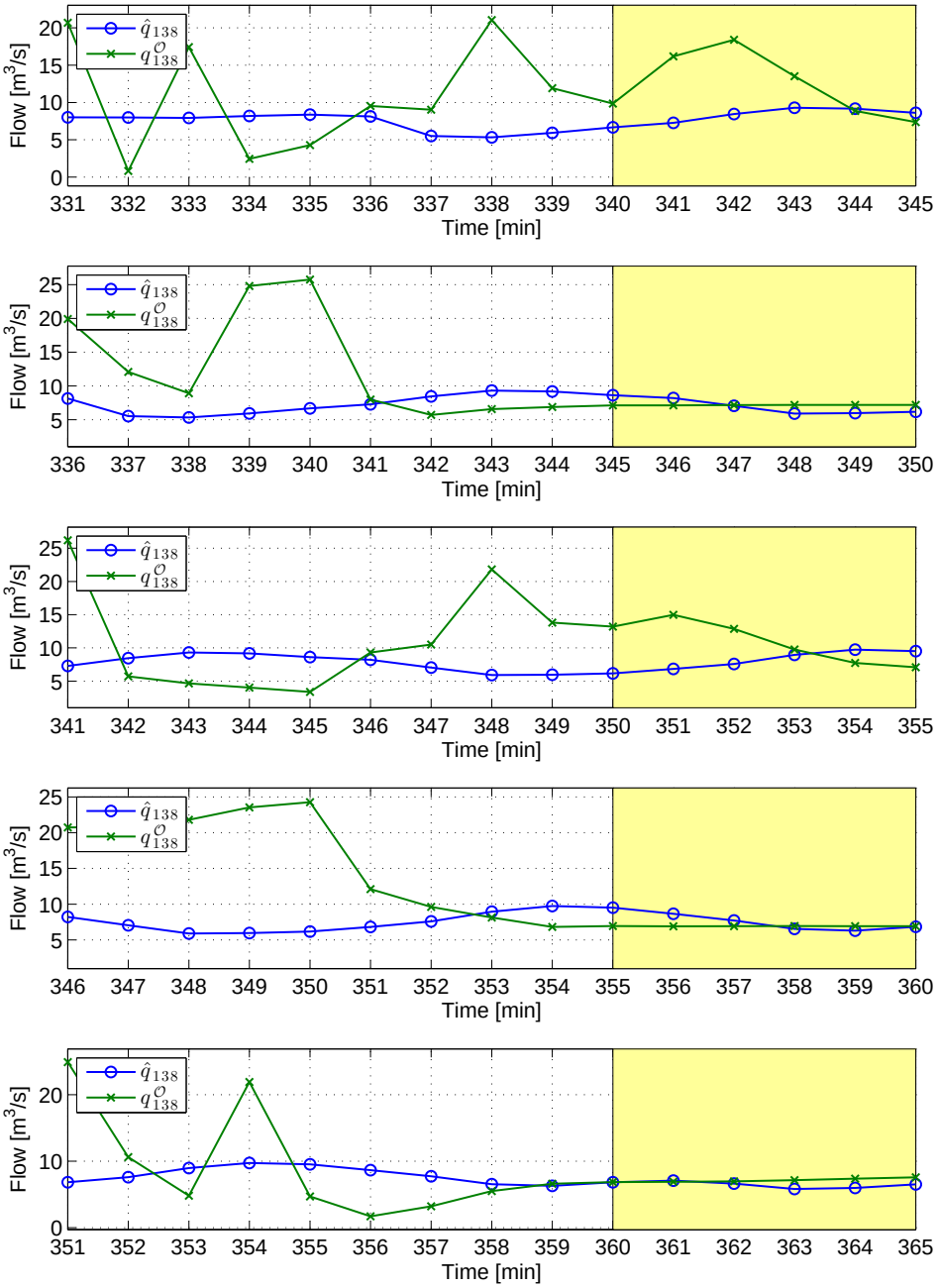

Figure 6.25: SEP solution corresponding to sewer pipe q_{138} for several consecutive problems in the MHEC scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

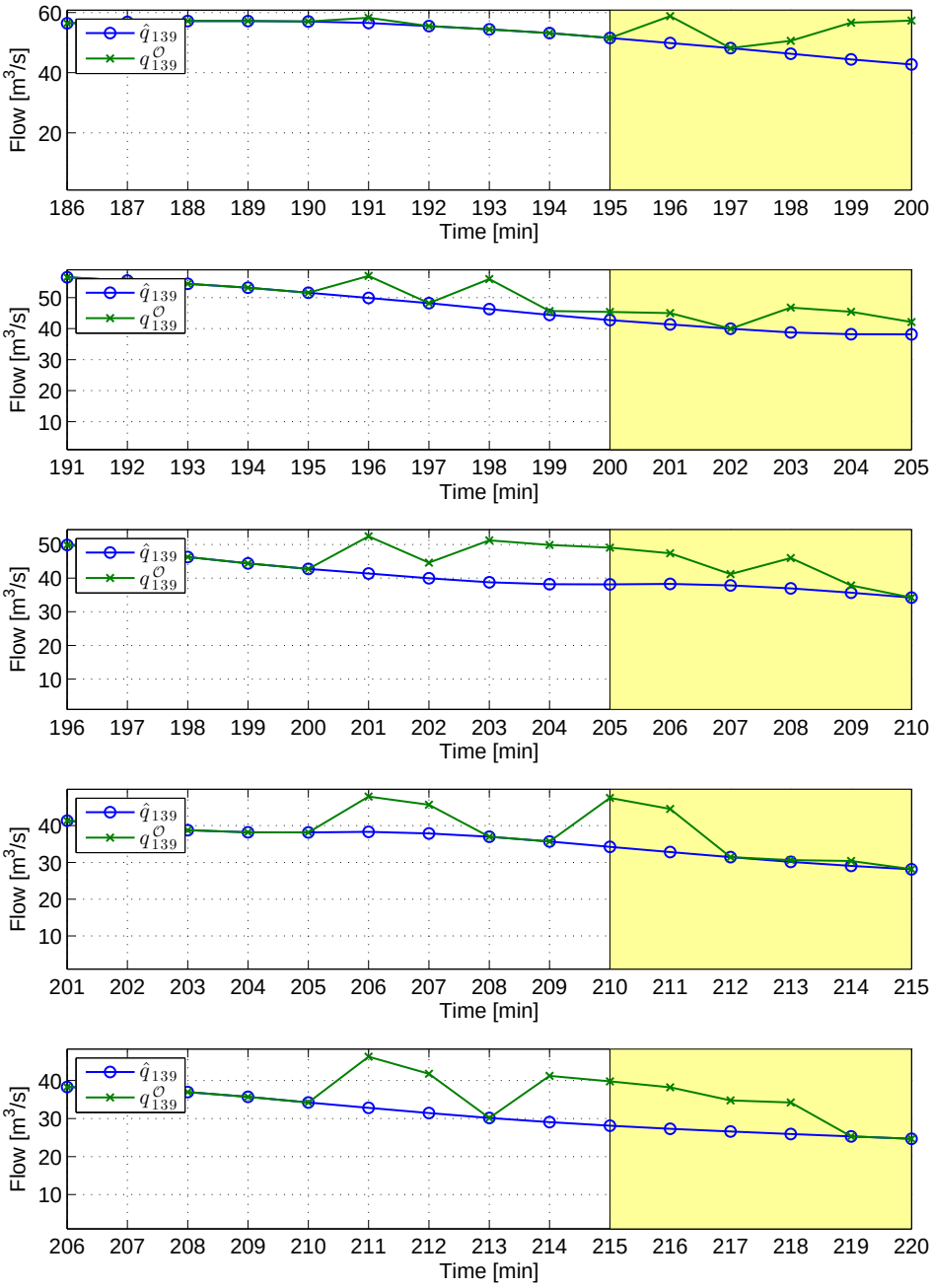

Figure 6.26: SEP solution corresponding to collector q_{139} for several consecutive problems in the MHEC scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

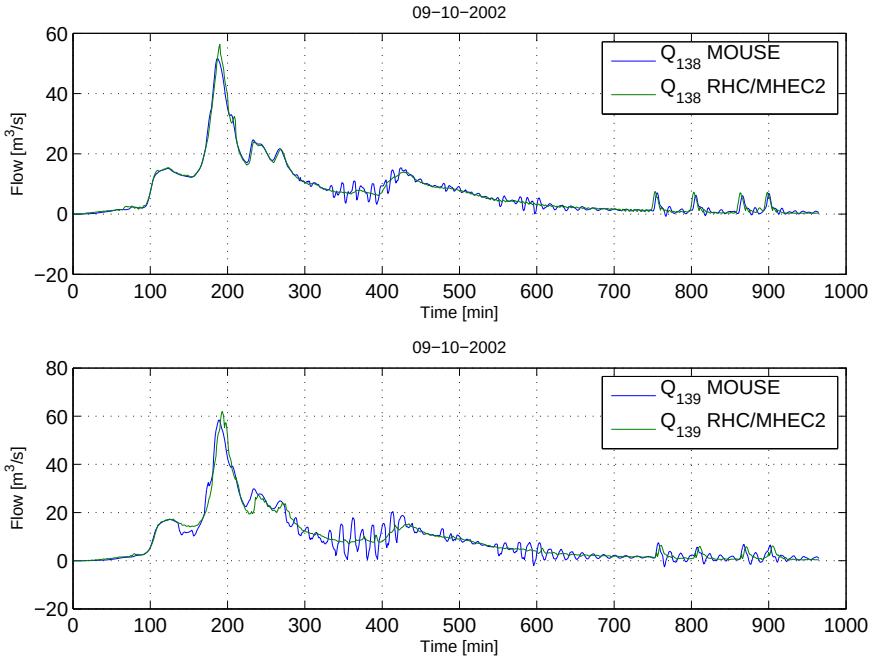

Figure 6.27: RHC prediction and MOUSE simulation values for the MHEC2 scenario for sewer pipes q_{138} and q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

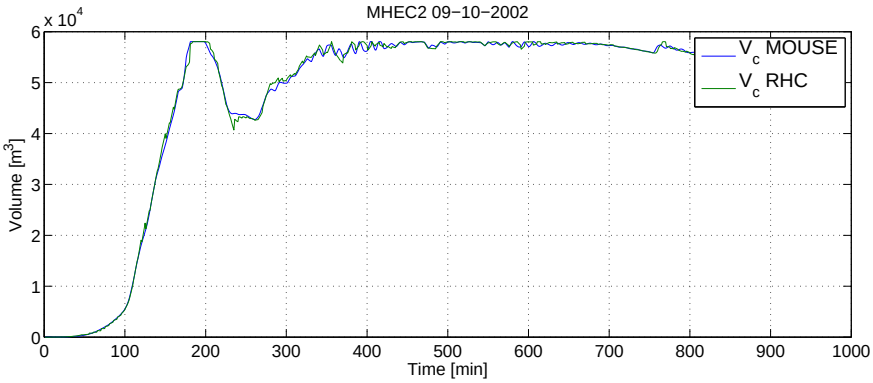

Figure 6.28: RHC prediction and MOUSE simulation values for the MHEC2 scenario for the volume contained in collector q_{139} corresponding to rain event 09-10-2002. Plots for the rest of rain events can be found in Appendix B.

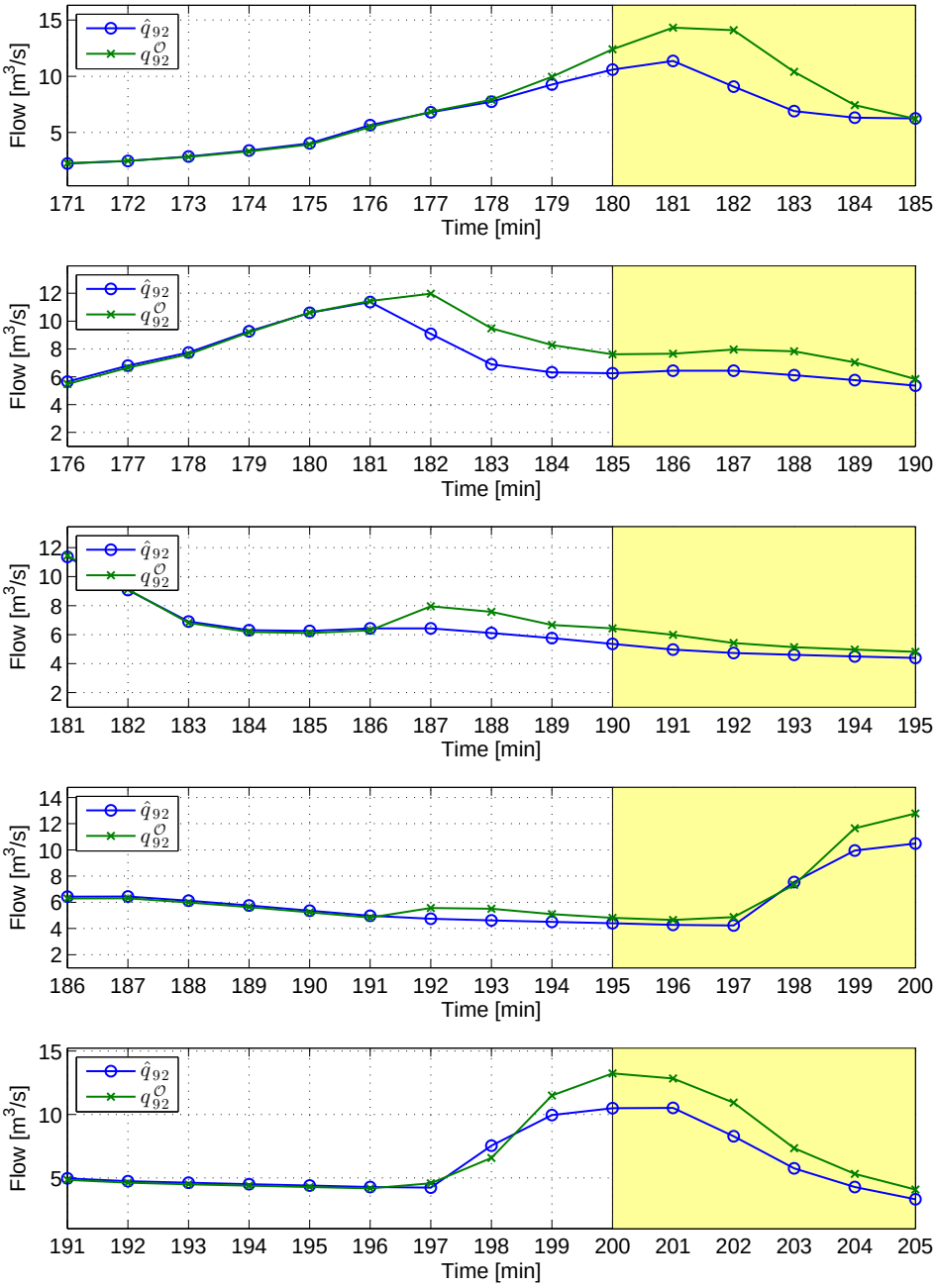

Figure 6.29: SEP solution corresponding to sewer pipe q_{92} for several consecutive problems in the MHEC2 scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

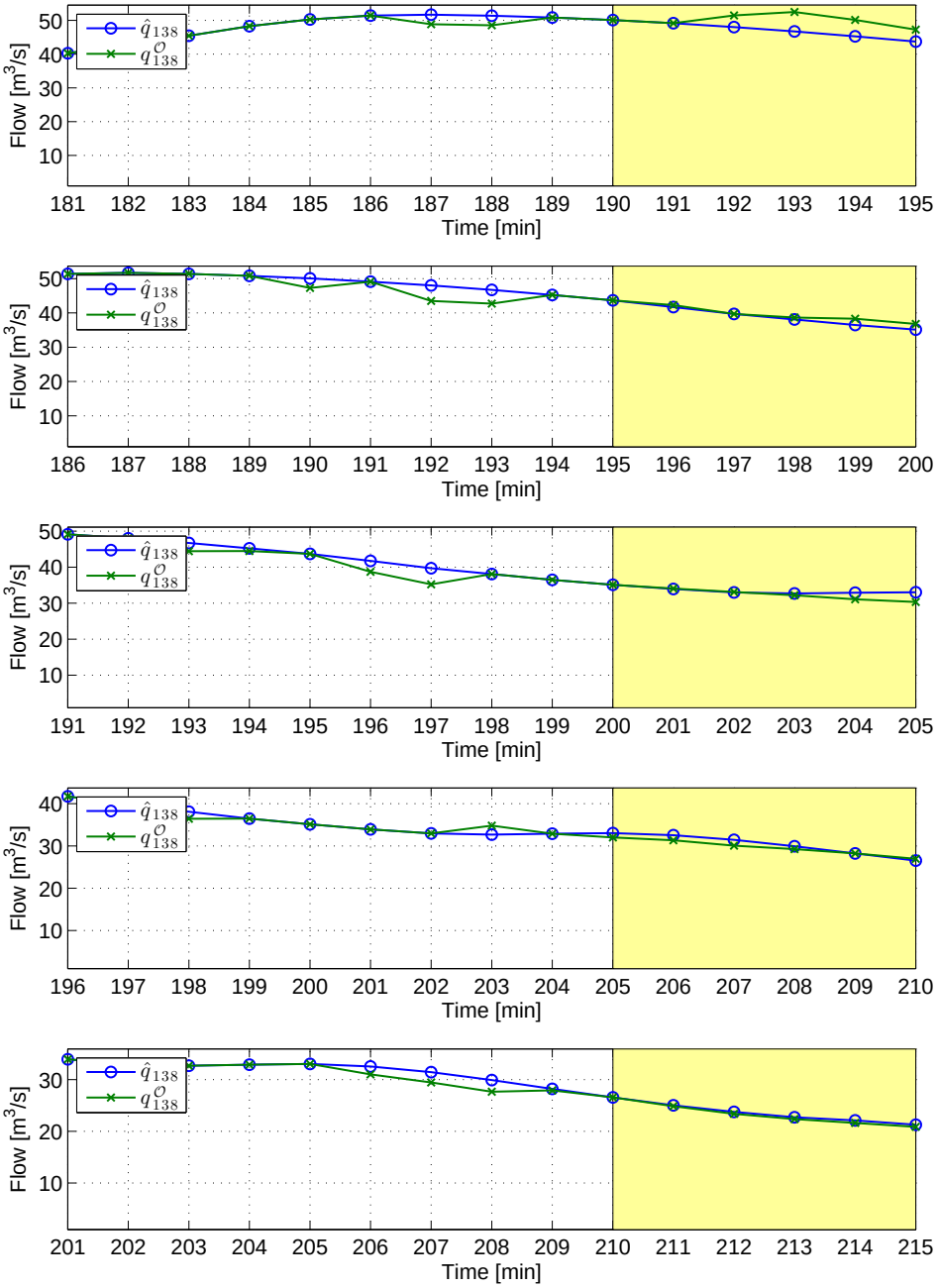

Figure 6.30: SEP solution corresponding to sewer pipe q_{138} for several consecutive problems in the MHEC2 scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

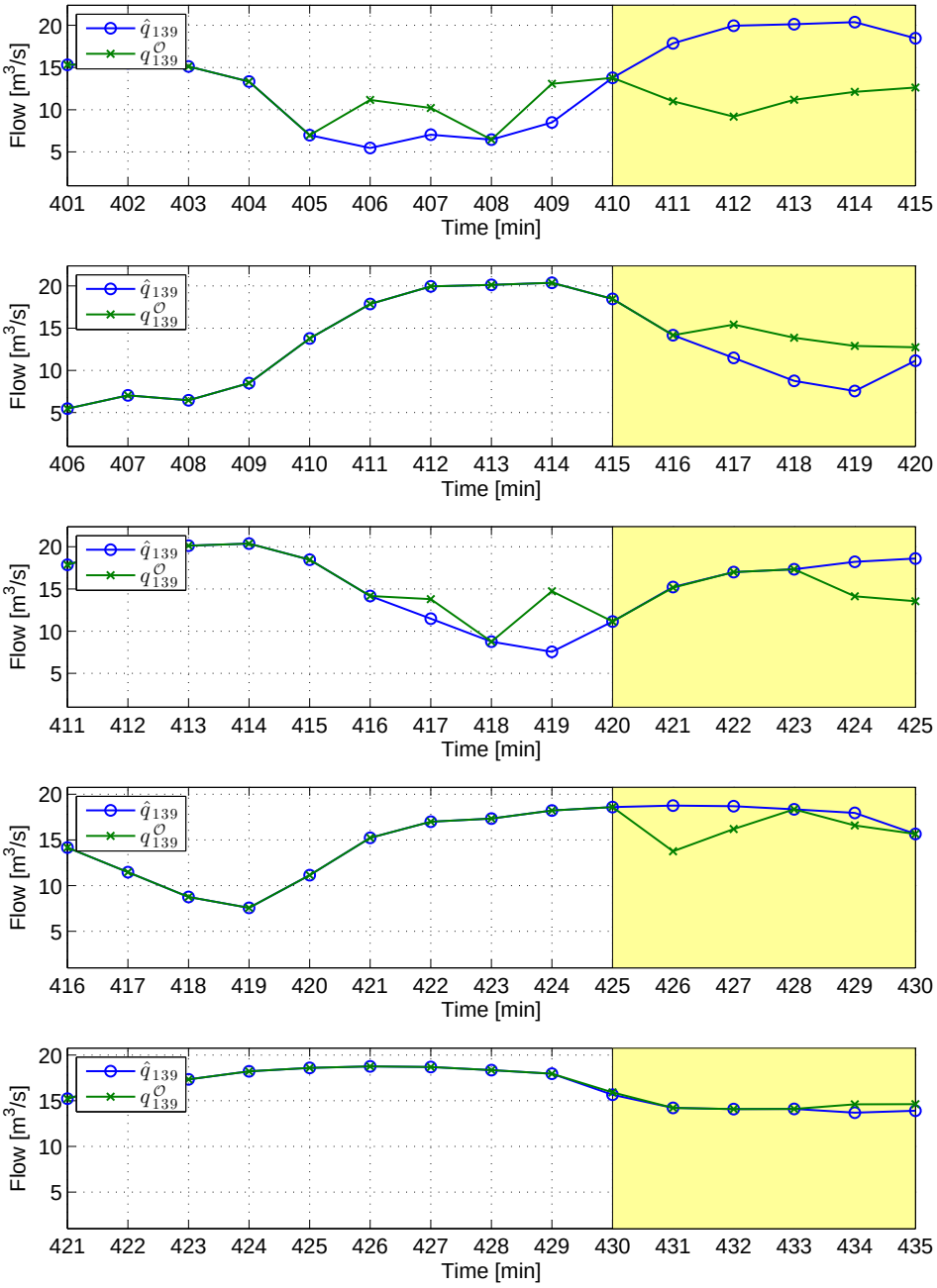

Figure 6.31: SEP solution corresponding to collector q_{139} for several consecutive problems in the MHEC2 scenario for rain event 09-10-2002. The first one corresponds to the maximum absolute error obtained among all the solved SEPs. SEP solution plots for the rest of the rain events can be found in Appendix D.

6.6 Results Discussion

For ease of comparison, the results of each of the MHE scenarios shown in Tables 6.3, 6.4 and 6.5 together with the results of the FSM (1TD N=10) scenario shown in Table 6.2 are collected in Table 6.7. It can be noticed that minimal variation of the overflow and WWTP objectives is obtained with the different measurement approaches. Overflows occur in the upper to middle part of the network and are avoided by redirecting part of the flow to the detention tanks. The presence of measurements at the sewer pipes upstream of the gates redirecting flow to the tanks and the accurate approximations by means of the flow-level relation at those locations guarantee a proper management of the tanks and an optimal mitigation of urban overflows. Gates g_8 and g_9 in Figures 6.32–6.35 correspond the RHC computed set-points and MOUSE inflows to the detention tanks. It can be noticed that in some instances backwater occurs at gate g_8 eventually leading to reversed flow (negative flow), which is currently not taken into account in the HLD model. However, even in those cases proper use of the detention tank to minimize overflows is achieved.

Table 6.7: RHE/MHE results and comparison with state feedback (FSM).

Episode	Measurements	Overflow $[\times 10^3 \mathrm{m}^3]$	CSO $[\times 10^3 \mathrm{m}^3]$	WWTP $[\times 10^3 \mathrm{m}^3]$
17-09-2002	FSM	0.16	9.21	107.20
	MHEF	0.16 (0.00%)	4.06 (-55.86%)	107.43 (0.21%)
	MHEL	0.16 (0.00%)	32.61 (254.16%)	106.06 (-1.07%)
	MHEC	0.16 (0.00%)	17.06 (85.30%)	106.02 (-1.11%)
	MHEC2	0.16 (0.00%)	3.78 (-58.95%)	106.46 (-0.69%)
09-10-2002	FSM	1.01	341.74	101.27
	MHEF	1.08 (7.52%)	340.90 (-0.24%)	101.56 (0.28%)
	MHEL	1.01 (0.48%)	364.10 (6.55%)	100.81 (-0.45%)
	MHEC	1.03 (2.26%)	354.83 (3.83%)	100.90 (-0.37%)
	MHEC2	1.01 (0.02%)	333.63 (-2.37%)	101.12 (-0.15%)
15-08-2006	FSM	0.25	4.87	100.71
	MHEF	0.25 (0.00%)	5.26 (8.21%)	100.61 (-0.10%)
	MHEL	0.25 (0.00%)	11.04 (127.01%)	99.57 (-1.13%)
	MHEC	0.25 (0.00%)	6.16 (26.71%)	99.58 (-1.12%)
	MHEC2	0.25 (0.00%)	5.74 (17.93%)	99.91 (-0.79%)
30-07-2011	FSM	0.75	39.38	108.14
	MHEF	0.75 (0.00%)	41.55 (5.51%)	108.18 (0.04%)
	MHEL	0.75 (0.00%)	67.49 (71.38%)	107.27 (-0.80%)
	MHEC	0.75 (0.00%)	56.36 (43.11%)	107.14 (-0.93%)
	MHEC2	0.75 (0.00%)	40.13 (1.90%)	107.56 (-0.54%)

Regarding the WWTP objective, results are quite similar in all measurement scenarios since in all cases the plant receives its maximum inflow all the time since soon after the start of the rain event. This behaviour is illustrated by gate g_{10}, the one redirecting flow to the WWTP, in Figures 6.32–6.35.

