UNIVERSITY OF NATAL

OPTIMAL OPERATION OF A WATER DISTRIBUTION NETWORK BY PREDICTIVE CONTROL USING MINLP

Cédric P. G. BISCOS

OPTIMAL OPERATION OF A WATER DISTRIBUTION NETWORK BY PREDICTIVE CONTROL USING MINLP

Cédric P. G. BISCOS

Submitted in partial fulfilment of the academic requirements for the degree of Master of Science in Engineering in the School of Chemical Engineering at the University of Natal, Durban.

Supervisor : Professor Michael Mulholland

Durban, September 03

ABSTRACT

THE objective of this research project is to develop new software tools capable of operational optimisation of existing, large-scale water distribution networks. Since pumping operations represent the main operating cost of any water supply scheme, the optimisation problem is equivalent to providing a new sequence for pumping operations that makes better use of the different electricity tariff structures available to the operators of distribution systems. The minimisation of pumping costs can be achieved by using an optimal schedule that will allow best use of gravitational flows, and restriction of pumping to low-cost power periods as far as possible.

A secondary objective of the operational optimisation is to maintain the desired level of disinfectant chlorine at the point of delivery to consumers. There is a steady loss of chlorine with residence time in the system. If the level drops too low there is a risk of bacterial activity. Re-dosage points are sometimes provided in the network. Conversely, too high a level produces an unacceptable odour.

The combination of dynamic elements (reservoir volumes and chlorine concentration responses) and discrete elements (pump stati and valve positions) makes this a challenging Model Predictive Control (MPC) and constrained optimisation problem, which was solved using MINLP (Mixed Integer Non-linear Programming). The MINLP algorithm was selected for its ability to handle a large number of integer choices (valves open or shut / pumps on or off in this particular case).

A model is defined on the basis of a standard element, *viz.* a vessel containing a variable volume , capable of receiving multiple inputs and delivering just two outputs. The physical properties of an element can be defined in such a way as to allow representation of any item in the actual network: pipes (including junctions and splits), reservoirs, and of course, valves or pumps. The overall network is defined by the inter-linking of a number of standard elements. Once the network has been created within the model, the model predictive control algorithm minimises a penalty function on each time-step, over a defined time horizon from the present, with all variables also obeying defined constraints in this horizon. This constrained non-linear optimization requires an estimate of expected consumer demand profile, which is obtained from historical data stored by the SCADA system monitoring the network. Electricity cost patterns, valve positions, pump characteristics, and reservoir properties (volumes, emergency levels, setpoints) are some of the parameters required for the operational optimisation of the system.

PREFACE

THIS project has been initiated by Professor Chris Buckley and the eThekwini Water Services, the authority responsible for the Durban water distribution system, to develop optimal strategies for the water distribution in this area. A strong emphasis was put on dynamic modelling and process control, which are the strong points of the supervisors of this project.

The investigations required a close co-operation with eThekwini Water Services staff to define a characteristic area of the network on which the methods developed could be applied. It was also necessary to get a better understanding of the way this target network was operated. Data, such as maps and operational procedures, were also collected from eThekwini Water Services Head Office in the Durban Central Business District.

Other studies and off-line simulations were conducted in the postgraduate offices of the School of Chemical Engineering at the University of Natal (Durban campus) under the supervision of Professor Michael Mulholland, Professor Chris Buckley, Mr Chris Brouckaert and Professor Marie-Veronique Le Lann of the Laboratoire d'Analyse et d'Architecture des Systèmes and the Institut National des Sciences Appliquées in Toulouse, France

In addition to the research work presented in this dissertation, two courses were completed in the School of Chemical Engineering at the University of Natal:

- Process Dynamics and Control [DNC4DC1]
- Real Time Process Data Analysis [DNC5RT1]

This work has also led to two journal papers and three conference presentations.

I certify that all of the work in this dissertation is my own work, except where otherwise indicated and referenced, and that it has not been submitted for a degree to any other university or institution.

Cédric P.G. BISCOS	Professor Michael MULHOLLAND Supervisor
Date: ………………..	Date: ………………………………..

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to all of the people who, directly or indirectly, contributed to this research by giving some of their precious time and advice.

A special word is due to:

- My supervisor, Professor Michael Mulholland, for his many contributions in this investigation, and for his encouragement, and final proof-reading of this dissertation;
- My co-supervisors Professor Chris Buckley and Mr Chris Brouckaert at the University of Natal, and Professor Marie-Véronique Le Lann at INSA and LAAS (Toulouse, FRANCE), for making this project possible and for their help during its different stages;
- Mr Reg Bailey from eThekwini Water Services for giving me access to the Research & Development department and for his assistance in the data gathering;
- My family and friends for their understanding, support and encouragement;

The financial support of this research from the Water Research Commission, the South African National Research Fund, the French Centre National de la Recherche Scientifique and the Natal University Research Fund is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 4 – MODELLING OF A WATER DISTRIBUTION NETWORK FOR OPERATIONAL OPTIMISATION PURPOSES

CHAPTER 5 – TOOLS INVOLVED IN THE MATHEMATICAL IMPLEMENTATION OF THE MODELLING

LIST OF SYMBOLS

a

a**:** value of the lower bound of the MINLP problem

D

D**:** backward difference operator

$\overline{\mathrm{D}\mathbf{f}_{\mathbf{pump}}}$: vector of flow contributions brought in by the pumps fitting the output legs of the standard modelling elements – available flows strategy only

$\mathrm{D}\mathbf{f}_{\mathbf{pump\ A}}$**:** flow contribution brought in by the pump fitting the upper output leg of the standard modelling element - available flows strategy only *($ML.d^{-1}$)*

$\mathrm{D}\mathbf{f}_{\mathbf{pump\ B}}$**:** flow contribution brought in by the pump fitting the lower output leg of the standard modelling element - available flows strategy only *($ML.d^{-1}$)*

$\mathrm{D}\mathbf{P}_{\mathbf{F\ A}}$**:** frictional pressure drop across the upper output leg of the standard modelling element – calculated flows strategy only *(m of water)*

$\mathrm{D}\mathbf{P}_{\mathbf{F\ B}}$**:** frictional pressure drop across the lower output leg of the standard modelling element – calculated flows strategy only *(m of water)*

$\mathrm{D}\mathbf{P}_{\mathbf{potential\ A}}$**:** available potential pressure in the upper output leg of the standard modelling element – calculated flows strategy only *(m of water)*

$\mathrm{D}\mathbf{P}_{\mathbf{potential\ B}}$**:** available potential pressure on the lower output leg of the standard modelling element – calculated flows strategy only *(m of water)*

$\overline{\mathrm{D}\mathbf{P}_{\mathbf{pump}}}$**:** vector of pump heads of the pumps fitting the output legs of the standard modelling elements – calculated flows strategy only

$\mathrm{D}\mathbf{P}_{\mathbf{pump\ A}}$**:** pump head of the pump fitting the upper output leg of the standard modelling element - calculated flows strategy only *(m of water)*

$\mathrm{D}\mathbf{P}_{\mathbf{pump\ B}}$**:** pump head of the pump fitting the lower output leg of the standard modelling element - calculated flows strategy only *(m of water)*

$\overline{\mathrm{D}\mathbf{P}_{\mathbf{static}}}$**:** vector of the heights difference at the end of the ouput legs – calculated flows strategy only

$\mathrm{D}\mathbf{P}_{\mathbf{static\ A}}$**:** difference of heights at the end of the upper output leg – calculated flows strategy only *(m of water)*

$\mathrm{D}\mathbf{P}_{\mathbf{static\ B}}$**:** difference of heights at the end of the lower output leg – calculated flows strategy only *(m of water)*

$DP_{V\,A}$**:** pressure drop across the valve fitted on the upper output leg of the standard modelling element – calculated flows strategy only *(m of water)*

$DP_{V\,B}$**:** pressure drop across the valve fitted on the lower output leg of the standard modelling element – calculated flows strategy only *(m of water)*

Dt**:** increment of time (or time-step) *(h)*

B

b_j**:** known constraints occurring in the general programming problem

C

C: current chlorine concentration in the vessel of the modelling element *(mg.L^{-1})*

$\overline{C}$ **:** vector of current chlorine concentrations in the vessels of the standard modelling elements

C_{aux}**:** chlorine concentration in the auxiliary input flow of the standard modelling element *(mg.L^{-1})*

$\overline{C_{aux}}$ **:** vector of chlorine concentrations in the auxiliary input flows of the standard modelling elements

C_{max}**:** high-emergency chlorine concentration in the vessel of the standard modelling element *(mg.L^{-1})*

$\overline{C_{max}}$ **:** vector of high emergency chlorine concentrations in the vessels of the standard modelling elements

C_{min}**:** low-emergency chlorine concentration in the vessel of the standard modelling element *(mg.L^{-1})*

$\overline{C_{min}}$ **:** vector of low emergency chlorine concentrations in the vessels of the standard modelling elements

C_{SP}**:** chlorine concentration setpoint in the vessel of the standard modelling element *(mg.L^{-1})*

$\overline{C_{SP}}$: vector of chlorine concentration setpoints in the vessels of the standard modelling elements

$\overline{C_V}$: vector of valve size coefficients of the output legs of the standard modelling elements – calculated flows strategy only

$C_{V\,A}$**:** valve size coefficient for the valve fitting the upper output leg of the standard modelling element - calculated flows strategy only *(ML.d^{-1}.(m of water)$^{-1/2}$.(valve fraction open)$^{-1}$)*

$C_{V\,B}$**:** valve size coefficient for the valve fitting the lower output leg of the standard modelling element - calculated flows strategy only *(ML.d^{-1}.(m of water)$^{-1/2}$.(valve fraction open)$^{-1}$)*

D

d: day

E

e: external disturbance (also known as measurement error) appearing in the MPC formulation

$\mathbf{e(t+i)_{|t}}$**:** values of external disturbance at time t + i, with available information at instant t

$\mathbf{E_{cost}}$**:** time-varying element used to represent the different electricity tariffs along the day

F

$\underline{\mathbf{f}}$**:** mathematical expression of the objective function appearing in the MINLP formulation (generally expressed as a combination of continuous and binary variables)

$\overline{\mathbf{f}}$ **:** vector of output flows of the standard modelling elements

$\mathbf{f_A}$**:** flow exiting the standard modelling element via its upper output leg *(ML.d^{-1})*

$\overline{\mathbf{f_{available}}}$ **:** vector of maximum available flows accessible in the output legs of the standard modelling elements – available flows strategy only

$\mathbf{f_{available\ A}}$**:** maximum available flow accessible in the upper output leg of the standard modelling element – available flows strategy only *(ML.d^{-1})*

$\mathbf{f_{available\ B}}$**:** maximum available flow accessible in the lower output leg of the standard modelling element – available flows strategy only *(ML.d^{-1})*

$\mathbf{f_{aux}}$**:** flow entering the standard modelling element via its auxiliary input leg *(ML.d^{-1})*

$\overline{\mathbf{f_{aux}}}$ **:** vector of auxiliary input flows of the standard modelling elements

$\mathbf{f_B}$**:** flow exiting the standard modelling element via its lower output leg *(ML.d^{-1})*

$\mathbf{f_{demand\ A}}$**:** users demand flow set for the upper output leg of a standard modelling element *(ML.d^{-1})*

$\mathbf{f_{demand\ B}}$**:** users demand flow set for the lower output leg of a standard modelling element *(ML.d^{-1})*

$\mathbf{f_{in}}$**:** flow sent to the input of the current modelling element, when several elements are connected together *(ML.d^{-1})*

$\overline{\mathbf{f_{max}}}$ **:** vector of maximum output flows of the standard modelling elements – calculated flows strategy only

$\mathbf{f_{max\ A}}$**:** maximum flow exiting the standard modelling element via its upper output leg - calculated flows strategy only *(ML.d^{-1})*

$\mathbf{f_{max\ B}}$**:** maximum flow exiting the standard modelling element via its lower output leg - calculated flows strategy only *(ML.d^{-1})*

$\overline{\mathbf{f}_{min}}$ **:** vector of minimum output flows of the standard modelling elements – calculated flows strategy only

$\mathbf{f}_{min\,A}$ **:** minimum flow exiting the standard modelling element via its upper output leg -calculated flows strategy only *(ML.d^{-1})*

$\mathbf{f}_{min\,B}$ **:** minimum flow exiting the standard modelling element via its lower output leg -calculated flows strategy only *(ML.d^{-1})*

G

$\mathbf{g_j}$**:** mathematical expression of the constraint function appearing in the MINLP formulation (generally expressed as a combination of continuous and binary variables)

I

i: position index

J

j: position index

K

k: first-order chlorine decay constant in the vessel of the standard modelling element *(.d^{-1})*

$\overline{\mathbf{k}}$ **:** vector of first-order chlorine decay constants in the vessels of the standard modelling elements

K: number of linearised points in the MILP master problem

$\overline{\mathbf{K}}$: vector of pipeline conductance of the output legs of the standard modelling elements – calculated flows strategy only

$\mathbf{K_A}$**:** conductance of the upper output leg of the standard modelling element- calculated flows strategy only *(ML.d^{-1}.(m of water)$^{-1/2}$)*

$\mathbf{K_B}$**:** conductance of the upper output leg of the standard modelling element- calculated flows strategy only *(ML.d^{-1}.(m of water)$^{-1/2}$)*

M

$\mathbf{m_f}$: flow interconnection parameter

M: length of the control horizon used for GPC techniques

$\mathbf{M_{flow}}$: flow interconnection matrix

$\mathbf{M_{pressure}}$: matrix of pressure selection – calculated flows strategy only

N

n: number of variables occurring in the MINLP formulation

N: length of the prediction horizon used for MPC

nel: number of standard modelling elements used in a model to represent a particular water supply system

$\mathbf{nel_{connect}}$**:** number of elements connected to the input of the current modelling framework

$\mathbf{N_{min}}$**:** minimum costing horizon used for GPC techniques

$\mathbf{N_{max}}$**:** maximum costing horizon used for GPC techniques

P

$\overline{\mathbf{P}}$ **:** vector of receiving and delivery pressures of the standard modelling elements – calculated flows strategy only

$\mathbf{P_{in}}$**:** receiving pressure on the input leg of the standard modelling element - calculated flows strategy only *(m of water)*

$\mathbf{P_{out\ A}}$**:** delivery pressure on the upper output leg of the standard modelling element - calculated flows strategy only *(m of water)*

$\mathbf{P_{out\ B}}$**:** delivery pressure on the lower output leg of the standard modelling element - calculated flows strategy only *(m of water)*

$\overline{\mathbf{Pump_1}}$ **:** vector of the first parameters required for the characteristics of the pumps fitting the output legs of the standard modelling elements

$\overline{\mathbf{Pump_2}}$ **:** vector of the second parameters required for the characteristics of the pump fitting the output legs of the standard modelling elements

$\mathbf{Pump_{1\ A}}$**:** first parameter of the characteristic of the pump fitting the upper output leg of the standard modelling element

$\mathbf{Pump_{2\ A}}$**:** second parameter of the characteristic of the pump fitting the upper output leg of the standard modelling element

$\mathbf{Pump_{1\ B}}$**:** first parameter of the characteristic of the pump fitting the lower output leg of the standard modelling element

$\mathbf{Pump_{2\ B}}$**:** first parameter of the characteristic of the pump fitting the lower output leg of the standard modelling element

R

$\mathbb{R}^n$: set of real numbers

$\mathbf{R_i}$**:** time-dependent weighting sequence, assumed to be positive and semi-definite

T

t: time

U

u: manipulated variables appearing in the MPC formulation

$u(t+i-1)_{|t}$: expected values of manipulated variables at time t + i - 1, with available information at instant t

u_{max}: maximum value authorized for the manipulated variables

u_{min}: minimum value authorized for the manipulated variables

V

v: disturbance variables appearing in the MPC formulation

V: current volume in the vessel of the standard modelling element *(ML)*

$\overline{V}$: vector of current volumes in the vessels of the standard modelling elements

V_{max}: high-emergency volume in the vessels of the standard modelling elements *(ML)*

$\overline{V_{max}}$: vector of high-emergency volumes in the vessels of the standard modelling elements

V_{min}: low-emergency volume in the vessel of the standard modelling element *(ML)*

$\overline{V_{min}}$: vector of low-emergency volumes in the vessels of the standard modelling elements

V_{SP}: volume setpoint in the vessel of the standard modelling element *(ML)*

$\overline{V_{SP}}$: vector of volume setpoints in the vessels of the standard modelling elements

W

W_C: time-varying coefficient used to keep the chlorine concentration in the vessel of the standard modelling element close to its setpoint

$\overline{W_C}$: diagonal weighting matrix used to store the time-varying coefficients used to keep the chlorine concentrations in the vessels of the standard modelling elements close to their setpoint

W_i: time-dependent weighting sequence, assumed to be positive and semi-definite

W_V: diagonal weighting matrix to keep the volume in the vessel of the standard modelling element close to its setpoint

$\overline{\mathbf{W_V}}$ **:** diagonal weighting matrix used to store the time-varying coefficients used to keep the chlorine concentrations in the vessels of the standard modelling elements close to their setpoint

X

x: state variables appearing in the MPC formulation **OR** continuous variables occurring in the MINLP formulation

$\dot{\mathbf{x}}$: derivative of the state variables appearing in the MPC formulation

$\mathbf{x(t+i)_{|t}}$**:** expected values of state variables at time t + i, with available information at instant t

$\overline{\mathbf{X}}$ **:** vector of valve positions or pump status of the standard modelling elements

$\mathbf{X_A}$**:** valve position or pump status on the upper output leg of the standard modelling element *(valve fraction open)*

$\mathbf{X_B}$**:** valve position or pump status on the lower output leg of the standard modelling element *(valve fraction open)*

$\underline{\mathbf{x}}^{\mathbf{L}}$**:** lower bound of the continuous variables occurring in the MINLP formulation

$\underline{\mathbf{x}}^{\mathbf{U}}$: upper bound of the continuous variables occurring in the MINLP formulation

$\underline{\mathbf{x}}^{*}$: optimal values of the continuous variables occurring in the MINLP formulation

Y

y: controlled variables (process outputs) appearing in the MPC formulation

$\mathbf{y(t+i)_{|t}}$**:** expected values of controlled variables at time t + i, with available information at instant t

$\mathbf{y_{max}}$**:** maximum value authorized for the controlled variables

$\mathbf{y_{min}}$**:** minimum value authorized for the controlled variables

$\mathbf{y_{SP}}$**:** reference value for process outputs

$\mathbf{y_{SP|t}}$**:** reference value for process outputs at time t

$\underline{\mathbf{y}}$: integer variables occurring in the MINLP formulation

$\underline{\mathbf{y}}^{\mathbf{k}}$: fixed integer variables occurring in the MINLP formulation

$\underline{\mathbf{y}}^{*}$: optimal values of the integer variables occurring in the MINLP formulation

Z

Z: objective function defined within the MINLP formulation

$\mathbf{Z_L^K}$: lower bound of the MINLP problem

$\mathbf{Z_U^k}$: upper bound of the MINLP problem

$\mathbb{Z}^m$: set of integer numbers

LIST OF ACRONYMS

A

AIChE: **A**merican **I**nstitute of **C**hemical **E**ngineers

AIDES: **A**daptive **I**nitial **DE**sign **S**ynthesizer

AP: **A**ugmented **P**enalty

ARX: **A**uto-**R**egressive model with e**X**ogenous inputs

B

BB: **B**ranch and **B**ound

C

CV: **C**ontrolled **V**ariable

D

DICOPT: **DI**screte **C**ontinuous **OPT**imiser

DMA: **D**urban **M**etropolitan **A**rea

DMC: **D**urban **M**etropolitan **C**ouncil / **D**ynamic **M**atrix **C**ontrol

DV: **D**isturbance **V**ariable

DWAF: **D**epartment of **W**ater **A**ffairs and **F**orestry

E

EHAC: **E**xtended **H**orizon **A**daptive **C**ontrol

ER: **E**quality-**R**elaxation

F

FL: **F**uzzy **L**ogic

G

GAMS: **G**eneral **M**odelling **A**lgebraic **S**ystem

GBD: **G**eneral **B**enders **D**ecomposition

$g.L^{-1}$: gram per litre

GPC: **G**eneralized **P**redictive **C**ontrol

GRG: Generalized **R**educed **G**radient

H

h: hour

I

IDCOM: IDentification and **COM**mand

K

KL: kilo litre

km: kilometre

km^2: square kilometre

km^3: cubic kilometre

L

L: litre

LP: Linear **P**rogramming

M

m: metre **OR** number of inequalities occurring in the MINLP formulation

MATLAB: MATrix **LAB**oratory

M / D: Modelling / **D**ecomposition

mg.L^{-1}: milligram per litre

MILP: Mixed-**I**nteger **L**inear **P**rogramming.

MIMO: Multiple **I**nputs, **M**ultiple **O**utputs

MINLP: Mixed-**I**nteger **N**on-**L**inear **P**rogramming

ML: mega litre

mm: millimetre

MPC: Model **P**redictive **C**ontrol

MPHC: Model **P**redictive **H**euristic **C**ontrol

MRAC: Model **R**eference **A**daptive **C**ontrol

MV: Manipulated **V**ariable

N

NLP: **N**on-**L**inear **P**rogramming

O

OA: **O**uter **A**pproximation

P

PIP: **P**rocess **I**nvention **P**rocedure

PLC: **P**rogrammable **L**ogic **C**ontroller

PROSYN: **PRO**cess **SYN**thesizer

R

R: Rand

S

SA: **S**imulated **A**nnealing

SDP: **S**tochastic **D**ynamic **P**rogramming

SISO: **S**ingle **I**nput, **S**ingle **O**utput

STAC: **S**elf-**T**uning **A**daptive **C**ontrol

STR: **S**elf-**T**uning **R**egulator

W

WHO: **W**orld **H**ealth **O**rganisation

WSSD: **W**orld **S**ummit on **S**ustainable **D**evelopment

CHAPTER 1 – INTRODUCTION

WATER has been identified as the main resource concern of the century. The world is running short of water and, as populations are growing bigger, they are becoming thirstier, with the result that freshwater resources are becoming depleted. Half the world's wetlands have disappeared during the last century, while studies suggest that the water used worldwide will rise by 50 % in the next 30 years. These figures were extracted from the 2002 World Bank's annual report, which was released during the World Summit on Sustainable Development (WSSD) in Johannesburg in September 2002. This report also estimates that as much as half of the world's population, concentrated in Africa, the Middle East and South Asia, will face severe water shortages by 2025. A third of the world's population (i.e. about 2 billion people) are living in countries which are already experiencing moderate to high water shortages. This proportion could rise to half or more in the next thirty years. It means that local water conflicts and the loss of freshwater ecosystems appear to be unavoidable in some regions. The report concludes that there is already ample evidence to justify immediate and coordinated actions to preserve freshwater supplies, as well as provide, distribute and use water more efficiently.

Looking at the efficiency problem for the distribution of water, existing infrastructures have not necessarily been designed optimally. Pumping operations, also regarded as one of the major operational costs in such systems, occur often without any restrictions regarding the different tariffs charged by electricity companies. Some pumping operations which could take advantage of off-peak tariffs, are still undertaken during periods in which electricity is more expensive. In addition, municipalities which are amongst major electricity consumers, often draw this huge amount of energy without regard to the demands from other consumers on electricity facilities, at various times of the day. This stress has repercussions on design and maintenance costs for electricity network operators. Their facilities must indeed be oversized to respond to a high demand during a very short period of the day. These costs are of course transferred onto consumers' bills. It is obvious that a more steady demand from the consumers would reduce the initial investment on infrastructures, reduce the amount of energy losses, and finally limit the maintenance costs on production facilities. Therefore, it would minimise the impact of electricity production on the environment. Nowadays, electricity providers are doing their best to educate their consumers, so that this goal can be achieved. This objective is particularly important in countries like South Africa, where the production of electricity is still responsible for major greenhouse effect gas emissions. Amongst all consumers, municipalities are the main targets of this new policy, as they have a strong influence on the sources of these emissions and have a major interest in reducing them.

Besides improved air quality and improved general quality of life for communities, it would also allow the achievement of huge financial savings and the access to foreign currencies, by way of sponsorships of environmentally friendly programmes.

Optimisation techniques are believed to be one of the promising solutions which could address the problem of water scarcity and efficient distribution. This particular research has developed a new tool, capable of optimising the day-to-day operations of large existing water distribution networks. This tool uses a Model Predictive Control (MPC) approach, and solves the optimisation problem using Mixed Integer Non-Linear Programming (MINLP). The non-linear nature of the problem arises from the simultaneous interest in delivered chlorine concentration, while integer programming is used to determine an optimised sequence for some variables of the system which can only access binary values.

This project was developed as a pilot study, to reveal potentials and requirements for a full simulation / optimisation application on the water distribution system of the Durban Metropolitan Area (South Africa). By reducing the electricity demand from pumps during peak tariffs periods, and respecting simultaneously water quality and security objectives (maximum and minimum volume), it plays a part in the general process of reducing pollution and preserving natural resources. It is also hoped that the savings achieved on the electricity consumption would contribute to make the free basic water approach (which provides 6 kilolitres of free water to every household each month) even more economically sustainable, and therefore, benefit previously disadvantaged communities in the long term.

1.1 Dissertation layout

AS an introduction, Chapter 1 gives background on water distribution, and ends with a look at South-African specific matters. A quick summary of the project objectives and of the methods developed to fulfil them are also presented.

Chapter 2 begins with a short description of the organisation of the Durban Metropolitan Area (DMA), geographically as well as politically. Then, previous and actual adopted water policies are detailed. Finally, important information about the distribution of water as a whole will be given. A link is then made with Durban water distribution network and its customers.

Chapter 3 consists of a comprehensive review of the work undertaken in the last decades in the field of process optimisation and concentrates on MPC and MINLP.

This chapter begins with a presentation of the existing applications of optimisation techniques to the water industry. It continues with information on MPC and MINLP. For each technique, backgrounds of the methods are given, followed by a literature survey of the research field. Finally, the theory underlying each method is presented.

Chapter 4 focuses on the modelling developed to achieve the operational optimisation of water distribution systems. It presents the two successive modelling approaches developed to create a virtual representation of the distribution networks before it is sent to the optimiser for optimisation. The chapter begins with the description of features that are common to both modelling approaches. Then, the specifics of each modelling strategy are presented separately.

Chapter 5 presents the different tools used to implement mathematically the modelling strategies presented in Chapter 4. The chapter begins by giving information on the software MATLAB which is used to create an interface between the models and the optimisation software GAMS. Then, it presents the basic features of the GAMS programming environment. Finally, information is given on the solvers used during the course of this research project.

The following two chapters (Chapters 6 and 7) present respectively small-scale and large-scale applications of the model. The small-scale application, based on a case study, is used to validate the models developed, before one of them is applied to a large-scale application. The large-scale application is based on a particular aqueduct of the Durban water distribution system. It is used to estimate the benefits of the operational optimisation of the system.

Chapter 8 highlights the benefits of the operational optimisation of water distribution networks. It is shown that an optimal solution benefits the system significantly, and that the present study is of scientific interest as a case study of the optimisation of a large hybrid system.

1.2 The origins of water distribution

QUANTIFYING the water reserves of our planet is not an easy task, particularly for the storages inside the Earth's crust. However, several recent studies agree that there is an estimated 1 385 990 800 cubic kilometres (km^3) of water on Earth, mainly stored in the oceans. This volume of water has not changed much over several millions of years. Despite these impressive storages, freshwater is scarce, as 97 % of the water resources are improper for human consumption because of their extreme salinity. Water is indeed considered safe for human consumption when its salinity level is less than 3 grams per litre ($g.L^{-1}$).

Worse, of the small 3 % qualifying for this salinity objective, only 20 to 25 % is effectively available, the rest being trapped at the poles, inside the ice cap. The exploitable freshwater resources are therefore estimated to be a small 9 million km^3. To complicate the scenario even more, freshwater sources are often located in remote areas, far away from those areas suitable for human life. Consequently, the necessity to convey water over long distances, to fulfil the basic needs of populations, appeared long ago, within the first organised societies.

The ancient Hindu and Assyrian civilizations had developed widespread adduction systems, based on the use of gravity, long before the Romans (well known for their skills in building Aqueducts). In major cities, they even succeeded in building basic water distribution networks, although at this stage, the water was provided to final consumers via open channels. A minority of privileged people had access to running water in their houses, but for the rest of the population, fresh water was delivered by mean of fountains or wells. Wastewater and garbage were dumped in gutters, carved in the middle of streets. These channels were linked to waterways, contributing to the overall increase of waterborne diseases in the population (mainly dysentery and diarrhoea).

Nothing changed much until a very interesting property of the water molecule was discovered, at the beginning of the 17th century in Europe. It was proved that the liquid water molecule had the ability to multiply its volume by 1 700, once transformed into its vapour state. Before the end of this century, Denis Papin, a French scientist, conducted numerous experiments, to prove that the huge amount of energy released during the change of state could be transformed into mechanical work. His major scientific discovery was further developed and brought the social and industrial worlds into a completely new dimension. From the beginning of the 19th century, it lead to the commercial development of the steam engine which marked the beginning of the industrial revolution. The rush towards urbanized areas which resulted from this scientific revolution, marked new developments in the field of water supply.

1.3 Water distribution in the modern era

IN Europe, until the beginning of the 19th century, equilibrium was still present between rural and urban populations. However, with the generalisation of the use of steam engines in the industrial sector in the 1840's, all of a sudden there was a great need for a task force in the cities. As steam also brought major developments in the area of transportation, making the travel easier for all, a migration from rural areas towards the city began to take place at great pace. The promises of better living conditions and access to better infrastructures had a strong appeal for people living outside the cities.

However, with industrial waste adding to agricultural and domestic garbage in waterways, the level of pollution reached new heights, creating an unprecedented level of waterborne diseases in the population. This massive migration of population, combined with an outbreak in diseases, improved the awareness of rulers concerning the necessity to improve urban sanitation, and coincided with the creation of the first state structures, responsible for distributing freshwater, removing the garbage in the streets and maintaining public gardens. In the oldest cities, this period saw the coverage of the remaining open-channels, the building of waterworks and storage reservoirs to improve buffer capacities in the distribution of water, the extension of existing distribution systems, the creation of separate sewage systems and the propagation of direct delivery of water to houses.

All this engineering work was conducted at great cost. However, it enabled the access to running water for all, which in return contributed to the overall increase in water consumption. With this huge increase in water consumption, planners became aware that there was more and more a need to secure huge amount of freshwater for dry periods, during which the capacities available inside natural reservoirs would not be adequate for consumer demands for drinking water. This became particularly obvious during periods of severe drought: consequently, it focused on such infrastructures as dams, reservoirs and aqueducts, and pushed forward their development from the last years of the 19th century.

Urbanisation, nevertheless, cannot be the only parameter held responsible for the water consumption increase. Lots of the new technologies which appeared during the industrial revolution were indeed based on the use of freshwater at some stage, and soon industries became the main drawers of water, leaving residential customers far behind. This trend has been even more reinforced with the general use of electricity which appeared following Thomas Edison’s work in the early 1880’s. This technology which was still young at the beginning of the 20th century; grew rapidly throughout the century and changed the face of cities forever, in terms of population growth, comfort, productivity, and of course, water distribution. With the development of electrical water pumps, it became possible to deliver even more water to existing consumers, and increase the number of connections with industrial and residential consumers living in previously unattainable areas of the cities. The discovery of power generation by means of water turbine focused even more interest on dams. Besides storing water for later use, these facilities would allow the production of cheap and clean electricity for the benefit of all.

The discovery of electricity at the end of the 19^{th} century coincided with other breakthroughs realised by science in the medical field which quickly elevated the average hygiene conditions. In 1881, Louis Pasteur discovered the existence of microbes, and therefore, pointed out the source of waterborne diseases. Soon after, waterworks integrated an operation of sand filtering, before the water was sent for storage, reducing dramatically the death rate arising from waterborne diseases. In the early years of the 20^{th} century, the water purification process has been further improved with the discovery of the disinfectant properties of the chlorine molecule.

At the same time, the generalisation of indoor plumbing and the promotion of cleanliness by authorities pushed people to install bathtubs in their households. The improvement of hygiene conditions allowed by these new living standards has been synonymous of delivering more water to the people and has put an additional stress on existing distribution systems.

In Europe and in the United States, the infrastructures continued to develop in urban and suburban areas until the beginning of the 1980's. However, this development has recently been conducted at a far slower pace than before, contrarily to the rural areas, where a lot of engineering had still to be undertaken to respond to the lack of infrastructure, particularly in the most remote areas. Concerning the consumers demand, there has been an important reverse in the general trend of consumption from the beginning of the 1970's. The successive and unexpected economic crises contributed, for a large part, to the drop in the demand from the industrial sector. It is indeed well known that during periods of slowdown of the economy, the oldest industries are always the ones to disappear first. With these industries being particularly reliant on water, their disappearance had significant consequences on water consumption, and created, in the worst cases, a situation of overcapacity for operators of water distribution systems.

Following the economic crises, the introduction of new purification treatments brought several major increases in the price of water: the industrial consumers which survived the economic slowing down from the 1970's had no other choice than to optimise their infrastructure so that it would reduce their water bill. With residential consumers finally becoming aware of the scarcity of water, Europe and the United States are still facing a decrease in their water demand, or at least stabilisation. This should go on over the next decades, thanks to further scientific progress and because of the implementation of more severe environmental laws, encouraging recycling operations more and more, particularly for the industrial sector. Consequently, for companies operating water distribution systems, the situation of overcapacity seems to persist: that is what makes the present a good time for these companies to optimise their existing infrastructure.

1.4 Developments of optimisation techniques applied to the water industry

WITH the reversing of the trend in water consumption, engineers began to think seriously of finding means of planning developments in a more logical way, and reducing operational costs to their minima. They began to improve the empirical techniques to operate the networks, which had been applied since the 1940's.

It was not until the appearance of computer-assisted techniques and the progress in mathematics and statistics that a big jump forward could be achieved. This made possible the development of tools capable of taking into account all of the factors affecting the water demand in one way or another. These specific tools have been created to account for the changes in demand over different intervals of time (going from one hour to several years).

Most of the recent water distribution systems operated worldwide were designed by using these new techniques. Water requirements are fully satisfied, for a combination of domestic, commercial, industrial and fire-fighting purposes. This means that these systems have the ability to meet users' demands at all times, with satisfactory levels of pressure, keeping a good margin for future developments. These powerful tools are helpful for managers who have to plan future investments. They remove a part of the inherent risk that exists in every investment. It also helps the water companies to operate their network in the most efficient way.

1.5 The situation of water distribution in South Africa

THE historical review of the above paragraphs shows clearly that the optimisation taking place in the operation of water distribution networks is a new discipline which is still under development. These techniques have been designed for countries that have followed the path of preserving their resources, or at least, using them in a better way.

South Africa is a country known for its diversity. However, even with all its rich resources and its cultural and natural diversity, water is probably its scarcest and most valuable resource. South Africa's average annual rainfall is well below the world average, and, at current levels of supply and demand, South Africa is predicted to run out of useful water by 2030. Until recent days, the total demand for potable water was still increasing tremendously in the country, following the trend in the population growth, mostly because of a lack of infrastructure in the past. As everywhere else in the world, the economy is almost entirely dependant on a sustained and continued supply of good quality water.

Without the necessary water supplies, sufficient food cannot be produced anymore, industry slows or closes down and tourism invariably suffers. Unfortunately, it is now dear that the water wasting which took place for decades in many countries of the Northern hemisphere because of believed unlimited resources, cannot be a long-term sustainable solution for South Africa.

Before 1994, in South Africa, there was relatively slow progress towards providing equitable and sustainable access to water supply and sanitation services. However, since the peaceful transition towards democracy at the beginning of the nineties, they have been many achievements. Nowadays, the Constitution guarantees everybody access to basic water supply and sanitation services, necessary for human health and well-being. In 1994, the White Paper on Water Supply and Sanitation put in new goals of providing water services to the 12 million people who did not have access to this service previously, and the 18 million not having access to basic sanitation. To implement this political will concretely, it has been necessary to develop the Draft National Sanitation Policy in 1996 and the Water Services Act in 1997. Consequently, there is now an efficient framework for the provision of water services and sanitation to customers. Wherever possible, it is tried to delegate the management of these new services to the lowest level possible.

In 2000, an inventory was made of the progress achieved by the water industry since the application of these new policies. Concerning the access of people to sanitation, it is clear that for various reasons (mainly financial and political), too little has been done to improve access to these services. This problem will become a major challenge for South Africa in the near future. For potable water access then, more than 50% of the population that did not have access to water in 1994 have now been connected to a freshwater source. By developing the free basic water policy (giving monthly a free amount of treated water to the poorest households), the government is trying to solve the problem of inequity inherited from the apartheid era. Unfortunately, these figures, extracted from the Water Services Progress Report 2001, published by the Department of Water Affairs and Forestry (DWAF) do not give any indication of the sustainability of the resources.

The future challenges for the potable water industry in South Africa are numerous. On the one hand, poor operation and maintenance are a problem for most rural schemes, with a population unevenly spread over a huge area. On the other hand, innovative approaches are urgently needed in urban areas to address the ageing infrastructure in the oldest parts of the network, the rapid urbanisation in other parts, and provide good quality and sustainable water to dense informal settlement areas that have mushroomed around all cities and towns since 1994.

This project on the operational optimisation of the Durban water distribution network is part of these innovative approaches concerning urban areas that have been initiated around South Africa.

1.6 Aims of the project

DURBAN, on the shores of the Indian Ocean, is the third biggest city in South Africa with a population of approximately 2.8 million inhabitants. The population of this region is growing continuously because of the arrival of rural people from other parts of Kwazulu-Natal, to enjoy the benefits of its urban environment. Whereas water consumption is still on the rise in the previously disadvantaged areas (mostly in townships and informal settlements), it is decreasing in well-established residential areas, as well as commercial and industrial areas. Consequently, although in some parts Durban water supply system is still expanding to respond to a growing demand, in others, it finds itself in a situation of overcapacity.

Extensions are always an expensive process and have to be undertaken carefully, so that the engineering work conducted is relevant and adapted to future evolution of the demand. To meet the needs of new customers, eThekwini Water Services which operates Durban water distribution network, have decided to integrate powerful prediction tools in the infrastructure design process. There is a general will to prevent the repetition of mistakes of the past: for instance, over-sizing new aqueducts.

Indeed, four years ago, eThekwini Water Services management launched a new initiative, contrasting completely with what was done in the past. At this time, Durban was losing 100 million Rand a year because of leaky pipes in the townships. The demand from new consumers was so high, and the general state of the network in previously disadvantaged areas so bad, that building a new multimillion-Rand dam seemed the only alternative. However, the optimisation of existing infrastructures was pursued, and crews were sent to informal settlements with instructions to repair burst pipes and leaking taps which allowed reducing the average consumption of townships by 30%.

Four years after the start of this innovative policy, the water losses from the city have been halved. By billing consumers who were receiving water illegally and connecting people who previously had no access to water, the municipality managed to increase its revenue related to water by almost 40 %. This reduction in the operational costs enabled the implementation, for the first time in South Africa, of the free basic water policy which has since been adopted countrywide.

With the implementation of this optimisation being a complete success, eThekwini Water Services management has been looking at new directions to reduce operational costs even more. This is particularly justified in a context where it is now clear that the consequences of the HIV / AIDS epidemic, particularly widespread in Kwazulu-Natal and the Durban area, will have strong repercussions on the water consumption, causing a significant decrease in the water consumption and therefore in the water revenue. Current predictions show a strong growth of the death rate which should lead to a decrease in the demand for the next five years, at a rate approaching 5% per year. Then, the demand is expected to stabilise for the next 30 years. This fall could reach dramatic levels, with the recent commissioning of the Durban water recycling plant which will allow industrial customers to reduce even more the amount of water they draw every year.

eThekwini Water Services have prepared since 1997 for this unfortunate reduction of the demand. They have found means to reduce their operational costs, and have limited investments and extensions to a strict minimum. Until now, they managed to prevent any major repercussions on their consumers' bill. Unfortunately, from this year, more and more stress will be put on remaining consumers because of external factors which cannot really be controlled. For instance, Umgeni Water, bulk supplier for the DMA (on which eThekwini Water Services is entirely reliant for the way it charges its consumers) has increased its tariffs each year since 1997. Earlier this year (2003), Umgeni proposed a consequent 28 % tariff hike for metropolitan areas. If new means of reducing operational costs are not found, customers will have to finance this new increase in water tariffs. Moreover, politically, utilities such as eThekwini Water Services must deliver services at the lowest cost, and this is not currently the case. That is where this project plays a part.

Pumping is one of the most important operational costs found in water distribution systems. This is particularly true in the DMA which is a very hilly area, where many pumping operations are required to transfer water from one area to another. Consequently, eThekwini Water Services has identified pumping operations as one of the possible domains requiring operational optimisation. Generally, pumps are electrically driven and fall into two categories: the fixed-speed pumps and the variable-speed pumps. While variable-speed pumps have several continuous operating points; fixed-speed pumps, on the contrary, are operated according to binary choices which means they are either on or off. Although operating policies are applied so that pumps do not switch on during electricity cost peak periods as far as possible, no strategy is currently under use to restrain pumping operations during intermediate and off-peak periods.

It is believed that filling storage capacities by means of pumping when the electricity is less expensive, would allow cutting operational costs by a non-negligible amount, to be evaluated. This research project was created to tackle this issue and conducted as a case study.

General principles have been validated on a small-scale application. Once validated, the methods have been implemented on an existing large-scale application, representing a particular part of the water distribution network of the DMA. The maximisation of the use of low-cost power for pumping operations is the main optimisation objective and can be achieved by using an optimal planning that will allow best use of gravitational flows, and restriction of pumping to low-cost power periods (e.g. overnight pumping) as far as possible. This objective is constrained by the maintenance of the minimum and maximum emergency volumes in all reservoirs, as well as a good water quality in every part of the distribution system. Even if pumping operations are optimised, the new sequence must ensure that a sufficient chlorination is provided in the treated water which is distributed.

In the DMA, the water is treated by means of chlorination. As the network is widespread over a vast area, there is a need for re-chlorination at some stages. The quality of water can therefore be guaranteed, even at final delivery points. Chlorine is a cost efficient, easy-to-use disinfectant, effective in killing bacteria and having residual properties far better than those of ozone. Unfortunately, it produces odours that are easily recognised by customers. When solving the problem of operational optimisation of the water distribution network, efforts are also made to find a compromise between sufficient chlorination (to ensure bacteriological quality) and providing the network with water which consumers find pleasant to drink.

The combination of dynamic elements (i.e. reservoirs volumes) and discrete elements (pumps or valves status, routing) makes this a challenging predictive control and constrained optimisation problem which is being solved by MINLP. The MINLP algorithm is used for its ability to handle integer choices. In this optimisation problem, the presence of integer choices comes from the use of binary valves, and fixed-speed pumps in the distribution network. Integer choices are associated with the status of these elements: valves open or shut / pumps on or off in this particular case.

A non-linear penalty function and several constraints are defined inside a model. The model is based on a standard modelling element, having one input and two outputs linked through a vessel of variable volume.

The physical properties of an element can be defined in such a way as to allow representation of any item in the actual network: plain pipes, reservoirs, and of course, valves or pumps. The overall network is defined by inter-linking a number of standard modelling elements. Once the network is represented within the model, the penalty function is minimised over a defined time horizon, with input and output variables obeying the constraints.

This non-linear optimisation requires the anticipation of a consumer demand profile which is obtained from historical data stored by the SCADA system monitoring the network. Valve positions, pump characteristics, reservoir volumes, and electricity cost patterns are some of the parameters required to achieve the operational optimisation of the system.

1.7 Concluding remarks to the chapter

IN this introduction chapter, a history of the distribution of water worldwide has been given. The recent changes undertaken by the water industry in South Africa have also been presented, with particular interest focused on the situation in the DMA. This background clarified the origin of the present project and its objectives. The tools used to tackle these objectives were also briefly introduced.

In the next chapter, the reader can find more details about the structure of the DMA, geographically, politically and economically. Essential background on the distribution of water from the raw source to the consumers is also given. As far as possible, the discussion refers to existing infrastructure of the DMA.

CHAPTER 2– WATER DISTRIBUTION IN THE DMA

DURBAN is the busiest port in Africa and one of the largest economic bases in Southern Africa, with strong recreational and tourism potential. It is the primary economic centre of the KwaZulu-Natal region. Six local councils and a regional metropolitan council, the Durban Metropolitan Council (DMC), govern this large metropolitan area. The Durban Metropolitan Area (DMA) covers an area of approximately 1 370 square kilometres (km^2) and has an estimated population of 2.77 million inhabitants residing within its bounds. It has been extended at the end of the year 2000 and now includes large areas of former neighbouring councils as presented on Figure 2.1.

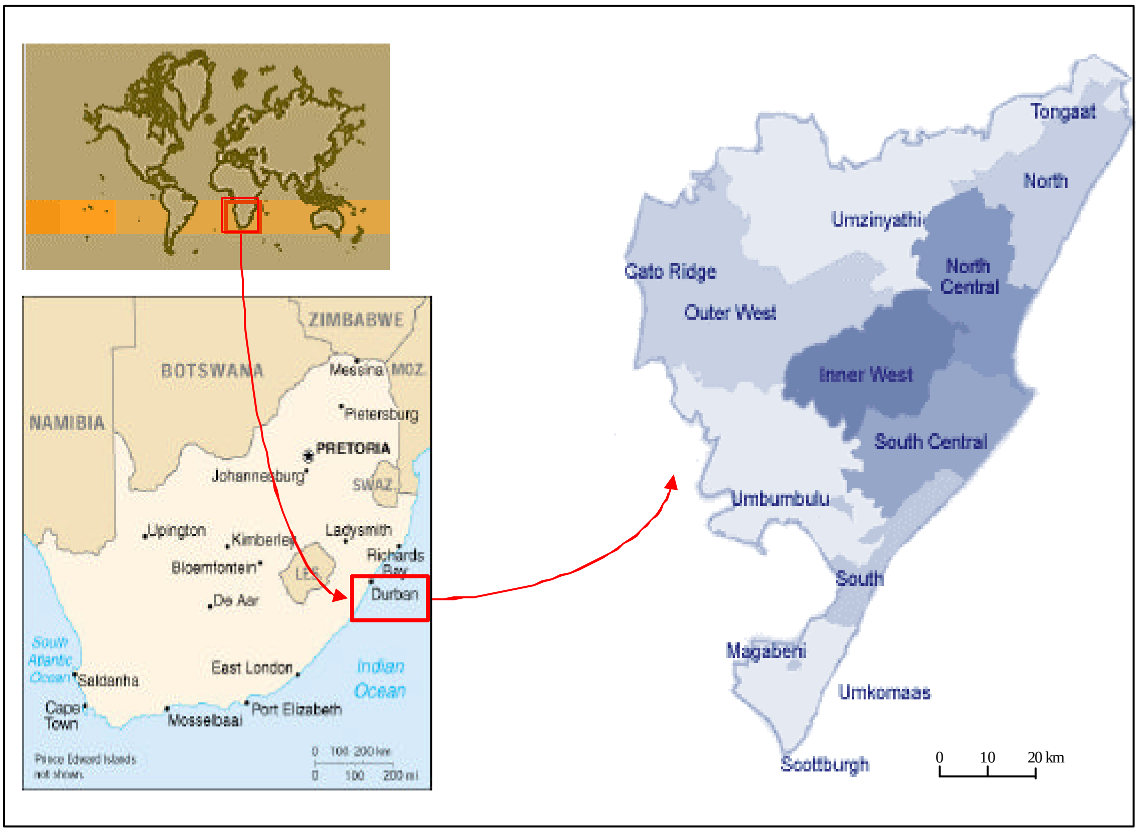


Figure 2.1: Durban Metropolitan Area (DMA)

2.1 The Durban Metropolitan Area and its environment management structures

THE DMA includes very diverse activity zones. Industrial areas are concentrated along the eastern seaboard, radiating westwards towards the interior. Agricultural areas are mainly present in the northern and western areas of the DMA, with sugar cane farming being the main activity. Recreational areas are concentrated in and around formal settlement areas. Sensitive areas, nature reserve and open spaces are found along river valleys and the costal strip of the DMA.

The terrain of this vast area consists of relatively flat zones along the coastal belt, then increasing from sea level to about 500 metres (m) above sea level within a strip of 25 kilometres (km) and finally rising steadily towards the western boundaries of the DMA. The mean annual precipitation of this area is between 700 and 1000 millimetres (mm), the third of which falls between November and March, during the summer season. The climate is typically sub-tropical with moderate winters and warm to hot summers.

Prior to 1996, the local government in Durban had a very limited role in environmental management. However, Durban municipality undertook a democratisation and restructuring process in 1996 and finally inherited the responsibility for environmental management in the metropolitan area. Initially, this function was undertaken on an informal basis by the environment branch of the metropolitan council and the remaining local councils, due to the lack of appropriate capacity and expertise within the metropolitan area. Then, the municipality initiated a program to review the new environmental management responsibilities of the DMA. A working group, consisting of officials representing environmentally related functions (parks, health, and of course water and waste), was created to achieve this task and be capable of giving guidelines on urban planning and management based on sustainable development.

The results found by this group showed that the DMA has currently a wide range of inequalities in terms of the level of services provided for both housing and social infrastructures. Attempts of restoring these historical imbalances in terms of service provision and social infrastructure has become currently amongst the highest priorities of the DMC. This policy applies of course to the water sector and to address it, the DMC, as the water services authority for the DMA, has included all water services functions into a single entity: eThekwini Water Services. This service unit consists of three technical departments (Water, Wastewater Management, and Solid Waste Management) and three service departments (Finance, Communications, and Human Resources). eThekwini Water Services missions are to convey and distribute treated water, as well as collect and recycle wastewater and provide services related to the provision of that water to the satisfaction of consumers within the DMA.

2.2 The water cycle of use adapted to the DMA

WATER used for consumption is mainly surface water or groundwater. Usually, raw water is captured from a lake (natural or artificial), a river, or a well. It is then transported to the urban area, through channels or pipes. If this water is not drinkable, it has to be treated in a purification plant before being delivered to consumers.

Once the water has become potable, it is distributed in the area to the customer tap using a water distribution network consisting of pumps, pipes and reservoirs. After use, this water is classified as wastewater and is collected in a sewage network to be treated inside several wastewater treatment plants. This step is essential not to pollute the environment and to enable downstream use of this water for diverse purposes. The purified water is then released to the receptive environment and diluted in it. The water participates finally to a natural phenomenon called auto-dilution. The residual organic and microbial pollutions are slowly removed during this process. Figure 2.2 below presents the simplified cycle of use of surface water and groundwater:

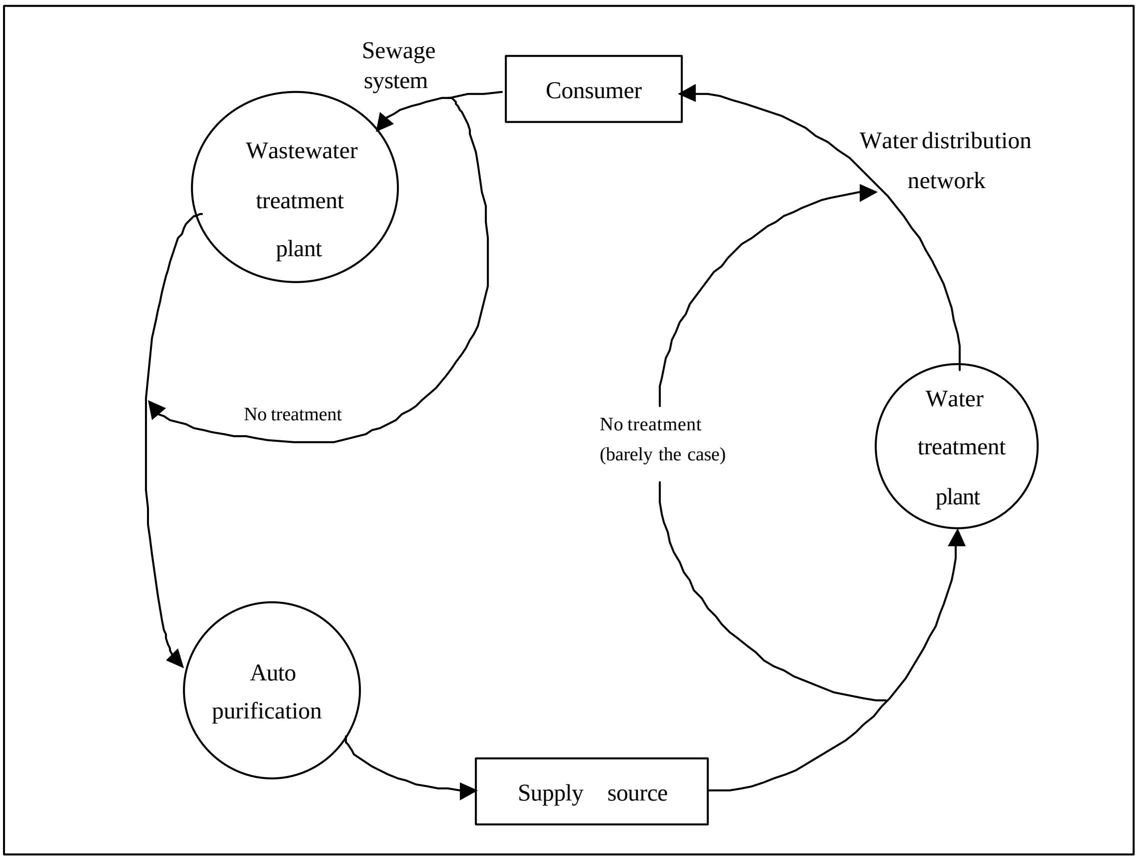


Figure 2.2: Surface water and groundwater cycle of use

The water distribution network occupies a strategic position between the phase of purification of the water and its consumption by the users. A closer examination will be made of the portion of the cycle lying between the supply source and the consumers.

2.2.1 Supply source

THE hydrologic cycle presented on Figure 2.3 shows that human life lies in a relatively small area of the cycle, between the fall of water in the form of rain and its return to the ocean. Clearly, the global water issue is not an overall quantity problem, but rather a flow problem.

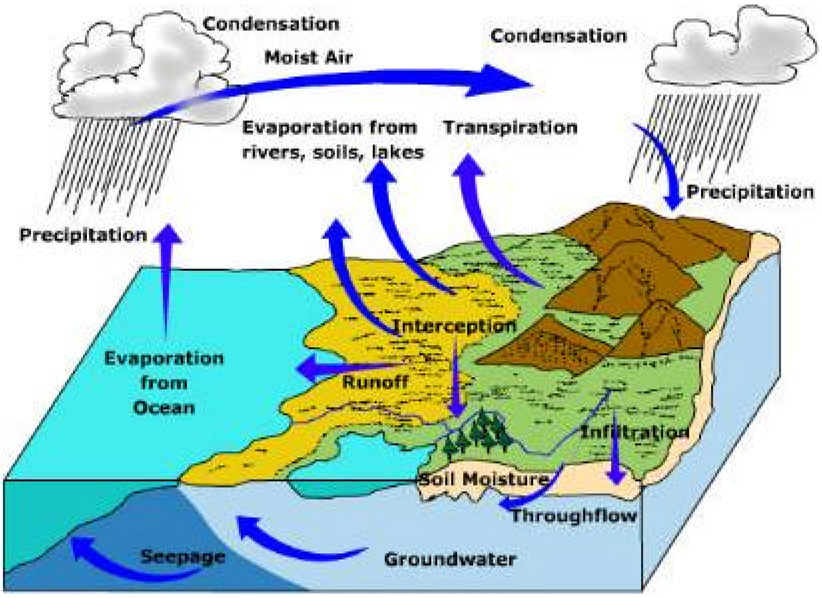


Figure 2.3: The hydrologic cycle (drawn after Gabler et al., 1999)

Indeed, the amount of water available on Earth is constant. However, 97 % of this water is salted and what is of an interest for human beings is to evaluate the amount of freshwater available in the parts of the hydrologic cycle which they are interfering with. When it comes to find a suitable supply source to deliver water to consumers, several choices are available. It is indeed possible to use surface water, groundwater, rainwater, or even salted water. In most cases, surface water and groundwater are the sources that are favoured, the two other sources being used only when surface water and groundwater are not available in sufficient quantities.

Surrounding Durban is an area which has a strong potential concerning surface water. On the contrary, no allocation and reserve of groundwater has been proclaimed for domestic water supply. Extracted groundwater is only used for industry, gardening, and irrigation purposes. All bulk water purchased by the DMC is stored in various places owned by the Department of Water Affairs and Forestry (DWAF) and operated by Umgeni Water. These storages are located along river systems, outside the administrative boundaries of the DMA. The larger of these river systems represent a total catchment's area of 8 200 km^2, and a combined river length of 940 km. Umgeni Water's operational area is presented on Figure 2.4.

The raw water purchased by eThekwini Water Services originates from Inanda, Nagle and Midmar dams. The total capacity of these supply sources represents 800 million megalitres (ML) per year.

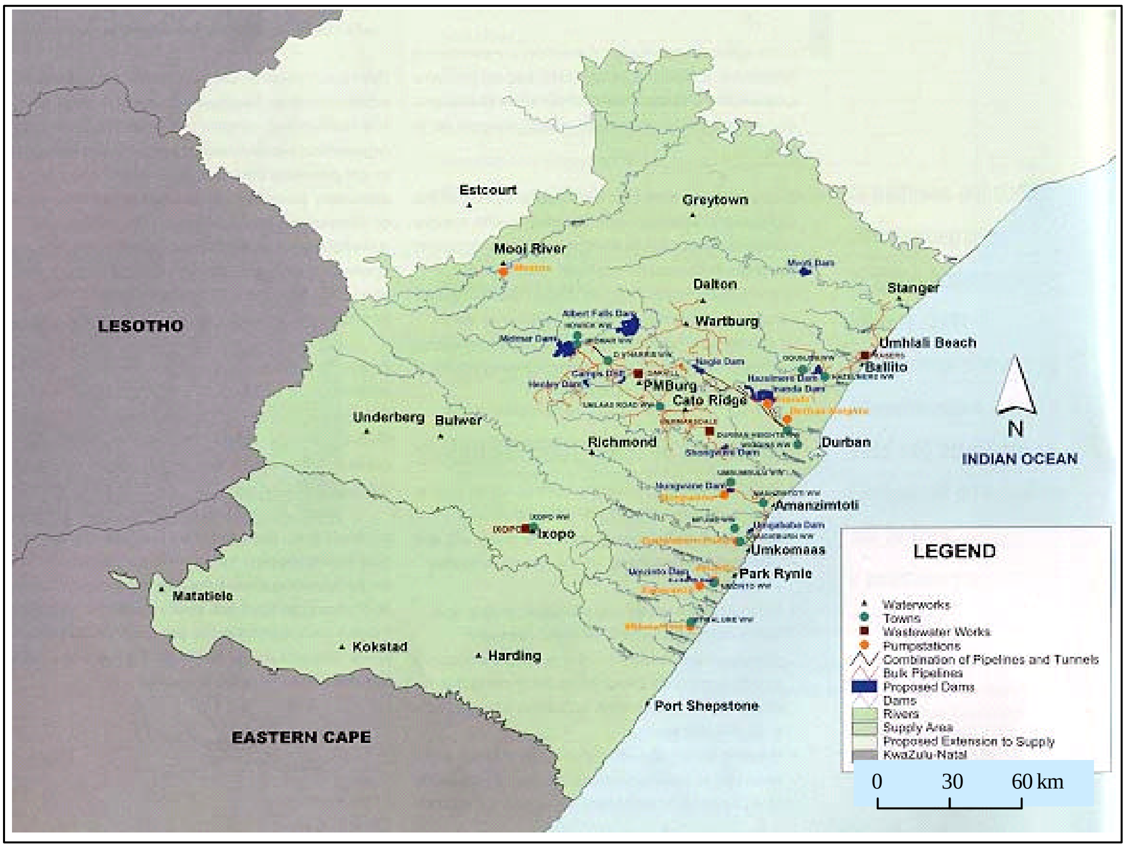


Figure 2.4: Umgeni Water operational area (courtesy of Umgeni Water)

2.2.2 The purification phase

THE water treatment processes developed in the 19th century and refined during the 20th century are simple in nature. However, engineers have since developed ways of making these processes happen faster, in a smaller area, in a more controlled way and at lower cost.

The transformation of surface water or groundwater into water suitable for human consumption needs the use of a combination of very diverse treatments, done in a specific sequence, to finally get a product which is conform to the international health norms promulgated by the World Health Organisation (WHO). Without entering into an excessively detailed description of the treatment process, the most common chain of treatment consists of six distinct steps.

Water extracted from the dams and rivers via pipelines and tunnels passes through wire screens at the waterworks intake points to remove any solid objects. Chlorine gas and powdered activated carbon are added to disinfect the raw water and remove bad tastes and odours. A polyelectrolyte is then mixed with the water and the particles coagulate with suspended dirt particles in the water. The resultant flocs sink to the bottom of the sedimentation tanks.

The clear water above is filtered through graded sand filters which remove all suspended matter. Finally, chlorine is added to kill any remaining germs before samples of the treated water are tested. The treated water is stored in reservoirs until required.

2.2.2.1 The intake

WHEN a water supplier takes untreated water from a river or reservoir, the water always contains dirt and organic matter, as well as trace amounts of certain contaminants. The biggest of these particles are removed at the intake of the waterworks. During this operation, the water goes through a filter, of which the apertures are small enough (between 0.5 and 2mm for the diameter) to remove undesirable particles.

2.2.2.2 Pre-treatment

WHILE methods in the pre-treatment process vary, the effect is always the same. At this stage of the process, the water is already relatively clean but it still contains diluted organic matters and suspended colloidal particles. In most cases, chlorine and powdered activated carbon are added to remove bad tastes and odours from the raw water.

Chlorine is a cheap reactant but, when used excessively, it produces components whose taste and odours are particularly unpleasant for consumers. Chlorine kills microorganisms and prevents the growth of algae at the treatment plant that may interfere with the treatment of water. In modern waterworks, the pre-chlorination is often replaced by ozonation which has the advantages of chlorine disinfection without its disadvantages.

2.2.2.3 Mixing, coagulation / flocculation and sedimentation

IN this process, Aluminium Sulphate,a polyelectrolyte, is mixed in the water. It forms tiny and sticky particles which aggregate to the suspended dirt particles in the water to form particles of bigger size. This phenomenon is called coagulation / flocculation. It enables particles that are slow to settle, or that are non-settling, to settle out more readily. Immediately after, the water enters a sedimentation basin, in which it moves very slowly due to the large size of the basin. The long time taken to exit the basin allows the flocs to settle to the bottom of the basin and to be evacuated out of the treatment process. Removal of most of the particles inside the sedimentation basin improves the operation of filtration which is next in the treatment process.

The coagulation / flocculation processes associated with sedimentation are very effective and are capable of removing up to 99.9 % of the bacteria and 99 % of the viruses still contained in the water at this stage of the treatment.

2.2.2.4 Filtration

FLOCS remaining from the previous operation can be effectively removed by passing water through a filter, in most cases, a bed of sand and gravel, generally not more than one meter thick. It is the last and most important step in purifying the water. Of course, the filters become dirty as contaminants are being trapped in the sand bed. To maintain an acceptable rate of filtration, it is necessary to have them washed from time to time. This cleaning occurs by using pulsed air combined with backwashing which must be monitored carefully to prevent particles of sand following the contaminants into the drain. Filters are backwashed on a rotating schedule to ensure that the plant can operate continuously.

Filtration is one of the oldest and simplest processes used to treat water. A combination of coagulation / flocculation / sedimentation and filtration is the most widely applied water treatment technology around the world, used routinely for water treatment since the early part of the 20th century.

In the middle part of the 20th century, engineers developed rapid sand filters which use high rates of water flow (speed filtration between 5 and 10 meters per hour, against 2 to 5 meters per day for the oldest filters) and sophisticated backwashing of the filter bed to remove the trapped contaminants. These new techniques are now considered the norm around the world.

2.2.2.5 Disinfection

WHEN the water exits the filters, it is generally perfectly clear but there is still a risk that it contains some pathogenic bacteria. A treatment similar to the one described for the pre-treatment operation is consequently applied to the water before it is sent into the water distribution network. At this stage, samples of treated water are tested to make sure it is safe for drinking. Generally, chlorine is preferred to ozone as its decay rate is very slow compared to ozone. It guaranties an excellent bacteriological quality in the water, even long after water has left the waterworks. When ozone is preferred for its strong virus-killing action, it is always combined with the introduction of a small amount of chlorine to prevent the development of pathogenic bacteria, once ozone has lost its beneficial effects.

2.2.2.6 Storage

THE clean drinking water is normally stored in a conditioning reservoir at the waterworks. When drinking water is needed, it is transferred through a network of pipes to the consumers.

2.2.2.7 The water treatment pattern in the DMA

DURBAN water distribution network is provided mainly with water arriving from two waterworks: Durban Heights Waterworks and Wiggins Waterworks. Durban Heights Waterworks is the oldest and the biggest water treatment plant of the DMA, capable of providing up to 720 ML/day of treated water. Wiggins waterworks has been designed in the mid-eighties at a time when Durban Heights waterworks was utilised at near full capacity. The initial capacity of this new plant was 175 ML/day, extended to 350 ML/day in 1995. Figure 2.5 presents the process used to disinfect water at Wiggins Waterworks.

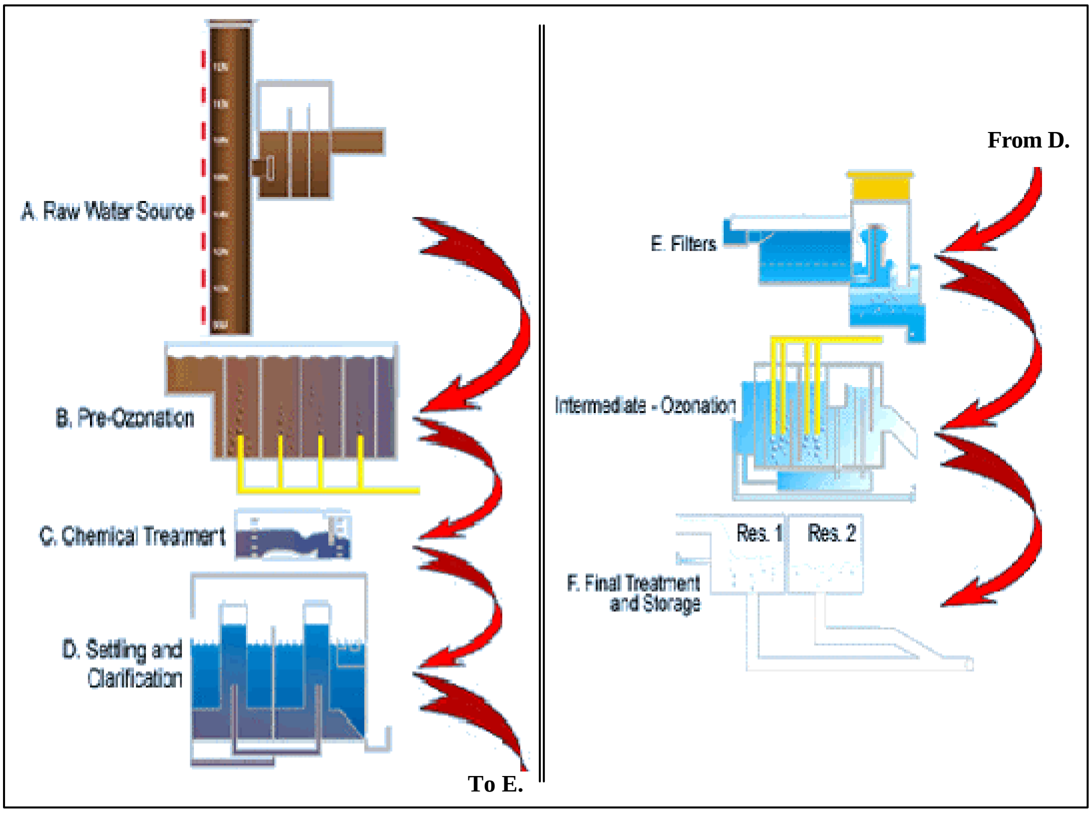


Figure 2.5: Wiggins Waterworks treatment process (courtesy of Umgeni Water)

This process is composed of all the different steps described in paragraphs 2.2.2.1 to 2.2.2.6. It can be seen that ozone is used for pre-treatment, and combined with chlorine disinfection for the final treatment.

Durban Heights and Wiggins Waterworks are capable of providing up to 1070 ML/day to the DMA, of the 1540 ML of treated water that Umgeni Water can produce each day in its different waterworks. During normal consumption days, only 790 ML/day of drinking water are purchased from Umgeni Water. This is supplemented by approximately 23 ML/day from three water-treatment plants that eThekwini Water Services own and operate.

2.2.3 Distribution of water

THE purpose of the water distribution system is to supply water demanded by customers at an adequate pressure, with a bacteriological quality which is in accordance with quality standards defined by the WHO. An underground network of pipes typically delivers drinking water to the homes and businesses served by the water system. Small systems serving just a handful of households may be relatively simple. On the other hand, large metropolitan water systems are extremely complex, sometimes with thousands of miles of piping serving millions of people.

2.2.3.1 Main elements of a water distribution system

WATER distribution systems are built on a combination of four elements. Distribution reservoirs are used to store the water in anticipation of future demand. Pipes interconnect them and deliver water to the consumer taps. Valves allow water pressure and flow direction to be regulated along the way to consumer taps. Finally, many pumps deliver the water with an adequate pressure, even in the most remote areas of the system.

2.2.3.1.a Storage reservoirs

DISTRIBUTION reservoirs have multiple attributes which are technical, as well as economical. From the technical point of view, five fundamental functions can be found. Firstly, a reservoir is an excellent means to regulate a flow, consequently enabling adaptation of the water production to its consumption. Since daily consumption has large fluctuations, it is necessary to use the buffering capacity of reservoirs, eg. as auxiliary input sources in the case of instantaneous peaks in the demand. A reservoir transforms a peak demand interval into an average demand interval and smoothes the demand on the waterworks.

Secondly, a reservoir can be used as an emergency source in the case of an incident (pollution of the raw water; breakdown of one of the components of the waterworks, breach in one of the pipes, disruption in the energy supply) for one or several elements playing a part in the water distribution process.

The third technical function of a reservoir is to regulate the pressure in the distribution system, reducing problems of equipment stress linked to constant changes of pressure.

The next advantage of reservoirs lies in a simplification in the day-to-day operations, allowing, for instance, equipment stoppage for routine maintenance.

Finally, the reservoirs positioned immediately downstream of a waterworks improve the water quality by enabling a sufficient contact time between the disinfectant agent and the drinking water. This guarantees an adequate treatment before water is distributed further down to the consumers.

From an economical view, reservoirs lead to a significant reduction in the capital cost of a water distribution system. The presence of reservoirs enables a reduction of main distribution pipe diameters. Of course, in networks for which pumping energy is used to distribute water to consumers, the presence of a reservoir on a pipe route allows significant energy cost savings, by storing water during off-peak periods for use during peak periods.

The main categories of storage reservoirs are surface reservoirs, standpipes, and elevated tanks. Surface reservoirs are the most widespread type of reservoirs. They can be covered or uncovered, but whenever possible, a cover is used for the prevention of contamination of the water supply. Standpipes or elevated reservoirs are used when the construction of a surface reservoir would not provide a sufficient head. A standpipe is usually a tall cylindrical tank whose storage volume includes an upper portion, the useful storage and a lower portion which only acts to support the useful storage and provide the required head. Often it becomes more economical to build a supporting structure for an elevated tank than to provide a supporting storage (generally for heights greater than 15 meters).

Distribution reservoirs have to be located strategically to get maximum benefits. It is better to have the reservoir near the centre of use, but in large metropolitan areas such as Durban, a number of distribution reservoirs are located at key points. Reservoirs providing storage must be high enough to develop adequate pressures in the system they have to provide with water. The amount of storage to be provided is a function of the capacity of the distribution network and of the location of the service storage.

As it is generally better to operate pumping units at a steady rate, demands on the distribution system in excess of these rates have to be accepted by storage facilities. During emergencies, requirements for fire-fighting purposes should provide water for at least 12 hours (h).

When the inflow to the reservoir is shut off, security requirements should account for a consumption of more than a couple of days. At all times, a good circulation of water inside the reservoir must be ensured to prevent degradation in the water quality during the time of retention inside the storage facility.

2.2.3.1.b Pipes

A distinction is made between the main distribution pipes (originating from a waterworks, a reservoir or a pump-station) and the secondary distribution pipes (linking the main distribution pipes to consumer reticulation). The main distribution pipes provide the secondary distribution pipes with water. They do not have any connections with the reticulation system and the branching with secondary distribution pipes is limited to a minimum to increase the reliability of the whole network. Considering the large flows they are supporting, their diameters are rarely lower than 300 mm. The average diameter of secondary distribution pipes varies between 100 and 150 mm, but a range of 40 to 300 mm is possible. The reticulation branchings from the secondary pipes distribute water to customers or groups of customers. Their diameters are in the range of 20 to 200 mm. It is on these pipes that the metering is done.

All pipe diameters take into account the strongest foreseen instantaneous flow. These diameters have to be large enough to respond to the fire-fighting needs occurring occasionally. The materials used are a function of the levels of pressure in the pipe. Generally, cast iron, steel or reinforced concrete is used for the main distribution pipes. For the secondary distribution pipes and reticulation pipes, besides these three materials, it is common to use polyethylene and PVC.

2.2.3.1.c Valves

VALVES are found commonly in water distribution systems, as it is impossible to design a practical fluid system without some means of controlling the volume and pressure of the fluid and directing the flow to the different operating units (reservoirs or diverse service areas).

Valves are classified according to their use: flow control, pressure control and directional control. Some valves have multiple functions that correspond to more than one classification. Flow and pressure control valves are the most common valves used in water distribution networks. The control of flows inside a water distribution network is essential to fulfil correctly the demand requirements in every parts of the system. Flow control valves can be operated manually, hydraulically, electrically or pneumatically. They are usually either binary (fully opened or fully closed only) or continuous (full range between fully opened and fully closed).

As there is a need to control flow in the system, there is also a need to control pressure levels. There are many types of automatic pressure control valves. Some of them provide an escape for pressure that exceeds a set pressure, whereas others only reduce the pressure to a lower pressure, and some keep the pressure in the system within a required range. Valves often lead to small pressure drops inside the system.

2.2.3.1.d Pumps

THE use of pumps has been widespread since the beginning of the 20th century, making all other methods of elevating water disappear. The generalisation of pumps has enabled the water delivery with an adequate pressure, even in the most remote areas of the service area. Pumps are normally driven by electrical power and can be classified according to the way they operate: either with fixed-speed or with variable-speed. The fixed-speed pumps are either switched on or switched off, and unlike variable-speed ones, only have limited flow adjustment by valve throttling. As it is always better to operate a pump at a steady rate, fixed-speed pumps will be generally preferred to variable-speed ones.

The success of pumps can be explained by the big advantages they offer. Their robustness and longevity are remarkable, mostly due to the simplicity of their conception and the limited maintenance they need. Furthermore, the use of a pump in a water distribution network increases stability, with the pump configuration able to adapt automatically to important variations of the characteristics of the network. Their electricitiy use efficiency is generally high, around 0.70 for an average pump. Finally, since pumps have become common machinery, they are sold at affordable prices.

The characteristics of a pump are usually summarized by three different types of curves with varying flow, giving its head, its efficiency and its absorbed electrical power. For each parameter, there is a curve corresponding to rotation speed, and consequently there is a family of characteristic curves of a pump when its rotation speed varies. Pump characteristics are well known for their non-linearities. As for valves, idling centrifugal pumps often cause small pressure drops in the supply system.

There are several advantages in the combination of reservoirs and pump-stations. Firstly, pressure does not change much inside the network, even during peak periods. The pumping flows can remain relatively uniform which allows specification of pumps to run at a level close to optimality. Further, it is possible to transfer water from one service area to another in case of an emergency at one of the water treatment plants.

2.2.3.2 Classification of water distribution networks

DISTRIBUTION systems can be ordinarily classified as grid systems, branching systems or a combination of these, all including the four types of elements presented in the above paragraph. The configuration of the system is decided primarily by topography, street patterns, degree and type of development of the area and of course the location of the treatment and storage waterworks. Figure 2.6 summarizes the different categories of water distribution networks.

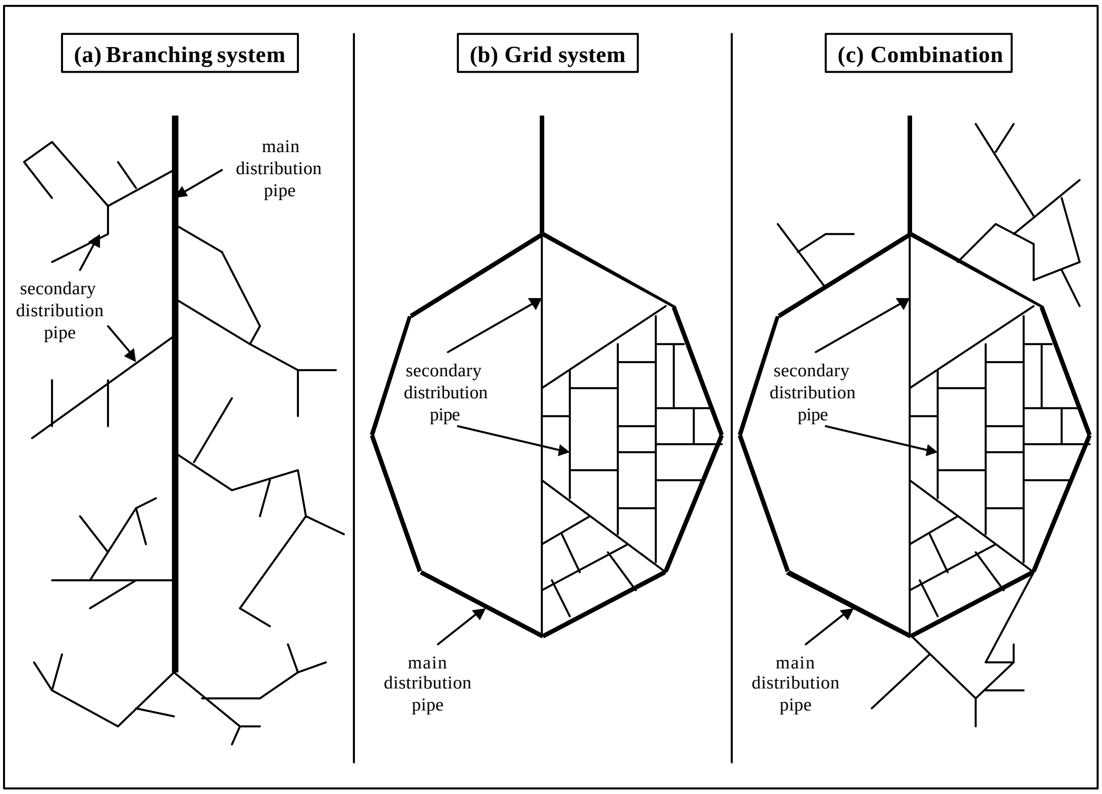


Figure 2.6: Different types of water distribution networks

A grid system is usually preferred to a branching system, because it can ensure that water will be provided to any point from at least two directions. This configuration limits the number of consumers deprived of water in case of a serious problem in the system. On the contrary, the branching system does not permit this type of circulation, because of the numerous terminals or dead ends it is composed of. Consequently, when damage occurs on one part of the system, all the customers located downstream of the faulty pipe lose their access to water. Sometimes a combination of branching and grid systems is also used.

In locations where sharp changes in topography occur (like within the DMA boundaries), it is common to divide the distribution system into several service areas or zones. This eliminates the difficulty of maintaining extremely high-pressure levels in low-lying areas in order to get convenient pressures for higher elevations.

Generally, such systems are interconnected to each other, the interconnections being closed off by valves during normal operation.

2.2.3.3 Water distribution in the Durban Metropolitan Area

APPROXIMATELY 800 km of bulk mains convey treated water to 220 storage reservoirs within the DMA. Water is further distributed via 4500 km of pipes, generally with a diameter less than 300 mm. These secondary pipes radiate outwards from the storage reservoirs or water treatment works. Approximately 160 test points in the piped potable water distribution system are continuously monitored by eThekwini Water Services to ensure compliance with the international health norms provided by the WHO. Bulk supply pipeline routes are routinely flown by helicopter and potential leak sites are identified from the air by looking for areas where grass and vegetation is particularly green and lush. Ground teams are then dispatched to investigate further.

eThekwini Water Services utilise a comprehensive fault management system which covers all components of the bulk infrastructure network and ensures that any necessary repair work is done. All jobs undertaken by the Operations Department are logged, together with details of material type, pipe diameter, fittings, etc. Area Managers monitor the condition of infrastructure in their areas of responsibility by extracting and analysing detailed fault reports. eThekwini Water Services do not presently utilise a pressure metering and monitoring programme in their reticulation network. There is indeed an awareness of the location of problem areas because of high water losses and resources are concentrated in those areas to detect and remove leaks.

Two hundred and thirty-five pump-stations are currently operational within the DMA. A cyclical programme for inspection of pump-stations, cleaning of screen and sumps or overhauling of pump motors has been set. All of the pump-stations presently operational within the DMA have had telemetry instrumentation fitted.

eThekwini Water Services have several bulk infrastructure projects currently, either in the planning, or construction phase. Over the next five years, these projects will represent an investment of about 440 million Rand (R), 90 % of this investment being used for the construction of a new aqueduct in the western areas of the DMA (bulk pipeline and reservoirs) by the end of the year 2005. Other projects (i.e. new storage reservoirs or trunk main duplication) have been identified but have not been costed yet.

2.2.4 The consumers

A number of elements can affect the amount and timing of water use. They must be taken into consideration when forecasts of future water uses are to be made. The elements most affecting the amount of water consumed are the population size, the climate, the composition of the region or community being dealt with, the cost of water, the economic conditions, the public attitude towards water conservation, and the existence of a strong tourist potential.

Concerning the population, the amount of water used in a locality is directly related to its size, its distribution and the composition of the local population. Errors in projecting population changes affect water use projections. The amount of water used in a region is also a function of the local climate. Gardening, bathing, irrigation or cooling affect consumption. Moreover, the type and scale of residential, commercial, and industrial developments in an area will have a profound effect on local water usage. On the economic aspect, inflation and other economic trends influence the availability of funds for water supply. Social attitudes toward environmental protection strongly affect water resources management and development processes. Water use forecasts must also take into account changing attitudes about what constitutes a beneficial use of water and how society perceives the importance of allocating water for environmental purposes. Finally, in regions with a strong tourist potential like Kwazulu-Natal, the impact of tourists must be acknowledged when forecasting future water demands.

2.2.4.1 Demand prediction

GENERALLY, the greater the number of people residing in an area, the more water will be used. Furthermore, it is not only the number of people that is important but also their ages, level of education and social backgrounds. Population issues include the way in which the current population is distributed, as well as the way the future population will be distributed. Besides this, they take into account the way the population growth rate affects the area economy, its natural resources, labour force, energy requirements, or urban facilities.

Historical data are the basis on which future levels of population are predicted, but unfortunately, these data are not always adequate. The problem is especially important in countries like South Africa where there is a big data deficiency inherited from the apartheid-era and many uncertain factors influencing population changes (fertility, mortality, migration towards cities, HIV / AIDS epidemic etc). That is why most forecasters suggest the use of at least three possible trends in population growth, based on plausible mixes of influencing factors.

In South Africa, and of course within the DMA, three factors complicate the population estimate. Firstly, the wealthiest people are leaving the central business district, to settle in luxury suburbs, so that the remainder are left alone to support the tax burdens. This trend is placing a great stress on traditional mechanisms of municipal finance. In the water supply area, this particular factor is obvious in the problem of maintaining a large infrastructure already in place with fewer resources to do so. The second problem is rural migration towards the cities. The rapid shifting of the population from remote areas towards the cities creates significant problems of providing needed services like water. Finally, the third affecting factor, the HIV / AIDS epidemic, particularly widespread within the DMA, because of its unpredictable nature, is probably the one which most complicates the task of planning.

The different possible population flows must be given serious consideration by planners and managers. The impact of rapid increases in population (as was the case until recently inside the DMA), or rapid decrease (as could become the case soon, with the HIV / AIDS epidemic becoming severe), has just been pointed out. Nevertheless, there is also a need to know how long these population trends will last, and when a reverse will eventually occur.

The types of population estimates required for the operation of a water distribution system are short-term estimates in the range 1 to 10 years, and a long-term estimate of 10 to 50 years (or even more). War, technological developments, new scientific discoveries, government operations and many other factors can drastically disrupt population trends. Usually, the population estimates are based primarily on past census records for the area and an extrapolation is predicted from the existing trends. As this does not take into account any external factors, it is improved by using all possible information regarding industrial growth, local birth rates, death rates, and tourism developments, etc.

2.2.4.2 Current situation in the DMA

CURRENTLY, the population of the DMA is estimated to 2767440 inhabitants. This value has been projected to year 2002 by applying a constant growth rate of 1.67% which is the “Middle AIDS scenario” growth rate as described in the “Demography and Demographic Projections to 2020” document published by the eThekwini Traffic and Transportation Department. In the year 2000, the DMA had an estimated 440 000 households on formally serviced sites, 133 600 households in informal settlements, and 27 400 households in backyard shacks.

As in every major urban centre around the country, there are still lots of inequalities amongst these households’ incomes and their access to water.

Regarding the access to water, 77 000 households were still lacking access to any source of potable water, 81 500 were accessing it via communal standpipes, 28 500 households were connected to yard tanks or taps and about 400 000 households had a proper in-house connection. On top of the private users, commercial, industrial and institutional users represent an approximate 15 000 added connections.

From a financial viewpoint, 36 % of these households earned less than R 1000 per month, 45 % got from R 1000 to R 6000 per month, and only 19 % earned more than R 6000 per month. Payment for water services in the DMA is higher in the established suburbs than in the lower income township areas. The current policy on consumer non-payment for services is to disconnect them. In order to encourage payment for services, eThekwini Water Services have instituted a system which assists the consumers in managing water usage. Once a customer has been disconnected, he is provided with a device which only allows him to use 200 litres (L) per day, since the first 6 kilolitres (KL) is free each month for domestic users. This device is removed once the consumer has brought his outstanding balance back to zero.

2.3 Concluding remarks to the chapter

IN this chapter, backgrounds have been given concerning the DMA and the current situation of water distribution inside this area. From a theoretical point of view, the surface water and groundwater cycle of use has been explained. The portion of the cycle which is of use for this project, has been explained and considered in relation to the specific issues of the DMA.

In the next chapter, the techniques used to achieve the operational optimisation of water distribution systems are presented.

CHAPTER 3 - PREVIOUS WORK

OPTIMISATION is a common task in engineering practice. In general, it can be classified into two different categories: the static optimisation, as opposed to the dynamic optimisation. Static optimisation concerns systems for which the variables to optimise have reached their steady state, whereas for dynamic systems, the variables to optimise are still evolving with time. Because of the challenges introduced by the second category of systems, the research in the field of dynamic optimisation has not been extended as far as it has been for the static optimisation, and numerous teams of researchers around the world are still undertaking a considerable amount of work (Schwartz and Chen, 2002).

Dynamic optimisation problems exist in the areas of simulation, parameter estimation for dynamic models and optimal control. For this last application, the goal of dynamic optimisation is to find the optimal control profile of one or more control variables (or control parameters) of a system. Optimality is defined as the minimisation or maximisation of a penalty function without violating given constraints (Wen J., 2002). This family of problems is part of a vast range of control methods developed around common ideas and gathered under the denomination of Model Predictive Control (MPC). The operational optimisation of a water distribution network is a very good example of MPC, as the problem consists in finding the best sequence for decision variables (like valves switching or pumps status), while it looks at fulfilling consumers' demands and respecting the integrity of water quality, security and financial constraints at the same time.

Numerical methods designed for solving optimal control, and in particular MPC problems, have not reached the stage of methods used widely for solving differential equations. Therefore, solving an optimal control problem can require, depending on the level of complexity of the problem, a significant user involvement in the solution process. In other words, the user of such techniques must understand the theory of optimal control, optimisation and, of course, numerical methods for dealing with differential equations. Until recently, due to computational limitations, most of the work linked to dynamic programming was often addressed by breaking up a complex optimisation problem into a reasonable number of simpler problems, as originally proposed by Bellman (1957). The solution of the simpler problems was then used to give the solution of the global formulation. With computers becoming more and more powerful, the use of dynamic programming to solve optimal control problems is more and more popular among researchers. When applied to the field of water distribution, dynamic programming must take into account the binary nature of some of the variables involved in the system.

Mixed-Integer Non-Linear Programming (MINLP) techniques are amongst the latest approaches which are used to solve this kind of complex combinational problem.

3.1 Process optimisation applied to water distribution networks

OPTIMISATION of water distribution systems has received enormous attention in the past few decades. Research in this field is still highly active, as the complexities associated with the operation of multiple interconnected reservoirs still exceeds the capabilities of existing optimisation tools in finding solutions easily (Teegavarapu and Simonovic, 2002). Nowadays, the optimisation efforts are found especially on reservoir operation. Numerous algorithms are being tested on distribution systems by researchers to get the most reliable solutions, using the least computational time possible. Fuzzy logic (FL), stochastic dynamic programming (SDP), simulated annealing (SA) algorithms and, of course, MINLP, are the most promising methods (Mousavi and Ramamurthy, 1999).

3.1.1 Design optimisation

IN the literature, the topic of optimisation in water systems is mainly focused on the design of optimised configurations for pipe-interconnected reservoirs. Numerous trials have been attempted to get a better design of reservoir interconnections during construction or extension phases of new or existing networks. The optimisation problem is often posed as maximising the benefits, or at least, minimising the cost of distributing the water, by searching for optimal values of a set of variables subject to intrinsic system constraints. This problem is usually of a large size, and known to be non-linear. This non-linear nature originates from the presence of equipment, like pumps, of which the characteristics include significant non-linearities. As optimisation of water distribution problems combines dynamic elements with discrete ones, it is clear the problem must also be solved by using an approach based on integer programming, when it is required to keep a model close to reality (Walters and Cembrowicz, 1993). However, with the complexities associated with integer programming, it is not surprising that the models developed to tackle these problems have been deliberately simplified. Only recent developments in the computational field have enabled the consideration of approaches that are more comprehensive.

The creation of linearised models has been a frequent option to facilitate the use of Linear Programming (LP). Schaake and Lai (1969) were amongst the first to create an application of LP techniques for the design of water distribution networks.

Fujiwara *et al.* (1987) applied the methods further to a looped distribution system; while Kessler and Shamir (1989) developed new techniques specifically designed for water supply systems, and based on the LP gradient methods. These researchers have improved their method in 1991 by developing decomposition techniques for the optimal design of distribution networks.

Another path has consisted of forgetting about the discrete nature of some variables. Research teams who used this approach managed to solve systems more easily, still keeping them non-linear, but no longer combinational. In extreme examples, the discrete elements were even completely removed. For instance, Cunha and Sousa (1999) developed an approach based on the SA method for the optimisation of the Hanoi distribution network. However, to test the performance of their model, they used simple networks which did not contain pumps.

Few of the original approaches were implemented for the design of complete new systems, mainly for computational reasons. This is why they have frequently been limited to the scale of a case study. As described by Walski (1996) in his discussion, using test problems to evaluate new optimisation techniques does not necessarily prove that the methods can be generalised to a large-scale system. As he shows, more often than not, most of these new methods could not be implemented in reality for practical reasons. With the enormous progress being constantly achieved on computers, this statement is likely to become less and less exact.

Optimisation models applied to the design of water distribution problems can be classified depending on whether or not they are formulated by taking into account hydraulics of the network (Wu, 2002). Nowadays, including hydraulics in a model is very common, but this was not the case in the 1970's, partly because the first computers had limited ability to handle complex problems. Details about models that do not include hydraulics can be found in Alperovits and Shamir (1977). The original and innovative approach they proposed was to reduce the complexity of the non-linear nature of the problem by solving a sequence of linear sub-problems. One optimisation is run to solve for the optimum flow status for a given network and is a non-linear optimisation problem. A second optimisation follows to determine the optimal solution of the variables for this optimal flow distribution. This step is considered as linear. To improve the solution, an iterative procedure takes place. Several research teams, like Quindry *et al.* (1981), Goulder and Morgan (1985), and more recently Eiger *et al.* (1994), have also adopted this sequence and have subsequently improved it.

The first models including the hydraulics of the network appeared in the 1980's, again to respond to the problem of providing optimised design of water distribution systems.

Usually, an hydraulic network solver is integrated and used to find the best alternative designs of water distribution networks during the optimisation process. Specific hydraulic configurations and different types of constraints can be integrated within the formulation. Woodburn *et al.* (1987), for instance, combined a non-linear programming procedure with an hydraulic simulation program in a model designed to determine which pipes should be replaced, rehabilitated, or left alone in order to minimise operational costs. Soon after, Su *et al.* (1987) introduced probabilistic considerations to the Woodburn *et al.* (1987) work. The reliability of such models has since been improved thanks to the work of researchers, like Lansey and Mays (1989), Duan *et al.* (1990), and Kim and Mays (1994).

Nowadays, the latest approaches proposed for optimal design of water distribution systems can be considered reliable enough to find their way into practice, and do not necessitate unattainable computational requirements anymore. Still, besides optimising the topology of distribution networks, relatively few researchers have tried to optimise day-to-day operations of existing water supply systems.

3.1.2 Operational optimisation

THE objective for this second family of problems is mainly to generate control strategies ahead of present time, using predictive techniques, to guarantee a competent service in the network, while achieving certain quality objectives at the same time. Minimisation of supply and pumping costs, maximisation of water quality and leak prevention are among the main targets (Cembrano *et al.*, 2000).

Creasey (1988) reviewed the mathematical techniques which are appropriate to solve the problem of operational optimisation for water distribution networks, insisting particularly on the pump-scheduling problem. He concludes that a better pump scheduling could save £ 10 million a year to the UK water industry alone. Although he reckons that there is a range of valid methods for the implementation of scheduling tools, he also states that integer-based programming seems to be the single way to achieve these savings for a wide range of network sizes and complexities.

Unfortunately, with the limited available level of technology which existed on integer programming at the time of the first studies on operational optimisation of water distribution systems, alternative methods had to be developed. It appears that most of these initial approaches have been based on dynamic programming at some stage.

Conducted as case studies, they would not have produced acceptable computational times on a large-scale application. For instance, at the end of the 1970's, Rao and Bree (1977) presented a procedure to perform an extended-period simulation of a water distribution network, with up to 1500 nodes. The tool they developed is aimed at maintaining flow rates and pressures within specified bounds, at various points in the system. It also looks at managing the storage so that the supply and the distribution are balanced more efficiently. To solve this optimisation problem, the two researchers used sparsity-oriented programming techniques. The simulation period is cut into several intervals on which a static optimisation giving changes in water levels in the reservoirs and pumping schedules is conducted. To reconstruct the solution over the full simulation interval, they developed an integration scheme which links several static solutions. With this method, it has been possible to evaluate heuristic control schemes for several distribution networks, over a 24 to 48 h simulation period. The main drawback of this method is the fact that it can only be applied off-line, to systems which have reached a steady state, so that a static optimisation can be performed.

Similarly, Wood and Rayes (1981) concentrated on the problem of analysis of pressure and flow in piping systems. They reviewed the reliability of five algorithms for applications to piping systems of general configuration, containing pumps and other common hydraulic components. They summarize a pipe network as a number of pipe sections which have constant diameter sections, containing pumps and fittings such as bends and valves. The end-points of the pipe sections are nodes which can be connected either to another pipe section, or to a reservoir where a constant pressure head is maintained. Their study is aimed at identifying which of the five algorithms is the best to determine optimal flowrates and heads to be used within a target piping system. They acknowledged that to achieve significant convergence on the five methods they tested; they had to estimate an initial set of flowrates or heads which was close to the optimal values. This constitutes an important limitation to these algorithmic methods, since there is no reliable means to determine efficiently better initial values. Worse, for three of the algorithms, starting with an excellent set of initial conditions does not even assure convergence.

Goldberg and Kuo (1987) preferred orientating their work to the optimal pumping of a serial pipeline, where the objective was to minimise energy cost and apply adequate pressure boundaries at all pump stations. The search for an optimal solution was done by manipulating 40 binary variables, but no objective was specified on reservoir setpoints or user demand flows.

Soon after, Coulbeck (1988) reviewed the different methodologies for the modelling and control of a water supply.

He acknowledges that the required optimisation computations to be used for such problems can only be performed for sufficiently simplified system models, since the computational resources at his disposal were still limited. Consequently, he presents three simplified modelling approaches: a linear dynamic model, a non-linear dynamic model, and finally an operational cost model. These models are solved using dynamic programming, but also use decomposition methods which cater most efficiently for large-scale application. Applications of these models are presented for single / multi-source, single / multi-reservoir systems, with fixed / variable-speed pumps. These applications are limited to the scale of a case study.

The first approach of dynamic programming applied successfully on a large-scale system can be found in Fallside and Perry (1975). Although the method originated in the mid-1970's, it really became operational at the end of the 1980's, after successive improvements. These two researchers derived an hierarchical price-decomposition algorithm for the optimisation of a water supply system in the Cambridge area (United Kingdom), chosen because it was already extensively fitted with telemetry equipment. The objective of their optimisation was to provide an automatic scheduling of pumps, to meet consumption demand at minimum cost. Their project produced a predictor-optimiser control scheme in which consumptions were predicted 24 hours ahead and a minimum-cost pumping schedule was calculated for these predictions. The target application has been decomposed into several sub-systems and an optimisation was conducted around the operating point on each of these subsystems. However, this attempt did not achieve satisfactory results in the long term which lead Coulbeck (1977) to offer a more general and improved version of the original decomposition techniques developed by Fallside and Perry (1975). Once the initial method became more robust and reliable, Fallside (1988) conducted a study to identify what would be the requirements to implement the same methodology on water distribution systems not as highly monitored as the one they used initially, since these systems were clearly the norm in United Kingdom at the end of the 1980's. They reached the conclusion that without the same degree of telemetry as the one used in the Cambridge area, it would be difficult to successfully apply the initial method they developed. These observations discouraged any widespread use of this method in other water distribution systems in the country.

In Spain, almost simultaneously with Fallside and Perry's research, Solanas (1974) created an approximation of an existing dynamic programming algorithm which was improved until 1988, and finally successfully implemented on the Barcelona distribution network (Solanas and Montoliu, 1988). This approach was preferred to existing decomposition techniques because of the highly interconnected nature of Barcelona network which makes inefficient the use of these techniques.

In the 1980's, one of the most relevant applications was the implementation of the model developed by Carpentier (1983), for the southern areas of Paris distribution network. His method was based on the use of dynamic programming techniques which take into account the particular star structure of this supply system. The tool he describes calculates the optimal control of valves and pumps, with regard to economic criteria. It is based on decomposition-coordination, aggregation-disaggregation and dualization methods. The process of optimisation is split into two steps: one offline, for the main calculation; the other online which uses the main calculation and the actual reservoir levels to realise an hourly feedback. Carpentier's method has since been used successfully on other areas of Paris distribution system. For instance, Alla and Jarrige (1988) present economical results concerning the use, since 1985, of this system on Paris western areas. On the positive side, they highlight that the benefits which can be realised with the optimised sequence, lie between 0 and 5% of the total operational cost of this particular part of the network. However, on the negative side, these savings are expensive in terms of safety, since reservoir levels are often reaching their lower emergency levels during periods where the electricity tariffs are high. Furthermore, pump and valve operations occur very often which increases operator workload significantly, since this tool does not provide automatic control yet.

As discussed above, dynamic programming has been extensively used in the non-linear models occurring in the field of the operational optimisation of water distribution networks. However, while it is a good method for systems with less than five reservoirs, it requires much computer time for a system with high dimensionality. Consequently, to apply dynamic programming to large-scale application, it has been necessary to develop secondary approaches like the decomposition approach or the aggregation-disaggregation model mentioned earlier. Although the combination of dynamic programming with one of these two methods has proved to be successful for some special network structures, the convergence cannot be guaranteed for networks that are more complicated. At the end of the 1980's, the need to develop approaches that had a more general field of application than the existing ones became imperative.

Tatjewski's work (1988) constitutes the first attempt to develop formulations that are more general. This researcher preferred to develop a sub-optimal approach to the scheduling of reservoir levels for a multi-reservoir water distribution network. His work started from the observation that solving rigorously such an aggregated problem with rigorous algorithms is only possible for networks with very few reservoirs, due to computational requirements. Consequently, he proposed an approach where he approximates the non-continuous curves of pumping stations, by continuous ones. In other terms, he eliminated integer variables from the problem.

This leads to a problem which only approximates the original problem, but is radically easier for numerical solution. To solve the resulting non-linear discrete time optimal control problem with constraints, he used standard mathematical programming algorithms. However, it appears in reality that Tatjewski's method is only applicable to a very particular class of distribution networks, in which pumping stations consist of identical fixed-speed pumps connected in parallel. Since the approximation process is not automatic, a good knowledge of the network is also required to apply the method. Hence, the first attempt to develop a general formulation for the scheduling problem applied to water distribution networks fell short.

Soon after, a new trend arose due to improvement of computational resources which enabled the use of LP to solve for the pump-scheduling problem. Researchers gained interest in LP because of the availability of new robust and efficient methods to solve for this class of problems. Jowitt *et al.* (1988) described for instance the use of LP for the real-time forecasting and control of water distribution systems. Jowitt and Germanopoulos (1992) continued to work in this particular field and produced a new paper in which they combined the use of LP with an extended-period network-simulation approach. Their research was aimed at determining the minimum cost schedule of pumping, on a 24 hours basis, for a given distribution network. The extended-period network-simulation model was used to provide parameter estimates of the linearised network equations and constraints, before they are sent for optimisation. Their study showed that LP could be used for the solution of the scheduling problem in a water-supply network. The method was applied successfully to an existing network, for which substantial savings on the pumping costs have been achieved.

In the long term however, it appeared that the resulting accuracy and reliability of such models could be quite poor. Consequently, one of the latest approaches consists in taking fully into account the non-linear relationships which are part of any pump-scheduling problem. As for LP, the appearance of improved Non-Linear Programming (NLP) algorithms on the market, convinced researchers to rather apply NLP techniques to solve for this problem. Chase and Ormsbee (1989), Lansey and Zhong (1990) and Brion and Mays (1991) linked network-simulation models with non-linear optimisation algorithms to determine optimal operations. Yu *et al.* (1994) also proposed a method based on NLP to determine the optimal operation of general water distribution systems containing multiple sources and reservoirs. Their method does not require any network simplification, since full network simulations are performed on the basis of optimal strategies calculated by the NLP algorithm they used. These researchers state that they could achieve a profit of approximately 10 % on the operational costs, for a 50 nodes case-study network.

As NLP techniques are recent, few references are available in the literature demonstrating successful application of these methods to existing installations. The work conducted in this field is still at the stage of research. For instance, recently, a number of European research programmes have attempted to produce tools applicable to a variety of water utilities across Europe. The most advanced one, WATERNET, aims to demonstrate the benefits of introducing optimisation techniques to the water distribution sector. With their work, Cembrano *et al.* (2000) are part of this research programme. Their work proves the advantages of optimal control methods on the water distribution network of a Portuguese district. However, although they integrated their model successfully within the existing supervisory control system, they have not been able to obtain integer strategies for pumping operations which is the final objective of the WATERNET programme.

It is nowadays clear that recent developments in the field of optimisation allow the resolution of water distribution problems, without applying major simplifications or decomposition methods on the elements they include. Among researchers who began to explore this new path, relatively few have reached a point where it is possible to use predictive techniques to obtain optimal operation; using one optimisation step only, in a multi-reservoir system fitted with several binary valves and pumps, keeping the binary nature of some variables as an essential part of the formulation. The present project proposes an original approach to address this important issue.

3.2 Model Predictive Control techniques

AS described by Luyben (1990), most chemical processing plants were run essentially manually prior to the 1940's. Only the most elementary types of controllers were used. Many operators were needed to keep an eye on the numerous variables on the plant. More than often, large tanks were used to act as buffer reservoirs between various units in the plant. These tanks enabled a reduction of some of the dynamic disturbances resulting from manual operations, by isolating one part of the process from problems occurring in other parts. With increasing labour and equipment costs, and with the development of higher capacity and higher performance processes, in the 1940's and early 1950's, it became impossible to run plants without automatic control equipment. At this stage, controllers were employed with little consideration for the dynamics of the process itself. In the 1960's, chemical engineers began to apply dynamic analysis and control theory to processes. In addition to designing better control systems, processes and plants were designed or modified so that they became easier to control. This concept of examining the many parts of a complex process, with all of the interactions included, and finding ways to control the entire system is called system engineering.

The rapid increase in the energy prices in the 1970's gave to system engineering and effective control systems a new impetus. The design and redesign of many types of equipment, to reduce energy consumption, resulted in more complex and integrated processes that became much more interactive. Consequently, the challenges to process control engineers have continued to grow over the last 40 years.

As MPC is probably the most general way of putting the control of a process within the time domain, it has gained real success in the last few years. It is now used widely by chemical engineers worldwide. This section is divided into four subsidiary parts. Firstly, the spirit of Model Predictive Control is presented. Secondly, a comprehensive literature review on this particular subject is conducted. Then, implementation strategies for the most common families of MPC controllers are given. Finally, one of the most common MPC algorithm is presented with full details.

3.2.1 Introduction to Model Predictive Control

MPC does not designate a specific control strategy but refers to numerous control methods developed around common ideas (Camacho and Bordons, 1995). Three principles always appear for formulations based on MPC techniques. Firstly, a model is created so that the output of the process to be optimised can be predicted over a future period (referred to as a prediction horizon). The use of an explicit model allows the controller to deal with all of the significant features of the process dynamics. Secondly, a penalty function (also referred to as an objective function) has to be defined and minimised. It generally takes into account constraints intrinsic to the problem. Constraint violations are therefore less likely, and optimality is preserved in the presence of constraints. Finally, the problem is solved over the prediction horizon. At each new time step, the horizon moves one step forward.

This strategy allows anticipation of disturbances, handles dead time and is multivariable in nature, driving the plant closely to the desired trajectory. Several MPC approaches are based on these principles. The difference in the formulations of these families of problems only resides in the way the model is built and the objective function defined. As every process is unique, it needs a specific model to represent it. The objective function to be minimised is defined according to the objectives of the optimisation, on a particular process. It is therefore also specific to the problem considered.

MPC presents numerous advantages over other methods. Firstly, it is applicable to a wide range of processes, including the more complex ones.

Secondly, even in the most complex cases, tuning the controller can usually be done. It is also relatively easy to use when dealing with multivariable problems, or treating intrinsic constraints of the process. Finally yet importantly, MPC intuitive concepts make it particularly attractive to staff with limited skills in the control area.

Although MPC has numerous advantages, it also presents some inconveniences. One of the disadvantages refers to the amount of computation required for processes which see their dynamics changing with time. In this particular case of MPC, referred to as adaptive control, heavy computational work may be conducted at every time step. The operational optimisation of a water distribution system has adaptive control features. However, the computing resources available in control rooms supervising these systems are generally reserved for tasks other than the control strategy itself (for instance, logging, alarms or dialogue with the operators). Therefore, the use of these strategies usually requires the installation of additional computer power.

However, the most important limiting factor found with MPC is surely the ability to find an appropriate model of the process to be optimised. The benefits given by the optimisation depend strongly on a good fit between the real process and the model used. The definition of an adequate model is therefore the most important task when using MPC.

3.2.2 Literature review

AUTOMATIC control is a driver of industrialisation and process upgrade. Control theory has been subject to impressive developments throughout the 20th century. Originating with the stability theory of Lyapunov in the early 1900's, it developed into the modern control theory through the conception of a three-terms controller in the 1910's (later known as PID), through the discovery of feedback amplifier in the 1920's and through the development of Nyquist and Bode charts in the 1930's. Developments reached new edges with Richard Bellman's work on the principle of optimality from 1957 and the discovery of the statistical filtering theory by Dr Rudolph Kalman and Dr Richard Bucy in the early 1960's (Bellman, 1957; Kalman and Bucy, 1961).

In the literature, the first appearance of research work related to MPC can be traced as long ago as the late 1950's. The early developments of control strategies developed by Gregory were linked to the design of autopilots for new military aircraft and rockets (Gregory, 1959). Considerable research activities were conducted at this time to develop self-adaptive systems that would automatically adjust to changing flight conditions.

Concerning the use of MPC for industry purposes, 1959 seems also to be the year of reference. For instance, Aström and Wittenmark (1984) have been able to determine the 12th March 1959 as the first day when a computer control system was introduced at a Texaco refinery in Texas. In the course of the same year, an on-line computer control and optimisation system was also implemented in one of the Texan Union Carbide ethylene's production units. Unfortunately, because of a non-existent theory and limited hardware resources, these attempts proved unsuccessful. As described by Seborg *et al.* (1986), a more rigorous theoretical basis was needed to transform these attempts from failure to success.

The 1960's and 1970's saw researchers developing modern control theories, like the receding horizon principle (Propoi, 1963), the use of quadratic costs associated with linear constraints, as well as a moving horizon (Rafal and Stevens, 1968) and finally, the self-tuning control system (Aström and Wittenmark, 1973). Associated with a decrease in price of digital control hardware, these new theoretical developments really strove for the use of MPC within the industrial world.

As described in Jacobs (1981), controversies were born at this time because of the use of numerous alternative definitions for similar algorithms, creating a lot of confusion. This can be explained because, like many technical inventions, MPC was implemented in industry long before a clear understanding of its main theoretical properties was available. During a conference on parameter estimation in 1976, Richalet *et al.* (1976) were the first researchers to give a comprehensive overview of successful MPC industrial applications, based on a new approach of their own. They managed to establish a consensus in the control area, by gathering the numerous existing approaches, still subject to discussion and controversy, under a unique problem. Their work was later summarized in a 1978 *Automatica* paper (Richalet *et al.*, 1978) and was referenced as Model Predictive Heuristic Control (MPHC). As the control law they used was non-linear and could not be expressed by means of a transfer function, it was referred to as heuristic, but nowadays the term non-linear MPC would be more appropriate.

The solution software derived from this approach was referred to as IDCOM, an acronym standing for IDentification and COMmand. It has been applied successfully on several industrial applications like the distillation column of a fluid catalytic cracker unit or for the steam generation of a power plant. Several distinguishing features characterize the IDCOM approach. Firstly, a model is defined specifically for the plant to be optimised and is subject to a perturbation expressed mathematically as an impulse response, with inputs and internal variables chosen linear.

Secondly, a quadratic performance objective is introduced over a finite prediction horizon. The formulation of this mathematical function includes also input and output constraints. Finally, the desired behaviour of future plant output is specified by a reference trajectory. These three features have been taken up by all of the research teams who continued to improve this initial method. As described in the paragraph introducing MPC techniques, they constitute the MPC framework, as it is still known nowadays. For the IDCOM application, the optimal inputs are determined by use of a heuristic iterative algorithm.

Following the Richalet *et al.* (1978) work, Cutler and Ramaker introduced in 1979 a new method referred to as Dynamic Matrix Control (DMC). In this approach, a dynamic process model is also created in order to predict the effect of the future control actions on the output, but the impulse response of IDCOM is replaced by a step response. Details of their work can be found in the proceedings of two conferences held between 1979 and 1980 (Cutler and Ramaker, 1979; Cutler and Ramaker 1980).

The initial IDCOM and DMC algorithms represent the first generation of MPC technology. They had an enormous impact on industrial process control and served to define the industrial MPC framework. Academic interest in MPC started only in the second half of the 1980's after two workshops organized by Shell (Prett and Morari, 1987; Prett *et al.*, 1990). Thanks to the research conducted by Morari and Garcia (1982) or Rawlings and Muske (1993), a clear understanding of what MPC is has emerged. Practitioners and theoreticians now have a strong foundation available when working with MPC techniques, although several issues still need to be addressed in the existing framework (Nikolaou, 2001).

Following the initial developments in the field of MPC, adaptive control saw its ideas developed independently from the other MPC families. It soon became an active area of research, and it still is nowadays. Since approximately two decades ago, an extensive interest appeared for feedback control systems that automatically adapt their controller settings to compensate for change in the process behaviour.

As described by Seborg *et al.* (1986), the original idea of this family of MPC was to identify the control parameters of a process as they change and find the new settings for the manipulated variables accordingly. This approach is pushed to its full potential when the control parameters adaptation is conducted online (i.e. in real-time). Unfortunately, an online adaptation is always complex and can lead to poor performance. Furthermore, because of the considerable time it takes to reach a result, it is not always adequate.

The parameters of the plant may have indeed changed to a different condition before the calculation of the new manipulated variable settings reaches its end. Nevertheless, online identification is not essential in the chemical industry, as it is usually possible to predict general plant behaviour with offline tests, and controller parameters can, if necessary, simply be scheduled as conditions change.

Two general categories of adaptive control problems exist, differing essentially in the way the controller parameters are evaluated. The first category of problems refers to situations where the process changes cannot be directly measured or anticipated. On the contrary, for the second category of adaptive control problems, the process changes can be deduced from process measurements. Consequently, if the process is well understood, the controller settings can be adjusted in a predefined manner as process conditions change. This second approach is referred to as gain scheduling and has been tackled by several teams of researchers over the past years (Andreiev, 1981; Aström, 1983; Wong and Seborg, 1985; and Shinskey, 1996).

From the beginning of the 1970's, researchers realised the limitations introduced by gain scheduling, and preferred to concentrate their work on the first family of problems which sounded more promising for industrial applications, even if more difficult to handle originally. For this particular family of problems, the most popular methods are Model Reference Adaptive Control (MRAC), based on the global stability theory; and the Self-Tuning Adaptive Control (STAC), based on the use of a quadratic cost function.

Recent research has demonstrated that, although MRAC and STAC have been developed from different viewpoints, the techniques are closely interrelated, and can even be analyzed from a unified theoretical framework. Examples of applications of MRAC have been reported by Parks *et al.* (1980) in an extensive literature review. However, significant applications of MRAC are still scarce in the process industries, as the problem of showing global stability is a nontrivial one. On the contrary, the STAC approach has received more attention than any other adaptive control strategy over the past decades.

The first developments of STAC arose with Peterka and Aström's work (1973). They developed the predictor-based self-tuning control, by deriving it directly from Kalman's original idea (1958) of a self-optimising control system. This control system, presented for the first time at a conference on parameter estimation, minimises the expected value of a quadratic objective function, for the most recent predicted values, over a given control horizon. Aström and Wittenmark (1973) detailed this approach in an *Automatica* paper soon after.

Following this pioneering research, Hoopes *et al.* (1983); Kraus and Myron (1984); and Bengston and Egardt (1984) contributed to the greater acceptance of this particular adaptive control technique in industry. They developed commercial versions of initially pilot-scale control systems, based on the self-tuning control algorithm. Unfortunately, their algorithms could only handle the problem of controlling linear and time-invariant systems, with unknown or uncertain models, and subject to random dynamic disturbances. Linear and time-invariant systems barely occur in chemical processes, and these applications were soon shown to lack robustness.

The original method of self-tuning control was therefore further developed by Ydstie *et al.* (1985). The method they developed enabled their regulator to cope more easily with the features of industrial control in the chemical industry. The most significant modifications to Peterka and Alström's, and Aström and Wittenmark's work concerns the use of a variable forgetting factor for the handling of past data, the replacement of the minimum variance control objective by an extended horizon controller and the expression of the algorithm in an incremental form to eliminate the offset. The use of an extended horizon gives more time to the controller and therefore allows it to look beyond process time delays which was the limiting factor of the previous approach. Despite all of these improvements, it appeared that this method was still not applicable to suit a majority of existing chemical processes.

Clarke *et al.* (1987) are the first researchers who designed a robust regulator capable of controlling either a simple plant (e.g. open loop stable), or a more complex one, such as open loop unstable and having variable dead time. Their method, known as Generalized Predictive Control (GPC), presented in two separate papers (one for the presentation of the algorithm, the other for the extension and interpretations of the method), is still one of the most popular in the adaptive control family, both in industry and academic worlds. As summarized by Camacho and Bordons (1995), the basic idea of GPC is to calculate a sequence of future control signals by minimising a multistage cost function defined over a prediction horizon. The function to be minimised is the expectation of a quadratic function measuring the distance between the predicted system output and some predicted reference sequence over the horizon, plus a quadratic function measuring the control effort.

GPC has many ideas in common with the other predictive controllers previously mentioned, since it is based on the same initial concepts. It has also some differences. The most noticeable one concerns the possibility to weight control increments in the cost function. This leads to a greater variety of control objectives compared to other approaches.

Since the original applications, GPC techniques have been successfully applied in industry. An account of these applications can be found in Clarke (1988). The main stream of research has concerned the integration of non-linear process models within the formulation. The non-linearity of chemical processes is well documented and the reader can refer to Bequette (1991), Allgöwer and Doyle (1997), Ogunnaike and Wright (1997) and Baffi *et al.* (2002) for further information on this particular subject.

With further improvements still going on in the field of computation, it is reckoned that MPC will experience an even greater dissemination in both academic and industrial worlds in the future. However, it is generally acknowledged that MPC algorithms can still be improved. For instance, it has to become even more robust and therefore gain more reliability. As described by Nikolaou (2001), a discussion with MPC practitioners reveals indeed that the most expressed improvement need has been to increase the time for which the MPC controller is on-line. The reasons why the controller is switched to manual are numerous. Generally, besides problems with the MPC theory directly, practitioners encounter problems of model development, computation, programming or user interfacing. Nikolaou's study (2001) reveals that the MPC algorithm becomes nowadays only one of the numerous elements which can make an MPC product successful or not in the marketplace. This means that improvement efforts have now to be made on several fronts simultaneously.

3.2.3 Model Predictive Control strategy

THE methodology of all of the controllers belonging to the MPC class of algorithm is characterized by the strategy summarized on Figure 3.1. As presented in Camacho and Bordons (1995), the future outputs for a determined horizon N (called the prediction horizon) are predicted at each instant t using the process model. These predicted outputs depend on the known values for past inputs and outputs, up to the instant t, and on the future control signals which are those to send to the system and to be calculated. The set of future control signals is calculated by optimising a determined criterion, in order to keep the process as close as possible to the reference and conform to constraints, when they exist. This criterion usually takes the form of a quadratic function of the errors between the predicted output signal and the specified reference trajectory (or setpoint), over an N-step horizon. Including the control effort in the objective function is optional, and often this term is rather replaced by an economic objective.

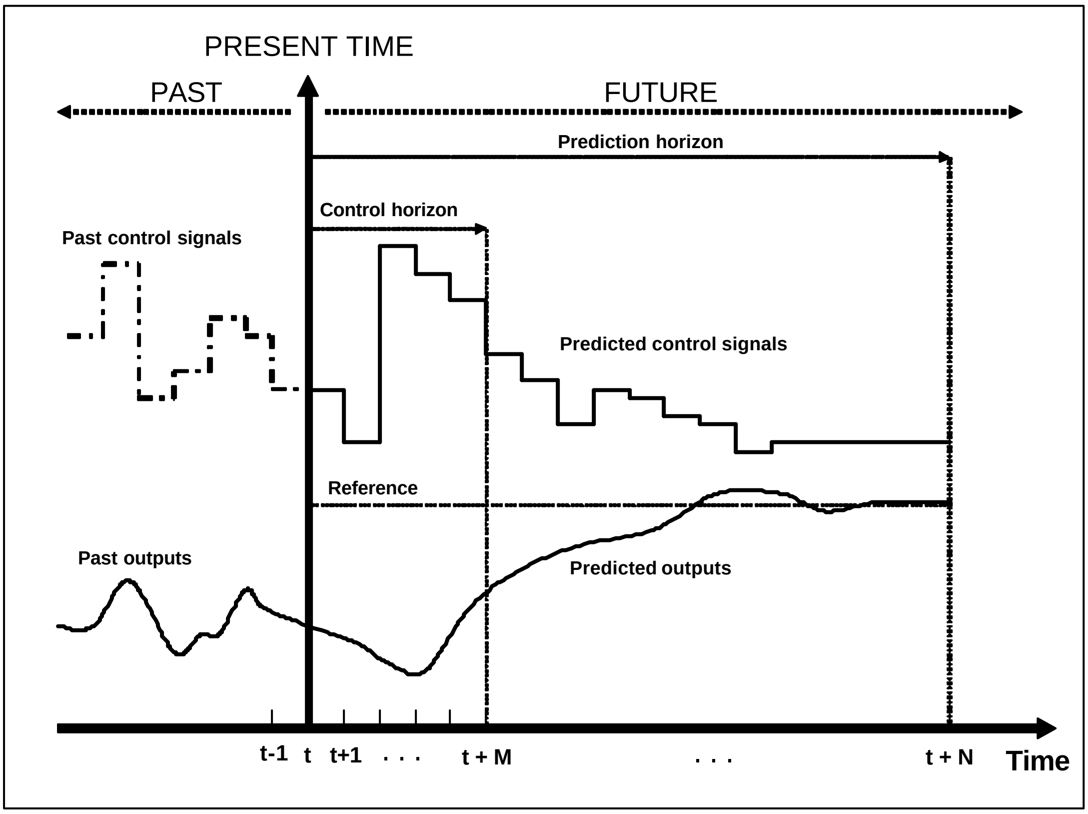


Figure 3.1: Model predictive control strategy

The length M of the manipulated variables sequence represents the control horizon. When the horizon is moved forward, only the first element of this predicted control sequence of M steps is sent to the real process. Since the rest of this control sequence was calculated for, now outdated, plant conditions, it is not adequate anymore and needs to be brought up to date with the new plant conditions. This explains why the rest of control signals stored in the control horizon are rejected.

The whole sequence described above is repeated on each time-step. The period between controller intervention is commonly referred to as a time step or sampling instant. When the previous strategy is applied to a constrained non-linear model, an iterative optimisation method has to be used for each time-step to solve for the set of future control signals. Figure 3.2 summarizes the MPC calculation undertaken at each time-step.

The structure shown in Figure 3.3 presents how to implement this strategy with a controller. This concept is behind most of the model predictive controllers. Figure 3.3 shows that the prediction of the system output is based on two different components:

- the free response which represents the predicted behaviour of the output, based on old outputs and inputs, assuming a future control action of zero;

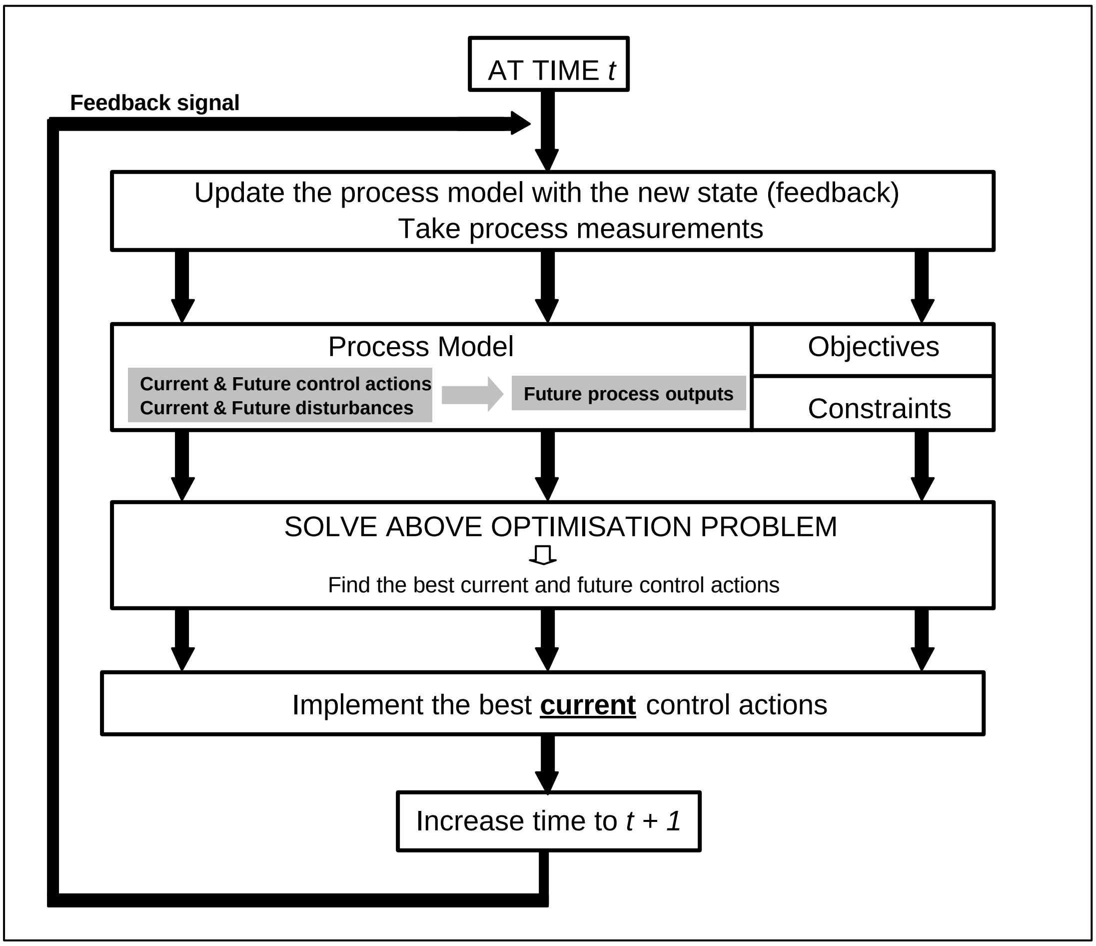


Figure 3.2: MPC calculations undertaken at each control execution

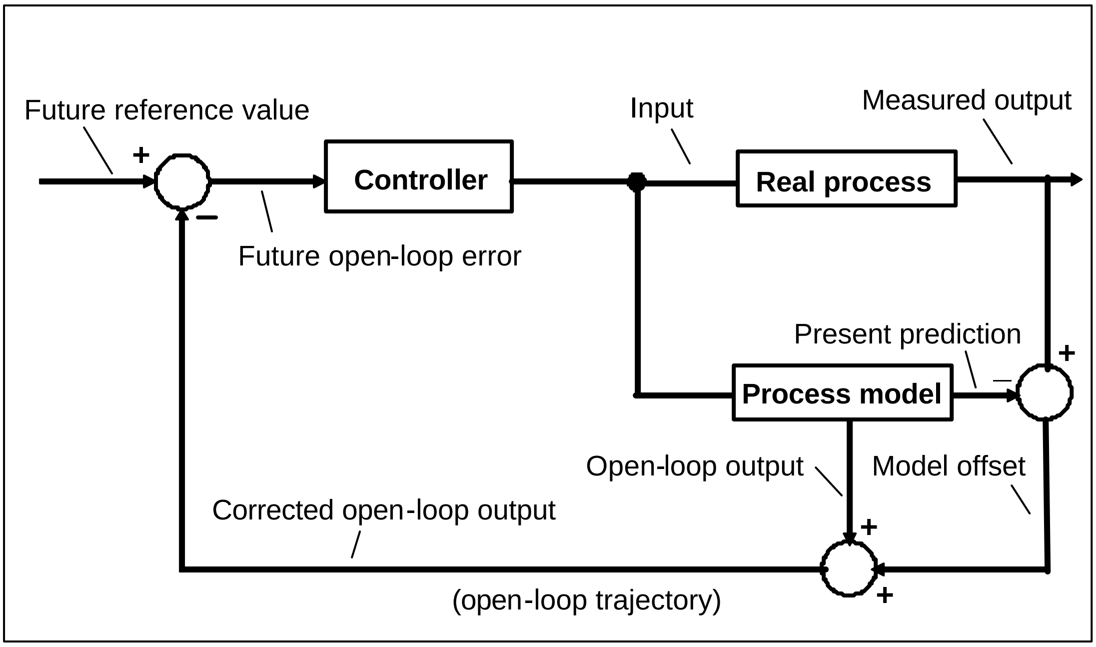


Figure 3.3: Basic structure of Model Predictive Control controllers

- the forced response which represents the additional component of the output resulting from the optimisation.

The total prediction is the sum of both components. The future error term is deduced by subtracting the future reference value with this total prediction term. It is sent to the optimiser for minimisation.

Although all of the MPC formulations use this common framework, slightly changes are introduced, depending on the adaptive scheme chosen for the controller: gain scheduling, model reference adaptive control, or self-tuning adaptive control.

3.2.3.1 Gain scheduling

WHEN it is possible to find measurable variables that correlate well with changes in the process dynamics, the control strategy is referred to as gain scheduling. A block diagram of this scheme is presented Figure 3.4.

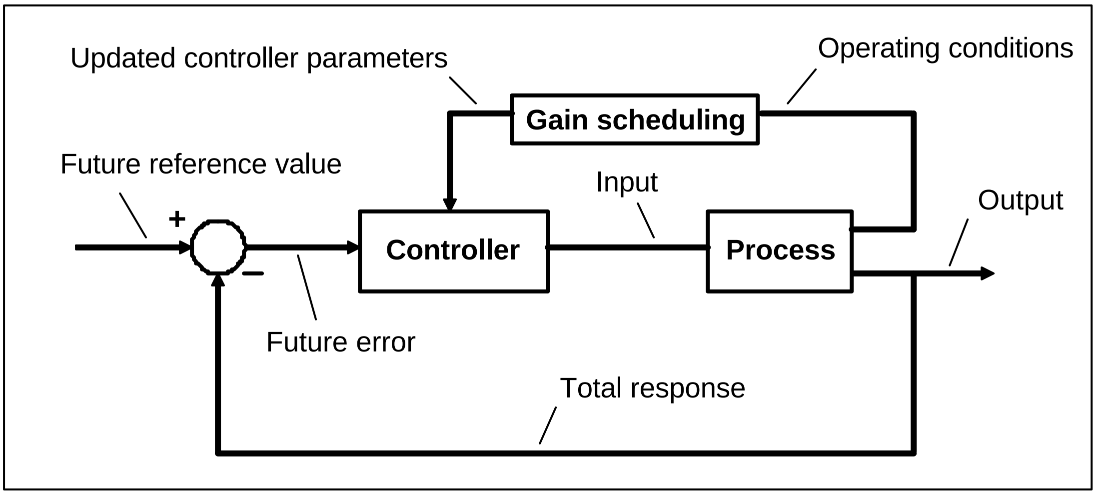


Figure 3.4: Structure of gain scheduling

Such variables can be used to change the controller parameters in a predetermined manner, by monitoring the operating conditions. The most common approach is to store in a table several sets of process parameters, for different plant conditions; and choose amongst these sets, the most adequate settings as operating conditions vary. As no parameter estimation occurs for this particular approach, the controller adapts quickly to changing process conditions which constitutes the main advantage of this method. However, a good knowledge of the process is required to apply it which constitutes one of the main drawbacks.

3.2.3.2 Model Reference Adaptive Control

HISTORICALLY, the first attempts to design stable adaptive control systems appeared with the development of the MRAC approach. In MRAC, the basic objective is to force asymptotically the output of a process towards a given reference. A block diagram of this scheme is presented Figure 3.5.

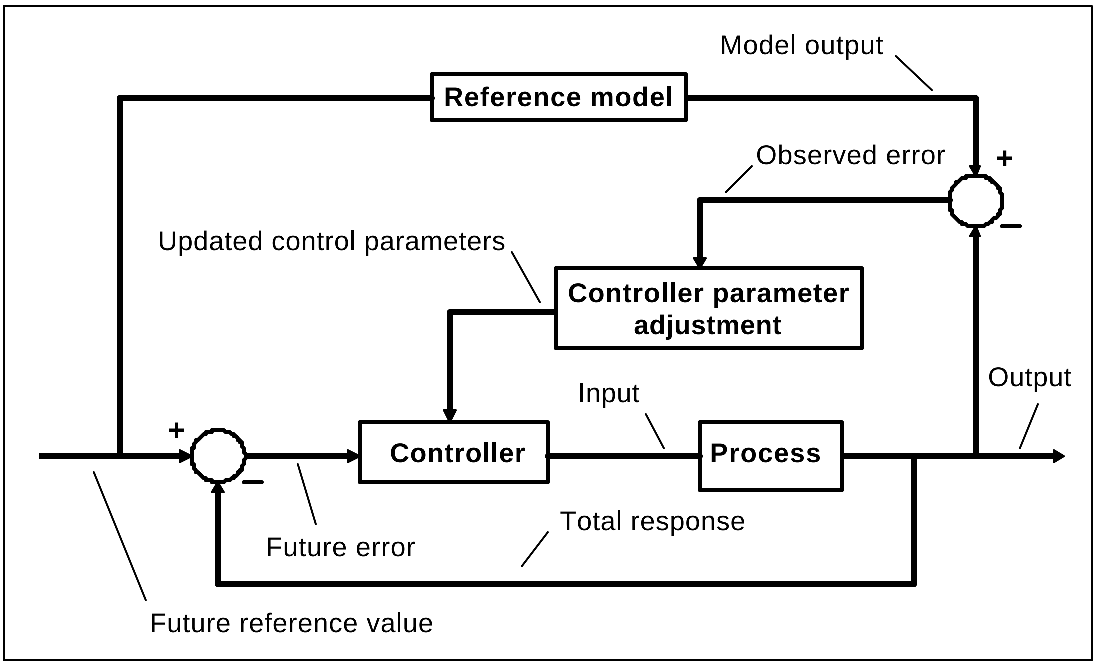


Figure 3.5: Structure of Model Reference Adaptive Control

The key technical problem of this family of applications is to determine the structure of the parameter adjustment mechanism so that the overall system is globally stable. In this particular context, the stability criterion is equivalent to having process input and output remaining bounded for all time-steps, with the observed error tending to be equal to zero when time goes to infinity.

3.2.3.3 Self-Tuning Adaptive Control

STAC can be defined as the strategy which estimates model parameters and adjusts the controller settings with these parameter estimates. The self-tuning control configuration is flexible enough to accommodate a large number of parameter estimation techniques and controller design strategies. A block diagram of this scheme is presented in Figure 3.6.

It is common to classify self-tuning regulators (STR) into two general classes, referred to as explicit or implicit classes. In the explicit approach, a process model is used and the update of control moves is based on an estimation of the model parameters.

In this particular case, the parameters do not appear in the control law. On the contrary, for the implicit class of problems, the model parameters are used in the control algorithm.

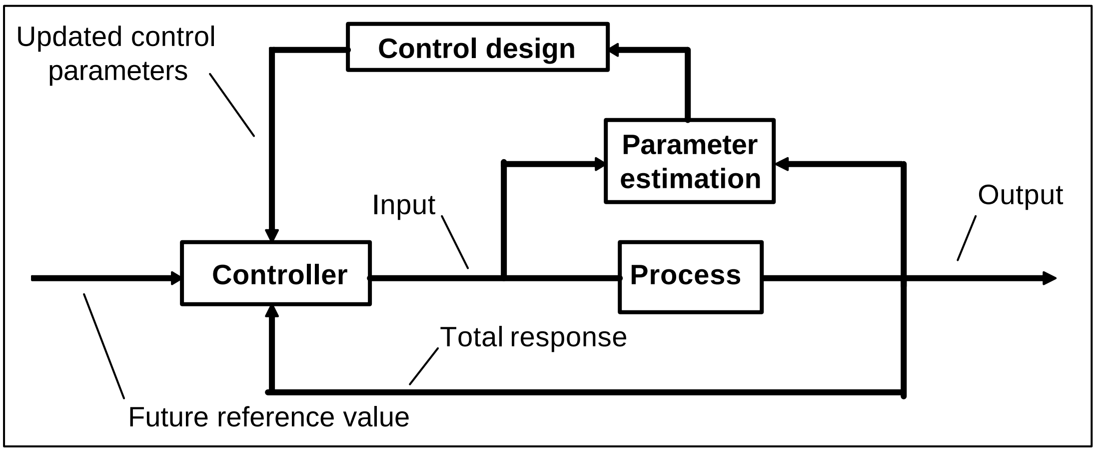


Figure 3.6: Structure of Self-Tuning Adaptive Control

The original process model is here converted to a predictive form that allows the future process outputs to be predicted from current and past input and output variables. For both classes of STR, it is the model which determines the effectiveness of the controller. Consequently, the most essential feature here is to have robust model identification. This is particularly true for Generalized Predictive Controllers which are part of the STR family

3.2.3.4 General recommendations for all formulations

MPC systems based on non-linear models present an increased complexity due to two main factors. Firstly, the non-linear programming required for the solution of the problem does not produce exact solutions, but rather solutions that are optimal within a pre-specified tolerance. The solution can even be only locally optimal if the problem is non-convex. Secondly, even if the global optimum of the optimisation problem is assumed to be reached, the controller behaviour can show patterns that would not be expected initially.

As Aström and Wittenmark show in their book (1995), an adaptive controller is inherently non-linear. It is therefore more complicated to use adaptive controllers than fixed-gain controllers. Consequently, before relying on adaptive control techniques, a study has to be conducted to check if the control problem could not be solved by using a fixed-gain approach. Figure 3.7 represents the procedure the two authors propose to undertake to define the type of controller to use for a given problem. When the choice of an adaptive controller is made, two elements must be taken into consideration. Firstly, as the proposed optimal future control actions are directly evaluated by using the model, the model chosen must be able to represent process dynamics precisely, so that the prediction of future outputs is accurate. It is also better to work with models which are easy to understand and to implement.

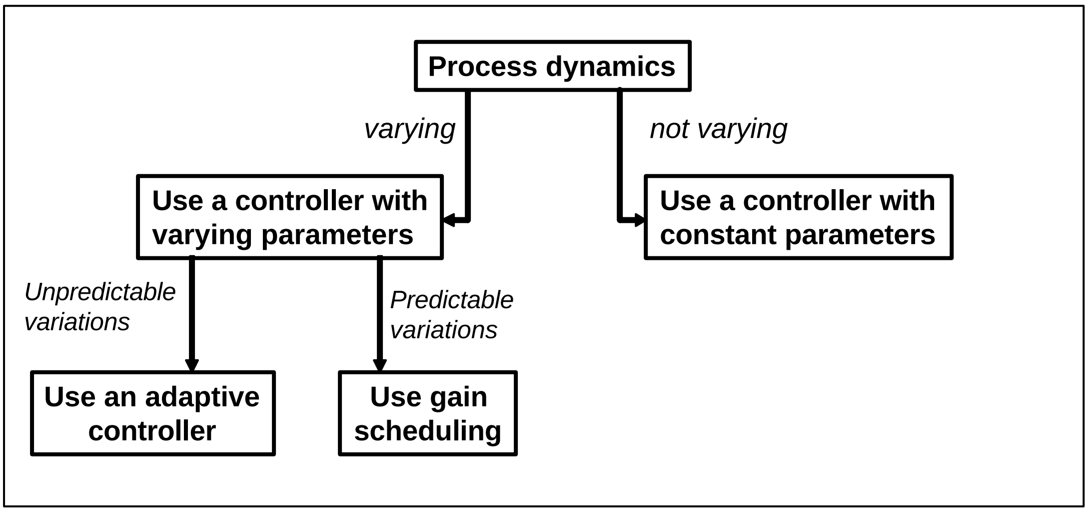


Figure 3.7: Procedure used to determine the most adequate type of controller

Secondly, the optimiser is also another fundamental part of the strategy, as it provides the control actions. When inequality constraints are present, the solution is obtained by means of computationally taxing optimisations. Regarding the last statement, it is easy to realise that adaptive control techniques require time-consuming optimisation algorithms to achieve satisfactory results. Therefore, it is only applicable to slow processes such as chemical processes, where the changes in state variables do not usually occur rapidly. Water distribution networks are part of this category of processes. The objective of applying adaptive control techniques to these systems is mainly to predict well in advance the binary valve and pump sequences needed to respond to demand objectives (users' consumption). The control sequences are optimised to guarantee at the same time quality objectives (water quality must conform to WHO standards), security objectives (sufficient storage in the reservoirs) and financial objectives (making the best use of off-peak electricity).

Non-linear GPC seems to be the most appropriate algorithm to be used here. Besides the use of Non-Linear Programming (NLP) for the resolution of the problem, the specific nature of water distribution networks (e.g. presence of binary choices in the optimisation) requires state-of-the-art computer codes which can solve for a combination of continuous and binary variables.

3.2.4 Model Predictive Control algorithm

TO implement their IDCOM approach, Richalet *et al.* (1978) chose a black-box representation for the process they wanted to optimise. In the case of a Multiple Inputs, Multiple Outputs system (MIMO), process inputs are divided into manipulated variables (MV) which the controller adjusts; and disturbance variables (DV) which are not available for control. Process outputs are referred to as controlled variables (CV).

With this approach, presented in Figure 3.8, the process inputs influence the process outputs directly. The relationship between process inputs and outputs is assured by the model used to represent the process.

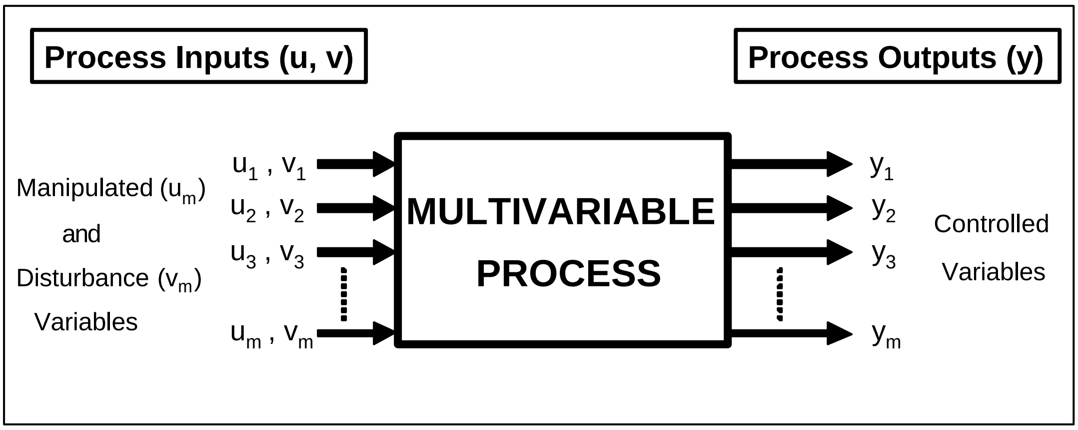


Figure 3.8: Richalet et al. black-box approach for Multiple Inputs, Multiple Outputs systems

At the dynamic optimisation level, an MPC controller must compute a set of MV adjustments that will drive the CV to desired operating point, without violating the constraints. Industrial MPC controllers generally use three different methods to parameterise the MV sequence: multiple moves, single moves or use of a basis function (Figure 3.9).

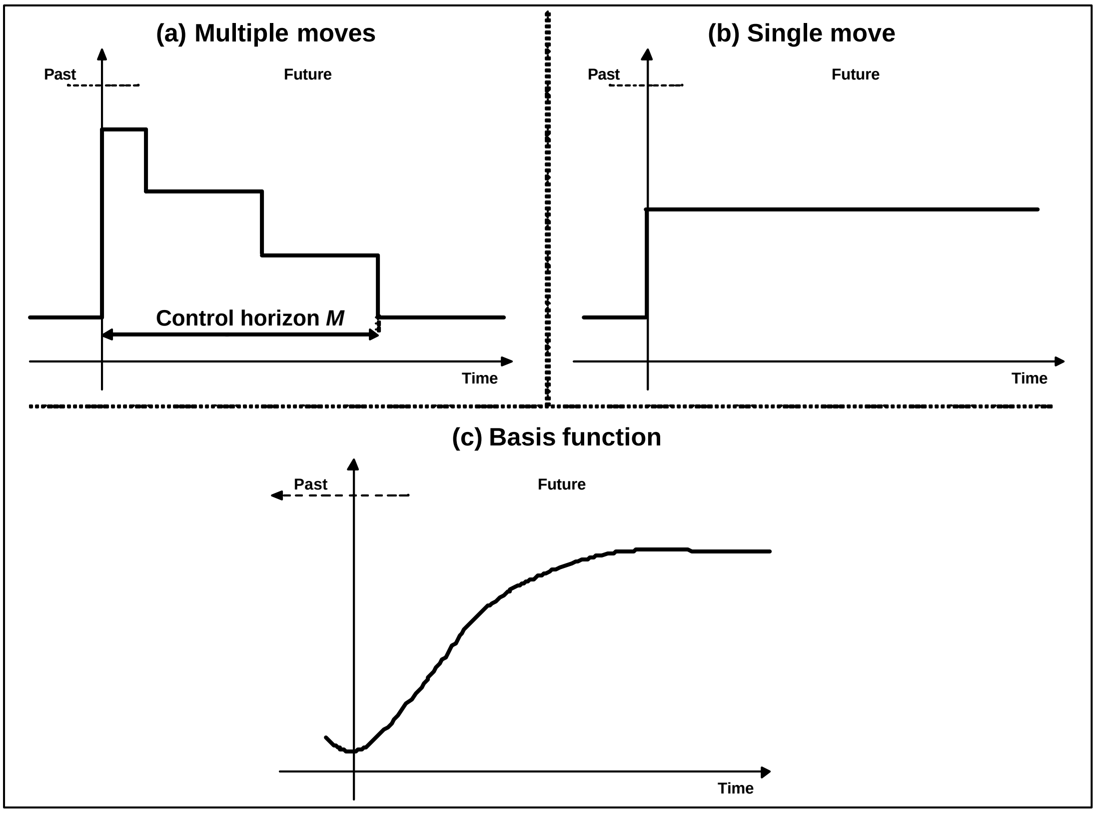


Figure 3.9: Input parameterisation options

For an easier representation of the control algorithm, a Single Input, Single Output system (SISO) is preferred to the more complex MIMO approach. The SISO system is taken at steady state, with inputs u, v and output y, and the simulation is conducted over a prediction horizon of length N. This value is a basic tuning parameter for MPC controllers and it must be taken big enough to take into account the full effects of the calculated MV's moves.