The most noticeable variations that can be observed in Table 6.7 are regarding the CSO values. The fulfillment of this objective is closely related to the proper use of the in-line detention capacity of the collector, which is in turn related to the accuracy of the flow approximations at its inflow and upstream sewer pipe. The estimated collector inflows provided by the SEPs have a direct effect on the model values for the collector volume in the first time steps of the prediction horizon, since due to the delay added in the model by means of additional tanks, those values are directly used to compute the collector volume. On the other hand, the estimated flow values upstream of the collector are used through the transport equation (3.8) to compute the inflow to the collector in the first prediction time steps and the collector volume at further time steps.

As mentioned in previous sections, flow values obtained from level measurements turn into important overestimation of the collector inflows due to backwater effects. This inflow overestimation leads also to an overestimation of the collector volume causing its in-line storage capacity not to be fully used. Measuring the flow at the collector inflow provides suitable initial conditions for the collector volume thus improving its in-line capacity usage and leading to a reduction of the CSO volume. However, the effect of the overestimated inflows at the upstream sewer pipe still causes bad volume approximations in the first prediction time steps. To completely correct the effect of the bad flow-level approximations at the collector and its upstream sewer pipes, flow measurements at both locations must be used.

To illustrate the mentioned phenomena along the whole simulated events, plots of the collector inflows and contained volume as generated by the RHC model and MOUSE for the different measurement scenarios have been provided in Section 6.5 for rain event 09-10-2002. RHC approximations have been built using the first five minutes of each OCP problem solved, that is, the time corresponding to the time lapse of the applied control actions (i.e., the control interval). It can be noticed that for the MHEL scenario there is a difference of about $10^4\,\mathrm{m}^3$ between the volume computed by the

control model and the volume computed using the physically-based model. The same overestimation in the collector volume can also be observed, to a lesser extent, in the MHEC scenario. These overestimations cause the solution of the OCPs to force the downstream gate to release more volume than actually needed to fulfill the volume constraint, thus leading to higher CSO values.

From a more general point of view, it is worth noticing that the set-points for the gate flows computed by the OCPs based on the HLD model can be suitably achieved by the local controllers, as shown in Figures 6.32–6.35 and in Appendix E, proving that the flow approximations provided by the model are sufficiently accurate. Two exceptions, with minor impact to the final performance results, can be noticed. First, the already commented backwater effect occurring at gate g_8. Secondly, it can be noticed that the set-points at the collector downstream gate g_7 at the beginning of each rain event are not achieved and a few time steps later a peak flow over the desired 2 m^3/s flow (WWTP maximum inflow) occurs. This is due to the extended delay occurring with the first flow wave in each rain episode. The PID set-points are computed taking into account shorter delays, which are actually more accurate for the rest of the rain event, and ask the gate to release some positive amount of flow when no flow is actually available at the gate location. As a response, the local PID controllers force the gate to open more and more. Then, when the first flow flush reaches the gate, it opens beyond needed and a peak flow over the desired rates occurs before the gate has enough time to close again. Of course, this problem could be overcome by using different parameters at the beginning of the event. However, the focus of this thesis is to show the suitability of the proposed model for real-time control through its ability to accurately predict flows and techniques regarding on-line parameter calibration have not been covered, although they are an important future research topic that can surely help improve the controller performance.

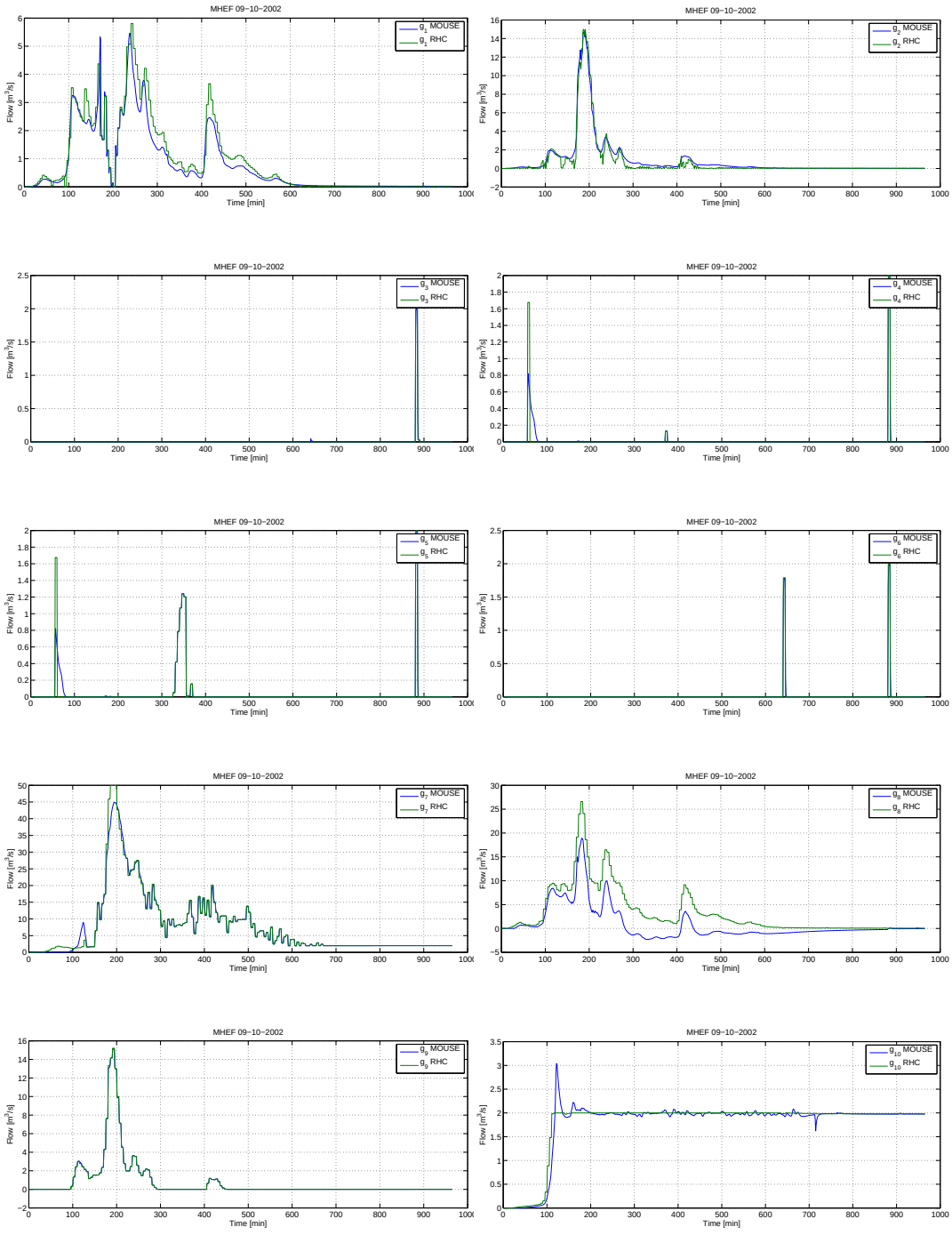

Figure 6.32: Gate flows and set-points for the MHEF scenario for rain event 09-10-2002. Similar plots for the rest of rain events are provided in Appendix E.

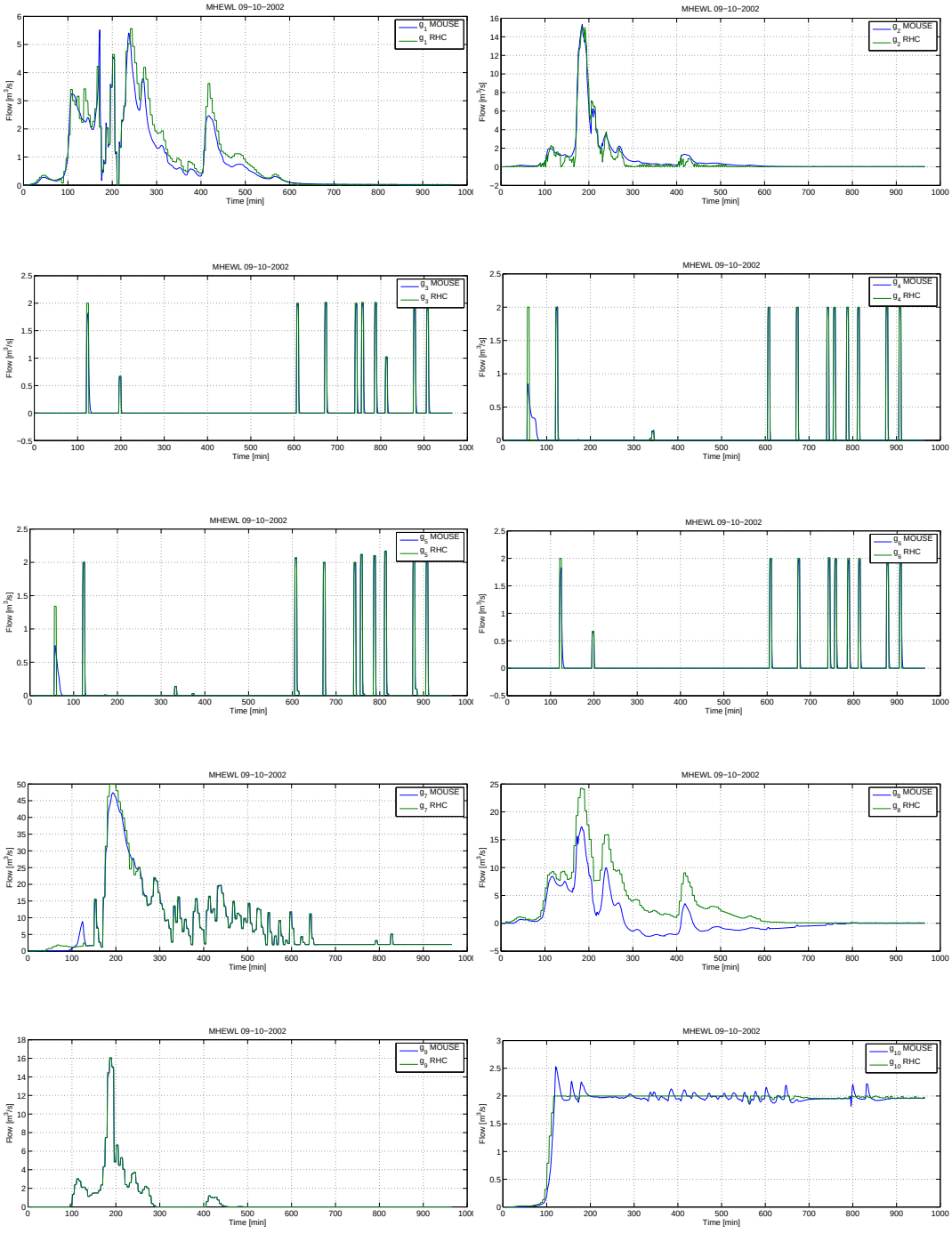

Figure 6.33: Gate flows and set-points for the MHEL scenario for rain event 09-10-2002. Similar plots for the rest of rain events are provided in Appendix E.

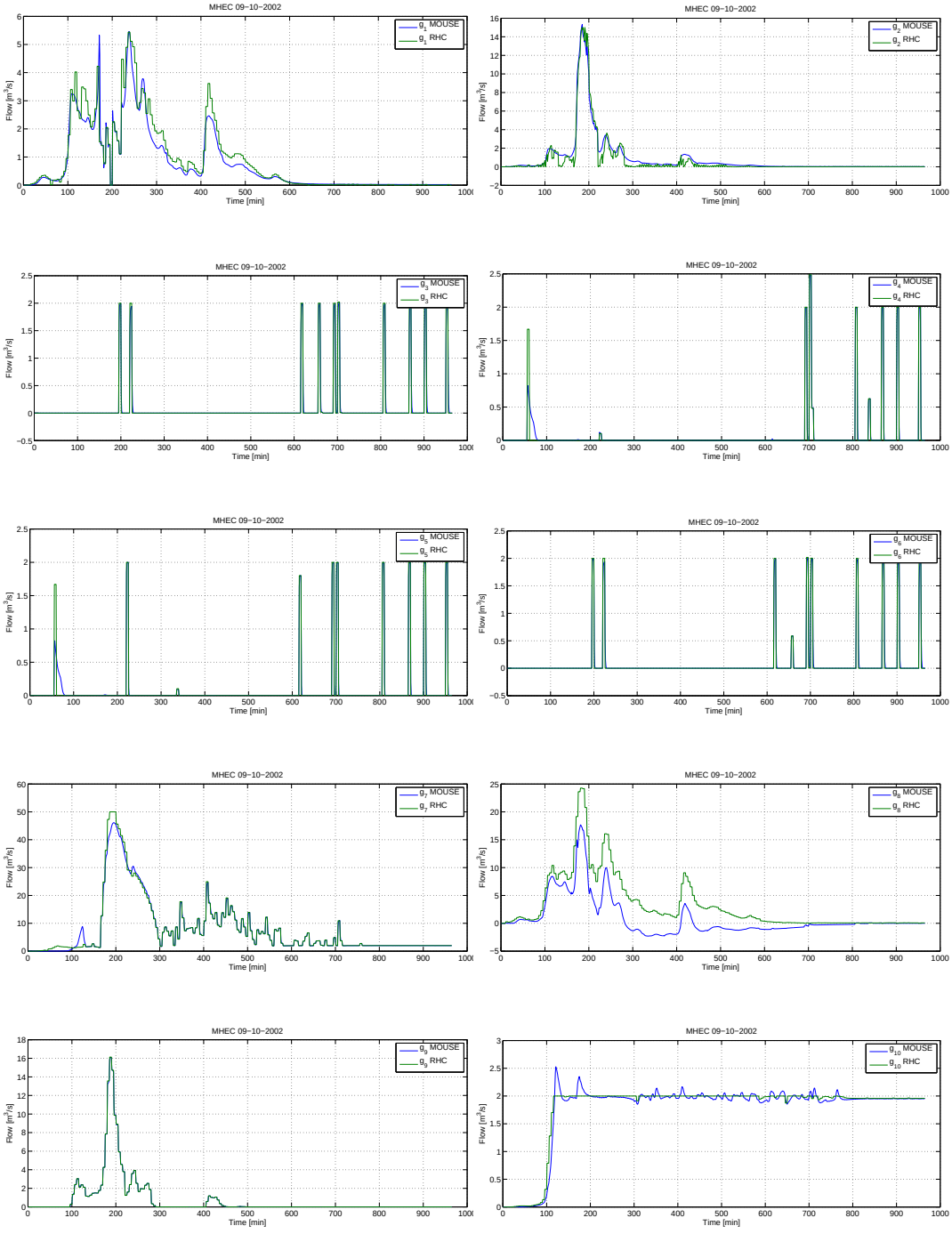

Figure 6.34: Gate flows and set-points for the MHEC scenario for rain event 09-10-2002. Similar plots for the rest of rain events are provided in Appendix E.

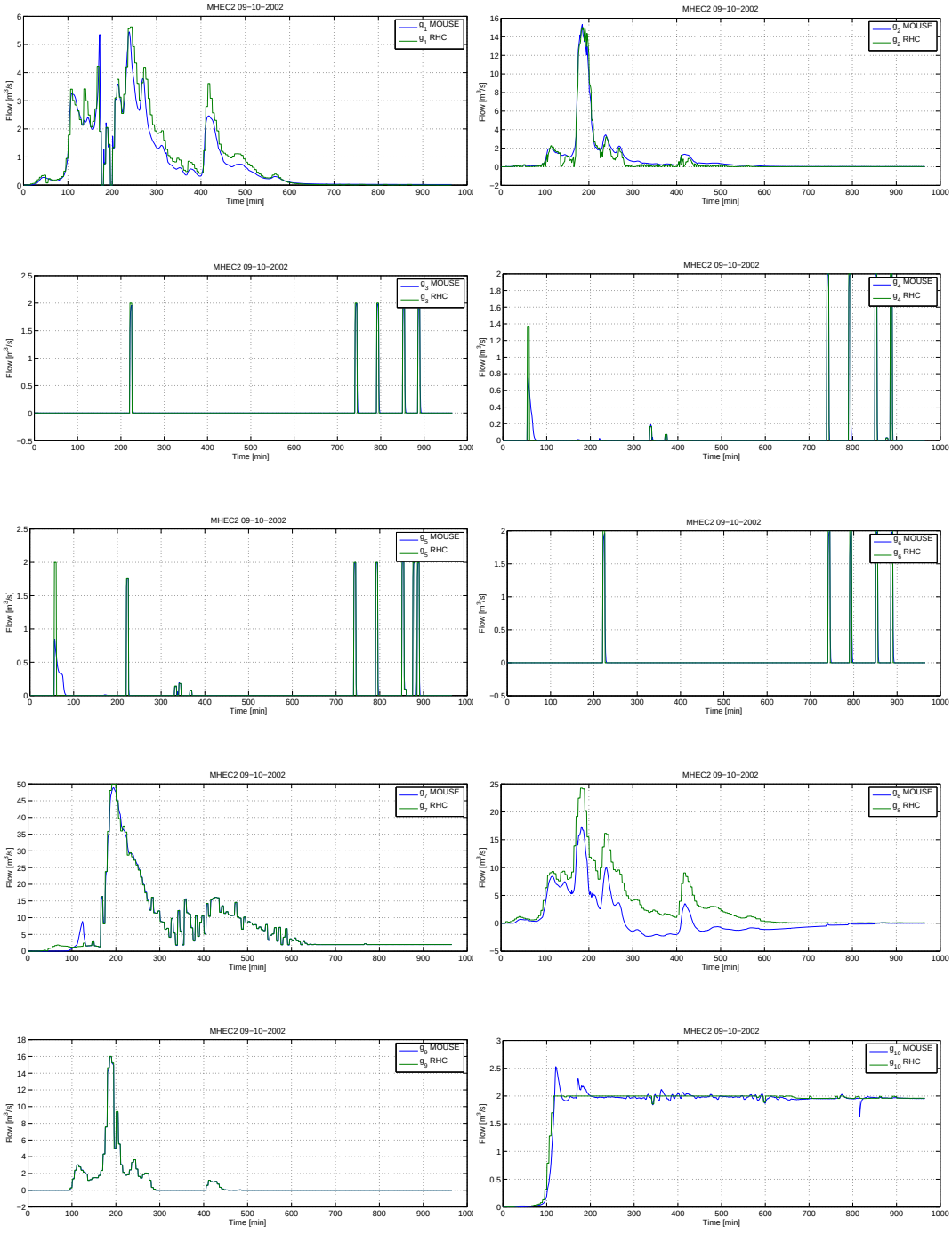

Figure 6.35: Gate flows and set-points for the MHEC2 scenario for rain event 09-10-2002. Similar plots for the rest of rain events are provided in Appendix E.

6.7 Computational Details

All optimization problems were solved using CPLEX v12.5 [CPLEX™, 2011] MILP solver with standard settings, available thanks to IBM Academic Initiative [IBM ILOG, 2013], on a desktop with an Intel Core 2 Duo CPU with 3.33 GHz and 8 GB RAM and a laptop with an Intel Core i7 CPU with 2.2 GHz and 8 GB RAM.

6.7.1 OCP Computational Details

Tables 6.8 and 6.9 show respectively the size and times needed to solve the OCP for the different rain events. It is a very important feature of the whole modelling and control approach that these problems can be solved within short times so that the whole real-time RHC strategy can be implemented. It can be noticed that all the maximum times needed to solve the problems are below 10 seconds, which are suitable times for a real-time controller taking into account that the control interval is of five minutes.

Table 6.8: Number of variables and constraints of the OCPs for the different considered models with a prediction horizon of $H = 40$.

Model	Equality Constraints	Inequality Constraints	Continuous Variables	Binary Variables
1T	6720	7200	7800	1040
1TD N=5	7040	7240	8120	1040
1TD N=10	7440	7240	8520	1040
1TD N=15	7840	7240	8920	1040
NT N=5	6920	8320	8160	1200
NT N=10	7120	9720	8560	1400
NT N=15	7320	11120	8960	1600

Table 6.9: OCPs computation times for the different measurement scenarios.

(a) 17-09-2002

Model	Mean Time [s]	Max Time [s]
1T	0.30	1.22
1TD N=5	0.58	1.45
1TD N=10	0.59	1.61
NT N=5	0.67	2.56
NT N=10	0.91	7.42

(b) 09-10-2002

Model	Mean Time [s]	Max Time [s]
1T	0.36	1.55
1TD N=5	0.66	1.73
1TD N=10	0.65	1.70
NT N=5	0.83	3.20
NT N=10	1.11	4.85

(c) 15-08-2006

Model	Mean Time [s]	Max Time [s]
1T	0.29	0.95
1TD N=5	0.58	1.84
1TD N=10	0.59	2.20
NT N=5	0.63	2.25
NT N=10	0.83	3.82

(d) 30-07-2011

Model	Mean Time [s]	Max Time [s]
1T	0.28	0.92
1TD N=5	0.58	1.89
1TD N=10	0.58	1.67
NT N=5	0.65	2.04
NT N=10	0.83	4.56

6.7.2 SEP Computational Details

Even though SEPs have less than half the number of variables than the OCPs (c.f. Table 6.10), they have proven to be harder to solve due to stronger conflict among the individual objectives in the cost function (a proper fitting at a particular measurement point can cause a poorer one at another point) which requires a higher number of iterations before optimality can be guaranteed, therefore taking longer computational times. To guarantee the computation times within each RHEC/MHE iteration to be suitable for a RTC application, a time limit of 1 minute has been set for all the SEP, provided a feasible suboptimal solution is available. Table 6.11 shows the mean and maximum computational times for the SEP for all the considered MHE scenarios. When the maximum time is above 60 second it means that the optimization has been stopped due to violation of the time limit constraint. The fourth and fifth columns in table 6.11, show the maximum percentage of suboptimality of the SEPs in which the optimization has prematurely stopped due to the time limit constraint and the number of times this situation has occurred out of 193 SEP instances for each rain event.

According to the CPLEX™ [2011] documentation, the suboptimality index, called the *Relative MIP Gap*, is computed taking into account the solutions of intermediate

Table 6.10: Details of the SEPs.

Continuous variables	3645
Discrete variables	390
Equality constraints	1783
Inequality constraints	2510

Table 6.11: SEP computation times, maximum MIP gap and number of time limit violations out of 193 SEP instances for each rain event.

Episode	Scenario	Mean Time	Max Time	Max MIP Gap	Time Limit Violations
17-09-2002	MHEF	3.36	60.12	0.68 %	2
	MHEL	3.76	60.05	0.47 %	3
	MHEC	3.48	60.03	0.31 %	1
	MHEC2	3.27	60.17	0.38 %	2
09-10-2002	MHEF	4.44	60.03	0.75 %	3
	MHEL	6.92	60.05	0.92 %	6
	MHEC	5.13	60.03	0.76 %	4
	MHEC2	5.01	60.05	0.32 %	3
15-08-2006	MHEF	2.81	36.47	0.01 %	0
	MHEL	2.71	42.72	0.01 %	0
	MHEC	3.07	60.05	0.08 %	1
	MHEC2	2.92	60.03	0.02 %	1
30-07-2011	MHEF	3.03	51.25	0.01 %	0
	MHEL	3.46	60.03	0.09 %	2
	MHEC	2.81	60.03	0.12 %	1
	MHEC2	2.57	60.02	0.02 %	1

subproblems solved during the branching algorithm used to solve the corresponding MILPs. Such algorithms relax the binary variable constraints $\delta \in \{0, 1\}$ to $\delta \in (0, 1)$ to obtain LPs. During the optimization process further subproblems are solved by forcing some of the binary variables to take one of their two possible values $\delta = 0$ or $\delta = 1$. This process is performed following a tree-like structure where the nodes are subproblems with different configurations of relaxed and fixed values for the binary variables. It is usual that in the solution of some of the node subproblems all the binary variables of the initial MILP already take either the value 0 or 1, even though the condition $\delta \in \{0, 1\}$ is not enforced for some of them. This solutions are called *integer* solutions, and are candidates to be the optimal solution to the problem. Notice that in a minimization

problem, while exploring the problem tree it always holds that

$$J_{INT} \geq J_{OPT} \geq J_{NODE},$$

where J_{INT} is the best integer solution objective value found, J_{OPT} is the optimal solution objective value and J_{NODE} is the best node solution objective value. The optimal solution objective value is always greater or equal than the best node solution found since the latter is the solution of a problem with less constraints. On the other hand, for the best integer solution found, the same constraints as for the optimal solution hold. However, although the best integer solution found is a candidate to be the optimal solution, it is still possible that a better one will be found while further exploring the problem tree. Taking this into account, the Relative MIP Gap is defined as the relative difference between the best integer solution and the best node solution:

$$RMIPG = \frac{|J_{INT} - J_{NODE}|}{10^{-10} + |J_{INT}|}$$

where the 10^{-10} is added to avoid division by zero. Taking into account the previous inequalities, the Relative MIP Gap provides an upper bound of the relative difference between the best integer solution and the optimal solution.

The number of violations of the time constraint and their corresponding values of the Relative MIP Gap (as a percentage) in Table 6.11 show that the situation is not common, and even in those cases the obtained suboptimal solution is sufficiently close to the optimal one to be used without problems in the RHC/MHE iterations. Notice that when the time constraint is not violated, the Relative MIP Gap value is always of 0.01% (i.e., $RMPIG = 10^{-4}$), since this is the default value below which the solver considers that the best integer solution is already the optimal one and the algorithm terminates.

6.8 Summary

In this chapter results of the performance of the model-based control techniques described in previous chapters for the case study sewer network have been presented. The model-based controller has been designed to minimize overflows and CSOs and to maximize the use of WWTP capacity. To assess the performance of the controller in

a realistic basis, a physically-based model of the case study sewer network has been used as virtual reality. By iteratively solving SEPs, OCPs and performing physically-based model simulations, a Receding Horizon Control strategy with Moving Horizon Estimation has been simulated. Overall, the results prove that the network objectives are fulfilled leading to a significant improvement with respect to passive control.