From now on, only the GPC controller discussed previously will be considered, since it is the one which was chosen for the particular case of operational optimisation of a water distribution network. The GPC algorithm can be described by implementing the following strategy:

For $t \in [1, N]$; find at time t, all the $u(t+i-1)$ and $u(t+i-2)$ (for $i = N_{\min}, \ldots, M$), such that:

$$\min \sum_{i=N_{\min}}^{N_{\max}} W_i \left[y(t+i)_{|t} - y_{SP\ |t} \right]^2 + \sum_{i=N_{\min}}^{M} R_i \left[u(t+i-1)_{|t} - u(t+i-2)_{|t} \right]^2 \quad (3.1)$$

When the backward difference operator (Δ) is used, this formulation becomes:

For $t \in [1, N]$; find at time t, all the $u(t+i-1)$ and $u(t+i-2)$ (for $i = N_{\min}, \ldots, M$), such that:

$$\min \sum_{i=N_{\min}}^{N_{\max}} W_i \left[y(t+i)_{|t} - y_{SP\ |t} \right]^2 + \sum_{i=N_{\min}}^{M} R_i \left[\Delta u(t+i-1)_{|t} \right]^2 \quad (3.2)$$

In equation (3.2), the term $\sum_{i=N_{\min}}^{N_{\max}} W_i \left[y(t+i)_{|t} - y_{SP\ |t} \right]^2$ represents the quadratic deviation of the controlled variable y from its reference value y_{SP}. It is calculated at time t, until the prediction horizon N is reached. $N_{\min}$ and $N_{\max}$ are respectively the minimum and maximum costing horizons. $N_{\min}$ can generally be taken as 1, whereas the value for $N_{\max}$ is fixed, so that it takes into account all the responses which are significantly affected by the current control actions at t. It is obvious that $N_{\max}$ is taken less than or equal to N. Once fixed, $N_{\min}$ and $N_{\max}$ give the length of an interval which is referred to as the moving window. At every new time step, this window of fixed length is moved forward by one unit, and a new calculation takes place. Usually, the length of this window is fixed before the optimisation is run and does not change until the process is completed.

Industrial MPC controllers, and therefore GPC controllers, have four basic choices for process output references: a setpoint, a zone, a reference trajectory, or a funnel. The differences between these behaviours is summarized Figure 3.10.

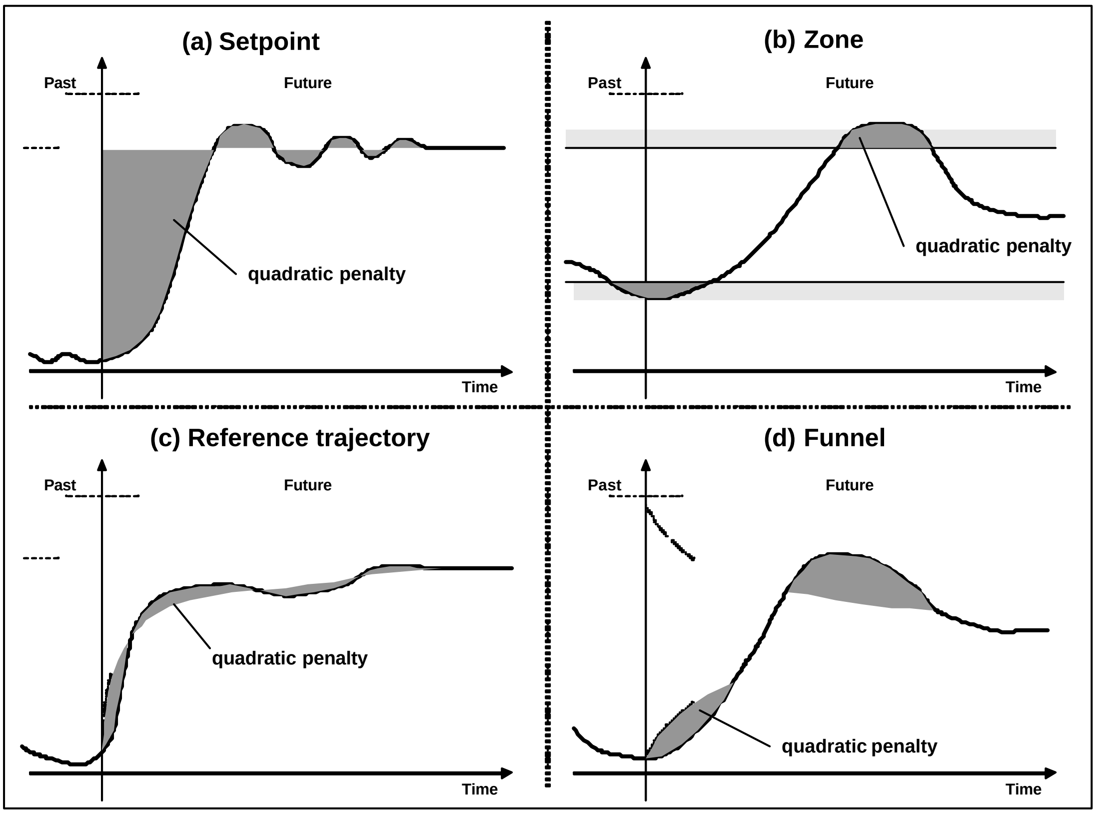


Figure 3.10: Different types of process output references

In the expression of the quadratic penalty $\sum_{i=N_{\min}}^{N_{\max}} W_i \left[y(t+i)_{|t} - y_{SP\ |t} \right]^2$, W_i represents a time-dependent weighting sequence, assumed to be positive and semi-definite.

Still in (3.2), the term $\sum_{i=N_{\min}}^{M} R_i \left[\Delta u(t+i-1)_{|t} \right]^2$ is referred to as the move suppression term. It provides a softer control by preventing the MV's values from changing tremendously between two successive steps. It is an optional term. In the particular case of the operational optimisation of a water distribution network, an economic factor will be preferred to this control term. The move suppression is calculated at time t, for a limited sequence of manipulated process inputs. If $N_{\min}$ is equal to 1, this calculation is conducted over a fixed-length interval of length M, referred to as the control horizon. For a traditional MPC formulation, the move suppression term would be evaluated over the complete moving window. In contrast, with a GPC algorithm, the real power of the method lies in the assumptions which state that after an interval $M < N_{\max}$, all of the projected control increments are assumed to be equal to zero.

This idea, borrowed from the DMC of Cutler and Ramaker (1979), is expressed as follows:

For $i \in [1, N_{\max}]$:

$$\Delta u\left(t+i-1\right)_{|t} = 0 \quad \mathit{for} \;\; i > M \tag{3.3}$$

The control horizon M is another basic tuning parameter for MPC controllers and increased control performance is obtained when its value increases. However, these improvements can only be achieved at the expense of the computational time.

In the formulation of the move suppressor term, R_i represents a time-dependent weighting sequence.

The minimisation of equation (3.2) is subject to the respect of three sorts of inequality constraints:

For $t \in [1, N]$; respect at time t:

$$y_{\min} \leq y(t+i)_{|t} \leq y_{\max} \quad \mathit{for} \;\; i = N_{\min},\ldots, N_{\max} \tag{3.4}$$

$$u_{\min} \leq u(t+i-1)_{|t} \leq u_{\max} \quad \mathit{for} \;\; i = N_{\min},\ldots, N_{\max} \tag{3.5}$$

$$\Delta u_{\min} \leq \Delta u(t+i-1)_{|t} \leq \Delta u_{\max} \quad \mathit{for} \;\; i = N_{\min},\ldots, N_{\max} \tag{3.6}$$

These constraints can be of three basic types: hard, soft, or setpoint approximation of soft constraint. The differences between these types of constraints are presented Figure 3.11.

It is obvious that the formulation including equations (3.2), (3.4), (3.5) and (3.6) requires the prediction of future outputs $y\left(t+i\right)_{|t}$. This prediction is achieved using a model which takes into account the process inputs and external disturbances. The mathematical form of the process model defines the scope of an MPC algorithm. In the most common cases, mathematical models used for MPC are time-invariant.

For the particular case of water distribution network optimisation, a general non-linear discrete-time state space model has to be used. It can be described as follows:

For $t \in [1, N]$; at time t:

$$x(t+1) = f\left[\,x(t), u(t), v(t)\,\right] \tag{3.7}$$

$$y(t) = g\left[x(t)\right] + e(t) \tag{3.8}$$

The model update step is essential to any MPC algorithm. Here, x represents the set of process states. This set is evaluated from state values, MV's values u; and DV's values, v; all taken at the previous time step.

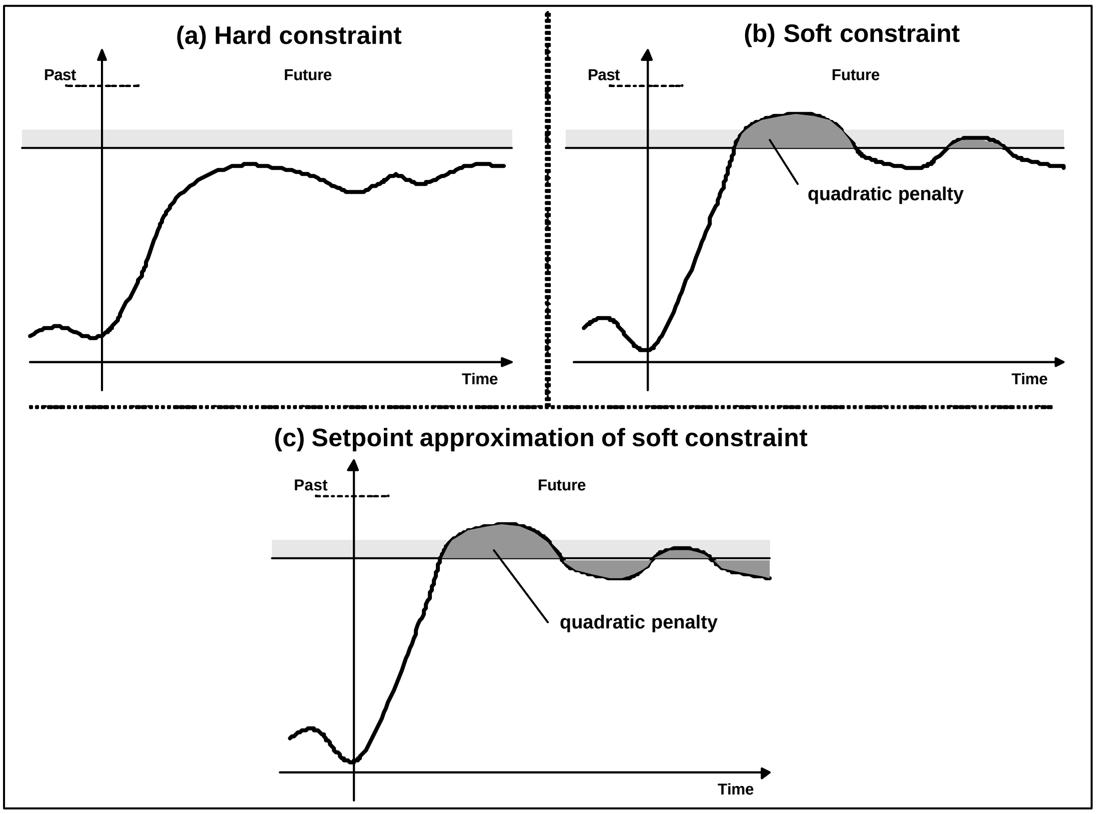


Figure 3.11: Different types of inequality constraints

In (3.8), $y(t)$ represents the value of the process outputs at time t, while $g[x(t)]$ refers to the internal model which describes the dynamics of the controlled process. $e(t)$ is the measurement error which usually represents the gap between the current measured process output and the current predicted output. That is where the feedback enters the control loop.

From (3.7) and (3.8), it can be deduced:

For $t \in [1, N]$; at time t:

$$y(t+i)_{|t} = g\left[x(t+i)_{|t}\right] + e(t+i)_{|t} \qquad for\ \ i = N_{\min}, \ldots, N_{\max} \tag{3.9}$$

Equation (3.9) shows that the prediction of future outputs $y(t+i)_{|t}$ is evaluated at time $t+i$, from the model update $g\left[x(t+i)_{|t}\right]$ and the measurement error e, both estimated with available information at time t. From (3.9), it is clear that the prediction of the future disturbance $e(t+i)_{|t}$ cannot be conducted with certainty. However, with the form of feedback used for the control loop (comparison between predictions and current state), it is equivalent to having a constant output disturbance for all future time. Consequently, the approximation assuming that $e(t)$ is constant, is a convenient simplification.

This gives therefore:

For $t \in [1, N]$; at time t:

$$e(t+i)_{|t} = e(t) = y(t) - g\left[x(t)\right] \quad \textit{for} \;\; i = N_{\min},\ldots,N_{\max} \tag{3.10}$$

From (3.9) and (3.10), it is easily deduced that:

For $t \in [1, N]$; at time t:

$$y\left(t+i\right)_{|t} = y(t) + g\left[x(t+i)_{|t}\right] - g\left[x(t)\right] \quad \textit{for} \;\; i = N_{\min},\ldots,N_{\max} \tag{3.11}$$

Equation (3.11) expresses mathematically that the past model output estimates are fed back to the model. This particular identification method is referred to as output error approach: it results in expressing $y\left(t+i\right)_{|t}$ as part of an Auto-Regressive model with eXogenous inputs (ARX). In an ARX model, the model parameters appear nonlinearly which requires non-linear parameter estimation. This method is considered as a long-range prediction approach, as it looks at the complete process output estimation set. It is therefore particularly suited for the GPC concepts developed previously, unlike the equation error approach developed initially which was achieving a feedback from the last output measurement only (therefore equivalent to a one-step ahead prediction approach).

When equation (3.11) is substituted in (3.2), the GPC formulation is expressed as follows:

For $t \in [1, N]$; find at time t:

$$\min \sum_{i=N_{\min}}^{N_{\max}} W_i\left[y(t) - y_{SP\;|t} + g\left[x(t+i)_{|t}\right] - g\left[x(t)\right]\right]^2 + \sum_{i=N_{\min}}^{M} R_i\left[\Delta u\left(t+i-1\right)_{|t}\right]^2$$

subject to:

$$\begin{aligned} y_{\min} &\leq y(t) + g\left[x(t+i)_{|t}\right] - g\left[x(t)\right] \leq y_{\max} && \textit{for } i = N_{\min},\ldots,N_{\max} \\ u_{\min} &\leq u(t+i-1)_{|t} \leq u_{\max} && \textit{for } i = N_{\min},\ldots,N_{\max} \\ \Delta u_{\min} &\leq \Delta u(t+i-1)_{|t} \leq \Delta u_{\max} && \textit{for } i = N_{\min},\ldots,N_{\max} \end{aligned} \tag{3.12}$$

The above optimisation problem is non-linear quadratic programming problem. In the case of the operational optimisation of a water distribution network, some of the variables have a discrete nature. Therefore, the problem is solved by mixed-integer non-linear programming, at each time t.

3.3 Process optimisation by means of Mixed-Integer Non-Linear Programming

THE previous paragraph proposed GPC as an adequate strategy for optimisation and control of water distribution systems. The non-linear nature of this algorithm is well acknowledged, and therefore, justifies the use of NLP techniques to solve the problem. The subject of non-linear programming is much broader and much more difficult to unify than the one of Linear Programming (LP).

At present, there is still no real unifying theory for what can be referred to as NLP techniques, as this particular field is still in a state of rapid change. It is commonly admitted that non-linear programming deals with the problem of optimising a non-linear objective function in the presence of non-linear equality and inequality constraints. If all the functions are linear, a LP program needs to be solved. Otherwise, the problem is called a non-linear program and is generally a lot more complex to solve. Many realistic problems cannot be adequately represented as LP because of the non-linear nature of the objective function or of any of the constraints. Efforts to solve non-linear problems more efficiently have made constant and rapid progress since the appearance of programming techniques at the end of the 1950's. MINLP is one of the latest approaches which was developed to improve the efficiency of non-linear optimisation applied to systems of a combinatorial nature, involving both discrete and continuous choices.

This paragraph is divided into three different sub-sections. Firstly, an overview of process optimisation is given. Then, a comprehensive literature review, concentrating particularly on integer programming, is presented. Finally, the MINLP algorithm is presented with full details.

3.3.1 Introduction to process optimisation

OPTIMISATION is the science of selecting the best of many possible decisions in a complex real-life environment. An optimisation problem can be described as any problem which seeks to maximise or minimise a numerical function of one or more variables (referred to as penalty or objective function), with the variables being independent or related in some way through the specification of certain constraints. A solution to an optimisation problem specifies the values of the decision variables and, therefore those of the objective function. It is referred to as feasible when it satisfies all of the constraints. It will be qualified as optimal if it is feasible and if it provides the best value for the objective function. It can be noted that several optimal solutions can exist for the same problem. A solution can also be near-optimal which means that it is feasible and also provides a superior objective function value, but not necessarily the best.

Optimisation theory has had a long and distinguished history as a major area of applied mathematics. The earliest work on this particular subject began at the end of the 17th century with the development of a theory called the calculus of variations. It includes results from many of the greatest mathematicians of the last 300 years. The possibility of using the classical methods of differential calculus and the calculus of variations to solve certain types of optimisation problems, arising in geometry and physics, has followed these earliest developments and has been applied since the middle of the 18th century. However, since the 1950's, there has been a remarkable growth of interest in a new class of optimisation problem which are usually not solvable by the classical methods of the calculus. These problems are referred to as programming problems.

The rapid increase in the size and complexity of problems resulting from the technological growth since World War II has led to the development of a systematic approach to programming problems. This approach generally involves three steps. Firstly, a mathematical model representing the decision problem under consideration must be developed. This model is built by identifying the variables of the problem, by collecting relevant data, by formulating the objective function to be optimised and, finally, by arranging the variables and data into a set of mathematical relations, such as equations or inequalities. Then, the process of decision-making continues by analysing the mathematical model and by selecting an appropriate numerical technique for finding the optimal solution of the problem. If one or more of the constraints or the objective function of the model are non-linear, the optimisation problem is a non-linear program. Finally, the last stage of the approach consists of finding an optimal solution, usually using a computer. The optimisation is considered complete when the numerical solution has been reinterpreted in terms of the practical choices to be made.

Mathematically, the general programming problem is formulated as follows:

Determine values for n variables $x_1, \ldots, x_n$ which satisfy the m inequalities or equations:

$$g_j(x_1, \ldots, x_n) \; [\leq, =, \geq] \; b_j \quad \textit{for } j = 1, \ldots, m \tag{3.13}$$

Maximising or minimising the function:

$$Z = f(x_1, \ldots, x_n) \tag{3.14}$$

The restrictions in the set of equations (3.13) are called the constraints, while equation (3.14) corresponds to the objective function. In (3.13), the $g_j(x_1, \ldots, x_n)$ are assumed to be specified functions, and the b_j are known constants.

Further in this set of equations, only one of the signs $[\leq, =, \geq]$ is used for each constraint, but the sign can vary from one constraint to another. Finally, the values of m and n are not related in any way, and m can be less than, equal to or greater than n.

Programming problems are frequently linked to economic criteria, rather than geometrical or physical ones, as it is more the case for classical optimisation problems. Often a programming problem consists in allocating a scarce resource for the manufacture of one or several products meeting certain specifications; while an objective function, such as profit or cost, is maximised or minimised. Programming problems have received huge interest because they do not occur only in theoretical economics, but also concern important practical problems in industry, commerce, or even in the military domain. Process optimisation has become one of the main applications of these techniques.

Process engineers and operations researchers are indeed traditionally involved in problem solving. The problem may involve reaching an optimal design, allocating resources in a better way, or finding the most suitable choices for a particular parameter. In the past, a wide range of solutions was considered acceptable and, in engineering design, it was common to include a large safety factor. However, in the competitive market, nowadays it is no longer adequate to find only an adequate solution to any given problem: the solution must be as close as possible to optimality. Process optimisation is one of the new tools available for engineers in their quest for optimality. It is widely applied in the chemical industry, where processes are well known for their non-linearity.

3.3.2 Literature review

ORIGINALLY, optimisation techniques applied to chemical processes mainly concerned the phase of synthesis, or design, of utility plants. Nishida *et al.* (1981) define process synthesis as the act of determining the optimal interconnection of processing units, as well as the optimal type and design of the units, within a processing system. Grossman (1989) gives a similar definition from a conceptual point of view. He describes the process synthesis problem as a problem for which input specifications as well as output ones are given: the problem then consists in finding a flowsheet capable of converting the inputs into the outputs, meeting several specifications at the same time. In other words, the operation of synthesis is equivalent to conducting a directed search for all the alternatives of configurations, and for each of these configurations, performing a search on all of the process variables.

The process synthesis problem is complex. For Hendry *et al.* (1973), the complexity associated with process synthesis come from the fact that the phase of synthesis must determine simultaneously the nature of the process components, including the way they interact; as well as the operating conditions for these components. It is obvious that this particular problem has a strong combinatorial nature, since all of the interconnections between pieces of equipment are represented with discrete variables. This naturally gives rise to an MINLP problem. The lack of sufficient theoretical background for this particular type of problem prevented a widespread usage of the method before the beginning of the nineties. However, it did not prevent researchers from showing interest in the process synthesis problem and developing alternate methods, making best use of the available level of technology.

As stated by Bruno *et al.* (1998), even if numerous methods of process synthesis have been reported in the literature since the earliest developments and applications in 1968, they generally follow three basic approaches: those using decomposition methods, those based on thermodynamics targets and heuristic rules, and more recently, those based on mathematical programming. The initial decomposition and thermodynamics approaches mainly focused on the synthesis of small sub-systems, like sequencing of distillation columns or heat exchanger networks. Since then, new developments have enabled their use on larger systems. In contrast to process synthesis, methods based on mathematical programming have been immediately applied to processes that are more complex. Mathematical programming methods constitute nowadays the main stream of research in the field of process synthesis.

3.3.2.1 Process optimisation based on decomposition techniques

AT the end of the sixties, Rudd (1968) proposed the first systematic approach to process synthesis. His initial research originated from the observation that too many equipment arrangements are possible for most processing tasks. Because the conventional optimisation techniques capable of handling discrete variables were originally only applicable to small-sized problems, it was obvious that it was not possible to examine each of these arrangements. Consequently, he proposed a synthesis based on decomposition. In the approach he developed; a problem, for which the level of complexity is too large for the existing optimisation techniques, is broken down into a sequence of sub-problems. However, this method never made its way into practice because of the simplified models it was using.

Masso and Rudd (1969) improved this initial approach soon after by making use of heuristic problem-solving techniques.

Heuristics are rules to search to find optimal or near-optimal solutions. Heuristics can be constructive which means that the solution is built piece by piece; or improved which means that a non-optimal solution to the problem is used as a good starting point, and is modified until a better solution is found. Generally, an heuristic method seeks the solution to a problem by means of plausible but fallible guesses. It is based on rules of thumb resulting from long experience. This approach immediately received large interest, as empirical decisions were common rules in chemical engineering and directly available for many processes. Heuristic methods have mainly been applied on separation processes, as they constitute an important part of the total annual operational cost of many processes. The optimal synthesis of separation processes represents therefore a significant factor in the decrease of the product costs. It has consequently received great attention.

Masso and Rudd's original method consisted in representing the set of separations by a binary tree of possible solutions, composed of operators (separators) and operands (intermediate or final products). Processing steps of interest are obtained from acknowledged design criteria and introduced in the system structure. The computer used is supplied with these processing steps and the optimised synthesis process is obtained by having the computer select a relevant sequence of steps. This is done by determining a selection weight for each of the design criteria entered in the system. These weights are adjusted to take into account the experience gained from previous synthesis trials. The design criteria and the manipulation of the selection weights constitute the structuring rules of the heuristics.

This method has been successfully implemented on numerous processes by other researchers; however, the additional work they conducted can be divided into two sub-classes of problems. The first sub-class is based on rules of thumb which help in the determination of optimal design characteristics and operating conditions for separation operations. Problems belonging to this category of methods can be found in Freshwater and Ziogou (1976), Morari and Faith (1980), Minderman (1981), Minderman and Tedder (1982) or Stephanopoulos *et al.* (1982). The second sub-class of decomposition techniques based on heuristic methods is aimed at defining the optimal process structure, and is more related to the aspect of optimising process flowsheets in the field of component separations. This second family of heuristic optimisation problems has been extensively tackled by research teams. The first known application of combined heuristic design-decomposition approach to flowsheet optimisation is due to Siirola *et al.* (1971). However, even if heuristic methods have allowed a reduction in the tree of possible solutions, an optimal solution cannot be guaranteed. Moreover, as described by Nishida *et al.* (1981), it has been quickly realised that some conflicts can occur in the combined use of some of the rules of thumb.

This did not prevent some additional research from being conducted, with success, in this particular area of optimisation for the next 10 years or so. For additional references, the reader can refer to Freshwater and Henry (1975), Seader and Westerberg (1977), Doukas and Luyben (1978), or Nath and Motard (1981).

Following the Siirola *et al.* (1971) work, a new trend has arisen: this has tended to improve heuristic methods, using algorithmic procedures. At the end of the eighties, new developments concerning decomposition techniques marked the first occurrence of mathematical programming in the field of chemical engineering, well before the advent of the powerful programming commercial packages,. It has therefore to be highlighted. Several researchers simultaneously published their work on this particular aspect. They all applied existing optimisation techniques and mathematical programming to the flowsheeting of chemical processes. As far as possible, they only tried to use heuristic rules of very wide applicability, so that the risks of conflict are reduced when several rules of thumb are combined. The first attempt to use algorithmic methods for the combined heuristic design-decomposition approach can be credited to Rudd, Siirola and their co-workers (Siirola and Rudd, 1971; Powers, 1972; Rudd, Powers and Siirola, 1973). They developed an heuristic approach for the synthesis of multi-component separation sequences, as a part of a general strategy that generated process flowsheets in an adaptive manner. The program they created, known as the AIDES system (Adaptive Initial DEsign Synthesizer) is based on linear programming and makes use of heuristics at two distinct levels. Firstly, it selects the best separation method. Then, it decides on the best sequencing to have the separation objectives reached. The drawback of this approach concerns the impossibility of incorporating equipment costs within the optimisation.

Soon after, at an AIChE meeting in Dallas in 1972, Ichikawa and Fan (1972) presented an alternative approach where all of the flowsheets existing for a given process are combined in a unique flowsheet, by defining all of the interconnections possibly existing between the different items or systems. All of the interconnections are assigned with split fractions of value between zero and one. The solution to the synthesis problem is obtained by determining the optimal split fractions for each interconnection using existing optimisation techniques, like linear programming, dynamic programming or non-linear programming. Optimal design parameters for the different process items are also evaluated. This approach has been successfully implemented by Umeda *et al.* (1972) for the design of a process composed of chemical reactors, distillation columns and heat exchangers. However, because of the poor quality of the optimisation techniques chosen to solve the problem, this approach can easily lead to non-optimal solutions. As stated by Hendry *et al.* (1973), applying continuous solutions to integer programming often leads to non-optimality.

Ichikawa and Fan's initial approach (1972) falls into this frame as discrete variables are represented by continuous ones. The method is also heavily reliant on the starting values chosen for the split fractions and design variables, with a strong probability of local optima resulting from inadequate choices during the initialisation phase.

Also in 1972, Hendry and Hugues (1972) proposed a different approach to optimal flowsheeting, based again on the combined heuristic design-decomposition approach. Their method is based on the use of dynamic programming to scan the search space, represented by a ranked list of decisions. This method was extended to the synthesis problem with heat integration by Rathore *et al.* (1974a, 1974b). They used list-processing techniques to determine all of the feasible heat matches between the hot and cold streams. The main drawback of their approach is that it requires the exploration of all the possible sequences to reach a result which requires huge computer resources and makes it impossible to apply the method on an entire chemical process.

Mahalec and Motard (1977a, 1977b) have managed to remove the important limitation of the AIDES system, linked to the need to provide good starting values to have satisfactory results. They especially developed the BALTAZAR procedure for the synthesis of a complete process. This computer program has been successfully used to produce alternate process flowsheets for the design of monochlorodecane by the direct chlorination of decane. The main difference between AIDES and BALTAZAR consists in the fact that for the latter, there is no need to have any preliminary specifications for process species allocation. It removes the problem of tight constraint on the possibilities for structure generation which occurs with the AIDES procedure. The BALTAZAR software still cannot integrate any equipment costs but uses heuristics as indicators of economic performance.

From the mid-eighties, a new concept, still based on heuristics, gave some renewed success to the ageing decomposition techniques initially developed. This concept, known as hierarchical decomposition, is still used nowadays, due to its powerful computer implementation, referred to as Process Invention Procedure (PIP). The hierarchical decomposition appeared in 1985, following the research undertaken by Douglas (1985). Since then, it has been successfully improved several times by the same researcher (Douglas; 1988, 1990). The initial approach created a framework to categorize any particular design problem into six different categories which are referred to as discrete decision levels. For each decision level, the economic potential of the project is evaluated and a decision is taken concerning the further viability of the project. The criteria which determine whether it is worth going on to the next decision level are based on heuristics, short-cut design procedures and physical properties.

This approach is motivated by the claim made by Douglas that less than 1% of the ideas for new design ever become commercialised. By using decision levels, it permits an early termination of poor design procedures. The economic evaluations at each level are conducted in terms of the dominant variables occurring for the process considered. Originally applied to a limited class of petrochemical processes, the method has been extended to a wider class of processes (like multiple reaction steps or plant complexes) by redefining some of the decision levels.

Following Douglas' initial developments, Kirkwood *et al.*(1988) originated the PIP computer implementation at the end of the eighties. PIP is based on the same concept of hierarchical decomposition, using decision levels. It attempts to invent a flowsheet using a strategy where one of the goals is to see if there is some reason why none of the available alternatives will ever be profitable. Therefore, it tries to create a design before any alternatives are considered. If profit is observed out of the design, PIP proceeds to the next level and adds more details to the flowsheet. Kirkwood and his co-workers present nine applications of successful flowsheet design, using the PIP implementation. However, relying on heuristics, this approach cannot rigorously produce an optimal design. Moreover, because of the sequential nature of the synthesis, interactions between the design variables cannot properly be taken into account. For instance, Duran and Grossman (1986a) and Lang *et al.* (1988) show that simultaneous optimisation of flowsheets generally produces significant improvements in the profit, compared to any of the sequential approaches.

After the work of these two research teams, decomposition methods (especially the combined heuristic design-decomposition approach) lost most of their appeal because of their intrinsic limitations. With the impressive developments in the field of computation enabling simultaneous optimisation, it is not justified anymore to break a process into several sub-systems. Furthermore, these methods have a very limited field of application which do not make possible their use for systems more complex than basic separation processes. The impossibility to quantify the quality of the solution is also a major problem.

3.3.2.2 Process optimisation by mean of thermodynamics targets

THERMODYNAMIC studies are aimed at finding the maximum possible work output or minimum work input when a system is taken by some process from a given first specified state to a second specified state, in the presence of a specified environment. Process optimisation by means of thermodynamics methods appeared in the 1970's, at a time when decomposition techniques based on heuristics and mathematical programming showed their first limitations.

The main concept of this approach is based on the rule of thumb which states that designs featuring strong energy efficiency are generally very close to optimality, at least from an economic point of view. Consequently, performing a thermodynamic analysis often leads to the determination of the lower bound for energy consumption which therefore becomes the target to be reached for the design engineer (Linnhoff *et al.*, 1982). This concept has been mainly applied to heat-exchange systems; partly to respond to the deadlock in which researchers were, since they were trying to apply mathematical programming to these systems without major success. The pinch analysis, still used nowadays, originated with this particular concept.

Optimisation of heat-exchange systems with fixed arrangements has been studied since the mid-sixties. However, with the doubling of energy costs at the beginning of the 1970's, it became more important to increase heat recovery ratios and therefore provide systems that are more economical. Various methods have been proposed to achieve these goals, but synthesis based on the available energy concept seems to have encountered a big success. This approach had been initially proposed by Umeda *et al.* (1978). For this concept, the problem consists in finding an optimal structure which minimises the fixed costs of heat-exchange and heat supply, as well as the operating costs of the system. This objective is achieved by performing a thermodynamic analysis of the process. Temperatures, heat duties and the physical properties of process streams are fixed before the optimisation procedure begins. Practically, the overall problem can be solved by constructing an hierarchical structure, comprising three sub-problems. The first sub-problem is used to determine the optimal value of the total amount of heat-exchange taking place within the process. This parameter is estimated by minimising the sum of the annual costs for the given heat-exchange network, constructed from the result of the second sub-problem. The second sub-problem consists therefore in constructing this network, for which the total heat-exchange area is minimised. This optimal heat-exchange network is determined by minimising the total investment cost resulting from the amount of heat-exchange determined by the first sub-problem. Therefore, it is obvious that the two first sub-problems are closely inter-related. Finally, the last sub-problem consists in modifying the optimal heat-exchange network, so that it takes into account the considerations of real-life applications, like operability, flexibility and safety.

Nishio *et al.* (1980) followed the path opened by Umeda *et al.* (1978), by applying the above method to the design of a steam-power system in the petroleum refinery. Steam-power systems are commonly found in utility plants, where they constitute major sub-systems. The method Nishio and his co-workers developed is aimed at selecting the optimal header pressure levels of the equipment constituting these sub-systems.

This work originates from Nishio's initial research which showed the limitations of mathematical programming for this particular type of problem, by proving too demanding for the computational resources available at the time he conducted his first research (Nishio, 1977). Nishio and his team justify their new study by the need to produce technical information on energy conservation that is more detailed, in the design phase of any process. That is where the thermodynamic approach plays a part.

Chou and Shih (1987) extended the method of Nishio *et al*. (1980) to the synthesis of a complete plant utility system. Besides the steam cycle problem, they also took into consideration the problem of regenerative gas turbine cycles, increasing the window of application of thermodynamic-based methods as well as the level of complexity of the optimisation. Their approach can be very powerful if energy is the dominant cost item in the process whose design needs to be optimised. However, with this kind of approach, the capital costs of the optimised design may be too high, even if a high thermal efficiency is reached. This constitutes the main drawback of thermodynamic-based methods. Furthermore, thermodynamic targets provide only guidelines which mean that considerable trial and error could be required to find solutions that are closed from the predicted targets. The impressive developments which occurred at the end of the 1980's in the field of computation, made computational resources become more and more easily available to all. As for the decomposition methods mentioned earlier, these improvements which have not stopped since, did not justify anymore the efforts of simplification in which thermodynamic-based methods originated. Even if physical properties continue to play an essential role for the understanding of process synthesis, the limited field of application of these approaches has also contributed to their abandonment in favour of more general and powerful automated synthesis procedures. Marechal and Kalitventzeff (1997, 1999), and Mavromatis and Kokotis (1998) have nevertheless proposed recently new interesting applications for thermodynamic based techniques. However, it is still true to say that nowadays the current state of flowsheet synthesis is only represented efficiently by mathematical programming.

3.3.2.3 Process optimisation based on mathematical programming

AS stated by Grossman (1985), when mathematical programming is used for process synthesis, the main idea of the approach is to formulate the synthesis of the flowsheet as an optimisation problem. Generally speaking, this requires a representation of a specified set of flowsheets among which the optimal solution will be selected. This set is referred to as a superstructure.

The definition of a superstructure is the most important aspect of this family of problems, as the optimal solution that is obtained can only be as good as the process representation which is being used.

As described earlier, the process synthesis problem is intrinsically combinatorial and non-linear which means that the optimisation problem, when solved by mathematical programming, requires the use of a MINLP approach. Until very recently, no tools were available to handle such a complex formulation and that is what justifies the development of alternative concepts, maybe less accurate than a MINLP formulation, but capable of achieving optimised results for a few specific process configurations. From the point of view of numerical analysis, integer programming cannot be solved with the conventional strategy used usually for continuous numerical optimisation. This strategy which consists in taking a trial solution and seeing if a small change improves it, cannot be used. Continuous methods usually only produce a local optimum, whilst integer programming problems typically have many local optima, that are often much worse than the global optimum. The last statement justifies the need to develop specific methods to tackle the integer-programming problem.

With strong developments occurring in the field of integer programming since the beginning of the 1980's, a consensus emerged in favour of a more widespread use of algorithmic methods for the particular class of process synthesis problems. This new concept completely changed the perception of process synthesis, since the spirit of algorithmic methods allowed the handling of numerous process configurations using the same approach. Resolving the main limiting factor of decomposition and thermodynamic-based methods, both having a limited field of application, mathematical programming has been willingly adopted worldwide. However, it is not until better knowledge was acquired on integer programming and improved computational resources made available that algorithmic approaches found their way efficiently into the industrial world.

3.3.2.3.a Pioneering research in mathematical programming

IT is generally acknowledged that the origins of mathematical programming have been quite laborious. The general approach to process synthesis described above is not recent. It appeared first with Umeda and his co-workers' research (Umeda et al., 1972). Their paper referenced earlier in this chapter proposed one of the first algorithmic methods for process synthesis. The direct search technique they used is referred to as the Box's Simplex method, and has been used to synthesize a system consisting of two reactors, two distillation columns and several heat exchangers.

However, although their approach has been implemented on a full-scale process, the limited computational resources at their disposal forced them to make use of heuristics in the model they developed, therefore losing accuracy in the solution of the problem. Furthermore, they never managed to have their model running on a process composed of more than a few basic elements.

Mishra *et al.* (1973a, b) used the simplex pattern search technique for the synthesis of a biological wastewater treatment system, while Mehrotra and Kpur (1974) preferred to apply an integrated approach for the determination of the optimal configuration of a flotation circuit, for the separation of mineral species. Soon after, Himmelblau (1975) tested several available NLP codes to test their applicability for the solution of the non-linear optimal synthesis problem. He tried to implement these codes on Mishra and co-workers' case study and reported that three of these codes behaved poorly. He only managed to have a successful run with one of the remaining codes.

This explains why, following this pioneering research, Nishio and Johnson (1979), and Petroulas and Reklaitis (1984) rather preferred to develop models based on LP approaches. The latter research has been aimed at developing a synthesis procedure for the preliminary design of utility systems. Petroulas and Reklaitis' algorithm determines the optimal header pressure levels, the distribution of steam turbines in the network and the steam flows between all devices. The optimisation objective is about maximising the amount of real work recovered from all sources. Two major assumptions have been made to facilitate linear programming. The steam header states have been indeed represented by continuous variables, while the optimal allocation of drivers was represented by linear variables. Besides these assumptions, steam sources, heating and drive horsepower needs must be known given parameters at the beginning of the optimisation.

3.3.2.3.b The Mixed-Integer Linear Programming approach

IN the meantime, Papoulias and Grossman (1983) originated a new concept referred to as Mixed-Integer Linear Programming (MILP), in which the combinatorial aspect of the process synthesis problem is taken into account, whereas its non-linear nature is not. The approach is directly derived from the original MINLP formulation by fixing some of the operating conditions (such as pressure and / or temperature). They successfully implemented the method to select among all the alternative units included in a proposed utility plant superstructure.

Their optimisation criterion referred to the minimisation of the linearised cost function (including capital costs, fixed charges and operating costs) which was solved using a standard Branch and Bound (BB) computer code which was originally the standard procedure for solving MILP formulations. This algorithmic method originated in Little and his co-workers' work (1963) which developed the framework of the BB method, as it is known nowadays. This pioneering research lead to the development of an algorithm referred to as the Last In / First Out algorithm. Garfinkel and Nemhauser (1972) and Beale (1977) successively reshaped this initial research and finally gave birth to the BB method. The algorithm has been enumerated soon after as a computer code by Balas (1979). The method consists in enumerating a tree whose nodes represent different combinations of partial assignment of binary variables, represented by fixed integer values. Each node in the tree involves the solution of a linear programming sub-problem that has some of the binary variables with fixed integer values, while the remaining binaries are treated as continuous variables. These linear programming sub-problems can be updated very efficiently and their solutions provide bounds to the solution of the original problem. By making use of these bounds, an intelligent search can be performed and as a result, only a small subset of nodes in the tree needs to be enumerated to find the global optimum solution.

Due to efficient commercial computer codes, this algorithm has showed promising results which prompted Gupta and Ravindran (1985) to study the feasibility of applying this formulation to a convex non-linear integer-programming problem. Their attempt proved successful, but at the expense of considerable computational resources which originally limited the generalisation of the method to small-scale applications (generally including less than 200 variables in the optimisation).

During the same period, Grossman (1985) developed separately a general-purpose method based on MILP to solve for large-scale synthesis problems. His work was published at a time where LP was still popular amongst researchers. He justifies the use of a MILP formulation for synthesis problems rather than a formulation based on continuous variables by the fact that the latter formulation often gives rise to a non-convex optimisation problem which can have multiple local solutions. As he shows, this phenomenon can easily annihilate all the advantages of removing the combinatorial aspect of the problem, as there is a need to analyse every single configuration to check that it is not a local optimum which is time-consuming. With the way a MILP formulation is posed, it avoids this exhaustive enumeration of all possible configuration alternatives. Although the problem to be solved is convex, MILP is still performing better than a formulation based on continuous variables, as it is capable of determining directly the global optimum solution.

To achieve the same results with continuous programming, global optimisation techniques have to be used, but they have never been proven rigorous enough for real-life applications.

Grossman has in fact adapted the Generalized Benders Decomposition (GBD) procedure for the particular case of MILP problems. The class of procedures using the GBD principle was developed after the initial work of Benders (1962). Benders' algorithm is aimed at solving problems involving non-linear discrete variable and linear continuous variables. Geoffrion (1972) extended the Benders method to more general problems with non-linear integer and continuous variables which gave birth to the GBD procedure. The principle of the method is to partition mixed-integer programming problems. The idea is that the whole problem is represented as a set of independent non-linear programming sub-problems, except for some linking variables which may be integer variables. A "master problem" is built: it consists of constraints involving only the linking variables and the value of the objective function. These constraints are derived by solving the whole problem for different sets of fixed values for the linking variables and examining the resulting value of the objective function. The problem is solved when the solution to the "master problem" also solves the whole problem.

Originally, this method was rather difficult to implement, as it needs the solution of both a NLP sub-problem and a pseudo-integer programming master problem. For instance, in Grossman's approach, the GBD procedure is simplified and only applied to a problem where the integer and continuous variables occur linearly in the objective function and constraints. Since the original development of MILP, the method has appeared to be particularly well suited to the synthesis of utility systems. It has been used with success for the design optimisation of utility plants by Hui and Natori (1996) and Iyer and Grossman (1997). Chang and Hwang (1996) also solved the problem of waste minimisation in utility plants with a similar approach.

However, any MILP formulation is limited by the fact that it cannot really handle explicitly non-linearities which prevents using the method on a large class of problems. Although MILP proved successful for the design of utility systems, it is easy to understand why research continued in the area of mathematical programming.

3.3.2.3.c Development of new algorithms to solve comprehensive MINLP formulations

AFTER the previously mentioned attempt by Gupta and Ravindran (1985) to use a BB code for a convex non-linear problem, the late 1980's saw the rise of new developments, finally leading to comprehensive approaches, where the process synthesis problem is explicitly formulated as an MINLP problem.

In the context of process synthesis, a MINLP optimisation model usually involves continuous variables representing process parameters (such as pressures, temperatures or flows), as well as binary variables to identify the potential presence of units (0, not included; 1, included). The objective function is normally a cost function. Mass and energy balances are written as equality constraints whereas the design specifications and logical constraints are represented by inequalities. The successful application of such formulations to a real-life environment has led to the development of a new computer code by Duran and Grossman (1986a). They conceived an algorithm, referred to as Outer Approximation (OA), to handle explicitly the non-linearities in process models.

The OA concept is similar in nature to the GBD method described earlier. The differences between the two methods occur mainly in the definition of the MILP "master problem". Generally speaking, the OA algorithm partitions the problem into two separate parts. Firstly, it solves for an NLP sub-problem, where the continuous variables are initially optimised. Then, non-linear equations are linearised to obtain a MILP master problem which is used to determine the optimal configuration for the next NLP sub-problem. This algorithm has been successively improved (Kocis and Grossman, 1987; Viswanathan and Grossman, 1990) and is nowadays available in the program DICOPT++ (DIscrete Continuous OPTimizer), under the denomination AP / OA / ER (standing for Augmented Penalty / Outer Approximation / Equality Relaxation). The latest version of DICOPT++ is capable of handling both convex and non-convex non-linear problems and, to solve for large-scale MINLP problems, it has been coupled with the powerful modelling language General Algebraic Modelling System (GAMS), edited by the GAMS Corporation.

Duran and Grossman (1986a, b, c) applied the initial OA algorithm for simultaneous optimisation problems, as well as heat integration in process flowsheets, while Kocis and Grossman (1987) were the first researchers who succeeded in successfully implementing the OA method on the design of a process flowsheet. They nevertheless had to develop a new version of the OA algorithm, referred to as the OA / ER approach since (Outer Approximation / Equality Relaxation). The latter team of researchers continued their research efforts on the OA approach and showed that the quality of the modelling has a great impact on the reliability of the solutions which are obtained (Kocis and Grossman, 1989). Regarding the last observation, they presented a modelling / decomposition (M / D) scheme to improve the performance of the OA / ER algorithm. In their approach, zero flows are avoided and non-linearities reduced as far as possible to minimise the computational burden and increase the chance of reaching a global optimum.

The M / D approach showed good results on several case studies; it has been used for implementation by Kravanja and Grossman (1990) in the computer package PROSYN (PROcess SYNthesizer). This software has been developed to handle simultaneously the synthesis and the heat integration of process flowsheets. It consists in an automated procedure which achieves the decomposition of the general superstructures used within the MINLP formulation to reduce computational efforts. Initially, PROSYN was only applicable to convex non-linear problems. However, following the application of an OA approach to non-convex formulations, using the new AP / OA / ER algorithm developed by Viswanathan and Grossman (1990), Kravanja and Grossman (1994) also extended their approach to this particular family of problems. They based their new developments on this new OA algorithm, after it had been finally implemented in DICOPT+ +, again by Viswanathan and Grossman (1990). They used the occasion of releasing a new version of their software to consider also the complaints from PROSYN users who criticized the need to provide a complex model representation with the earlier versions of the software. To remove this important drawback, they included a model generator and a comprehensive library of models for basic process units in the new PROSYN package.

Like Kocis and Grossman (1989), Yuan *et al.* (1989) also tried to develop an extension of the initial OA method, by developing the generalised OA algorithm for solving large-scale problems. In the proposed algorithm, the general MINLP problem is decomposed into a sequence of NLP sub-problems and the algorithm has been proved to converge in a finite number of iterations. In this solution procedure, finding the solution of the NLP sub-problems is the most difficult task. To implement this MINLP code, a Generalized Reduced Gradient (GRG) method was used for the NLP sub-problems, whereas, the MILP relaxed problems were solved by using the BB approach. The GRG algorithm has been first proposed by Abadie and Carpentier (1969) and implemented in the CONOPT commercial package by Drud (1985). As for DICOPT+ +, CONOPT has been coupled with GAMS, to solve for large-scale NLP problems.

The GRG approach extends the reduced gradient method, due to Wolfe (1963, 1967), to the case where not only the objective function, but also the constraints, are non-linear. This iterative method consists in partitioning the variables into dependent and independent variables. Then, it computes the gradient in Wolfe's sense, determines the direction of progression of the independent variables and, finally, modifies the dependent variables in order to verify the constraints. In CONOPT, a matrix approach is followed. All matrix operations are implemented by using sparse matrix techniques to allow very large models. The algorithm uses many dynamically set tolerances and therefore runs, in most cases, with default parameters.

The system is continuously being updated, mainly to improve reliability and efficiency on large models.

The efficiency of the extended OA algorithm has been proved by Yuan *et al.* (1989) on mathematical problems and on the gas transmission network synthesis, originally presented by Duran (1984). On this particular example, the authors note that using a MINLP formulation allows optimising simultaneously the investment cost and the rate of heat recovery. Recently, Fletcher and Leyffer (1994) improved the extended OA algorithm of Yuan *et al.* (1989), by developing the Quadratic Outer Approximation which allowed integer variables to occur non-linearly. Their termination criterion is also simplified compared to the original OA approach of Duran and Grossman (1986a).

Although it is obvious from above that the OA algorithm has constituted the main area of research for MINLP formulations of process synthesis, the other algorithms developed initially (like BB and GBD) have also been successfully applied to MINLP problems. The implementations of BB (referred to as LP / NLP-based BB and due to Quesada and Grossman, 1992) and GBD from MILP to MINLP problems have been relatively easy. This is expected since the initial BB and GBD approaches have been directed to the MINLP class of problem since their conception. This is only because of technological limitations on computers that these approaches never solved rigorously comprehensive MINLP formulations initially, and had to be originally restricted to the MILP class of problems. Recent developments concerning the improvement of BB codes suiting a comprehensive MINLP formulation can be found in Adjiman *et al.* (1997). This team of researchers introduced two new NLP-based BB algorithms which guarantee that the global optimum is reached, even with non-convex functions. They show the potential application of these global optimisation algorithms on a few case studies and realistic examples. Smith and Pantelides (1997) have also pursued the problem of handling non-convexities with BB algorithms and they proposed an alternative approach to Adjiman and co-workers' extended BB approach. They state that they have been able to locate the global optimum of a number of test problems from the field of engineering design, in reasonable computational time.

Although new algorithms emerged recently to solve the integer-programming problem, they have been derived, in one way or another, from one of the three families of algorithms described earlier (GBD, OA, and LP / NLP-based BB) and they have not encountered much success yet. These theoretical developments are still going on to improve the ability of MINLP formulations to handle large-scale systems more easily.

For more information on this, the reader is directed to Westerlund *et al.* (1994: Extended Cutting Plane), Westerlund and Petterson (1995: Binary Separable Programming), Turkay and Grossman (1998: Generalized Disjunctive Programming), Van den Heever and Grossman (1999: Hybrid Disjunctive Bi-level Decomposition Algorithm), Bhatia and Biegler: (1999: Reduced Hessian Successive Quadratic Programming). More conveniently, the reader can also consult Saunion (2001), since he reviewed all of these methods in a single study.

3.3.2.3.d Industrial applications of MINLP algorithms

GBD, OA, and LP / NLP-based BB (and the commercial packages associated with these methods) still constitute the three major families of algorithmic methods for integer-programming problems. These three algorithms have been widely validated at the scale of case studies by numerous research teams, and emerged as robust and reliable methods. They are currently being validated for applications on an industrial scale. It is therefore obvious that all of the approaches that can be referred to as industrial applications of MINLP for process optimisation are quite recent. It also means that only a few papers are currently available in the literature.

A good example of industrial application of MINLP technique to process optimisation can be found in Harsh *et al.* (1989). In their study, they presented a systematic strategy for dealing with a restricted class of retrofit problems, where the functional topology of the flowsheet remains the same but individual units can be replaced or modified in order to implement the retrofit. Their approach, based on the OA algorithm of Duran and Grossman (1986a, b, c), has been applied to the retrofit of an ammonia process. The result showed almost US$ 3 to 4 million increase in venture profit, with a capital investment as low as US$ 100,000. This study has been among the first to demonstrate the potential of MINLP approaches with conventional simulation models.

Dikewar *et al.* (1992) combined the OA and GBD codes to develop an automatic process synthesis environment with the ASPEN chemical process simulator. They have tested the applicability of their synthesizer to the structural optimisation of the hydrodealkylation of the toluene process. Their approach consists in selecting the best flowsheet out of several alternatives contained in a superstructure, so that the annualised profit is maximised. Operating conditions must also be evaluated accordingly.

Diaz and Bandoni's research (1996) concerned the structural and operational optimisation of a real ethylene plant in operation by outer approximation.

Compared to the best continuous solutions available at the time of their research, their optimisation showed an increase of approximately US$ 300,000 per year in gross profit. This was done by simultaneous parameter and structural optimisation, using an MINLP model. This result also showed that the OA method can work very well on real-life problems, as it managed to reach a convergence in a few iterations.

Bruno *et al.* (1998) also performed a structural and parameter optimisation of a utility plant that satisfies given electrical, mechanical and heating demands of industrial processes. In their model, non-linear equations are extensively used for the cost of equipment and for the plant performance in terms of enthalpies, entropies and efficiencies. They proved that the solution to the optimisation problem can be obtained efficiently with the computer code DICOPT+ +.

Similarly, Saunion (2001) optimised the steam distribution at an oil refinery, in which imbalances between the various steam distribution levels required the venting of steam from the lowest level. The problem originating from an excess of high-pressure steam production for too few medium-pressure steam users and turbines, he proposed to solve the problem by considering the replacement of selected steam turbines with electrical drives. To find optimal configuration changes, he introduced a MINLP formulation interfaced with GAMS which maximises the return on investment, meeting the electrical, mechanical and steam demands of the refinery. Using DICOPT+ + as the MINLP solver of the problem, this researcher proposed to replace two high-pressure turbines and, with this new configuration, projected a profit of R38.3 million per year in the best case, with an investment of R9.2 million (South African Rands, year 2000). Saunion acknowledges that modest computational effort was required.

These pioneering applications have opened the way to the development of mathematical programming methods for optimisation of industrial installation flowsheets, and the integration of these methods as standards. Since the first applications to process synthesis, algorithmic methods have also awoken interest in a new category of problems, referred to as the sequencing and scheduling problem. The optimisation of the Durban water distribution system can be categorised as part of this new family of problems.

3.3.2.3.e Recent developments using algorithmic methods: sequencing and scheduling problem

THE recent developments in the field of mathematical programming have opened the application of these techniques to solve a very important problem in chemical engineering, referred to as the sequencing and scheduling problem. This problem is equivalent to determining a correct production sequence to respond to one or several objectives.

With the recent tendency which consists in expanding the complexity of plant structures, the search by a human operator for a feasible production sequences responding to an increasing number of objectives becomes more and more difficult. If the sequence has to be optimal, the search becomes almost impossible, and computer resources have to be employed instead.
The sequencing and scheduling problem is also part of the general family of optimisation problems referred to as combinatorial, and is ideally solved by MINLP. A typical feature of most scheduling problems is that the determination of a production sequence requires a large calculation effort, as the number of solutions grows exponentially with the size of the process. This explains why it stimulated the interest of researchers only after major breakthroughs had been accomplished in the field of algorithmic optimisation, as well as computer resources.

The sequencing and scheduling problem really became a matter of interest after the early developments in the field of combinatorial optimisation proved successful. Being similar in nature to the optimal flowsheet synthesis problem described earlier, the initial algorithms have reemployed the experience acquired earlier on this particular class of problems. The first references to the sequencing and scheduling problem only appeared in the second half of the 1970's, the pioneering work in this field being due to Rivas and Rudd (1974a, b). These two researchers modelled a chemical engineering process as a network of valves and connectors. The connectors can be represented by pipes, vessels, reactors or any other equipment through which materials flow from input to output. The flow through certain connectors can be stopped by the action of valves. The determination of operational procedures is made according to three levels of goals: respectively, objective to be reached, tasks to be done to meet the objective and actions to be taken to achieve these tasks. Originally, the computer-aided system they developed was mainly used to prevent the use of valve sequences leading to dangerous operating conditions. Eventually, this method was extended by O'Shima (1978) to find the best path for the materials between two locations in the chemical plant. As the number of cases for the sequence of operation of the valves increases exponentially, O'Shima (1978) recommends defining all of the operating conditions that should be avoided, so that no search is conducted for valve sequences leading to these conditions. The computational burden is therefore reduced to its minimum.

Another approach to the problem of optimising operating procedures has been given by Kinoshita *et al.* (1981). He proposed to generate a sequence of state transitions to bring the plant from the initial state to the final goal state. The plant was optimised by decomposing it into sub-systems of reasonable size, because of the limited computational resources available at that time.

This approach is actually quite similar to the decomposition method described earlier in this literature survey for the flowsheet synthesis class of problems, even if it is based on algorithmic procedures.

Later on, Tomita *et al.* (1986) attempted to develop a computer system based on heuristics for automatic synthesis of plant operations. In this approach, the topology of the plant is represented by means of a directed graph, where all of the equipment, except valves, is expressed as nodes. Pipes are expressed as arcs to which the valves are attached. The flow path is defined in the graph as a continuous sequence of arcs between the inputs and the outputs of the process.

These early computer codes have confirmed the interest in using optimisation methods to solve for the sequencing and scheduling problem. In the decade that followed these initial developments, researchers successively improved these codes and mainly adapted them for application to batch systems. Batch operations usually consist in the use of general-purpose equipment for the manufacture of a variety of products, following different routes through the plant. They are highly non-optimal by nature which explains why they have been originally regarded as inefficient in production, and labour consuming, compared to the more efficient continuous processes. However, producing diverse products in small quantities with the same plant has become more and more common over recent years, particularly for the high-benefit components produced in the pharmaceuticals, cosmetics or biochemical industry. This kind of production strategy is particularly suited to batch processes, hence the need for optimising the operations of these plants.

Based on the pioneering developments described earlier, new computer codes for production control have since been discovered and have encouraged even further the development of batch operations. In the literature, the sequencing and scheduling problem for batch processes has been extensively tackled. From the short-term scheduling of Egli and Rippin (1986), to the short / medium-term strategy of Làzaro *et al.* (1989); from the simplified multi-units / multi-products process of Ku and Karimi (1988), later improved to take due-dates penalty into account (Ku and Karimi, 1990), to the more complex approach developed by Birewar and Grossman (1990) adding inventory costs, clean-up times and production shortfalls to the optimisation, the sequencing and scheduling problem of batch processes has taken highly diverse forms. Over the years, the formulations have seen their complexity increasing to make better use of improved computational resources. Initially using poor performance heuristics-based methods, they have since integrated the latest algorithmic developments, allowing recently the development of comprehensive MINLP formulations.

Reklaitis (1992), or Rippin (1993) have recently reviewed extensively this area of research, and the reader is invited to refer to these two references to find more details concerning this particular area of research.

Nowadays, successful applications of sequencing and scheduling techniques can be found in almost every single process based on batch operations. The experience accumulated from these implementations has opened the application of these techniques to a broader class of processes which are not exactly batch in nature, but present some of the characteristics of a batch production. For instance, determining an optimal schedule for the pumping operations of the Durban water distribution network falls into this new area of research.

Azzaro-Pantel *et al.* (1998) give a good example of application of sequencing and scheduling techniques to a non-batch process. It is one of the rare articles available in the literature on the subject. Starting from the observation that the wafer manufacturing process used in the micro-electronic industry has important similarities with batch chemical processing, these researchers developed a discrete-event simulation model to solve successfully for the industrial-size scheduling problems of wafer manufacturing.

Similarly, Joly and Pinto (2003) developed mixed-integer programming techniques for the scheduling of a real-world fuel oil and asphalt production scheduling problem at the PETROBAS REVAP Refinery which produces approximately 80 % of all fuel oil consumed in Brazil. The complexity of the problem resides in the large number of combinatorial alternatives due to the operational decisions that must be taken in order to satisfy all product requirements. Therefore, they first modelled the system as a non-convex MINLP problem. A rigorous MILP was then derived from the initial MINLP formulation. Although this linearization caused an increase in the model size, it allowed solving the problem to global optimality. Their results showed that the solution of the non-convex MINLP formulation can successfully be achieved with the augmented penalty version of the OA method, while the MILP model was solved with a commercial implementation of the LP-based BB method. The two researchers proved that these methods, while susceptible to infeasible times to achieve global optimal solutions, are ready for the exploration of market opportunities, mainly in the short term, providing therefore an efficient tool for the scheduler.

For the particular case of optimising the scheduling of pumps inside a water distribution system, once again, the main characteristics of batch chemical processing can be found.

Water distribution involves indeed a long sequence of processing steps before it is provided to the final consumers (dispatching in the network, re-treatment, storage in the reservoirs), as well as many items of equipment (pipes, reservoirs, valves pumps). Since these items are highly interconnected, the water can follow different paths throughout the network. In addition to these operations, cleaning, measurement or inspection operations are performed by operators. Finally, the network behaviour can be affected by sporadic equipment breakdown or maintenance operations. Consequently, mathematical programming, and more precisely MINLP, seems particularly well suited to find an optimal scheduling, responding to water quality, security and economic considerations.

3.3.2.3.f Shortcomings of mathematical programming methods

UNTIL now, a rather optimistic picture of algorithmic methods has been drawn. However, as for decomposition methods and approaches based on thermodynamics targets, several drawbacks can be enumerated.

The main criticism of process optimisation is that these methods tend to be computationally expensive. This was true originally, and it still is nowadays, as the complexity of the problems tackled grows at the same speed than progress is achieved in the field of computer resources. Another complaint concerns the fact that design engineers feel they are removed from the process of decision when they use these techniques. For them, mathematical programming provides too few explanations concerning the decisions which are taken. Unlike the two other approaches, where improvements are obtained by using common physical knowledge, some researchers feel that the somewhat more complex decisions that need to be taken to obtain good results are not necessarily accessible to design engineers who may not have a good mathematical background in optimisation methods. Finally, others are concerned by the difficulty in evaluating the quality of the solution.

There is no denying that these criticisms have some validity but, as Grossman (1985) states, it is also true that they come from a lack of appreciation of capabilities and limitations of optimisation techniques. With chemical processes becoming more and more complex, it is becoming increasingly difficult to get a design engineer to take all of the decisions himself which is the case if he has to use decomposition or thermodynamic techniques. Furthermore, algorithmic methods should not be disregarded too early, as neither decomposition methods, nor thermodynamic ones can offer alone an efficient way to tackle process optimisation problems.

Although it is true that it can be difficult to guarantee finding a global solution when using mathematical programming, experience shows that the best solution found with algorithmic methods usually improves the objective function by several thousand of US dollars per year which justifies the effort of an optimisation project. In addition, considering that algorithmic methods managed to provide a common framework for a large class of optimisation problems composed of different types of components, their shortcomings should be regarded as minimal. The important "standardization" needed in the field of process optimisation was desperately awaited and was lacking until recently. It is obvious that with major developments still going on in this field of research, improved algorithms should appear on the market in the near future. The number of adverse comments regarding an extensive use of mathematical programming for process optimisation may reduce accordingly.

3.3.2.4 Concluding remarks to the literature survey

FOR this literature survey, a chronological approach has been adopted so that the reader could get an organised representation of the successive developments which occurred in the field of process optimisation. This choice was clearly made to improve the comprehension by classifying the numerous approaches available on the market, within only two or three major categories. As a result, it can appear that all of the methods enumerated here have been developed separately and never co-existed or interacted. The reader should not be mistaken into believing that the development of a new method caused the disappearance of previous ones; in other words, believing that algorithmic methods took over the thermodynamic-based approach which themselves took over decomposition techniques. In the reality, things have been more "chaotic" and they can still appear to be. In the diversity of approaches existing on the market, it becomes more and more difficult to make a choice. When working in the field of process optimisation, one should take into account one's own needs. Nowadays, the differences between the families of approaches become less marked. Knowing that every single method has its own advantages and drawbacks, the new popular idea is to combine them and take the best of each. For instance, both thermodynamic and algorithmic methods make use of heuristics in one way or another, while some of the heuristics methods are algorithmic. Kovac and Glavic (1995) have for instance combined thermodynamic with MINLP, for optimal flowsheeting. Unpromising structures are first eliminated thanks to thermodynamic choices, while the resulting superstructure is optimised using MINLP. More recently, Hostrup *et al.* (2001) have also opted for a similar combination of methods and presented a comprehensive integrated approach, based on thermodynamic choices and structural optimisation for the solution of process synthesis, design and analysis.

Although process optimisation has received considerable attention over the past decades, it is still an area of interest for researchers. The complexity of this problem is demonstrated by the fact that it is currently the subject of extensive research efforts thirty years after the original developments. Besides MINLP, probabilistic optimisation algorithms like fuzzy logic, genetic algorithms and the simulated annealing approach are the domains in which most research is under way. It is hoped to improve existing process optimisation methods in the near future, using these innovative approaches.

3.3.3 The Mixed-Integer Non-Linear Programming approach

MIXED-INTEGER Non-Linear Programming deals with optimisation techniques in which an objective function is minimised (or maximised), while subjected to both equality and inequality constraints, and where two types of variables can be specified: continuous or integer variables. Continuous variables can take any real value within given bounds, while integer variables can either take 0 – 1 values or access a range of integer values. The unique feature of the method is precisely the capability of handling binary variables which makes it particularly suited to represent discrete decisions in combinatorial problems. MINLP is a class of these problems for which the functions involved in the problem are non-linear.

In this paragraph, the most basic form of an MINLP problem will be presented. It will be followed by a rapid overview of the different strategies which can be used to solve this family of optimisation problems and which have been mentioned earlier. The work presented here has been based mainly on the paper published by Grossman (1996) which gives an extensive overview of mixed-integer techniques for algorithmic process optimisation. The information presented about the different algorithmic families used to solve MINLP problems has been extracted from Geoffrion (1972), Gupta and Ravindran (1985), Duran and Grossman (1986a), Grossman (1989) and Quesada and Grossman (1992).

3.3.3.1 Basic representation of a MINLP problem

THE most basic form of an MINLP problem, when represented in algebraic form is as follows:

$$\begin{aligned} &\min\ Z = \underline{f}(\underline{x},\underline{y}) \\ &subject\ \ to\ \ \underline{g}_j(\underline{x},\underline{y}) \le 0 \quad j \in J \\ &\underline{x} \in X,\ \ \underline{y} \in Y \end{aligned} \tag{3.15}$$

In (3.15), $\underline{f}(\cdot)$ and $\underline{g}(\cdot)$ are convex, differentiable constraint functions, while $\underline{x}$ and $\underline{y}$ are the continuous and discrete variables respectively. The set X is commonly assumed to be compact, while Y is generally discrete. These sets can be defined as shown below:

$$X = \left\{ \underline{x} \mid \underline{x} \in \mathbb{R}^n,\, D\underline{x} \le d,\, \underline{x}^L \le \underline{x} \le \underline{x}^U \right\} \tag{3.16}$$

$$Y = \left\{ \underline{y} \mid \underline{y} \in \mathbb{Z}^m,\, A\underline{y} \le a \right\} \tag{3.17}$$

In (3.16), the continuous variables $\underline{x}$ are generally represented by process parameters, such as flow rates, pressure levels, temperature, or even equipment sizing characteristics. The inequalities $D\underline{x} \le d$ represent linear equalities or inequalities that involve these process parameters. As represented by $\underline{x}^L \le \underline{x} \le \underline{x}^U$, the set X is generally constrained by known lower and upper bounds on the continuous variables (respectively $\underline{x}^L$ and $\underline{x}^U$).

In (3.17), the discrete variables are generally restricted to $0 - 1$ values in most applications which means that the corresponding set can be transformed as follows:

$$Y = \left\{ \underline{y} \mid \underline{y} \in \{0,1\}^m,\, A\underline{y} \le a \right\} \tag{3.18}$$

In the above equation, $A\underline{y} \le a$ correspond to the linear constraints that must be satisfied by the $0 - 1$ variables. In the context of process optimisation, the binary variables generally represent the potential existence of units, or the status of the pieces of equipment. In (3.15), the constraints functions $\underline{g}_j(\underline{x},\underline{y})$ represents performance relationships for the process. In the field of process optimisation, these correspond to mass, flow or energy balances, physical constraints or operational specifications. Still considering the same equation, the objective function Z is typically a cost function, also involving process performance. In most applications of interest, the objective and constraint functions $\underline{f}(\cdot)$ and $\underline{g}(\cdot)$ are linear in $\underline{y}$ to account for fixed cost charges or logic constraints.

As described earlier, numerous algorithms have been developed to solve the problem presented in equation (3.15). However, with the application of MINLP techniques to large-scale systems, only three major approaches have proved successful: the Generalized Benders Decomposition, the family of Outer Approximation approaches and finally, the LP / NLP-based Branch and Bound. All of these algorithms have in common the fact that the optimisation problem is partitioned into a MILP master problem and several NLP sub-problems. The approaches differ by the way the MILP and NLP problems interact with each other.

3.3.3.2 Definition of the NLP sub-problems

FOR fixed integer variables $\underline{y}^k$, the definition of the NLP sub-problem is as follows:

$$\begin{aligned} &\min Z_U^k = \underline{f}(\underline{x}, \underline{y}^k) \\ &\textit{subject to } \underline{g}_j(\underline{x}, \underline{y}^k) \leq 0 \qquad j \in J \\ &\underline{x} \in X \end{aligned} \tag{3.19}$$

The formulation above produces an upper bound Z_U^k for the original MINLP problem represented by equation (3.15), provided the problem in (3.19) has a feasible solution which is not always the case.

3.3.3.3 Definition of the MILP master problem

THE new predicted values $\underline{y}^k$ (or even, $(\underline{y}^k, \underline{x}^k)$) are obtained from a relaxed MILP that is based on the K linearization points, $(\underline{y}^k, \underline{x}^k)$ *for* $k = 1, \ldots, K$ generated at the K previous steps. The relaxed problem is expressed as follows:

$$\begin{aligned} &\min Z_L^K = a \\ &\textit{subject to } \begin{cases} a \geq \underline{f}(\underline{x}^k, \underline{y}^k) + \nabla \underline{f}(\underline{x}^k, \underline{y}^k)^T \begin{bmatrix} \underline{x} - \underline{x}^k \\ \underline{y} - \underline{y}^k \end{bmatrix} \\ \underline{g}_j(\underline{x}^k, \underline{y}^k) + \nabla \underline{g}_j(\underline{x}^k, \underline{y}^k)^T \begin{bmatrix} \underline{x} - \underline{x}^k \\ \underline{y} - \underline{y}^k \end{bmatrix} \leq 0 \end{cases} \quad \textit{for } k = 1, \ldots, K \\ &\underline{x} \in X, \ \underline{y} \in Y, \ a \in \mathbb{R}^1 \end{aligned} \tag{3.20}$$

The formulation above produces Z_L^K which is a valid lower bound for the problem presented in equation (3.15).

3.3.3.4 Modes of interaction between the MILP and NLP problems

IT has been already mentioned that the algorithmic methods developed to solve the MINLP optimisation problems are mainly differentiated by the way the NLP and MILP problems interact. In this paragraph, the different means of interactions will be presented for the three major families of algorithms used in mixed-integer non-linear programming.

3.3.3.4.a Generalized Benders Decomposition

USING Benders' approach (1962) for exploiting the structure of mathematical programming problems with "complicating variables", Geoffrion (1972) originated the Generalized Benders Decomposition which generalises Benders' initial method to a broader class of programs. "Complicating variables" can be defined as variables which, when temporarily fixed, render the remaining optimisation problem considerably more tractable. For the class of problems specifically considered by Benders, fixing the value of "complicated variables" reduces the given problem to an ordinary linear program, parameterized by the value of the "complicating variables" vector. Geoffrion generalised the initial method to problems in which the parameterized sub-problem is not necessarily a linear one.

In the GBD, the MINLP problem is split into two sub-problems: a MILP master problem and a NLP sub-problem. The spirit of the method consists in generating the points $(\underline{x}^k, \underline{y}^k)$ by solving successively NLP sub-problems and MILP master problems, in a cycle of iterations. The NLP sub-problems have the role of optimising the continuous variables and producing an upper bound that corresponds to the best current solution, while the MILP master problems provide a sequence of lower bounds for the objective function, as well as new 0 – 1 variables values for each new cycle of iterations. The cycle of iterations is stopped when the lower and upper bounds are within a specified tolerance.

The main feature of this method is the way the MILP master problem is defined. In this particular case, the master problem is given by a dual representation of the continuous space where only active inequalities are considered and the set $\underline{x} \in X$ is disregarded. Duran and Grossman (1986a) proved that compared to other algorithmic methods, GBD generally requires a larger number of iteration cycles to reach the optimised result. It can become problematic when the number of 0 – 1 variables increases.

3.3.3.4.b Outer Approximation

FROM a conceptual point of view, algorithms based on outer-approximation describe the solution region of a given problem as the intersection of an infinite collection of sets. Duran and Grossman (1986a) derived an algorithm based on this principle for a class of mixed-integer non-linear programs. The OA method is similar to the GBD algorithm in the way it is defined. Here again, NLP sub-problems and MILP master problems are solved in a cycle of iterations.

The NLP sub-problems are solved using fixed binary variables and produce an upper bound to the objective function. Then, relaxed versions of the MILP master problems give a sequence of lower bounds for the objective function. The search is stopped when both bounds have reached a satisfactory level of accuracy. The only one difference with the GBD approach resides in the way the MILP is defined: a primal approximation of the continuous space is preferred to the dual representation used for the GBD.

In both methods, the MILP master problems accumulate new constraints as iterations go on. This is one of the main drawbacks of these methods, particularly when they are applied to large-scale applications, because when the MILP master problems grow in size, it has a major impact on the computational effort to be provided. However, as the master problems of the OA method are richer in information than those of the GBD are, OA produces lower bounds which are greater than or equal to those of the GBD. Consequently, the OA method generally requires less cycles of iterations to reach the solution. It is also important to note that the MILP problem is not necessarily solved to optimality when using an outer approximation.

3.3.3.4.c LP / NLP-based Branch and Bound

THE Branch and Bound method developed by Gupta and Ravindran (1985) for the non-linear integer programming is a direct extension of the linear case (MILP). This method starts by relaxing the integrality requirements of the 0 – 1 variables which leads to a continuous NLP optimisation problem. If all of the integer variables $\underline{y}$ take discrete values, the search is stopped. Otherwise, it performs a tree search in the space of the integer variables $\underline{y}$ and solves a sequence of relaxed NLP sub-problems. In other words, it performs a tree enumeration where a subset of the 0 – 1 variables are successively fixed at each node. The solution of the corresponding NLP at each node provides a lower bound for the optimal MINLP objective function value. The search is stopped when the lower bound exceeds the current upper bound, or when all integer variables $\underline{y}$ take discrete values. Clearly, the size of the tree is dependent on the number of binary variables in the formulation

The extension of the method to NLP problems is due to Quesada and Grossman (1992). The LP / NLP-based BB avoids solving a complete MILP master problem at each major iteration which is its main advantage. The basic idea here is to perform a LP-based BB method for the MILP master problem, solving NLP sub-problems, only for the nodes in which feasible integer solutions are found.

By updating the representation of the MILP master problem in the nodes of the tree with the addition of the corresponding linearization, the need to restart the tree search is suppressed. The major disadvantage of the LP / NLP-based BB method is that it usually requires the solution of a relatively large number of NLP sub-problems which cannot be updated as easily as in the linear case.

This method can be applied to both GBD and OA algorithms, and generally, allows an appreciable reduction of the number of nodes enumerated. Although the number of NLP sub-problems to be solved increases significantly, the use of a LP / NLP-based BB algorithm within GBD or OA formulations generally brings significant computational savings in the search for an optimal solution.

3.4 Concluding remarks to the chapter

IN this chapter, a comprehensive review of the previous work which can be linked to the operational optimisation of the Durban water distribution network, has been conducted. Firstly, past and more recent advances in the field of optimisation applied to the water industry have been presented. Then, the theoretical bases needed to understand the tools developed for this particular research project have been detailed. These theoretical developments have been organised in two categories, referring to the techniques involved within the scheduling problem, that is to say a model predictive approach, associated with mixed-integer non-linear programming. For each theoretical aspect, backgrounds of the method have been given and have been followed by a literature survey centred on the methods used. Finally, the theory itself has been presented with the associated equations involved in the method.

In the next chapter, a presentation of the model developed to solve for the operational optimisation of Durban water distribution network is given.

CHAPTER 4 - MODELLING OF A WATER DISTRIBUTION NETWORK FOR OPERATIONAL OPTIMISATION PURPOSES

IN the present study, the necessity to achieve optimal planning within a system combining dynamic elements (reservoir volumes and chlorine concentration responses) and discrete elements (pump and valve switching or routing) makes this a constrained optimisation and a challenging non-linear Model Predictive Control (MPC) problem which can be solved using Mixed-Integer Non-Linear Programming (MINLP) techniques.

Non-linear MPC techniques for constrained and multi-variable systems have been developed and applied by many workers (e.g. Mulholland and Narotam, 1996; Le Lann, 1999; or Deghaye et al., 2000). However, these researchers have not considered the possibility of discrete (integer) variables in their structures which constitutes the major interest of the work presented here. During the past 15 years, considerable progress has been made in the theory and practice of non-linear MPC. These techniques are most appropriate for constrained multi-variable processes which are non-linear enough to make conventional control techniques inadequate. The commercial market for this technology is wide and its future depends on continued development of effective modelling techniques, alternative problem formulations, efficient solution strategies and novel process applications (Henson, 1998). As described in Chapter 3, non-linear MPC requires the availability of a suitable non-linear dynamic model of the process for the prediction of the vector of manipulated input variables. Therefore, the development of non-linear process models is of paramount importance. Unfortunately, due to the complexity of non-linear systems, it is rarely possible to develop non-linear system identification techniques by extension of the linear theory. Alternatively, the non-linear MPC controller is usually based on a fundamental model derived by applying transient mass, energy or momentum balances to the process (Ogunnaike and Ray, 1994).

In the absence of spatial variations, the resulting models have the general form presented below:

$$\dot{x} = f(x, u) \tag{4.1}$$

$$0 = g(x, u) \tag{4.2}$$

$$y = h(x, u) \tag{4.3}$$

In the above formulation, x is a vector of process state variables, u is a vector of manipulated input variables and y is a vector of controlled output variables.

The ordinary differential equations in (4.1) and the algebraic equations in (4.2) are derived from conservation laws, while the output equations in (4.3) are usually chosen by the control system designer.

A non-linear discrete model has been developed for the operational optimisation of water distribution networks. The purpose of this model is to provide a comprehensive representation of any existing water distribution network worldwide which is then used to predict future reservoir volumes and chlorine concentrations, using economic, quality and security considerations.

The model presented in this chapter creates a computer version of a water supply system. In other terms, it creates a bridge between the physical realities and the mathematical relations associated with these realities. By translating physical behaviours into mathematical relations, it serves as an interface with the optimisation tools used to solve the problem of optimal sequencing of valves and pumps. In the course of this project, two different modelling strategies based on this model had to be developed, the results of the first one not being satisfactory. Referred to as "calculated flows" strategy and "available flows" strategy, these modelling strategies refer to different concepts, but are based on a similar "modelling framework". Firstly, the model is presented from a general point of view, where it is possible to find information about features of the modelling framework that are common to both strategies. Then, the two different modelling strategies are detailed separately, concentrating on the unique features of both approaches, and the impact that these differences have on the modelling framework.

4.1 General presentation of the model

THE non-linear model developed especially for the operational optimisation of water distribution networks belongs to the family of fundamental models. Henson (1998) has recently provided an interesting overview of current non-linear MPC technology and applications. In his paper, he devoted a paragraph to the significant effort that has to be dedicated to the elaboration of a model for non-linear MPC applications. This researcher states that non-linear fundamental models have several advantages in comparison with the non-linear empirical models against which they position themselves. For instance, he reckons that less process data is required for the development of fundamental models, since they are usually highly constrained with respect to their structure and parameters. In particular, model parameters may be estimated from laboratory experiments and routine operating data, instead of the time-consuming plant tests required by empirical models.

Furthermore, as long as the assumptions made during the development phase remain valid, fundamental models can be expected to extrapolate to operating regions which are not represented in the data set used for model development. This property is particularly important when a process operates over a wide range of conditions, as a water distribution may.

A good model must represent with accuracy all of the elements included in the system that is modelled, as well as all the physical phenomena taking place within the system. Henson (1998) cites Ricker and Lee (1995) as one of the few references available in the literature presenting a model capable of reflecting the large-scale nature of typical industrial applications. He links the absence of work in this field to the difficulties involved in deriving fundamental dynamic models for large-scale processes.

In the particular case of the operational optimisation of water distribution networks, an original model, based on new concepts, had to be developed to take into account the large-scale aspect of these systems, without provoking unsolvable complications. This model has been developed in such a way that all of the typical features existing in a water distribution network can be represented within the same common framework, referred to as the "modelling element". The principal elements constituting a water distribution network have been presented in Chapter 2: they are the storage reservoirs, the pipes, the valves and the pumps. It is recalled here that storage reservoirs are positioned at strategic points in the network, mainly to regulate flows and pressures, as well as to provide buffer capacities during emergencies. Pipes have the ability to convey the water from one place to another, while valves are used as a means to control and direct the flows to the different operating units of the network. Finally, pumps help to convey the water more efficiently, and with adequate levels of pressure, even to the most remote part of a distribution system.

4.1.1 Standard features of the modelling element

THE modelling element developed for the operational optimisation of water distribution networks is based on a standard framework This framework is common to the two modelling strategies followed during the course of this project, and which are presented later. Even if each modelling strategy is based on different concepts, several features of the modelling element are found in both approaches. These standard features are presented in this section.

4.1.1.1 General description of the modelling element

THE modelling element is a mixed-flow vessel equipped with one input and two separate outputs. An optional auxiliary input can also be defined. This modelling element is presented in Figure 4.1.

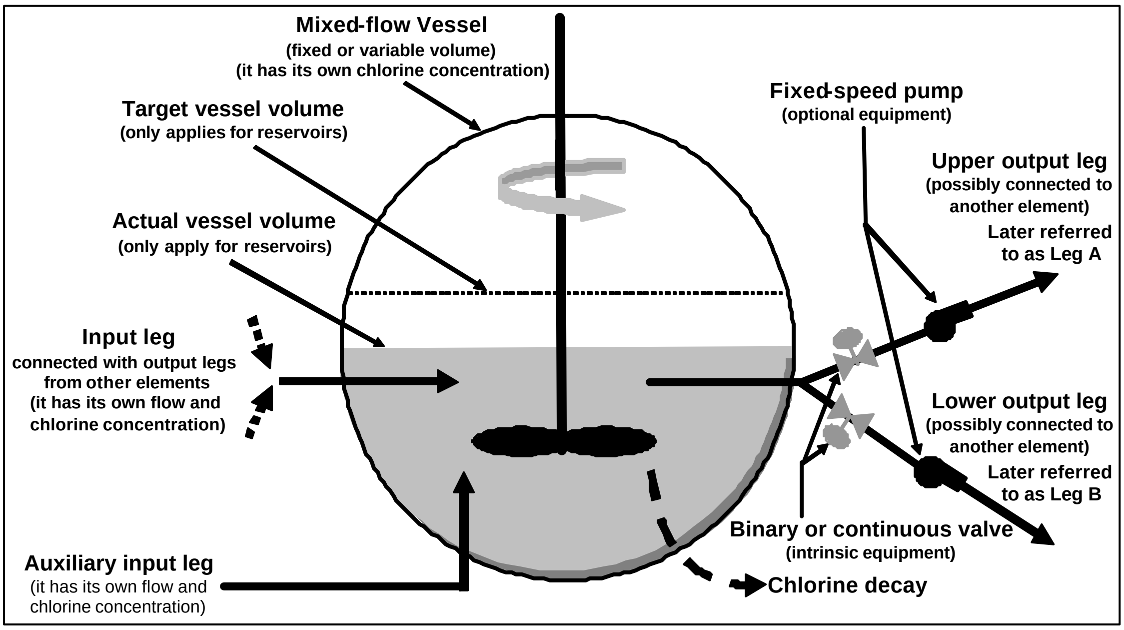


Figure 4.1: General representation of the standard modelling element

Figure 4.1 shows that this standard modelling framework can be divided into three different zones: the mixed-flow vessel, the input zone, and the output zone. Data have to be captured for each of these constituent parts. However, the physical properties needed for the characterisation of an auxiliary input leg are different from those representing the behaviour of the mixed-flow vessel or of the output legs.

4.1.1.1.a Mixed-flow vessel

THE inputs and the outputs of the standard modelling element are linked through a mixed-flow vessel which represents the most important part of the modelling element. The vessel's main parameter is its volume which can either be fixed or variable. Depending on how this volume is defined, the vessel can represent an open storage tank, a reservoir under pressure or a pipe offtake.

When an open storage tank or an under-pressure reservoir is modelled with a standard element, an initial volume has to be given. For an open storage tank, the volume is a variable of time, while for an under pressure reservoir this volume is fixed. In both cases, this volume must be different from zero.

Setpoint, low-emergency and high-emergency volumes have to be provided too. Besides the reservoir volume, the chlorine concentration in the tank also requires initialisation. A setpoint, a low-emergency level, a high-emergency level and a chlorine decay constant also need to be defined for the chlorine property.

Within the framework of a modelling element, a pipe off-take is equivalent to a reservoir under pressure with a zero volume. Consequently, when a pipe off-take is defined, the volume of the mixed-flow vessel is taken equal to zero. No other parameter of the vessel requires definition for this particular configuration

4.1.1.1.b Input zone

THE input zone consists of the input leg and the auxiliary input leg. Unlike the input leg and the two output legs mentioned earlier, the auxiliary input is not a connectible feature of the modelling framework. It is an optional feature which is only used to represent external water supply or chlorine re-dosing points.

The input zone is the one among the three constituting parts of a standard element for which the least data has to be provided. No data is indeed required for the compulsory input leg, while when an auxiliary input leg is added, information is only provided for the auxiliary input flow and for the chlorine concentration in the auxiliary input .

4.1.1.1.c Output zone

THE output zone consists of one line which is split into two separate legs, each fitted with a valve. Valves are intrinsic equipment of a standard element and can operate according to two different modes: binarily or continuously. In the formulation of this modelling element, a line free of valves is represented by a line on which a fully opened valve is fitted. Each output leg can also be fitted with fixed-speed pumps but, in this case, pumps constitute optional equipment. The possibility to include variable-speed pumps is not included in the model, since this family of pumps is barely used in a water distribution network. It is remembered here that binary valves and fixed-speed pumps can only access 0 – 1 values, while continuous valves have the ability to access the full range between 0 and 1.

The compulsory information required for the output legs concerns the definition of initial values for the flow on each leg, as well as the type of valves used (binary or continuous) and their initial position (fully opened, fully closed or intermediate). The initial valve positions depend of the category of valves present on the leg.

Binary valves can only be initialised as fully shut or fully opened, while the initial position of a continuous valve usually corresponds to a value chosen between 0 and 1. However, when a pump is added on one of the output legs, the valve present on the same leg is automatically opened fully, and no information must be provided for this particular valve.

Depending on the presence of pumps on the output legs, additional information can be required for the output zone. In this case, specification is required for the initial status of these pumps (on or off). It is also necessary to provide a pump parameter from which an operating cost can be calculated based on the pump flow.

4.1.1.2 Definition of the properties of a standard element

SEVERAL general rules have been defined to make the definition of the properties of the modelling element easier. These rules apply to the three constituting zones presented in the above paragraph.

4.1.1.2.a Definition of a data table

ALL of the properties of a standard element are defined within the same numerical table. This table consists of as many columns as there are properties to define in a modelling element. If use is made of the framework presented in Figure 4.1, it appears that the corresponding data table consists of nineteen separate columns, as presented in Table 4.1.

Table 4.1: General data table for the definition of the properties of the modelling element

Physical property	X_A	X_B	f_A	f_B	f_{aux}	C_{aux}	C	C_{Sp}	C_{min}	C_{max}	k	V	V_{Sp}	V_{min}	V_{max}	$Pump_A$ 1st param.	$Pump_A$ 2nd param.	$Pump_B$ 1st param.	$Pump_B$ 2nd param.
Corresp. value																			

Glossary:

X_A – Valve position or pump status on leg A
X_B – Valve position or pump status on leg B
f_A – Outpout flow on leg A (ML/d)
f_B – Outpout flow on leg B (ML/d)
f_{aux} – Auxiliary input flow (ML/d)
C_{aux} – Chlorine concentration in the auxiliary input flow (mg/L)
C – Chlorine concentration in the vessel (mg/L)
C_{Sp} – Chlorine concentration setpoint in the vessel (mg/L)
C_{min} – Minimum chlorine concentration in the vessel (mg/L)
C_{max} – Maximum chlorine concentration in the vessel (mg/L)
k – First-order chlorine decay constant (d^{-1})
V – Volume in the vessel (ML)
V_{Sp} – Volume setpoint in the vessel (ML)
V_{min} – Minimum volume in the vessel (ML)
V_{max} – Maximum volume in the vessel (ML)
$Pump_A$ – Pump characteristic coef. on leg A
$Pump_B$ – Pump characteristic coef. on leg B

This table only includes numerical values relative to features that are common to the two modelling approaches developed during the course of this project. Features that are specific to each modelling approach will be presented later in this chapter, as well as the updated versions of the general data table resulting from these special features.

4.1.1.2.b Organisation of the data table

FROM a general point of view, it can be observed that this table includes two different types of information. The first category of data refers to initialisation values which serve as a good starting point for the optimisation, while the second category of data corresponds to values required for the definition of intrinsic characteristics of the standard modelling element. Valve positions, pump status, auxiliary input flow, output flows, chlorine concentration in the auxiliary input flow and in the vessel, as well as volume in the vessel are parameters that are part of the first category. Constraints on chlorine concentration and volume in the vessel, as well as the chlorine decay constant and the pump characteristic parameters belong to the second category. While some of the data belonging to the first category can be either constant or variable in the course of the optimisation, those from the second category are necessarily constant parameters.

Another important feature of this table consists in the way the information is organised, especially for data referring to the output zone described earlier. For this part of the standard element, the data capture is done separately for each output leg. Hence, valve positions, pump status, output flows and the two parameters of the pump characteristics require two columns each in the data table.

Taking a closer look, the two first columns of the data table are reserved for the definition of the initial position of valves, or the initial status of pumps. The model allows the definition of constant or variable valve positions, while the statuses of pumps are always variable. It can appear surprising that initial valve positions and pump status share the same column in the table. However, to justify this modelling choice, it is remembered here that valves and pumps cannot occur simultaneously on the same output leg.

A valve can be added alone on a leg, but, as soon as there is a pump on a leg, any valve present on the same leg is automatically fully opened. Since the geometry of the table is common to any of the equipment associated with the output legs, a special method had to be devised to differentiate, without any confusion, continuous valves from binary valves or fixed-speed pumps. This scheme is presented later in this chapter.

The following three columns of the table serve for the definition of initial values for the auxiliary input flow and for the two output flows. These flows can be defined either as variables, or as constant parameters. When they are considered variable, the initial values provided in the table will be updated at each new time-step, up to the prediction horizon, according to the changing conditions in the vessel. On the contrary, when defined as constants, these values are referred to as feed flows when used for the auxiliary input flow, or as user demand flows when used for output flows. Feed flows have to be provided when an auxiliary input flow is used to represent the connection between an external source of water and the current standard element. User demand flows have to be provided when one of the outputs of any standard element is connected to the reticulation system. In this model, predicted user demand flows are captured before the optimisation is run. A value is required at each new time-step, for each user demand flow, until the optimisation has reached an end. Providing accurate user demand flows constitutes one of the essential prerequisites of the model presented in Chapter 3. Usually, user demand flows are based on statistics derived from historical data.

The six next columns all refer to the chlorine concentration in the vessel. The initial chlorine concentration in the auxiliary input flow and the initial chlorine concentration in the vessel are defined in separate columns and can be considered as either constants or variables. Finally, the chlorine concentration setpoint in the vessel, low-emergency and high-emergency levels, as well as the chlorine decay constant are entered in the four following columns. These values are necessarily constant parameters. Note that no chlorine concentrations need to be defined for the output legs, since they are automatically equalled to the chlorine concentration inside the vessel. Concerning the decay of chlorine, most network modelling tools generally assume that it can to be represented by first-order kinetics (Chambers *et al.*, 1995). Although Powell *et al.* (2000) criticise the fact that these models fail to account for any temporal variability in the decay characteristics, he admits that second-order kinetics, although fitting observations better, would be far more difficult to implement with the current stage of the research in this field. Consequently, for this model, it has been decided to adopt first-order kinetics.

The following part of the table corresponds to the characterisation of the volume of the vessel: initial volume in the vessel, volume setpoint, low-emergency volume and finally high-emergency volume are successively defined. While the initial volume is either a variable or a constant parameter, the three other properties are necessarily constants.

Finally, the four last columns of the data table are reserved for the definition of pump characteristic parameters.

The pump characteristics used here represent the power consumed by a pump for the flow going through this particular pump. It is assumed that the power consumed can be expressed by:

$$Power = Flow \times Pump1 + Pump2 \tag{4.4}$$

Therefore, two parameters $Pump1$ and $Pump2$ have to be defined for each pump: these are constants parameters.

4.1.1.2.c Quantity of data to provide for a particular modelling element

SINCE the quantity of data to provide for a particular modelling element depends of the type of equipment (open reservoir, under pressure tank, or pipe offtake fitted or not by valves or pumps), some of the data appearing in the storage table may not have to be defined. Consequently, when this is the case, a rule states that the numerical value 0 is set in the columns corresponding to the properties which do not require definition. For instance, the modelling of a pipe offtake would require zero values in all of the columns linked to data concerning the mixed-flow vessel (chlorine concentration initialisation, setpoint and emergency levels or volume initialisation, setpoint or emergency levels). Similar examples may be considered with different equipment. However, an exception to this rule occurs when there is a will to represent a line free of valves. Since a valve is an intrinsic item of the modelling element, when this situation occurs, the numerical value 1 is set in the column of the numerical table corresponding to the valve position. This choice is obvious, since a valve position of 1 corresponds to a fully-opened valve.

The rule presented in this paragraph is of particular significance with respect to the geometry of the modelling element. Until now, this element has always been represented with an output line splitting into two output legs. However, with the introduction of the previous rule, this is not necessarily the case, since one of the output legs can be neutralised by adding zero values in all of the columns referring to this particular leg.

4.1.1.2.d Difference between variable and fixed parameters

THE difference made between variable and fixed parameters is probably the most important rule of all. In this model, negative values have been chosen to represent parameters that have to be modified to reach an optimised state. Absolute values of the negative entries are taken as starting points by the optimisation program. These must be within constraints for the reaching of a feasible solution. On the contrary, positive (or zero) values are used for fixed parameters. Fixed parameters are not changed during the optimisation: they are considered as constants.

Variable parameters can be used for output flows (only for output legs which are not connected to the reticulation system), reservoir volumes, chlorine concentration in the reservoirs, as well as for the positions of variable continuous valves.

Fixed parameters are mainly used for setpoints (volume and chlorine concentration setpoints), constraints (low / high emergency volumes in the reservoirs, low / high emergency chlorine concentration in the reservoirs), pump characteristic parameters, as well as for the chlorine concentration in the auxiliary input flow and for the first-order chlorine decay constant. However, they are also required for the definition of the position of fixed continuous valves and are necessary to define feed flows (auxiliary input flows) and user demand flows (output flows).

An exception to the rule differentiating variables from fixed parameters can be found for the definition of the initial position of binary valves and of the initial status of fixed-speed pumps. Similarly to continuous valves, the definition of these items requires the capture of an initial position, or an initial status. By definition, binary valves and fixed-speed pumps can only access positions or status represented by zero or one values. Therefore, their initial position or status in the table must necessarily be equal to zero or one. As presented earlier, the use of a zero, a one or an intermediate value in the two first columns of the table is reserved for continuous valves. However, no matter which kind of equipment is fitted on the output legs, the definition of its initial status is done in these two first columns of the data table. Consequently, special rules have been developed to prevent any confusion in the definition of these different items. These rules are based on the use of fixed parameters, to which special values are given. Hence, when there is a will to add a binary valve on one of the legs, the initial valve position must be equalled to the numerical value 2 or 3 in the corresponding column of the data table. These two values each refer to a different initialisation: 2 for shut (0), 3 for open (1). Similarly, when a fixed-speed pump is present on one of the output legs, the initial pump status must be equalled to 4 or 5 in the corresponding column of the data table. Once again, these two values each refer to a different initialisation state: 4 for off (0), 5 for on (1).

In this particular case, fixed parameters are used in a very special context. The use of the numerical values 2, 3, 4 and 5 in the two columns of the table referring to valve openings or pump status allows differentiating continuous valves from binary valves or fixed-speed pumps. The use of fixed parameters in this context does not mean at all that the status of binary valves and fixed-speed pumps has to remain unchanged, since by definition the status of these variable items must be actuated during the course of the optimisation process, at each time-step, up to the prediction horizon.

4.1.1.2.e Definition of feed flows and users demand flows

FEED flows and user demand flows occur when a positive value is entered in the corresponding columns of the data table. With these flows being defined as fixed parameters, their values are considered constant during the course of the optimisation process. This is not a very realistic situation, since most of the feed and user demand flows are variables with time. Consequently, a method had to be devised to represent feed flows and user demand flows that are variable, using fixed parameters in the data table presented Table 4.1. This method consists in the definition of an auxiliary data table, in which time-varying coefficients are entered for feed flows as well as user demand flows, for each time-step of the optimisation process, up to the prediction horizon. The auxiliary data table consists of as many columns as there are positive flows in the main data table (which gives a maximum of three columns per element), while the number of rows is equal to the length of the prediction horizon N to which the length of the control horizon M is added. Both prediction and control horizons have been introduced in Chapter 3. The structure of the auxiliary data table is presented Table 4.2.

The values of the time-varying coefficients entered in the auxiliary data table must lie between 0 and 1, since fixed flows are specified as maxima in the main data table. At the beginning of each new time-step, an update process takes place for the calculation of actuated feed flows and user demand flows. The updated values of these flows is obtained by multiplying the value of the time-varying coefficient at the current time-step with the initial value of the feed flow or users demand flow, defined in the main data table presented as Table 4.1. It is obvious that the values of feed flows and user demand flows that are defined in the main data table represent base values and must therefore be equal to the maximum feed or user demand possibly achieved during the course of the optimisation process.

Table 4.2: Auxiliary data table used for the time-varying coefficients of feed and demand flows

Time-Step	Time-varying coefficient for feed flow	Time-varying coefficient for demand flow on leg A	Time-varying coefficient for demand flow on leg B
t			
t + 1			
t + 2			
t + 3			
...			
t + N			
...			
t + N + M			

The rule presented in this paragraph constitutes another exception to the notion of fixed parameters as it was defined in paragraph 4.1.1.2.d.

4.1.1.2.f Definition of time-of-use tariffs for electricity

REDUCING the operational costs linked to the electrical consumption of pumps during pumping operations is the main optimisation objective of this project. To achieve this objective, it is obviously necessary to provide data relative to the cost of an electrical unit. It is well known that electricity providers frequently devise complex tariff structures for industrial users, with prices varying according to the time of day and even the day of week and the month of the year.

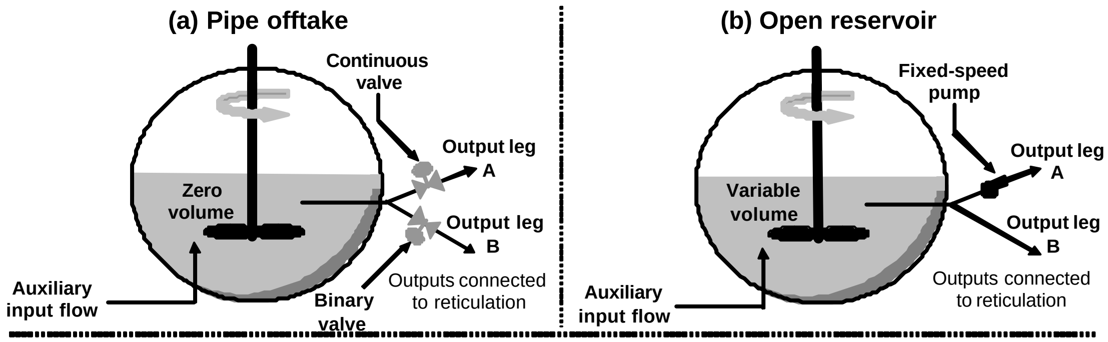
Within this model, the time-of-use tariffs are represented by time-varying coefficients, similarly to the approach presented in the above paragraph for feed flows and users demand flows. Therefore, for each time-step of the optimisation, a new coefficient, referred to as E_{cost} and representing the current tariff, is captured in a numerical table referred to as the table of time-of-use tariffs. The calculation of the cost of pumping operations is based on these time-varying coefficients.

The table of time-of-use tariffs consists of a single column. The number of rows is equal to the length of the prediction horizon N to which the length of the control horizon M is added.

4.1.1.2.g Application examples

LET us consider the two simple configurations presented in Figure 4.2 to illustrate the rules presented in the above paragraphs. Firstly, a modelling element is used to represent a pipe off-take fed directly by an external source of water (configuration (a) on Figure 4.2). This pipe is then split into two different legs. The upper leg is fitted by a continuous valve, while a binary valve is added on the lower leg. The binary valve is initialised to a fully opened position. Then, another element is used to represent an open reservoir fed directly by the output of a water treatment plant (configuration (b) on Figure 4.2).

The output of the reservoir splits into two legs. A fixed-speed pump is positioned on the upper leg and initialised to off, while neither a valve, nor another pump is added on the lower leg. The pump that is defined is a booster pump which means that it allows a flow in a line even when it is in the off position. For both examples, the output legs are connected to reticulation which means that user demand flows have to be specified for each output leg. In these examples, the user demand flows do not change with time.



(c) Main data table & initialisation values

Physical property	X_A	X_B	f_A	f_B	f_{aux}	C_{aux}	C	C_{Sp}	C_{min}	C_{max}	k	V	V_{Sp}	V_{min}	V_{max}	$Pump_A$ 1st param.	$Pump_A$ 2nd param.	$Pump_B$ 1st param.	$Pump_B$ 2nd param.
Pipe offtake	0.3	3	8.5	13.2	15.75	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Open reservoir	4	1	10	5	20	2.0	-1.8	1.8	1.3	2.4	2	-17.6	20.3	6.8	32.1	6.6	434.6	0	0

Figure 4.2: Examples of definition of the properties of a standard element

The pipe offtake is probably the configuration which requires the least data. An example of the values that need to be entered in the main table to represent this particular configuration can be found on Figure 4.2. Let us review the table column per column. Initial valve positions and the statuses of pumps are defined in the two first columns. Since a continuous valve is present on the upper output leg, an initialisation value of 0.3 is chosen. This value is positive since the leg on which this continuous valve is fitted is connected to a reticulation system, where the users demand flow is constant. Hence, the valve position does not need to be updated during the course of the optimisation process. The valve on the lower leg is binary and initialised as fully opened. According to the rules defined in paragraph 0, a numerical value of 3 must be entered for this parameter in the data table.

Flows are defined in the three next columns of the table. Since the output flows are both connected to reticulation, they are equivalent to user demand flows and according to what has been defined in paragraphs 0 and 4.1.1.2.e, their values in the main data table must be declared positive. In this example, the flow in the upper output leg is taken equal to 8.5 ML.d^{-1}, while the flow in the lower output leg is equal to 13.2 ML.d^{-1}. These values may have been derived from statistics based on logged demand flows. However, since the position of the continuous valve on the upper output leg is fixed at a value below 1, the flow going through this leg is inferior to the value defined in the main data table. Here, the actual flow will be equal to the result of the product between the valve position in the line and the users demand flow in the same line. This gives a value of 2.55 ML.d^{-1} going through the upper output leg. Similarly, since the pipe offtake is fed directly from a connection with another aqueduct of the same distribution system, an auxiliary input flow has to be defined.

This auxiliary flow corresponds to a feed flow, and therefore a positive value must be entered in the corresponding column of the main data table. With this configuration, since the mixed-flow vessel has a zero volume, the auxiliary feed flow is necessarily equal to the flow drawn by the two outputs. The auxiliary input flow can be expressed as follows:

$$f_{aux} = \left| X_A \quad f_A \right| + \left| X_B \quad f_B \right| \tag{4.5}$$

The calculation with the data entered in the table gives an auxiliary input flow equal to 15.75 $ML.d^{-1}$. The next six columns refer to properties relative to the chlorine concentration in the auxiliary input flow and in the vessel. Besides the first-order chlorine decay constant which need to be defined and which is taken equal to 2 d^{-1}, no other chlorine concentration specifications are required for a pipe offtake. According to the rule presented in paragraph 4.1.1.2.c, zero values are added in the corresponding columns of the main data table. The following four columns serve for the definition of the characteristics of the mixed-flow vessel. A pipe offtake is defined as a zero-volume reservoir in this model. Consequently, the numerical value 0 is entered in the corresponding column of the main data table. No other specifications are required on the volume properties for a pipe offtake, and zero values also have to be entered in the three remaining columns referring to the volume characteristics of the vessel. Finally, the last four remaining columns of the main data table serve for the definition of pump characteristics. Since there is no pump on any of the output legs, zero values must be added in the four last columns of the main data table.

The open reservoir configuration is probably the one for which the most data must be provided in the main data table. An example of the values that need to be entered in the main table to represent this particular configuration can be found on Figure 4.2. Let us review the table column by column. Initial valve positions and the statuses of pumps are defined in the two first columns. To represent a fixed-speed booster pump initialised to off on the upper output leg, the numerical value 4 is entered in the first column, accordingly to what has been said in paragraph 0. Since there is neither a valve, nor a pump on the lower output leg of this element, the numerical value 1 is chosen for the initial valve position and pump status of this leg, as presented in paragraph 4.1.1.2.c. Flows are defined in the three next columns of the table. Similarly to the pipe offtake case-study presented before, the output flows in the second configuration are also both connected to reticulation. Therefore, their values in the main data table must be declared positive. In this example, the flow in the upper output leg is taken equal to 10 $ML.d^{-1}$, while the flow in the lower output leg is equal to 5 $ML.d^{-1}$. Similarly, since the open reservoir is fed directly from the output of a water treatment plant, an auxiliary input flow has to be defined. This auxiliary flow corresponds to a feed flow, and therefore a positive value must be entered in the corresponding column of the main data table.

Here, the auxiliary input flow is taken equal to 20 $ML.d^{-1}$. For this particular configuration, it is not necessarily taken equal to the sum of the output flows, since the reservoir can act as a buffer capacity. The next six columns refer to properties relative to the chlorine concentration in the auxiliary input flow and in the vessel. A constant chlorine concentration of 2.0 $mg.L^{-1}$ is fixed for the feed flow, while a variable chlorine concentration is applied in the vessel. This concentration is initialised at 1.8 $mg.L^{-1}$. The negative sign in front of this numerical value is used to highlight the variable character of this particular parameter. The chlorine concentration setpoint in the vessel was fixed at 1.8 $mg.L^{-1}$, while the low-chlorine level was taken equal at 1.3 $mg.L^{-1}$ and the high-chlorine level at 2.4 $mg.L^{-1}$. Finally, a first-order chlorine decay constant of 2 d^{-1} was chosen. The following four columns serve for the definition of the characteristics of the mixed-flow vessel. Since in the second configuration the vessel represents an open reservoir, its volume is a variable parameter. In this example, it is initialised at 17.6 ML. The negative sign which is added in the table in front of this numerical value, is used to represent the variable nature of this parameter. The volume setpoint was fixed at 20.3 ML, while the low-emergency volume was taken equal to 6.8 ML and the high-emergency volume at 32.1 ML. Finally, the two parameters of the characteristic of the pump located on the upper output leg are entered in the table. Since there is no pump on the lower output leg, zero values must be added in the two last columns of the main data table, accordingly to the rule defined in paragraph 4.1.1.2.c.

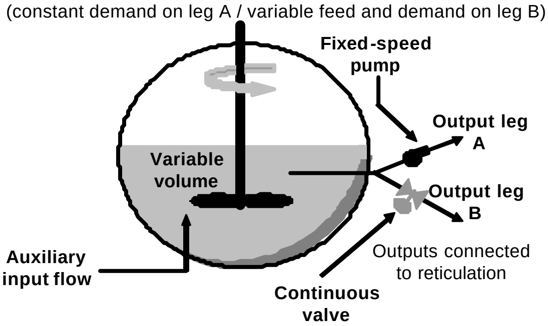
It is recalled here that, in both configurations presented in this paragraph, the output flows represent constant user demand flows. Therefore, there is no need for an auxiliary data table. However, having constant feed and user demand flows does not represent the most frequent situation in a water distribution system.

Let us consider what would change in the data definition protocol if the feed flows and user demand flows were defined as variable. For this example, only the open reservoir configuration is considered. The feed flow and the lower output flow are changed into variable flows, while the upper output flow is still considered constant. The prediction interval N is fixed at 5 hours while the control M horizon is taken equal to 2 hours. In other words, 7 hours of feed and user demand flows need to be provided before the optimisation is run. Consequently, the auxiliary data table that can be derived consists of seven rows.

The new configuration with the corresponding data tables is presented in Figure 4.3. Since the user demand flow on the lower output leg has become variable, the binary valve must be replaced by a continuous device, declared as a variable in the main data table.

(a) Open reservoir

(constant demand on leg A / variable feed and demand on leg B)



(b) Auxiliary data table

N

Time-Step	T-V coef. for feed flow	T-V coef. for demand flow on leg A	T-V coef. for demand flow on leg B
1	0.521	1.0	1.0
2	0.558	1.0	1.0
3	0.592	1.0	0.5
4	0.362	1.0	0.5
5	0.409	1.0	0.5
6	1.0	1.0	0.0
7	0.980	1.0	0.0

(c) Main data table & initialisation values

Physical property	X_A	X_B	f_A	f_B	f_{aux}	C_{aux}	C	C_{Sp}	C_{min}	C_{max}	k	V	V_{Sp}	V_{min}	V_{max}	$Pump_A$ 1st param.	$Pump_A$ 2nd param.	$Pump_B$ 1st param.	$Pump_B$ 2nd param.
Open reservoir	4	-1	10	5	20	2.0	-1.8	1.8	1.3	2.4	2	-17.6	20.3	6.8	32.1	6.6	434.6	0	0

Figure 4.3: Properties of an open reservoir of which feed and lower output flows are variable

However, it can be added that with this particular configuration the numerical values chosen initially for the valve positions have no significance. For convenience purposes, the numerical expression "-1" is added in the corresponding column of the main data table, but it does not necessarily mean that the valve is opened fully. This constitutes the unique change on the configuration of the open reservoir case study presented earlier. Therefore, in the main data table (table (c) on Figure 4.3), the valve position corresponding to the lower output leg is initialised to -1, accordingly to the rule defined above. The variable demand on the lower output leg is rather represented by the variable time-varying coefficients stored in the last column of the auxiliary data table (table (b) on Figure 4.3).

In this example, the auxiliary data table reveals that the initial user demand flow that is requested must be equal to the maximum possible demand flow defined in the main data table (and equal to 5 $ML.d^{-1}$). This translates into occurrence of a time-varying coefficient equal to 1 in the first row of this column.

The variable demand on the feed flow is represented by the time-varying coefficients stored in the first column of the auxiliary data. Unlike the definition of a variable user demand flow, the introduction of a variable feed flow does not cause any configuration change on the modelling element.

Although the demand flow in the upper output leg is constant, it can be observed that time-varying coefficients have nevertheless been defined. This is to show that a constant demand can also be represented within the auxiliary table by coefficients equal to a constant parameter.

4.1.1.3 Constitution of an entire water distribution network using the common framework

THE above paragraphs have presented the geometry, the features, and the different types of application proposed for the standard modelling element developed for the operational optimisation of existing water distribution systems. Until now, this standard framework has been represented as a single entity, capable of representing a storage reservoir or a pipe offtake, fitted or not with a valve or a pump. However, among all the properties of this standard framework, one has been deliberately omitted until now. The use of this particular feature leads to the creation of a comprehensive representation of an existing water distribution system that is based on the single element modelling framework presented in the above paragraphs.

The rule presented in this paragraph states that it is possible to link the output leg of a modelling element with the input of any other element. With the introduction of this new rule, the number of network connections possibly represented by an element increases. Besides the connection to reticulation or no connections at all, the two output legs can be linked to the input of another single element, or to the input of two different elements.

4.1.1.3.a Definition of the different linkage configurations available between two elements

FIGURE 4.4 presents the different linkage configurations available for each modelling element. The presence of equipment (valves or pumps) on the output legs does not change the frame of applications of these configurations.

On this figure, configuration (a) corresponds to a situation where the outputs of an element are connected to the input of another single element. This configuration is generally used to represent parallel lines.

Configuration (b) represents the configuration where the outputs of an element are connected to the inputs of two different elements. This configuration is generally the one used the most frequently, since it is used to represents splits in the main line.

Configuration (c) illustrates a configuration where a modelling element is used to represent the end of a main line. The output legs are therefore connected to a reticulation system which means that user demand flows must be defined on these legs. A good illustration of this configuration can be found in paragraph 4.1.1.2.f, where two case studies based on a similar configuration have been presented. Finally, configuration (d) corresponds to a scheme where no connections with any other device are realised. It is generally used to represent a closed line.

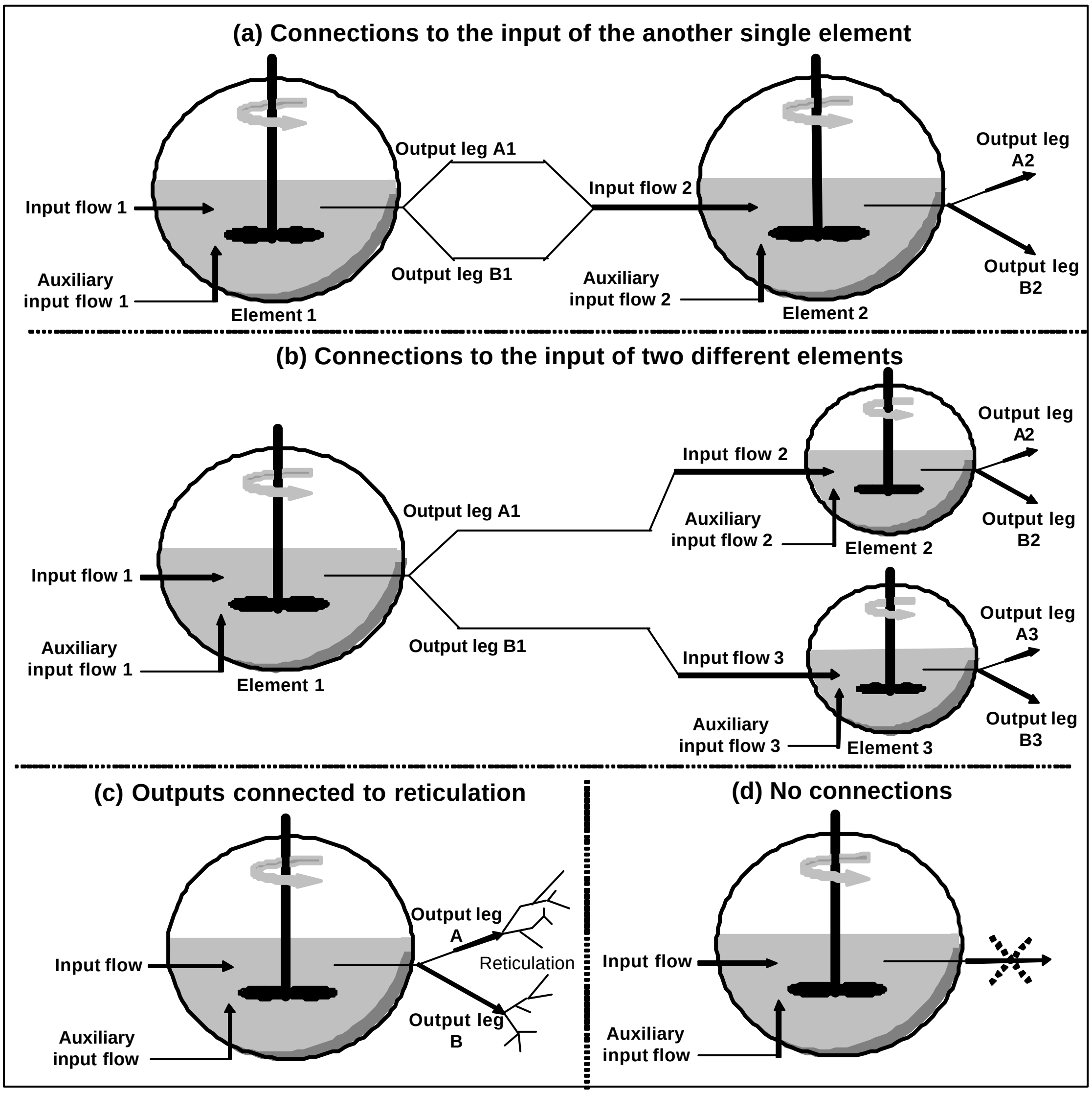


Figure 4.4: Different linkage configurations accessible to a modelling element

This configuration is important regarding the geometry of the modelling element. Until now, it may have seemed that it is compulsory to connect the output of an element to the input of another element. This is not necessarily the case, since the topology rules that have been defined enable the setting of no connections for one or both output legs. On the contrary, it is also possible to connect the outputs of several elements to the input of the same element: this constitutes an extension of the configuration (a) presented on Figure 4.4.

It is obvious that any water distribution network can possibly be constructed by linking several modelling frameworks together. From now on, the interactions between the elements become another property that needs to be defined during the phase of data capture. To build an exact representation of the topology of the network, each standard element is numbered separately. The total number of standard modelling elements needed to represent the water distribution system within the model is represented by the parameter nel .

Once the numbering is completed, a table is created to represent the topology numerically. This table consists of two columns and of *nel* rows. The upper output leg of each element is associated with the first column of the table, while the lower output leg of each element is associated with the second one. Each element is reviewed separately and a tracing process is conducted to identify towards which other elements the outputs of the current element are pointing.

4.1.1.3.b Application example

LET us consider the two configurations studied in paragraph 4.1.1.2.f. With the property defined in the above paragraph, it is possibly to link these two modelling elements together. As presented in Figure 4.5, the lower output of the pipe off-take is linked to the input of the open reservoir, while the upper output of the open reservoir is connected to the input of the pipe off-take. This configuration is equivalent to having a pipe off-take linked to an open reservoir via a pipe in which the flow can circulate in both directions. It represents a good example of a complex configuration that can be achieved with this model, since it is possible to feed the reticulation system linked to the upper output leg of Element 1 with water coming from the open reservoir represented by Element 2. For this particular configuration, several changes can be observed, compared to when the two modelling elements were taken separately.

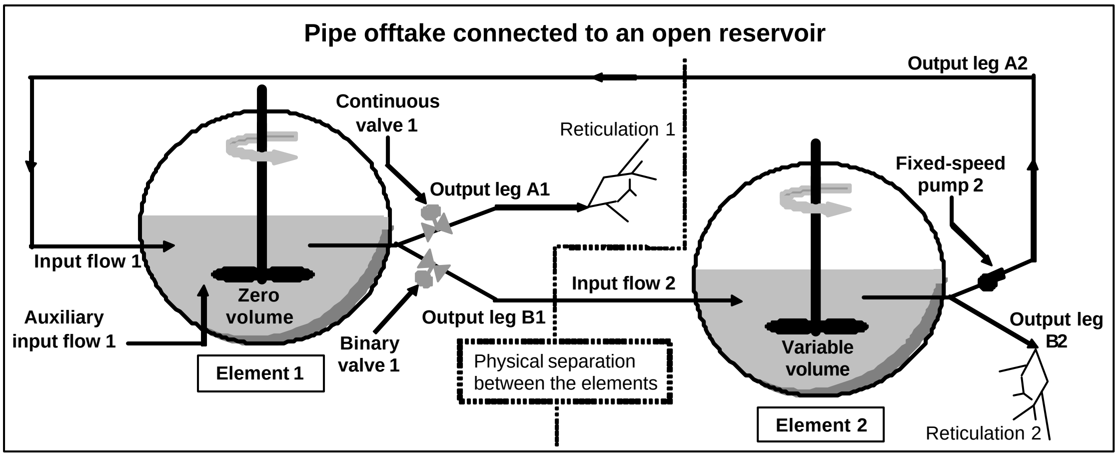


Figure 4.5: Illustration of a possible linkage between two modelling elements

On the first element, the main change occurs by the definition of an input leg which is in fact an extension of the upper output leg of the second element. The use of an input leg on an element has been ignored until now intentionally. An input leg only appears on a modelling element when there is an interconnection between the current element and one or several other elements. The main function of an input leg is to transmit the data from one element to another, decoding these data from the multiple legs which can attach to the input of the current element.

The input leg is the one part of the standard element that does not need the definition of any data in the main data table, since its properties are directly deduced from those set for the output legs of preceding elements from which the current element is fed. Time-dependent variables, like flows or chlorine concentration are the main variables that are transmitted between elements for the calculation of the new volumes or chlorine concentration in the reservoirs.

Still on the first element, the new feed flow arriving from the second element is defined as variable. Furthermore, the lower output leg is not connected to the reticulation system anymore, but rather to the input of the second element.

For the second element, an input leg is also defined: it represents an extension of the lower output leg coming from the first element. Furthermore, the upper output leg of Element 2 is linked with the input of the first element. It can be observed that the auxiliary input flow has been removed for this particular element.

When a water distribution system needs to be represented within this model, the first task is always to represent schematically the network in terms of standard elements, as has been done in Figure 4.5. All of the devices fitting the line must be identified and represented on the schematic diagram. The next step consists in carrying on with the numbering of the elements. This is the most dangerous task of all since the way the modelling elements interact between each other and exchange information is determined by this numbering. The presence of mistakes in the capture of the data associated to the topology of the network would obviously lead to results with no validity, since they would correspond to a distribution network which does not adequately represent the existing application on which the optimisation must be realised. A good hint consists in trying to keep, as far as possible, an upstream to downstream order in the numbering of the elements, so that the interactions between them are traced more easily. On a case study composed of two elements, this can appear easy, but it becomes rather complicated for complex or interconnected networks.

Once all of the standard modelling elements representing the water distribution network have been numbered, the topology table can be filled. In the particular example of a pipe off-take linked to an open reservoir represented in Figure 4.5, the topology table is composed of two rows and two columns. The lower output leg of Element 1 is linked to the input of Element 2. Hence, the numerical value 2, representing the second element, must be inserted on the first row and in the second column of the topology table. Similarly, the upper output leg of Element 2 is connected to the input of the first element.

Therefore, the numerical value 1, representing the first element, must be added on the second row and in the first column of the data table. Its content is shown in Table 4.3 (Table (a)).

Table 4.3 Topology, main, and auxiliary tables for the 2 modelling element - configuration

(a) Topology table

	Element number to which leg A is connected	Element number to which leg B is connected
Pipe offtake (element 1)	0	2
Open reservoir (element 2)	1	0

(b) Auxiliary data table

	T-V coef. for feed flow 1	T-V coef. for demand flow on leg A1	T-V coef. for demand flow on leg B2
Element number	1	1	2
Position in the main table	5	3	4
Time step 1	0.521	1.0	1.0
2	0.558	1.0	1.0
3	0.000	1.0	0.5
4	0.000	1.0	0.5
5	0.409	1.0	0.5
6	1.0	1.0	0.0
7	0.980	1.0	0.0

(c) Main data table & initialisation values

Physical property	(1) X_A	(2) X_B	(3) f_A	(4) f_B	(5) f_{aux}	(6) C_{aux}	(7) C	(8) C_{Sp}	(9) C_{min}	(10) C_{max}	(11) K	(12) V	(13) V_{Sp}	(14) V_{min}	(15) V_{max}	(16) $Pump_A$ 1st param.	(17) $Pump_A$ 2nd param.	(18) $Pump_B$ 1st param.	(19) $Pump_B$ 2nd param.
Pipe offtake	0.3	3	8.5	-5.66	15.75	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Open reservoir	4	-1	-0.01	5	0	2.0	-1.8	1.8	1.3	2.4	2	-17.6	20.3	6.8	32.1	6.6	434.6	0	0

Since there are no connections with any other elements, zero values must be added in the remaining boxes of the topology table. When zero values are captured in the topology table, they can represent either an output leg with no connections at all (configuration (d) on Figure 4.4), or an output leg connected to reticulation (configuration (c) on the same figure). The differentiation between the two configurations is made by an analysis of the initialisation values entered in the main data table for the corresponding output flows (Table (c) on Table 4.3). If these values are above zero, then the corresponding output is connected to a reticulation system and time-varying coefficients have to be defined for these flows in the auxiliary data table. Otherwise, the output leg is not connected to any other element and therefore, no further information is required.

In the particular example presented in Figure 4.5, the two remaining output flows that are not connected to the input of another element are linked to reticulation systems. Hence, they represent user demand flows which justifies the definition of time-varying coefficients for these flows in the auxiliary data table (Table (b) on Table 4.3). Definition of time-varying coefficients is also required for the auxiliary input flow defined on the first element.

As presented, it may be noticed that the auxiliary data table has been slightly modified compared to the ones presented earlier for configurations based on the use of a single element. When several elements are linked together, there is a need to trace exactly which variables of the main data table these time-varying coefficients apply. This is done by adding two rows at the top of the auxiliary data table. The tracing method developed here is very similar to the one developed to represent the topology of a water distribution system within the frame of the common modelling element. Besides the numbering of each standard modelling element, all the columns of the main data table are also represented by numbers. Therefore, all of the variables stored in this table (valve positions, pump status, flows, chlorine and volume properties, pump characteristics) can be easily traced thanks to the number representing the current element and the column position of the appropriate variable.

When the task of defining time-varying coefficients is undertaken in a configuration with several elements linked to each other, the two first rows of the modified auxiliary data table must be filled first. The first one is used to identify to which element the time-varying coefficients refer, while the second row of the modified auxiliary data table is used to identify on which input or output leg of the element these time-varying coefficients must be applied. User demand flows on the upper and lower output legs, as well as feed flows are each represented by the column number allocated to them in the main data table (respectively 3, 4 and 5).

From the geometry of the particular example presented in Figure 4.5, it is deduced that three flows require the definition of time-varying coefficients: the auxiliary input flow and the upper output flow on Element 1, as well as the lower output flow on Element 2. Any auxiliary input flow must be defined in the fifth column of the main data table presented Table 4.1. In this particular example, the auxiliary input flow belongs to the first element: hence the capture of the numerical value 1 on the first row and of the value 5 in the second line of Table (b) on Table 4.3. Similarly, any numerical value representing an upper output flow must be defined in the third column of the main data table. Since in this example the upper output flow requiring the definition of time-varying coefficients belongs to Element 1, it is necessary to capture a 1 and a 3 in the second column of Table (b) presented on Table 4.3.

Finally, any initialisation relative to a user demand flow leaving via the lower leg of an element must be captured in the fourth column of the main data table. On the particular arrangement represented in Figure 4.5, the lower output flow connected to reticulation belongs to Element 2. Therefore, the need to add a 2 and a 4 in the third column of the Table (b) shown in Table 4.3.

Once the topology of feed and users demand flows has been represented in the auxiliary data table, the rest of the definition protocol remains unchanged compared to the initial definition of the auxiliary data table in paragraph 4.1.1.2.e.

The final step in the phase of definition of the model is to capture the data of the main data table (Table (c) on Table 4.3). In this particular example, this table is still very similar to the one described earlier, when the two elements were considered separately. Changes have only been made in the three columns relative to the definition of flows of the main data table.

Let us begin with the element representing the pipe off-take. The single change on the corresponding row of the table concerns the initialisation value of the lower output flow. Firstly, it is now initialised as a variable parameter (represented by the negative sign) which means that the optimiser will update this parameter at each new time-step, up to the prediction horizon, according to the quantity of water provided inside the system and to the demand in water in other parts of the network. Furthermore, the initialisation value of this flow has also been changed. This can be justified because, in this configuration, the auxiliary input flow is varying with time, according to the the sequence of time-varying coefficients used with this flow (Table (b) on Table 4.3). By using the first time-varying coefficient of this sequence, it is easy to calculate that the initial auxiliary flow will be equal to 8.21 $ML.d^{-1}$. Since the initial user demand flow on the upper output leg of the first element is still equal to 2.55 $ML.d^{-1}$, it can be deduced that the flow sent to the lower output leg of Element 1 is equal to 5.66 $ML.d^{-1}$.

For the element representing the open reservoir, two changes occur. The first one concerns the initialisation value of the upper output flow which is taken as close as possible to zero, since no water is initially sent back to the first element. Then, the second change is relative to the removal of the auxiliary input flow of the second element. Therefore, a zero value must be added in the corresponding column of the main data table.

4.1.2 Solving strategy

AS described in Chapter 3, the solving strategy is formulated as a Model Predictive Control problem (MPC) which uses a moving-horizon, ahead of present time, as presented in Figure 4.6. For each simulation, the length of the total period of solution N (prediction horizon) is fixed at 48 h. In other words, the solution begins at 0 h and proceeds to 48 h. The model is configured so that the algorithm looks ahead at immediate future demands over a 12 h horizon.

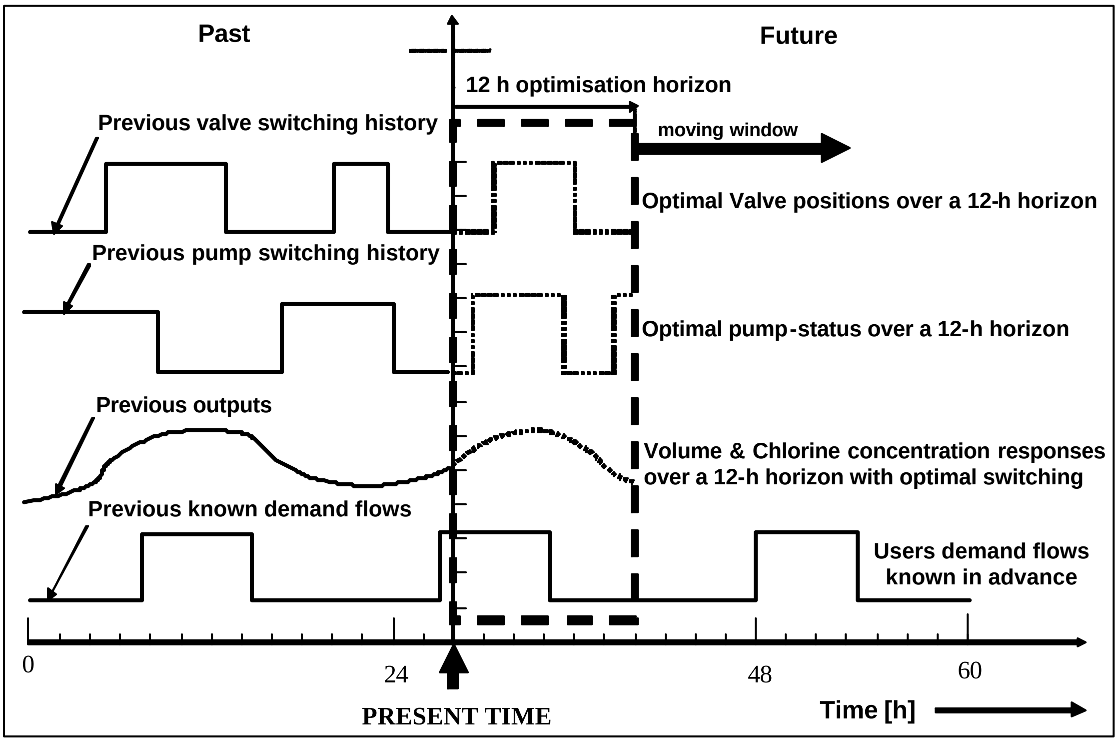


Figure 4.6: Model predictive control over 48 hours with a 12 hour prediction horizon

Thus, the length of the control horizon M is equal to 12 and a total of 60 h of known feed flow data and known user demand flow data has to be provided at the beginning of each new experiment. The sampling period is fixed to equal periods of 1 h.

This parameterisation of the optimisation problem gives a good compromise between the accuracy of the results, the amount of data to be provided and the computational time needed to achieve a result.

4.2 Modelling strategies

THE geometry of the general modelling element developed specifically for the operational optimisation of water distribution networks has been detailed in the above paragraphs, as well as its most general features. These features are common to the two different modelling strategies successively developed in the course of this project. However, even if these strategies share a common background, they are based on modelling concepts that do not use the same physical principles to reach the solution of the optimisation problem. It is obvious that once translated into mathematical terms, these two concepts also use specific sets of equations. Therefore, each approach has necessitated an adaptation of the standard framework which has mostly resulted in the development of a specific set of rules for each of the modelling strategies.

The first modelling strategy is referred to as the “calculated flows” strategy. It is based mainly on friction losses and available heads. The resulting hydraulic model gives a detailed pressure-balanced solution which is used to get a comprehensive representation of the physical phenomena taking place in the network. Although accurate, this modelling approach is rather complex to solve. The second modelling strategy, known as the “available flows” strategy, is much easier to solve, since it is based on a model where the pipe network hydraulics have been simplified.

In this paragraph, the two different modelling strategies are detailed separately, concentrating especially on the differences of concept existing between the two approaches. For each strategy, the repercussions on the modelling framework resulting from the use of different hydraulic models are looked at. The mathematical relations relative to these concepts are also presented.

4.2.1 Calculated flows strategy

THE main feature of the calculated flow strategy is the definition of a complex hydraulic model based on pressure balances.

4.2.1.1 Adaptation of the standard modelling element to the calculated flows strategy

THE standard modelling element used with this strategy is represented Figure 4.7.

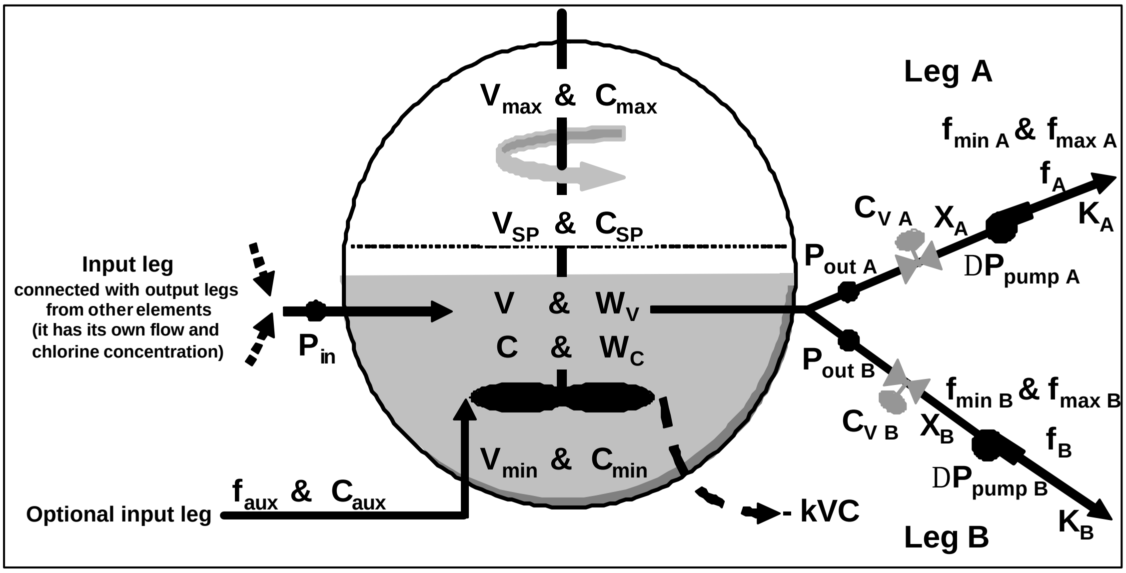


Figure 4.7: Representation of the standard modelling element for the calculated flows strategy

It can be observed that the geometry is the same as the general framework presented earlier on in Figure 4.1.

All of the properties defined in the general data table (Table 4.1) are part of this model. Hence, the mixed-flow vessel has a liquid volume V, a volume setpoint V_{SP}, a chlorine concentration C, and a chlorine concentration setpoint C_{SP}. It is constrained by the low-emergency and high-emergency volumes $V_{\min}$ and $V_{\max}$, as well as by the low-emergency and high-emergency chlorine concentration $C_{\min}$ and $C_{\max}$. A first-order chlorine decay constant k is also defined. The auxiliary input leg is represented by its flow f_{aux} and its chlorine concentration C_{aux}, while for the output legs the upper leg output flow is represented by f_A and the lower leg output flow by f_B. According to the rules defined previously, X_A and X_B are used to represent the initial valve positions or the statuses of pumps, depending on which type of equipment is fitted on the upper and lower output legs.

Besides the parameters which come as standard on the modelling element, additional information needs to be provided for this particular modelling strategy. The necessity to introduce several other parameters is mostly linked to the introduction of a comprehensively pressure-balanced hydraulic model. For instance, to represent the frictional pressure drop in the lines, a valve size coefficient C_V and a line conductance K are required for each output leg. Furthermore, when a pump is fitted on one of the output legs, its contribution towards the pressure levels in the system must be defined. This justifies the definition of two additional parameters which represent the pump head and are referred to as ΔP_{pumpA} and ΔP_{pumpB}. Besides pump heads, the pressure at each node must also be provided to establish an accurate pressure balance. For this modelling strategy, a distinction is made between the receiving pressure of the input node and the two delivery pressures on each of the output legs. Note that the receiving pressure P_{in}, and the delivery pressures P_{outA} and P_{outB} include an hydraulic potential pressure. For instance, a flow leaving a reservoir elevated at a height of 250 m and entering one of the output legs at atmospheric pressure, has a potential pressure of 250 m. This may feed to a line at an atmospheric outflow at an elevation of 150 m, giving an available driving ΔP of 100 m, if no pump is operating in the output line. All of the pressures defined within the “calculated flows” strategy can be considered either variable or constant. Thus, the set of rules developed to distinguish variables from fixed parameters and presented in paragraph 4.1.1.2.d apply also to node pressures. Fixed pressures are usually defined for the receiving and delivery pressures of a storage reservoir, since reservoir elevations are known parameters. Variable pressures are chosen for all of the other configurations possibly represented by a modelling element.

Besides parameters related to the definition of a pressure-balanced hydraulic model, it is also necessary to define minimum and maximum constraints for the output flows leaving the standard modelling element. These new variables are referred to as $f_{\min A}$ and $f_{\max A}$ for the upper output leg, and as $f_{\min B}$ and $f_{\max B}$ for the lower output leg. These values are used to build a set of flows that is representative of the capacities possibly supported by each pipeline present in the distribution system.

It can also be observed that the mixed-flow vessel requires the definition of two additional parameters, referred to as W_V and W_C. These parameters are only defined when the modelling elements represent an open reservoir. They are stored inside two separate matrices, represented by the notations $\overline{W_V}$ and $\overline{W_C}$. These matrices are similar to the time-dependent weighting sequence W_i introduced for the MPC formulation, described in Chapter 3. Therefore, they are also assumed positive and semi-definite and, in this configuration, they are also both diagonal. In the context of operational optimisation of a water distribution network, W_V parameters penalise deviations of the reservoir volumes from their respective setpoints, while W_C parameters have the same function for the chlorine concentrations in the reservoirs. The greater these values are, the tighter the control will be. However, a tight control is always achieved at the expense of the computational time.

All of the additional information required for the calculated flows strategy is also stored in the main data table presented Table 4.1. In other words, in addition to the 19 columns of this table, 13 additional columns are defined. In paragraph 4.1.1.2.b, a distinction has been made between two different categories of data: initialisation values and intrinsic characteristics of the standard modelling element. In this extended table, the new parameters that belong to the family of initialisation values are only represented by the pressures defined for each node. On the contrary, valve size coefficients, pipeline conductance, minimum and maximum flows of the output legs, pump heads, as well as weighting sequences must be considered as additional intrinsic parameters. For each of the new parameters defined specifically for the calculated flows approach, the rules presented in paragraph 4.1.1.2.c about the use of zero values in the main data table still apply.

4.2.1.2 Equations required for the calculated flows strategy

FOR every standard modelling element used to represent a computer version of a water distribution network, several equivalence constraints must apply.

The calculated flows strategy requires the definition of pressure, volume and chlorine concentration balances. Several constraints and an objective function are also defined.

4.2.1.2.a Pressure balance

THE pressure balance states that the frictional pressure drop across an output leg is equal to the available potential pressure existing on the same leg.

The frictional pressure drop ΔP_F across an output pipeline is represented by the sum of the pressure drop ΔP_V across the valve (when applicable) and the pressure drop ΔP_L through the pipeline. Mathematically, it is expressed by:

$$\Delta P_F = \Delta P_V + \Delta P_L \tag{4.6}$$

Let us consider how this expression can be transformed for the upper output leg of any standard modelling element. For this leg, the output flow can be expressed as follows if the analysis is based on the valve:

$$f_A = C_{V\,A}\ X_A\ \sqrt{\Delta P_{V\,A}} \tag{4.7}$$

From equation (4.7), it is possible to deduce: $\Delta P_{V\,A} = \dfrac{f_A{}^2}{C_{VA}{}^2\ X_A{}^2}$ (4.8)

If the analysis is based on the pipeline, another equation can be expressed:

$$f_A = K_A\ \sqrt{\Delta P_{L\,A}} \tag{4.9}$$

This leads to: $\Delta P_{L\,A} = \dfrac{f_A{}^2}{K_A{}^2}$ (4.10)

From (4.8) and (4.10), it is possible to deduce the frictional pressure drop in the upper output leg of any standard modelling element: $\Delta P_{F\,A} = f_A{}^2 \left(\dfrac{1}{C_{V\,A}{}^2\ X_A{}^2} + \dfrac{1}{K_A{}^2} \right)$ (4.11)

Similarly, on the lower output leg: $\Delta P_{F\,B} = f_B{}^2 \left(\dfrac{1}{C_{V\,B}{}^2\ X_B{}^2} + \dfrac{1}{K_B{}^2} \right)$ (4.12)

The available potential pressure on an output leg is represented by the difference of height at its ends (referred to as ΔP_{static}) plus the head of the pump fitting the current pipeline (if applicable). In mathematical terms, and for the upper output leg, it gives:

$$\Delta P_{potentialA} = \Delta P_{staticA} + \Delta P_{pumpA} \tag{4.13}$$

Similarly, on the lower output leg: $\Delta P_{potentialB} = \Delta P_{staticB} + \Delta P_{pumpB}$ (4.14)

Hence, by combining equations (4.11) to (4.14), the pressure balance can be expressed as follows:

- upper output leg:

$$\boxed{f_A{}^2\left(\frac{1}{C_{VA}{}^2 X_A{}^2} + \frac{1}{K_A{}^2}\right) = \Delta P_{staticA} + \Delta P_{pumpA}} \quad (4.15)$$

- lower output leg:

$$\boxed{f_B{}^2\left(\frac{1}{C_{VB}{}^2 X_B{}^2} + \frac{1}{K_B{}^2}\right) = P_{staticB} + \Delta P_{pumpB}} \quad (4.16)$$

4.2.1.2.b Volume balance

THE theoretical volume balance is applied on every standard modelling element, using the conservation principle:

$$Accumulation = Inflow + Generation - Outflow - Consumption \quad (4.17)$$

Since a volume cannot be either generated or consumed, the conservation principle becomes:

$$Accumulation = Inflow - Outflow \quad (4.18)$$

The accumulation term is represented mathematically by the expression $\frac{dV}{dt}$. The inflow of any standard element can be expressed as the sum of the auxiliary input flow of the current element (when applicable) with all the flows f_{in} arriving from other parts of the distribution system and connected to the input of the current modelling element. With the notation defined earlier, the inflow of any modelling element is defined by: $Inflow = f_{aux} + \sum^{nel_{connect}} f_{in}$ (4.19)

The parameter $nel_{connect}$ in equation (4.19) refers to the number of standard modelling elements which are connected to the input of the current modelling element.

To select the inputs that need to be included in the summing term of equation (4.19), use is made of a flow interconnection parameter, represented by the notation m_f. This parameter is equalled to 0 / 1 values only, depending on whether or not a connection is present between two elements.

The values of the flow interconnection parameter are automatically determined by the model thanks to the use of the topology table described in paragraph 4.1.1.3.a. Therefore, considering the standard element i, the expression represented by equation (4.19) can be transformed as follows: $Inflow_i = f_{aux_i} + \sum_{j=1}^{2*nel} m_{f\,ij}\, f_{in\,j}$ (4.20)

The index j used in equation (4.20) allows identifying the output legs of the other elements of the network to which the input of the element i is possibly connected. For instance, $m_{f\,i2}$ represents the value of the flow interconnection parameter for the link between the element i and the lower output leg of the first element of the network. Similarly, $m_{f\,i7}$ represents the value of the interconnection parameter between the element i and the upper leg of the fourth element of the network.