The problems arising from the lack of measurements in real applications have been considered by conducting closed-loop simulations considering four scenarios with different available measurements, including the ideal full-flow measurement case as a basis for comparison. To update the OCPs in the RHC iterations, estimates of the flow values all along the network are needed. However, in most real applications measurements are only available at several locations and they are not usually flow measurements, but water level ones. By means of a flow-level polynomial approximation, water level measurements can be used to compute flow values and by means of solving a SEP, estimates of the whole network flows can be obtained. As expected, due to errors in the flow-level approximations at the collector inflow, the water level measurements case is shown to provide the worst results in terms of CSOs. However, the water level measurements scenario with two additional flow measurements at the collector at the downstream end of the network, has proven to provide similar results as in the flow-measurement case.

Finally, details on the computational times needed to solve the OCPs and SEPs have been provided. All the OCPs have been solved within less than 10 seconds, therefore proving a suitable strategy for RTC. On the other hand, the SEPs have proven to be harder to solve with a few instances needing more than one minute to come up with a solution. To avoid problems in the RHC iterations for a real application, a time constraint of one minute has been added to the SEPs solver forcing the algorithm to terminate. In these situations the best feasible solution found is used to update the corresponding OCP. The degree of suboptimality of the solutions in those cases has been shown to be below 1%, thus it is concluded that the computational times for SEPs are also suitable for RTC.

Part IV

Concluding Remarks

Chapter 7

Conclusions and Future Work

7.1 Conclusions

According to the objectives stated in Chapter 1, the different stages for a model-based predictive controller set-up have been successfully developed and they also have been tested for a real case study using a physically-based model as virtual reality. It is worth to highlight the importance of having used these simulations to evaluate the accuracy of the model and performance of the controller, since they add much realism and reliability to the obtained results. In fact, this is the common procedure to assess control techniques in real applications (though it is not always in academic studies), before considering the actual implementation of such techniques in real operation. Therefore, much programming effort has been put in the development of the necessary software interfaces to communicate the physically-based model simulator and the optimization module.

The presented model has proven to fulfill the three requirements described in Chapter 1: accuracy, ease of calibration and suitable computational times for model-based Optimal Control and State Estimation problems. Unlike most models for large-scale networks proposed in the literature, the HLD model is based on modelling each network element individually, thus avoiding the need for topological aggregation. Using this modelling approach, the direct correspondence between the elements of a physically-based model and the control model strongly simplifies the calibration procedure, leads to accurate flow approximations and allows to separate the hydrologic submodel from the control model. Moreover, the definition of overflows at the junction level rather

than the at catchment one, helps predicting overflows with increased accuracy with respect to aggregate models.

Validation results comparing the model predictions with those generated by a physically-based model simulator have shown the model to be suitably accurate for control purposes. The accuracy of the model, however, depends on the intensity of the rain scenarios used for calibration. Using sensitivity analysis the robustness of the model has been tested for rain events stronger than those used for calibration, showing that moderate flows are still properly approximated while peak flows are not. Therefore, it is concluded that on-line calibration techniques or different parameter sets for different rain intensities should be used in a real implementation.

On the other hand, recent control-oriented models based on individual network elements seek to produce very accurate predictions and, therefore, are based on nonlinear formulations. This approach is not suitable for large-scale networks (e.g., the case study network used in this thesis, with more than 140 sewer pipes) since the resulting OCPs would become huge non-linear optimization problems or even mixed integer non-linear ones, if overflows were considered. The piecewise-linear framework used in the HLD model results in a suitable trade off between model accuracy and computational burden also taking advantage of the simplified topological and calibration features of non-aggregate models.

Thanks to the model accuracy and computational speed, the MHE strategy produces estimates of the network states at every control iteration without exceeding the available time, thus allowing for output feedback control. This is a very important feature of the model-based control strategy since, while optimal controllers require full-state initial conditions to formulate the optimization problems, only a few measurements are usually available in large-scale sewer networks. Therefore, a model-based controller taking into account individual network elements without a corresponding state estimation strategy would be impractical for real applications. Additionally, by means of the closed-loop output feedback simulations, assessment of sensor placing can be readily conducted.

The whole modelling and control approach has been developed to be readily applicable to general network instances without the need of expert knowledge or advanced mathematical or physical tools. Therefore, it can be seen as a starting point for further

development of advanced control and modelling techniques and as a benchmark for comparison for these newly developed techniques. On the other hand, it is also worth noticing the usefulness of this kind of modelling and control approach as a starting point to develop infrastructure and instrumentation planning techniques, as shown by the results provided in Chapter 6, where, by considering different measurement scenarios, it is shown that the installation of two flow meters might notably improve the control performance.

The main contributions of this thesis can be summarized as follows:

- A new sewer network model for real-time optimization-based receding horizon control has been developed and proven to fulfill three main requirements:
  - Accuracy
  - Calibration ease
  - Suitable computational times for optimal control and state estimation
- By means of a MLD systems reformulation, the model has proven to be easily adapted for the formulation of the following problems:
  - Constraint Satisfaction Problems for model simulation
  - Optimal Control Problems
  - State Estimation Problems
- A closed-loop simulation algorithm using a physically-based model simulator as virtual reality has been developed to evaluate the control performance of the model-based controllers and state estimators.
- Results of the closed-loop simulation according to different measurement scenarios have been shown to be useful for sensor placement.

7.2 Future Work

One first idea for future research is to improve collector flow approximations. Three different approaches for collector model taking into account delays and volume distribution have been tested with similar results due to the maximum volume constraint.

Although more accurate modelling can surely be achieved by using nonlinear modelling tools including both flows and water levels, the inclusion of such a submodel would force to reevaluate the whole control strategy since the optimization problems resulting from such a formulation are not guaranteed to have global optimal solution and the extended computational times required to numerically solve this kind of problems could possibly be unsuitable for real-time control. A non-centralized control approach with two independent models, one for the upper part of the network and one for the collector, leading to separate optimal control problems appears to be a promising approach to overcome the difficulties arising from a nonlinear model for collectors. However, it should be assessed whether such a strategy would still represent a performance improvement with respect to the hybrid linear modelling, taking into account the inherent suboptimality arising from trying to meet the control objectives by means of two separate optimal control problems. Coordinated strategies consisting in iteratively solving the two optimal control problems to decrease the suboptimality degree could be applied if the needed computational times allowed to. Notice, however, that with the latter approach, networks with complex topologies including several collectors communicating parts of the network with simpler dynamics would require the solution of not two, but several optimization problems in an iterative and coordinated way.

Regarding the control techniques, the proposed RHC and MHE algorithms can be said to be baseline approaches. Immediate improvements should take into account automatic on-line calibration techniques, OCP tuning and robustness techniques to deal with the rainfall forecast uncertainty.

From a more general point of view, information provided by rain forecasts and rainfall-runoff models should be used, not only as flow inputs to the network, but also to adapt the model and prioritize the control objectives accordingly. To this end, a wide range of real rain events would be needed to determine suitable classification criteria to later design different model and controller modes.

State estimation could be improved by means of better flow-level relations. Since these relations are not included into the optimization problems, more complex models can be used without risk of resulting into unsuitable computational times. Both physically-based models and identification-based models, for example including further

variables like gate position in a multivariable polynomial fitting, should be considered and compared.

Finally, a water quality model could be included to the HLD model in order to define quality-based objectives taking into account the WWTP state, according to an integrated model approach.

The following list includes the most relevant topics of future research to improve the proposed modelling and control approach, also according to the above discussion.

- Modelling
  - Collector modelling, including backwater and reverse-flow effects
  - Flow-level relations in presence of backwater effects
  - Rain event classification criteria for parameter selection
  - Water quality model
- Control
  - On-line calibration
  - OCP tuning
  - Robustness techniques to deal with the rainfall forecast uncertainty
  - Adaptive model and objective function according of the rain forecast

Since the objective of this thesis was to cover the main aspects for the set-up of a real-time model-based controller, focusing on in-depth development of the above topics has been left out of the scope. Each of these future research lines are complex enough to be subject of study on their own and the presented methodology and the corresponding developed software can be used as a basis to carry this research.

Appendix A

HLD System Expression of the Sewer Network Model

In Section 5.2 all the system equations have been reformulated into matrix expressions including MLD inequalities. The resulting expressions are in terms of vectors containing the different network elements at different time steps. To obtain the final model expression, a vector containing all the system variables needs to be defined so that all equations and inequalities can be expressed in terms of a single unknown at each time step.

Remark A.1. *Since the inequalities for the three collector models involve different matrices and variables three different final model expressions would be obtained. In the following, however, only the case of the N Tank model will be described, since the description of the three different cases does not add further insight to the technique used, and since the obtained general expression is analogous.*

The final vector of unkonwns includes all the systems variables at a given time step t:

$$X(t) = \Big(V(t)^\top, Q_{in}(t)^\top, W(t)^\top, \Delta_W(t)^\top, F(t)^\top, \Delta_F(t)^\top, V_f(t)^\top, \\ Q_f(t)^\top, \Delta_f(t)^\top, V_c(t)^\top, Q_c(t)^\top, \Delta_c(t)^\top, f_c(t)^\top, \delta_{f_c}(t)^\top, G(t)^\top\Big)^\top,$$

Notice that the rain inflow disturbances are not included in this vector, since for simulation, control and estimation purposes, they take predefined values that must be obtained by means of some forecasting technique.

Now equations (5.10), (5.9), (5.18) and (5.21) and inequalities (5.14), (5.17), (5.20), (5.22) and (5.23) can be expressed in compact matrix form in terms of $X(t)$ as

$$\begin{aligned} M_0\, X(t) &= b_{eq}(t) \\ N_0\, X(t) &\leq b_{ineq}(t) \end{aligned} \tag{A.1}$$

with

$$M_0 = \begin{pmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & -\tilde{A}_W & 0 & -\tilde{A}_F & 0 & 0 & -\tilde{A}_T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\tilde{A}_G \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$b_{eq}(t) = \begin{pmatrix} V(t-1)+\Delta t B_G G(t-1) \\ \sum_{i=1}^{T} \tilde{A}_i\, Q_{in}(t-i)+\tilde{A}_R\, R(t) \\ V_f(t-1)+\Delta t F(t-1)-\Delta t Q_f(t-1) \\ V_c(t-1)+K_{Q_c}\, Q_c(t-1)+K_{in}\, Q_{in}(t-1)+K_G\, G(t-1)+K_{F_c}\, F_c(t-1) \end{pmatrix}$$

and

$$N_0 = \begin{pmatrix} 0 & E_{Z_W} S_W A_Q A_0 & E_W+E_{Z_W} S_W A_W^+ & E_{\Delta_W} & 0 & 0 & 0 & \dots \\ 0 & E_{Z_F} S_F A_Q A_0 & E_{Z_F} S_F A_W^+ & 0 & E_F & E_{\Delta_F} & 0 & \dots \\ 0 & E_{Z_F} S_F A_Q A_0 & E_{Z_F} S_F A_W^+ & 0 & E_{F_f} & 0 & E_{V_f}-\frac{1}{\Delta t} E_{Q_f} & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \end{pmatrix}$$

$$\begin{pmatrix} \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & E_{Z_W} S_W A_G^+ \\ \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & E_{Z_F} S_F A_G^+ \\ \dots & E_{Q_f} & E_{\Delta_f} & 0 & 0 & 0 & 0 & 0 & E_{Z_F} S_F A_G^+ \\ \dots & 0 & 0 & 0 & 0 & 0 & E_{f_c} & E_{\delta_{f_c}} & 0 \\ \dots & 0 & 0 & H_{V_c} & H_{Q_c} & H_{\Delta_c} & 0 & 0 & H_G \end{pmatrix},$$

$$b_{ineq}(t) = -\begin{pmatrix} E_{Z_W} S_W \big(\sum_{i=1}^{T} A_Q A_i\, Q_{in}(t-i)+A_R\, R(t)\big)+E_{C_W} \\ E_{Z_F} S_F \big(\sum_{i=1}^{T} A_Q A_i\, Q_{in}(t-i)+A_R\, R(t)\big)+E_{C_F} \\ E_{Z_F} S_F \big(\sum_{i=1}^{T} A_Q A_i\, Q_{in}(t-i)+A_R\, R(t)\big)+E_{C_f} \\ E_{V_c}\, V_c(t-1)+E_{Q_c}\, Q_c(t-1)+E_{Q_{in}}\, Q_{in}(t-1)+E_c \\ H_c \end{pmatrix}.$$

Finally, to come up with an expression like (5.26) the left hand-sides of (A.1) are to be expressed in terms of the system variables at previous time steps $X(t-1), \dots, X(t-T)$:

$$b_{eq}(t) = -\sum_{i=1}^{T+1} M_i X(t-i) + m(t),$$

$$b_{ineq}(t) = -\sum_{i=1}^{T} N_i X(t-i) + n(t),$$

with

$$M_1 = -\begin{pmatrix} I & 0 & 0\,0 & 0 & 0\,0 & 0 & 0\,0 & 0 & 0 & 0 & 0 & \Delta t B_G \\ 0 & \tilde{A}_1 & 0\,0 & 0 & 0\,0 & 0 & 0\,0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0\,0 & \Delta t I & 0\,I & -\Delta t I & 0\,0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -K_{Q_{in}} & 0\,0 & 0 & 0\,0 & 0 & 0\,0 & -K_{Q_c} & 0 & -K_{F_c} & 0 & -K_G \end{pmatrix},$$

$$M_i = -\begin{pmatrix} 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & \tilde{A}_i & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \end{pmatrix}, \qquad i = 2 \ldots T,$$

$$m(t) = \begin{pmatrix} 0 \\ \tilde{A}_R R(t) \\ 0 \\ 0 \end{pmatrix}.$$

and

$$N_1 = \begin{pmatrix} 0 & E_{Z_W} S_W A_1 & 0\,0\,0\,0\,0\,0\,0\,0 & 0 & 0 & 0\,0\,0\,0 \\ 0 & E_{Z_F} S_F A_1 & 0\,0\,0\,0\,0\,0\,0\,0 & 0 & 0 & 0\,0\,0\,0 \\ 0 & E_{Z_F} S_F A_1 & 0\,0\,0\,0\,0\,0\,0\,0 & 0 & 0 & 0\,0\,0\,0 \\ 0 & E_{Q_{in}} & 0\,0\,0\,0\,0\,0\,0\,0 & E_{V_c} & E_{Q_c} & 0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0 & 0 & 0 & 0\,0\,0\,0 \end{pmatrix},$$

$$N_i = \begin{pmatrix} 0 & E_{Z_W} S_W A_i & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & E_{Z_F} S_F A_i & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & E_{Z_F} S_F A_i & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \end{pmatrix}, \qquad i = 2 \ldots T,$$

$$n(t) = -\begin{pmatrix} E_{Z_W} S_W A_R R(t) + E_w \\ E_{Z_F} S_F A_R R(t) + E_f \\ E_{Z_F} S_F A_R R(t) + E_{q_f} \\ E_c \\ H_c \end{pmatrix}.$$

Appendix B

RHC Results Figures

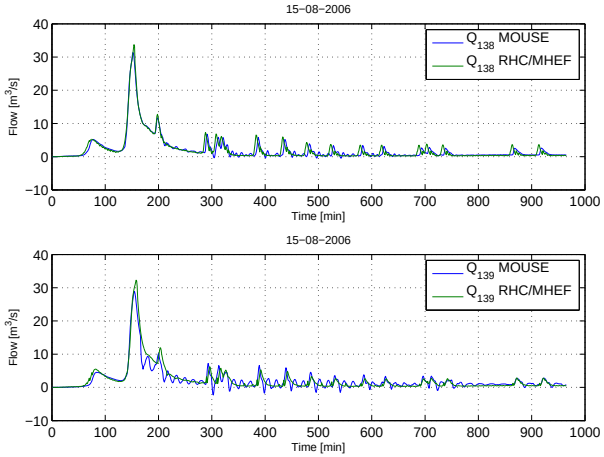

Figure B.1: RHC prediction and MOUSE simulation values for the MHEF scenario for sewer pipes q_{138} and q_{139}.

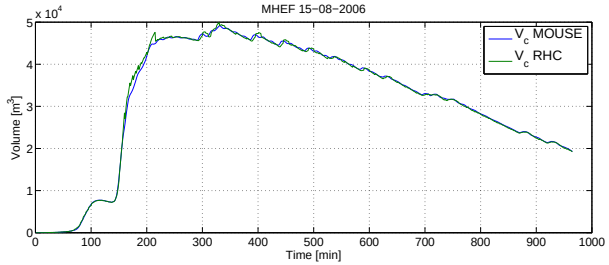

Figure B.2: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

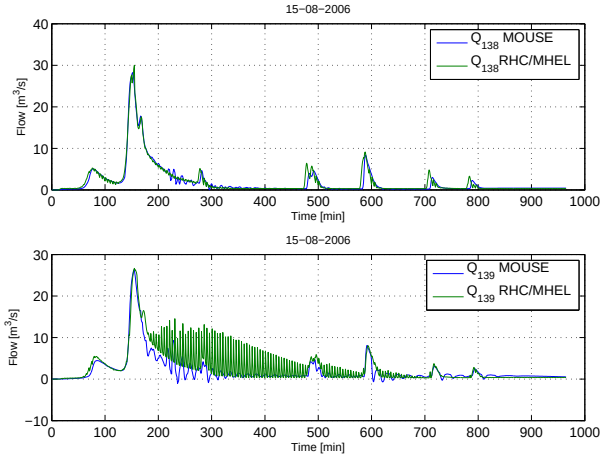

Figure B.3: RHC prediction and MOUSE simulation values for the MHEL scenario for sewer pipes q_{138} and q_{139}.

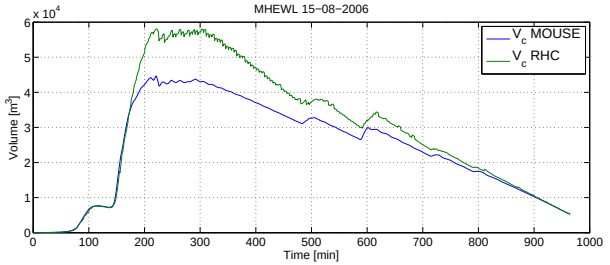

Figure B.4: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

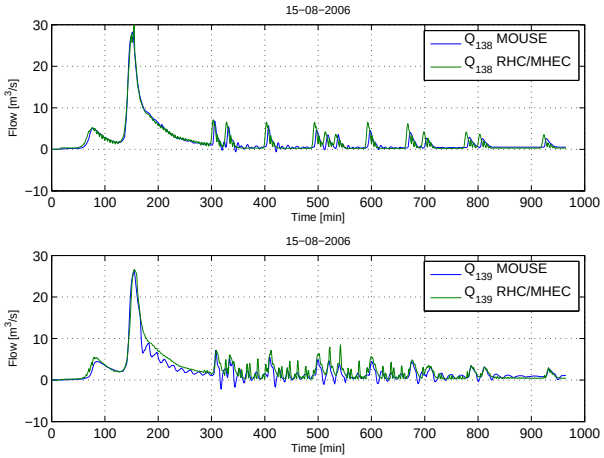

Figure B.5: RHC prediction and MOUSE simulation values for the MHEC scenario for sewer pipes q_{138} and q_{139}.

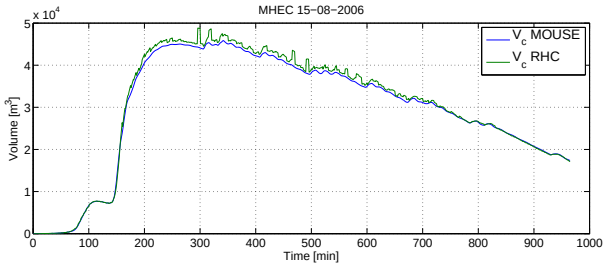

Figure B.6: RHC prediction and MOUSE simulation values for the MHEC scenario for collector q_{139} volume.

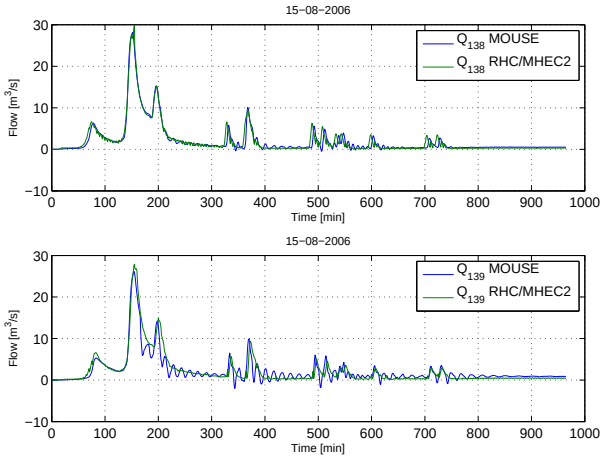

Figure B.7: RHC prediction and MOUSE simulation values for the MHEC2 scenario for sewer pipes q_{138} and q_{139}.

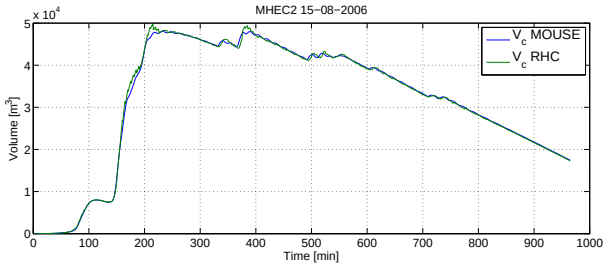

Figure B.8: RHC prediction and MOUSE simulation values for the MHEC2 scenario for collector q_{139} volume.

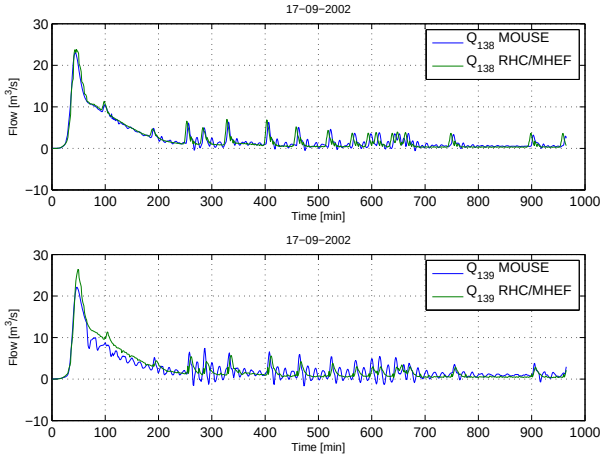

Figure B.9: RHC prediction and MOUSE simulation values for the MHEF scenario for sewer pipes q_{138} and q_{139}.

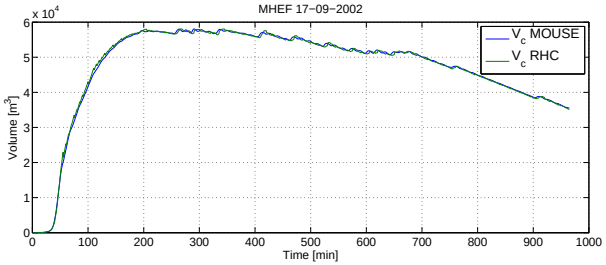

Figure B.10: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

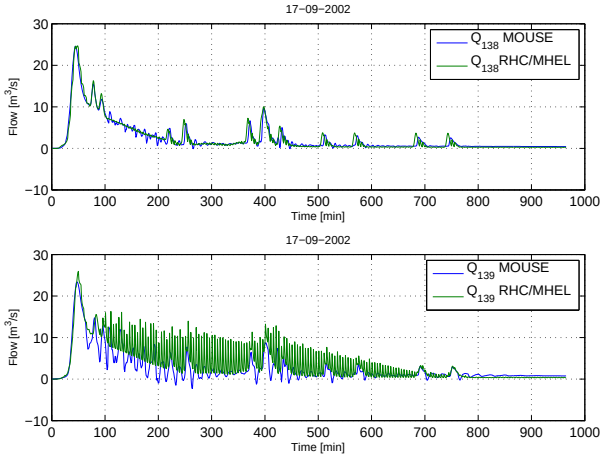

Figure B.11: RHC prediction and MOUSE simulation values for the MHEL scenario for sewer pipes q_{138} and q_{139}.

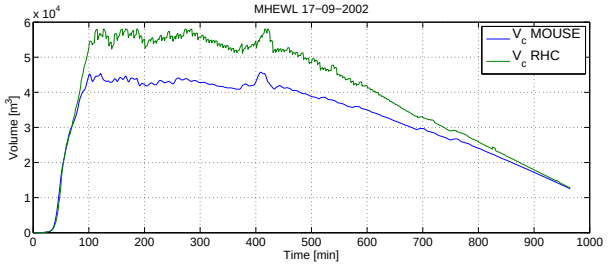

Figure B.12: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

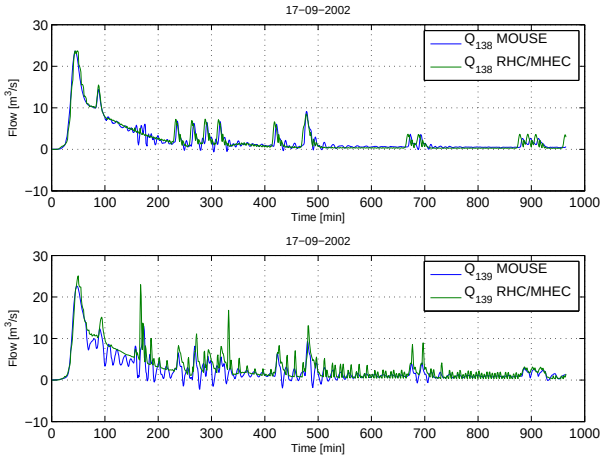

Figure B.13: RHC prediction and MOUSE simulation values for the MHEC scenario for sewer pipes q_{138} and q_{139}.

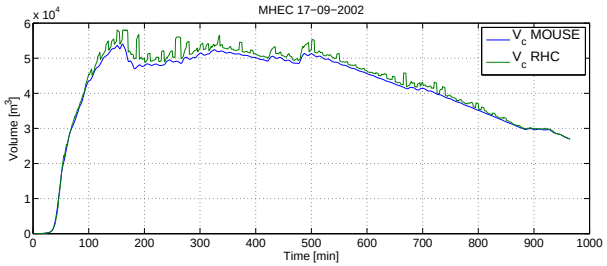

Figure B.14: RHC prediction and MOUSE simulation values for the MHEC scenario for collector q_{139} volume.