Concerning the output zone of a standard element, it is represented by the sum of both flows leaving the current element via its output legs. It is expressed by:

$$Outflow = f_A + f_B \tag{4.21}$$

The mathematical version of the volume balance is deduced from equations (4.18), (4.19) and (4.21):

$$\frac{dV}{dt} = f_{aux} + \sum_{i=1}^{nel_{connect}} f_{in_i} - f_A - f_B \tag{4.22}$$

When integrated over the period t to $t + \Delta t$, the volume balance is expressed by:

$$V(t+\Delta t)_{|t} = V(t) + \frac{dV}{dt}\Delta t \tag{4.23}$$

The expression of the volume balance used at time t is deduced from equations (4.22) and (4.23). It is presented below:

$$\boxed{V(t+\Delta t)_{|t} = V(t) + \left[f_{aux}(t) + \sum_{i=1}^{nel_{connect}} f_{in_i}(t) - f_A(t) - f_B(t) \right] \Delta t} \tag{4.24}$$

It can be noted that when the modelling element represents a pipe off-take or an under-pressure reservoir, a constant volume has to be specified in the main data table (represented by a positive value). From this, it is deduced that the term $\frac{dV}{dt}$ is equal to 0. Hence, the volume balance becomes: $V(t+\Delta t)_{|t} = V(t)$ (4.25)

4.2.1.2.c Chlorine concentration balance

THE conservation principle represented by equation (4.17) is also valid for the chlorine concentration. For this particular application, it can be transformed as follows:

$$Accumulation = Inflow - Outflow - Consumption \qquad (4.26)$$

To evaluate the chlorine concentration inside the vessel of the element, the inventories VC and the species flows $f_{aux}C_{aux}$, $f_A C$ and $f_B C$ have to be considered. These notations are representative of the property introduced in paragraph 4.1.1.2.b, stating that the chlorine concentration in the vessel is transmitted to both output legs. In other terms, $C_A = C$ and $C_B = C$.

The accumulation is equivalent to the rate of variation of the chlorine inventory and can be represented by the expression $\frac{dVC}{dt}$.

According to what has been presented in the above paragraph for the volume balance, the inflow is represented here by:

$$Inflow = f_{aux}C_{aux} + \sum_{i=1}^{nel_{connect}} (f_{in}C_{in})_i \qquad (4.27)$$

The input concentration C_{in} represents the resulting chlorine concentration that is obtained once all the flows arriving from other parts of the distribution systems have been mixed to constitute the input of the current element. Similarly to f_{in}, C_{in} is not a parameter that requires definition in the main date table, but is determined automatically.

Similarly, the outflow can be expressed by: $Outflow = f_A C + f_B C$ (4.28)

The consumption is represented by the natural chlorine decay inside the vessel. It has been seen before that this decay must be represented by first-order kinetics. Mathematically, this may be expressed by the term: $Consumption = kVC$ (4.29)

In this equation, k represents the first order chlorine-decay constant.

The chlorine concentration balance can then be deduced from equations (4.27) to (4.29):

$$\frac{dVC}{dt} = f_{aux}C_{aux} + \sum_{i=1}^{nel_{connect}} (f_{in}C_{in})_i - f_A C - f_B C - kVC \qquad (4.30)$$

The change in inventory can be approximately expressed using an average gradient term from equation (4.30). This gives: $V(t+\Delta t)C(t+\Delta t) = V(t)C(t) + \left(\frac{dVC}{dt}\right)\Delta t$ (4.31)

Using the arithmetic average of the gradient from equation (4.30), the chlorine concentration balance becomes:

$$\boxed{\begin{aligned} V(t+\Delta t)C(t+\Delta t) = {} & V(t)C(t) \\ & + f_{aux}(t)C_{aux}(t)\Delta t + \sum_{i=1}^{nel_{connect}}\left[f_{in}(t)_i\left(\frac{C_{in}(t+\Delta t)+C_{in}(t)}{2}\right)_i\right]\Delta t \\ & - \left[f_A(t)+f_B(t)\right]\left[\frac{C(t+\Delta t)+C(t)}{2}\right]\Delta t \\ & - k(t)\left[\frac{V(t+\Delta t)C(t+\Delta t)+V(t)C(t)}{2}\right]\Delta t \end{aligned}} \tag{4.32}$$

4.2.1.2.d Constraints

CONSTRAINTS are defined for each leg of the modelling element, on valve positions and on output flows. On the upper output leg, they give:

$$0 \le X_A \le 1 \tag{4.33}$$

$$f_{\min A} \le f_A \le f_{\max A} \tag{4.34}$$

Similarly, on the lower output leg:

$$0 \le X_B \le 1 \tag{4.35}$$

$$f_{\min B} \le f_B \le f_{\max B} \tag{4.36}$$

When the current standard element represents a storage reservoir, it is also necessary to add constraints on the volume and on the chlorine concentration inside the mixed-flow vessel:

$$V_{\min} \le V \le V_{\max} \tag{4.37}$$

$$C_{\min} \le C \le C_{\max} \tag{4.38}$$

4.2.1.2.e Reorganisation of data into vectors and matrices:

FOR convenience, all of the equations of this modelling strategy have been presented until now for a particular standard element isolated from the rest of the system to which it belongs. These equations still apply when several modelling elements are linked together.

However, the equations are more conveniently dealt with in matrix-vector form. In this paragraph, the notations used to represent all of the variables defined earlier as vectors and matrices is introduced. The organisation of the data inside the vectors and matrices is also detailed.

$\overline{C}$ represents the vector of chlorine concentrations in the mixed-flow vessel, while the volumes in each vessel are stored in $\overline{V}$. These vectors are composed of nel rows, since a value is specified for these variables on each modelling element. They are organised as follows:

$$\overline{C} = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_{nel} \end{pmatrix} \qquad \overline{V} = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_{nel} \end{pmatrix}$$

For instance, in the vector $\overline{C}(t)$, $C_1(t)$ represents the chlorine concentration inside the vessel of Element 1, at time t; while $C_{nel}(t)$ corresponds to the concentration inside the vessel of Element nel, at time t.The other vectors of variables organised similarly are: $\overline{f_{aux}}$, $\overline{C_{aux}}$, $\overline{C_{SP}}$, $\overline{C_{\min}}$, $\overline{C_{\max}}$, $\overline{k}$, $\overline{V_{SP}}$, $\overline{V_{\min}}$ and $\overline{V_{\max}}$.

For other variables, the organisation of the data can be slightly different, since some of the parameters defined on the output legs are represented by pairs (one value for each leg). Therefore, some vectors must be constituted with $2 \times nel$ rows. This rule applies for instance to $\overline{X}$, used to represent the vector of valve positions and pump status, or to $\overline{f}$ which is the vector of output flows. These vectors are organised as follows:

$$\overline{X} = \begin{pmatrix} X_{A_1} \\ X_{B_1} \\ X_{A_2} \\ X_{B_2} \\ \vdots \\ X_{A_{nel}} \\ X_{B_{nel}} \end{pmatrix} \qquad \overline{f} = \begin{pmatrix} f_{A_1} \\ f_{B_1} \\ f_{A_2} \\ f_{B_2} \\ \vdots \\ f_{A_{nel}} \\ f_{B_{nel}} \end{pmatrix}$$

Certain other variables are organised similarly. These are: $\overline{Pump_1}$, $\overline{Pump_2}$, $\overline{C_V}$, $\overline{K}$, $\overline{\Delta P_{pump}}$, $\overline{f_{\min}}$ and $\overline{f_{\max}}$.

Finally, the pressures are defined as triplets. Therefore, they need to be defined in a vector consisting of $3 \times nel$ rows. This vector is presented below:

$$\overline{P} = \begin{pmatrix} P_{in_1} \\ P_{out_{A_1}} \\ P_{out_{B_1}} \\ P_{in_2} \\ P_{out_{A_2}} \\ P_{out_{B_2}} \\ \vdots \\ P_{in_{nel}} \\ P_{out_{A_{nel}}} \\ P_{out_{B_{nel}}} \end{pmatrix}$$

The pressure vector is the only one organised in such a way. For instance, in the vector $\overline{P}$, P_{in_1} represents the receiving pressure of Element 1 at time t, $P_{out_{A_2}}$ is the delivery pressure on the upper output leg of Element 2 at time t, and $P_{out_{B_{nel}}}$ corresponds to the delivery pressure on the lower leg of Element nel at time t.

As mentioned earlier, the only variables that are stored inside matrices are the weighing coefficients W_V and W_C. These matrices consist of nel rows and of M columns. They are diagonal and filled as follows:

$$\overline{W_V} = \begin{pmatrix} W_{V_{11}} & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & W_{V_{nelM}} \end{pmatrix} \qquad \overline{W_C} = \begin{pmatrix} W_{C_{11}} & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & W_{C_{nelM}} \end{pmatrix}$$

This process of organisation of the data is done automatically by the software developed to implement this model. The constitution of the vectors and matrices is based on the data provided in all of the data tables described earlier in paragraph 4.1.1.

4.2.1.2.f Objective function

THE objective function relative to this particular optimisation problem is defined according to the work presented about MPC techniques in Chapter 3.

In the particular example of the operational optimisation of water distribution networks, the move suppression factor is replaced by an economical term associated with the cost of operating fixed-speed pumps during pumping operations.

The objective function is represented by equation (4.39), where there are $k = 1, \ldots, N$ to the optimisation horizon. The expression represented below has to be minimised.

$$f = \sum_{k=1}^{N} \Big[\left(\overline{V} - \overline{V_{SP}}\right)_k^T \overline{W_V} \left(\overline{V} - \overline{V_{SP}}\right)_k + \left(\overline{C} - \overline{C_{SP}}\right)_k^T \overline{W_C} \left(\overline{C} - \overline{C_{SP}}\right)_k + \left(\overline{f}_k^{\,T}\, \overline{Pump_1} + \overline{Pump_2}\right) \overline{E_{cost}} \Big] \tag{4.39}$$

In this equation, the term $\left(\overline{V} - \overline{V_{SP}}\right)_k^T \overline{W_V} \left(\overline{V} - \overline{V_{SP}}\right)_k$ corresponds to the weighted sum of squared deviations of reservoir volumes from their setpoints, for the time-step k of the present solution. Similarly, the term $\left(\overline{C} - \overline{C_{SP}}\right)_k^T \overline{W_C} \left(\overline{C} - \overline{C_{SP}}\right)_k$ penalises chlorine concentration deviations from setpoints.

The term $\left(\overline{f}_k^{\,T}\, \overline{Pump_1} + \overline{Pump_2}\right) \overline{E_{cost}}$ is an economic factor which represents the money spent over the simulation period for pumping operations. In this expression, $\overline{E_{cost}}$ is the cost of unit pump power (electrical cost), while $\overline{f}_k^{\,T}\, \overline{Pump_1} + \overline{Pump_2}$ represents the pump characteristics that give a representation of the pump power against the unit flow going through the pump.

4.2.1.3 Shortcomings of the calculated flows strategy

THE calculated flows strategy is based on a fundamental modelling approach which gives a comprehensive hydraulic model. According to Henson (1998), a potential disadvantage of the fundamental modelling is the complexity of the resulting dynamic model which can lead to the inadequacy of the model for the design of a non-linear MPC controller.

It has appeared on the first tests conducted on the model based on the calculated flows strategies that Henson’s remark proved right for configurations requiring more than five standard modelling elements. Unfortunately, in the particular case of the operational optimisation of water distribution networks, the number of elements to include within the same model is generally greater than this.

Therefore, this shortcoming has motivated the development of a simplified modelling technique, still largely constrained as before, but which provides a dynamic model with significantly fewer equations. This modelling strategy is referred to as the “available flows” strategy.

4.2.2 Available flows strategy

THE main feature of the available flows strategy is the definition of an hydraulic model which simplifies the pipe network hydraulics by removing all of the pressure balances. All of the network components are still represented using a standard modelling element, but the flows in these elements are determined in a linear way, using a valve position or a pump status. Essentially, the system hydraulics are solved using a mass balance approach without taking into account the pipe frictions. This new approach removes all of the problems of non-linearity linked to the use of pressure balances in the equations and simplifies the amount of data to be provided in the model. As a result, the definitions of receiving and delivery pressures, pipeline conductance and valve coefficients are no longer required.

4.2.2.1 Adaptation of the standard modelling element to the available flows strategy

THE standard modelling element on which this strategy is based is presented on Figure 4.8. It can be observed that the geometry is the same as the general framework presented earlier in Figure 4.1. All of the properties defined in the general data table (Table 4.1) are also part of this model. Hence, the mixed-flow vessel has a liquid volume V, a volume setpoint V_{SP}, a chlorine concentration C, and a chlorine concentration setpoint C_{SP}.

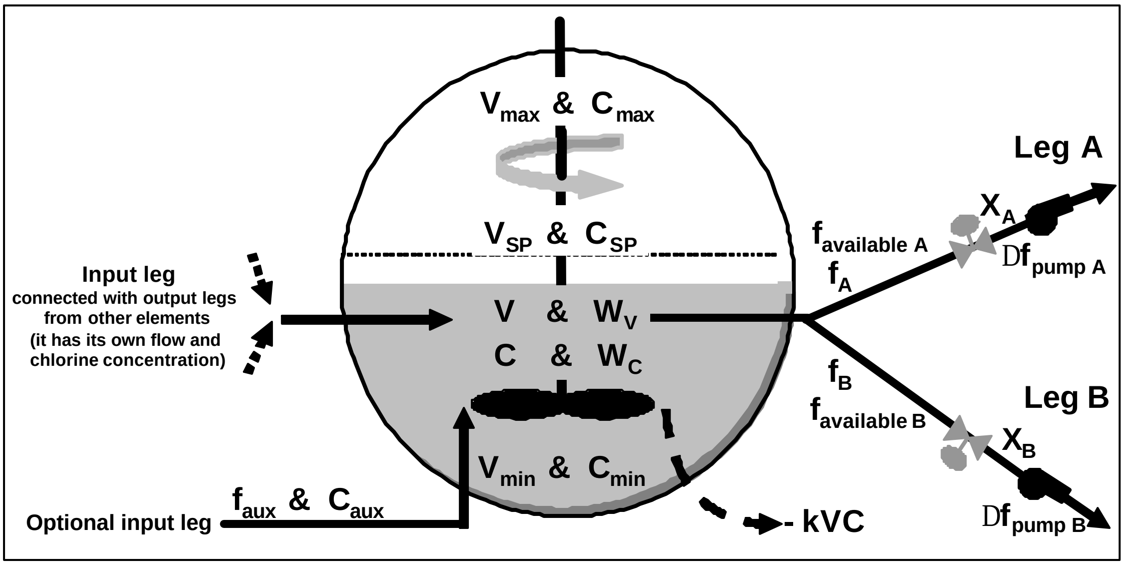


Figure 4.8: Representation of the standard modelling element for the available flows strategy

It is constrained by the low-emergency and high-emergency volumes $V_{\min}$ and $V_{\max}$, as well as by the low-emergency and high-emergency chlorine concentration $C_{\min}$ and $C_{\max}$. A first-order chlorine decay constant k is also defined. The auxiliary input leg is represented by its flow f_{aux} and its chlorine concentration C_{aux}, while for the output legs the upper output flow is represented by f_A and the lower output flow by f_B. According to the rules defined previously, X_A and X_B are used to represent the initial valve position or pump status, depending on which type of equipment is fitted on the upper and lower output legs.

Besides the parameters which come as standard on the modelling element, additional information needs to be provided to take into account the special features of the available flows strategy. The necessity to introduce several other parameters is mostly linked to the introduction of a conceptual notion for the calculation of flows in the output pipelines. When using this strategy, a maximum available flow, referred to as $f_{availableA}$ or $f_{availableB}$ depending on the line considered, is defined for each line. The current output flows are then calculated at each new time-step, up to the prediction horizon, as the product of the maximum available flow of the current pipeline and the actual valve position in this particular line. In this model, the maximum available flows are equivalent to the flows provided by each output leg, once the valve fitted on the leg is fully opened. If there is an online pump in the same pipeline, its contribution is translated into an additional flow, added to the initial maximum available flow. The pump contribution is represented by Δf_{pumpA} and Δf_{pumpB} depending on which of the output legs is considered.

Besides parameters that refer to the conceptual notion of maximum available flows, it can also be observed that the mixed-flow vessel requires the definition of two additional parameters referred to as W_V and W_C. They have exactly the same function as those required for the calculated flows strategy and already presented paragraph 4.2.1.1.

All of the additional information required for the calculated flows strategy is also stored in the main data table presented as Table 4.1. In other words, in addition to the 19 columns of this table, six additional columns are defined. As expected, the amount of data to provide is significantly reduced in comparison with the calculated flows strategy. In paragraph 4.1.1.2.b, a distinction has been made between two different categories of data: initialisation values and intrinsic characteristics of the standard modelling element. In this extended table, no parameter belongs to the family of initialisation values.

On the contrary, maximum available flows, flow contributions due to pumps, as well as weighting sequences must be considered as additional intrinsic parameters.

4.2.2.2 Equations required for the available flows strategy

FOR every standard modelling element used to represent a computer version of a water distribution network, several equivalence constraints must apply. The available flows strategy requires the introduction of rules for the calculation of output flows, as well as the definition of volume and chlorine concentration balances. Several constraints and an objective function are also defined.

4.2.2.2.a Output flows calculation

CALCULATIONS of output flows are obtained by multiplying the current valve position for the considered output leg by the maximum available flow defined for that pipeline. However, the balance equations differ slightly depending on the valve or pump configuration chosen for the current output leg.

For instance, when a fixed or a binary valve is defined on the upper output leg of the current modelling element, the balance equation can be written as follows:

$$f_A \leq X_A \, f_{availableA} \tag{4.40}$$

Similarly, for any lower output leg: $f_B \leq X_B \, f_{availableB}$ (4.41)

It is remembered that maximum available flows are fixed parameters corresponding to maximum possible flows when valves are fully opened. These are obtained in practice from reservoir volume derivatives. Thus, in these two particular cases, to allow for upstream restrictions, the calculated output flows must be lower than or equal to the product between the current valve position and the maximum available flow defined for the pipeline being considered. With user demand flows changing continuously in the network, the output flows of any standard modelling element must be adapted at each new time-step, and up to the prediction horizon. Therefore, they are not always necessary equal to their maximum possible values.

When a pipeline is fitted with an inline pump, the flow contribution of the pump must be added to the maximum available flow defined with the pump offline.

In the particular configuration of a pump fitting the upper output leg of the current modelling element, the output flow equations are transformed in the following way:

$$f_A = f_{availableA} + \Delta f_{pumpA} \tag{4.42}$$

Similarly, on the lower output leg: $f_B = f_{availableB} + \Delta f_{pumpB}$ (4.43)

Finally, when the definition of a fixed output flow is required to represent a user demand flow for instance, the valve position does not have any importance. If a user demand flow is defined on the upper output leg of the current standard modelling element, the output flow equation is expressed as follows: $f_A = f_{demand\ A}$ (4.44)

Similarly, for a user demand flow defined on the lower output leg: $f_B = f_{demand\ B}$ (4.45)

$f_{demand\ A}$ and $f_{demandB}$ represent set user demand flows linked to a reticulation system.

4.2.2.2.b Volume balance

THE volume balance used in the available flows strategy is the same as the one presented for the calculated flows strategy (equation (4.24)). The reader is invited to refer to paragraph 4.2.1.2.b for more information.

4.2.2.2.c Chlorine concentration balance

THE chlorine concentration balance used in the available flows strategy is the same as the one presented for the calculated flows strategy (equation (4.32)). The reader is invited to refer to paragraph 4.2.1.2.c for additional information.

4.2.2.2.d Constraints

BESIDES the suppression of all of the constraints relative to output flows (equations (4.34) and (4.36)), all constraints defined in paragraph 4.2.1.2.d for the calculated flows strategy still apply for the available flows strategy.

4.2.2.2.e Reorganisation of data into vectors and matrices:

ALL of the rules defined on this particular subject for the calculated flows strategy are still valid for the available flows strategy. All the new parameters required especially for the available flows strategy belong to the family of vectors constituted by $2 \times nel$ rows.

Two new vectors, referred to as $\overline{f_{available}}$ and $\overline{\Delta f_{pump}}$, are defined. The way these vectors are organised can be guessed from information presented in paragraph 4.2.1.2.e.

4.2.2.2.f Objective function

THE objective function used in the available flows strategy is the same as the one presented for the calculated flows strategy (equation (4.39)). The reader is invited to refer to paragraph 4.2.1.2.f for additional information.

4.2.2.3 Shortcoming of the available flows strategy

IT is obvious that the available flows strategy, while simpler and easier to solve, does not provide an accurate model of the complex hydraulic behaviour of a water distribution system. Such a simplified model can only be relied upon if validated using a full hydraulic model of the system. This explains why it is recommended to combine the use of the modelling strategy presented in this paragraph with the public domain software package ePANET (Rossman, 2000). ePANET is widely accepted as the world standard in hydraulic and water quality modelling of water distribution systems. However, ePANET cannot be used alone to meet the objectives of this project, since it is not an optimisation tool. Most of the existing water distribution systems have been modelled within ePANET which makes the combination of both models easier. In this approach, the ePANET version of the water distribution to be optimised is used to determine accurately the values of the maximum available flows appearing in the new version of the model presented here. It is considered that the values extracted from the ePANET file represent the actual situation in the targeted application. Consequently, no other calibration of the model based on the available flows strategy is required.

4.3 Concluding remarks to the chapter

IN this chapter, the modelling developed to respond to the problem of operational optimisation of water distribution systems has been presented. The model that was developed takes the shape of a standard modelling element which is repeated as often as necessary to build the complete target application. Two different modelling strategies have been developed successively during the course of this research project, since the results of the first one would not have been satisfactory on a large-scale application, due to the extensive computational resources it requires. All of the equations relative to the theory developed Chapter 3 have been presented for both modelling strategies.

The next chapter presents the different tools that need to be used to achieve the implementation of these strategies within the selected modelling software.

CHAPTER 5 – TOOLS INVOLVED IN THE MATHEMATICAL IMPLEMENTATION OF THE MODELLING

THE non-linear Model Predictive Control (MPC) optimisation problem related to the operational optimisation of a water distribution network requires the calculation of the optimised sequence of manipulated input variables. As presented Chapter 4, this sequence is obtained using a non-linear discrete model. The standard framework on which this model is based has been constructed within the MATLAB software package to interpret mathematically all of the features of the model presented in the previous chapter.

Besides the prediction of the states of some of the variables, this optimisation problem requires simultaneously the solution of a non-linear programming (NLP) problem. The NLP problem is usually non-convex because the equations appearing in the model are non-linear. However it can also be convex with non-linear equations. These families of problems are generally solved using powerful algorithms based on mathematical programming techniques. Consequently, the success of non-linear MPC techniques applied to the operational optimisation of a water distribution network is also partially linked to the widespread availability of efficient and robust solution techniques for these types of NLP optimisation problems. When non-linear MPC techniques are applied to operational optimisation problems where discrete choices have to be made, the determination of the optimised sequence of manipulated input variables can be achieved using the MINLP algorithm (e.g. Pahor and Kravanja, 1995; or Zamora and Grossman, 1998). With the development of the Augmented Penalty / Outer Approximation / Equality Relaxation (AP / OA / ER) algorithm presented in Chapter 3 (associated with recent advances in Non-Linear Programming (NLP) solvers and Mixed-Integer Linear Programming (MILP) solvers), MINLP problems can be solved more efficiently. MINLP algorithms are usually readily available for use in powerful optimisation packages. One of the most popular optimisation packages where MINLP algorithms are included is the General Algebraic Modelling System (GAMS).

GAMS is a modelling language which is used to realise an interface between the non-linear discrete model proposed earlier and the MINLP algorithm, needed to solve the non-linear Model Predictive Control (MPC) optimisation problem. Unlike MATLAB and Excel which can only handle small-scale non-linear models, the GAMS software has been developed to solve large non-linear applications. However, although GAMS is extremely powerful at providing an interface between a model and the solver used to solve the optimisation problem, it is not the ideal tool for the essential phase of data capture.

This shortcoming explains why MATLAB has been preferred to GAMS for the creation of the general modelling framework used during the phases of data capture and data manipulation. The MATLAB-based program is only developed for the capture and the setting-up of the data required by the model. The phase of manipulation of the data consists of transforming them into a form suitable for the optimisation package. Once this is done, the data are written into a GAMS file. This transfer is achieved with MATLAB, using its ability to write instructions in secondary files. Besides the transfer of model data under a form suitable to GAMS, MATLAB is also used to transform the mathematical relations that are at the core of the model, into a form easily interpreted by GAMS. The GAMS file is then run separately within the optimisation package. Once the MINLP optimisation has converged, GAMS saves the optimised results in a result file. Finally, this file is accessed in Excel for further exploitation of the results. Figure 5.1 gives an overview of the optimisation procedure used to achieve the operational optimisation of a water distribution system by, using the model presented in Chapter 4.

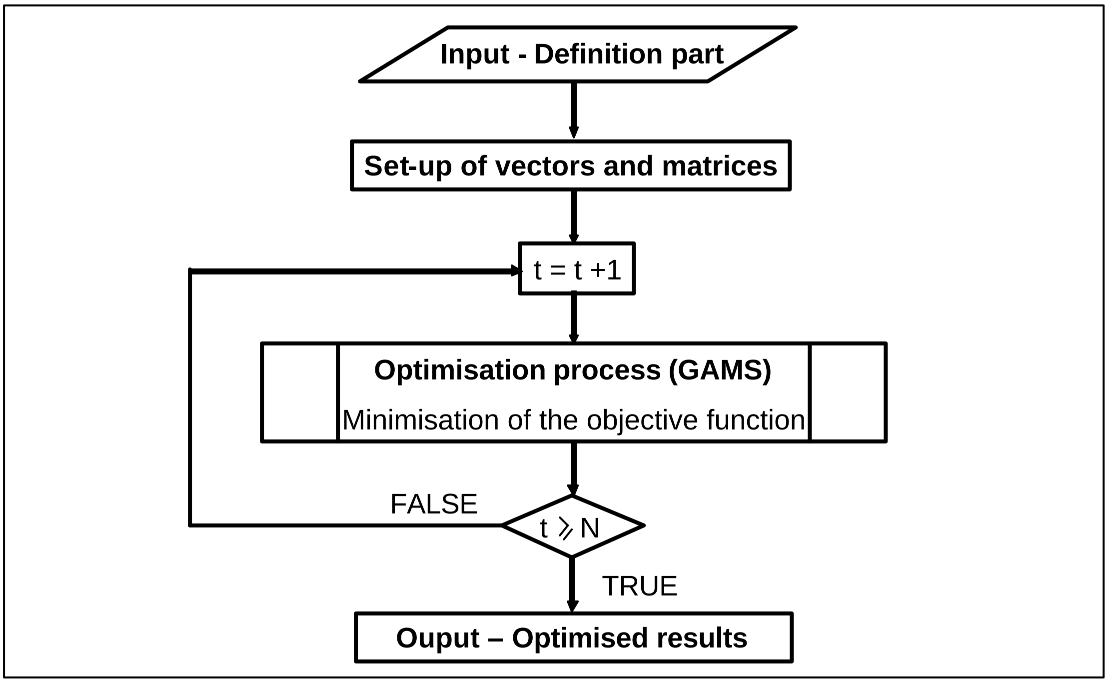


Figure 5.1: Organisation chart of the optimisation process

5.1 MATLAB-based programming

FROM the user’s point of view, the efficiency of an algorithm or the quality of the programming is not necessarily the critical factor in determining the usefulness of the code. The ease with which it can be interfaced to a particular application is often more important. The most suitable interface depends strongly on the particular application and on the context in which it is solved. For instance, for users that are accustomed to a spreadsheet interface, a code that can accept input from this source presents a big advantage over others.

On the contrary, for users used to handling modelling codes that set up and solve optimisation problems by means of subroutine calls, the substitution of a more efficient package that uses more or less the same subroutine interface may be the best option.

Since MATLAB is a software package which is very popular due to its user-friendly interface, its use has been preferred instead of GAMS for all of the functions for which a user intervention is required. An interface between the two modelling packages has been developed so that users without any knowledge on GAMS can nevertheless build a water distribution network within the model and achieve its operational optimisation.

5.1.1 Main features of the MATLAB software package

MATLAB which stands for MATrix LABoratory, is an interactive software package for numerical computations and graphics. It has been edited by The Mathworks since 1984 and, since it was first released more than 15 years ago, it has been adopted as a standard for technical computing by major companies and research laboratories. An estimated 400000 engineers use this software regularly worldwide. The latest version that has been released is known as MATLAB 6.5 Release 13.

MATLAB is one of the fastest and most enjoyable ways to solve problems numerically. As the name suggests, MATLAB is especially designed for matrix computations. It can be used in place of other languages like C and C++, with equal performance but less programming. For instance, the MATLAB language enables the easy manipulation of scalars, vectors and matrices. In addition, the software has a variety of powerful graphical capabilities.

MATLAB can be accessed either by entering commands directly after the prompt (in that case it is used interactively) or by defining scripts (in which case, the information is collected in separate programs known as “M-files”, nothing more than a text file with the extension “.m”). Those scripts can be programs or functions with special parameters. The functions are very convenient since every user can extend the capabilities of MATLAB to his own domain of application. Because the syntax for using MATLAB interactively is the same for writing programs, a code can be quickly converted into a reusable, automated analysis routine. Unlike most traditional languages, MATLAB gives the freedom to focus on technical concepts rather than on programming details like memory management and variable declarations. Furthermore, M-files require no compiling or linking which allows editing and step-by-step debugging of a program without having to leave MATLAB.

MATLAB provides the standard constructs, such as conditionals and loops, to reduce the amount of repetitive tasks. For instance, the software has a standard IF-ELSEIF-ELSE conditional, while the logical operators are expressed by < (less than), > (greater than), <= (less than or equal to), >= (greater than or equal to), == (logical equal), and ~= (not equal). The logical operators are considered as binary variables and they only return the values 0 and 1. In MATLAB, two types of loops can be provided: a FOR-loop (comparable to a FORTRAN DO-loop or a C-DO loop), and a WHILE-loop. A FOR-loop repeats the statements in the loop as the loop index takes on the value in a given row vector. The WHILE-loop repeats as long as the expression positioned after the WHILE instruction is true (non-zero).

The MATLAB technical computing environment is presented Figure 5.2.

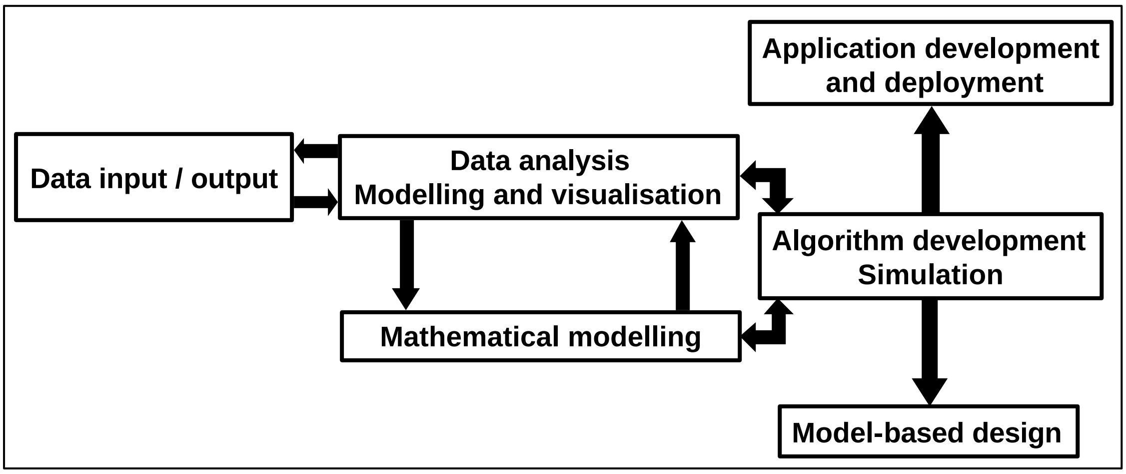


Figure 5.2: MATLAB technical computing environment (courtesy of The Mathworks)

By minimising human time, MATLAB is particularly useful in the initial investigation of real problems, even though they may eventually have to be solved using more computationally efficient ways. This functionality of the software is represented in Figure 5.2 by the path linking the data input box with the model-based design box. This is the path that has been followed for this particular project, since MATLAB has been used as an interface to the powerful optimisation package GAMS.

The modelling developed for the operational optimisation of a water distribution network has already been introduced in Chapter 4, as well as the different types of data that need to be provided. Since preliminary manipulation needs to be done on these data before they can be sent for optimisation, MATLAB appeared to be the best software for this task.

5.1.2 Data pre-processing by MATLAB

THE MATLAB pre-processing program is stored on the CD attached to this thesis. The modelling strategies have been created within separate files which are stored at the following location:

\Programming\MATLAB programming

Most of the tasks allocated to this software are common to both strategies. These tasks are presented on the organisation chart below (Figure 5.3). This chart presents the sequencing followed within the M-file.

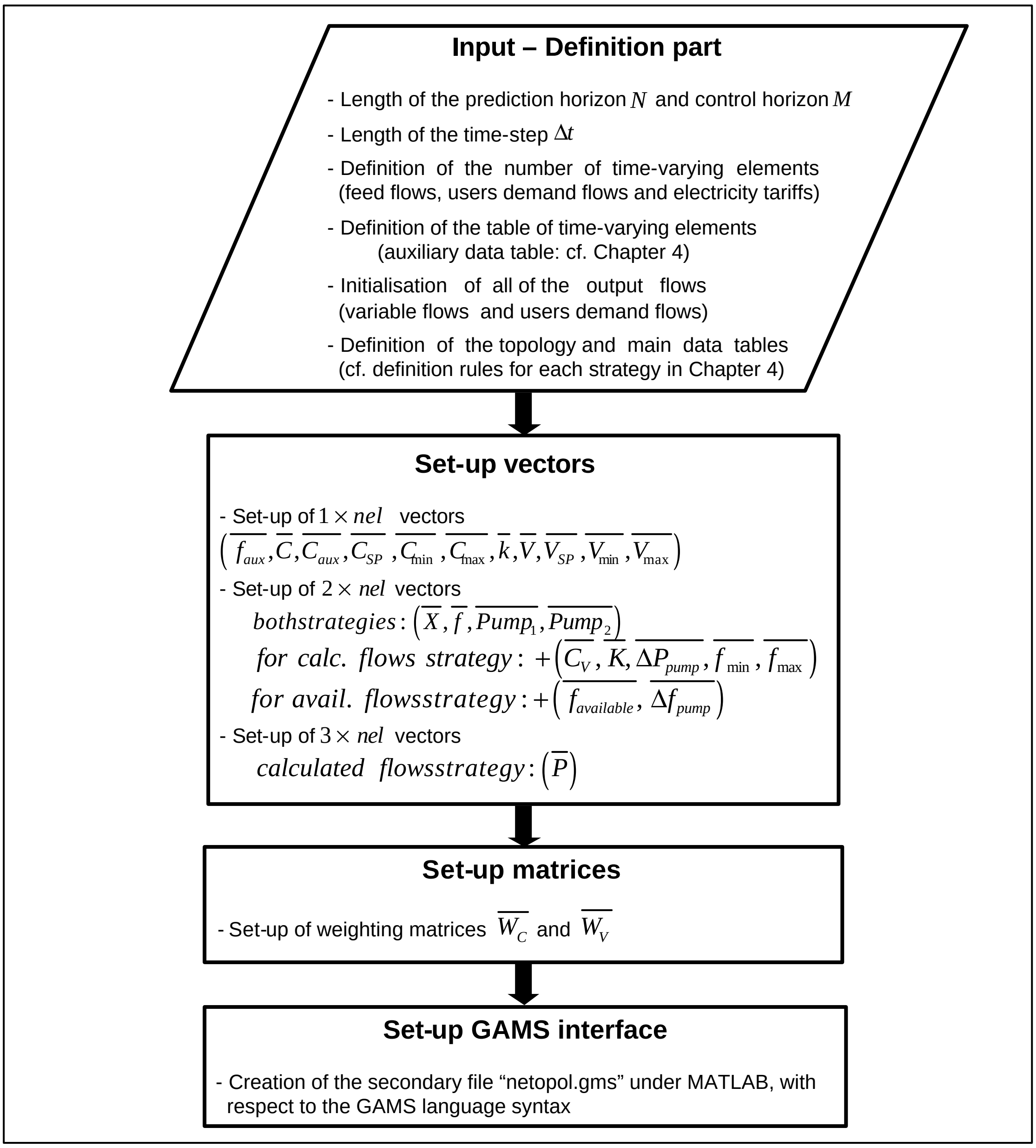


Figure 5.3: Organisation chart of the MATLAB-based program

The MATLAB program begins with the definition of the water distribution within the standard modelling element presented in Chapter 4. Essential parameters like the lengths of the prediction period, of the control horizon and of a sampling interval are defined first, since they determine the exact number of data points to provide. Then, all of the data tables noted in Chapter 4 are constructed. Within MATLAB, these tables are represented by matrices. For convenience, the auxiliary data table is combined with the table of electricity tariffs in a single matrix; while, similarly, the topology table is combined with the main data table.

The capture of data within the different tables is realised according to the rules that have been defined earlier. The amount of data to be provided is also a function of the modelling strategy that has been chosen to achieve the operational optimisation of the water distribution network being considered. Besides the topology and some others parameters that need to be fixed, most of the data captured in the M-file represent initialisation states for the optimisation problem. It is important to understand that MATLAB is not used for its optimisation abilities, but rather to give a picture, at a given instant, of the water distribution system being considered. In other words, MATLAB has no responsibility in the parameter adjustment that needs to take place to reach the optimised state of the supply system. The time for which the initial set of data captured in MATLAB is available is taken as the origin of the prediction horizon ($t = 0$).

The next steps of the sequencing process presented in Figure 5.3 consist in setting-up the initialisation data into vectors and matrices, whose structures are suitable for later use in the optimisation package GAMS. For more information, on the way the rearrangement of the data is achieved, the reader is invited to refer to what has been presented on the subject in Chapter 4.

Finally, the last step of the sequencing achieved by MATLAB consists in the creation of a secondary file, further exploited by GAMS. This file, entitled “netopol.gms”, and interpretable by the optimisation package, contains a representation of the initial state of the water distribution system, as well as the equations of the model which will allow it to move from the initial state to the optimised state. This secondary file is developed using the ability of MATLAB to write formatted instructions into a separate file.

It is interesting to note that the sequencing of the MATLAB program is the same for both modelling approaches described in Chapter 4: the only one difference resides in the number of parameters to be provided initially to the model. A selection is also operated on the equations of the model, since they differ between the two modelling approaches.

5.2 GAMS programming

THE above paragraph has presented the role given to MATLAB in the general process of optimisation. This role is passive, as no implementation against the time takes place within the mathematical software. The implementation as in real-time is rather effected within GAMS which conducts a new optimisation at each new time-step, until the prediction horizon has been reached.

GAMS has been preferred to MATLAB for this particular phase, since GAMS has been developed especially to create an interface between a model and the powerful codes that have appeared on the market to solve large-scale optimisation problems.

5.2.1 Presentation of the General Algebraic Modelling System

THE work presented in this paragraph is mainly based on the GAMS user's guide (Brooke *et al.*, 1998), as well as on the paper prepared by Phimister (2000).

5.2.1.1 Introduction to the modelling system

AS presented in Chapter 3, substantial progress was made in the 1950's and 1960's with the development of algorithms and computer codes to solve large mathematical programming problems. However, the number of applications of these tools in the 1970's was less than expected because the solution procedures formed only a small part of the overall modelling effort. A large part of the time required to develop a model involved data preparation and transformation. Each model required many hours of analysis and programming to organize the data and write the programs that would transform the data into the form required by the mathematical programming optimisers. Furthermore, it was difficult to detect and eliminate errors because the programs that performed the data operations were only accessible to the specialist who wrote them and not to the analysts in charge of the project. The General Algebraic Modelling System (GAMS) was developed to improve on this situation by:

- providing a new language for the representation of large and complex models,
- allowing changes to be made in model specifications simply and safely,
- allowing unambiguous statements of algebraic relationships,
- permitting model descriptions that are independent of solution algorithms.

GAMS is a programming language developed by the GAMS Development Corporation.

It provides a flexible framework used for formulating and solving linear, non-linear and mixed-integer optimisation problems. The system is particularly useful with large and complex problems and can handle dynamic models involving time sequences, lags or leads. Among other functionalities, its syntax allows the declaration of variables, constants and constraints using sets. Through this syntax, input files are written compactly and similarly to the typical formulations of optimisation problems.

In addition, GAMS provides a large number of solvers to optimise a variety of problem formulations, including linear programs, non-linear programs, mixed-integer linear programs and mixed-integer non-linear programs. GAMS is available for use on personal computers, workstations and super-computers: the latest version that has been released is known as GAMS IDE 2.5.

5.2.1.2 Basic principles of the modelling language

THE design of the initial version of GAMS has been based on several principles, derived from ideas taken from the field of database analysis and mathematical programming. The following principles were used in designing the system:

- the access to all existing algorithmic methods should be available to solve the optimisation problem, without changing the user's model representation;
- the expression of the optimisation problem should be independent of the data it uses. This separation of logic and data allows a problem to be increased in size without causing an increase in the complexity of the representation;
- the allocation of computer resources should be automated. This means that large and complex models can be constructed without the user having to worry about details such as arrays sizes and scratch storage.

The GAMS model representation is done in a form that can be easily read by people and computers. This means the GAMS program itself is the documentation of the model and that the separate description required in the past is no longer needed. In addition, a GAMS model representation is concise, and all information needed to understand the model is stored in one document. Explanatory text can be made part of the definition of all symbols and is reproduced whenever associated values are displayed.

Consequently, GAMS lets the user concentrate on the modelling. By eliminating the need to think about purely technical machine-specific problems such as address calculations, storage assignments, subroutine linkage or input-output control, GAMS increases the time available for conceptualising the model, running it, and analysing the results.

Of course, some discipline is needed to take full advantage of these design features, but the aim is to make models more accessible, more understandable, more verifiable and hence more credible.

Finally, the GAMS language is similar to commonly used programming languages. Therefore, it is familiar to anyone with programming experience.

5.2.1.3 Presentation of the basic features of the modelling language

IN its simplest form, GAMS operates on a user-supplied input file, normally denoted with a ".gms" extension to the filename. This file is used to encode the mathematical formulation of the optimisation problem being examined. Selection of the word processor for use in editing the input file is left to the discretion of the user. In this project, MATLAB has been chosen for this particular task, as presented earlier.

GAMS programs consist of one or more statements (sentences) that define data structures, initial values, data modifications and symbolic relationships (equations). While there is no fixed order in which statements have to be arranged, the order in which data modifications are realised is important. Symbols must be declared before they are used and must have values assigned before they can be referenced in assignment statements.

GAMS operates in two stages on program files. Thus, the process of optimisation always begins by a phase of compilation. This stage ensures that the input file is understandable to GAMS. The compiler checks for error in the input file to ensure that the file respects a specific layout, does not contain syntax errors, and uses an appropriate solver. However, the compiler does not solve the problem or indicate that a solution exists. When the compiler locates errors in the input file, they are highlighted and written in the output file before GAMS terminates. The user must modify the input file accordingly before GAMS proceeds to the next phase. Once the input file is readable, GAMS achieves the second phase, referred to as the execution process. During this phase, GAMS carries out the optimisation using an appropriate solver for the problem formulation: Linear Programming (LP), Non-Linear Programming (NLP), Mixed-Integer Linear Programming (MILP), or Mixed-Integer Non-Linear Programming (MINLP). The solver declared by the user must be applicable to the formulation and, for instance, a LP solver cannot be used to solve a NLP problem. At the end of the execution process, GAMS writes to the output file, and provides information about the quality of the solution. If an optimised solution has been reached, GAMS gives the solution values.

The structure of a GAMS input file must be defined in a particular way. In its simplest form, an input file must consist of four successive steps:

a) declaration of variables;
b) declaration of equations;
c) definition of equations;
d) declaration of the model, with an appropriate solve statement.

It is recommended that all statements are ended with a semi-colon, as statements without semi-colons may cause compiler errors.

5.2.1.3.a Declaration of variables

EACH variable must be declared in the input file. It is important to note that a variable for the objective function must also be declared.

GAMS has also a powerful feature that allows the definition of variable into sets. Sets allow the grouping of variables of the same kind. They can be used in variable and constant declarations, as well as equation definitions. Sets are fundamental building blocks in any GAMS model and permit the model to be succinctly stated and easily read. The sets are defined by the following statement: $\mathrm{SET}\ \ SET\_NAME\ \ /\ \ SET\_VARIABLES\ \ /$. In the present model, the main function of sets is to define the index ranges nel, $2 \times nel$ and $3 \times nel$ for arrays containing variables such as reservoir volumes, chlorine concentrations and flows.

5.2.1.3.b Declaration of equations

EACH constraint as well as the objective function must be defined with a name. For every declared name, a corresponding equation must be defined. Similarly to variables, when use of sets is made with equations, they allow equations to be declared more concisely. For instance, rather than declaring five equations relative to the same property and defining each of them, a single set can be declared and defined to encompass all of the five equations

5.2.1.3.c Definition of equations

TO define an equation, its name is restated. Then it is followed by two dots and at least one space. The equation is then stated using the declared variables and constants, mathematical operators and GAMS integrated function (sine, cosine, sum, etc…). Since each equation is defined as a statement, it is necessary to terminate each with a semi-colon.

Relational operators accessible in GAMS are defined below in Table 5.1.

Table 5.1: Different relational operators accessible in GAMS

Relation	**Syntax**
Equality constraint (=)	=E=
Less than or equal to (=)	=L=
Greater than or equal to (=)	=G=

There is no definition for the strict inequalities less than (<) or greater than (>). This omission is intentional and does not result in any loss of generality by GAMS. Similarly, the most commonly used operators are presented in Table 5.2.

Table 5.2: Commonly used operators in GAMS

Operator	**Syntax**
Addition	+
Subtraction	-
Multiplication	*
Division	/
Exponent (to the power of)	**

5.2.1.3.d Declaration of the model, with an appropriate solve statement

THE model declaration statement defines a name for the model and declares the equations that have to be included in the model. For initial GAMS users, it is recommended to include all of the equations that have been defined. The most general way to express the model declaration statement is presented below:

MODEL *MODEL_NAME* / ALL / ;

The model declaration statement is generally followed by the solve statement. The solve statement defines the model to be optimised. It refers to the one defined previously in the model declaration statement. The solve statement also specifies the type of solving procedure (LP, NLP, MILP or MINLP), the type of optimisation (minimisation or maximisation) and the variable that need to be optimised (objective function). The most general way to express the solve statement is presented below:

SOLVE *MODEL_NAME* USING *PROBLEM_TYPE* MINIMIZING *OBJ_FUNCT_VARIABLE*
or MAXIMIZING *OBJ_FUNCT_VARIABLE*

The key words SOLVE and USING must be present in the solve statement, as well as either MINIMIZING or MAXIMIZING .

The common solved formulations are presented in Table 5.3.

Table 5.3: Most common optimisation formulations available in GAMS

Formulation	**GAMS syntax**
Linear programming	LP
Mixed-integer linear programming	MIP
Non-linear programming	NLP
Mixed-integer non-linear programming	MINLP

5.2.1.4 Presentation of important expanded features of the modelling language

LARGE optimisation problems can be solved using the basic features of GAMS presented above. However, several other features greatly improve the formulation and readability of the input and output files.

5.2.1.4.a Syntax

THE GAMS compiler does not distinguish between uppercase and lowercase characters. Hence, to allow for texts that are more readable, both cases may be used in an input file.

5.2.1.4.b Documentation

IT is important to add documentation to the input file to simplify debugging, as well as to clarify the formulation of the optimisation problem. Documentation can be used throughout the input file by placing an asterisk (*) in the first column of a line.

5.2.1.4.c Bounds and initialisation of variables

IT is generally a good thing to bound variables and to provide good initial conditions, particularly for NLP problems. It is indeed often difficult for a non-linear solver to locate a feasible solution which satisfies all of the constraints, especially when the initial guessed values are poor. This situation can be improved by the user, who can locate where a solution cannot exist, and hence provide tighter bounds on the search space. Furthermore, when the user provides a feasible starting point, it is more likely that the solver will achieve an optimal solution.

It is also possible to write input files that are more concise by redeclaring some of the variables with options integrated directly into GAMS. The key words to redefine bounds on a variable are presented Table 5.4.

Table 5.4: GAMS integrated bounds that can be defined on variables

GAMS Syntax	Range on Variable
FREE (default)	-8 to +8
POSITIVE	0 to +8
NEGATIVE	-8 to 0
BINARY	0 or 1 only
INTEGER	0,1, … , 100

Bounding and initialisation statements appear after the variable declaration statements, but before the equation declaration statements.

5.2.1.5 Possible ways of organisation of a GAMS input file

THERE are two common ways of organising GAMS programs. They are shown on Figure 5.4 below.

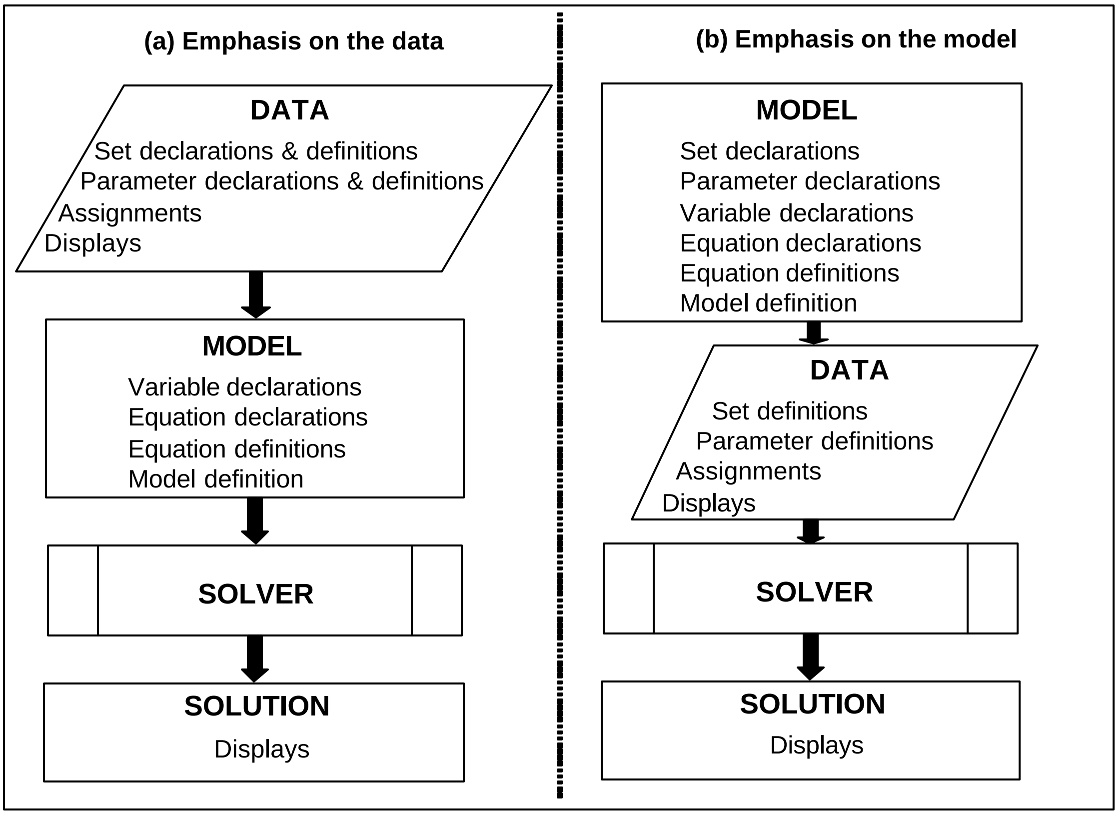


Figure 5.4: Different types of organisation for a GAMS input file

The first configuration (organisation chart (a) on Figure 5.4) places the data first, followed by the model and then the solution statements. In this kind of organisation, the sets are placed first, then the data are specified with parameter, scalar and table statements. Next, the model is defined with the variable, equation declaration, equation definition and model statement. Finally, the model is solved and the results are displayed.

The second configuration (organisation chart (b) on Figure 5.4) emphasizes the model by placing it before the data. This is a particularly useful order when the model needs to be solved repeatedly with different data sets. For this particular configuration, there is a need to separate the declaration phase from the definition phase.

5.2.1.6 Debugging

TWO types of errors can be encountered during the phase of debugging: errors during compilation and errors during execution. All errors are written to the output file.

5.2.1.6.a Compilation errors

MANY errors are usually reported in the output file the first time an input file is run. Since errors further down in the input file are often a result of an error earlier in the file, it is recommended to begin the process of error corrections from the top of the file and to rerun the input file after an error has been corrected. Compilation errors indicate that the input file contains statements that are not recognised by the compiler. Common compilation errors include syntax errors or incorrect reference to a variable name. Until the compilation errors are corrected, GAMS is unable to execute the model.

5.2.1.6.b Execution error

EXECUTION errors occur after the program has been compiled successfully, during the phase when the solver is attempting to locate the optimum solution. It is usually more difficult to correct these errors because they are related to the optimisation algorithm. When trying to resolve an execution error, it is generally a good idea to begin checking the optimisation formulation and its transcription in the GAMS input file. Then, adding new bounds or altering existing ones is also something to be looked at seriously. When the solver reports that the problem is infeasible, it may be necessary to consider removing some of the constraints or slackening some of the variables. In the particular case of a NLP problem, the solver may not be able to locate a feasible solution because of the way the equations have been formulated.

It may be necessary to reformulate them to give the solver an alternate route to feasible solutions and the optimum.

5.2.2 Application of the GAMS modelling language to the current optimisation problem

5.2.2.1 Definition of the MINLP problem to solve

THE MINLP problem that needs to be solved to achieve the operational optimisation of water distribution networks depends on which modelling strategy has been chosen during the phase of construction of the model (refer to Chapter 4).

$$Min\ \ f = \sum_{k=1}^{N} \Big[\left(\overline{V} - \overline{V_{SP}}\right)_k^T \overline{W_V} \left(\overline{V} - \overline{V_{SP}}\right)_k + \left(\overline{C} - \overline{C_{SP}}\right)_k^T \overline{W_C} \left(\overline{C} - \overline{C_{SP}}\right)_k + \left(\overline{f}_k^{\ T}\ \overline{Pump_1} + \overline{Pump_2}\right) \overline{E_{cost}} \Big]$$

$$\text{subject to:} \quad \begin{cases} \overline{\Delta P_F} = M_{pressure}\, \overline{P} + \overline{\Delta P_{pump}} \\ \overline{V}(t+\Delta t) = \overline{V}(t) + \overline{f_{aux}}(t)\,\Delta t + M_{flow}\, \overline{f}(t)\,\Delta t \\ \overline{V}(t+\Delta t)_{|t}\, \overline{C}(t+\Delta t)_{|t} = \overline{V}(t)\,\overline{C}(t) \\ \qquad + \overline{f_{aux}}^T(t)\, \overline{C_{aux}}(t)\,\Delta t \\ \qquad + \overline{f_{in}}^T(t) \left(\dfrac{\overline{C_{in}}(t+\Delta t) + \overline{C_{in}}(t)}{2} \right) \Delta t \\ \qquad + M_{flow}\, \overline{f}^T(t) \left[\dfrac{\overline{C}(t+\Delta t) + \overline{C}(t)}{2} \right] \Delta t \\ \qquad - \overline{k}^T(t) \left[\dfrac{\overline{V}(t+\Delta t)\,\overline{C}(t+\Delta t) + \overline{V}(t)\,\overline{C}(t)}{2} \right] \Delta t \end{cases}$$

$$\begin{cases} \overline{X} \geq 0 \\ \overline{X} \leq 1 \\ \overline{f} \geq \overline{f_{\min}} \\ \overline{f} \leq \overline{f_{\max}} \\ \overline{V} \geq \overline{V_{\min}} \\ \overline{V} \leq \overline{V_{\max}} \\ \overline{C} \geq \overline{C_{\min}} \\ \overline{C} \leq \overline{C_{\max}} \end{cases}$$

Problem 5.1: MINLP problem to solve for the operational optimisation of water supply systems (calculated flows strategy)

Problem 5.1 summarizes the MINLP problem that is solved when the calculated flows strategy is used. It conforms the formulation of a general MINLP problem, as presented in Chapter 3. In this problem, $M_{pressure}$ represents the matrix of pressure selections which accesses the full pressure vector to differentiate the pairs of output pressures from the rest of the data contained in $\overline{P}$. Similarly, M_{flow} corresponds to the flow interconnection matrix. It is used to select a sum of flows from the full flow set to the input leg of each element.

Similarly, Problem 5.2 represents the MINLP formulation that is solved when the model is based on the available flows strategy. It is also conform to the general MINLP problem formulation presented Chapter 3.

$$Min\ f = \sum_{k=1}^{N}\Big[\left(\overline{V}-\overline{V_{SP}}\right)_k^T \overline{W_V}\left(\overline{V}-\overline{V_{SP}}\right)_k + \left(\overline{C}-\overline{C_{SP}}\right)_k^T \overline{W_C}\left(\overline{C}-\overline{C_{SP}}\right)_k + \left(\overline{f}_k^{\,T}\,\overline{Pump_1} + \overline{Pump_2}\right)\overline{E_{cost}}\Big]$$

subject to:

$$\begin{cases} \overline{f} \le \overline{X}^T\ \overline{f_{available}}\ \left(+\ \overline{\Delta f_{pump}}\right) \\ \overline{V}(t+\Delta t) = \overline{V}(t) + \overline{f_{aux}}(t)\Delta t + M_{flow}\ \overline{f}(t)\ \Delta t \\ \overline{V}(t+\Delta t)_{|t}\,\overline{C}(t+\Delta t)_{|t} = \overline{V}(t)\overline{C}(t) \\ \qquad + \overline{f_{aux}}^T(t)\,\overline{C_{aux}}(t)\Delta t \\ \qquad + \overline{f_{in}}^T(t)\left(\dfrac{\overline{C_{in}}(t+\Delta t)+\overline{C_{in}}(t)}{2}\right)\Delta t \\ \qquad + M_{flow}\,\overline{f}^T(t)\left[\dfrac{\overline{C}(t+\Delta t)+\overline{C}(t)}{2}\right]\Delta t \\ \qquad - \overline{k}^T(t)\left[\dfrac{\overline{V}(t+\Delta t)\overline{C}(t+\Delta t)+\overline{V}(t)\overline{C}(t)}{2}\right]\Delta t \end{cases}$$

$$\begin{cases} \overline{X} \ge 0 \\ \overline{X} \le 1 \\ \overline{V} \ge \overline{V_{\min}} \\ \overline{V} \le \overline{V_{\max}} \\ \overline{C} \ge \overline{C_{\min}} \\ \overline{C} \le \overline{C_{\max}} \end{cases}$$

Problem 5.2:MINLP problem to solve for the operational optimisation of water supply systems (available flows strategy)

In the formulation above, M_{flow} still represents the flow interconnection matrix, as described above.

5.2.2.2 GAMS input file developed to solve the current optimisation problem

THE GAMS program developed for the operational optimisation of water distribution networks is stored on the CD attached to the thesis, at the following location:

\Programming\GAMS programming

The file presented on the CD is based on the formulation of Problem 5.2 which corresponds to the available flows strategy presented in Chapter 4. It has been preferred to the first formulation, since the available flows strategy, being more robust and solving more rapidly, is the strategy which was chosen for the application on large-scale networks.

Since the set of data provided to the model does not change much once the topology of the supply system has been fixed, the GAMS input file is based on configuration (a) presented in Figure 5.4. Therefore, the emphasis is held on the data rather than on the model.

5.2.2.3 Choice of an adequate MINLP solver

THE two modelling strategies represented by Problem 5.1 and Problem 5.2 involve integer variables. They are therefore typical examples of applications where a MINLP solver is required. Few MINLP solvers are provided in GAMS since the modelling and optimisation of MINLP problems has not yet reached the maturity and reliability found in LP, MILP and NLP modelling. DICOPT++ is probably the MINLP solver which has gained the most interest among researchers: it has been chosen for this particular project.

5.2.2.3.a Presentation of DICOPT++

THE GAMS / DICOPT++ system has been designed with two main goals in mind. Firstly, it builds on existing modelling languages and provides compatibility to ensure easy transition from existing modelling applications to non-linear mixed-integer formulations. Then, it uses existing optimisers to solve the DICOPT sub-problems. This makes possible the use of the most adequate solvers to the current problem and guarantees that any new developments and improvements in the NLP and MILP solvers become automatically and immediately available in DICOPT++.

The actual DICOPT program is based on the extensions made on the initial outer-approximation algorithm (OA) developed in the middle of the 1980's. In its current version, DICOPT is known as DICOPT++ and is based on the Augmented Penalty / Outer Approximation / Equality Relaxation algorithm (AP / OA / ER) developed by Viswanathan and Grossman in the 1990's at the Engineering Design Research Centre, at Carnegie Mellon University in the United States. The MINLP algorithm inside DICOPT++ solves a series of NLP and MILP sub-problems. It is important to note that these sub-problems can be solved using any NLP or MILP solvers that run under GAMS. Therefore, besides choosing a MINLP solver, it is also necessary to choose a suitable solver for the NLP and MILP sub-problems. For this particular optimisation problem, CONOPT2 has been identified as the best NLP solver, while CPLEX was used to solve the MILP sub-problems. Figure 5.5 summarizes the interactions that occur between the different solvers during a normal MINLP optimisation process.

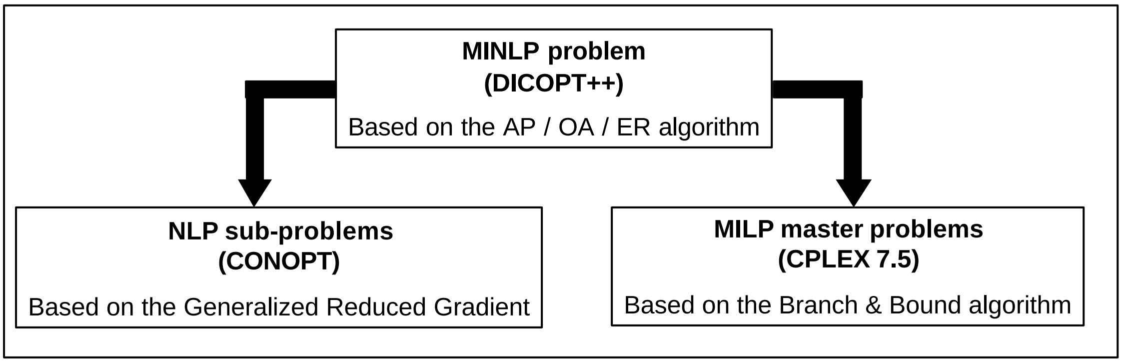


Figure 5.5: Interactions between the solvers involved in the optimisation of the MINLP problem

5.2.2.3.b Presentation of the solving process followed by DICOPT++

THE DICOPT++ algorithm starts by solving the NLP in which all the of 0 – 1 conditions on the binary variables are relaxed. If the solution to this problem leads to an integer solution, the search stops. Otherwise, it continues with an alternating sequence of non-linear programs (called sub-problems) and mixed-inter linear programs (called master problems). The NLP sub-problems are solved for fixed 0 – 1 variables that are predicted by the MILP master problem at each iteration. The default stopping criterion in the implementation of DICOPT++ for non-convex systems relies on the use of heuristics, since the optimisation process is generally stopped as soon as the NLP sub-problems start worsening. In other terms, the optimisation is stopped when the current NLP sub-problem has an optimal objective function that is worse than the previous NLP sub-problem.

5.2.2.3.c Options to be specified in the DICOPT++ solver

SEVERAL options can be specified for the DICOPT++ solver. Some of the options are specified through the GAMS option statement, while others are entered through a separate option file generally referred to as "dicopt.opt". The options allow, among other choices, the determination of the iterations limit, the setting of both NLP and MILP solvers or the stopping criterion to be used.

For more information on this subject, the reader is invited to refer to the GAMS/DICOPT++ manual provided with the GAMS software.

5.2.2.3.d Choice of an adequate solver for the NLP sub-problems

THE non-linear sub-problems that are created during a MINLP optimisation need to be solved with a NLP algorithm. Currently, there are two standard NLP algorithms available in GAMS: MINOS and CONOPT. Both algorithms attempt to find a local optimum.

The algorithms in MINOS and CONOPT are based on different mathematical algorithms, and they behave differently on most models. This means that, while MINOS is superior for some models, CONOPT is superior for others. However, it is almost impossible to predict how difficult it is to solve a particular model with a particular algorithm, especially for NLP models. Therefore, GAMS has not been designed to select the best algorithm automatically which means that the modeller needs to perform tests on his model to determine the most suitable NLP solver.

Several rules of thumb are available for the modellers, so that it is possible to identify which of the two algorithms is the best for the optimisation problem being considered. These rules are helpful, since they provide an idea of the best NLP solver, even before tests are performed on the model with both solvers. For instance, it is acknowledged that CONOPT is well suited to models with constraints that are highly non-linear. On the other hand, when the model does not present major non-linearities outside of the objective function, MINOS is probably the best solver. Here, both modelling strategies present strong non-linearities in the constraints: according to the first rule of thumb, it would seem that CONOPT is the more suitable.

CONOPT also has the ability to find a first feasible solution quickly on models with few degrees of freedom. Therefore, for models having roughly the same number of constraints and variables, CONOPT should usually be preferred.

On the contrary, if the number of variables is much larger than the number of constraints, MINOS usually performs better. This analysis must be performed for each modelling strategy.

Firstly, for the calculated flows strategy, the number of variables is estimated at $9 \times nel$ (where nel represents the number of standard modelling elements included in the model), while the number of constraints is estimated at $12 \times nel$. Therefore, the number of constraints is greater than the number of variables which does not vote in favour of MINOS. A deeper analysis shows that the number of constraints is greater to the number of variables by a factor of 25 %. Thus, it can be considered than the models developed with the calculated flows strategy have roughly a similar number of variables and constraints. According to the second rule of thumb, non-linearities in the model present another argument in favour of the use of CONOPT. Similarly, an analysis on the number of variables and constraints for models based on the available flows strategy shows that the number of constraints ($8 \times nel$) is greater than the number of variables ($6 \times nel$) by the same factor of 25 %. Therefore, CONOPT should to be better than MINOS according to the second rule of thumb applied to the available flows strategy.

The next rule of thumb states that CONOPT generally takes advantage of models where many equations can be solved one by one. This occurs in the pre-processing step where CONOPT can and remove this kind of equation from the model. In each of the modelling approaches, it is possible to find equations which can be solved separately from the others ($2 \times nel$ pressure equations for the calculated flows strategy, and $2 \times nel$ output flows equations for the available flows strategy). However, most of the other equations required for the volume and chlorine concentration balances do not fall into this category. It is therefore difficult to evaluate here whether the solving process of the two modelling strategies will benefit from the pre-processing step used by CONOPT.

Finally, CONOPT has many other advantages over MINOS. It has been especially developed for large and sparse models which means than it is suited to models where most functions only depend on a small number of variables. This is the case for the two modelling strategies proposed to solve the current optimisation problem. Furthermore, CONOPT is more user-friendly in the sense that it has many built-in tests and messages which help modellers to improve their model. It is a helpful debugging tool during the model development phase.

By summarising all of the arguments presented above, it is obvious that CONOPT seems to be better adapted than MINOS to the two modelling strategies developed for this project.

The preliminary tests conducted during the phase of validation of the calculated flows strategy have progressed well in this direction, showing strong performance using CONOPT. Unlike CONOPT, MINOS has shown problems maintaining feasibility during the full optimisation process and, for instance, it has not been possible to get a complete feasible solution with MINOS on configurations including more than 10 standard modelling elements.

These results have justified the use of CONOPT as the NLP solver, since the model presented in Chapter 4 has been developed especially for application on large-scale configurations.

5.2.2.3.e Choice of an adequate solver for the MILP sub-problems

THE master problems that are created during a MINLP optimisation need to be solved with a MILP algorithm. Several MILP algorithms are available in GAMS. These algorithms are known as BDLMP, OSL, CPLEX, XA or XPRESS. However, BDLMP is the only one algorithm which comes as standard with the GAMS software, and licences must be purchased separately to get access to the remaining solvers. Since BDLMP has been designed for small-scale applications, it is not applicable for the operational optimisation of a water distribution network.

The initial modelling developments for this particular project have been conducted with OSL, since a licence was available for this solver at the School of Chemical Engineering of the University of Natal (Durban, South Africa), where this research has been conducted. However, as the developments progressed to applications involving 10 standard modelling elements, it appeared that OSL could not cope any more with the large size of the applications. Unlike for the NLP solvers, no rules of thumb exist to determine the best MILP solver for a particular application. Therefore, it was decided to get a trial licence from the GAMS Development Corporation, for each of the remaining MILP solvers to determine if one of them was more suitable than OSL. During this trial period, the same configuration involving 10 standard modelling elements on which OSL failed has been tested with CPLEX, XA and XPRESS.

Since CPLEX was the only one solver capable of achieving a feasible solution on this particular configuration, a license has been purchased from the GAMS Development Corporation for the full use of CPLEX. The solver CPLEX 7.5 has been kept as the basis of the MILP optimisation, for the rest of the developments undertaken during this project.

5.2.2.3.f Additional information on CONOPT and CPLEX

THIS paragraph presents the general spirit of the CONOPT and CPLEX solvers. For detailed information on how the two solvers are used to solve the MINLP sub-problem, the user is invited to refer to the documentation provided by the GAMS Development Corporation.

As described before, CONOPT is used to solve the NLP sub-problem of the general MINLP optimisation process. This algorithm is based on the Generalized Reduced Gradient (GRG) approach suggested by Abadie and Carpentier (1969) and already presented in Chapter 3. More details on the algorithm can be found in Drud (1985). CONOPT has a considerable built-in logic that selects a solution approach that seems to be suited for most models. This solution approach is adjusted dynamically as information about the behaviour of the model is collected and updated. This feature generally prevents the definition of an option file, even if it is a possibility that can be considered for advanced users solving very large and complex models. CONOPT is therefore a solver that is very convenient to use.

CPLEX is used to solve the MILP master problems of the general MINLP optimisation process. It was released for the first time in 1988 and successively improved since. CPLEX was the first commercial optimiser developed in the C programming language. Since the early developments, successive algorithmic innovations have improved solution times by more than a factor of 100. When CPLEX is used to solve mixed-integer programming problems, the solver uses a Branch and Bound (BB) algorithm which solves a series of LP sub-problems, as already presented in Chapter 3. Because a single mixed-integer problem generates many LP sub-problems, even small mixed-integer problems can be very computation intensive and require significant amounts of physical memory. Running out of memory is precisely the most common difficulty when solving mixed-integer programming problems. Furthermore, one frustrating aspect of the BB technique for solving mixed-integer programming problems is that the solution process can continue long after the best solution has been found. It is here remembered that the BB tree may be as large as 2^n nodes, where n represents the number of binary variables. For instance, a problem containing only 30 binary variables could produce a tree having over one billion nodes. While numerous solving options are available, CPLEX automatically calculates and sets most option at the best values for the current optimisation process, similarly to the CONOPT solver. It is however possible to set these options manually.

5.3 Concluding remarks to the chapter

IN this chapter, the different tools used to implement mathematically the modelling strategies presented in Chapter 4 have been presented. The mathematical software MATLAB has been combined with the optimisation package GAMS. While MATLAB is used to create an interface for the capture of the data required by the model, GAMS is used to find the optimal adjustments of the system. For this particular project, the MINLP solver DICOPT++ has been used: it is associated with CONOPT and CPLEX to solve for the NLP and MILP sub-problems created by the general MINLP optimisation problem.

In the next chapter, presentation is made of a small-scale application. It is used to validate all of the features of the model presented in Chapter 4, before the method is applied to an existing large-scale water supply system.

CHAPTER 6 – APPLICATIONS OF THE MODELLING STRATEGIES ON SMALL-SCALE EXAMPLES

BEFORE applying the model on a real target application, it was rather decided to develop several case studies of limited scale to check that the two modelling strategies give logical and reliable results.

6.1 Introduction to the results presented

SEVERAL tests have been conducted on both modelling approaches to determine their performances, before deciding which one to apply on large-scale applications. The calculated flows strategy was developed first. It was then tested on several configurations which were used to highlight the ability of the calculated flows strategy to provide operational optimisation of a typical water distribution network, while responding simultaneously to users' demands and water quality objectives, as well as security constraints on reservoir volumes and chlorine concentrations in the reservoirs.