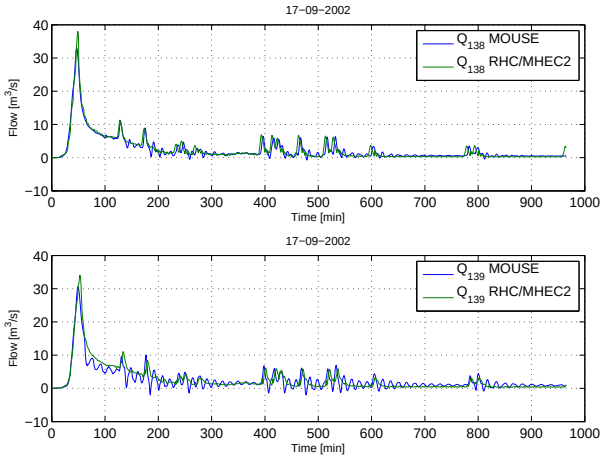

Figure B.15: RHC prediction and MOUSE simulation values for the MHEC2 scenario for sewer pipes q_{138} and q_{139}.

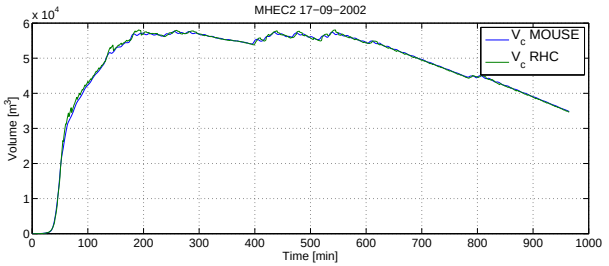

Figure B.16: RHC prediction and MOUSE simulation values for the MHEC2 scenario for collector q_{139} volume.

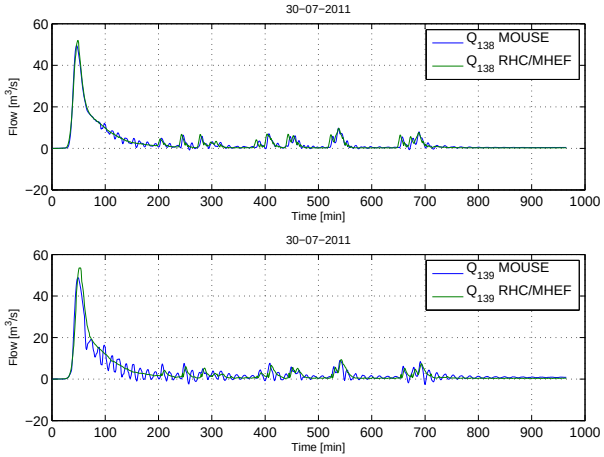

Figure B.17: RHC prediction and MOUSE simulation values for the MHEF scenario for sewer pipes q_{138} and q_{139}.

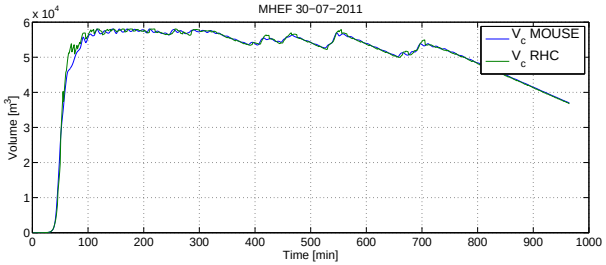

Figure B.18: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

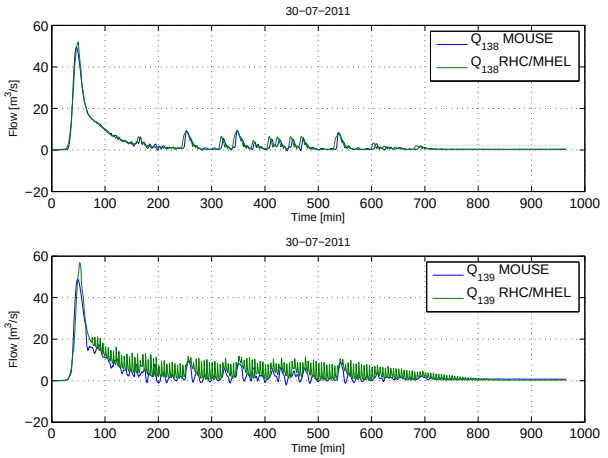

Figure B.19: RHC prediction and MOUSE simulation values for the MHEL scenario for sewer pipes q_{138} and q_{139}.

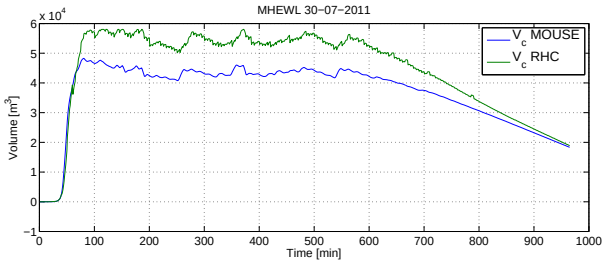

Figure B.20: RHC prediction and MOUSE simulation values for the MHEF scenario for collector q_{139} volume.

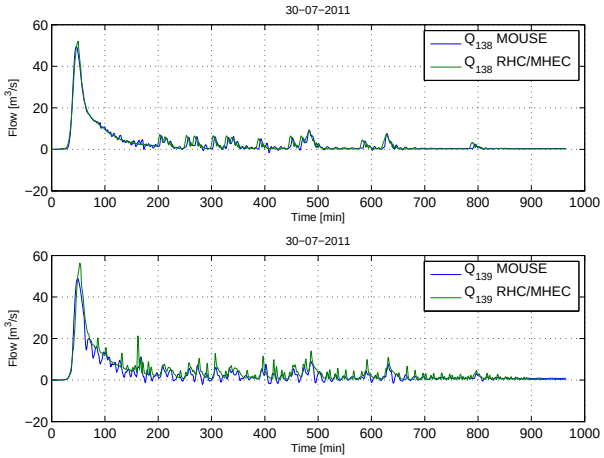

Figure B.21: RHC prediction and MOUSE simulation values for the MHEC scenario for sewer pipes q_{138} and q_{139}.

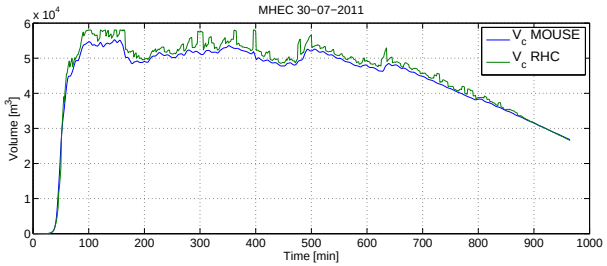

Figure B.22: RHC prediction and MOUSE simulation values for the MHEC scenario for collector q_{139} volume.

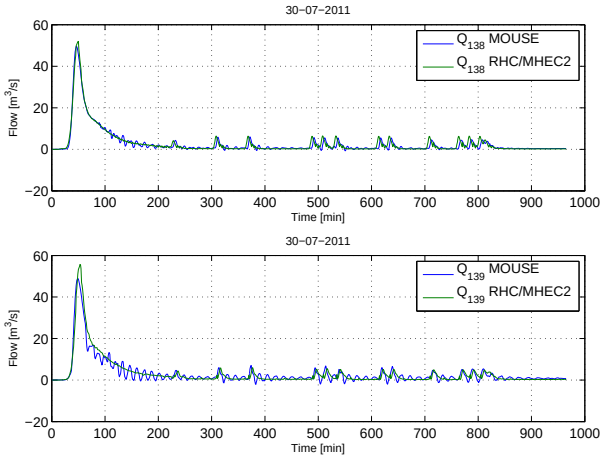

Figure B.23: RHC prediction and MOUSE simulation values for the MHEC2 scenario for sewer pipes q_{138} and q_{139}.

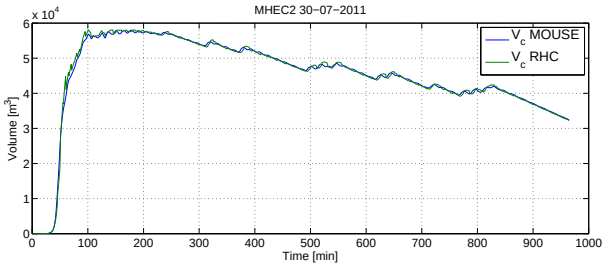

Figure B.24: RHC prediction and MOUSE simulation values for the MHEC2 scenario for collector q_{139} volume.

Appendix C

Flow-Level Polynomial Fitting Figures

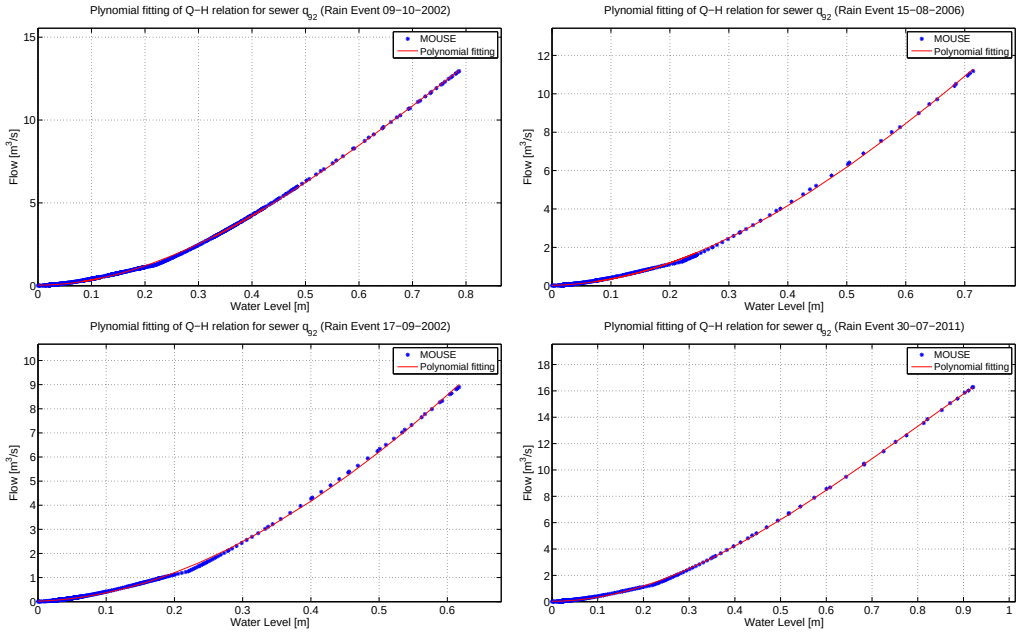

Figure C.1: Calibration data and polynomial fitting for sewer pipe q_{92} for each rain scenario.

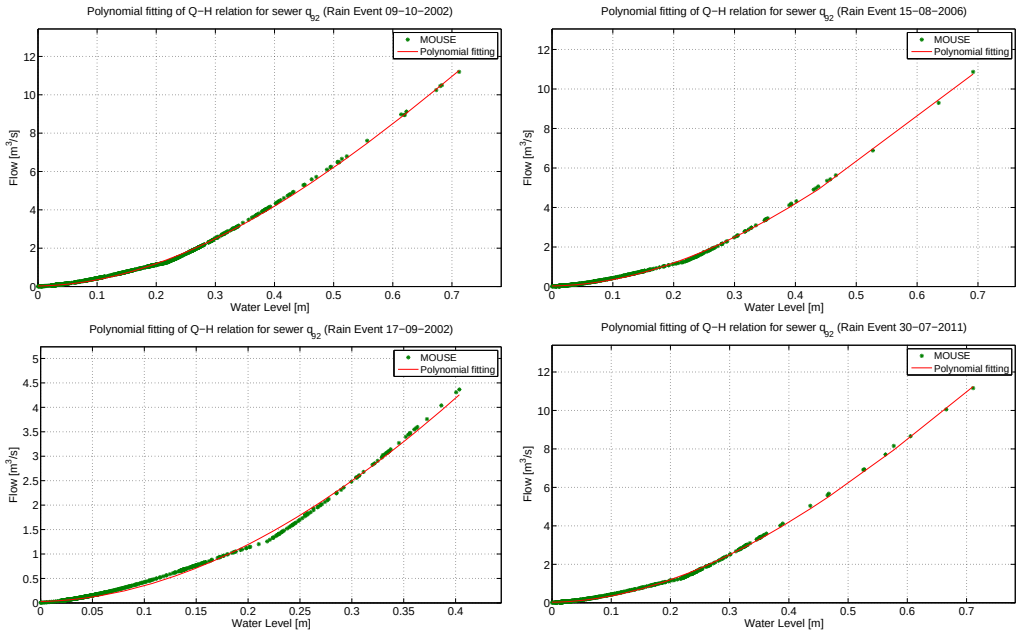

Figure C.2: Closed-loop simulation data and polynomial approximation for sewer pipe q_{92} for each rain scenario.

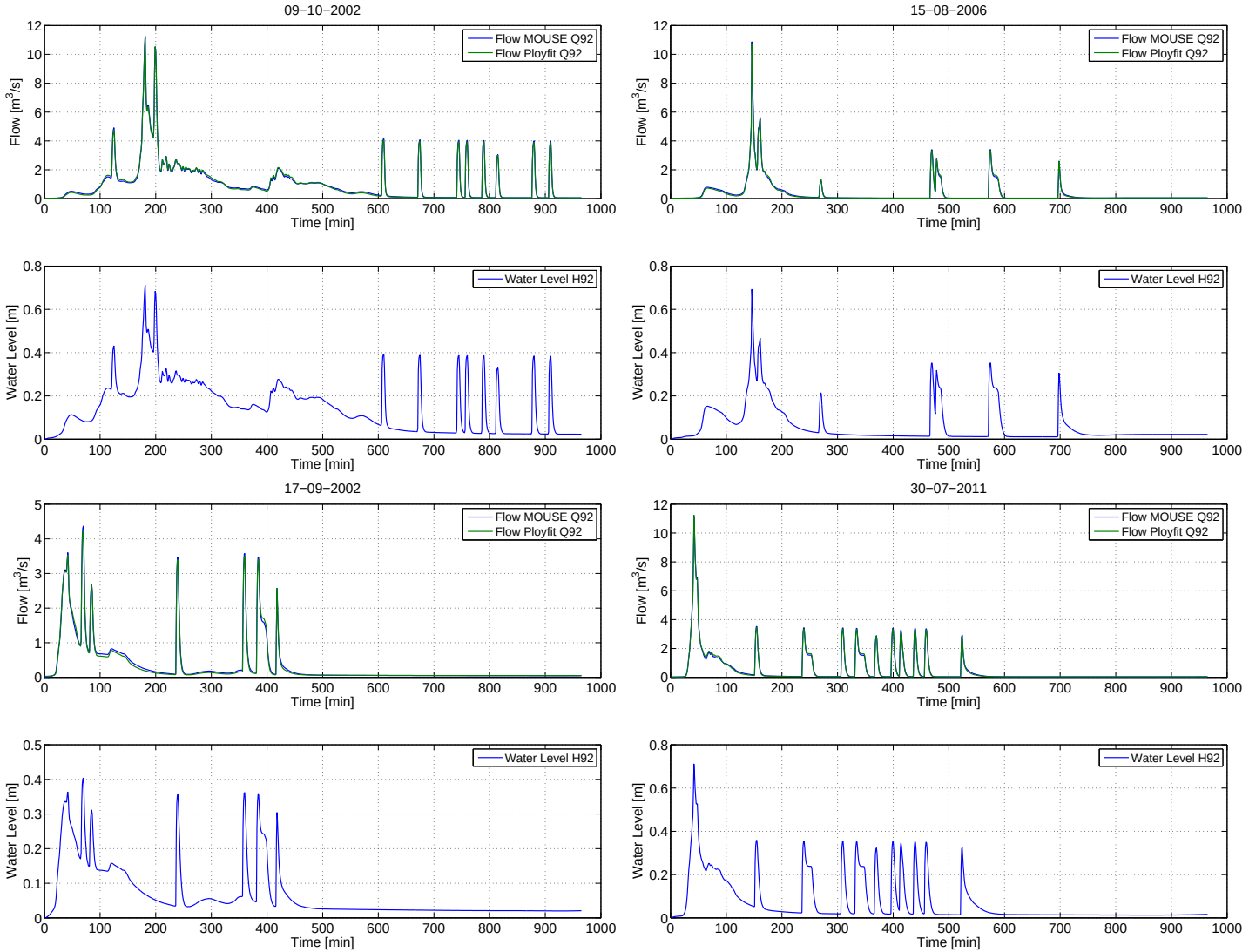

Figure C.3: Flow-level transformation of the closed-loop simulations water level for sewer pipe q_{92} for each rain scenario.

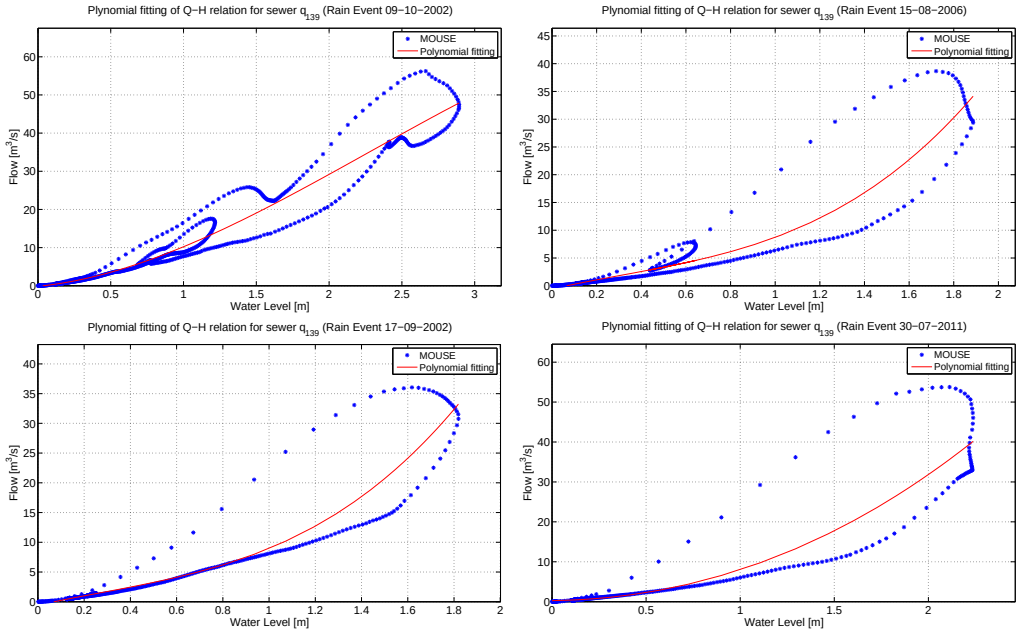

Figure C.4: Calibration data and polynomial fitting for collector q_{139} for each rain scenario.

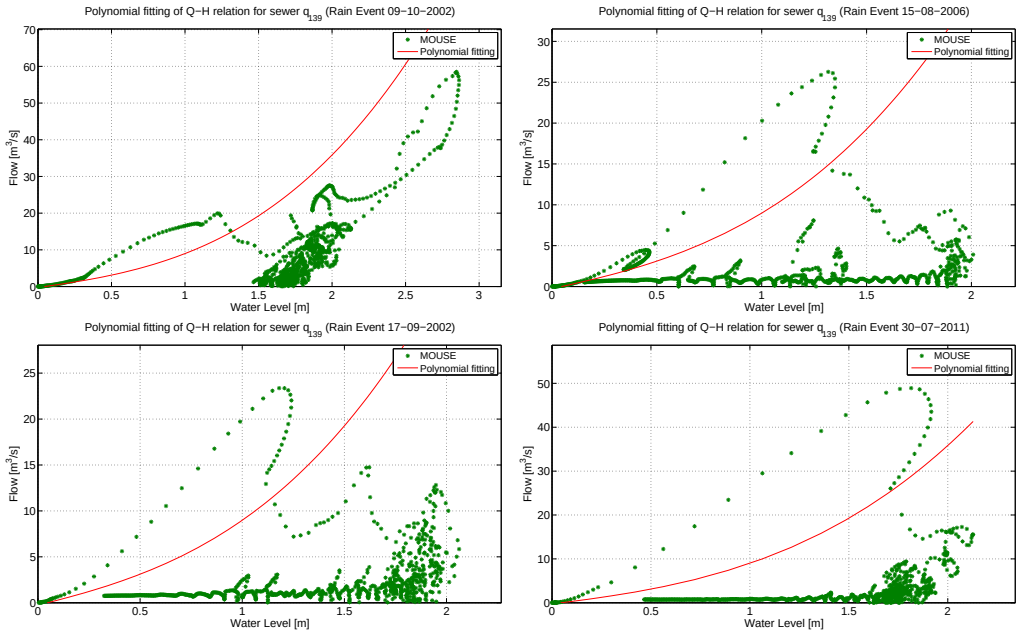

Figure C.5: Closed-loop simulation data and polynomial approximation for collector q_{139} for each rain scenario.

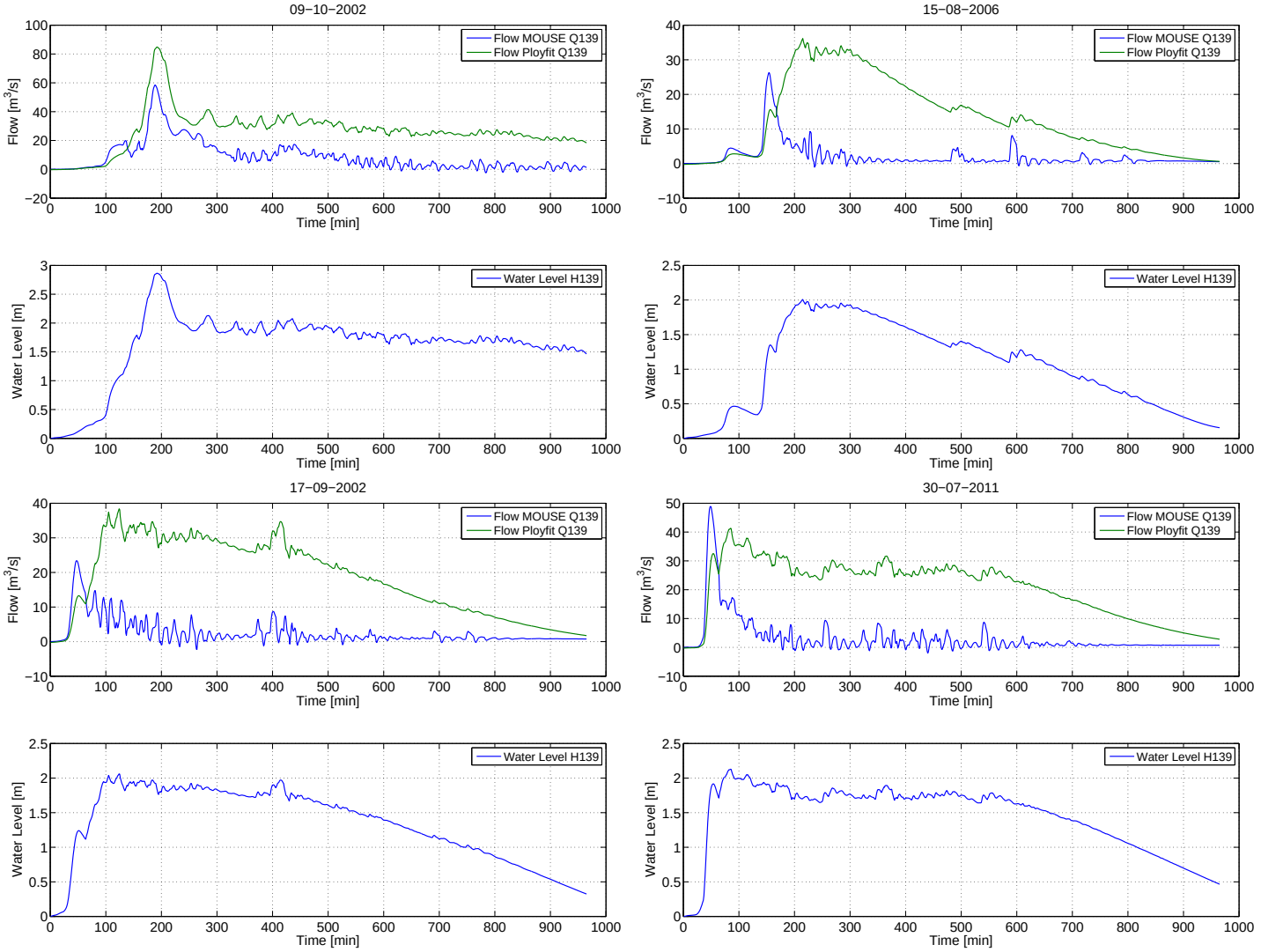

Figure C.6: Flow-level transformation of the closed-loop simulations water level for sewer pipe q_{139} for each rain scenario.

Appendix D

SEP Solution Figures

SEP solutions MHEF 15-08-2006

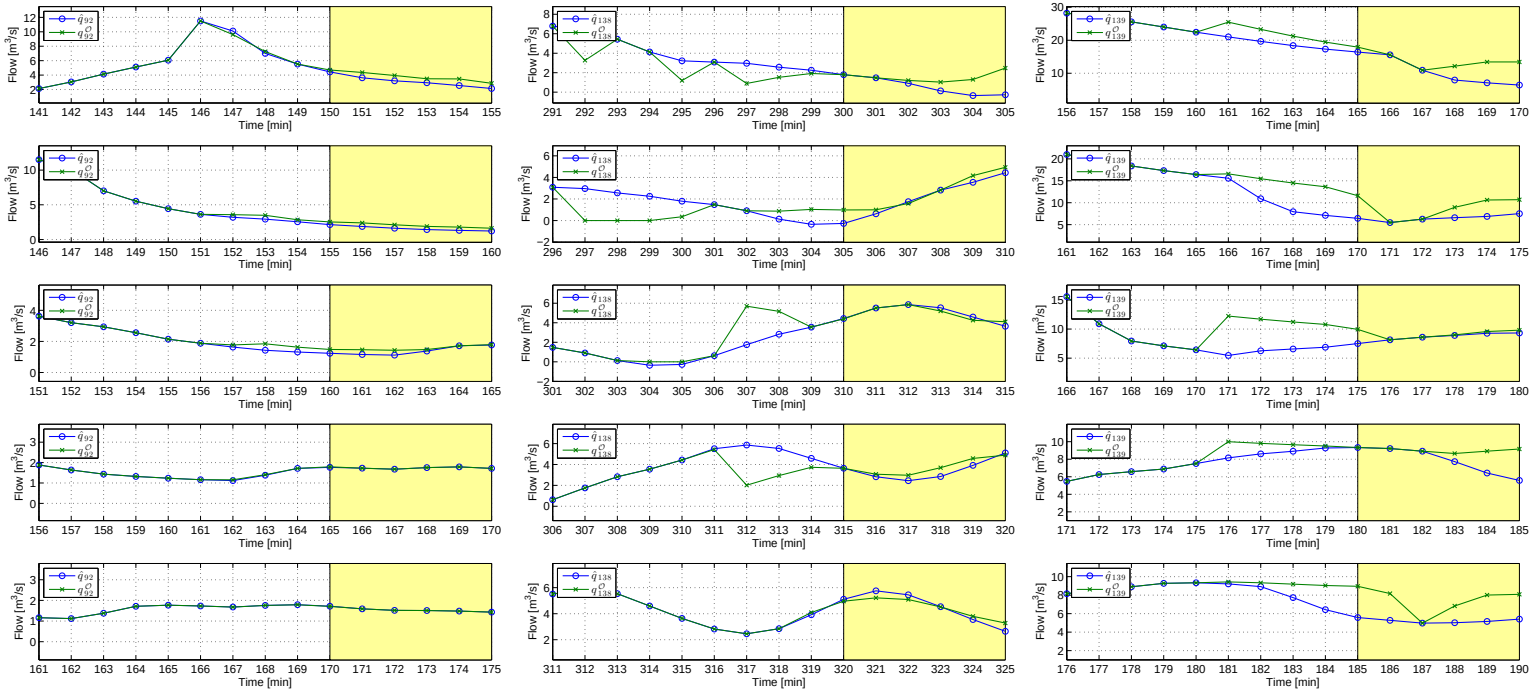

Figure D.1: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEF scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

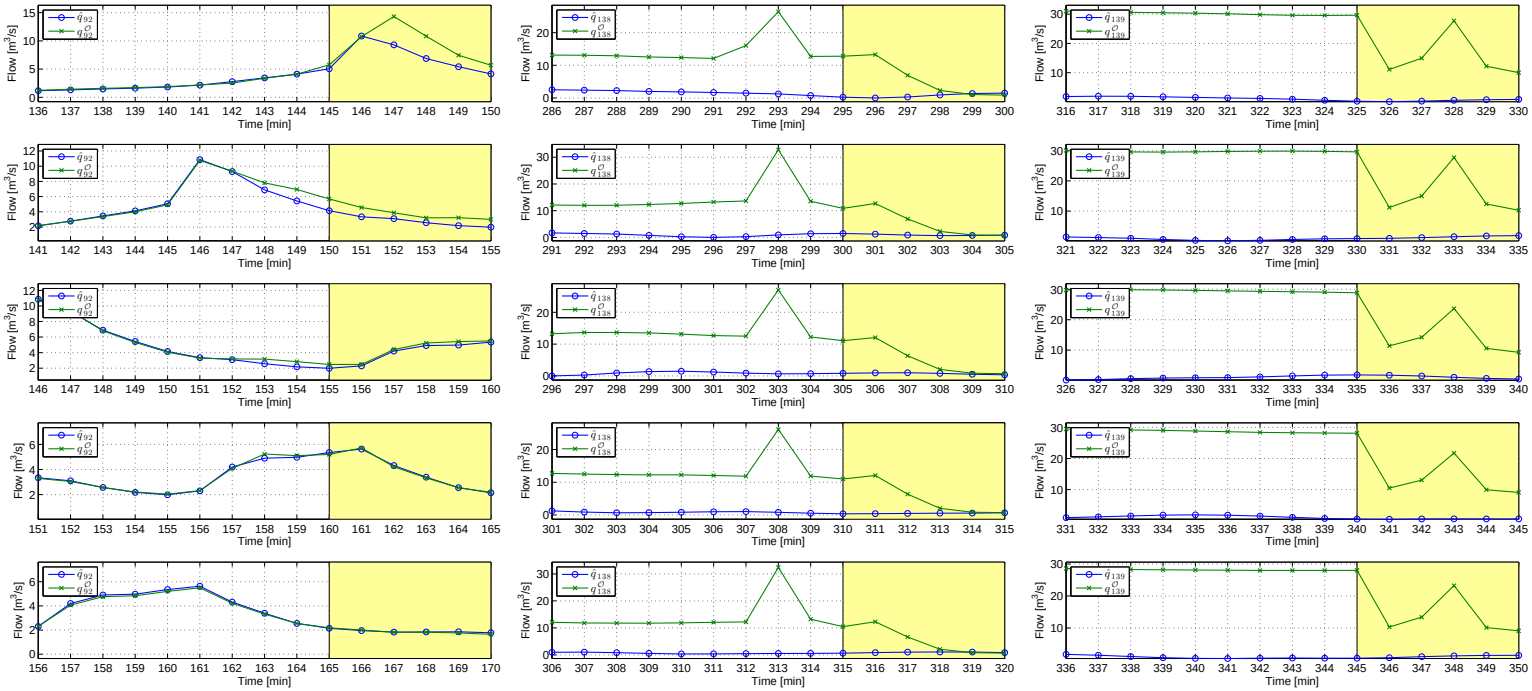

Figure D.2: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEL scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

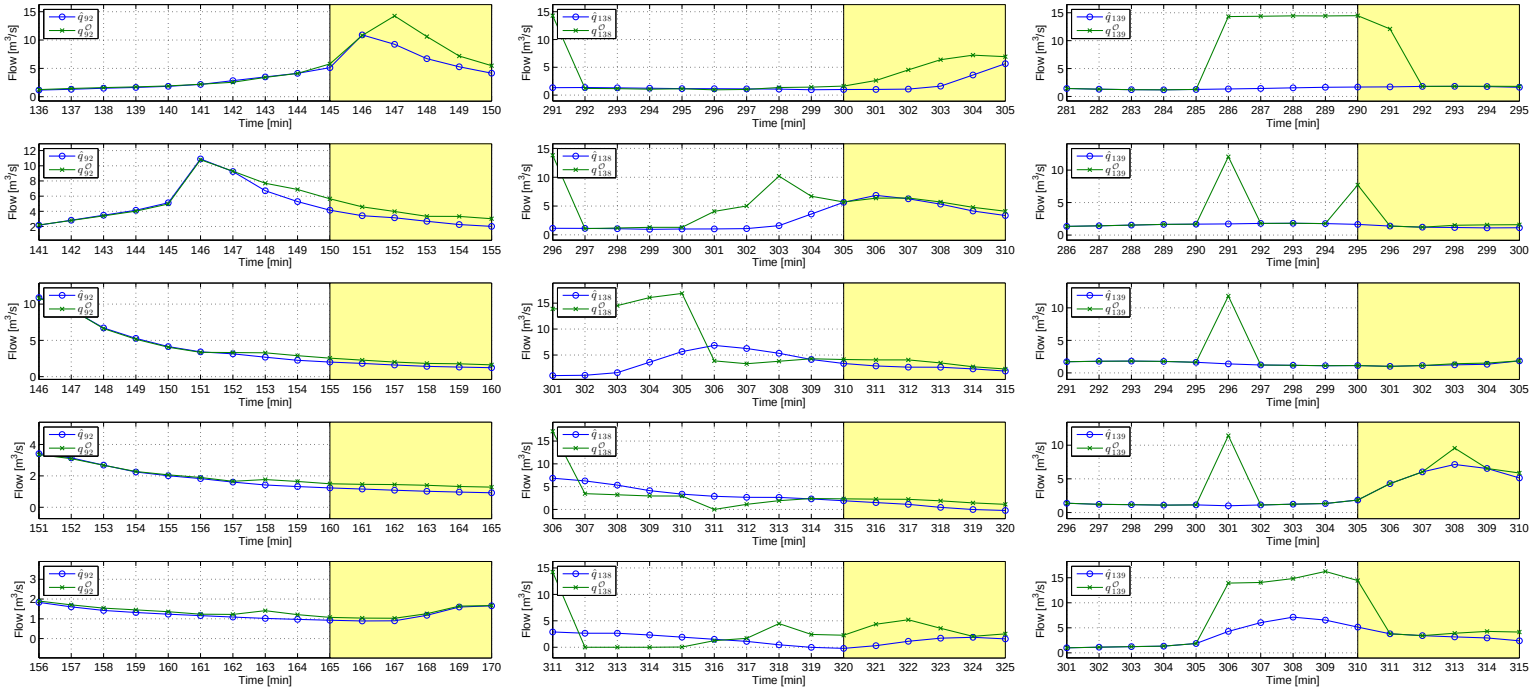

Figure D.3: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

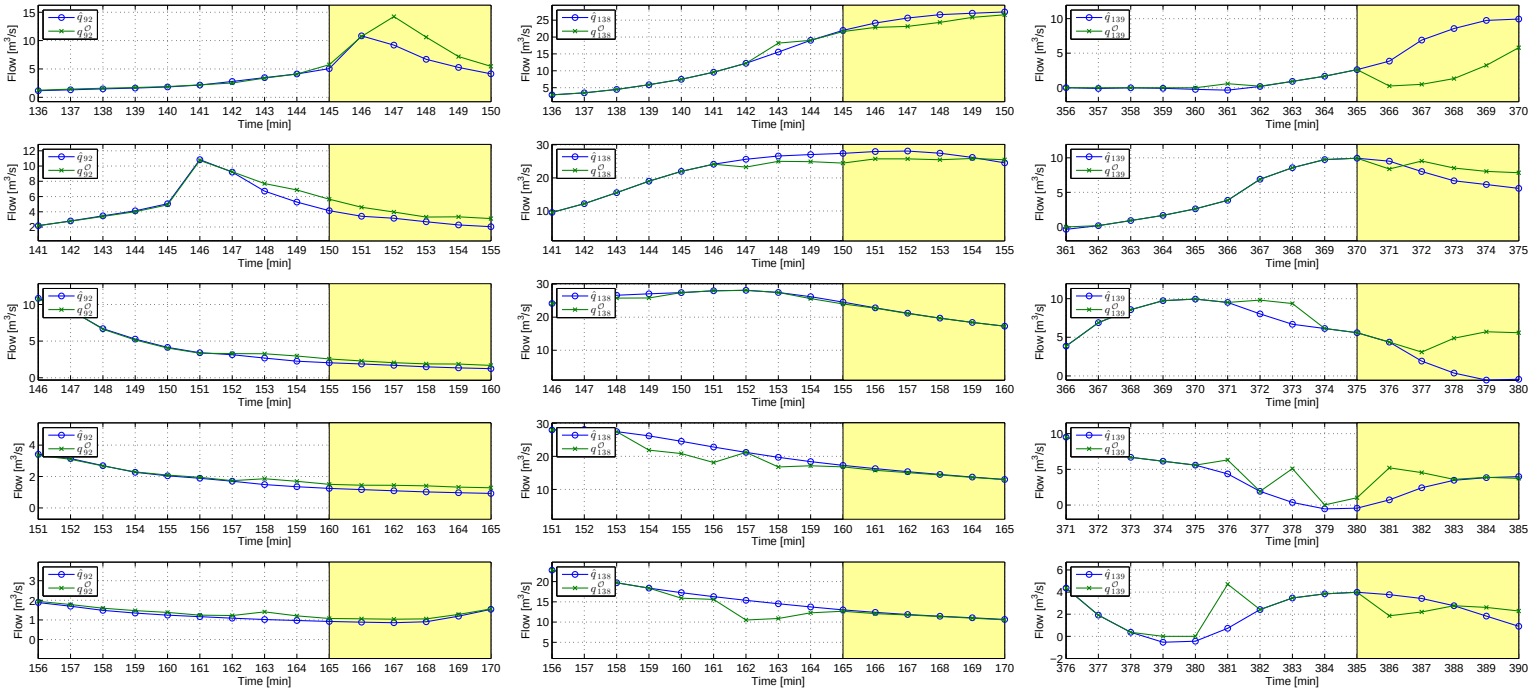

Figure D.4: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC2 scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

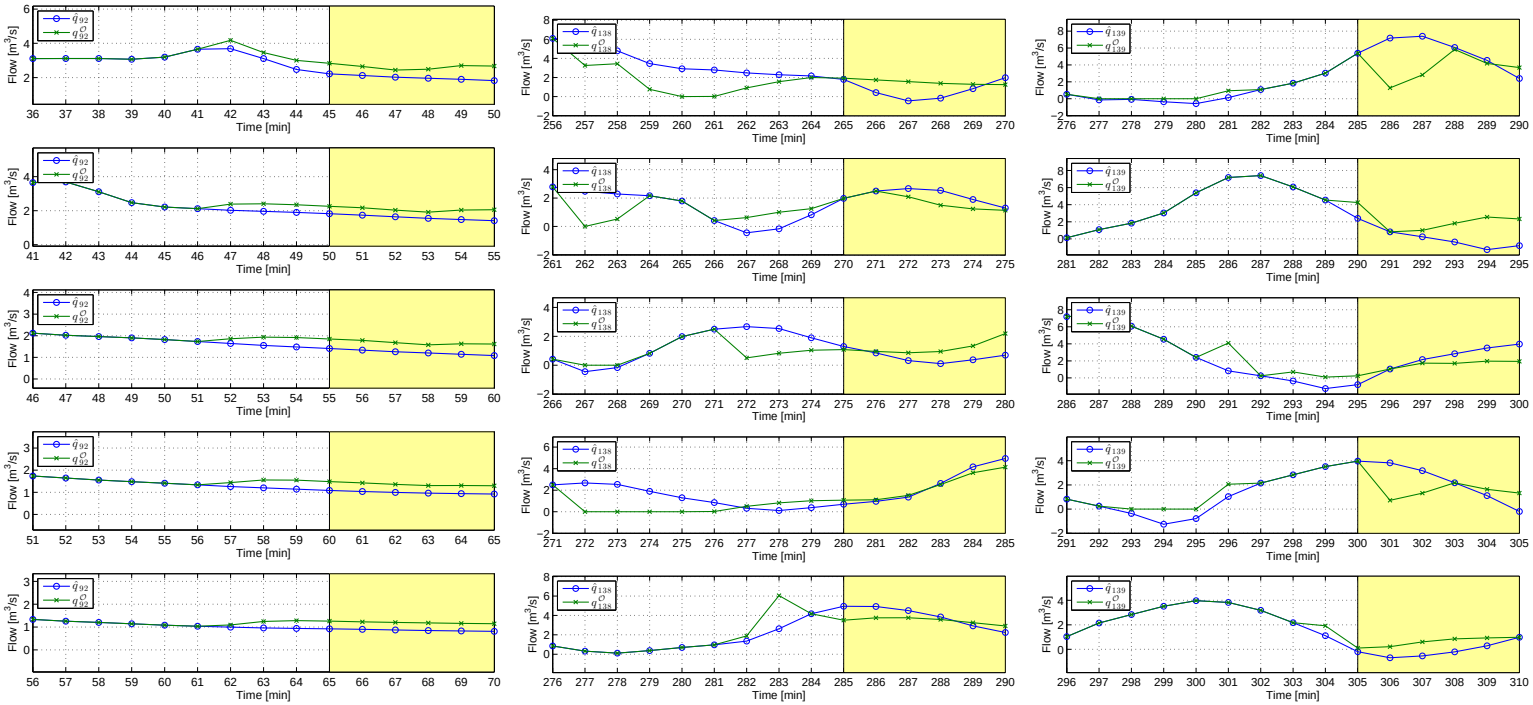

Figure D.5: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEF scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

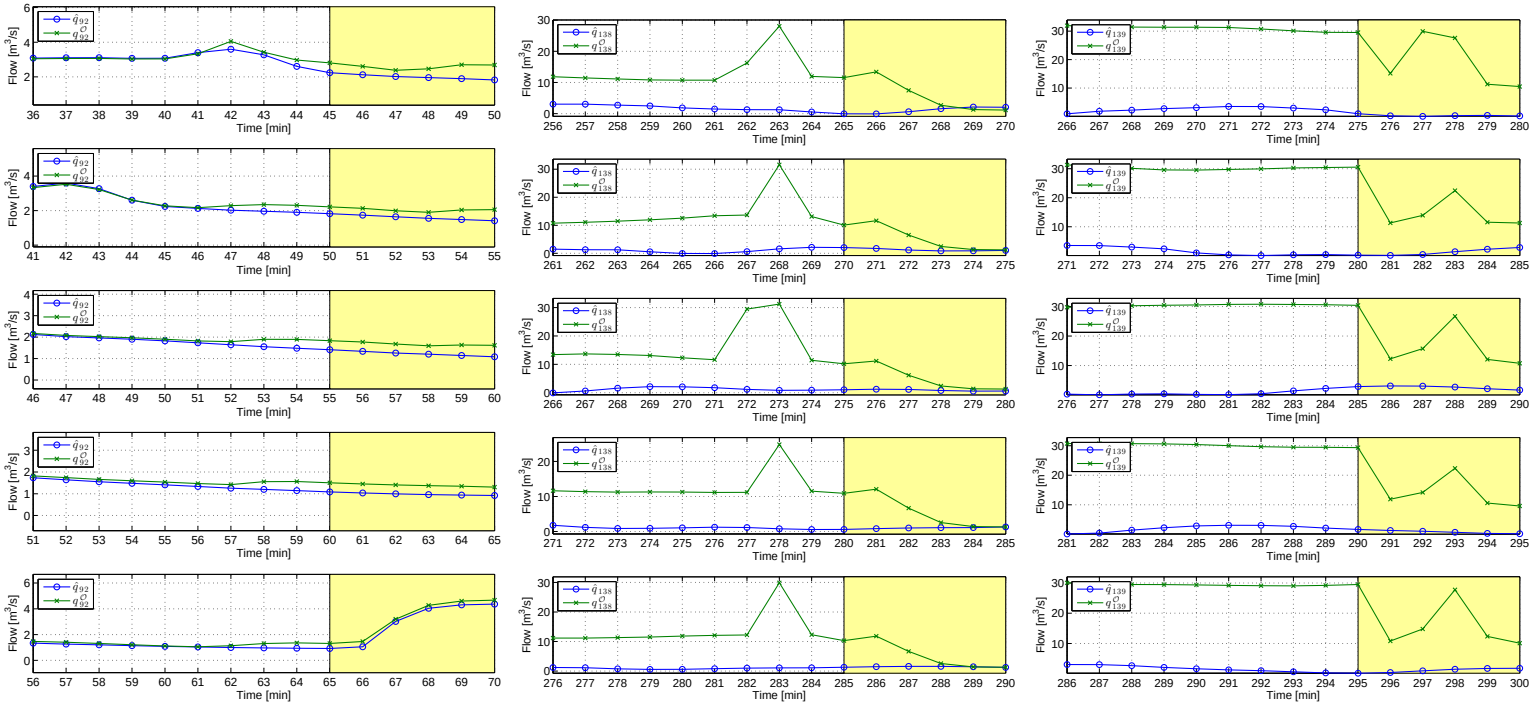

Figure D.6: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEL scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

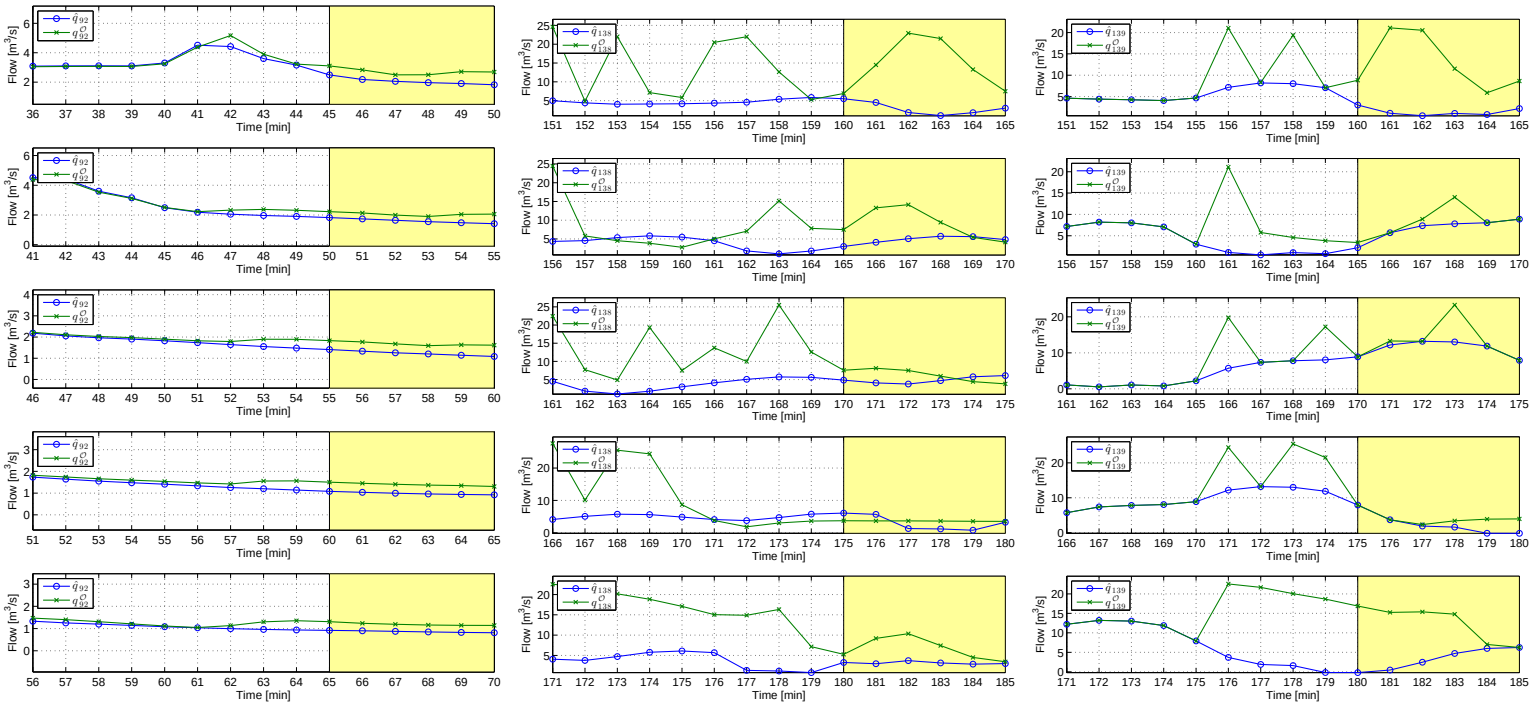

Figure D.7: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

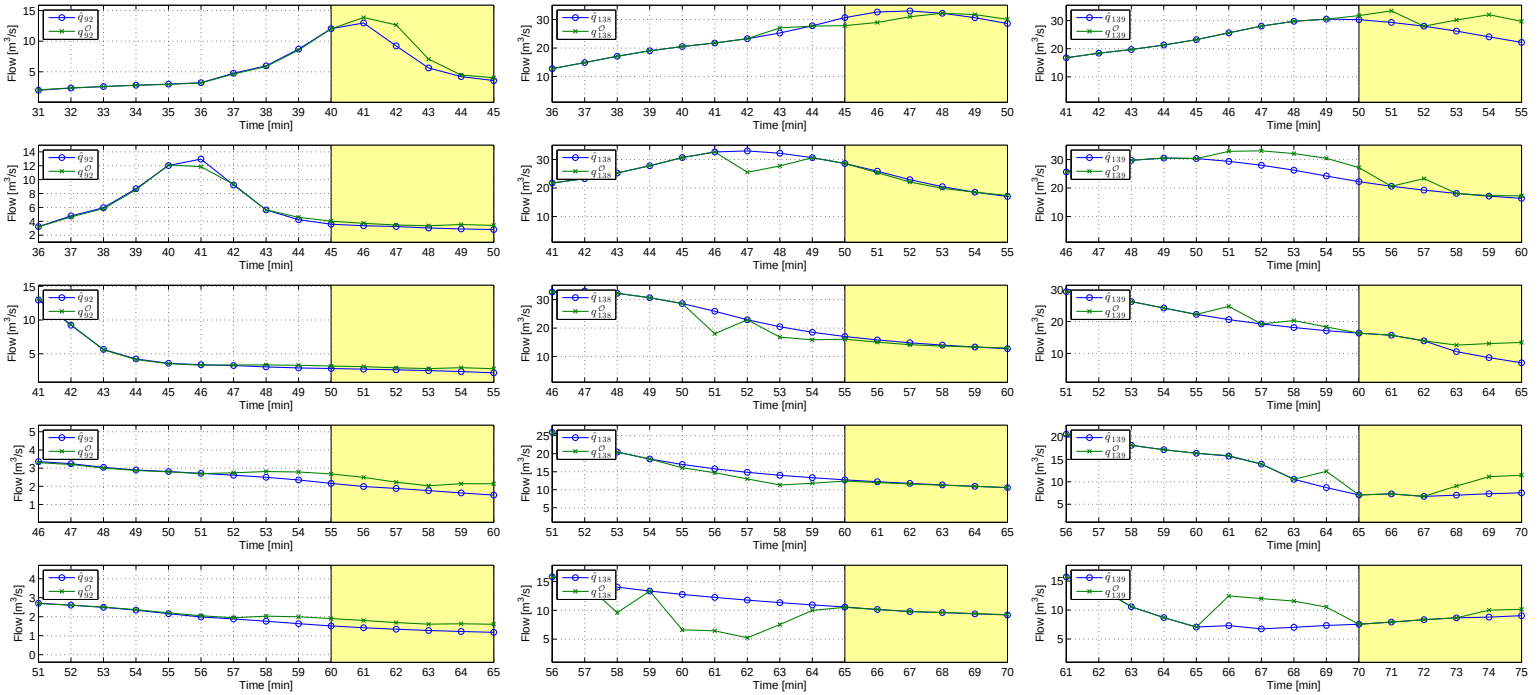

Figure D.8: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC2 scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