Since it has not been possible to obtain feasible results with this strategy easily, even on small-scale applications, it has been necessary to develop a less complex model: the available flows strategy. This new modelling approach immediately gave satisfactory results on the largest configuration tested initially. This has constituted a major improvement compared to the initial modelling strategy.

This chapter presents the application examples on which both modelling strategies have been tested. The results obtained on these applications will give reasons for preferring the available flows strategy to the calculated flows strategy.

6.2 Examples of application with the calculated flows strategy

THE calculated flows strategy is based on a fundamental modelling approach which gives a comprehensive hydraulic model. It requires the definition of pressure, volume and chlorine concentration balances. Several constraints and an objective function are also defined.

Initially, a configuration including all of the features of a typical water distribution system was created to test the ability of this modelling approach to tackle operational optimisation problems.

Since the initial results on this approach were discouraging, it was decided to break the initial problem into two sub-problems to dissociate the constraints and users' demand objectives from the quality objective. In the following paragraphs, the first sub-problem will be referred to as the "water case study", while the second one will be called the "chlorine case study".

6.2.1 Operational optimisation of a water supply system

INITIALLY, a water distribution system composed of 10 standard modelling elements was created. It included a waterworks composed of one storage reservoir, connected to a network consisting of several pipelines linking one pump-station (composed of a fixed-speed pump) with four storage reservoirs (preceded each by a binary valve). This network had seven different connections with the reticulation system.

6.2.1.1 Purpose of the case study

THE purpose of this case study was to highlight the ability of the modelling approach, referred to as the calculated flows strategy, to produce the operational optimisation of a typical water distribution network.

The main objective here was to reduce the pumping operations to their minimum, satisfying simultaneously volume setpoints in storage reservoirs and user's demand flows out of the reservoirs, using valve arrangements and flow redirections. Water quality objectives were also considered, but this optimisation objective was secondary.

A typical electricity cost pattern was also defined and users' demand flows were specified on output legs connected to reticulation. All of the simulations were conducted over a period of 48 h, using a control horizon of 12 h. Therefore, according to the Model Predictive Control (MPC) theory presented in Chapter 3, it was necessary to provide data tables with 60 h of known time-varying coefficients for electricity cost and demand flows. This allowed the algorithm to look ahead at future demands over the full period.

6.2.1.2 Presentation of the configuration

THE 10-element configuration is presented in Figure 6.1.

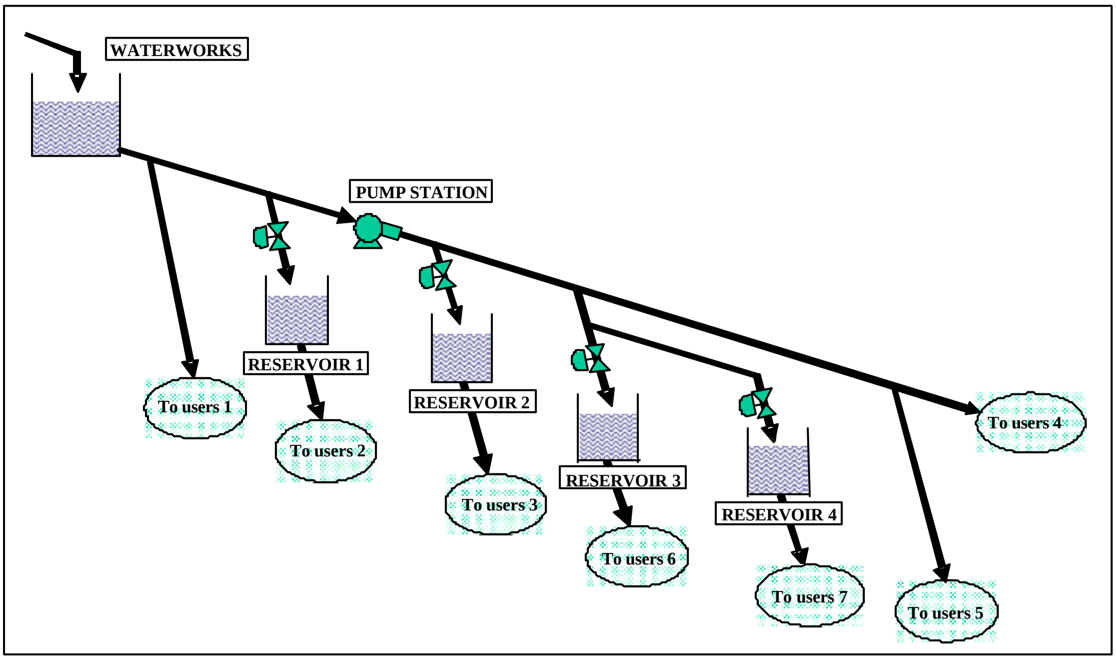


Figure 6.1: Operational optimisation of water supply systems - 10-element configuration

Figure 6.2 represents the same configuration in terms of standard elements.

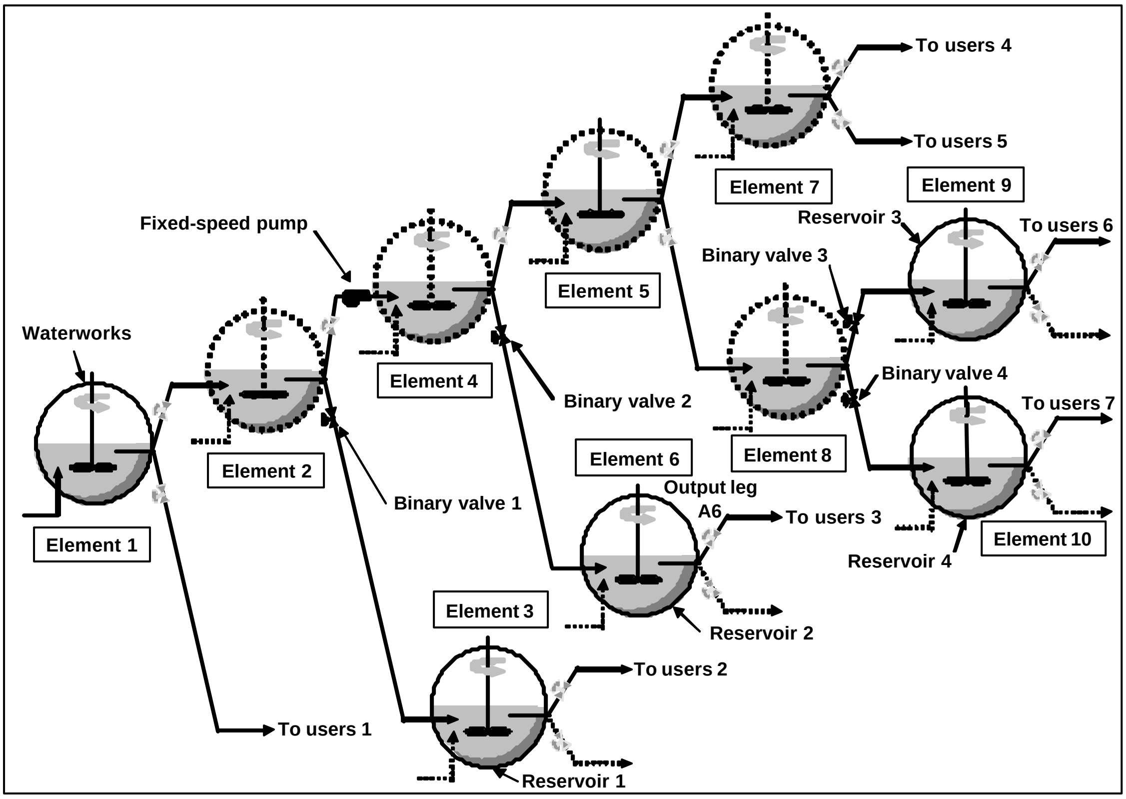


Figure 6.2: Representation of the 10-element configuration in terms of modelling elements

It required the combination of 10 modelling elements and, as presented in Chapter 4, each of these elements was associated with a number.

On Figure 6.2, it is necessary to differentiate connections that are plain from connections that are dotted. The plain connectors represent links that are active: in other words, links that are required for the construction of the distribution system being considered. Conversely, the dotted connectors are used for the connections that come as standard because of the use of a common modelling element, but that are not necessary for the definition of the particular configuration studied. Let us take the example of the second modelling element. Its vessel and its input pipeline are represented by a dotted contour or a dotted line. With the notation presented above, this means that this element which is used to represent a junction pipeline according to Figure 6.1, requires neither the definition of any of vessel properties, nor the definition of an auxiliary input flow. Similarly, by considering the third modelling element, it can be observed that its auxiliary input pipeline and lower output leg are represented by dotted connectors: this is explained by looking again at Figure 6.1, and observing that the Reservoir 1 is not fed by any auxiliary input flow and is composed of a single output pipeline.

The scheme presented Figure 6.2 is then used for the constitution of the topology table as well as of the main and the auxiliary data tables, following the rules presented in Chapter 4.

6.2.1.3 Definition of the data tables associated to the 10-element configuration

THE data tables required for the constitution of the 10-element configuration are presented in the CD attached to the thesis. The reader is invited to open the file entitled “10-element configuration – data.m” which can be found under the following path:

\Small-scale applications\Calculated flows strategy\Operational optimisation of a water supply system

It can be noted that the topology table has been combined with the main data table. Furthermore, a binary structure has been chosen for the parameter representing the electricity cost. In other words, the time-varying parameter used for this particular configuration combines a low value for off-peak periods with a high value, for peak periods. This cost structure is shown in Figure 6.3.

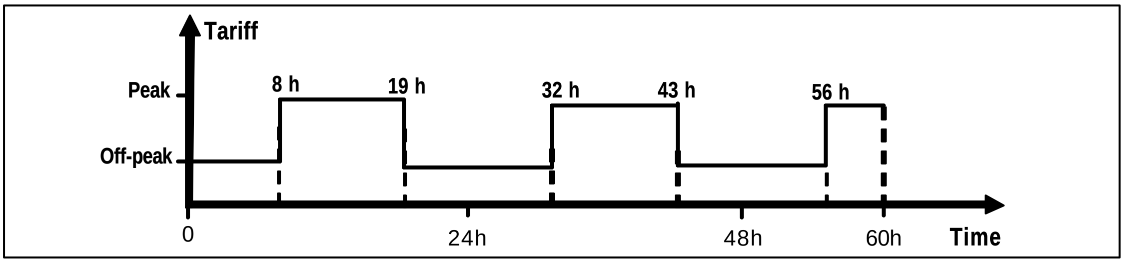


Figure 6.3: Calculated flows strategy – 10-element configuration: electricity cost structure

6.2.1.4 Choice of the solver and definition of the parameters of the optimisation

THE configuration presented on Figure 6.1 corresponds to the first example tested initially in the course of this research project. This particular example has been solved with the MINLP (Mixed Integer Non Linear Programming) solver DICOPT++, associated with the MILP (Mixed Integer Linear Programming) solver OSL and the NLP (Non Linear Programming) solver CONOPT, at a time when a licence for the CPLEX solver was not yet available.

The parameters of the optimisation can be found on the CD attached to the thesis, in the file already mentioned in paragraph 6.2.1.3. In addition, numerous options can be specified by the user for each of the three solvers mentioned above. The options chosen for this particular configuration can be found in the files entitled "dicopt.opt", "conopt.opt" and "osl.opt" stored on the CD attached to the thesis, under the location already mentioned in paragraph 6.2.1.3. . For more information on the significance of all of these parameters, the reader is invited to refer to the GAMS (General Algebraic Modeling System) manual (Brooke *et al.*, 1998)

6.2.1.5 Results

THE results of the optimisation for the 10-element configuration, solved with the calculated flows strategy, can be found in Biscos *et al.* (2002a). For convenience, an electronic version of this document can be found on the CD attached to the thesis. This file is available under the name "Water Science and Technology Paper.pdf" and is stored at the location already mentioned in paragraph 6.2.1.3.

This paper presents the optimised sequence of operations that was achieved with the calculated flows strategy, over a 48 h simulation period, with a 12 h control horizon, for this particular configuration. The graphs presented in this document show that the optimised sequence has been achieved with respect to the constraints specified in the main data table. An example of this can be found on Figure 6.4 below.

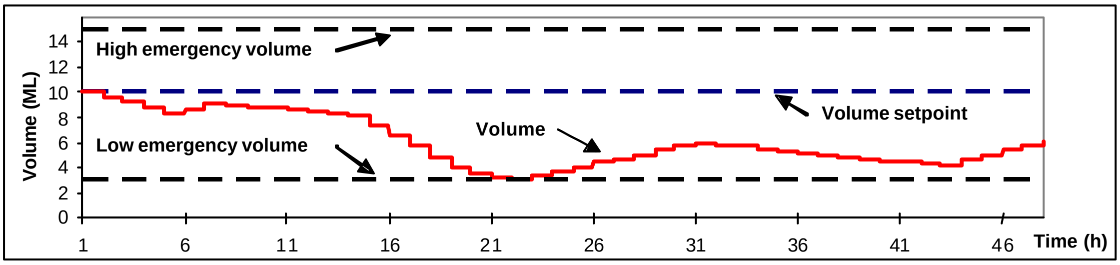


Figure 6.4: Evolution of the volume in Reservoir 1 (10-element configuration)

On this graph representing the evolution of the volume in Reservoir 1 against time, the volume always remains within the low and high emergency volumes specified by the user. The low emergency volume is reached for the 22nd hour of simulation, but something is done by the optimiser so that this limit is never crossed. In this particular case, the pump is switched on from the 20th hour of simulation to extract more water from the feed reservoir, as shown in Figure 6.5. This water is then directed towards Reservoir 1 to compensate the deficit occurring at this particular instant in this reservoir.

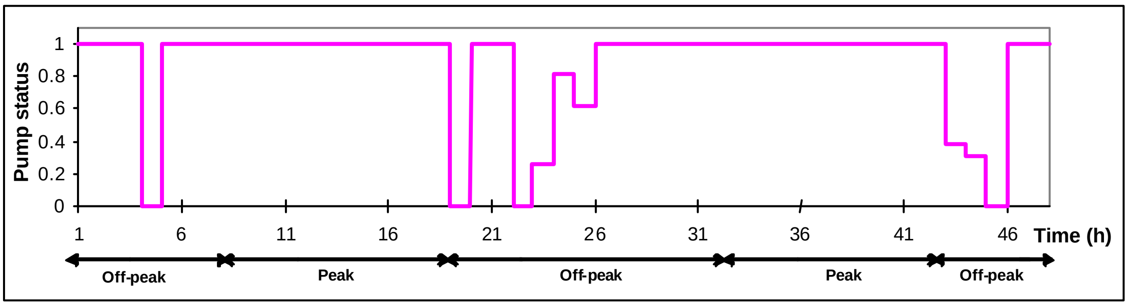


Figure 6.5: Pump status at the pump station of the 10-element configuration

It is also interesting to note the existence of an anticipation phenomenon. This consists in filling reservoirs while the cost of the power unit is the cheapest so that it is possible to use this water during electricity cost peak periods. This can be seen by looking at Figure 6.5 which shows that the pump is switched on in the middle of a period where the electricity is charged at its off-peak tariff (from the 23rd of simulation onwards). Unlike the previous use of the pump, where it responded to an emergency situation relating to the volume of one of the reservoirs, this new action of the pump does not coincide with any problems on one of the reservoirs. Therefore, all the additional low-cost energy provided by the pump within the network is used to fill most of the reservoirs well above their setpoints, in preparation for the next high tariff consumption period that will begin from the 32nd hour of simulation. This anticipation phenomenon can be seen for instance for Reservoir 1, whose volume is strongly corrected towards its setpoint from the 23rd hour of simulation onwards, as presented on Figure 6.4. This particular result is very important since it demonstrates the viability of this project which was mainly to develop a new tool capable of achieving optimal operation of a water distribution system by making best use of low-cost power for pumping operations.

However, these initial results are not fully satisfactory. A careful look at Figure 6.5 shows indeed that it is not possible to achieve an optimal sequence involving complete binary switching for the pump. Therefore, the solution presented in this paper had to be relaxed for some of the time-steps of the prediction horizon to get a convergence, with the result that intermediate values are sometimes chosen for binary variables.

In other words, the use of the calculated flows strategy for the resolution of this water case study gives intractable results very quickly, as the number of modelling elements becomes too high. It has been therefore necessary to test other configurations in an attempt to get an integer strategy, with this particular modelling approach. Since no other MILP solver was available at the time this initial research was conducted, the only other parameters that could be modified were the size of the problem and the number of constraints included in the optimisation.

Therefore, the best option appeared to break the initial problem into two different problems of smaller size. While the first one concentrates on the users' demands and the respect of security constraints on the volume in the reservoirs, the second tackles the quality constraints in the reservoirs.

6.2.2 Reduction in the size of the initial problem

THE 10-element configuration presented in the above paragraphs has confirmed the soundness of the modelling choices that have been made to achieve the operational optimisation of water distributions networks. This particular configuration is a typical example of a supply system, since it consists of several storage reservoirs, linked to each other by pipelines. Pumps and valves ensure the conveyance of water to areas where it is needed the most and the optimisation sequence that was determined takes into account the security margins defined on volume in the reservoirs, as well as the water quality objectives. Therefore, this example has allowed the justification of most of the features of modelling approach based on the use of several standard modelling frameworks linked together.

However, it has not been possible to achieve results that were fully satisfactory especially for the behaviour of binary pump and valve switching. Therefore, a closer investigation considered the reservoir volume constraints and setpoint objectives, and the composition constraints and setpoint objectives in separate examples.

6.2.2.1 Water case study

THE water case study involved three standard modelling elements. Its main objectives were to satisfy the demand in water originating from two different sources of customers, as well as to redirect the flow leaving a storage reservoir, so that the volumes in two demand reservoirs were as close as possible to their setpoints.

Since this case study concentrated on the satisfaction of user demand flows, reservoir volume setpoints and emergency volumes, no specifications were required on chlorine concentration. It was hoped to provide an integer strategy for all of the binary variables, over the complete prediction horizon.

6.2.2.1.a Presentation of the configuration

THE water distribution system represented by the 3-element configuration is presented in Figure 6.6. This figure associates the normal representation with the representation in terms of standard modelling elements.

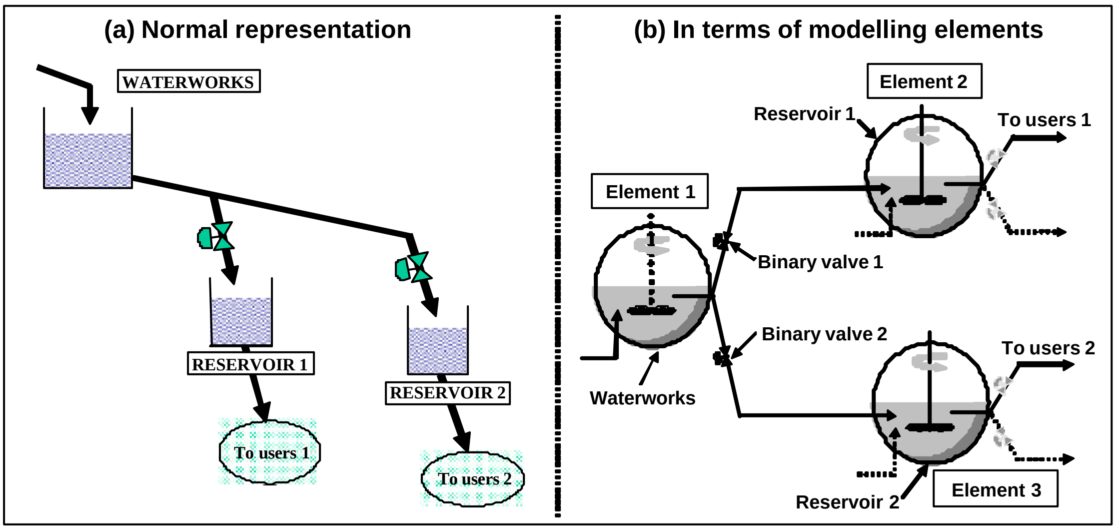


Figure 6.6: Water case-study – 3-element configuration

This case study is a reduced version of the distribution system presented in Figure 6.1. This new configuration was still fed by a unique storage reservoir, but the number of storage reservoirs and binary valves has been decreased to two, while the pump-station has been removed. The number of connections with the reticulation to users has also been reduced to two.

The demand pattern has been artificially created to represent a typical water consumption scheme with peak-demand periods and low-demand periods. The maximum capacity of the demand reservoirs has been set to 20 ML and their initial contents to 10 ML. On the contrary, the maximum capacity of the feed reservoir was chosen large enough (50 ML) to be sure that water could be provided in sufficient quantities in the system at any time. The initial charge was set to 30 ML. On each output leg of the feed reservoir, one binary valve is available to redirect the flow between the two demand reservoirs according to the level that is required in these reservoirs.

6.2.2.1.b Presentation of the data tables associated with the water case study

THE data tables required for the constitution of the 3-element configuration are presented in the CD attached to the thesis. The reader is invited to open the file entitled “3-element configuration – data.m” which can be found under the following path:

\Small-scale applications\Calculated flows strategy\Water case study

Since there was no pump in the system, it is obvious that the definition of time-varying coefficients for the parameter representing the electricity cost is not required.

6.2.2.1.c Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, still associated with the MILP solver OSL and the NLP solver CONOPT, at a time when a licence for the CPLEX solver was not yet available. The parameters of the optimisation can be found in the file already mentioned in the above paragraph.

The options chosen for the different solvers for this particular configuration can be found in the files entitled “dicopt.opt”, “conopt.opt” and “osl.opt”. These file are stored on the CD attached to the thesis, under the same location as the one mentioned in paragraph 6.2.2.1.b.

6.2.2.1.d Results

THE results of the optimisation for the 3-element configuration, solved with the calculated flows strategy, can be found in Biscos *et al.* (2002b). For convenience, an electronic version of this document can be found on the CD attached to the thesis. This file entitled “2002 WISA Conference Paper.pdf” is stored under the same location as the one already mentioned in paragraph 6.2.2.1.b.

This paper presents the optimised sequence of operations that was established with the calculated flows strategy, over a 48 h simulation period, with a 12 h control horizon, for this particular configuration.

The graphs presented in this document show that the optimised sequence has been achieved with respect to the security constraints on volume specified in the main data table. A good example of this can be found in Figure 6.7 which represents the evolution of the volume against time in the first demand reservoir

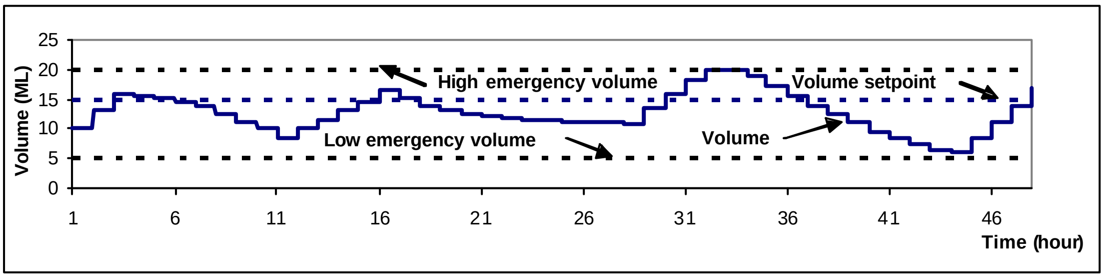


Figure 6.7: Evolution of the volume in Reservoir 1 (3-element configuration)

A detailed description of the evolution of the volumes in the feed and demand reservoirs can be found in Biscos *et al.* (2002b). What should be remembered from this paper is that, in this particular configuration, it is not possible to fill the two reservoirs simultaneously, even when there is a strong need of water in both reservoirs. As a result, it is difficult to have the volumes in demand reservoirs following closely their respective setpoints and the evolution is rather oscillatory, as shown in Figure 6.7 and Figure 6.8.

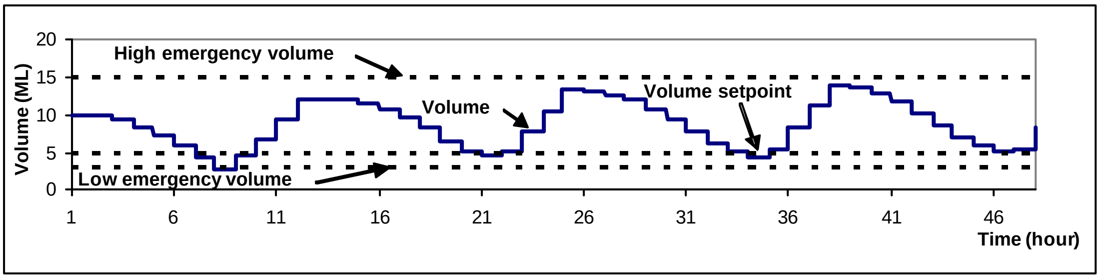


Figure 6.8: Evolution of the volume in Reservoir 2 (3-element configuration)

This highlights an interesting feature of any optimisation problem which is that, any result reflects the necessity of making trade-offs to reach a solution satisfying all objectives and constraints.

However, it has once again not been possible to achieve results that were fully satisfactory since the optimised sequence of valve switching does not correspond rigorously to a binary one, as shown by Figure 6.9 and Figure 6.10.

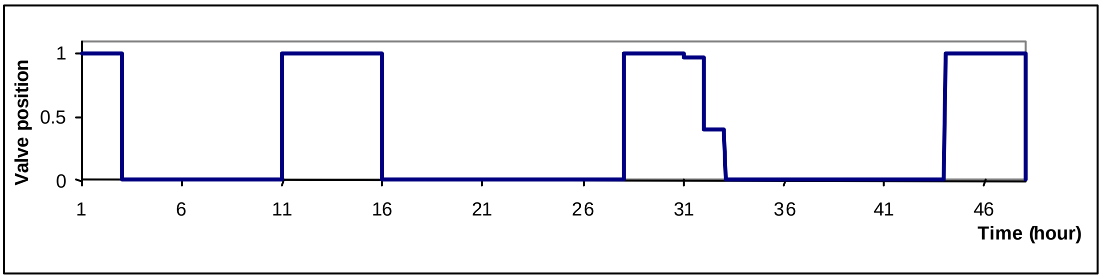


Figure 6.9: Valve position relative to the binary valve positioned at the input of Reservoir 1 (3-element configuration)

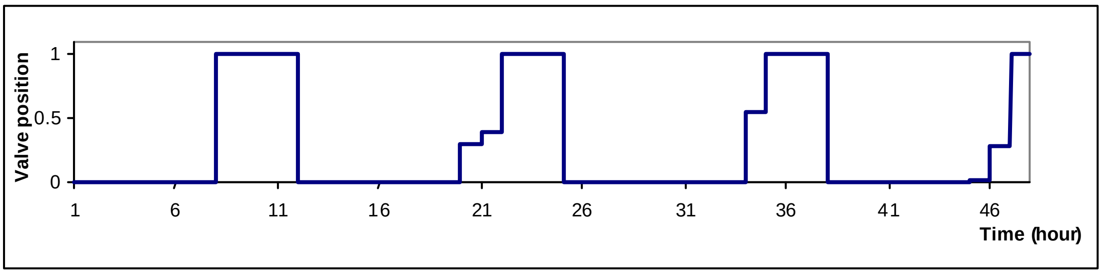


Figure 6.10: Valve position of the binary valve positioned at the input of Reservoir 2 (3-element configuration)

The solution presented in this paper had to be relaxed for some of the time-steps of the prediction horizon to get a convergence. This result is disappointing since, in this particular configuration, the number of standard elements used is very small. Therefore, this example strengthens the idea already mentioned that the calculated flows strategy is not adequate and too complex to solve for the kind of optimisation problems aimed for.

6.2.2.2 Chlorine case-study

THE chlorine case study involved four standard modelling elements. Its main objective was to fulfil chlorine concentration setpoints in two demand reservoirs by mixing water arriving from two different storage reservoirs. Since all of these reservoirs were interconnected, water could be routed from storage reservoirs to demand reservoirs using several valve arrangements.

As this case study concentrated on the satisfaction of chlorine setpoints, no specifications were required on volume setpoints. However, constraints relating to reservoir volume limits and demand flows of users were still defined. It was hoped to provide an integer strategy for all of the binary variables over the complete prediction horizon.

6.2.2.2.a Presentation of the configuration

THE water distribution system represented by the 4-element configuration is presented in Figure 6.11. This figure associates the normal representation with the representation in terms of standard modelling elements.

Similarly to the configuration used for the water case study, the number of storage reservoirs and the number of connections with the reticulation were equal to two. The number of binary valves was four. Finally, it can be observed that each of the storage reservoirs was fed by an auxiliary input flow.

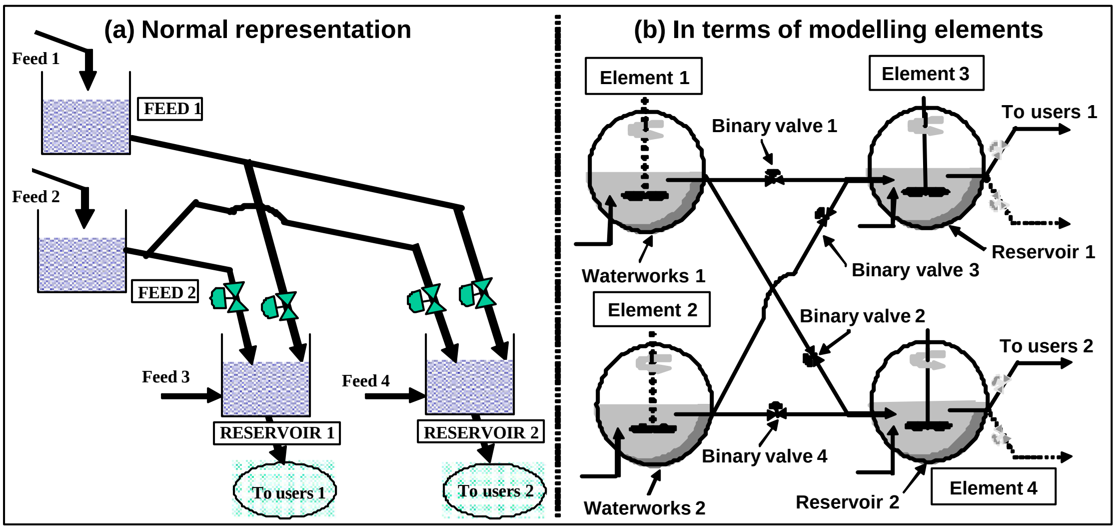


Figure 6.11: Chlorine case study – 4-element configuration

The first feed reservoir has been designed so that there is always a high chlorine concentration available at any time: therefore, a high volume and a chlorine concentration of 1.5 $mg.L^{-1}$ have been fixed. On the contrary, the second feed reservoir was used to supply low chlorine concentration water to the network (chlorine concentration fixed at $0.5mg.L^{-1}$). The initial capacity of both feed reservoirs was fixed at a relatively large volume of 50 ML, so that the volume constraints of these reservoirs had no effect. For each feed reservoir, two binary valves were available, one for each output leg.

At the other side of the network, the two demand reservoirs have been deliberately designed with a small initial volume (5 ML each) to speed up the concentration response in these reservoirs. The setpoint for the concentration in both demand reservoirs was fixed to an intermediate value (1 $mg.L^{-1}$). The first demand reservoir was initialised with a high chlorine concentration, while the second one was designed with an initial low chlorine concentration.

In all reservoirs, a first order chlorine decay constant was taken equal to $2d^{-1}$ (Powell *et al.*, 2000). All demand flows were taken equal to 40 $ML.d^{-1}$, and 10 $ML.d^{-1}$ was provided in each demand reservoir by the auxiliary feed flow.

6.2.2.2.b Presentation of the data tables associated with the chlorine case study

THE data tables required for the constitution of the 4-element configuration are presented in the CD attached to the thesis. The reader is invited to open the file entitled “4-element configuration – data.m” which can be found under the following path:

\Small-scale applications\Calculated flows strategy\Chlorine case study

Since there was no pump in the system, it is obvious that the definition of time-varying coefficients for the parameter representing the electricity cost is not required.

6.2.2.2.c Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, still associated with the MILP solver OSL and the NLP solver CONOPT, at a time where a licence for the CPLEX solver was not yet available.

The parameters of the optimisation can be found in the MATLAB file already mentioned in paragraph 6.2.2.2.b. Furthermore, the options chosen for all the solvers for this particular configuration can be found in the files entitled "dicopt.opt", "conopt.opt" and "osl.opt". These files are stored on the CD attached to the thesis, under the same location as the one mentioned in paragraph 6.2.2.2.b.

6.2.2.2.d Results

THE results of the optimisation for the 4-element configuration, solved with the calculated flows strategy, can be found in Biscos *et al.* (2002b). For convenience, an electronic version of this document can be found on the CD attached to the thesis. This file is available under the name " 2002 WISA Conference Paper.pdf" and is stored at the location already mentioned in paragraph 6.2.2.2.b.

This paper presents the optimised sequence of operations that was achieved with the calculated flows strategy, over a 48 h simulation period, with a 12 h control horizon, for this particular configuration.

The graphs presented in this document show that the optimised sequence has been achieved with respect to the security constraints specified on the volume in the reservoirs and on the chlorine concentration in the reservoirs. A good example of this can be seen on Figure 6.12 and Figure 6.13.

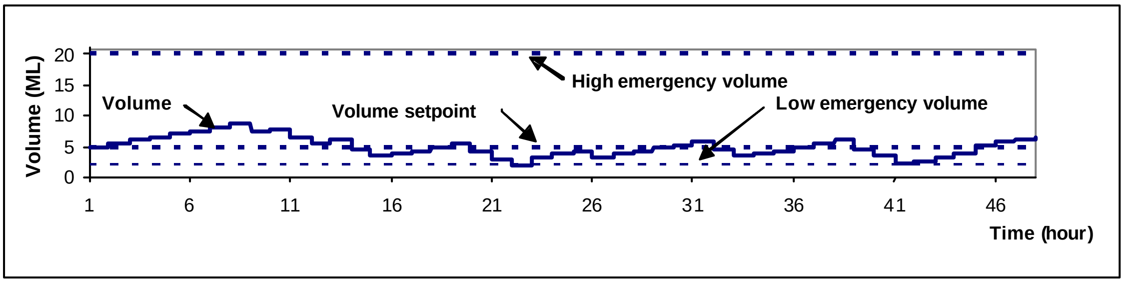


Figure 6.12: Evolution of the volume in Reservoir 1 (4-element configuration)

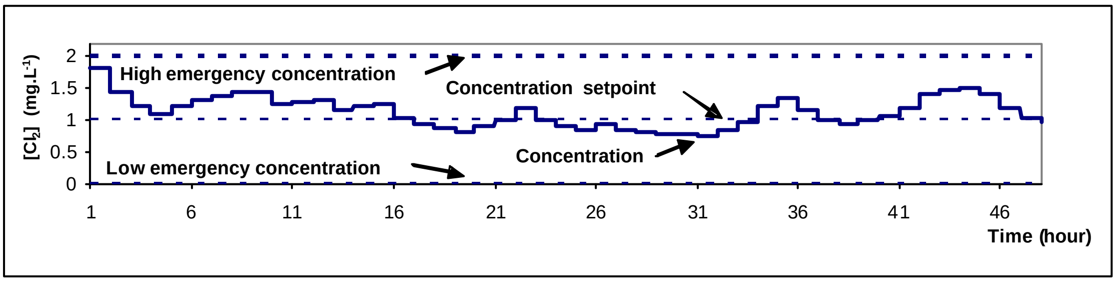


Figure 6.13: Evolution of the chlorine concentration in Reservoir 1 (4-element configuration)

A detailed description of the evolution of the chlorine concentration in the demand reservoirs can be found in Biscos *et al.* (2002b). The reader is invited to refer to this paper to find the explanations justifying these evolutions.

Unfortunately, similarly to the two previous case studies, the optimised sequence of valve switching does not correspond rigorously to a binary one, since the solution presented in this paper had to be relaxed for some of the time-steps of the prediction horizon. This can be seen on Figures 6.14 and 6.15 which represent respectively the valve status at the input of Reservoir 1 and Reservoir 2.

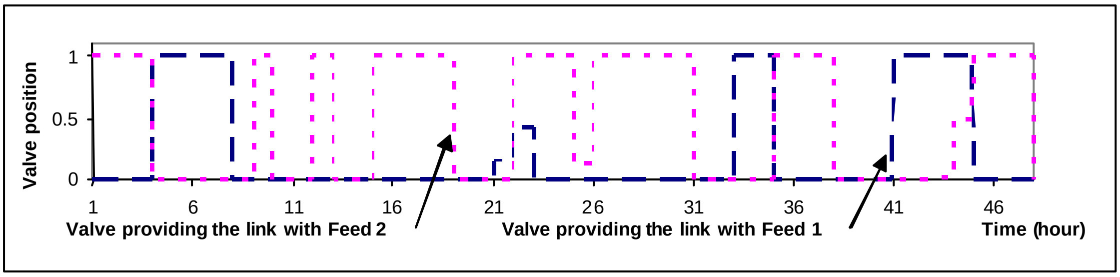


Figure 6.14: Valve positions of the binary valves positioned at the input of Reservoir 1 (4-element configuration)

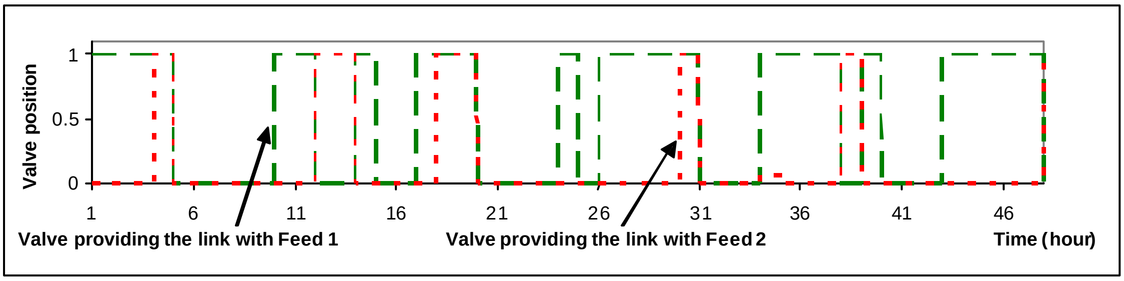


Figure 6.15: Valve positions of the binary valves positioned at the input of Reservoir 2 (4-element configuration)

This result is disappointing since, in this particular configuration, the number of standard elements used is very small. Therefore, it is now obvious that the complex pressure balanced approach which represents the core of the calculated flows strategy, must be simplified.

6.2.3 Concluding remarks for the calculated flows strategy

THREE examples of optimisation problems have been presented in the above section. Several constraints relative to security aspects on the volume in the reservoirs and to water quality aspects were included in the definition of the three problems. These constraints have played an important role in the determination of the optimal sequence of operations, since they restrict the states of the variables to certain values only.

All of the problems were defined by building several data tables and specifying numerous parameters following Chapters 4 and 5. Then, these problems were solved using the MINLP solver DICOPT++. Several MILP sub-problems needed to be solved during the course of the optimisation: this was made possible by the use of the OSL solver. Similarly, the CONOPT solver has been required for the solution of the NLP sub-problems.

It is important to note that optimised results have been reached for all three configurations. More importantly, these results have taken into account the constraints of each of the problems. In addition, several interesting behaviours like the anticipation phenomenon, or the existence of trade-offs between opposing objectives, have been highlighted.

However, it is obvious that these preliminary results are not fully satisfactory. The main objective of this research project being to develop a new tool providing an integer strategy for the binary valves and pumps of a water distribution network, and it is clear that it is not the case yet with the results presented in this section. In an attempt to decrease the level of difficulty on a configuration involving 10 modelling elements, it has been decided to break the initial problem into two separate applications of smaller size. While the first one concentrated on aspects of security regarding the volume in the reservoirs, the second one was more related to water quality problems. However, once again, a problem of infeasibility occurred. Starting from the fact that the solutions have to be relaxed for some of the time-steps of the prediction horizons, even on the smallest configurations, it has been necessary to admit that the calculated flows strategy is too complex to tackle the kind of problems aimed for. Hence, the development of a second strategy, still based on the use of a standard modelling element, but in which the number of non-linearities was reduced to its minimum, by removal of all of the pressure balances.

6.3 Results obtained with the available flows strategy

AS already mentioned in Chapter 4, the main feature of the available flow strategy is the definition of an hydraulic model which simplifies the pipe network hydraulics by removing all of the pressure balances.

The network components are still represented using a standard modelling element, but the flows in these elements are determined in a linear way, using a valve position or a pump status. Essentially, the system hydraulics are solved using a mass balance approach without taking into account the pipe frictions. All of the problems of non-linearity linked to the use of frictional pressure losses are therefore removed. This also helps to reduce the amount of data to be provided in the model.

The 10-element case study developed initially for the calculated flows strategy and including all of the features of a typical water distribution system has been tested with the available flows strategy to evaluate the ability of this modelling approach to tackle operational optimisation problems. Initially, it has been decided to simulate a simplified configuration where the security objectives on the volume of the reservoirs were separated from the water quality objectives. Because of the computational burden found on this example, results could only be obtained over a simulation period of 24 h, with an 8 h control horizon.

However, with the purchase of a licence granting access to a more powerful MILP solver (known as CPLEX), it has been possible to reunite later all of the objectives within the same optimisation problem and to extend the simulation period to 48 h, and the control window to 12 h, keeping exactly the same modelling approach.

6.3.1 Purpose of the case study

THE purpose of this case study is to highlight the ability of the modelling approach referred to as the available flows strategy to produce the operational optimisation of a typical water distribution network. The main objective here is to reduce the pumping operations to their minimum, satisfying simultaneously volume setpoints in storage reservoirs and user demand flows out of the reservoirs, using valve arrangements and flow redirections. Water quality objectives are also considered, but this optimisation objective is secondary, as the weighting factors associated with the chlorine concentration term in the objective function were deliberately reduced. A typical electricity cost pattern is also defined and user demand flows are specified on output legs connected to reticulation.

6.3.2 Presentation of the configuration

THE 10-element configuration used with the calculated flows strategy is the same as the one presented on Figure 6.1 and Figure 6.2.

It represents a water distribution system which includes a waterworks composed of one storage reservoir, connected to a network consisting of several pipelines linking one pump-station (composed of a fixed-rate pump) with four storage reservoirs (preceded each by a binary valve). This network has seven different connections with the reticulation system to users.

6.3.3 Operational optimisation of the configuration based on the MILP solver OSL

INITIALLY, the OSL solver was the only MILP solver for which a licence was available to solve the MILP sub-problems occurring during the normal solving process of any MINLP optimisation. Therefore, all of the preliminary tests on the available flows strategy were conduced with this particular solver.

The first tests conducted with the available strategy showed problems of convergence for some of the time-steps of the prediction horizon, but more importantly, the computational burden soon became unbearable. As mentioned earlier, to remove all of the problems of infeasibility and to quicken the convergence towards an optimised solution, it has been necessary to dissociate the objectives concerning security volumes in the reservoirs from the objectives associated with water quality considerations. Simultaneously, the simulation interval has been decreased from 48 h to 24 h, while the control horizon was changed from 12 h to 8h. As a result, only 32 h of data needed to be provided.

6.3.3.1 Consideration of reservoir volume security objectives only

IN this particular example, the main objective was to reduce the pumping costs to their minimum, by using gravity where possible, and favouring pump use during off-peak periods which had a lower cost of electrical power. Two different electricity tariff structures were tested to show the ability of the modelling strategy to handle several configurations. Each storage reservoir was defined with a volume setpoint which it tried to satisfy using valve arrangements and flow redirection from upstream parts of the water supply system. Apart from this, the user demand flows and the minimum and maximum volume constraints must be complied with. No requirements were held on chlorine concentration variables.

6.3.3.1.a Definition of the data tables

THE data tables required for the constitution of the configuration constrained by requirements on the reservoir volumes only are presented in the CD attached to the thesis.

The reader is invited to open the file entitled "10-element configuration – OSL - Security objectives data.m" which can be found under the following path:

\Small-scale applications\Available flows strategy\OSL solver\Security objectives

Two separate simulated experiments were conducted with different electricity cost structures, as represented on Figure 6.16.

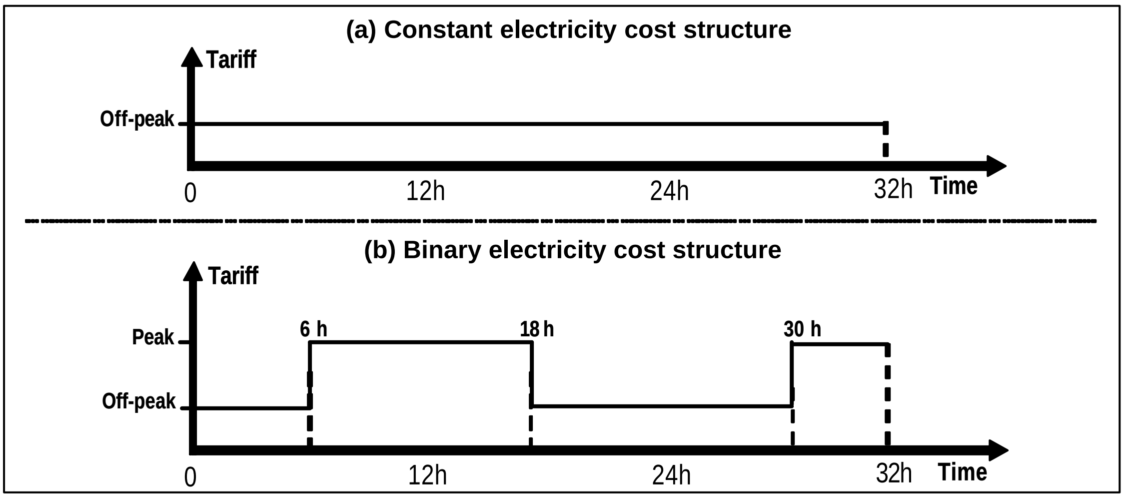


Figure 6.16: Available flows strategy/OSL – 10-element configuration: electricity cost structure

The first pattern is referred to as "constant electricity cost": it uses the same electricity cost factor over the whole prediction horizon. The second structure uses time-varying parameters for the electricity cost, similarly to what was presented in paragraph 6.2.1.3.

6.3.3.1.b Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, associated with the MILP solver OSL and the NLP solver CONOPT, at a time when a licence for the CPLEX solver was not yet available.

The parameters of the optimisation can be found in the MATLAB file already mentioned in the above paragraph. Furthermore, the options chosen for this particular configuration can be found in the files entitled "dicopt.opt", "conopt.opt" and "osl.opt", files which are stored on the CD attached to the thesis, at the same location than the one mentioned in paragraph 6.3.3.1.a.

6.3.3.1.c Results

THE results of the optimisation for the 10-element configuration, solved with the available flows strategy, with reservoir volume considerations only, can be found in Biscos *et al.* (2003).

For convenience, an electronic version of this document can be found on the CD attached to the thesis. This file is available under the name "WaterSA Paper.pdf" and is stored on the CD attached to the thesis, at the location already mentioned in paragraph 6.3.3.1.a.

This paper presents the optimised sequence of operations that was achieved with the available flows strategy, over a 24 h prediction horizon, with a 8h control horizon, for this particular configuration.

The graphs presented in this document show that the optimised sequence has been achieved for the two electricity structures, with respect to the security constraints relative to reservoir volumes. A good example of this can be seen on Figure 6.17 for the configuration based on a constant cost, and on Figure 6.18 for the configuration using a binary cost structure. These graphs represent the evolution of the volume in Reservoir 2 for the two electricity structures.

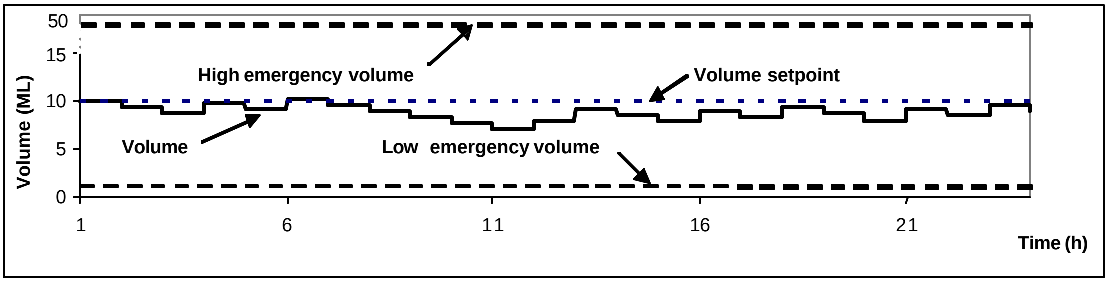


Figure 6.17: Evolution of the volume in Reservoir 2 (10-element configuration – constant cost)

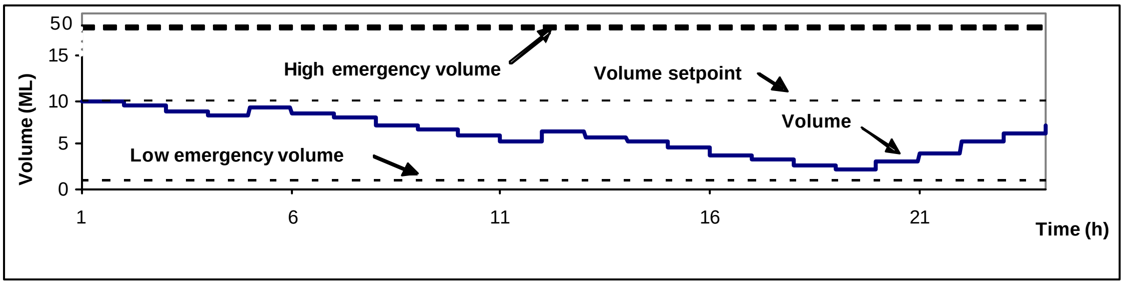


Figure 6.18: Evolution of the volume in Reservoir 2 (10-element configuration – binary cost)

A detailed description of the evolutions of the volumes in the feed and demand reservoirs can be found in Biscos *et al.* (2003). The reader is invited to refer to this paper to find the explanations justifying these evolutions.

What should be remembered from this paper is that the change in the modelling strategy has allowed the provision of an integer strategy for all the binary variables, over the whole prediction horizon. This can be seen on Figure 6.19 which represents the pump status for the different electricity cost structures.

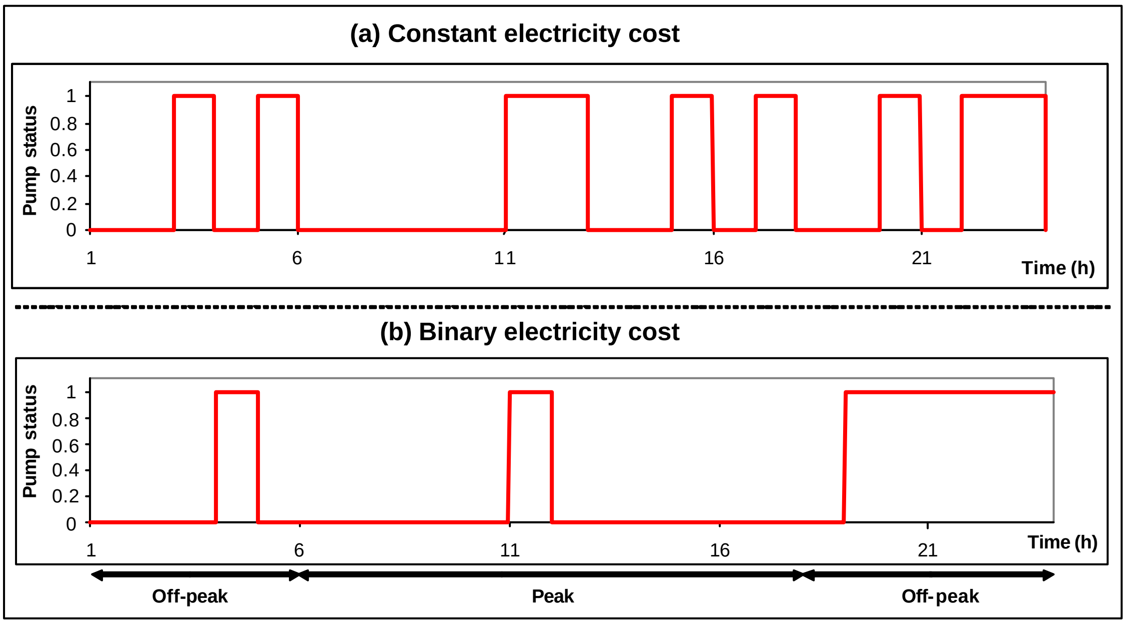


Figure 6.19: Pump status at the pump station for different electricity cost structures (10-element configuration)

This result is also valid for the position of the valves located at the input of the reservoirs. A good example of this is presented on Figure 6.20 and Figure 6.21 which give the evolution of the valve position of the binary valve located at the input of Reservoir 2. Note that similar results are obtained for other reservoirs.

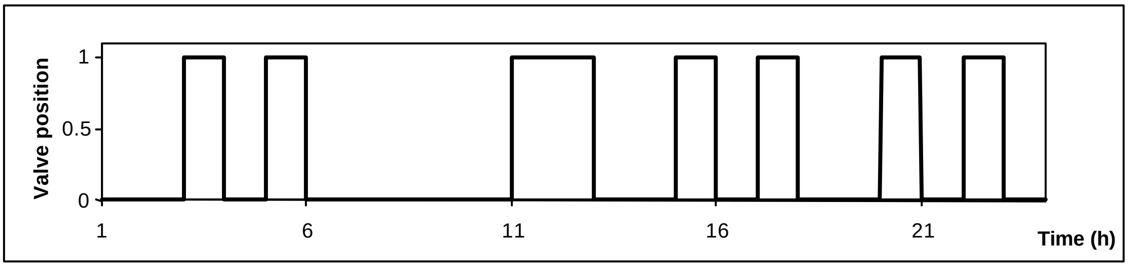


Figure 6.20: Valve position of the binary valve positioned at the input of Reservoir 2 (10-element configuration – constant cost)

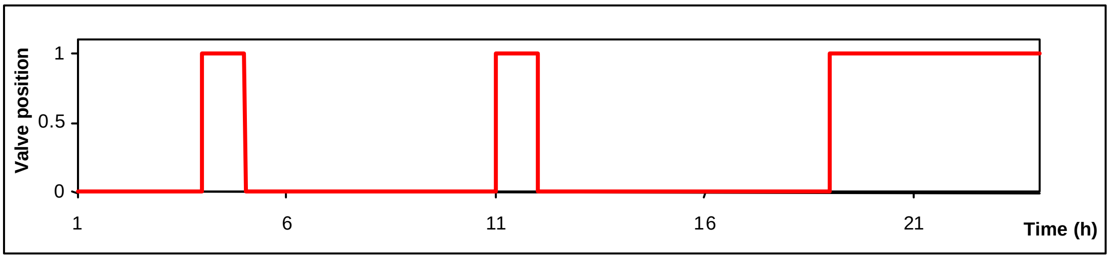


Figure 6.21: Valve position of the binary valve positioned at the input of Reservoir 2 (10-element configuration – binary cost)

These results are very encouraging for the extension of the method to configurations involving a larger number of standard elements.

Furthermore, it is interesting to note, as expected, that the structure of the electricity tariffs has a strong influence on the optimised sequence of valve and pump switching resulting from the operational optimisation of the system. As Figure 6.19 shows, while a constant electricity structure clearly results in a non-restricted use of the pump, it is obvious on the contrary that a variable electricity cost leads to a restriction of the use of the pump. As a matter of fact, the use of several tariff zones also implies a deterioration in the respect of the volume setpoints in the demand reservoir. In this example, for the particular weights chosen, the penalty incurred from the electricity cost during the peak periods clearly outweighs any benefit of attempting to keep the reservoir volumes close to their setpoints. A good example of this can be seen on Figure 6.17 and Figure 6.18, by looking at the differences occurring in the evolution of the volume in Reservoir 2 when different electricity patterns are chosen. However, the use of the pump during peak periods can become a priority when it is used to avoid infringing one of the emergency volume constraints of one of the reservoirs.

This case study is also very interesting in the way that it highlights once again the anticipation phenomenon mentioned already during the presentation of the results relative to the calculated flows strategy. This can be seen on Figure 6.19 (configuration (b)), from the 11^{th} hour of simulation, when the pump switches on in the middle of a period when the electricity is billed at its peak rate. This preventive action is realised to respond to a problem of volume in Reservoir 2 which would only occur from the 19^{th} hour of simulation if nothing had been done earlier.

Finally, this case study shows once again that trade-offs often have to be made in the field of optimisation..

The only negative aspect of this study concerns the fact that it has not been possible to extend the simulation period to a two day period, as originally desired. This study has however made it clear that the problems of convergence encountered with the calculated flows strategy were not only due to the complex modelling choices that were made, but also probably to the presence of some weaknesses in the optimisation tools deployed initially.

6.3.3.2 Consideration of water quality objectives only

IN this second example, the purpose of the study is to highlight the ability of the available flows strategy to handle chlorine concentration problems. The case study is conducted with the 10-element configuration presented on Figure 6.1 and Figure 6.2. For this simulation, a high chlorine concentration was initially defined in the feed reservoir, while a lower chlorine concentration was chosen for the rest of the network.

By increasing greatly the value of the first-order chlorine decay constant in each storage reservoir, it is hoped to get quick changes in the chlorine concentration, despite the smaller time of simulation. Since water quality objectives are the priority, no restrictions are applied on the use of the pump: this is done by selecting a constant parameter for the electricity cost, as presented already in Figure 6.16 (configuration (a)). Similarly, although each storage reservoir is specified with a volume setpoint, no penalty is applied on the deviation from setpoint. However, the user demand flows and the minimum and maximum volume constraints must still be complied with.

6.3.3.2.a Definition of the data tables

THE data tables required for the constitution of the configuration constrained by water objectives only are presented in the CD attached to the thesis. The reader is invited to open the file entitled "10-element configuration – OSL – Water quality objectives data.m" which can be found under the following path:

\Small-scale applications\Available flows strategy\OSL solver\Water quality objectives

6.3.3.2.b Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, associated to the MILP solver OSL and the NLP solver CONOPT, at a time where a licence for the CPLEX solver was not yet available.

The parameters of the optimisation can be found in the MATLAB file already mentioned in the above .Furthermore, the options chosen for this particular configuration can be found in the files entitled "dicopt.opt", "conopt.opt" and "osl.opt" which are stored under the same location as the one mentioned in paragraph 6.3.3.2.a.

6.3.3.2.c Results

THE results of the optimisation for the 10-element configuration, solved with the available flows strategy, can be found in Biscos *et al.* (2003). For convenience, an electronic version of this document can be found on the CD attached to the thesis. This file is available under the name "WaterSA Paper.pdf" stored at the location already mentioned in paragraph 6.3.3.2.a.

This paper presents the optimised sequence of operations that was achieved with the available flows strategy, over a 24 h simulation period, with an 8 h control horizon, for this particular configuration.

The graphs presented in this document show that the optimised sequence has been achieved with respect to the constraints defined on the emergency volumes and concentrations in the reservoirs. A good example of this can be seen on Figure 6.22 and Figure 6.23,which represent respectively the evolution of the volume and of the chlorine concentration in Reservoir 2.

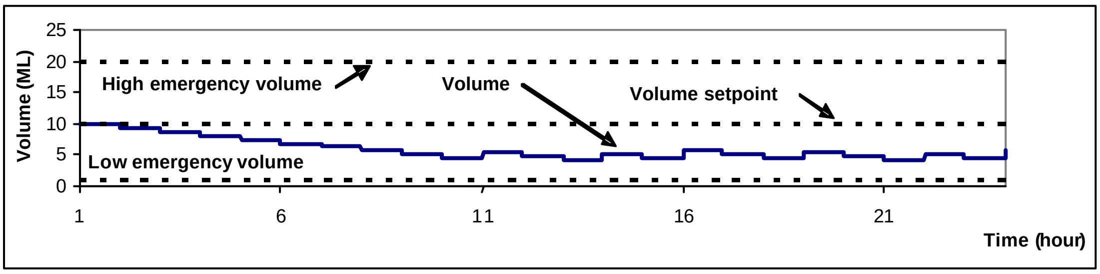


Figure 6.22: Evolution of the volume in Reservoir 2 (10-element configuration)

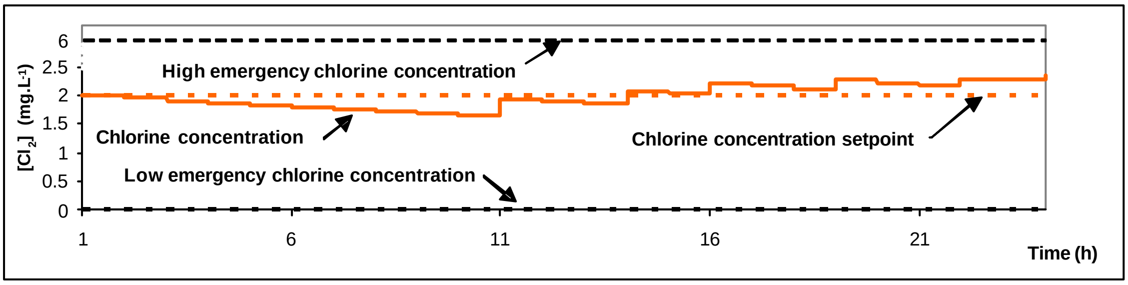


Figure 6.23: Evolution of the chlorine concentration in Reservoir 2 (10-element configuration)

A detailed description of the evolution of the volumes in the feed and demand reservoirs can be found in Biscos *et al.* (2003). The reader is invited to refer to this paper to find the explanations justifying these evolutions.

Furthermore, the optimised sequence gives an integer strategy for all the binary variables. A good example of this can be seen on Figure 6.24 which presents the evolution of the valve position of the binary valve located at the entrance of Reservoir 2.

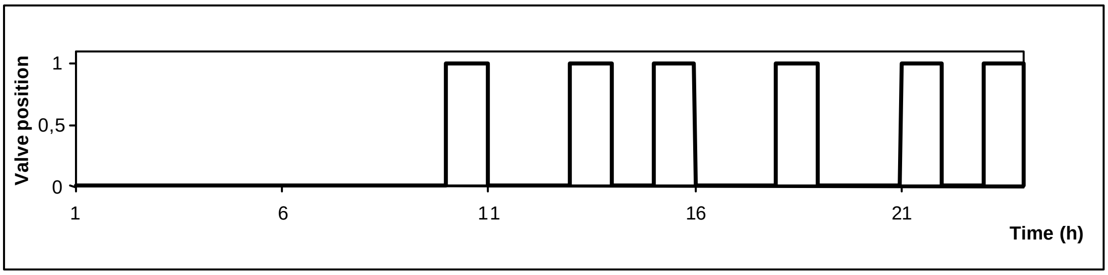


Figure 6.24: Valve position of the binary valve positioned at the input of Reservoir 2 (10-element configuration – constant cost)

The interesting aspect of these results is that there is a poorer tracking of the volume setpoints when chlorine considerations are introduced into the problem. This is seen easily seen by looking simultaneously at Figure 6.22 and Figure 6.23. Once again trade-offs have to be made, since the central objectives are in conflict with each other. Similarly to the previous example, it has not been possible to extend the simulation period to two days. This has confirmed that the problems of convergence encountered with the calculated flows strategy were linked to the presence of some weaknesses in the optimisation tools deployed initially. The next phase of work has therefore been to try to improve the convergence on the 10-element configuration by providing solvers that are more efficient and more adapted to the family of optimisation problems treated within this research project.

6.3.3.3 Concluding remarks about the use of the MILP solver OSL for the available flows

THE two experiments presented in the above paragraphs are good example of operational optimisation, since they were conducted on a generic type of water distribution network. Since optimised results have been reached for all three configurations, it has been possible to validate the modelling choices that were made during the development of the modelling strategy referred to as the available flows strategy. More importantly, these results have taken into account the constraints of each of the problems. Unlike the situation with the calculated flows strategy tested initially, integer strategies have been easily obtained for all of the binary variables. In addition, several interesting behaviours, like the anticipation phenomenon, or the existence of trade-offs between opposing objectives, have been highlighted.

However, it is obvious that these results can still be improved. Although the new modelling strategy appeared to be more adapted to tackle the class of operational optimisation problems applied to water distribution networks, it became clear that the solvers chosen initially were not necessarily the best suited. After the results presented in the above section were obtained, the GAMS Corporation was approached. Shortly thereafter, a temporary licence valid for 30 days and granting full access to all of the solvers included in the optimisation software was provided to the School of Chemical Engineering at the University of Natal in Durban. This allowed the testing of several other combinations of solvers, finally providing the available flows strategy with optimisation tools that are more efficient in the solving process. The results of this study showed that, while DICOPT++ and CONOPT could be validated as the most efficient solvers in their category, the solver OSL proved to be weaker than CPLEX in the resolution of the MILP sub-problems part of the MINLP optimisation. Consequently, the licence originally available for this solver was dropped in favour of CPLEX which was available in its version 7.5 at the time of the purchase.

6.3.4 Operational optimisation of the configuration based on the MILP solver CPLEX

THE initial tests conducted on the 10-element configuration with the available flows strategy using the MILP solver CPLEX showed tremendous improvements compared to what it was possible to obtain with OSL. For instance, in the two application examples presented in paragraph 6.3.3.1 and 6.3.3.2, it has been possible to work with a simulation period set at 48 h, and a control horizon extended back to 12 h. This was obtained without a single change in the data tables associated with these examples and, more importantly, all of the problems of convergence previously seen (for some of the time-steps), disappeared completely. In addition, the computational burden was reduced tremendously.

As a result, it was decided soon after to reintegrate the reservoir volume and water quality objectives within the same optimisation problem as originally described in paragraph 6.3.1.

6.3.4.1 Definition of the data tables

THE data tables required for the constitution of the configuration including both reservoir volume and chlorine concentration constraints and set-points are presented in the CD attached to the thesis. The reader is invited to open the file entitled “10-element configuration – CPLEX - Data.m” which can be found under the following path:

\Small-scale applications\Available flows strategy\CPLEX solver

The electricity cost structure is based on the use of time-varying parameters for the electricity cost, similarly to what was presented in paragraph 6.2.1.3. The configuration chosen for this particular example is presented on Figure 6.25.

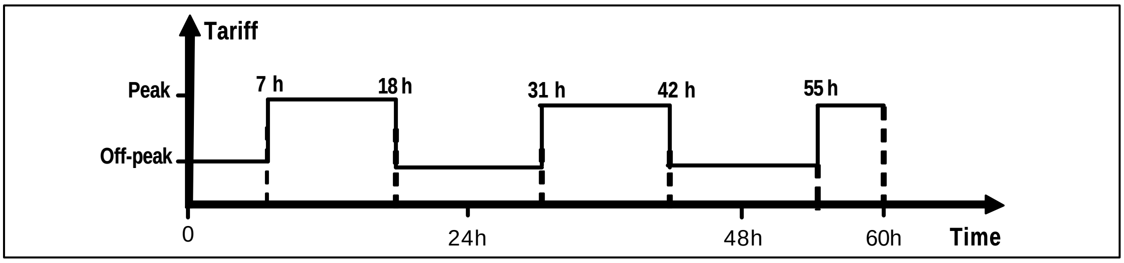


Figure 6.25: Available flows strategy/CPLEX – 10-element configuration: cost structure

6.3.4.2 Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, associated with the MILP solver CPLEX and the NLP solver CONOPT.

The parameters of the optimisation can be found in the MATLAB file already mentioned in the above paragraph. For more information on the significance of these parameters, the reader is invited to refer to the GAMS manual (Brooke *et al.*, 1998).

The options chosen for this particular configuration can be found in the files entitled “dicopt.opt”, “conopt.opt” and “cplex.opt”. These files are stored on the CD attached to the thesis, at the same location as the one already mentioned in paragraph 6.3.4.1.For more information on the significance of these options, the reader is invited to refer to the solvers manuals provided with the GAMS software.

6.3.4.3 Results

THE complete sequence of results of the optimisation for the 10-element configuration, solved with the available flows strategy, and constrained by emergency volumes in the reservoirs and water quality considerations, can be found in Appendix B. This appendix presents the optimised sequence of operations that was determined with the available flows strategy using the MILP solver CPLEX, over a 48 h simulation period, with a 12 h control horizon, for this particular configuration.

The graphs presented in this document are very similar to those presented in Biscos *et al.* (2002a) and mentioned already in paragraph 6.2.1.5. However, the major difference, compared to what was obtained for the same example with the calculated flows strategy, is that all of the binary variables evolve according to a complete integer strategy. This particular limitation of the original approach is the main reason that has led to the development of the modelling approach based on the concept of available flows. It appears with the new results presented in Appendix B that this was definitely a good decision.

It can also be observed that all of the interesting behaviours of the optimisation tool developed for this project (restriction of the use of the pump largely during off-peak periods, anticipation phenomena, trade-offs made to meet competing objectives) mentioned earlier, are here gathered within the same application example. Until the successful convergence on this particular application was reached, it is recalled that it had been necessary to develop several application examples to identify these behaviours separately. This is a major improvement which shows that the available flows strategy and the simultaneous use of the newly acquired CPLEX solver have allowed extending the field of application of the optimisation method a lot further than was initially possible.

6.3.5 Concluding remarks for the available flows strategy

THREE examples of optimisation problems have been presented in the above section. Several constraints relative to the emergency volumes in the reservoirs and the water quality aspects were included in the definition of the three problems. These constraints have played an important role in the determination of the optimal sequence of operations.

All of the problems were defined by building several data tables and defining numerous parameters according to what was presented in Chapters 4 and 5. Then, these problems were solved using the MINLP solver DICOPT++. Several MILP sub-problems needed to be solved during the course of the optimisation: this was made possible thanks to the newly acquired solver CPLEX. Similarly, the CONOPT solver has been required for the solution of the NLP sub-problems.

These results show that all of the constraints of each of the problems have been respected, particularly the achievement of an integer strategy for all of the binary variables included in the formulation. In addition, several interesting behaviours, like the anticipation phenomenon, or the existence of trade-offs between competing objectives, have been highlighted in the same way as previously found with the calculated flows strategy.

What should be remembered from the above paragraphs is that, unlike all of the applications examples presented before, the available flows strategy has allowed highlighting of all of these standard behaviours of the optimisation tool in the same application example. This has led to the validation and the confirmation of the available flows strategy as suitable for future modelling applications. This validation was awaited before the method could be applied to applications that are more realistic and of bigger scale.

6.4 Concluding remarks to the chapter

IN this chapter, the results associated with the two different modelling strategies developed during the course of this research project have been presented. While it has been impossible to get satisfactory results with the calculated flows strategy, even on the smallest application examples, the results obtained with the available flows strategy were entirely successful. It is clear that using a less complex approach has facilitated such a result. However, the use of a more effective MILP solver (at least for this problem), viz. CPLEX, should not be forgotten either.

The satisfactory results obtained with the available flows strategy have led to its choice for future modelling of water distribution systems, including a large-scale application involving more than 60 standard modelling elements.

In the next chapter, presentation is made of this large-scale application which is based on a significant portion of Durban's water distribution network. The modelling strategy based on the available flows concept is used to prove that there is room for the optimisation of the routine operations on the aqueduct providing Durban's southern area with potable water.

CHAPTER 7 – LARGE-SCALE APPLICATION OF THE AVAILABLE FLOWS STRATEGY

DURING the course of this research, an industrial application that required operational optimisation presented itself. It was thought at the time that this was an ideal opportunity to test the mixed-integer predictive controller on a much larger scale. This industrial application is part of the Durban water distribution network operated by eThekwini Water Services: it is referred to as the Southern Aqueduct. This scheme which supplies potable water to more than half a million people in southern Durban and surrounding areas was working close to maximum capacity at the end of the 1980's. This was due to the steady growth in water demand experienced in the area of supply, during the first twenty years of service of the aqueduct. The construction of a new scheme was therefore planned to augment the initial capacity of the system from 189 $ML.d^{-1}$ (1988 figures) to 307 $ML.d^{-1}$ (projection conducted in 1988 for the year 2005). The new scheme was phased in over a period of four years through a variety of contracts and the construction took place from November 1989 until the end of 1993.

However, this expected demand never materialised and, nowadays, the water demand seldom reaches 200 $ML.d^{-1}$. The over-design of this scheme is due to several factors. Firstly, during the 1980's, a rapid urbanisation was taking place with rural people moving into backyard shacks in the formal townships which were mushrooming at the time in the Durban southern areas (Umlazi area essentially). This resulted in an increase of the real water demand. Then, the poor maintenance of township reticulation system led to increased water losses through burst pipes and leaks. These two factors were responsible for the high growth in the general water demand which took place in the Umlazi area during the 1980's.