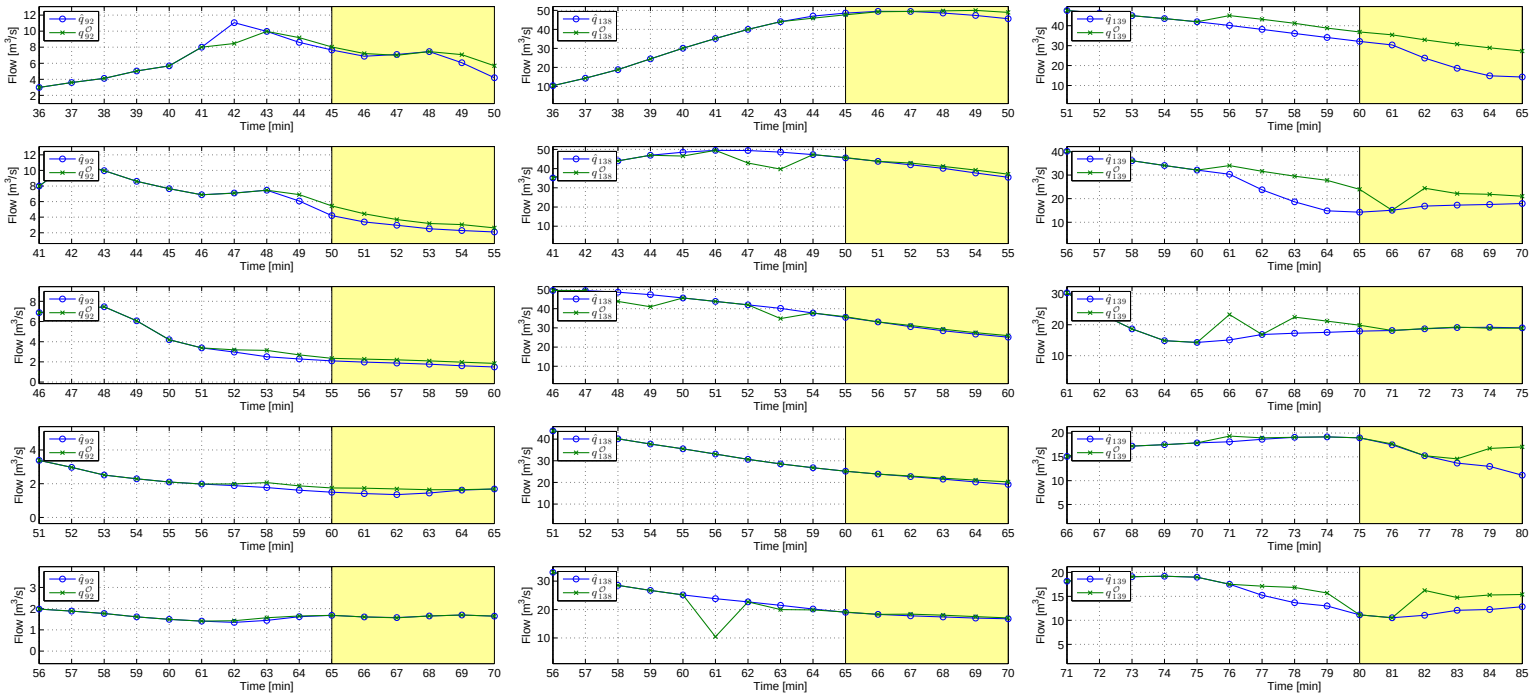

Figure D.9: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEF scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

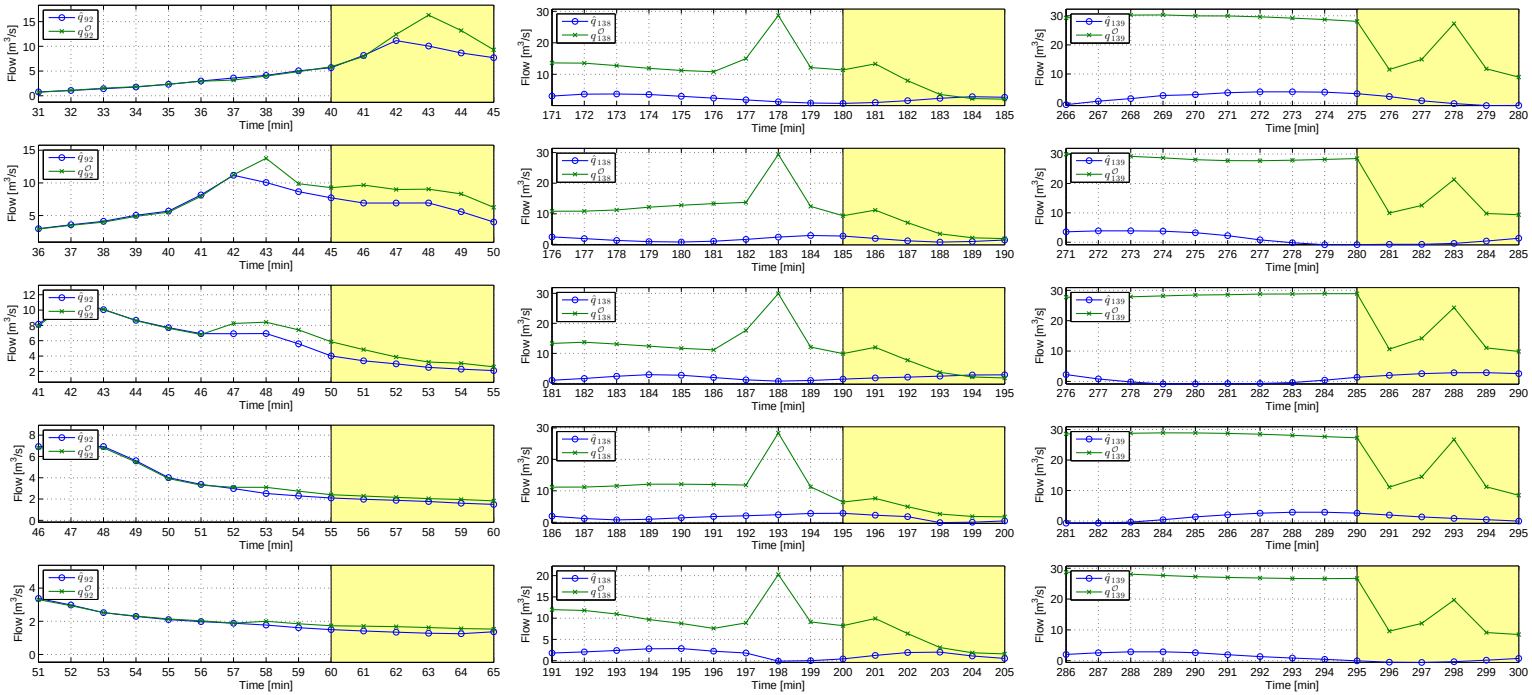

Figure D.10: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEL scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

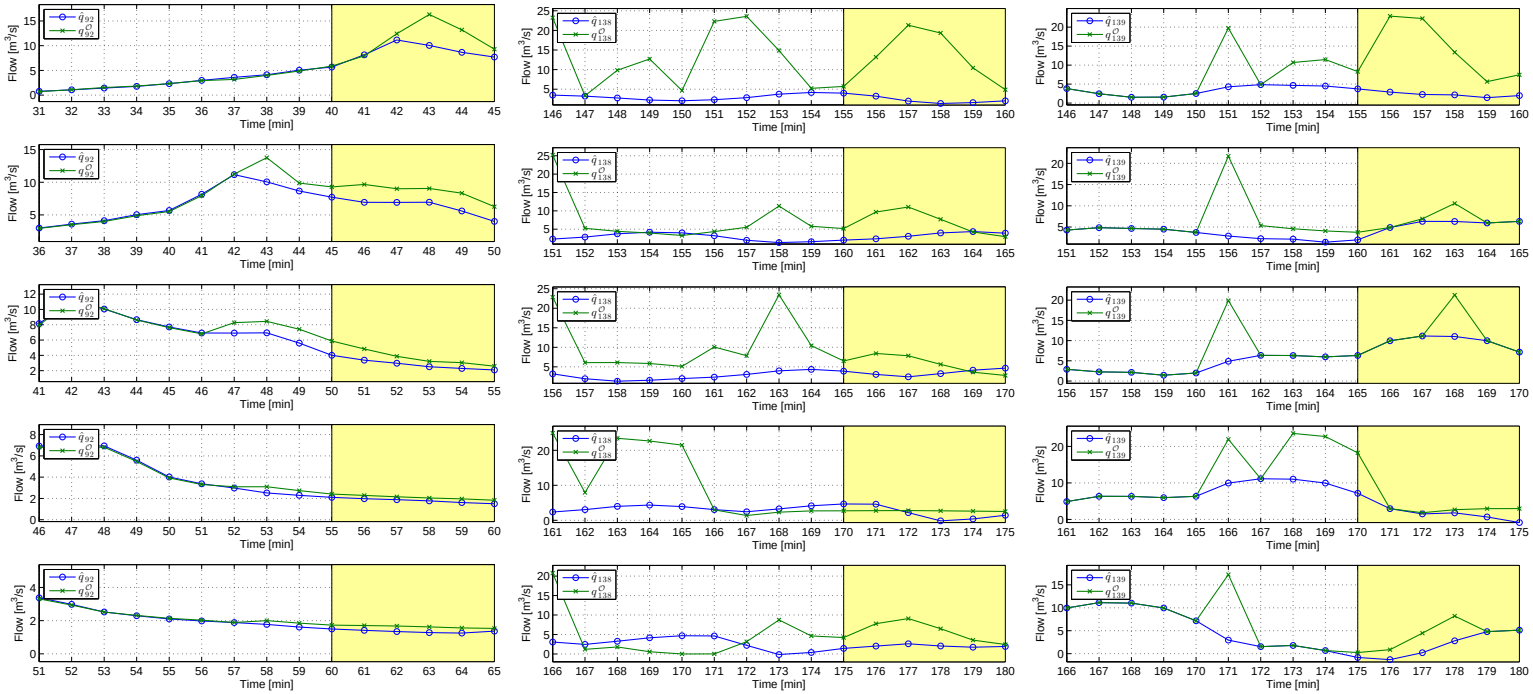

Figure D.11: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

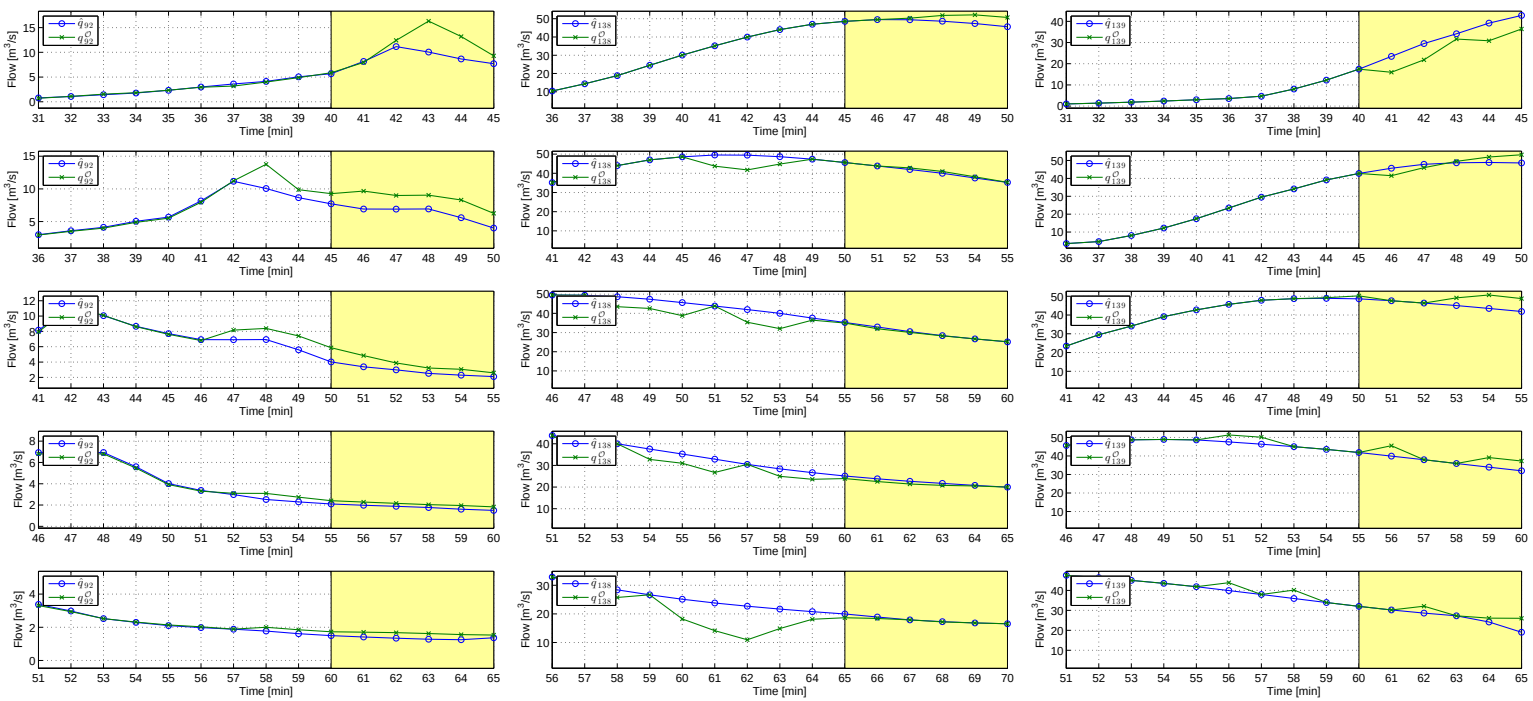

Figure D.12: SEP solution corresponding to sewer pipes q_{92}, q_{138} and q_{139} for several consecutive problems in the MHEC2 scenario. The first SEP of each sequence corresponds to the maximum absolute error obtained among all the solved SEPs.

Appendix E

Gate Flows and Setpoints Figures

Gate Flows and Set-points MHEF (15-08-2006)

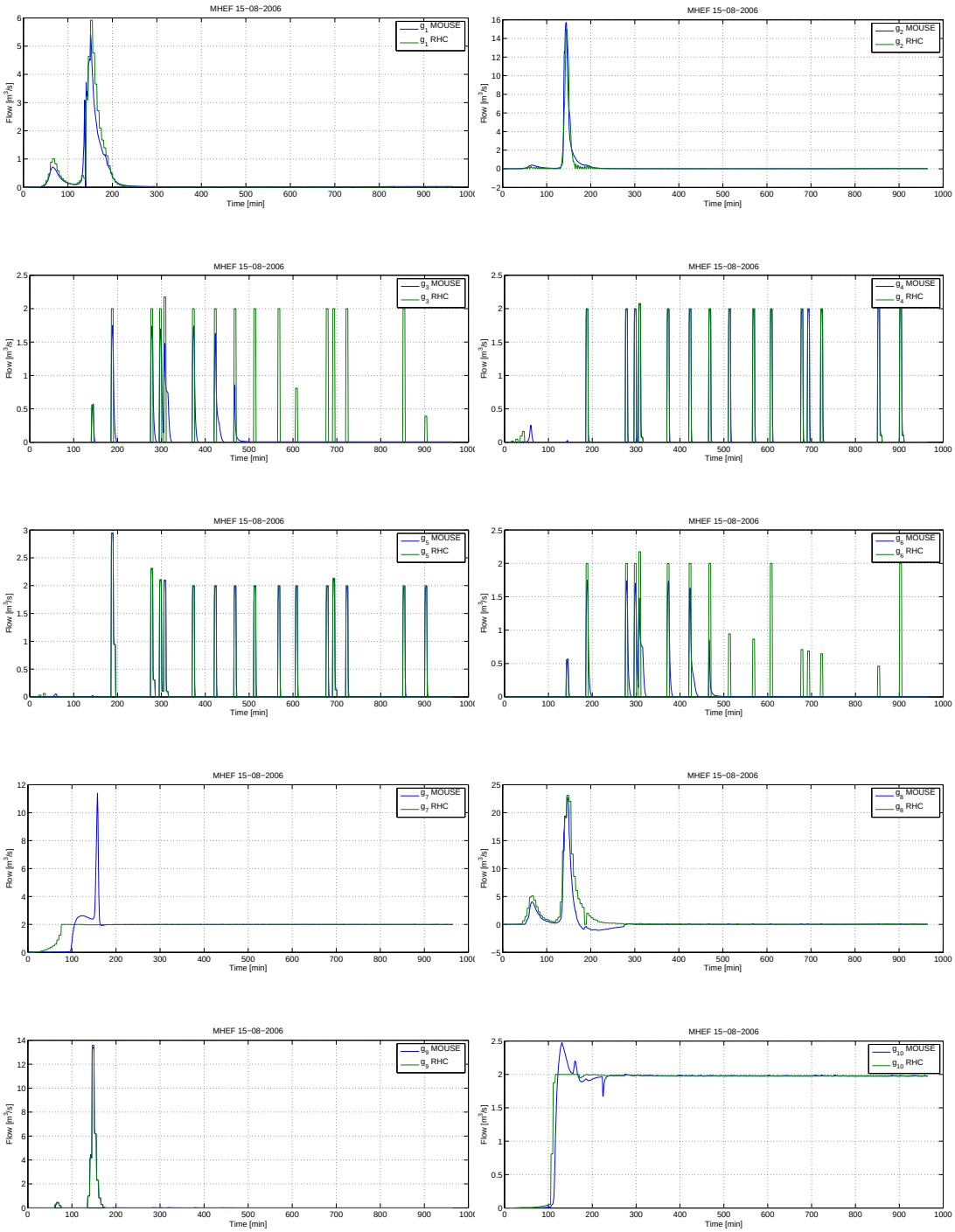

Figure E.1: Gate flows and set-points for the MHEF scenario 15-08-2006.

Gate Flows and Set-points MHEL (15-08-2006)

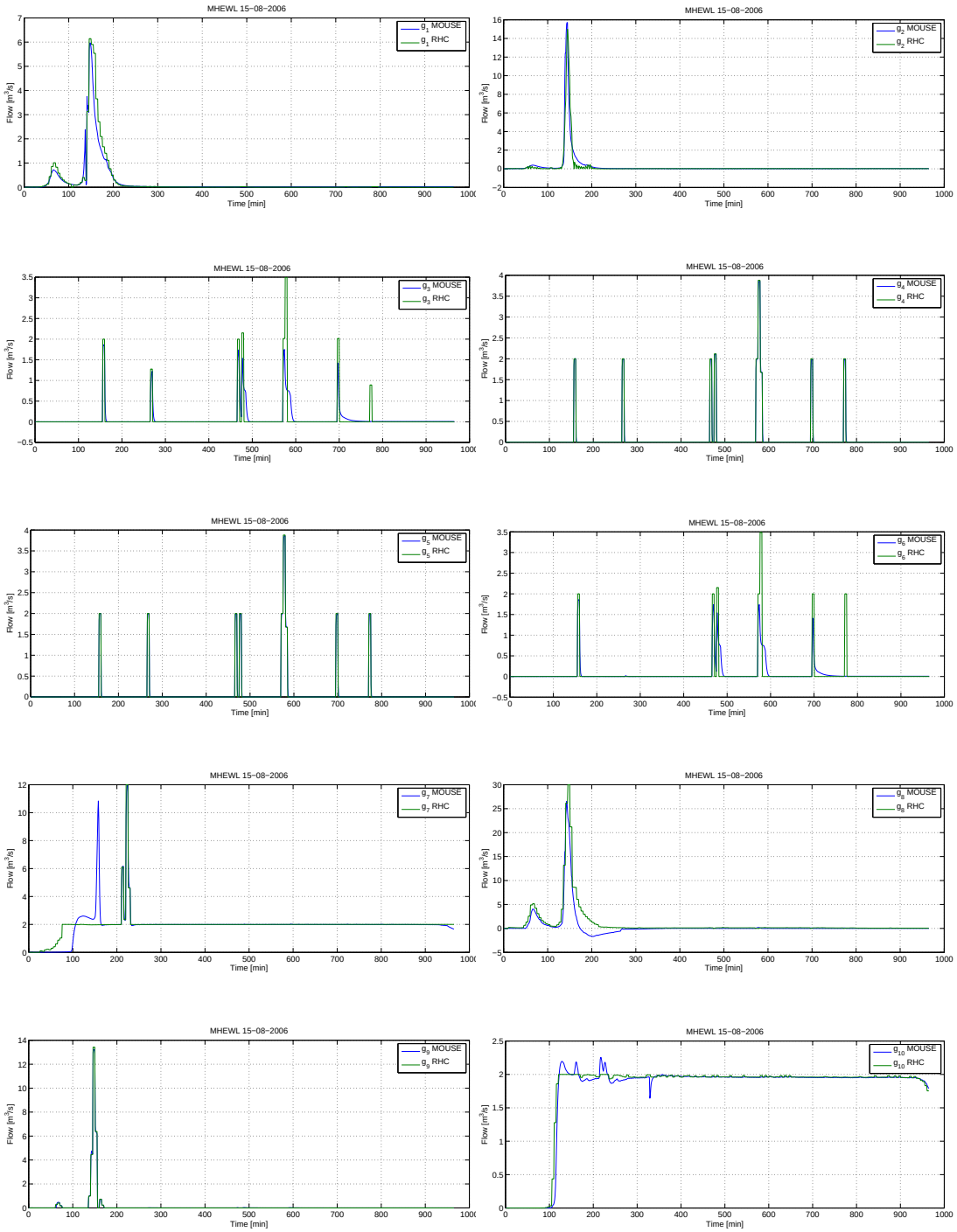

Figure E.2: Gate flows and set-points for the MHEL scenario 15-08-2006.

Gate Flows and Set-points MHEC (15-08-2006)

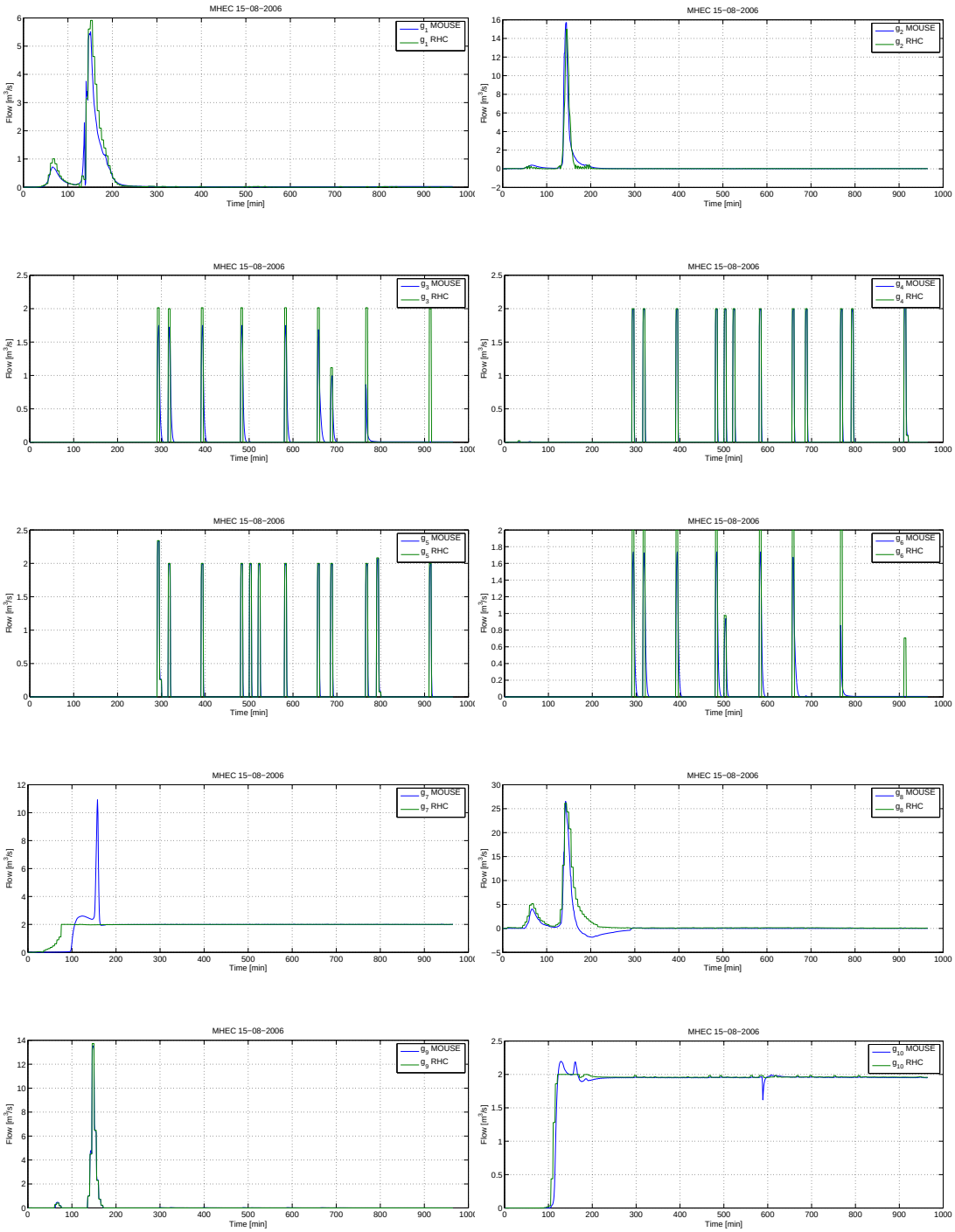

Figure E.3: Gate flows and set-points for the MHEC scenario 15-08-2006.

Gate Flows and Set-points MHEC2 (15-08-2006)

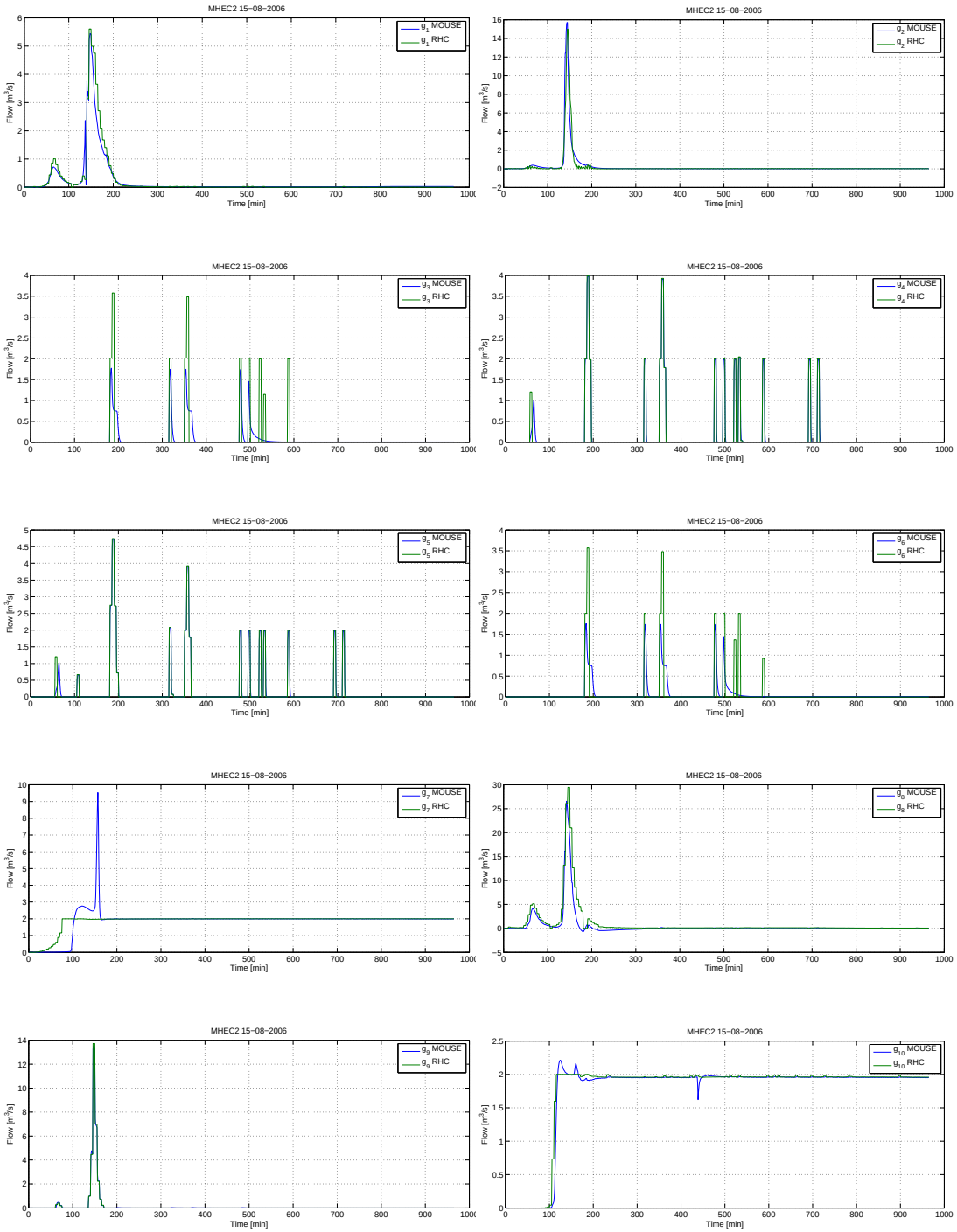

Figure E.4: Gate flows and set-points for the MHEC2 scenario 15-08-2006.

Gate Flows and Set-points MHEF (17-09-2002)

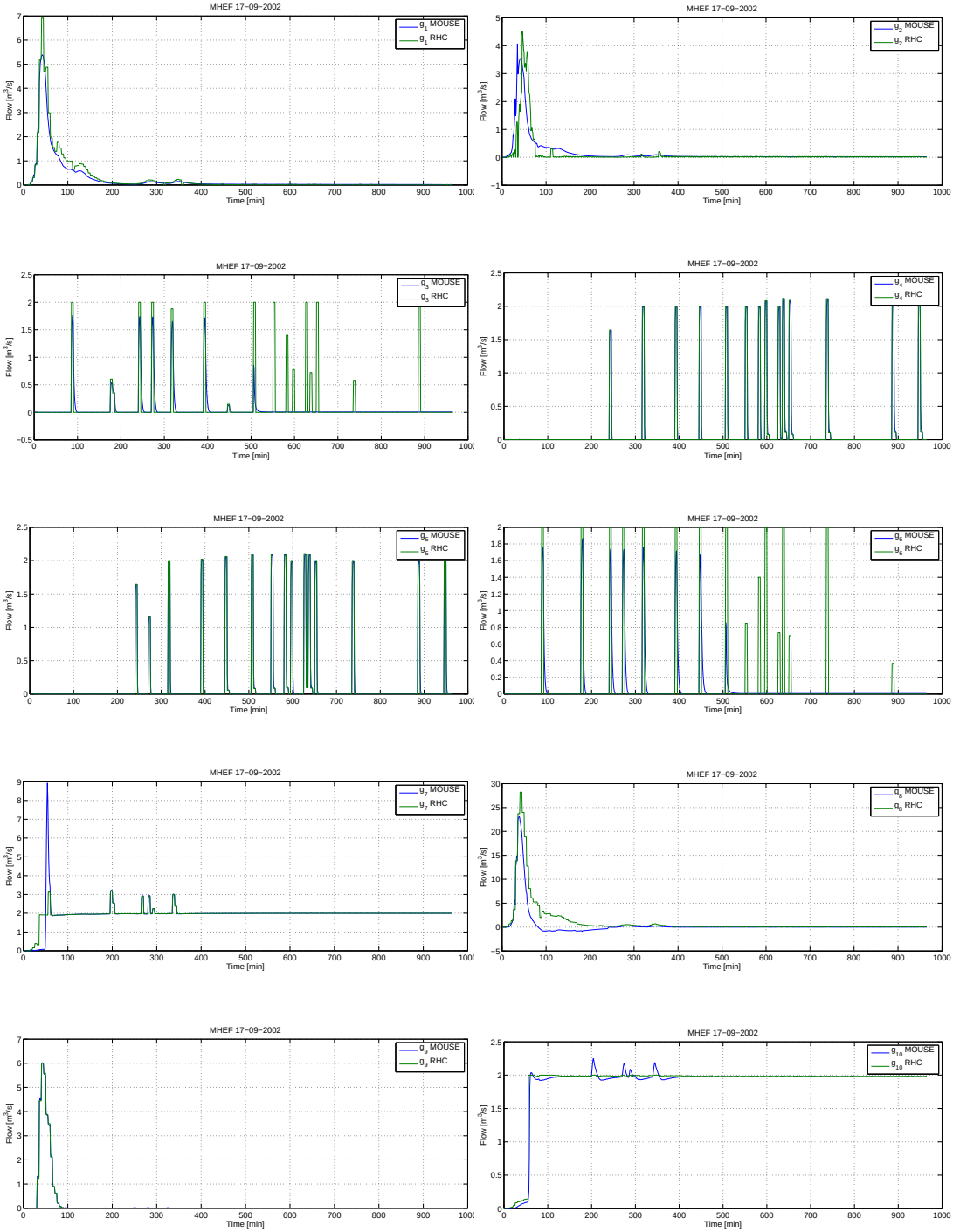

Figure E.5: Gate flows and set-points for the MHEF scenario 17-09-2002.

Gate Flows and Set-points MHEL (17-09-2002)

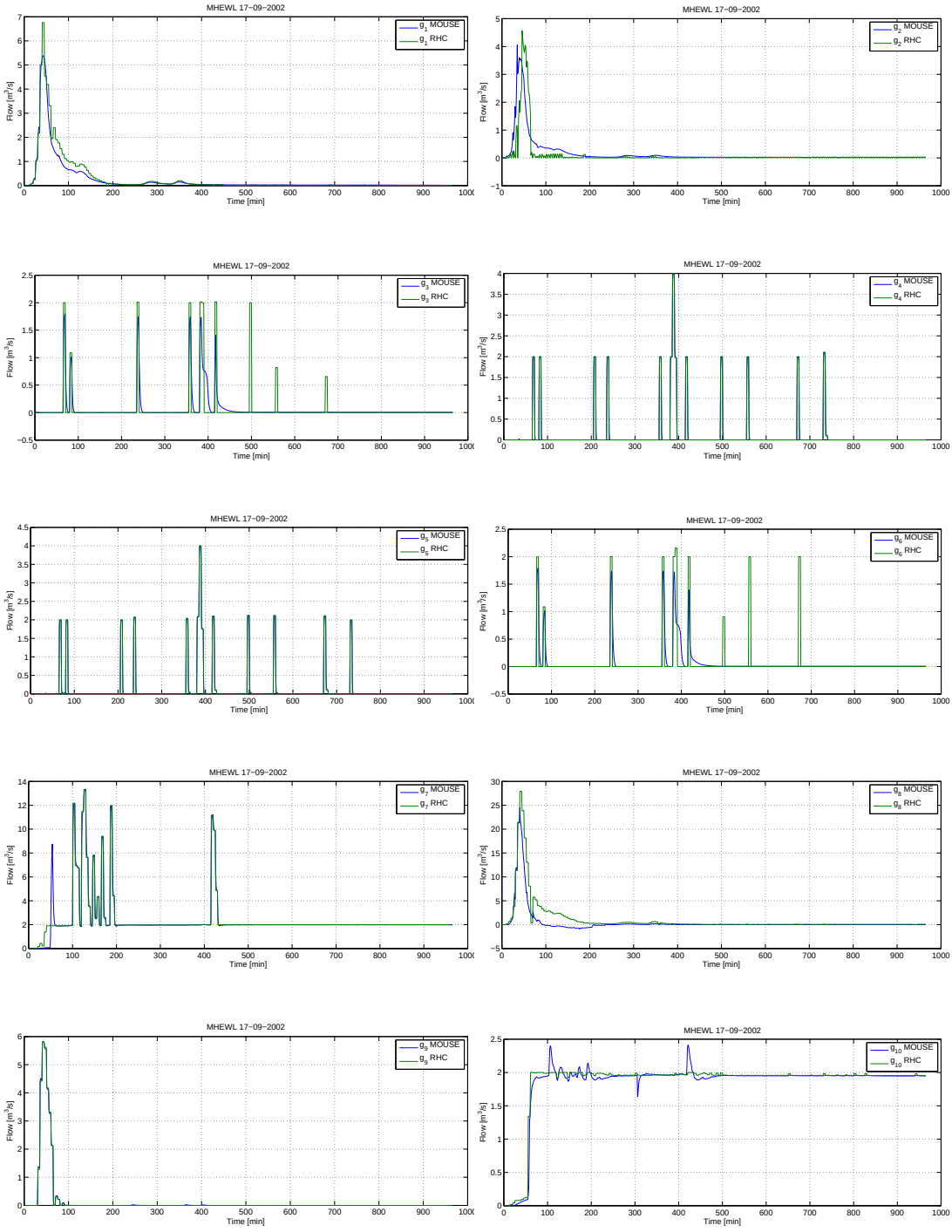

Figure E.6: Gate flows and set-points for the MHEL scenario 17-09-2002.

Gate Flows and Set-points MHEC (17-09-2002)

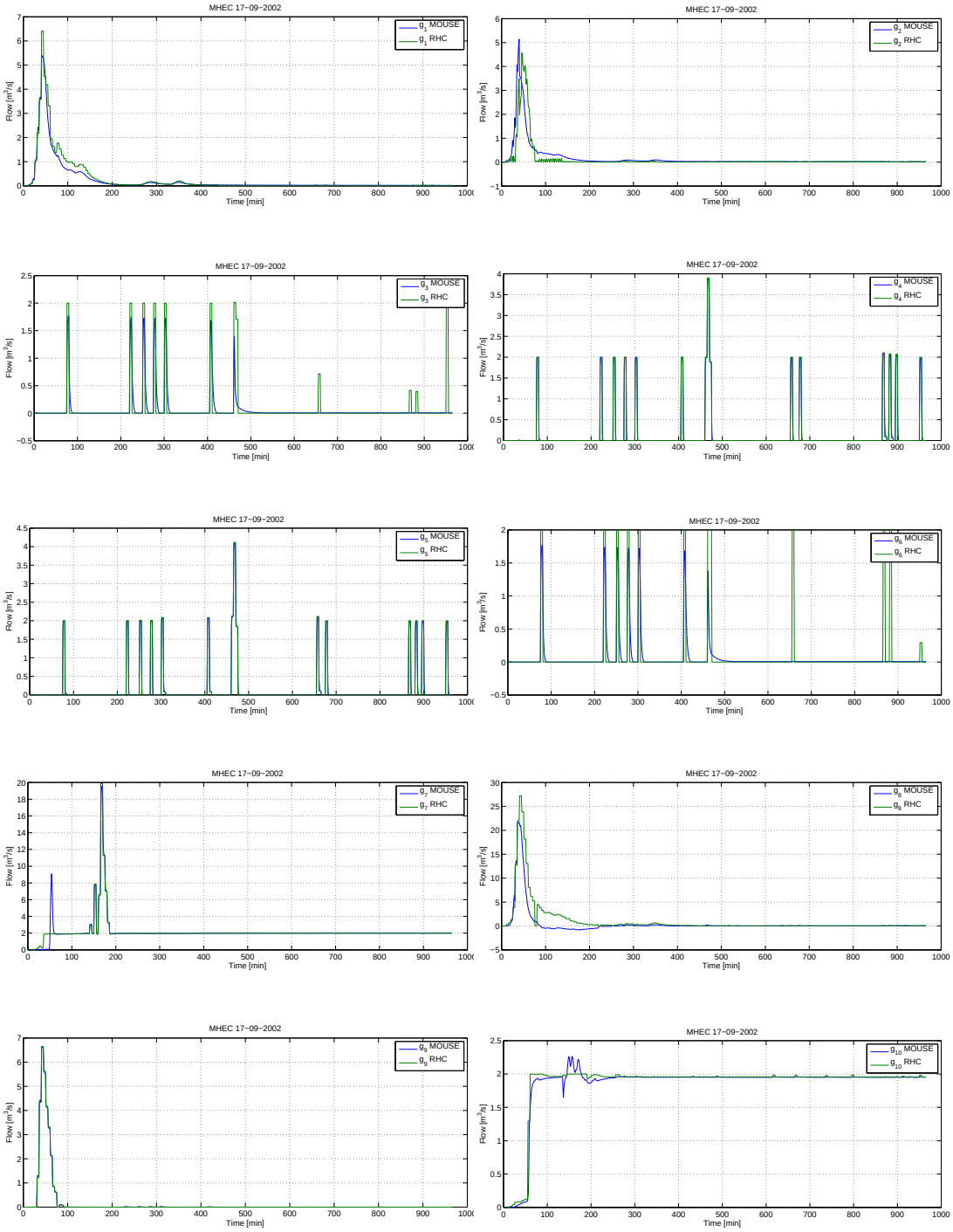

Figure E.7: Gate flows and set-points for the MHEC scenario 17-09-2002.

Gate Flows and Set-points MHEC2 (17-09-2002)

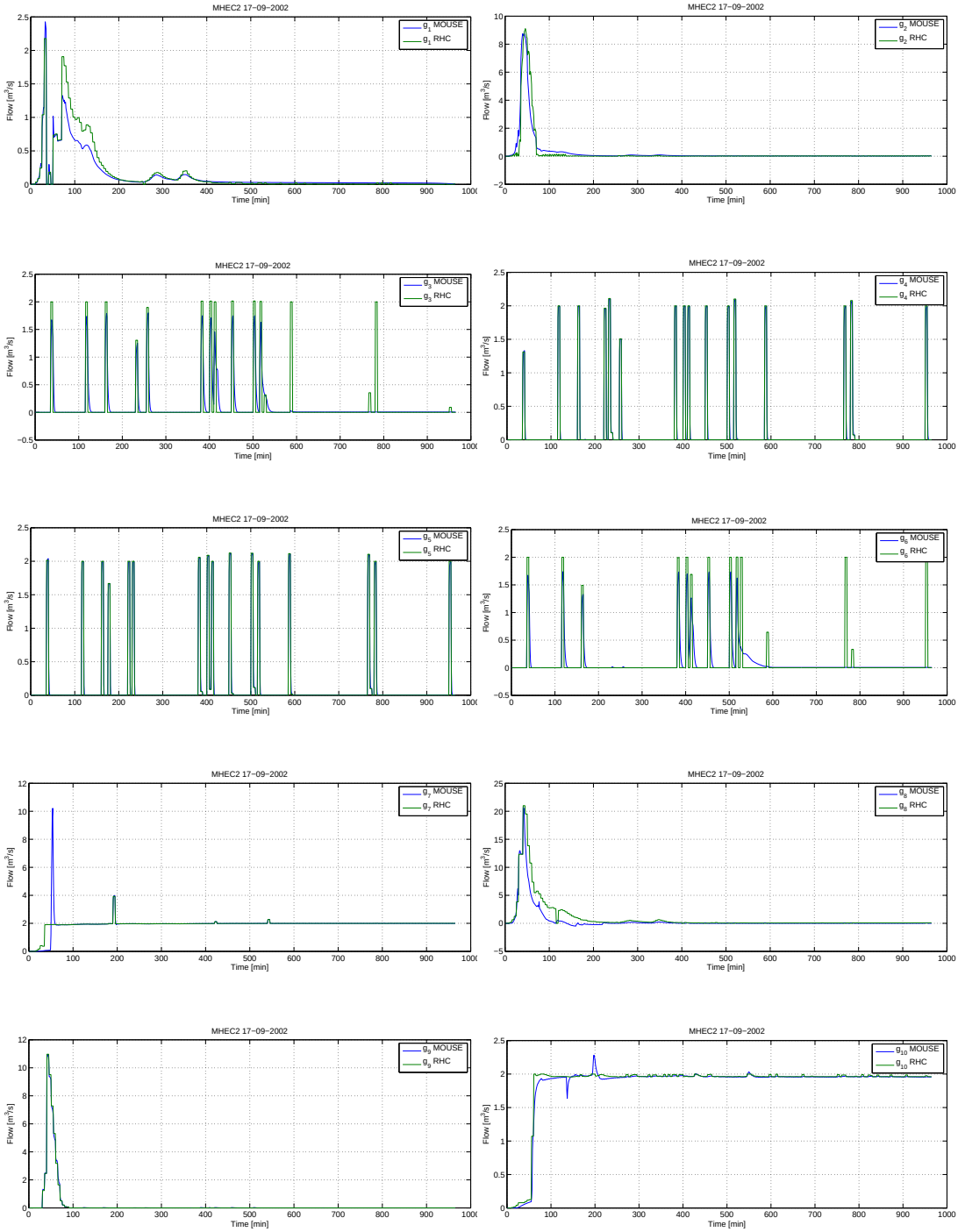

Figure E.8: Gate flows and set-points for the MHEC2 scenario 17-09-2002.

Gate Flows and Setpoints MHEF (30-07-2011)

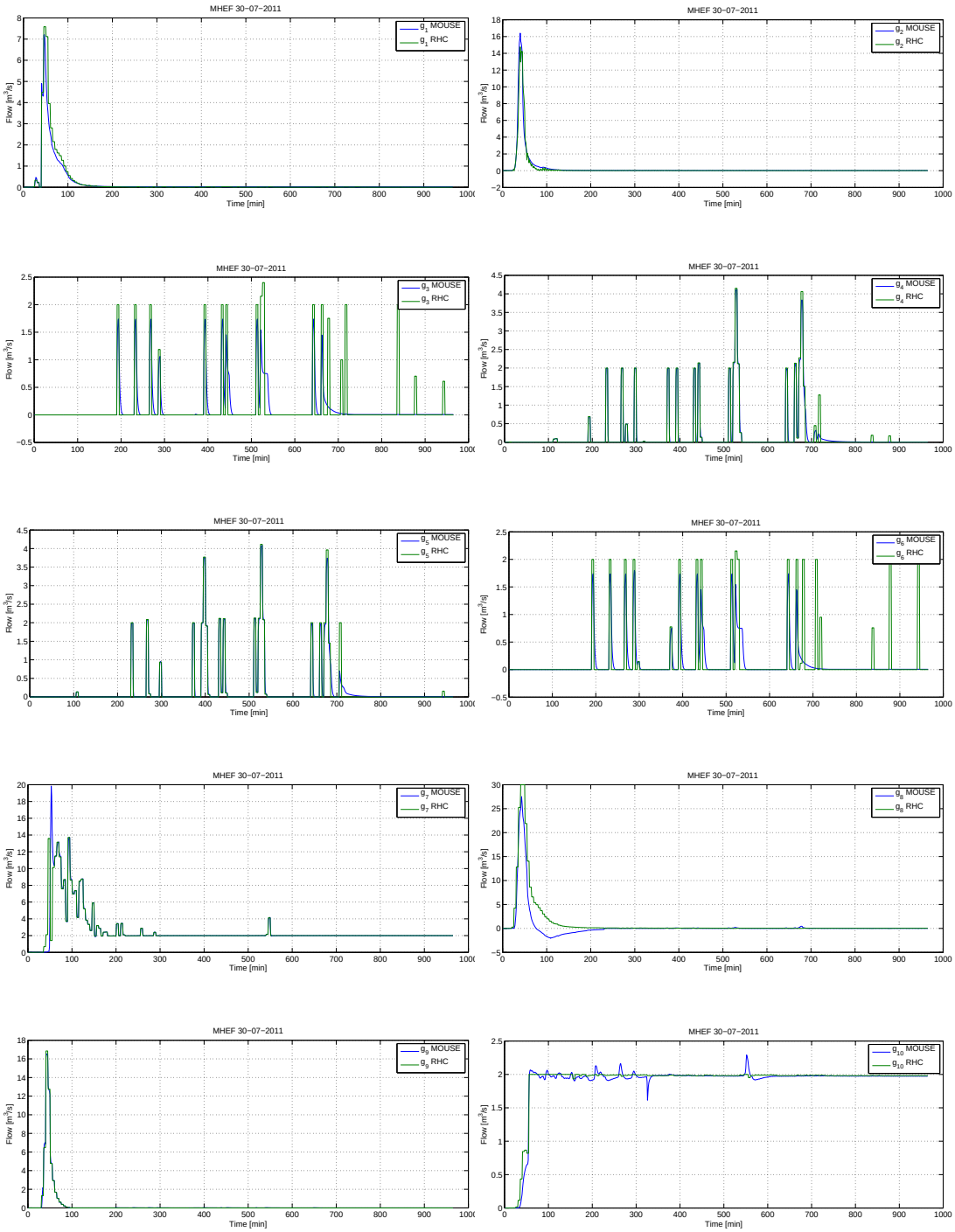

Figure E.9: Gate flows and set-points for the MHEF scenario 30-07-2011.

Gate Flows and Set-points MHEL (30-07-2011)

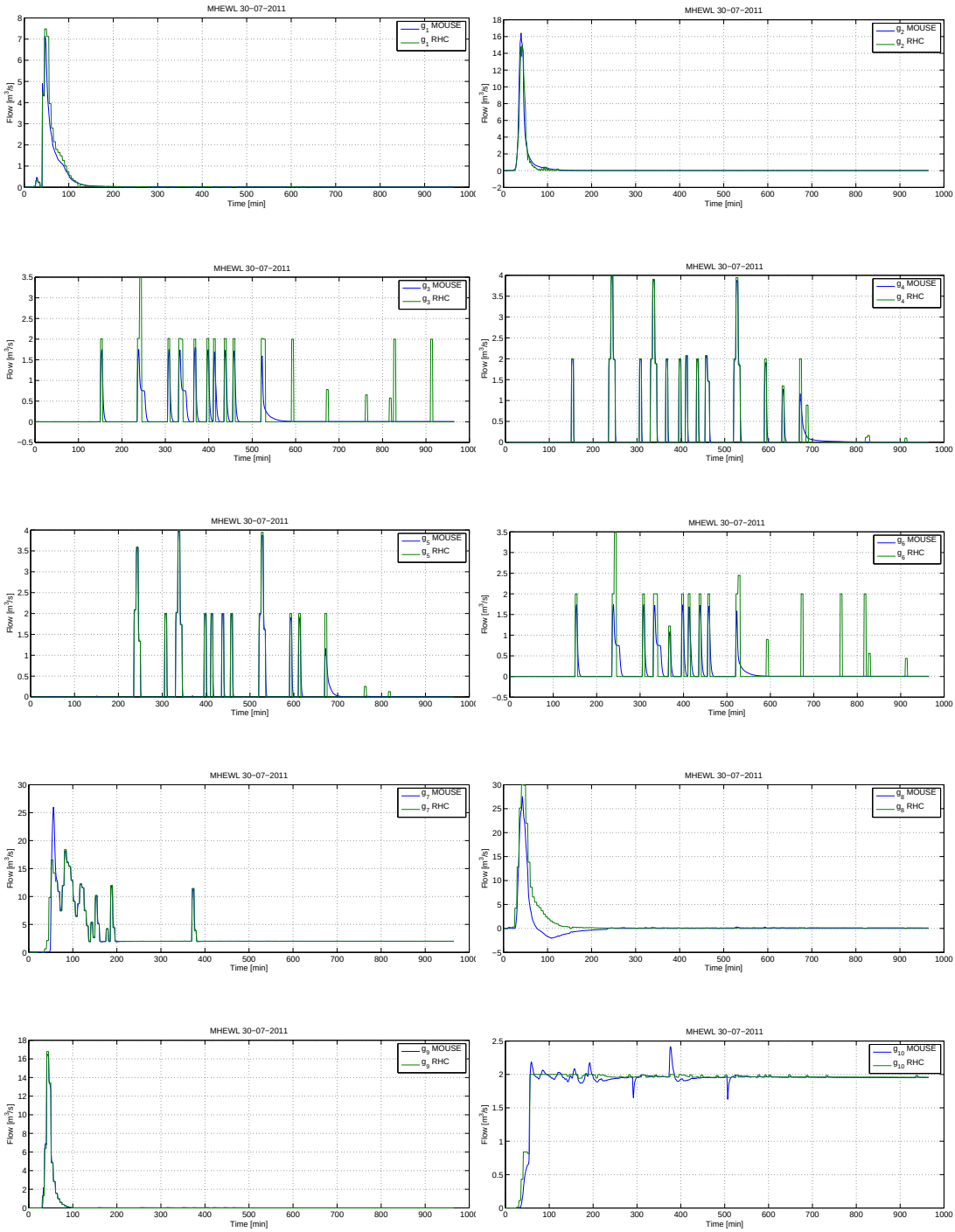

Figure E.10: Gate flows and set-points for the MHEL scenario 30-07-2011.