This growth was expected to continue at the same rate through the next decade. But during the 1990's, with the peaceful transition to democracy that took place nationwide, the Group Areas Act was abolished, allowing people to move out of the backyard shacks in the townships and to settle closer to work opportunities. At the same time, the Durban Municipality took over the administration of the township water reticulation system and repaired all the leaks. Finally, the HIV / AIDS epidemic started taking affect on population growth in the late 1990's. This has left eThekwini Water Services with a situation where they have excess capacity in the Southern Aqueduct, allowing the flexibility to consider various operating rules. Therefore, there is currently space left to achieve the operational optimisation of the system, with more permutations being possible than might be expected in a fully-utilised network.

Once it appeared that of the two modelling strategies proposed in Chapter 4, the available flows strategy was the only one that could be used with large-scale applications, the Southern Aqueduct was identified as the ideal target to test the behaviour of this modelling approach on an existing system.

This chapter presents an attempt at using Model Predictive Control (MPC) to provide the operational optimisation of an existing water supply scheme. Since the operational optimisation of the Southern Aqueduct was much more complex than expected, the attempt at using a mixed-integer predictive controller based on the concept of available flows was not pursued further than the off-line stage.

7.1 Process overview

THE Southern Aqueduct provides Durban southern areas with potable water. It has been identified as a good target application to test the ability of the modelling strategy developed during this project for the operational optimisation of water distribution systems. This system has been in operation since the early 1960's, but has undergone a major upgrade at the end of the 1980's. To respond to a strong increase in the general water demand, the initial capacity has almost been doubled. The upgrade has also been the occasion to review completely the electronic and telemetry control of the original system.

This section gives background about the Southern Aqueduct. Most of the information presented in this section is extracted from the paper prepared by Rodrigues and Carrie (1992).

7.1.1 Presentation of the initial configuration of the Southern Aqueduct

AS mentioned already in Chapter 2, the potable water supply to the Durban Metropolitan Area (DMA) is mainly provided from two purification works which are controlled by Umgeni Water. These waterworks, known as Durban Heights and Wiggins, abstract water from the Umgeni River. While Wiggins is located in the Cato Manor area, Durban Heights is situated in Reservoir Hills, above Durban. Durban Heights is the largest purification plant operated by Umgeni Water and it supplies Durban via the Northern, Central and Southern Aqueduct systems, as presented Figure 7.1. The purified water is purchased in bulk from Umgeni Water and is distributed throughout the city by eThekwini Water Services.

The Southern Aqueduct supplies more than half a million people and some large industrial areas in the southern environs of Durban.

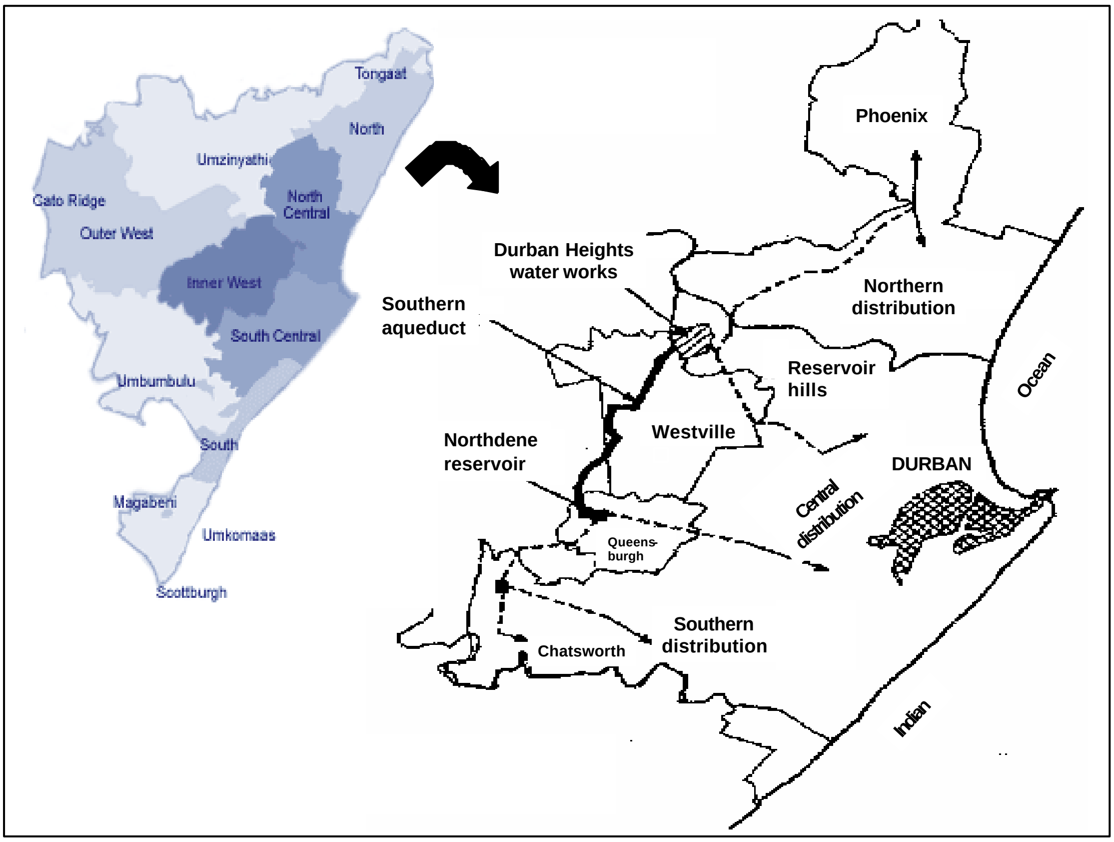


Figure 7.1: Localisation of the Southern Aqueduct within the Durban Metropolitan Area (courtesy of Rodrigues and Carrie, 1992)

It links Durban Heights to Northdene, Queensburgh, and Chatsworth, and allows the provision of water to the Umlazi and Amanzimtoti areas. The original aqueduct which was installed in the early 1960's, is equipped with a pre-stressed concrete pipeline, having a 965 mm internal diameter. It passes mainly through suburban residential areas, with a section traversing the Paradise Valley Nature Reserve. It is 9 km long.

The throughput of this main was originally equal to 128 $ML.d^{-1}$. Since sections of the southern zone of supply can also be fed by the central aqueduct system via a circuitous route, the initial configuration could give a total peak capacity of 189 $ML.d^{-1}$ for the south of Durban (1988 figures).

7.1.2 Upgrade process of the Southern Aqueduct

THE initial capacity of the Southern Aqueduct has been increased to 307 $ML.d^{-1}$ during the 1990's, following an upgrading phase of the initial configuration which took place between November 1989 and December 1993. At the time the augmentation of capacity was planned, the projected peak demand for the region was expected to be equal to 307 $ML.d^{-1}$ by the year 2005.

The projected demands for domestic water supply arising from the rapid urbanisation in the previously disadvantaged areas south of Durban was the main contributing factor for the expected increase in the water demand, while the growth in industrial water consumption was calculated to be of lesser significance.

Unfortunately, several factors were limiting the capacity of the original Aqueduct, the main one probably being the topographical barrier of Cowies Hill which the main has to traverse on its way south. Another limiting factor could be found with the tunnel which passes through this geographical ridge. Initial investigations considered the construction of a second tunnel near the existing one and the feasibility of avoiding the ridge by laying a new pipeline along a circuitous route. However, both these ideas were discarded for economic reasons. The solution that was finally implemented has consisted in laying a new steel pipeline in parallel to the existing one, in converting the existing tunnel to a pressure conduit and finally in constructing an "in-line" booster pump station in the Westville area.

With the observation that an augmentation of the Southern Aqueduct was required, an economic analysis was carried out by the Research and Development Department of eThekwini Water Services (formerly Durban Water) to consider the optimum combination of gravity-fed or booster-pumped flows associated with different pipe sizes and consequent civil, mechanical and electrical consumption costs. The most economic solution required the construction of a 1200 mm diameter duplicate steel aqueduct, combined with a booster pump station in Westville. In addition, an alternative pipeline route along Blair Atholl Road, in the Westville area, was chosen to provide a more favourable condition, from an hydraulic and accessibility point of view, than that of the existing main. In this section of the system, a 1400 mm diameter main was laid to replace the original 965 mm diameter concrete pipeline.

The cost of the Southern Aqueduct upgrading scheme was estimated in 1988 as being R47 million, with construction work phased over a 4-year period, starting in late 1989 and finishing at the end of 1993. It has been of paramount importance that the design and phasing of the new scheme corresponded with the continued supply of water to the southern areas of Durban, since this supply zone could only tolerate a 6 to 8 hour shutdown period on the supply side, before the system required replenishment.

The programming of the upgrading of the scheme has been divided into two different phases. The first phase concerned the augmentation of the Aqueduct on the upstream side of the proposed Westville pump station to prevent open-channel flow occurring in the aqueduct leading to the pump station.

This phase has involved the pumping of water over the Cowies Hill ridge via a temporary tunnel pipework system, for a period of about 16 months, while the original tunnel was being converted to a steel-lined pressure conduit. Once the pump station and tunnel bypass pipework were complete, the existing tunnel was taken out of commission for the required modifications, while the water continued to flow to the south via the tunnel bypass and pumps. Then, the second phase involved the installation of two high volume, low head pumps for pumping through the refurbished tunnel. This second phase criterion required that the refurbishment and conversion of the existing tunnel to a pressure conduit be commissioned at an early stage of the works. Indeed, this needed to be done during the years of relatively low flow through the tunnel bypass system, in order to keep pumping costs linked to the pumping of water over the Cowies Hill ridge to a minimum.

Figure 7.2 represents schematically the upgraded version of the Aqueduct, as it is known nowadays.

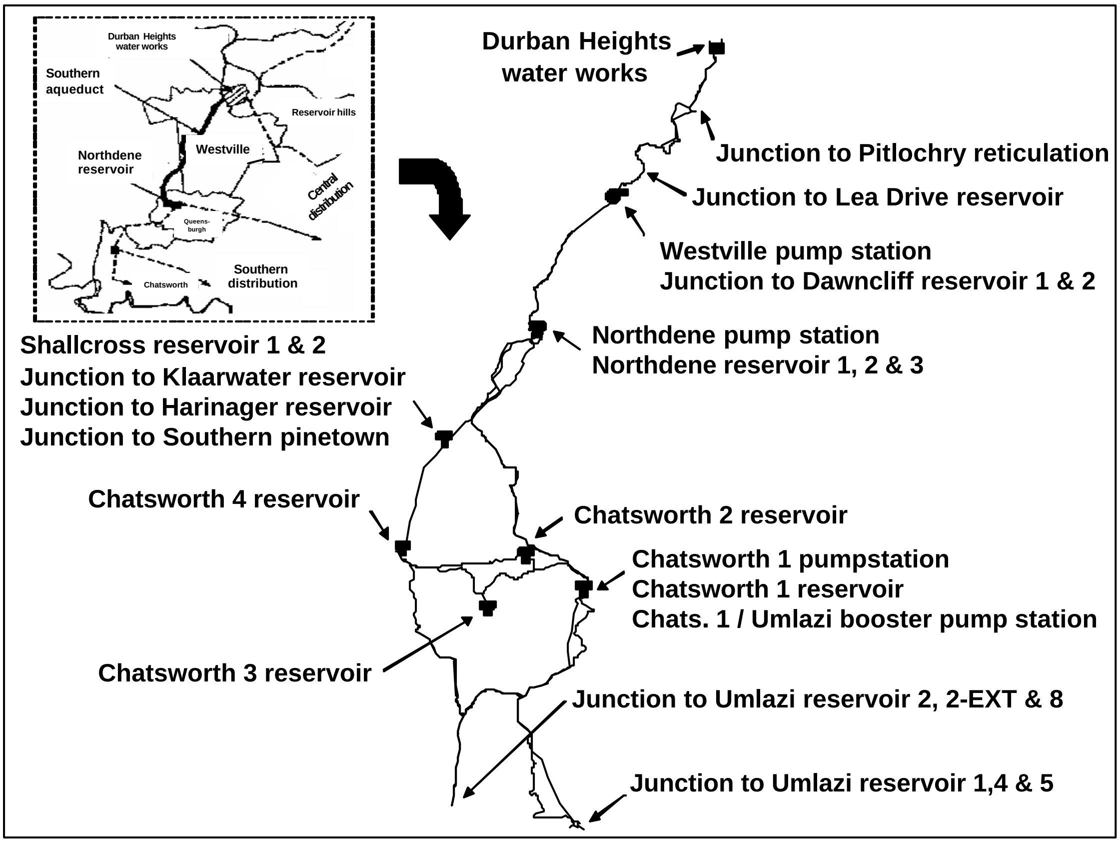


Figure 7.2: Schematic representation of the Southern Aqueduct

7.1.3 Past and present electronic and telemetry control of the Southern Aqueduct

ONE of the most important aspects of the management and operation of the revised Southern Aqueduct is the electronic and telemetry control of the different facets of the system.

eThekwini Water Services' existing telemetry system which links more than 60 telemetry outstations at the various reservoirs and pump stations, has been utilised on completion of the scheme to monitor and control the complete Southern Aqueduct system from the central control room at eThekwini Water Services head office in Central Durban.

Initially, the logic control diagram required for the first phase of the upgrade process allowed fully automatic control of the pumped system which fed the tunnel by-pass scheme. This control was carried out by a Programmable Logic Controller (PLC) situated at the Westville pump station which would automatically run the system on an instruction to commence pumping issued from the control room in Durban via telemetry. Numerous checks, controls and redundancies were built into the electronic equipment to retain automatic control of the system in the event of a fault occurring on the electronic hardware, or in the pumps or motors. An intercom communication system was also installed between all of the strategic equipment locations of the scheme so that an operating crew could run the entire control system manually should a major failure of the electronic hardware or computer software occurs. As the water supply via the Southern Aqueduct was of strategic importance during the upgrading process of the scheme, this requirement was essential in the event of a total failure of the automatic control. However, during the 16-month running period of the first phase of the upgrade of the Southern Aqueduct, this facility never had to be utilised.

The logic control system put in place during the second phase of the upgrading of the Aqueduct is simpler than that of the first phase, particularly at Westville pump station. However, with the interaction of the Westville logic control with that of the Northdene pump station, there is a requirement for the two PLC's at the remote pump stations to communicate with each other. This ensures that the required flows demanded in the south at the Northdene reservoir system are communicated to the valve and pump control at the Westville pump station. This link is essential due to the potential damage which could be caused to the Aqueduct should the coordination of operation between the two pump stations be lost. If the communication link should fail for any reason, the telemetry system will generate an alarm to alert the operator and, simultaneously, the PLC will immediately initiate a controlled shutdown of the entire system to prevent any possible damage.

A problem which had to be overcome concerned the linking of the new computer controlled system to the existing PLC and telemetry system. The problem of interfacing with the existing electronics and amending the software on the various existing control units has been overcome through the close liaison between the installers of the existing system and the contractors which have installed the new equipment.

The resulting system is based on a SCADA interface. All of the major elements constituting the Aqueduct are represented by a mimic graphic display. Figure 7.3 gives a good example of the look of the interface used for the representation part of the Southern Aqueduct within the supervisory control system. This figure shows the particular example of Westville pump station.

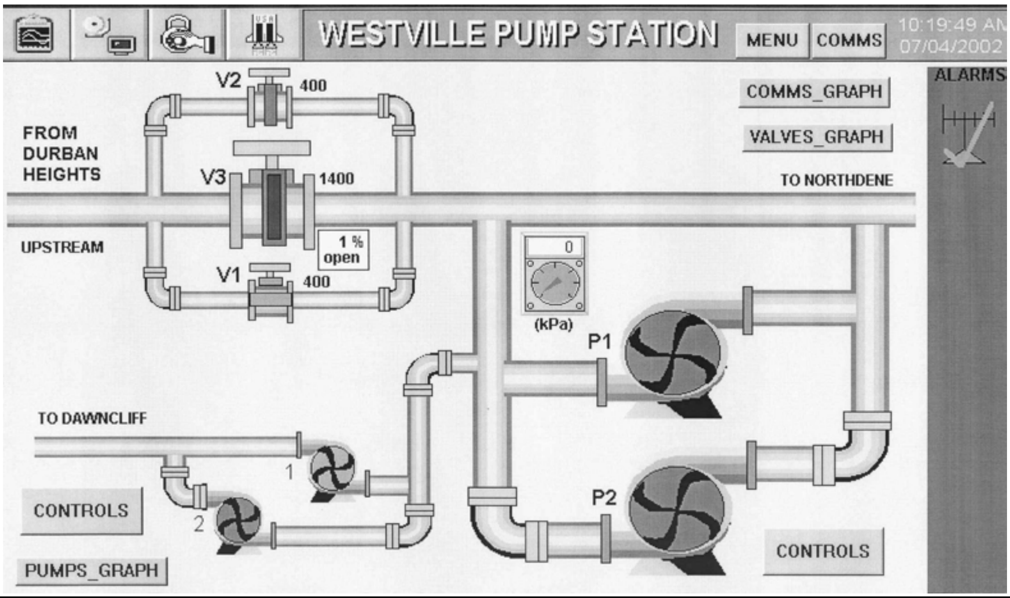


Figure 7.3: Representation of Westville pump station within the supervisory control system (courtesy of eThekwini Water Services, a non-return valve appears to be missing on the main line)

This "screenshot" shows that the supervisory control system represents accurately the geometry of the equipment. Here for instance, it shows that the pump station input is fitted by 2 binary valves (referred to as V1 and V2) and one continuous one (referred to as V3). Then the main line arriving from Durban Heights water works splits into two secondary pipelines: the first one leads to the pumps, while the second one bypasses the pump station. Just before entering the pumps, it can be observed than the flow is split again, since a portion of this flow is required for further use in the Dawncliff area. Then, the pump station is represented by two pumps, referred to as P1 and P2 on Figure 7.3.

Besides geometry, the SCADA system also displays, in real-time, the current state of the network. This feature is enabled by the telemetry equipment which is fitted as standard on all the major facilities of the supply system. The data are gathered in the central control room at eThekwini Water Services head office, before they are interpreted by the supervisory system. The example represented on Figure 7.3 for instance, shows that the first binary input valve (V1) at the entrance of the pump station is shut at the time the screen shot was taken, while the second one (V2) is fully opened. Finally, the continuous valve (V3) is positioned at a 1% opening. Colour displays are used to represent the state of the pumps. Hence, an online pump would be highlighted in green, while an offline one would be shown in red.

More importantly, it is also possible to access the controls of the system directly from the graphic sketch, using a mouse click on the buttons entitled "CONTROLS". This leads to a new screen on which different control options are available for the operators of the Aqueduct, depending on the conditions in which the global system is operated.

Let us have a look at the way reservoirs are represented within the system. Figure 7.4 gives the example of the Dawncliff reservoirs and reticulation system.

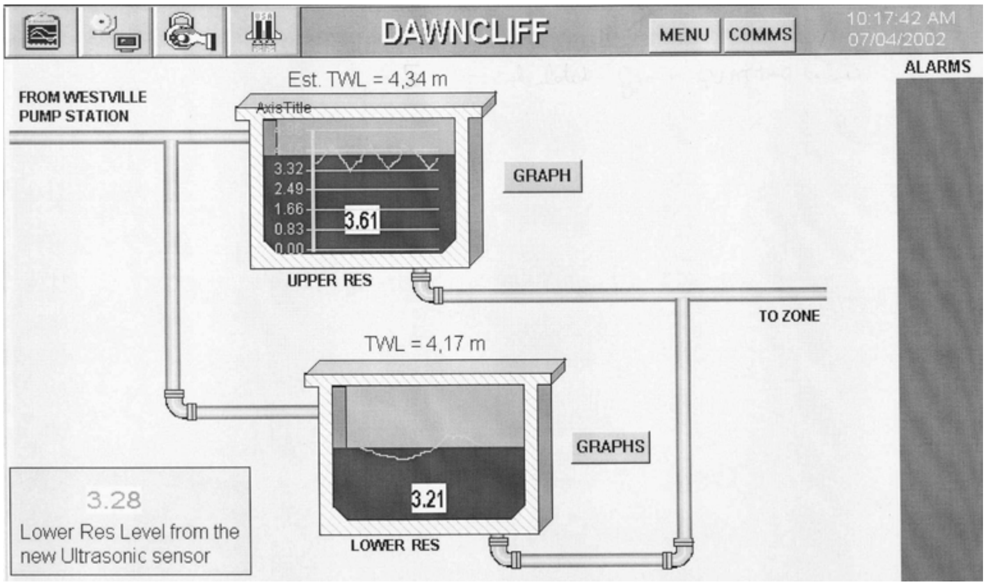


Figure 7.4: Representation of Dawncliff reservoirs and reticulation system within the SCADA system (courtesy of eThekwini Water Services)

This screenshot shows that the line arriving from the Westville pump station is split into two secondary pipelines, each of them feeding a reservoir separately. Then, the two reservoir output lines merge to convey the water to the Dawncliff reticulation system. Similarly to what was presented for the state of the pumps and valves on the snapshot referring to Westville pump station, the SCADA system is capable of displaying in real-time the amount of water stored within the reservoirs (parameter expressed as a level).

Figure 7.5 and Figure 7.6 present another feature of the supervisory control system which is that it displays an history of the state of the variables of interest. This is facilitated by logging the values of these variables at regular time intervals. This history can be accessed by a mouse click on the buttons on which the word "GRAPHS" appears. This leads to another screen where the logged data are plotted against a predetermined horizon. This feature is available for reservoir levels, as well as for the statuses of pumps.

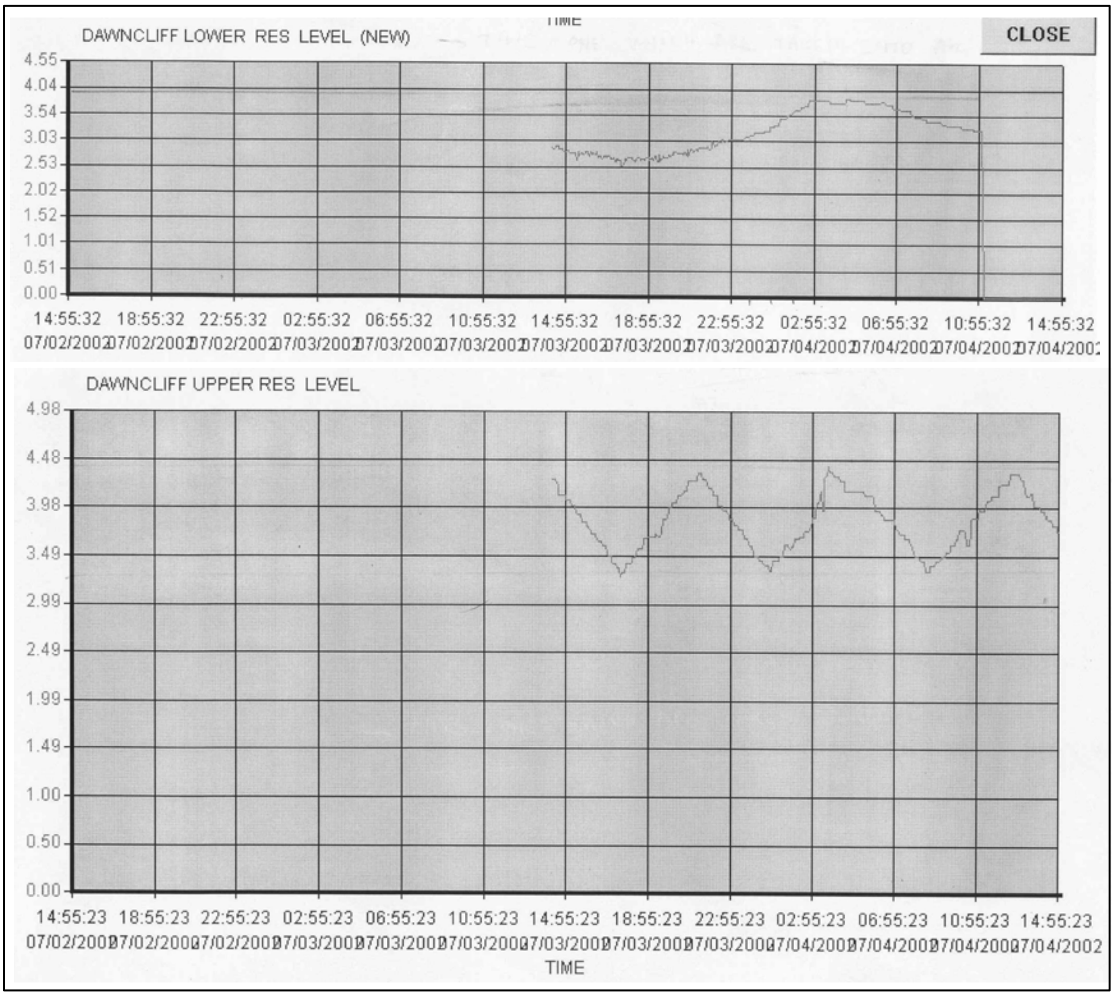


Figure 7.5: History of levels in Dawncliff lower and upper reservoirs given by the SCADA system (courtesy of eThekwini Water Services)

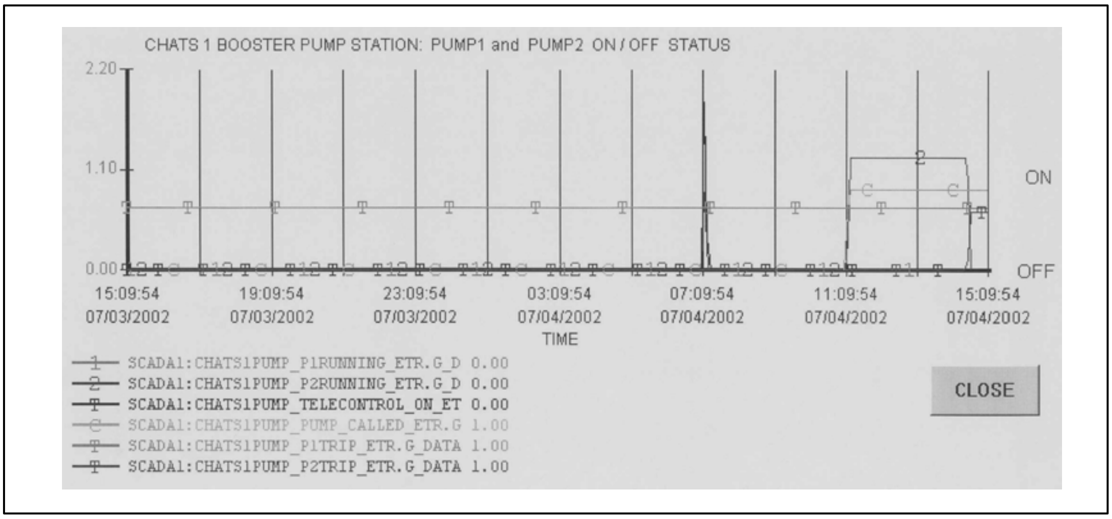


Figure 7.6: History of the pump status at Umlazi / Chatsworth 1 booster pump station (courtesy of eThekwini Water Services)

Similar displays can be obtained for the other reservoir systems and pump stations being part of the Southern Aqueduct. All of the additional displays can be found in Appendix C.

7.1.4 Current control strategy within the Southern Aqueduct

THE current control strategy applied to the Southern Aqueduct is relatively simple. Hydraulic considerations have led to the creation of a set of rules that prevent non-desirable configurations which could cause damage to the facilities. Furthermore, an analysis of the day-to-day operations has led to the development of empirical rules to operate the system in the best way possible, mainly taking into account the limitation of the pumping to electricity off-peak billing periods only.

These rules and constraints are now part of the supervisory control system which means that there is a pre-defined response for all of the known behaviours of the Aqueduct. For instance, when a pressure drop is detected in the Umlazi area, the response of the supervisory system is to switch on one of the pumps of the Chatsworth 1/ Umlazi booster pump station to see the pressure reaching normal levels again. Similarly, when the volume in Chatsworth 4 reservoir reaches its low-emergency level, one of the pumps located at Northdene pump station is started until the volume in the reservoir reaches a satisfactory level again.

Thanks to the telemetry system presented in the above paragraph, it is possible to keep track in real time of all of the variables of interest, and therefore, to determine the most suitable response to a particular event, according to the rules and constraints. Since they are based on empirical decisions, it is obvious that the current way of operation is not optimal. There is therefore space for optimisation on the Southern Aqueduct. This task is made easier by the situation of over-capacity which allows a greater degree of freedom for the development of new and more efficient operating rules.

7.1.5 Current modes of operation of the Southern Aqueduct

FIGURE 7.2 gives the schematic representation of the Southern Aqueduct. This Aqueduct uniquely does not require post-chlorination at any stage of the distribution of the water. This is explained by a short retention time in the reservoirs which guarantees good water quality, even at final delivery points. For instance, the retention inside the reservoirs located at Northdene never exceeds 12 h, even during low consumption periods. As a consequence, the chlorine concentration within the pipelines of the Aqueduct usually lies between 0.2 and 0.8 $mg.L^{-1}$, depending on the location and on the conditions of operation of the scheme (quality of the raw water entering Durban Height waterworks, levels of consumption, climatic conditions, etc…).

On Figure 7.2, it is obvious that the water can be directed to the different reticulation zones connected to the Southern Aqueduct via several routes. Each route represents a different way of operation of the scheme: currently, there are three different scenarios to operate the Southern Aqueduct. Several features are common to the three scenarios.

The main scenario is that the Westville pump station is by-passed at all times which means that the water arriving from the Durban Heights waterworks flows by gravity through two pumps that are not switched on anymore. Secondly, among the three valves positioned at the entrance of Westville pump station, the first one (referred to as V1 on Figure 7.3) is always open, the second one (V2 on Figure 7.3) is usually closed and only opened during high consumption periods, while the third one (V3 on Figure 7.3) is only opened when there is still a need of water when V1 and V2 are both open simultaneously. However, V3 is never opened more than 25 % since a greater opening would lead very quickly to a complete emptying of the reservoir at the output of the Durban Heights waterworks. Finally, in pump stations where several identical pumps are assembled in parallel, only one pump will usually be operated at a time. As far as possible, it is tried to alternate their use. Obviously, all scenarios share the same electricity cost structure, emergency volume constraints and user demand flows.

The first scenario consists in maximising the use of the gravity between the Durban Heights waterworks and the Umlazi reticulation system. The second scenario conveys the water from Durban Heights to Umlazi by transferring it via the Chatsworth 4 reservoir. Finally, the third scenario consists in simultaneously feeding Chatsworth 1 and Chatsworth 4 reservoirs with the water arriving from Northdene.

7.1.5.1 First scenario: maximisation of the use of gravity

IN this scenario, the use of the pumps is reduced as far as possible in favour of the use of gravity, where and when possible. To achieve this objective, the pump station located in the Northdene area is only used to supply Shallcross and Chatsworth 4 reservoirs with potable water. The water entering Chatsworth 4 reservoir does not go further than this reservoir, as shown on Figure 7.7.

In this scenario, the pumping taking place at the Chatsworth 1/ Umlazi booster pump only occurs in case of unexpected consumption or during emergency situations.

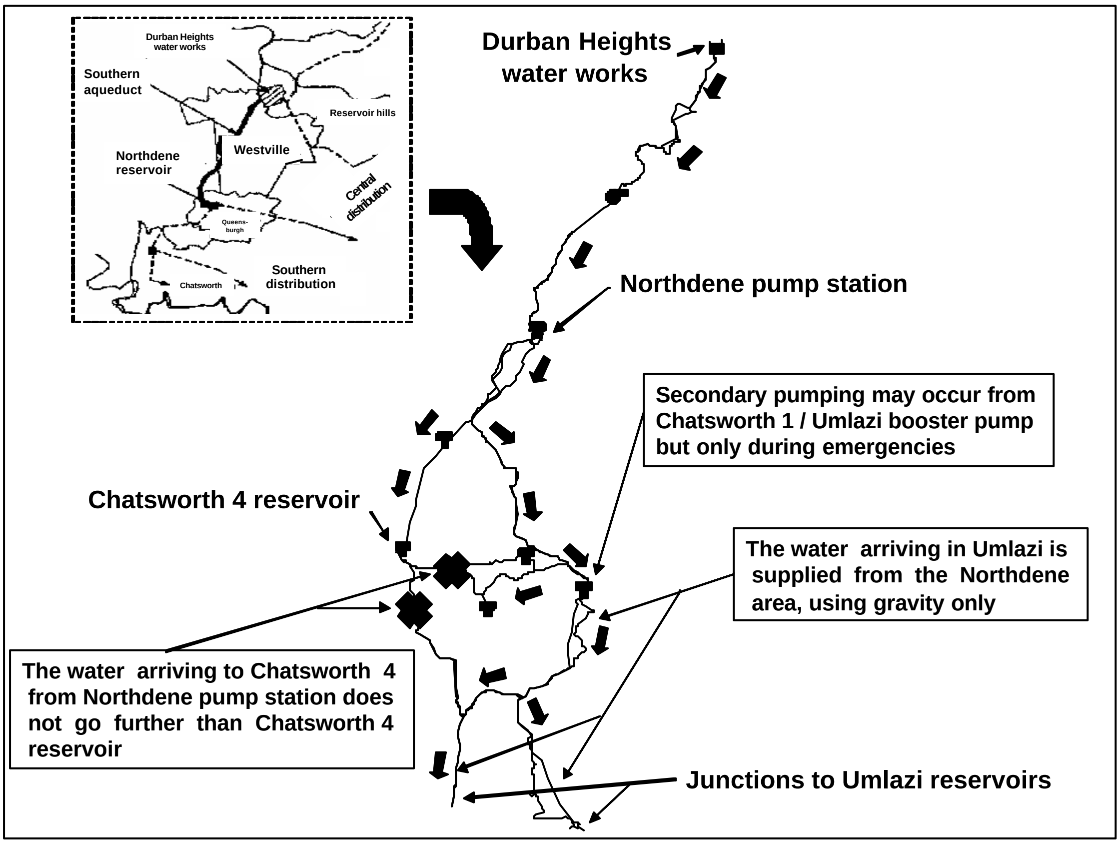


Figure 7.7: First scenario of operations of the Southern Aqueduct

7.1.5.2 Second scenario: Durban Heights to Umlazi via Chatsworth 4

IN this scenario, the water is provided to customers in the Umlazi area by transferring it via the Chatsworth 4 reservoir, thereby bypassing the two pump-stations located in the Chatsworth 1 area. In this scenario, the flows exiting Northdene reservoirs must be boosted to reach Chatsworth 4 reservoir with adequate pressure. However, once the water has reached this reservoir, it is conveyed until Umlazi by gravity. The circulation of the water during this scenario is summarised on Figure 7.8.

This scenario is very expensive in terms of pumping and results in having pumps switched on for 22 h per day, on average, at the Northdene pump station. As a result, it is only used in case of a problem on the pipeline linking the Northdene reservoir system with Chatsworth 1 reservoir, or during maintenance operations on this particular pipeline.

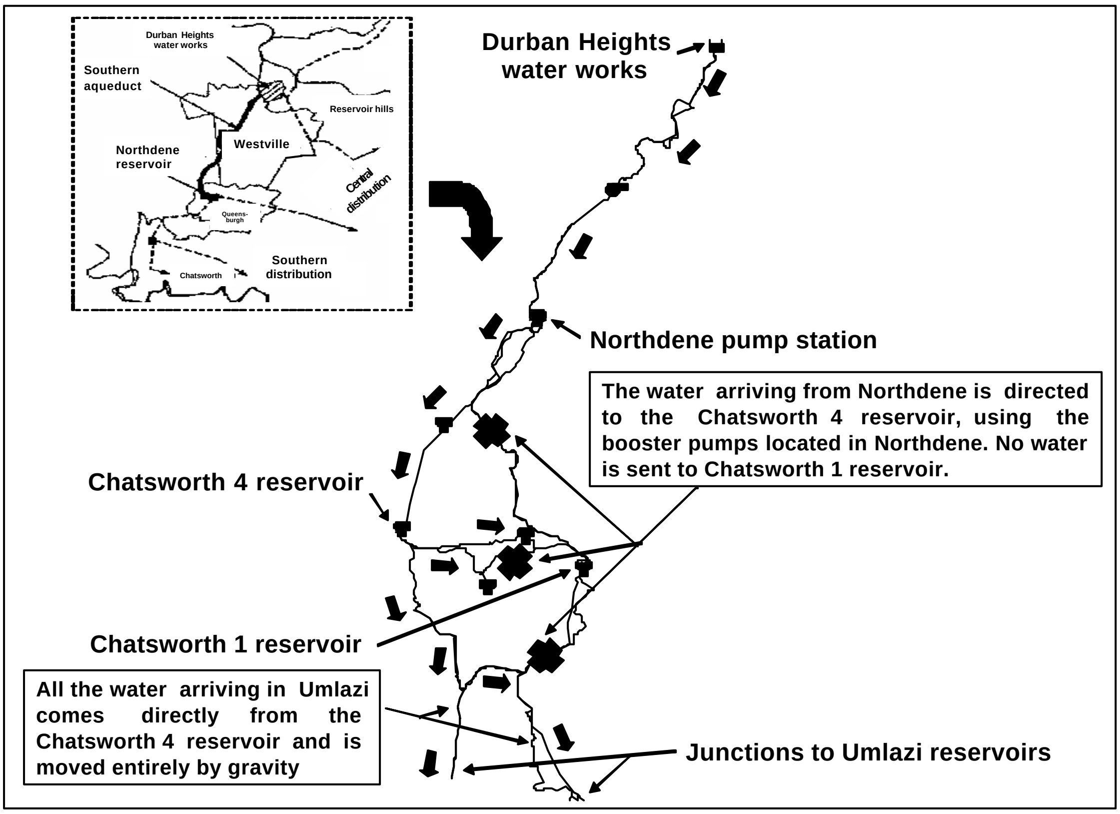


Figure 7.8: Second scenario of operations of the Southern Aqueduct

7.1.5.3 Third scenario: simultaneous alimentation of Chatsworth 1 and 4 reservoirs from Northdene

IN this scenario, the water flow arriving from Durban Heights is directed towards separate destinations from the Northdene reservoir system, so that it is possible to provide a sustained supply in Chatsworth 1 and 4 reservoirs. The water circulation taking place with this particular scenario is presented on Figure 7.9.

This scenario is very similar to the first one presented in paragraph 7.1.5.1: the main difference concerns the use of the Chatsworth 1 / Umlazi booster pump station. In this scenario, its use is not restricted anymore, and the pumps switch on alternatively when the volume in Umlazi 2 reservoir becomes too low, or when the pressure in the system drops significantly.

7.1.5.4 Choice of one of the scenario suitable for operational optimisation

THE first scenario of operations of the Southern Aqueduct is interesting in the way that it reduces the number of pumps that must be used to convey the water as far as the Umlazi area. This explains why this scenario, represented by Figure 7.7, was kept as the reference for the work of operational optimisation conducted on this particular scheme of the Durban water distribution system and presented in the following section.

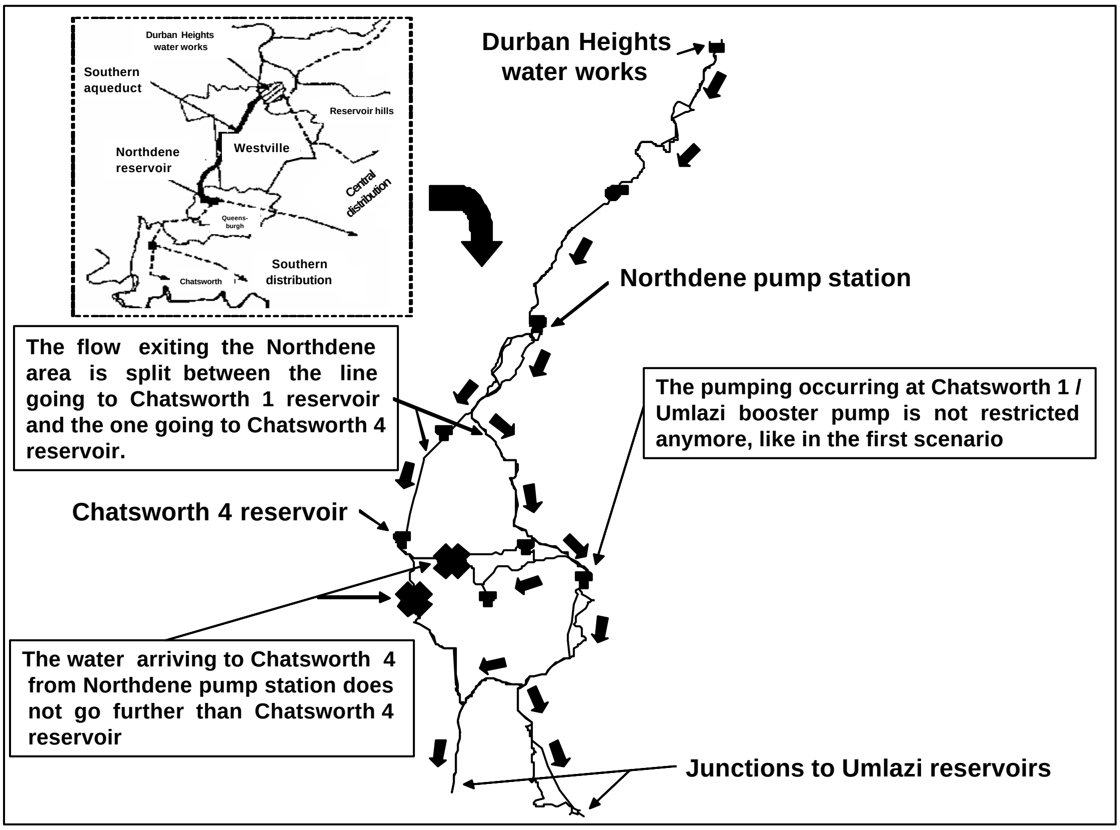


Figure 7.9: Third scenario of operations of the Southern Aqueduct

7.2 Operational optimisation of the Southern Aqueduct

AFTER the Southern Aqueduct was identified as a perfect large-scale application of the available flows modelling strategy, it has been necessary to recreate this application within the model, according to all of the principles presented in Chapter 4. To constitute all the data tables required for the operational optimisation of the system, a large amount of data collection had to be undertaken. Fortunately, an electronic version of the Southern Aqueduct, under the ePANET format, was available at eThekwini Water Services. This file gives a very accurate representation of the Southern Aqueduct as it is known nowadays. It is mainly used to test the hydraulic feasibility of the different operating scenarios of the scheme, or test the influence of any maintenance operations on the hydraulics of the system. However, it does not give any hints to the operators of the system about how to operate it as close to optimality as possible. Hence, it appeared immediately that ePANET was the perfect tool to access all the data required, before an optimisation process could be launched.

After construction of a version of the Aqueduct under the MATLAB format, the operational optimisation of the distribution system has been conducted. The main objective here was to find a more economical pumping sequence to make best use of the particular electricity cost structure used for the Southern Aqueduct.

As already mentioned in paragraph 7.1.5, no post-chlorination is required after the potable water has left the reservoir positioned at the output of Durban Heights waterworks. Therefore, water quality objectives have no real impact in this particular example of operational optimisation, and therefore they were dropped in favour of objectives concentrating on the satisfaction of volume setpoints and the respect of low / high emergency volumes. In other words, the results given by the optimisation software had to only comply with users' demand flows, as well as all of the constraints specified for low and high emergency volumes.

All of the simulations were conducted over a period of 48 h, using a control optimisation horizon of 12 h. Therefore, according to the Model Predictive Control (MPC) theory presented in Chapter 3, it was necessary to provide data tables with 60 h of known time-varying coefficients for electricity cost and demand flows. This allowed the algorithm to look ahead at future demands over the full period.

7.2.1 Reconstitution of the Southern Aqueduct within the modelling strategy

THE Southern Aqueduct has been created within the model according to the set of rules defined for the available flows strategy, in Chapter 4. Main and auxiliary data tables have been filled with all of the data available in the ePANET file available from eThekwini Water Services. A copy of this file has been included in the CD attached to the thesis. It is entitled "Southern Aqueduct.net" and can be found at the following location:

\Large-scale application\Available flows strategy

This file however does not include all of the elements represented within the supervisory control system, elements which are summarised in Appendix C. Since these elements are part of the Southern Aqueduct, it was nevertheless decided to include them in the MATLAB file used to create a version of the Southern Aqueduct compatible with the available flows strategy. As it was not possible to obtain these data in a reasonable amount of time from the eThekwini Water Services staff, it has been decided to take arbitrary values for all of the missing parameters, mainly represented by pump characteristics. This was done in an attempt to prove the ability of the available flows modelling approach to provide the operational optimisation of a large-scale distribution system.

It is therefore obvious that the results that are presented further in this chapter are not based on a calibrated version of the Aqueduct, and unfortunately, some additional work of calibration would be required before new and more economical operating rules can be presented to eThekwini Water Services management.

This explains why in the analysis of the results that is made later in this Chapter no attempt will be made to compare the pumping sequence used currently, with the new and optimal one proposed by the optimisation package.

7.2.1.1 Presentation of ePANET

MOST of the information contained in this paragraph is extracted from the ePANET users manual (Rossman, 2000), published by the United States Environmental Protection Agency.

ePANET is a public domain software package accepted as the world standard in hydraulic and water quality modelling of water distribution systems. It is a computer program developed by the United States Environmental Protection Agency which performs extended period simulation of hydraulic and water quality behaviour within pressurised pipe networks.

Under ePANET, a network consists of pipes, nodes (corresponding in fact to pipe junctions), pumps, valves and storage tanks or reservoirs. The software tracks the flow of water in each pipe, the pressure at each node, the height of water in each tank, and the concentration of a chemical species throughout the network during a simulation period comprising multiple time steps. Running under Windows, ePANET provides an integrated environment for editing network input data, running hydraulic and water quality simulations, and viewing the results in a variety of formats. These include colour-coded network maps, data tables, or time series graphs.

ePANET is designed to be a research tool for improving the understanding of the behaviour of drinking water constituents within distribution systems. It can be used for many different kinds of applications in distribution system analysis. Sampling program design, hydraulic model calibration, and chlorine residual analysis are some examples. Although ePANET can help assess alternative management strategies for improving water quality throughout a system, it is not an optimisation tool. This explains why a separate optimisation technique had to be developed in parallel with ePANET to achieve the operational optimisation of the Southern Aqueduct.

7.2.1.2 Detailed representation of the system

A general representation of the Southern Aqueduct is presented by Figure 7.2. However, to understand the contents of the data tables used to represent the Aqueduct within MATLAB, and presented in the next paragraph, it is necessary to have a closer look at the infrastructure of the Southern Aqueduct.

Therefore, all of the points of interest highlighted with an arrow on Figure 7.2 have been detailed in Appendix D, so that it is possible to determine which equipment needs to be included in the data tables of the MATLAB file used to represent the Aqueduct within the optimisation tool. All of the schematics presented in Appendix D are extracted from the ePANET file used to model the Southern Aqueduct at the Research and Development Department of eThekwini Water Services.

As mentioned already in the introductory part of paragraph 7.2.1, some of the elements of the Southern Aqueduct are not represented within the ePANET file. The decision was nevertheless taken to include all of these missing items in the optimisation process, although it has often meant the estimation of data applying to these. Furthermore, sometimes, the ePANET file represents elements in a different way to that adopted by the SCADA supervisory system. Therefore, several modelling choices had to be made during the phase of construction of the data tables of the MATLAB file used to represent the Southern Aqueduct. The main differences that were found between the actual network and the modelled one, as well as all of the modelling choices required, are presented in Appendix E.

The final version of the Southern Aqueduct that conforms to these modelling choices includes 14 reservoirs, 4 pump stations equipped each with a single fixed-speed pump, 8 binary valves, 10 continuous valves and 15 connections with the reticulation system. This version is kept as a reference for the construction of the data tables required by the available flows strategy.

7.2.1.3 Definition of the data tables

THE data tables required for the construction of the version of the Southern Aqueduct compatible with the modelling choices detailed in Appendix E are presented in the CD attached to the thesis.

The reader is invited to open the file entitled “Southern Aqueduct.m” which can be found at the location already mentioned in the introductory part of paragraph 7.2.1. This file shows that the representation of the Southern Aqueduct according to the principles of the available flows strategy is obtained by combining 67 standard modelling elements together.

In the main data table for this simulation, all of the columns relative to chlorine (setpoint, emergency levels, initial value) are filled with zero values, since all water quality objectives were excluded for the Southern Aqueduct. Furthermore, as mentioned in Appendix E, all the pumps are represented by a SULZER SM 402 – 720 (1490 RPM), for which a pump characteristic was downloaded from the vendor’s website.

The electricity cost structure chosen for this particular example is represented on Figure 7.10.

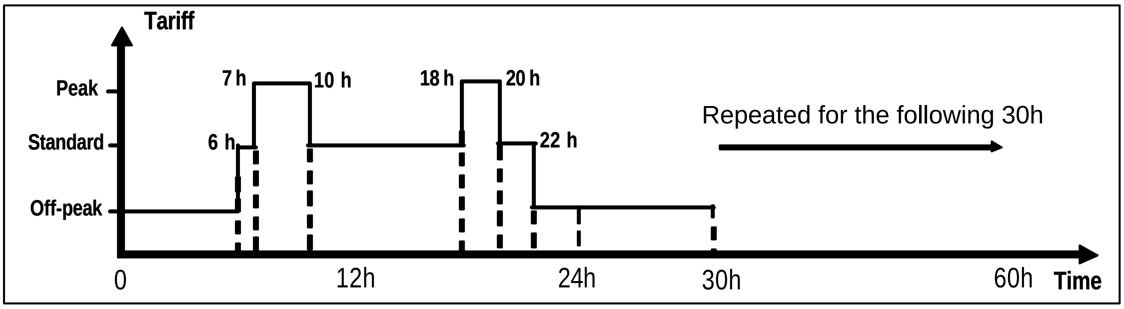


Figure 7.10: Electricity cost structures used with the Southern Aqueduct configuration

This tariff structure, referred to as the “true electricity cost”, is based on the bulk supply agreement offered to industrial customers by Durban Metro Electricity. It is equivalent to having a 3-zone tariff structure divided into off-peak, standard and peak periods. While the off-peak power is charged at R 0.09610.kWh^{-1}, the standard power costs R 0.150009.kWh^{-1} and the peak power is billed R 0.268030.kWh^{-1} (2002 figures, courtesy of Durban Metro Electricity).

In other words, with this tariff structure, the normal power is one and a half times more expensive than the off-peak power and the peak power is almost three times more expensive than the off-peak power.

7.2.2 Choice of the solver and definition of the parameters of the optimisation

THIS particular example has been solved with the MINLP solver DICOPT++, associated with the MILP solver CPLEX and the NLP solver CONOPT. The parameters of the optimisation can be found in the file entitled “Southern Aqueduct.m”, already mentioned in the introductory part of paragraph 7.2.1. The solver options chosen for this particular application can be found at the same location, in the files entitled “dicopt.opt”, “conopt.opt” and “cplex.opt”.

7.2.3 Results

THE result tables relative to the optimisation of the Southern Aqueduct configuration, solved with the available flows strategy, with no considerations on water quality, can be found on the CD attached to the thesis, in the EXCEL file entitled "Southern Aqueduct final – True cost.xls". This file, stored at the location mentioned in the introductory part of paragraph 7.2.1, gives the optimised sequence of operations that was achieved with the available flows strategy on the configuration presented in Appendix D. The simulation was conducted over a 48 h simulation period, with a 12 h control horizon.

An attentive look at these result shows that the optimised sequence has been achieved with respect to the constraints specified on the emergency volumes in the reservoirs, as well as the user demand flows.

Let us look at the behaviour of the pumps in each of the four pump stations constituting the Aqueduct to evaluate the performances of the optimiser on the aspect of operational optimisation. The first pump station of the supply system is located in the Lea Drive area, as summarised on Figure 7.11.

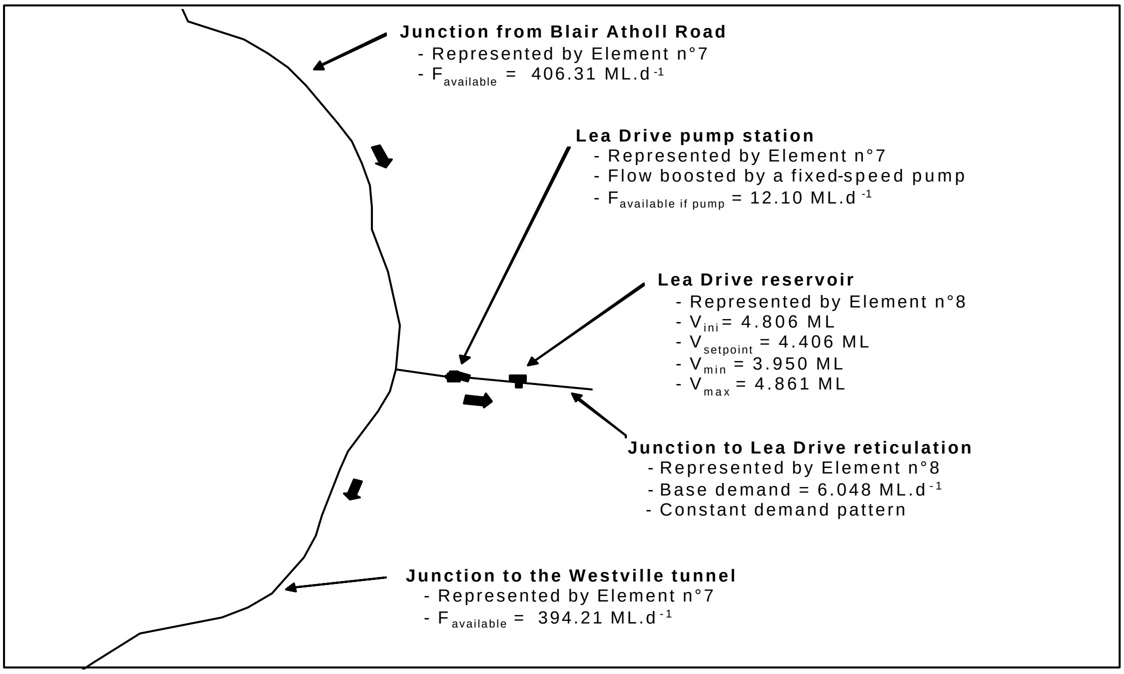


Figure 7.11: ePANET representation of the Lea Drive reservoir and reticulation system

In this configuration, this element of the supply system is represented by a fixed-speed pump which needs to be switched on to provide the reservoir with water. The figure also shows that the maximum available flow supplied to the reservoir is equal to 12.10 ML.d^{-1} and that the user demand out of this reservoir is constant and equal to 6.048 ML.d^{-1}.

The behaviour of the Lea Drive pump for the "true electricity cost" tariff structure is presented on Figure 7.12.

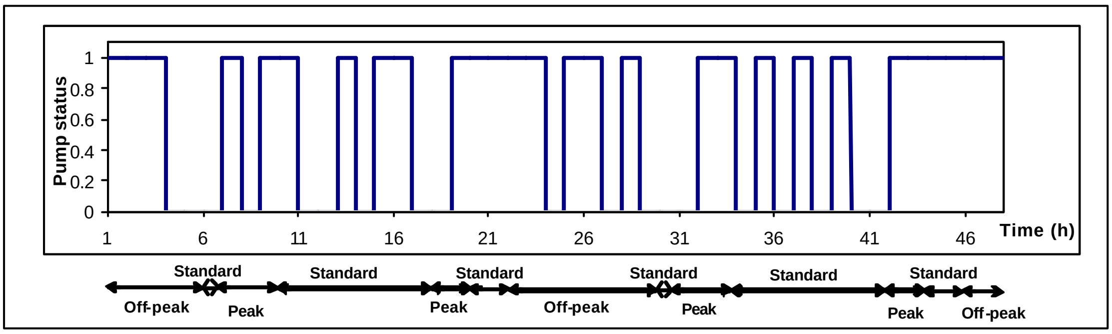


Figure 7.12: Pump status at the Lea Drive pump station for the true electricity cost structure

This figure shows that the use of the pump is as far as possible reduced with the tariff based on the bulk supply agreement, since in this configuration it is only working for 28 h over the 48 h of simulation. Let us look in detail at the behaviour of the pump for this particular electricity structure. This analysis can only be conducted by having a simultaneous look at the evolution of the volume in Lea Drive reservoir (Figure 7.13).

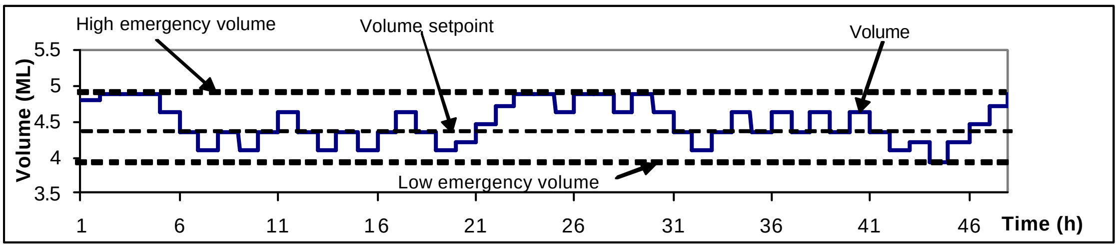


Figure 7.13: Evolution of the volume in Lea Drive reservoir (true electricity cost)

The pump is on for the first four hours of the simulation: this is to make best use of the off-peak electricity and fill the reservoir while the electricity is cheaper. However, since the volume reaches its top limit before the electricity tariff is switched to standard, the pump has to be turned off to respect this particular constraint. The pump remains in the off position until the 7th hour, where it is used again for one hour. It can be observed at this point that the electricity tariff is charged at the peak rate. Therefore, turning the pump on at this particular time can appear odd, but there is unfortunately no choice since the volume has become very close to its low-emergency constraint. This action brings the volume closer to its setpoint which justifies stopping the use of the pump. However, the demand out of the reservoir is steady, and interrupting the reservoir feed, by turning the pump off, causes an immediate a drop in the volume. This makes it necessary to switch on the pump again at the next time-step, although the tariff is still at the peak rate. This on / off situation goes on until the 19th hour of simulation and, inside this interval, the optimiser needs to find a compromise between the amount of electricity consumed and the respect of the low-volume constraint. Then, the pump remains on until the end of the 23rd hour of simulation.

This is to make best use of the cheaper electricity which is charged in standard or off-peak rates. However, since the volume is quickly corrected by the use of the pump to reach its high emergency constraint, the pump must be turned off, although at the end of the 23rd hour the electricity cost still belongs to the off-peak tariff. When the second day of simulation starts, the situation in the reservoir is very similar to that at the beginning of the simulation. Therefore, the pump behaviour and volume evolution over the next 24 h are much the same as has just been described for the first day of simulation.

To summarise the results available from the Lea Drive reservoir and reticulation system with the true electricity tariff, the pump is used mainly during off-peak and standard periods, meaning that overnight pumping is favoured as far as possible. However, the pump may have to be turned off during low-cost periods when the volume reaches its high emergency level. This is clearly a limitation, since more water could be pumped into the reservoir with off-peak electricity. Similarly, the pump may have to be switched on during peak periods to respond to a lack of water in the reservoir. Thus, there may be an advantage in revising all of the emergency constraints of the reservoir should a margin exist.

Let us now describe the pump behaviour at the next pump station located in the Northdene area (Figure 7.14). This behaviour is quite different from that occurring at Lea Drive, as shown in Figure 7.15. The results obtained here are indeed less oscillatory. This can be explained, since the pump here is an in-line pump that is used to boost the flow leaving the Northdene reservoir system, until the remote areas of the water supply system (Shallcross, Harinager drive, Klaarwater and Chatsworth 4 reservoirs). On the contrary, the pump used at Lea Drive was only used to sample water out of the main distribution trunk and to convey it to a unique reservoir of small capacity.

Furthermore, the water going out of Northdene pump station can be directed to several different areas. As a result, when one of the reservoirs reaches its high emergency volume, there is no need to switch off the pump, if it appears that the electricity is still charged at the off-peak rate. The water may indeed be conveyed to one of the other reservoirs of the area. Therefore, for this particular pump station, the results show that the proposed optimal pumping sequence makes best use of overnight pumping, with the pump rarely switching on during peak and standard periods. From an operational point of view, since the retention time in all the reservoirs of the Southern Aqueduct is low, the reservoirs empty rather quickly: therefore, this explains why the pump in the facility of Northdene pump station has to be working for 30 h over the 48 h of the simulation.

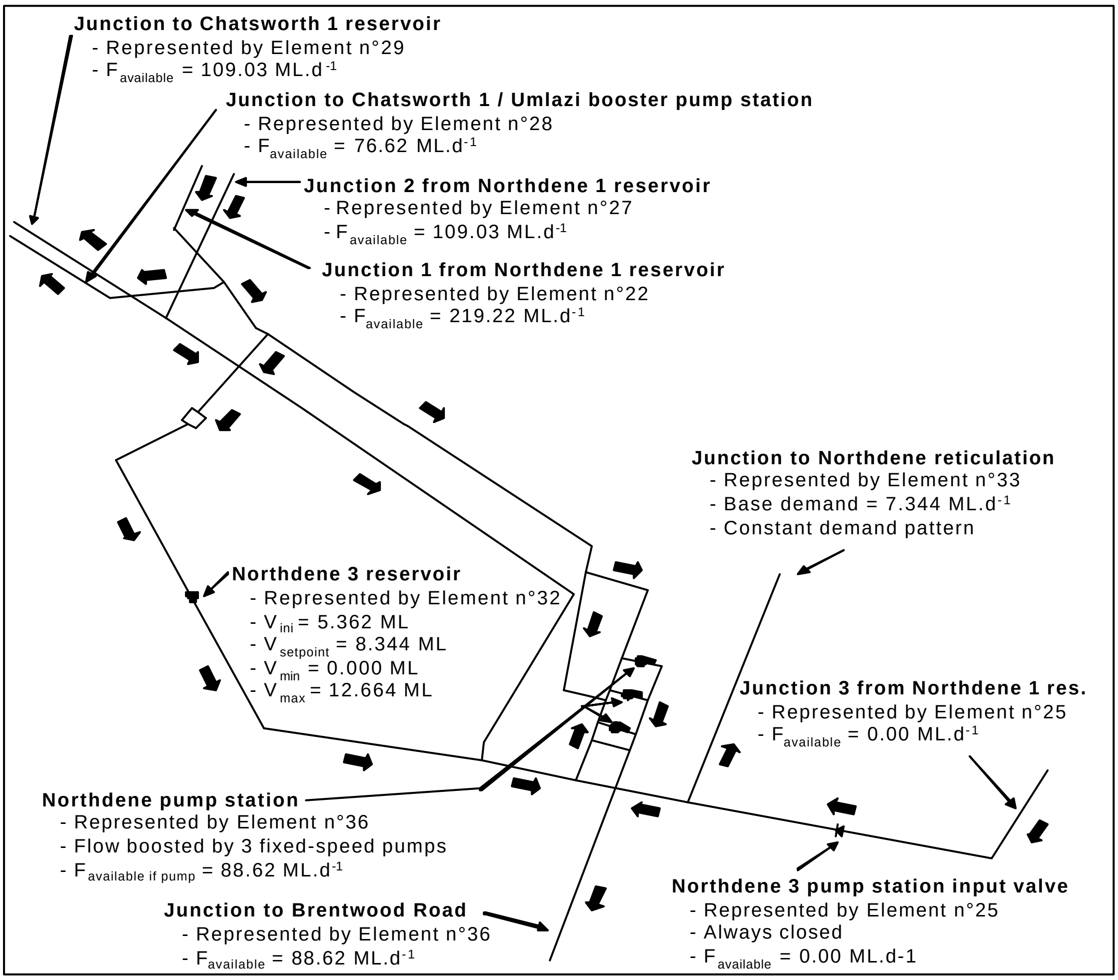


Figure 7.14: Northdene 3 reservoir, pump station and reticulation system

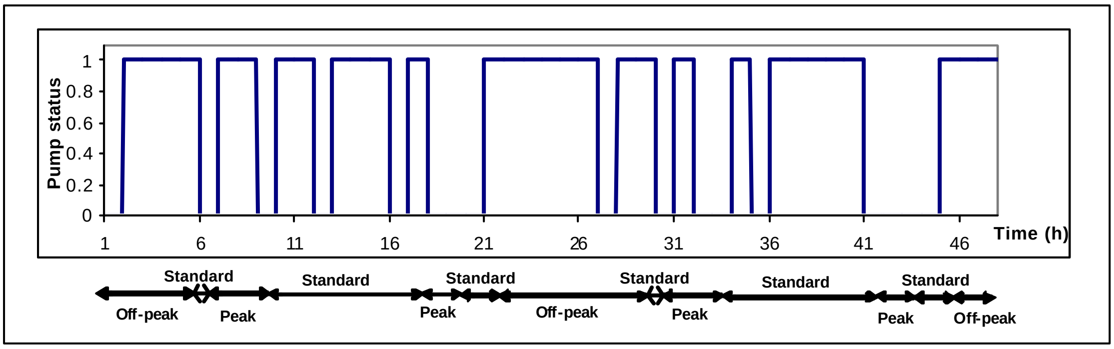


Figure 7.15: Pump status at the Northdene pump station for the true electricity cost structure

The following pump station is located in the facility of Shallcross, just before the input of Klaarwater reservoir, as presented on Figure 7.16. Once again, the behaviour of this pump is different from those already described for the Lea Drive and Northdene areas, as presented Figure 7.17 .

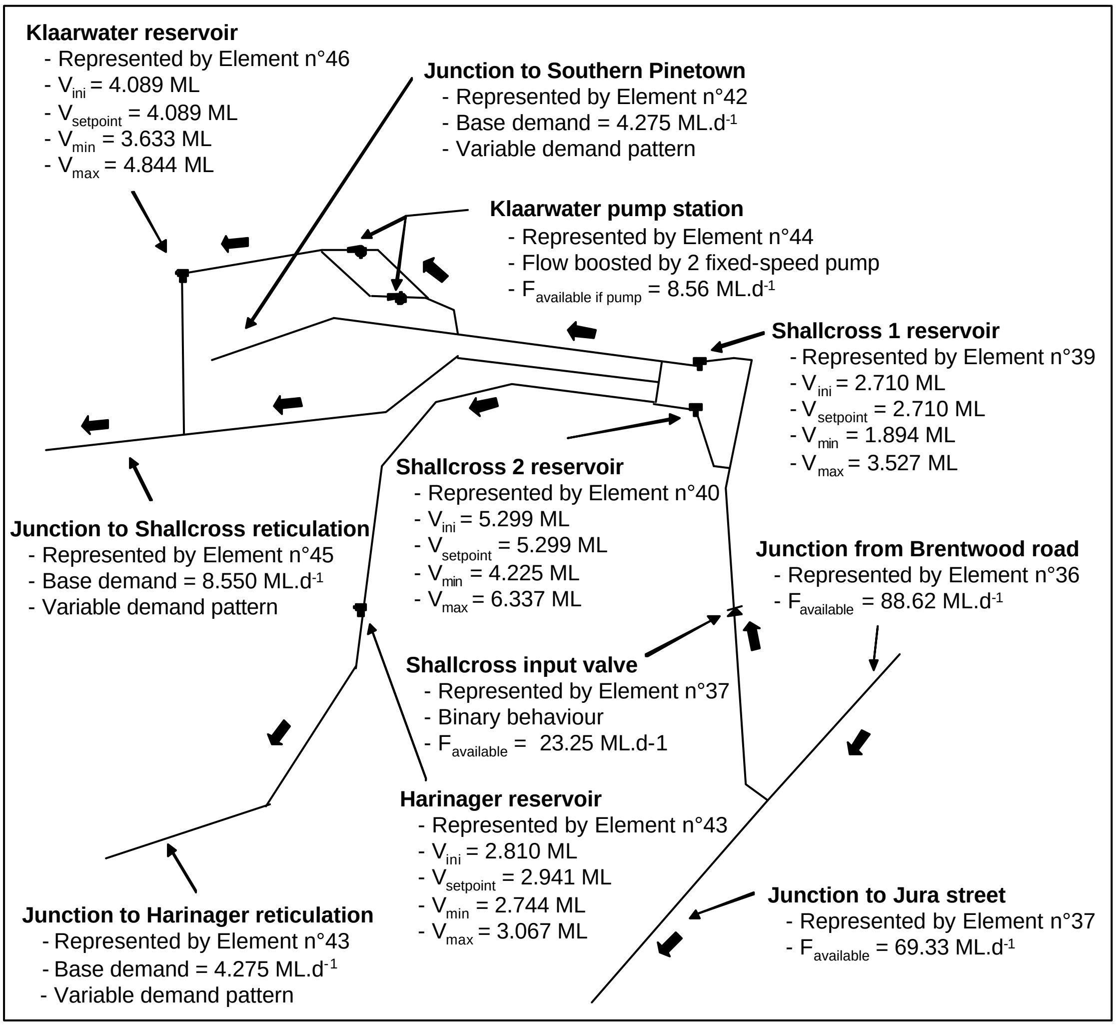


Figure 7.16: Shallcross area reservoirs and reticulation system

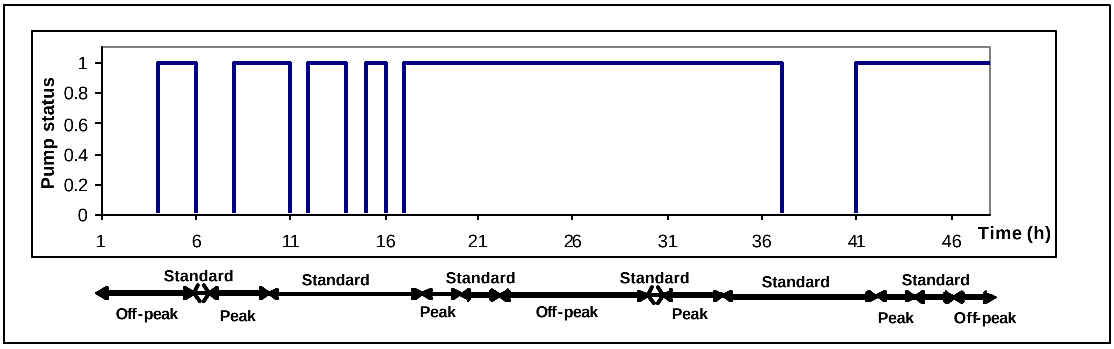


Figure 7.17: Pump status at the Klaarwater pump station for the true electricity cost structure

As shown on Figure 7.17, the use of the tariff structure conforming to the bulk supply agreement gives a pumping sequence that is even less oscillatory than that of Northdene pump. Since this particular behaviour is linked to having Klaarwater pump draw water out of a main line to convey it to a storage reservoir, the analysis of the optimised sequence of operation in the facility of Klaarwater can only be conducted by looking simultaneously at the evolution of the volume in Klaarwater reservoir (Figure 7.18).

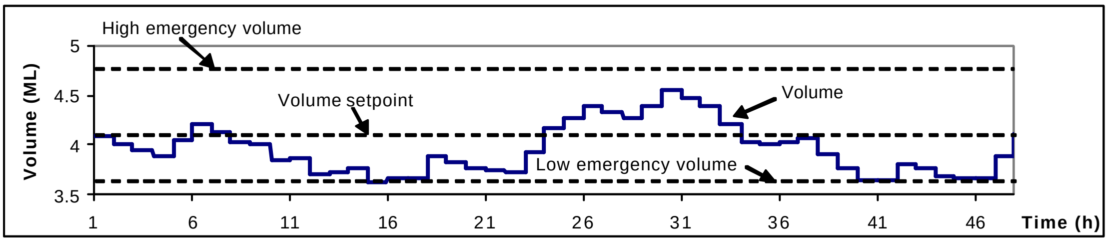


Figure 7.18: Evolution of the volume in Klaarwater reservoir (true electricity cost)

The optimised pumping sequence starts with the pump in the offline mode for the first four hours of the simulation, although the electricity is billed at the off-peak rate. Since Klaarwater reservoir volume is initialised at its setpoint, the optimiser prefers to distribute the water arriving from Northdene pump station to the Shallcross reservoirs and reticulation system where it is most needed. However, once water is again available for Klaarwater reservoir, it is sent to the input of the reservoir to try to correct the deviation from the setpoint that resulted from interrupting the feed of the reservoir for the first four hours. Therefore, the Klaarwater pump is switched on for two hours, until the tariff switches to the normal rate. This reduces the gap between the current volume in the reservoir and its setpoint. From this point, it could have been expected that the pump would not switch on before the next off-peak period occurring from the 22nd hour of simulation onwards. However, this is not the case, and as Figure 7.17 shows, the pump is switched on almost for the rest of the simulation, not taking into account anymore the different tariff zones. Furthermore, despite the action of the pump, there is a lot of difficulty in reducing the gap between the reservoir volume and its setpoint.