Gate Flows and Set-points MHEC (30-07-2011)

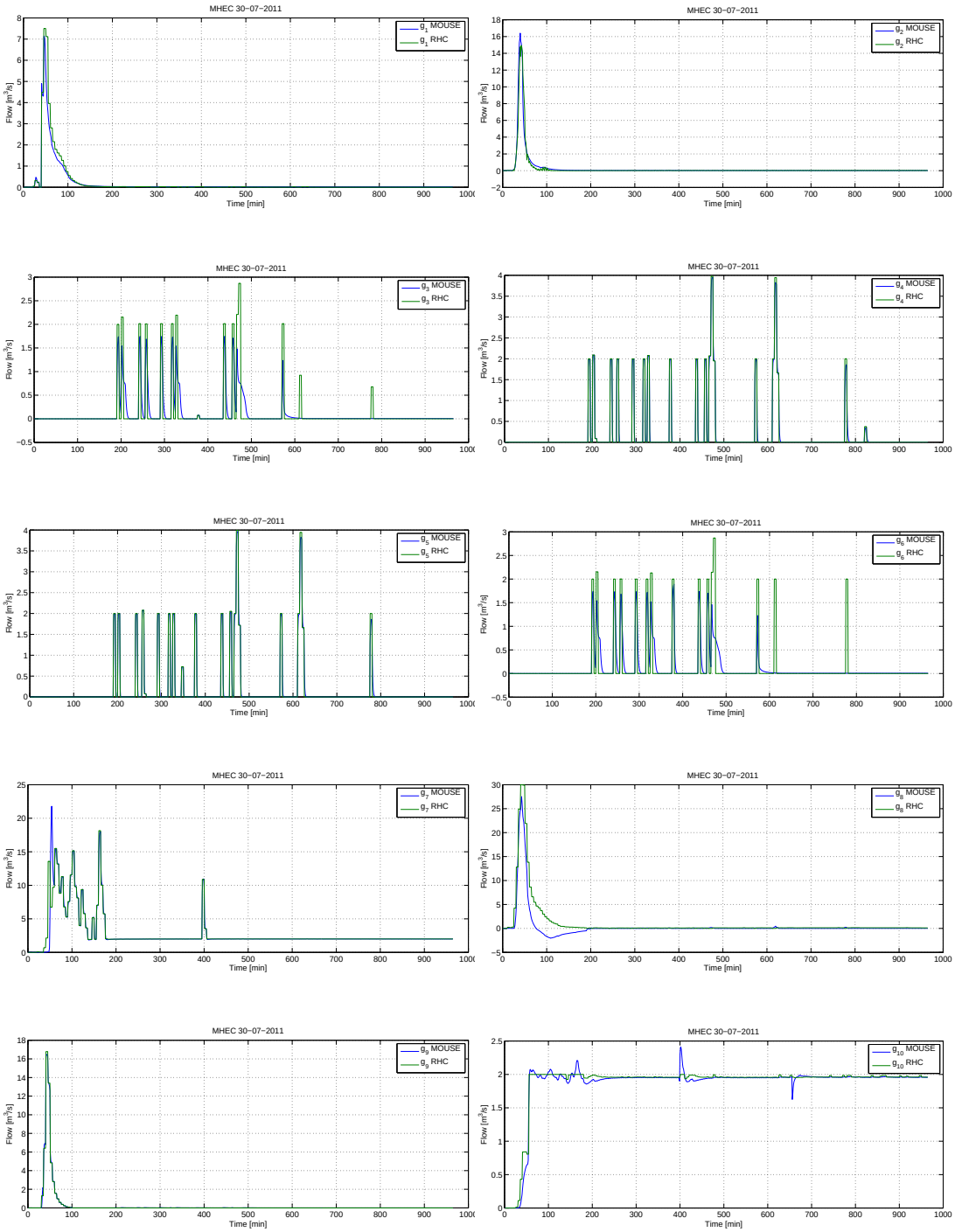

Figure E.11: Gate flows and set-points for the MHEC scenario 30-07-2011.

Gate Flows and Set-points MHEC2 (30-07-2011)

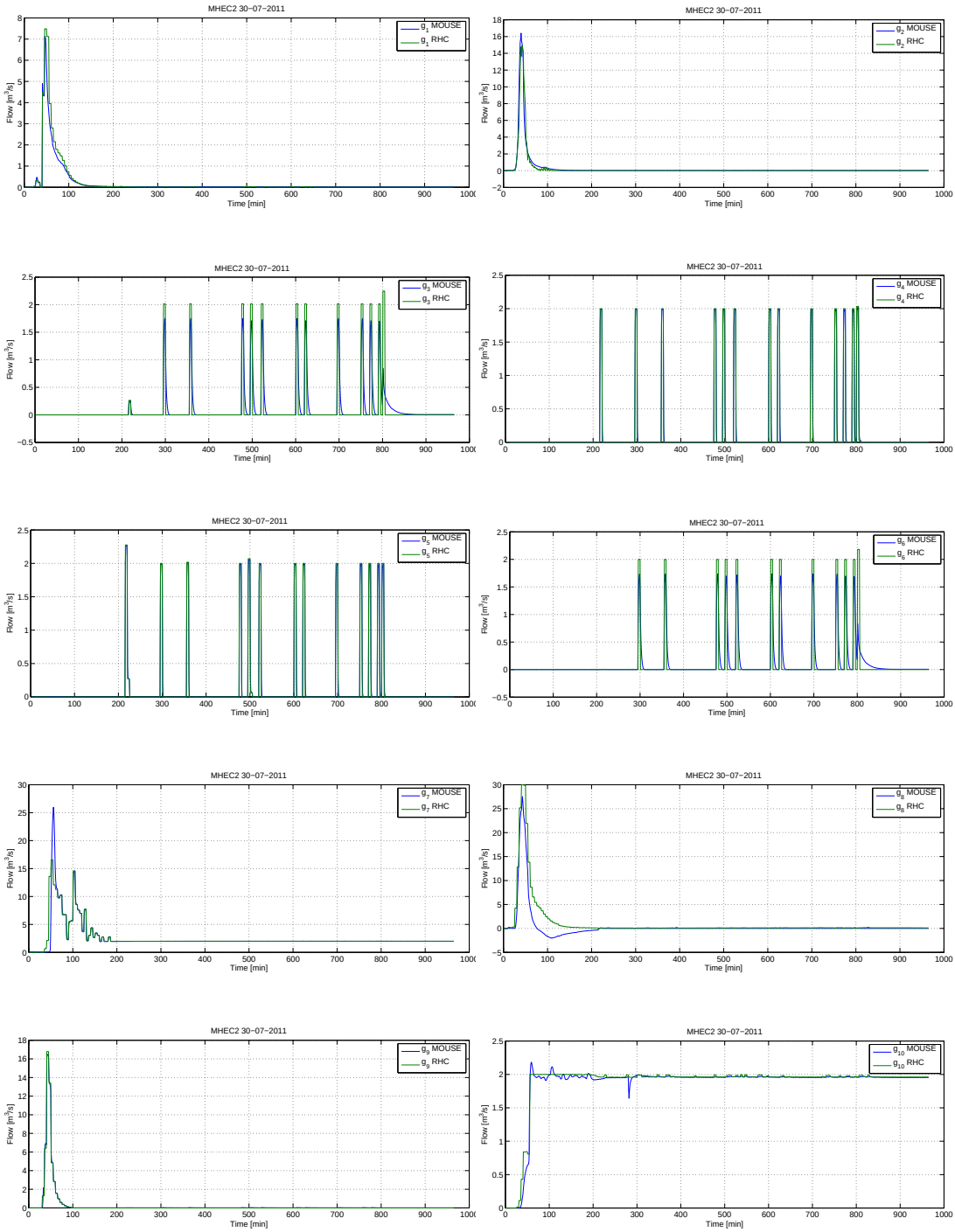

Figure E.12: Gate flows and set-points for the MHEC2 scenario 30-07-2011.

Bibliography

M.B. Abbott. *An Introduction to The Method of Characteristics*. Thames and Hudson, London, 1966.

A.O. Akan. *Open channel hydraulics*. Butterworth-Heinemann, Oxford, 2006.

P.M. Bach, W. Rauch, P.S. Mikkelsen, D.T. McCarthy, and A. Deletic. A critical review of integrated urban water modelling – Urban drainage and beyond. *Environmental Modeling & Software*, 54:88–107, 2014.

Ll. Ballester Rodés, J. Martí Marqués, and M. Salamero Sansalvadó. Control de compuertas de derivación en la red de alcantarillado de Barcelona. *Inegeniería del Agua*, 5(4):37–46, 1998.

J-P Baume, Jacques Sau, and P-O Malaterre. Modelling of irrigation channel dynamics for controller design. In *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*, volume 4, pages 3856–3861. IEEE, 1998.

T. Beeneken, V. Erbe, A. Messmer, C. Reder, R. Rohlfing, M. Scheer, M. Schuetze, B. Schumacher, M. Weilandt, and M. Weyand. Real time control (rtc) of urban drainage systems–a discussion of the additional efforts compared to conventionally operated systems. *Urban Water Journal*, 10(5):293–299, 2013.

A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427, 1999.

A. Bemporad, G. Ferrari-Trecate, and M. Morari. Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control*, 45 (10):1864–1876, 2000.

L. Benedetti, J. Langeveld, A. Comeau, L. Corominas, G. Daigger, C. Martin, P.S. Mikkelsen, L. Vezzaro, S. Weijers, and P.A. Vanrolleghem. Modelling and monitoring of integrated urban wastewater systems: review on status and perspectives. *Water Sci. Technol*, 68(6):1203–1215, 2013.

Keith J. Beven. *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2011.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

M. Breckpot, T. Blanco Barjas, and B. De Moor. Flood control of rivers with nonlinear model predictive control and moving horizon estimation. In *IEEE Conference on Decision and Control*, pages 6107–6112, Atlanta, USA, 2010.

J. Busch, P. Kühl, J.P. Schölder, H.G. Bock, and W. Marquardt. State estimation for large-scale wastewater treatment plants. *Water Research*, 47(13):4774–4787, 2013.

D. Butler and M. Schütze. Integrating simulation models with a view to optimal control of urban wastewater systems. *Environmental Modelling & Software*, 20(4):415–426, 2005.

E.F. Camacho and C. Bordons. *Model Predictive Control*. Springer, London, 2004.

A. Campisano, J. Cabot Ple, D. Muschalla, M. Pleau, and P.A. Vanrolleghem. Potential and limitations of modern equipment for real time control of urban wastewater systems. *Urban Water Journal*, 10(5):300–311, 2013.

G. Cembrano, J. Quevedo, M. Salamero, V. Puig, J. Figueras, and J. Martí. Optimal control of urban drainage systems. A case study. *Control Engineering Practice*, 12 (1):1–9, 2004.

H. Chanson. *The Hydraulics of Open Channel Flow: An Introduction*. Elsevier Butterworth-Heinemann, Oxford, 2004.

M. H. Chaudhry. *Open-Channel Flow*. Springer, New York, 2008.

V. T. Chow. *Open-Channel Hydraulics*. McGraw-Hill, New York, 1959.

CPLEX™. *version 12.5 (2012)*. IBM ILOG, Sunnyvale, California, 2011.

A. J. C. Barré de Saint-Venant. Théorie du mouvement non-permanent des eaux avec application aux crues des rivières et à l'introduction des marées dans leur lit. *Comptes Rendus des séances de l'Académie des Sciences*, 73:237–240, 1871.

G. Dirckx, M. Schütze, S. Kroll, Ch. Thoeye, G. De Gueldre, and B. Van De Steene. Cost-efficiency of rtc for cso impact mitigation. *Urban Water Journal*, 8(6):367–377, 2011a.

G. Dirckx, M. Schütze, S. Kroll, Ch. Thoeye, G. De Gueldre, and B. Van De Steene. RTC versus static solutions to mitigate CSO's impact. *12nd International Conference on Urban Drainage*, 2011b. Porto Alegre, Brazil.

S. Duchesne, A. Mailhot, E. Dequidt, and J.-P. Villeneuve. Mathematical modeling of sewers under surcharge for real time control of combined sewer overflows. *Urban Water*, 3(4):241–252, 2001.

S. Duchesne, A. Mailhot, and J.-P. Villeneuve. Predictive real time control of surcharged interceptors: impact of several control parameters. *Journal of the American Water Resources Association*, 39(1):125–135, 2003.

S. Duchesne, A. Mailhot, and J.-P. Villeneuve. Global predictive real-time control of sewers allowing surcharged flows. *ASCE Journal of Environmental Engineering*, 130 (5):526–534, 2004.

K. Eurén and E. Weyer. System identification of open water channels with undershot and overshot gates. *Control Engineering Practice*, 15(7):813–824, 2007.

G. Ferrari-Trecate, D. Mignone, and M. Morari. Moving horizon estimation for hybrid systems. *IEEE Transactions on Automatic Control*, 47(10):1663–1676, 2002.

R. Field, D. Sullivan, and A.N. Tafuri. *Management of combined sewer overflows*. CRC Press, 2004.

J. Figueras, G. Cembrano, V. Puig, J. Quevedo, M. Salamero, , and J. Martí. Coral off-line: an object-oriented tool for optimal control of sewer networks. In *IEEE International Symposium on Computer-Aided Control System Design*, pages 224–229, 2002.

O. Fradet, M. Pleau, A. Desbiend, and H. Colas. Theoretical field validation of solutions based on simplified hydraulic models for real-time control of sewer networks. *NOVATECH International Conference on Sustainable Techniques and Strategies for Urban Water Management*, 2010.

G. Fu, D. Butler, and S.-T. Khu. Multiple objective optimal control of integrated urban wastewater systems. *Environmental Modelling & Software*, 23(2):225–234, 2008.

G. Fu, D. Butler, and S.-T. Khu. The impact of new developments on river water quality from an integrated system modelling perspective. *Science of the total environment*, 407(4):1257–1267, 2009.

M.S Gelormino and N.L. Ricker. Model-predictive control of a combined sewer system. *Int. J. Control*, 59(3):793–816, 1994.

Ignacio E. Grossmann and Juan P. Ruiz. Generalized disjunctive programming: A framework for formulation and alternative algorithms for MINLP optimization. In Jon Lee and Sven Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumes in Mathematics and its Applications*, chapter 4, pages 93–116. 2012.

W. P. M. H. Heemels, B. De Schutter, and A. Bemporad. Equivalence of hybrid dynamical models. *Automatica*, 37(7):1085–1091, July 2001.

Mogens Henze. *Activated sludge models ASM1, ASM2, ASM2d and ASM3*. Number 9. IWA publishing, 2000.

IBM ILOG. IBM Academic Initiative. `http://www.ibm.com/academicinitiative`, 2013.

L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis*. Computational Intelligence and Complexity. Springer-Verlag, London, 2001.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Receding horizon control of hybrid linear delayed systems: Application to sewer networks. *IEEE Conference on Decision and Control*, 2013a. Firenze, Italy.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. A control-oriented hybrid modelling approach for sewer networks: Barcelona case study. *IWA Conference on Instrumentation, Automation and Control*, 2013b. Narbonne, France.

B. Joseph-Duran, M. Jung, C. Ocampo-Martinez, S. Sager, and G. Cembrano. Minimization of sewage network overflow. *Water Resources Management*, 28(1):41–63, 2014a.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Output-feedback control of sewer networks thorugh moving horizon estimation. *IEEE Conference on Decision and Control*, 2014b. Los Angeles, USA.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid control-oriented modeling of combined sewer networks: Barcelona case study. *Hydroinformatics Conference*, 2014c. New York, USA.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. Hybrid modeling and receding horizon control of sewer networks. *Water Resources Research*, 2014d.

J. Lau, D. Butler, and M. Shütze. Is combined overflow spill frequancy/volume a good indicator of receiving water quality impact? *Urban Water*, 4:181–189, 2002.

X. Litrico and V. Fromion. Frequency modeling of open channel flow. *Journal of Hydraulic Engineering*, 130(8):806–815, 2004a.

X Litrico and V Fromion. Simplified modeling of irrigation canals for controller design. *Journal of Irrigation and Drainage Engineering*, 130(5):373–383, 2004b.

X. Litrico and V. Fromion. *Modelling and Control of Hydrosystems*. Springer, London, 2009.

X. Litrico and D. Georges. Robust continuous-time and discrete-time flow control of a dam-river system. (I) Modelling. *Applied Mathematical Modelling*, 23(11):809–827, 1999.

X. Litrico and J.-B. Pomet. Nonlinear modelling and control of a long river stretch. *European Control Conference*, 2003.

X. Litrico, J.-B. Pomet, and V. Guinot. Simplified nonlinear modeling of river flow routing. *Advances in Water Resources*, 33(9):1015–1023, 2010.

L. Ljung. *System Identification: Theory for the User, 2nd Ed.* Prentice Hall, New Jersey, 1999.

J.M. Maciejowski. *Predictive Control with Constraints.* Prentice Hall, New Jersey, 2002.

P-O Malaterre and J-P Baume. Modeling and regulation of irrigation canals: existing applications and ongoing researches. In *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*, volume 4, pages 3850–3855. IEEE, 1998.

M. Marinaki and Papageorgiou. Nonlinear optimal flow control for sewer networks. *Proc. American Control Conference*, 2:1289–1293, 1998.

M. Marinaki and Papageorgiou. Rolling-horizon optimal control of sewer networks. *Proc. IEEE International Conference on Control Applications*, pages 594–599, 2001.

M. Marinaki and Papageorgiou. Linear-quadratic regulators applied to sewer network flow control. *Proc. European Control Conference*, pages 1–4, 2003.

M. Marinaki and M. Papageorgiou. *Optimal Real-time Control of Sewer Networks.* Springer, London, 2005.

J. Meirlaen. *Immission based real-time control of the integrated urban wastewater system.* PhD thesis, Ghent University, Belgium, 2002.

J. Meirlaen, B. Huyghebaert, L. Benedetti, and P. Vanrolleghem. Fast, simultaneous simulation of the integrated urban wastewater system using mechanistic surrogate models. *Water Science and Technology*, 43(7):301–309, 2001.

J. Meirlaen, J. Van Assel, and P.A. Vanrolleghem. Real time control of the integrated urban wastewater system using simultaneously surrogate models. *Water Science and Technology*, 45(3):109–116, 2002.

H. Michalska and D.Q. Mayne. Moving horizon observers and observer based control. *IEEE Transactions on Automatic Control*, 40(6):995–1006, 1995.

MOUSE. *MOUSE Pipe Flow*. DHI Software, 2007a.

MOUSE. *MOUSE Surface Runoff Models*. DHI Software, 2007b.

MOUSE. *MOUSE User Guide*. DHI Software, 2007c.

J.E. Nash. The form of the instantaneous unit hydrograph. *International Association of Scientific Hydrology, Publ.*, 3:114–121, 1957.

C. Ocampo-Martínez. *Model predictive control of wastewater systems*. Advances in industrial control. Springer, London, 2011.

C. Ocampo-Martínez and V. Puig. Piece-wise linear functions-based model predictive control of large-scale sewage systems. *Control Theory and Applications*, 4(9):1581–1593, 2010.

C. Ocampo-Martínez, A. Bemporad, A. Ingimundarson, and V. Puig. On hybrid model predictive control of sewer networks. In *Identification and Control: The Gap between Theory and Practice*, chapter 4, pages 87–114. Springer London, 2007.

C. Ocampo-Martínez, V. Puig, G. Cembrano, and J. Quevedo. Application of predictive control strategies to the management of complex networks in the urban water cycle [applications of control]. *IEEE Control Systems*, 33(1):15–41, 2013.

M. Pleau, H. Colas, P. Lavallée, G. Pelletier, and R. Bonin. Global optimal real-time control of the Quebec urban drainage system. *Environmental Modelling & Software*, 20(4):401–413, 2005.

M. Pleau, O. Fradet, H. Colas, and Ch. Marcoux. Giving the rivers back to the public. ten years of real time control in quebec city. *NOVATECH International Conference on Sustainable Techniques and Strategies for Urban Water Management*, 2010.

F. Previdi, M. Lovera, and S. Mambretti. Identification of the rainfall–runoff relationship in urban drainage networks. *Control Engineering Practice*, 7(12):1489–1504, 1999.

R. Price. Hydroinformatics and urban drainage: an agenda for the beginning of the 21st century. *Journal of Hydroinformatics*, 2:133–147, 2000.

V. Puig, G. Cembrano, J. Romera, J. Quevedo, B. Aznar, G. Ramón, and J. Cabot. Predictive optimal control of sewer networks using CORAL tool: application to Riera Blanca catchment in Barcelona. *Water Science and Technology*, 60(4):869–878, 2009a.

V. Puig, J. Romera, J. Quevedo, C.M. Cardona, A. Salterain, E. Ayesa, I. Irizar, A. Castro, M Lujan, P. Charbonnaud, P. Chiron, and J.-L. Troubat. Optimal predictive control of water transport systems: Arrêt-darré/arros case study. *Water Science and Technology*, 60(8):2125–2133, 2009b.

C.V. Rao, J.B. Rawlings, and J.H. Lee. Constrained linear state estimation–a moving horizon approach. *Automatica*, 37:1619–1628, 2001.

W. Rauch and P. Harremöes. Genetic algorithms in real time control applied to minimize transient pollution from urban wastewater systems. *Water Resources*, 33(5): 1265–1277, 1999a.

W. Rauch and P. Harremöes. On the potential of genetic algorithms in urban drainage modeling. *Urban Water*, 1(1):79–89, 1999b.

W. Rauch, J-L. Bertrand-Krajewski, P. Krebs, O. Mark, W. Schilling, M. Schütze, and P.A. Vanrolleghem. Mathematical modelling of integrated urban drainage systems. *Water science and technology*, 45(3):81–94, 2002.

J.B. Rawlings and D.Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, 2009.

Peter Reichert. *River water quality model*. Number 1. IWA Publishing, 2001.

M.R. Schütze, D. Butler, and M.B. Beck. *Modelling, Simulation and Control of Urban Wastewater Systems*. Springer, London, 2002.

J. Schuurmans, O.H. Bosgra, and Brouwer R. Open-channel flow model approximation for controller design. *Applied Mathematical Modelling*, 19(9):525–530, 1995.

J. Schuurmans, A.J. Clemmens, S. Dijkstra, A. Hof, and R. Brouwer. Modeling of irrigation and drainage canals for controller design. *Journal of irrigation and drainage engineering*, 125(6):338–344, 1999.

D. Schwanenberg, G. Verhoeven, and L. Raso. Nonlinear model predictive control of water resources systems in operational flood forecasting. *55th Internationales Wessenchaftliches Kolloquium*, 2010.

M. Schütze, A. Campisano, H. Colas, W. Schilling, and P.A. Vanrolleghem. Real time control of urban wastewater systems - where do we stand today? *Journal of Hydrology*, 299(3–4):335–348, 2004.

A.M. Solvi. *Modelling the sewer-treatment-urban river system in view of the EU Water Framework Directive*. PhD thesis, Ghent University, Belgium, 2006.

J. Szilagyi. State-space discretization of the Kalinin–Milyukov–Nash–cascade in a sample data system framework for streamflow forecasting. *Journal of Hydrologic Engineering*, 8(6):339–347, 2003.

J. Szilagyi. Discrete state-space approximation of the continuous Kalinin–Milyukov–Nash cascade of noninteger storage elements. *Journal of Hydrology*, 328(1–2):132–140, 2005.

A. Szöllösi-Nagy. The discretization of the continuous linear cascade my means of state space analysis. *Journal of Hydrology*, 58(3–4):223–236, 1982.

USEPA. Real Time Control of Urban Drainage Networks. Report 600-R-06-120, 2006.

P.J. van Overloop. *Model Predictive Control on Open Water Systems*. PhD thesis, Delft University, Netherlands, 2006.

P.J. van Overloop, J. Schuurmans, R. Brouwer, and C.M. Burt. Multiple-model optimization of proportional integral controllers on canals. *Journal of irrigation and drainage engineering*, 131(2):190–196, 2005.

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

P.A. Vanrolleghem, L. Benedetti, and J. Meirlaen. Modelling and real-time control of the integrated urban wastewater system. *Environmental Modelling & Software*, 20 (4):427–442, 2005.

L. Vezzaro and M. Grum. A generalized dynamic overflow risk assessment (dora) for urban drainage rtc. In *9th International Conference on Urban Drainage Modelling*, 2012. Belgrade, Serbia.

L. Vezzaro and M. Grum. A generalised dynamic overflow risk assessment (dora) for real time control of urban drainage systems. *Journal of Hydrology*, 515(0):292 – 303, 2014.

L. Vezzaro, M.L. Christensen, C. Thirsing, M. Grum, and P.S. Mikkelsen. Water quality-based real time control of integrated urban drainage systems: A preliminary study from copenhagen, denmark. *Procedia Engineering*, 70:1707–1716, 2014.

E. Weyer. System identification of an open channel. *Control Engineering Practice*, 9 (12):1289–1299, 2001.

H.P. Williams. *Model building in mathematical programming.* Wiley, New York, 1999. ISBN 0471997889.

M. Xu. *Real-time Control of Combined Water Quantity & Quality in Open Channels.* PhD thesis, Delft University, Netherlands, 2013.

M. Xu, P.J. van Overloop, and N.J. van de Giesen. On the study of control effectiveness and computational efficiency of reduced sint-venant model in model predictive control of open channel flow. *Advances in Water Resources*, 34(2):282–290, 2011.

M. Xu, R.R. Negenborn, P.J. van Overloop, and N.J. van de Giesen. De saint-venant equations-based model assessment in model predictive control of open channel flow. *Advances in Water Resources*, 49:37–45, 2012.