The explanation for this odd behaviour can be found in Figure 7.19 representing the evolution of the user demand flow out of the Shallcross reticulation system, to which the output of Klaarwater reservoir is connected.

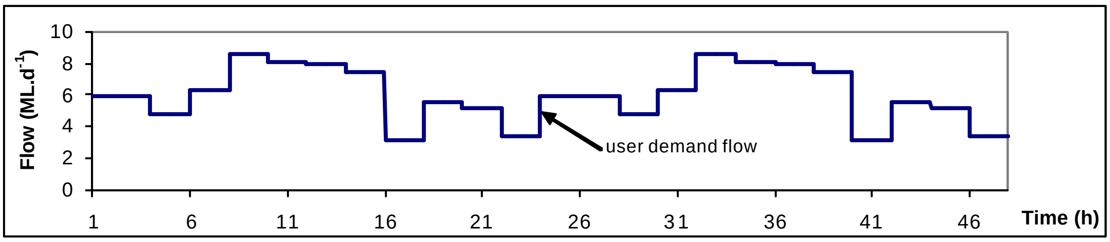


Figure 7.19: User demand flow out of the link to Shallcross reticulation system

This figure shows clearly that at the time there is a need of water in Klaarwater reservoir, there is also a strong demand in water from the users located in the Shallcross area. Therefore, all of the water pumped into the Klaarwater reservoir is immediately sent to the reticulation system which explains why the volume remains steady in spite of the action of the pump.

In other words, the action of the pump is not sufficient to correct the deviation of the volume from its setpoint, and the level is not far from the lower constraint which explains why the pump is on for so long, even during periods of peak electricity cost. The situation in Klaarwater reservoir only gets back to normal at the end of the 25th hour of simulation, with the volume crossing over its setpoint. This is made possible by the use of the pump, in conjunction with a strong decrease in the consumption by the Shallcross reticulation system. Since at this time, the tariff is charged at the off-peak rate, the pump remains on for some time after the setpoint has been crossed. This is to make best use of overnight pumping which allows storing water for further use. Unfortunately, the relief on this particular reservoir does not last for long, with the increase in the Shallcross user demand which materialises again from the 30th hour of simulation. With the reservoir volume dropping quickly towards it low-emergency constraints, there is no other choice but to use the pump once again to prevent the volume crossing this low limit, with no regard to the current cost of energy.

To summarise the analysis conducted on Klaarwater reservoir, it can be said that once again these results suggest that the emergency volumes that have been applied on Klaarwater reservoir are not optimal. Similarly to what was presented on the Lea Drive reservoir, important savings would be achieved on the pumping costs by allowing a more flexible range of operation on this particular reservoir.

Finally, the last pump station integrated in the optimisation problem is in the area surrounding Chatsworth 1 reservoir. It is used to boost the flow used to feed Chatsworth 2 and Chatsworth 3 reservoirs (Figure 7.20). It is recalled that the pumps located in the facility of Chatsworth 1 / Umlazi booster pump station are never switched on in the particular scenario of operations that was considered for the current problem of operational optimisation (cf. paragraph 7.1.5.1).

For this particular pump station, the use of the electricity structure that conforms to the bulk supply agreement gives results that are more complex to analyse, as shown on Figure 7.21. The pump is indeed switched on over almost the full length of the simulation (38 hours of use for a 48 h simulation), whereas the volume in Chatsworth 2 reservoir (reservoir to which this pump station is linked) never reaches its low-emergency constraint, as shown on Figure 7.22. Furthermore, despite the continuous action of the pump, the volume of this reservoir remains relatively steady and around its setpoint. It could be thought that this comes from the strong user demand requested in the Chatsworth 2 area at the time, but as a combined look at Figure 7.21, Figure 7.22 and Figure 7.23 shows, the use of the pump does not allow any correction in the volume of the reservoir, even during low consumption periods.

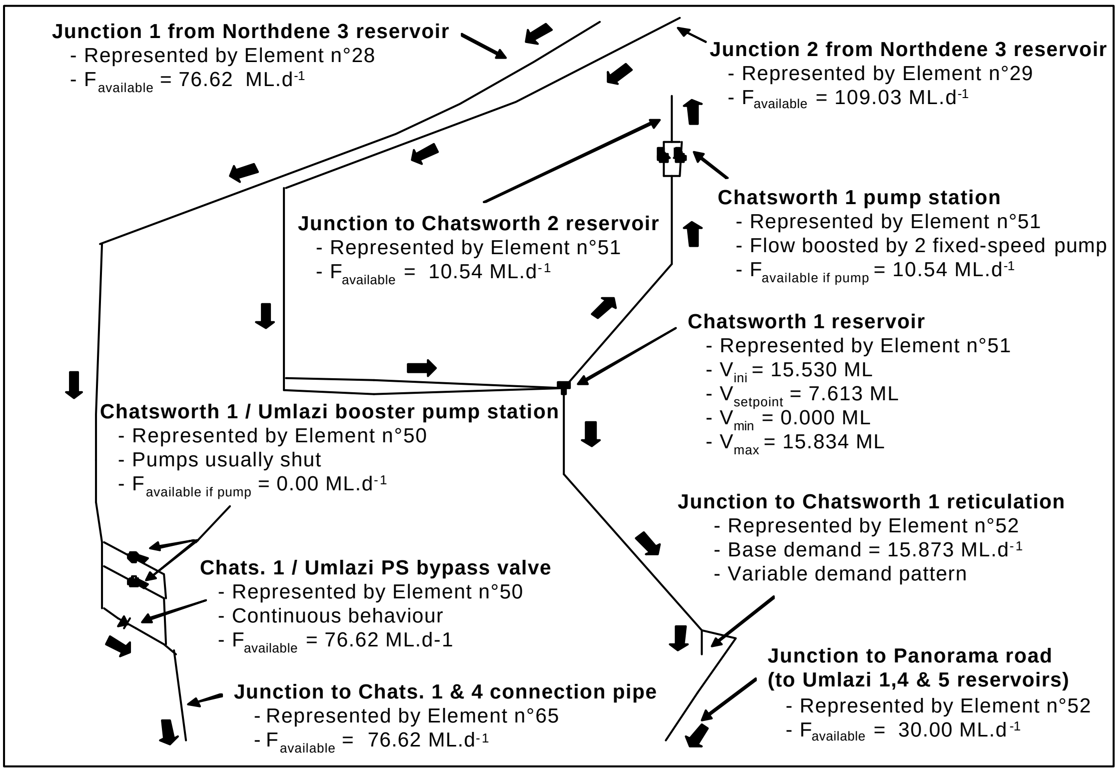


Figure 7.20: Chatsworth 1 reservoir, pump stations and reticulation area

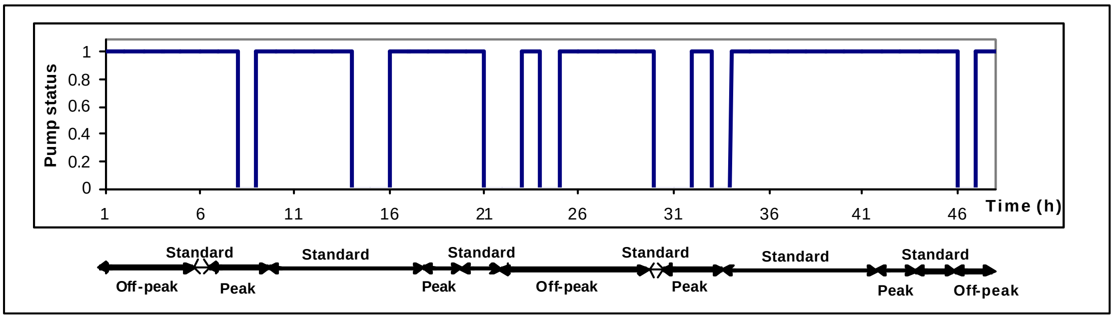


Figure 7.21: Pump status at the Chatsworth 1 pump station for the true electricity cost structure

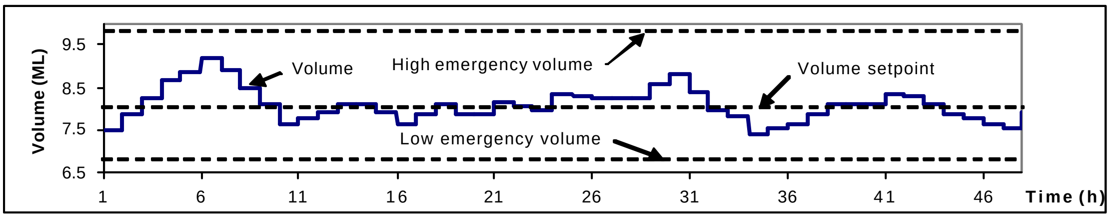


Figure 7.22: Evolution of the volume in Chatsworth 2 reservoir (true electricity cost)

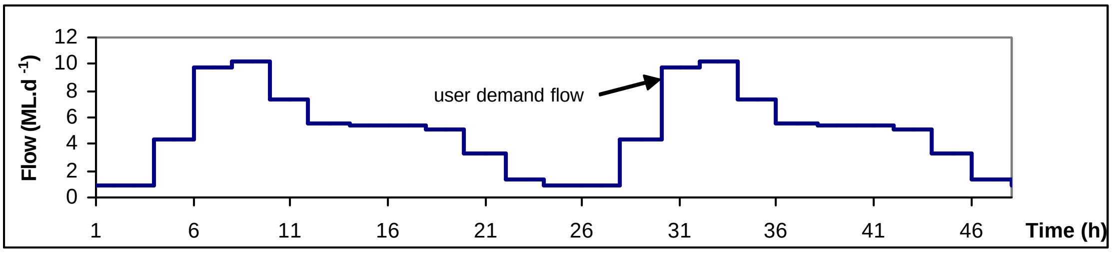


Figure 7.23: User demand flow out of the link to Chatsworth 2 reticulation system

In fact, the explanation resides in the fact that there is a lack of water in this part of the Aqueduct at the time the pump tries to correct the volume in Chatsworth 2 reservoir. Therefore, the flow entering Chatsworth 2 reservoir by the action of the pump is far below its maximum available flow and does not really help to bring the volume in the reservoir far above its setpoint. This behaviour is an important feature of the available flows modelling strategy: in other words, the maximum available flow defined for each pipeline cannot be guaranteed at each-time step. This is another example which shows that in general not all of the objectives of an optimisation problem can be satisfied simultaneously. Unfortunately, this behaviour will have strong consequences on the overall electricity bill.

7.2.4 Conclusions arising from the operational optimisation of the Southern Aqueduct

THE above paragraph presents the optimised sequence of operations of the Southern Aqueduct for two different electricity tariff structures. The modelled version of the water supply system is the one which corresponds to the network presented in Appendix D. This version was extracted from an ePANET file available at eThekwini Water Services which has required some adaptation since it did not rigorously conform to the version of the system represented in the supervisory control system used to monitor the Aqueduct in real-time (cf. Appendix C). Since some important parameters required for the optimisation were missing and could not be provided by eThekwini Water Services staff, it has been necessary to make several modelling choices before the operational optimisation of the system could be conducted. These modelling choices are summarised in Appendix E. It is obvious that the corresponding version of the Southern Aqueduct that results from these modelling choices does not give a rigorous picture of the existing water distribution system. This is why no comparisons were made between the results stored by the SCADA system and the two optimised sequences of operations obtained for the different electricity tariff structures.

The experiment that has been conducted on this particular portion of the Durban supply system has required the constitution of several data tables for the definition of time-varying parameters according to the principles of the available flows strategy, defined in Chapters 4 and 5. This problem was solved using the MINLP solver DICOPT++. Several MILP sub-problems needed to be solved during the course of the optimisation: this was made possible by the newly acquired solver CPLEX. Similarly, the CONOPT solver has been required for the solution of the NLP sub-problems.

The results show that all of the constraints of the problem have been respected, particularly the achievement of an integer strategy for all of the binary variables included in the formulation.

In addition, although it has not been possible to compare the results given by the optimiser with the current strategy of operation which could have led to an adapted pumping sequence making best use of the existing electricity structure, it has nevertheless been possible to highlight some interesting behaviours of the optimiser. Hence, several examples of the presence of the anticipation phenomenon have been given. This is one of the most important features of the optimiser concerning the operational optimisation objective, since it makes best use of overnight pumping by storing water in the reservoir when the electricity is cheaper. Furthermore, it was shown that, by extending the modelling strategy to a large-scale application, there is even more chance of having to compromise between conflicting objectives.

More generally speaking, the experiment conducted on this large-scale application has shown similar behaviours of the pump stations to those obtained with small-scale examples and presented in Chapter 6. Therefore, the results presented in the above section allow a final validation of the modelling method known as the available flows strategy. The use of a mixed-integer predictive controller has indeed provided satisfactory results, even on an existing distribution system.

Besides the final validation of the modelling choices presented earlier in this document, the results obtained on the approximate version of the Southern Aqueduct have also opened a new field of application for the particular modelling tool developed here. For instance, to improve the economic savings, it is clear that the software may be used to test the sensitivity to the emergency constraints chosen for the reservoirs. Furthermore, it could be envisaged to use this software tool to test the influence of different electricity patterns on the operational costs, therefore helping the management during their negotiations with electricity providers. Finally, this software could also be used for the development of new scenarios of operation that are more efficient from an economic point of view, by testing several configurations, on the same supply system. This would however require that the hydraulics of the new scenario be validated before by software similar to ePANET for instance (to obtain "available flows").

However, before envisaging further developments, like a commercial version of the modelling tool for instance, the next phase would be to calibrate finally the model used to represent the Southern Aqueduct. This step would be compulsory if one would like to estimate how large the savings on the operational costs could be on the Southern Aqueduct, with the new sequence of operations provided by the modelling tool. To evaluate that this calibration is reliable, a good test would be to check that the rates at which the reservoirs fill are similar in both the model and the actual network.

It is indeed true that the available flows strategy, on which this optimisation tool was based, while easier to solve, does not provide an accurate representation of the hydraulic behaviour of a complex distribution system. Such a simplified model can only be relied upon if it is validated using a full hydraulic model of the system. This is why it is recommended to use this modelling tool in combination with the public domain software ePANET, to determine all the required maximum available flows, as explained in the above sections.

Nevertheless, it is now clear that the modelling method developed especially for the operational optimisation of an existing water distribution system needs some adaptation before it can provide on-line optimisations of such systems. The computational complexities escalate rapidly as the network size grows. It is now clear that the algorithms currently used, even with the most powerful PC computer technology available nowadays, would not allow the solving of such big problems in amounts of time compatible with an on-line application.

7.3 Concluding remarks to the chapter

IN this chapter, the results for the extension of the available flows strategy to an existing large-scale distribution system have been presented. They have confirmed the soundness of the modelling choices made at the time this approach was initially developed. Furthermore, the robustness of the method has also been verified, since it has been possible to obtain an integer strategy for all of the binary variables of the application, in spite of its large size. However, some weaknesses have appeared in the method, since it requires large computational resources when applied on a large-scale system.

Nevertheless, these results have opened the way for new interesting applications of the optimisation tool, different from those initially envisaged at the origin of the project. This at least has given real satisfaction, since it has become clear that the method would not be applicable, in its current form, for the creation of an integrated on-line system, due to the complexities associated with solving in real-time huge optimisation problems including integer variables.

In the next chapter, all of the developments achieved during the course of this particular research project will be summarised. Furthermore, final recommendations concerning possible future development will be given.

CHAPTER 8 – CONCLUSIONS AND RECOMMENDATIONS

THIS chapter is divided into two parts: firstly, the conclusion summarises the different findings during the course of this research project, then the recommendation paragraph suggests areas of concern regarding these developments and gives possible methods of solution.

8.1 Conclusions

THIS study has focused on the off-line operational optimisation of water distribution networks, by Model Predictive Control (MPC), using Mixed Integer Non-Linear Programming (MINLP), for application to the Durban water supply system. This project was developed as a pilot study, to reveal potentials and requirements for a full simulation or optimisation application on the Durban water distribution system.

In order to achieve the operational optimisation of such a system, an MINLP model that incorporates rigorous equations has been proposed for the modelling and optimisation of any existing water distribution scheme, for a given electricity tariff structure, user demand flows and operating parameters. Initially, the model provided an accurate and comprehensive representation of the physical phenomena taking place in the network. This approach, known as "calculated flows strategy", was based on friction losses and available heads, so that a pressure-balanced solution could be provided. Unfortunately, the initial tests with this modelling strategy showed that it would not converge reliably, even on the smallest applications. This has lead to the reduction of the complexities associated with the calculated flows model and to the development of a new approach, termed as the "available flows" strategy. The new hydraulic model relative to this second modelling strategy has employed the concept of the definition of a fixed maximum flow in a pipe section, depending on whether a pump is active or not. The current flows were calculated as the product of the maximum flow defined for the line and the actual valve position in this particular line. It is clear that this new model simplifies the pipe network hydraulics. Essentially, the system hydraulics are solved using volume and chlorine balances. Since the influence of pipe friction is not looked at anymore, all of the problems of non-linearity linked to the use of pressure balances in the equations of the initial model are removed.

For both modelling approaches, the model employs a standard modelling element which is repeated until all of the different parts of the water distribution system are built within the model.

The modelling element was represented by a vessel of variable or fixed volume, having one input and two outputs. An auxiliary input flow could also be used on each element to represent external water supply or a chlorine re-dosing point. Constraints and initial values needed to be specified for standard elements representing storage reservoirs. Although the geometry of the standard element is common to the two modelling approaches, the quantity of data to provide is different from one approach to the other. The available flows strategy has reduced the amount of data to be provided for the model.

Because the water distribution system that requires optimisation consists of a combination of modelling elements, an entire network problem is specified in a data table, where one element is represented by one line in the table. Initialisations have to be provided before the process of optimisation can be launched. Valve positions, statuses of pumps, flows, storage reservoir volumes, and chlorine concentrations are examples of the type of data that need to be provided. Since some of these parameters are represented by an integer, the problem of operational optimisation of a water distribution network gives rise to a MINLP problem. The user interface is realised with the mathematical software MATLAB. However, since solving a MINLP optimisation problem requires the use of algorithms and computer codes that are external to MATLAB, it is necessary to create a link between this software and the optimisation package GAMS, in which these algorithms are available.

Recent interest in the development and application of MINLP optimisation algorithms has made it feasible to solve such problems, even of significant size. DICOPT++ was used as the MINLP solver of this large model. However, it was associated with the solver CPLEX for the MILP sub-problems and with the solver CONOPT for the NLP ones. The first phase of tests of the modelling strategy has been held on a small-scale configuration, involving 10 modelling elements. This has allowed the highlighting of some interesting behaviours of the optimisation tool based on the available flows strategy. Integer strategies have been achieved without major difficulties, even on simulation periods lasting for 48 hours, with a prediction interval of 12 h. The optimised sequences that have been reached were obtained with respect to all of the operational constraints and objectives. More importantly, on the operational side, it was proved that the optimiser could provide a new pumping sequence making best use of overnight pumping, guaranteeing therefore important savings. These promising results have encouraged the extension of the method to the domain of large-scale applications. Therefore, a version of the aqueduct providing the Durban southern areas with potable water was created within the optimisation tool.

An optimisation was conducted on this particular distribution network to check that all of the principles and behaviours highlighted during the first phase of tests could be extended to this larger application. The results obtained on the Southern Aqueduct have confirmed this, since eg. an anticipation phenomenon appeared on several reservoirs.

It is unfortunate that the analysis of the results on large-scale applications could not be pursued fully. Since data were missing for the final calibration of the model, no attempt was made at comparing the current pumping sequence with the optimised one provided by the optimisation tool based on the available flows strategy. Because of this problem of calibration, no economic figures could either be given on the Southern Aqueduct to confirm the benefit of applying such a method. However, the kind of results presented give strong hope that applying these optimisation techniques on the Southern Aqueduct would lead to a major cut in the operational costs.

Although the method developed especially for the operational optimisation of large-scale distribution systems gave encouraging results, even on the largest applications, a disappointment appeared when it became clear that this method could not be extended for on-line application, in its current state. The difficulties associated with the real-time solving of non-linear mixed integer problems can exceed the computational resources available for the supervision of a water distribution network. The huge computational effort required for the operational optimisation of the Southern Aqueduct is indeed not compatible with an on-line application. However, to finish on a more positive note, even if no further development is conducted on the method detailed here, it can be used as a the modelling tool in its current form for the definition of a more economically efficient operation or for the design of new water delivery routes that are more economical. It could even serve as a means to choose the most efficient electricity tariff structure among the numerous structures proposed by electricity providers.

8.2 Recommendations

THE empirical modelling approach based on non-linear MPC that was used during the course of this research project has appeared promising. For instance, detailed process knowledge was not required to achieve the operational optimisation, and it has been possible to choose the model form, in an attempt to reduce the amount of computation. Although not all of the initial objectives of this project have been fulfilled during the course of the research, the concept employed here broadly demonstrated its ability to resolve the operational optimisation problem.

However, there is still space for improvement and new developments would be required to develop an on-line application of the modelling tool. For instance, it is obvious that the available flows strategy does not provide in itself an hydraulic representation of the network that is accurate enough. Therefore, it is recommended that its use be associated with the ePANET software which should serve as the main database for the constitution of the data tables required by the optimisation tool.

It is also clear that an on-line application of the empirical model developed here would be limited by the “curse of dimensionality” linked to the application of non-linear programming techniques on a large-scale application. To overcome this problem, future research development should probably concentrate on keeping the same approach but running it off-line for different plant conditions representing the most common operating conditions of the system to optimise. The optimised operating conditions resulting from these off-line optimisations would then be stored in a table for further use. Going on-line would be synonymous with choosing, in real-time, amongst these pre-stored configurations, by estimating to which of the stored simulated operating conditions the current state is the closest. This new development would not lead to major changes of the current modelling tool. More interestingly, it would also remove situations when the optimiser cannot find online an integer solution for one of the control cycles, a situation that would lead to a “crash” of the controller, since, in real-time control, the plant awaits an input from the controller. By picking up the most suitable solution from a list of optimal behaviours, there would be the assurance that an integer solution is available at all times. Finally, as no other estimation occurs than determining which of the offline runs represents most closely the plant behaviour, the controller should adapt quickly to changing process conditions.

If the above “look-up” solution method proves to be successful, then a commercial application of the method may be envisaged. However, it would require the development of software tools that provide integrated dynamic modelling and real-time optimisation capabilities. Several vendors currently offer such software products, and it is likely that process economics will encourage the development of more efficient products. Also needed would be software support tools for long-term maintenance of non-linear models and controllers. This toolkit would enable the control engineer to analyse model accuracy and evaluate performance. The deployment of non-linear MPC technology would be advanced by establishing hardware and software standards. This would allow components from different vendors to be combined to achieve the best solution for a particular application.

APPENDIX A - REFERENCES

- Abadie J. and Carpentier J. (1969). Application of the GRG Algorithm to Optimal Control Problems. In *"Optimization: Symposium of the Institute of Mathematics and its Applications. University of Keele, England, 1968. Edited by R. Fletcher, Academic Press, London. ISBN 0-12-260650-7, 354 pages"*. pp 191 – 211.
- Adjiman C.S., Androulakis I.P. and Floudas C.A. (1997). Global Optimization of MINLP Problems in Process Synthesis and Design. *Computers & Chemical Engineering*. **21** (Supplement 1), S445 – S450.
- Alla P. and Jarrige P.A. (1988). Optimal Control of the West Parisian Area Water Supply Network. . In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 376 – 391.
- Allgöwer F. and Doyle J.F. (1997). Nonlinear Process Control: Which Way to the Promised Land? Fifth International Conference on Chemical Process Control. *AIChE Symposium Series, Kantor, Garcia and Carnahan Editors*. **93**, 24 – 45.
- Alperovits E. and Shamir U. (1977). Design of Optimal Water Distribution Systems. *Water Resources Research*. **13** (6), 885 – 900.
- Andreiev N. (1981). A New Dimension: A Self-Tuning Controller that Continually Optimizes PID Constants. *Control Engineering*. **28** (8), 84.
- Aström K.J. and Wittenmark B. (1973). On Self-Tuning Regulators. *Automatica*. **9**, 185 - 199.
- Aström K.J. (1983). Theory and Application of Adaptive Control – A Survey. *Automatica*. **19** (5), 471 – 486.
- Aström K.J. and Wittenmark B. (1984). Computer Control Systems: Theory and Design. *Prentice Hall*. 3rd Edition (1996). ISBN 0-13-314899-8, 555 pages.
- Aström K.J. and Wittenmark B. (1995). Adaptive Control, 2nd Edition. *Addison-Wesley Publishing Company, Inc.*
- Azzaro-Pantel C., Bernal-Haro L., Baudet P., Domenech S. and Pibouleau L. (1998). *Computers & Chemical Engineering*. **22** (10), 1461 – 1481.
- Baffi G., Morris J. and Martin E. (2002). Non-Linear Model Based Predictive Control Through Dynamic Non-Linear Partial Least Squares. *Trans ICheme*. **80** (Part A), 75 - 86.
- Balas E. (1979). Branch and Bound – Implicit Enumeration. *Annals of Discrete Mathematics*. **5**, 185.

- Beale E.M.L. (1977). Integer Programming. In *"The State of the Art in Numerical Analysis. Proceedings of the Conference on the State of the Art in Numerical Analysis Held at the University of York, April 12 – 15 1976, Organized by the Institute of Mathematics and its Applications. Edited by D. Jacobs, Academic Press, London. ISBN 0-12-378650-9, 978 pages"*. pp 409 – 448.
- Bellman R.E. (1957). Dynamic Programming. *Princeton University Press*.
- Benders J.F. (1962). Partitioning Procedures for Solving Mixed-Variable Programming Problems. *Numerische Mathematik*. **4**, 238 - 252.
- Bengtsson G. and Egardt B. (1984). Experiences with Self-Tuning Control in the Process Industry. *Proceedings of the 9th World IFAC Congress, Budapest.* Paper CS-7.
- Bequette B.W. (1991). Nonlinear Control of Chemical Processes: A Review. *Industrial and Engineering Chemistry Research*. **30**, 1391 – 1413.
- Bhatia T.K. and Biegler L.T. (1999). Multiperiod Design and Planning with Interior Point Methods. *Computers & Chemical Engineering*. **23** (7), 919 – 932.
- Birewar D.B. and Grossman I.E. (1990). Simultaneous Production Planning and Scheduling in Multiproduct Batch Plant. *Industrial & Engineering Chemistry Research*. **29**, 570 – 580.
- Biscos C., Mulholland M., Le Lann M.V., Brouckaert C.J., Bailey R. and Roustan M. (2002a). Optimal Operation of a Potable Distribution Network. *Water Science and Technology.* **46** (9), 155 – 162.
- Biscos C., Mulholland M., Le Lann M.V., Buckley C.A., Brouckaert C.J., Bailey R. and Roustan M. (2002b). Water Distribution Network Optimisation by MINLP. *Proceedings of the 2002 Biennal Water Institute of Southern Africa Conference and Exhibition.* International Conference Centre, Durban, South Africa, 19 – 23 May 2002. Session 3G.
- Biscos C., Mulholland M., Le Lann M.V., Buckley C.A. and Brouckaert C.J. (2003). Optimal Operation of Water Distribution Networks by Predictive Control Using MINLP. *Accepted for publication in WaterSA.*
- Brière F.G. (2000). Distribution et Collecte des Eaux. *Editions de l'Ecole Polytechnique de Montréal.* 2ème Edition. ISBN 2-553-00775-2, 400 pages.
- Brion L.M. and Mays L.W. (1991). Methodology for Optimal Operation of Pumping Stations in Water distribution Systems. *ACSE Journal of Hydraulic Engineering.* **117** (11), 1 – 19.
- Brooke A., Kendrick D., Meeraus A. and Raman R. (1998). GAMS: a User's Guide. *GAMS Development Corporation*.
- Bruno J.C., Fernandez F., Castells F. and Grossman I.E. (1998). A Rigourous MINLP Model for the Optimal Synthesis and Operation of Utility Plants. *Trans. IChemE*. **76** (Part A), 246 – 258.
- Camacho E.F. and Bordons C. (1995). Model Predictive Control in the Process Industry. *Springer Verlag New-York*. ISBN 3-540-19924-1, 239 pages.

- Carpentier P. (1983). Décomposition et Aggrégation dans la Conduite Optimale d'un Grand Réseau de Distribution d'Eau. *PhD Thesis. Ecole Nationale des Mines de Paris*. Paris, France.
- Cembrano G., Wells G., Quevedo J., Pérez R., Argelaguet R. (2000). Optimal Control of a Water Distribution Network in a Supervisory Control System. *Control Engineering Practice*. **8**, 1177-1188.
- Chambers V.K., Creasy J.D. and Joy J.S. (1995). Modelling Free and Total Chlorine Decay in Potable Water Distribution Systems. *Journal of Water Supply Research and Technology – Aqua*. **44** (1), 60 – 69.
- Chang C.T. and Hwang J.R. (1996). A Multiobjective Programming Approach to Waste Minimization in the Utility Systems of Chemical Processes. *Chemical Engineering Science*. **51** (16), 3951 - 3965
- Chase D.V., and Ormsbee L.E. (1989). Optimal Pump Operation of Water Distribution System with Multiple Storage Tanks. *Proceedings of the 1990 ASCE Conference on Water Resources Planning and Management. ASCE, New York*. 733 – 736.
- Chen Y.C. and Schwartz A.L. (2002). Riot-95, a MATLAB Toolbox for Solving General Optimal Control Problem and its Application to Chemical Processes. In "*Rein Luus: Recent Development in Optimization and Optimal Control in Chemical Engineering. Transworld Research Publishers. ISBN 81-7736-088-4*". pp 229-252.
- Chou C.C. and Shih Y.S. (1987). A Thermodynamic Approach to the Design and Synthesis of Plant Utility Systems. *Industrial & Engineering Chemistry Research*. **26**, 1100.
- City of Baltimore Department of Water website (http://www.baltimorecitydpw.com). How is Water Treated ? *City of Baltimore Department of Water*.
- Clarke D.W., Mohtadi C., Tuffs P.S. (1987). Generalized Predictive Control – Part I. The Basic Algorithm. *Automatica*. **23** (2), 137 – 148.
- Clarke D.W., Mohtadi C., Tuffs P.S. (1987). Generalized Predictive Control – Part II. Extensions and Interpretations. *Automatica*. **23** (2), 149 – 160.
- Clarke D.W. (1988). Application of Generalized Predictive Control to Industrial Processes. IEEE Control System Technology Magazine. April 1988, 49 – 55.
- Cooperative Research Centre for Water Quality and Treatment website (http://www.waterquality.crc.org.au). How is Water Treated ? *Cooperative Research Centre for Water Quality and Treatment*.
- Coulbeck B. (1977). Optimization and Techniques in Dynamic Control of Water Distribution Systems. *PhD Thesis. University of Sheffield*. Sheffield, United Kingdom.

- Coulbeck B. (1988). A Review of Methodologies for Modelling and Control of Water Supply. In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 80 – 109.
- Creasey J.D. (1988). Pump Scheduling in Water Supply: More Than a Mathematical Problem. In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 279 – 289.
- Cunha M.C. and Sousa J. (1999). Water Distribution Network Design Optimization: Simulated Annealing Approach. *Journal of Water Resources Planning and Management*. **125** (4), 215-226.
- Cutler C.R. and Ramaker B.L. (1979). Dynamic Matrix Control: a Computer Control Algorithm. *Proceedings of the 86th AIChE National Meeting, Houston (Texas)*. Paper S1b.
- Cutler C.R. and Ramaker B.L. (1980). Dynamic Matrix Control: a Computer Control Algorithm. *Proceedings of the Joint Automatic Control Conference, San Francisco (California)*. Paper WP5-B, 1 – 6.
- Deghaye S.W., Guiamba I. and Mulholland M. (2000). Enhanced Dynamic Matrix Control Applied to Liquid-Liquid Extraction. *South African Institution of Chemical Engineers*, 9th National Meeting, Secunda. Proceedings, October 9-12.
- Diaz M.S. and Bandoni J.A. (1996). A Mixed-Integer Optimization Strategy for a Large Scale Chemical Plant in Operation. *Computers & Chemical Engineering*. **20** (5), 531 – 545.
- Dikewar U.M., Grossman I.E. and Rubin E.S. (1992). An MINLP Process Synthesizer for a Sequential Modular Simulator. *Industrial & Engineering Chemistry Research*. **31**, 313 – 322.
- Doukas N. and Luyben W.L. (1978). Economics of Alternative Distillation Configurations for the Separation of Ternary Mixtures. *Industrial & Engineering Chemistry Process Design and Development*. **17**, 272 – 281.
- Douglas J.M. (1985). A Hierarchical Decision Procedure for Process Synthesis. *AIChE Journal*. **31**, 353 - 362.
- Douglas J.M. (1988). Conceptual Design of Chemical Processes. McGraw-Hill, New York. ISBN 0-07017-762-7, 601 pages.
- Douglas J.M. (1990). Synthesis of Multi-step Reaction Processes. *Proceedings of the 3rd International Conference on Foundation of Computer-Aided Process Design. Snowmass Village, Colorado. July 10 - 14 1989. Edited by Siirola J.J., Grossmann I.E., Stephanopolous G. . Amsterdam: Elsevier, 1990*. ISBN 0-444-88233-2, 568 pages.
- Drud A. (1985). CONOPT: A GRG Code for Large Sparse Dynamic Nonlinear Optimisation Problems. *Mathematical Programming*. **31**, 153 – 191.

- Duan N., Mays L.W. and Lansey K. E. (1990). Optimal Reliability-Based Design of Pumping and Distribution Systems. *Journal of Hydraulic Engineering*. **116** (2), 249-268
- Duran M.A. (1984). A Mixed-Integer Nonlinear Programming Approach for the Systematic Synthesis of Engineering Systems. *Ph.D. Thesis, Carnegie-Mellon University*. Pittsburgh, United States.
- Duran M.A. and Grossman I.E. (1986a). An Outer-Approximation Algorithm for a Class of Mixed-Integer Nonlinear Programs. *Mathematical Programming*. **36**, 307 – 339.
- Duran M.A. and Grossman I.E. (1986b). Simultaneous Optimization and Heat Integration of Chemical Processes. *AIChE Journal*. **32** (1), 123 - 138.
- Duran M.A. and Grossman I.E. (1986c). A Mixed-Integer Nonlinear Programming Approach for Process Systems Synthesis. *AIChE Journal*. **32** (4), 592 – 606.
- Durban Metro Water Services (2000).Water Services Development Plan. *Durban Metro Water Services*. Internal document, 41 pages.
- Egli U.M. and Rippin D.W.T. (1986). Short-Term Scheduling for Multiproduct Batch Chemical Plants. *Computers & Chemical Engineering*. **10** (4), 303 – 325.
- Eiger G., Shamir U. and Ben-Tal A. (1994). Optimal Design of Water Distribution Networks. *Water Resources Research*. **26** (4), 539-549.
- Environmental Protection Agency website (http://www.epa.gov/safewater). Groundwater and Drinking Water. *United States Environmental Protection Agency*.
- Fallside F. and Perry P.F. (1975). Hierarchical Optimization of a Water Supply Network. *Proceedings of the IEEE*. **122** (2), 387-401.
- Fallside F. (1988). Computer Techniques for On-Line Control of Water Supply Networks. In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 313 – 329.
- Fletcher R. and Leyffer S. (1994). Solving Mixed-Integer Nonlinear Programs by Outer Approximation. *Mathematical Programming*. **66**, 327 - 349.
- Freshwater D.C. and Henry B.D. (1975). Optimal Configuration of Multicomponent Distillation Systems. *Chemical Engineering*. 533 - 545.
- Freshwater D.C. and Ziogou E. (1976). Reducing Energy Requirement in Unit Operations. *The Chemical Engineering Journal*. **11**, 215 – 222.
- Fujiwara O., Jenchaimahakoon B. and Edirisinghe N.C.P. (1987). A Modified Linear Programming Gradient Method For Optimal Design of Looped Water Distribution Networks. *Water Resources Research*. **23** (6), 977-982.
- Gabler R., Sager, Wise & Petersen (1999). Essentials of Physical Geography. *Saunders College Publishing*.

- Garfinkel R.S. and Nemhauser G.L. (1972). Integer programming. *Series in Decision and Control. Wiley, New-York.* ISBN 0-471-29195-1, 427 pages.
- Geoffrion A.M. (1972). Generalized Benders Decomposition. *Journal of Optimization Theory and Applications*. **10** (4), 237 - 260.
- Goldberg D.E. and Kuo C.H. (1987). Genetic Algorithms in Pipeline Optimization. *Journal of Computing in Civil Engineering.* **1** (2), 128-141.
- Goulder I.C. and Morgan D.R. (1985). An Integrated Approach to the Layout and Design of Water Distribution Networks. *Civil Engineering Systems*. **2** (2), 104-113.
- Gregory P.C. (1959). Wright Air Development Centre Report 59 - 49. *Process Self-Adaptive Flight Control Symposium.*
- Grossman I.E. (1985). Mixed-Integer Programming Approach for the Synthesis of Integrated Process Flowsheets. *Computers & Chemical Engineering*. **9** (5), 463 – 482.
- Grossman I.E. (1989). MINLP Optimization Strategies and Algorithms for Process Synthesis. *Proceedings of the 3rd International Conference on Foundation of Computer-Aided Process Design. Snowmass Village, Colorado. July 10 - 14 1989. Editors, Siirola J.J., Grossmann I.E., Stephanopolous G. . Amsterdam: Elsevier, 1990*. ISBN 0-444-88233-2, 568 pages.
- Grossman I.E. (1996). Mixed-Integer Optimization Techniques for Algorithmic Process Synthesis. *Advances in Chemical Engineering*. **23**, 171-245.
- Gupta O.K. and Ravindran A. (1985). Branch and Bound Experiments in Convex Nonlinear Integer Programming. *Management Science.* **31** (12), 1533 – 1546.
- Harsh M.G., Saderne P. and Biegler L.T. (1989). A Mixed-Integer Flowsheet Optimization Strategy for Process Retrofits – The Debottlenecking Problem. *Computers & Chemical Engineering.* **13** (8), 947 – 957.
- Hendry J.E. and Hughes R.R. (1972). Generating Separation Process Flowsheets. *Chemical Engineering Progress*. **68**, 71.
- Hendry J.E., Rudd D.F., Seader J.D. (1973). Synthesis in the Design of Chemical Processes. *AIChE Journal.* **19** (1), 1 – 15.
- Henson M.A. (1998). Non-linear Model Predictive Control: Current Status and Future Directions. *Computers & Chemical Engineering*. **23** (2), 198 – 202.
- Himmelblau D.M. (1975). Optimum Design via Structural Parameters and Nonlinear Programming. *US – Japan Joint Seminar on Application of Process Systems Engineering to Chemical Technology Assessment.* Kyoto, Japan.
- Hoopes H.S., Hawk W.M. and Lewis R.C. (1983). A Self-Tuning Controller. *The Instrumentation, Systems and Automation Society Transactions.* **22**, 49.

- Hostrup M, Gani R., Kravanja Z., Sorsak A. and Grossman I.E. (2001). Integration of thermodynamic insights and MINLP Optimization for the Synthesis, Design and Analysis of Process Flowsheets. *Computers & Chemical Engineering*. **25** (1), 73 – 83.
- Hui C.W. and Natori Y. (1996). An Industrial Application Using Mixed-Integer Programming Technique: A Multi-Period Utility System Model. *Computers & Chemical Engineering*. **20** (Issue 972), S1577 – S1582.
- Ichikawa A. and Fan L.T. (1972) Necessary Condition for Optimal Process Structure ad Evolutionary Search for Optimal Structure. *Paper presented at the 1972 AIChE meeting.* Dallas, Texas.
- Integrated publishing website (http://www.tpub.com/fluid). Introduction to Fluid Power – Training Series on how Hydraulics and Pneumatics work. *Integrated publishing*.
- International Council for Local Environmental Initiatives (2000). *Members In Action: Durban Metropolitan Council, South Africa*. Biennial Report 1998 – 1999.
- Iyer R.R. and Grossman I.E. (1997). Optimal Multiperiod operational Planning for Utility Systems. *Computers & Chemical Engineering*. **21** (8), 787 - 800.
- Jacobs O.L.R. (1981). Introduction to Adaptive Control. Self-Tuning and Adaptive Control: Theory and Applications. *C.J. Harris and S.A. Billings Editions, London.* 2nd Edition (1985). ISBN 0-86341-036-7, 352 pages.
- Joly M. and Pinto J.M. (2003). Mixed-Integer Programming Techniques for the Scheduling of Fuel Oil and Asphalt Production. *Trans IChemE*. **81** (Part A), 427 – 447.
- Jowitt P.W., Garett R., Cook S. and Germanopoulos G. (1988). Real-Time Forecasting and Control for Water Distribution. In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 329 – 355.
- Jowitt P.W. and Germanopoulos G. (1992). Optimal Pump Scheduling in Water-Supply. *ASCE Journal of Water Resources Planning and Management*. **118** (4), 406 – 422.
- Kalman R.E. (1958). Design of a Self-Optimizing Control System. *Transaction ASME*. 80, 468 – 478.
- Kalman R.E. and Bucy R.S. (1961). New Results in Linear Filtering and Prediction Theory. *CBasic Engineering*, **83D**, 95-108
- Kessler A. and Shamir U. (1989). Decomposition Technique for Optimal Design of Water Supply Networks. *Water Resource Research*. **25** (7), 1469-1480.
- Kessler A. and Shamir U. (1991). Analysis of the Linear Programming Gradient Method for Optimal Design of Water Supply Networks. *Engineering Optimization*. **17** (1), 1 -19.
- Kim J.H. and Mays L.W. (1994). Optimal Rehabilitation Model for Water Distribution Systems. *Journal of Water Resource Planning and Management*. **120** (5), 982-1000.

- Kinoshita A.T., Umeda T. and O'Shima E. (1981). An Algorithm for Synthesis of Operational Sequences of Chemical Plants. *Proceedings of the 14th Symposium on Computerized Control and Operation of Chemical Plants*. Vienna, Austria.
- Kirkwood R.L., Locke M.H., and Douglas J.M. (1988). A Prototype Expert System for Synthesizing Chemical Process Flowsheets. *Computers & Chemical Engineering*. **12** (4), 329 - 343.
- Kocis G.R. and Grossman I.E. (1987). Relaxation Strategy for the Structural Optimization of Process Flowsheets. *Industrial and Engineering Chemistry Research*. **26** (9), 1869 – 1880.
- Kocis G.R. and Grossman I.E. (1989). A Modelling and Decomposition Strategy for the MINLP Optimization of Process Flowsheets. *Computers & Chemical Engineering*. **13** (7), 797 – 819.
- Kovac A. and Glavic P (1995). Retrofit of complex and energy intensive processes -- I. *Computers & Chemical Engineering*. **19** (12), 1255 - 1270.
- Kraus T.W. and Myron T.J. (1984). Self-Tuning PID Controller Uses Pattern Recognition Approach. *Control Engineering*. **31** (6), 106 – 111.
- Kravanja Z. and Grossman I.E. (1990). PROSYN – An MINLP Process Synthesizer. *Computers & Chemical Engineering*. **14** (12), 1363 – 1378.
- Kravanja Z. and Grossman I.E. (1994). New Developments and Capabilities in PROSYN – An Automated Topology and Parameter Process Synthesizer. *Computers & Chemical Engineering*. **18** (11 – 12), 1097 – 1114.
- Ku H. and Karimi I.A. (1988). Scheduling in Serial Multiproduct Batch Processes with Finite Interstage Storage: A Mixed Integer Linear Program Formulation. *Industrial & Engineering Chemistry Research*. **27**, 1840 – 1848.
- Ku H. and Karimi I.A. (1990). Scheduling in Serial Multiproduct Batch Processes with Due-Date Penalties. *Industrial & Engineering Chemistry Research*. **29**, 580 -590.
- Lang Y.D., Biegler L.T. and Grossman I.E. (1988). Simultaneous Optimization and Heat Integration with Process Simulators. *Computers & Chemical Engineering*. **12** (2/3), 311 -327.
- Lansey K.E. and Mays L.W. (1989). Optimization Model for Water Distribution System Design. *Journal of Hydraulic Engineering*. **115** (10), 1401-1418.
- Lansey K.E. and Zhong Q. (1990). A Methodology for Optimal Control of Pump Stations. *Proceedings of the 1990 ASCE Conference on Water Resources, Planning and Management. ASCE, New-York.*
- Lazàro M., Espuña A. and Puigjaner L. (1989). A Comprehensive Approach to Production Planning in Multipurpose Batch Plants. Computers & Chemical Engineering. **13** (9), 1031 - 1047.

- Le Lann M.V. (1999). Application de la Commande Avancée en Génie des Procédés, *Conférence inaugurale. Colloque : Applications de la Commande Avancée*, Paris. Proceedings Mars 5 – 12.
- Linhoff B.D., Townsend W., Boland D., Hewitt G.F., Thomas B.E.A., Guy A.R., Marsland R.H. (1982). A User Guide on Process Integration for the Efficient Use of Energy. *IChemE, London*. ISBN 0-85295-156-6, 247 pages.
- Little J.D.C., Murty K.C., Sweeny D.W. and Karel C. (1963). An Algorithm for the Travelling Salesman Problem. *Operations Research*. **11**, 972 – 989.
- Luyben W.L. (1990). Process Modelling, Simulation and Control for Chemical Engineers: Second Edition. *McGraw and Hill International Editions – Chemical Engineering Series*. ISBN 0-07-100973-8, 726 pages.
- Mahalec V. and Motard R.L. (1977a). Procedures for the Initial Design of Chemical Processing Systems. *Computers & Chemical Engineering*. **1** (1), 57 – 68.
- Mahalec V. and Motard R.L. (1977b). Evolutionary Search for an Optimal Limiting Process Flowsheet. *Computers & Chemical Engineering*. **1** (2), 133 -147.
- Marechal F. and Kalitventzeff B. (1997). Identifying the Optimal Pressure Levels in Steam Networks Using Integrated Combined Heat and Power Method. *Chemical Engineering Science*. **52** (17), 2977 – 2989.
- Marechal F. and Kalitventzeff B. (1999). Targeting the Optimal Integration of Steam Networks: Mathematical Tools and Methodology. *Computers & Chemical Engineering*. **Supplement**, 133 – 136.
- Masso A.H. and Rudd D.F. (1969). The Synthesis of System Designs. Part II: Heuristic Structuring. *AIChE Journal*. **15** (1), 10 – 17.
- Mavromatis S.P. and Kokossis A.C. (1998). Conceptual Optimisation of Utility Networks for Operational Variation. *Chemical Engineering Science*. **53** (8), 1585 – 1608.
- Mehrotra S.P. and Kpur P.C. (1974). Optimal Suboptimal Synthesis and Design of Flotation Circuits. *Separation Science*. **9**, 167.
- Minderman Jr. P.A. (1981). Heuristic Synthesis of Distillation Networks. *MSc(Eng) Thesis*. Georgia Institute of Technology. Atlanta, Georgia.
- Minderman Jr. P.A. and Tedder D.W. (1982). Comparisons of Distillation Networks; Extensively State Optimized vs. Extensively Energy Integrated. AIChE Symposium Series. **78**, 69.
- Mishra P.N., Fan L.T., and Erickson L.E. (1973a). Biological Wastewater Treatment System Design: Part I, Optimal Synthesis. *Canadian Journal of Chemical Engineering*. **51**, 694 – 701.

- Mishra P.N., Fan L.T., and Erickson L.E. (1973b). Biological Wastewater Treatment System Design: Part II, Effect of Parameter Variations on Optimal Process System Structure and Design. *Canadian Journal of Chemical Engineering*. **51**, 702.
- Morari M. and Faith D.C. (1980). The Synthesis of Distillation Trains with Heat Integration. *AIChE Journal*. **26**, 916.
- Morari M. and Garcia C.E. (1982). Internal Model Control: a Unifying Review and Some New Results. *Industrial & Engineering Chemistry Process Design and Development*. 21, 308 – 323.
- Mousavi H. and Ramamurthy A.S. (1999). Optimal Design of Multi-Reservoir Systems for Water Supply. *Advances in Water Resources*. **23** (2000), 613- 624.
- Mulholland M. and Narotam N.K. (1996). Constrained Predictive Control of a Counter-Current Extractor. In "*System Modelling and Optimisation, edited by J. Dolezal and J. Fidler. Chapman & Hall. ISBN 0-412-718804*". pp 251-258.
- Nath R. and Motard R.L. (1981). Evolutionary Synthesis of Separation Processes. *AIChE Journal*. **27** (4), 578.
- National Academy of Engineering website (http://www.greatachievements.org). Greatest engineering achievement of the 20^{th} century. *National Academy of Engineering.*
- Nikolaou M. (2001). Model Predictive Controllers: A Critical Synthesis of Theory and Industrial Needs. *Advances in Chemical Engineering Series*. Academic Press.
- Nishida N., Stephanopoulos G. and Westerberg A.W. (1981). A Review of Process Synthesis. *AIChE Journal*. **27** (3), 321 – 351.
- Nishio M. (1977). Ph.D. Thesis. *The University of Western Ontario*. London, Ontario (Canada).
- Nishio M., Itoh J., Shiroko K. and Umeda T. (1980). Thermodynamic Approach to Steam and Power System Design. *Industrial & Engineering Chemistry, Process Design and Development*. **19**, 306.
- Ogunnaike B.A. and Ray W.H. (1994). Process Dynamics, Modeling and Control. *Oxford University Press, New York*. ISBN 0-19-509119-1, 1280 pages.
- Ogunnaike B.A. and Wright R. (1997). Industrial Applications of Nonlinear Control. Fifth International Conference on Chemical Process Control. *AIChE Symposium Series, Kantor, Garcia and Carnahan Editors*. **93**, 46 – 59.
- O'Shima E. (1978). Safety Supervision of Valve Operations. *Journal of Chemical Engineering of Japan*. **11** (5), 390 – 395.
- Pahor B. and Kravanja Z. (1995). Simultaneous Solution and MINLP Synthesis of DAE / Process Problems: PFR Networks in Overall Processes. *Computers & Chemical Engineering*. **19** (Supplement 1), 181-188.
- Papoulias S. and Grossman I.E. (1983). A Structural Optimization Approach in Process Synthesis: Part I, II and III. *Computers & Chemical Engineering*. **7** (6), 695 – 734.

- Parks, P.C., Schaufelberger W., Schmid C. and Unbehauen H. (1980). Application of Adaptive Control Systems. *Springer-Verlag, New-York*. In *"Methods and Application in Adaptive Control, Lecture Notes in Control and Information Sciences, 24"*. pp 161 – 198.
- Peterka J. and Aström K. J. (1973). Control of Multivariable Systems with Unknown but Constant Parameters. *Proceedings of the 3rd IFAC Symposium on Identification and System Parameter Estimation*, 535 – 544.
- Petroulas T. and Reklaitis G.V. (1984). Computer-Aided Synthesis and Design of Plant Utility Systems. *AIChE Journal*. **30** (1), 69 - 78.
- Phimister J.R. (2000). General Algebraic Modeling System (GAMS). In *"Process Design Principles: Synthesis, Analysis, Evaluation. Edited by W.D. Seider, J.D. Seader, D.R. Lewin. Wiley, New-York. ISBN 0-471-21663-1, 816 pages"*. pp 726 – 739.
- Powell J.C., Hallam N.B., West J.R., Forster C.F. and Simms J. (2000). Factors which Control Bulk Chlorine Decay Rates. *Water Research*. **34** (1), 117 – 126.
- Prett D.M. and Morari M (1987). Shell Process Control Workshop. *Butterworths*.
- Prett D.M., Garcia C.E. and Ramaker B.L. (1990). Second Shell Process Control Workshop. *Butterworths*.
- Powers G.J. (1972). Heuristic Synthesis in Process Development. *Chemical Engineering Progress*. **68** (6), 88.
- Propoi A.I. (1963). Use of Linear Programming Methods for Synthesizing Sampled-Data Automatic Systems. *Automation Remote Control*. **24** (7), 837 – 844.
- Quesada I. and Grossman I.E. (1992). An LP/NLP Based Branch and Bound Algorithm for Convex MINLP Optimization Problems. Computers & Chemical Engineering. **16** (10 / 11), 937 – 947.
- Quindry G.E., Brill E.D., and Liebman J.C. (1981). Optimization of Looped Distribution Systems. *Journal of Environmental Engineering*. **107** (EE4), 665-679.
- Rafal M.D. and Stevens W.F. (1968). Discrete Dynamic Optimization Applied to On-Line Optimal Control. *AIChE Journal*. **14** (1), 85 – 91.
- Rao H.S. and Bree Jr. D.W. (1977). Extended Period Simulation of Water Systems – Part A. *ASCE Journal of the Hydraulic Division*. **103** (2), 97 – 108.
- Rathore R.N.S., Van Wormer K.A., Powers G.J. (1974a). Synthesis Strategies for Multicomponent Separation Systems with Energy Integration. *AIChE Journal*. **20** (3), 491 - 502.
- Rathore R.N.S., Van Wormer K.A., Powers G.J. (1974b). Synthesis of Distillation Systems with Energy Integration. *AIChE Journal*. **20**, 940.
- Rawlings J.B. and Muske K.R. (1993). The Stability of Constrained Receding Horizon Control. *IEEE Transaction on Automatic Control*. **38**, 1512 – 1516.

- Reklaitis G.V. (1992). Overview on Scheduling and Planning of Process Operations. *NATO Advanced Study Institute on Batch Processing Systems Engineering*. Antalya, Turkey.
- Richalet J., Rault A., Testud J.L., Papon J. (1976). Algorithmic Control of Industrial Processes. *Proceedings of the 4th IFAC Symposium on Identification and System Parameter Estimation*, 271 - 296.
- Richalet J., Rault A., Testud J.L., Papon J. (1978). Model Predictive Heuristic Control: Applications to Industrial Processes. *Automatica*. **14**, 413 – 428.
- Ricker N.L. and Lee J.H. (1995). Nonlinear Model Predictive Control of the Tennessee Eastman Challenge Process. *Computers & Chemical Engineering*. **19** (9), 961 – 981.
- Rippin D.W.T. (1993). Batch Process Systems Engineering: a Retrospective and Prospective Review. *Computers & Chemical Engineering*. **17** (Supplement 1), S1 – S536.
- Rivas J.R. and Rudd D.F. (1974a). Computer-Aided Safety Interlock System. *AIChE Journal*. **20** (2), 311 – 319.
- Rivas J.R. and Rudd D.F. (1974b). Synthesis of Failure-Safe Operations. *AIChE Journal*. **20** (2), 320 – 325.
- Rodrigues P. and Carrie J. (1992). Augmentation of Durban's Southern Aqueduct. *Imiesa*. **October 1992**, 3 – 17.
- Rossman L.A. (2000). Epanet 2 Users Manual. *United States Environmental Protection, Agency, Cincinnati.*
- Rudd D.F. (1968). The Synthesis of System Designs: Part 1, Elementary Decomposition Theory. *AIChE Journal*. **14** (2), 343 – 349.
- Rudd D.F., Power G.J. and Siirola J.J. (1973). Process Synthesis. *Prentice Hall*. ISBN 0-13723-353-1, 320 pages.
- Saunion R. (2001). Optimization of a Multi-Level Steam Distribution System by Mixed-Integer Non-Linear Programming. *MSc(Eng) Thesis*. University of Natal, Durban, South Africa.
- Schaake J., and Lai D. (1969). Linear Programming Dynamic Programming Applications to Water Distribution Network Design. *Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts*. Report n°116.
- Seader J.D. and Westerberg A.W. (1977). A Combined Heuristic and Evolutionary Strategy for Synthesis of Simple Separation Sequences. *AIChE Journal*. **23**, 951 – 954.
- Seborg D.E., Edgar T.F. and Shah S.L. (1986). Adaptive Control Strategies for Process Control: A Survey. *AIChE Journal*. **32**, 881 – 913.
- Shinskey F.G. (1996). Process Control Systems – 4th Edition. *Mc Graw-Hill Editions, New York*. ISBN 0-07057-101-5, 439 pages.
- Siirola J.J., Powers G.J., Rudd D.F. (1971). Synthesis of System Designs. Part III: Towards a Process Concept Generator. *AIChE Journal*. **17**, 677 - 682.

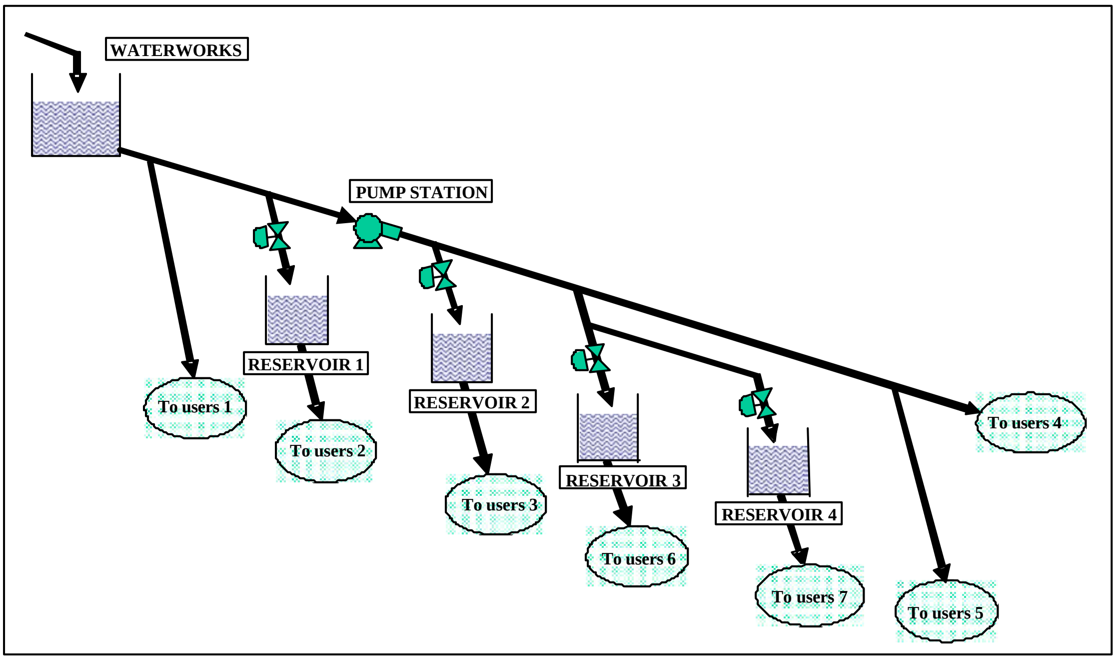
- Smith E.M. and Pantelides C.C. (1997). Global Optimisation of Nonconvex MINLPs. *Computers & Chemical Engineering*. **21** (Supplement 1), S791 – 796.
- Siirola J.J. and Rudd D.F. (1971). Computer-Aided Synthesis of Chemical Process Designs. *Industrial & Engineering Chemistry Fundamentals*. **10** (3), 353 - 362.
- Solanas J.L. (1974). Mando en Tiempo Real De Un Sistema de Suministro de Agua. *PhD Thesis*. *Universidad Politécnica de Cataluña.* Barcelona, Spain.
- Solanas J.L. and Montoliu J.M. (1988). The Optimal Operation of Water Systems In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 357 – 375.
- Stephanopoulos G., Linnhoff B. and Sophos A. (1982). Synthesis of Heat Integrated Distillation Sequences. *IChemE Symposium Series*. **74**, 11.
- Su Y.C., Mays L.W., Duan N. and Lansey K.N. (1987). Reliability-Based Optimization Model for Water Distribution System. *Journal of Hydraulic Engineering*. **114** (12), 1539-1556.
- Tatjewski P. (1988). A Suboptimal Approach to Scheduling of Reservoir Levels for a Multi-Reservoir Water Distribution Network. In *"Computer Application in Water Supply. Volume 2: Systems Optimization and Control. Edited by B. Coulbeck and C.H. Orr. Research Studies Press LTD. ISBN 0-86380-063-7, 438 pages"*. pp 225 – 239.
- Teegavarapu R.S.V. and Simonovic S.P. (2002). Optimal Operation of Water Resource Systems: Tradeoffs Between Modelling and Practical Solutions -Research Report R3T 2N2. *Department of Civil and Geological Engineering. University of Manitoba, Winnipeg, Canada.*
- Tomita S.M., Nagata M. and O'Shima (1986). On the Development of an Automatic Synthesizer of Operating Procedure for Chemical Plants. *Preprints of the IFAC Workshop, Kyoto*. 66 – 70.
- Turkay M. and Grossman I.E. (1996). Disjunctive Programming Techniques for the Optimization of Process Systems with Discontinuous Investment Costs – Multiple Size Regions. *Industrial & Engineering Chemistry Research*. **35**, 2611 – 2623.
- Umeda T., Hirai A, and Ichikawa A. (1972). Synthesis of Optimal Processing Systems by an Integrated Approach. *Chemical Engineering Science*. **27** (4), 795 – 804.
- Umeda T., Itoh J. and Shiroko K. (1978). Heat Exchange System Synthesis. *Chemical Engineering Progress*. **74** (7), 70 – 76.
- Umgeni Water website (http://www.umgeni.co.za/operations). Our Operations. *Umgeni Water - Amanzi*.
- Valiron F. (1990). Gestion des Eaux – Volume 1: principes, moyens, structures. *Presses de l'Ecole National des Ponts et Chaussés*. ISBN 2-85978-136-6, 352 pages.

- Valiron F. (1990). Gestion des Eaux – Volume 2 : Alimentation en Eau, assainissement. *Presses de l'Ecole des Ponts et Chaussés*. ISBN 2-85978-131-5, 506 pages.
- Valiron F. and Monod J. (1992). Mémento du Gestionnaire de l'Alimentation en Eau et de l'Assainissement – Tome 1: eau dans la ville, alimentation en eau. *Lyonnaise des Eaux - Lavoisier TEC & DOC*. ISBN 3-700-007761-9.
- Van den Heever S.A. and Grossman I.E. (1999). Disjunctive Multiperiod Optimization Methods for Design and Planning of Chemical Process Systems. *Computers & Chemical Engineering*. **23** (8), 1075 – 1095.
- Viessman W., Hammer J.R. & Hammer M.J. (1998). Water Supply and Pollution Control – Sixth Edition. *Addison-Wessley & Co Publication*. ISBN 0-321-01460-X, 878 pages.
- Viswanathan J. and Grossman I.E. (1990). A Combined Penalty Function and Outer-Approximation Method for MINLP Method. *Computers & Chemical Engineering*. **14** (7), 769 - 782.
- Walski T.M. (1996). Discussion of "Design Heuristic for Globally Minimum Cost Water Distribution Systems" by Loganathan *et al.* . *Journal of Water Resources Planning and Management*. **122** (4), 313-314.
- Walters G.A. and Cembrowicz R.G. (1993). Optimal Design of Water Distribution Networks. Water Supply Systems, State-of-the-Art and Future Trends. *Computational Mechanics Inc.*, 91-117.
- Wen J.T. (2002). *Introduction and Review of Static Optimization*. Lecture on Optimal Control.
- Westerlund T., Petterson F. and Grossman I.E. (1994). Optimization of Pump Configurations as a MINLP Problem. *Computers & Chemical Engineering*. **18** (9), 845 - 858.
- Westerlund T. and Petterson F. (1995). An Extended Cutting Plane Method for Solving Convex MINLP Problems. *Computers & Chemical Engineering*. **19** (Supplement 1), 131 - 136.
- Wolfe P. (1963). Methods of Nonlinear Programming. In *"Recent Advances in Mathematical Programming. Edited by Graves and Wolfe. McGraw-Hill, New York. ISBN 1-12414-137-5"*. pp 67 – 86.
- Wolfe P. (1967). Methods for Linear Constraints. In *"Nonlinear Programming. Edited by J. Abadie. North-Holland Publishing Company, Amsterdam. 316 pages"*. pp 120 – 124.
- Wong K.P. and Seborg D.E. (1985). A Theoritical Analysis of Smith and Analytical Predictors. *Submitted to AIChe Journal*.
- Wood D.J. and Rayes A.G. (1981). Reliability of Algorithms for Pipe Network Analysis. *ASCE Journal of the Hydraulic Division*. **107** (10), 1145 – 1161.
- Woodburn J. Lansey K.E., Mays L.W. (1987). Model for the Optimal Rehabilitation and Replacement of Water Distribution System Components. *Proceedings of the National Conference on Hydraulic Engineering*. Williamsburg, Virginia, United States of America.

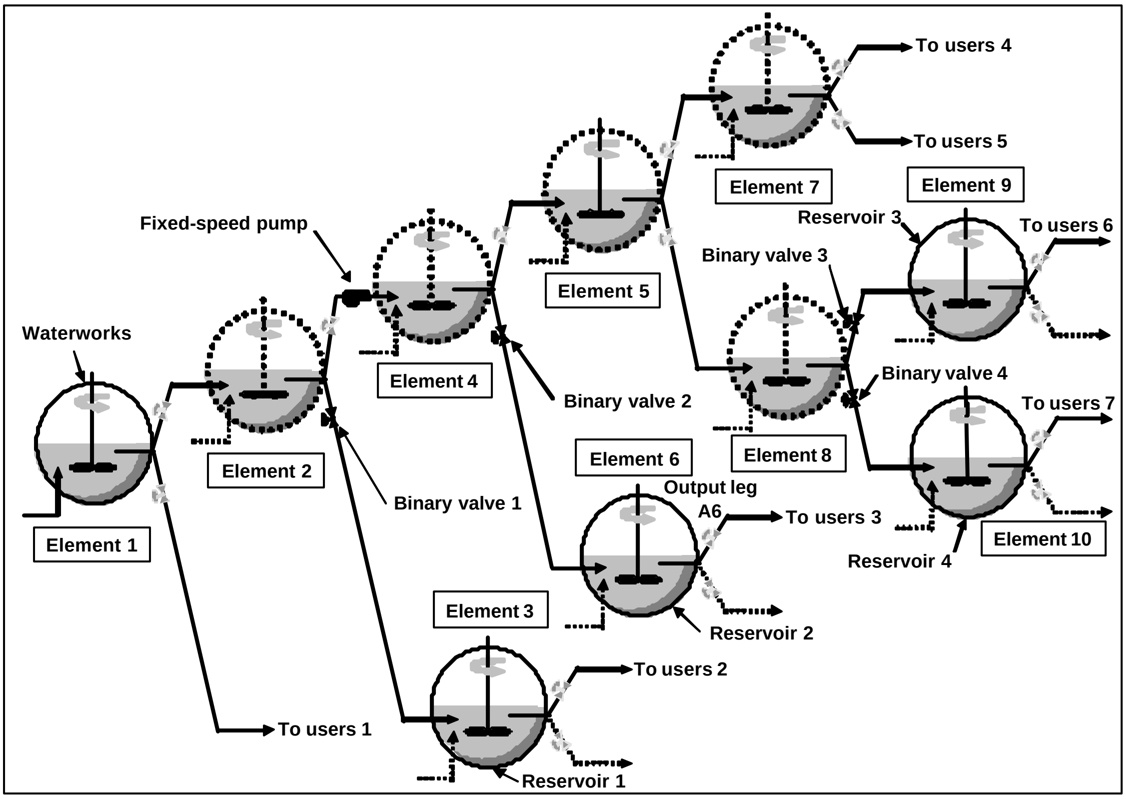
- World Bank (2002). World Development Report 2003 – Sustainable Development in a Dynamic World: Transforming Institutions, Growth and Quality of Life. *The World Bank*. ISBN 0-8213-5187-7.
- Wu Z.Y. (2002). A self-adaptive boundary search genetic algorithm and its application to water distribution systems. *Journal of Hydraulic Research*. **40** (2), 191-203.
- Yu G., Powell R.S. and Sterling M.J.H. (1994). Optimized Pump Scheduling in Water Distribution Systems. *Journal of Optimization Theory And Applications*. **83** (3), 463 – 488.
- Yuan X., Pibouleau L. and Domenech S. (1989). Experiments in Process Synthesis via Mixed-Integer Programming. *Chemical Engineering and Processing*. **25** (2), 99 – 116.
- Ydstie B.E., Kershenbaum L.S. and Sargent R.W. (1983). Theory and Application of an Extended Self-Tuning Controller. *AIChe Journal*. **32** (11), 1771 – 1780.
- Zamora J.M. and Grossman I.E. (1998). A Global MINLP Optimisation Algorithm for the Synthesis of Heat Exchanger Networks with no Stream Splits. *Computers & Chemical Engineering*. **22** (3), 367-384.

APPENDIX B – SMALL-SCALE APPLICATION: RESULTS WITH CPLEX ON THE 10-ELEMENT CONFIGURATION

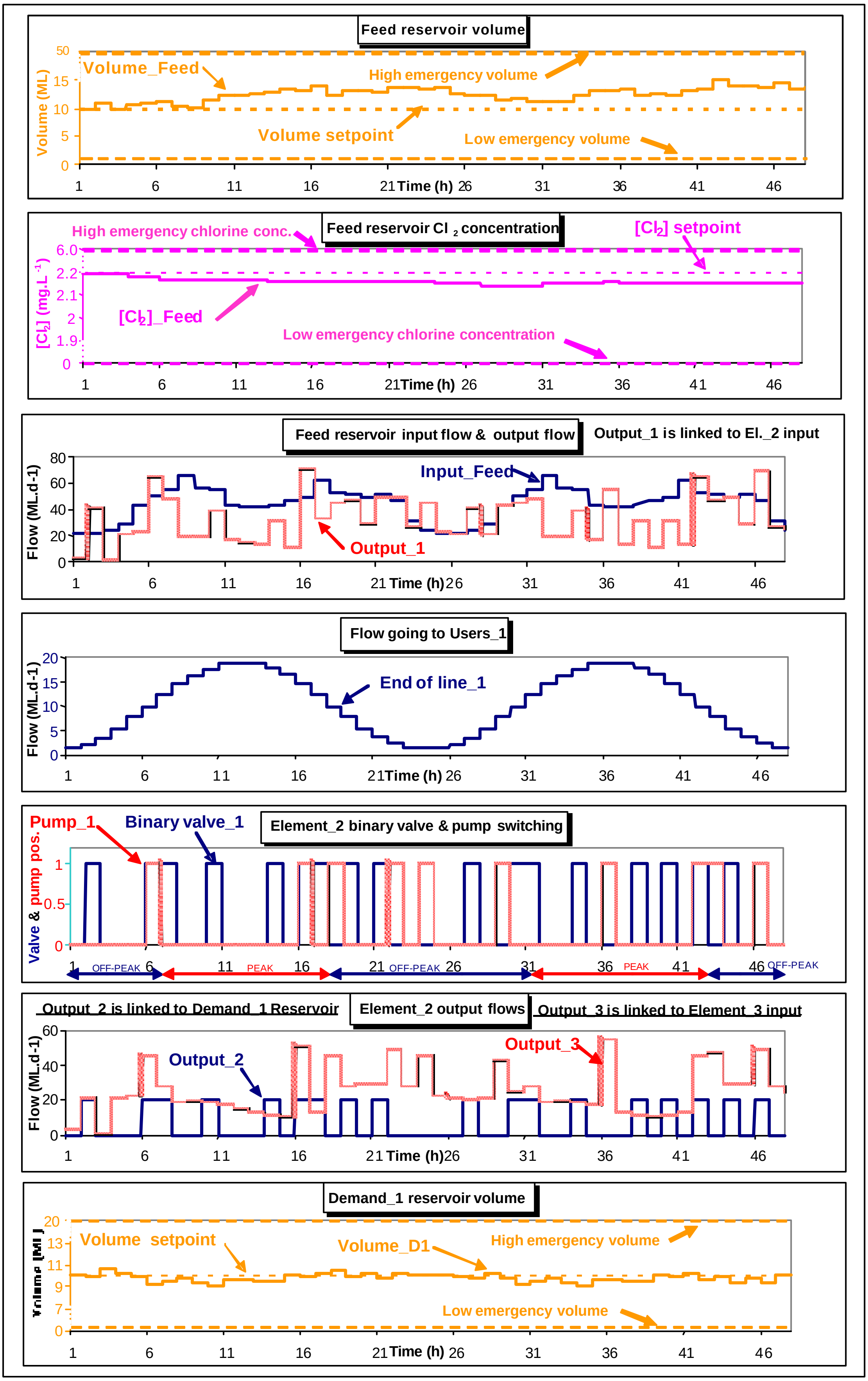
B.1 Standard representation of the 10-element configuration

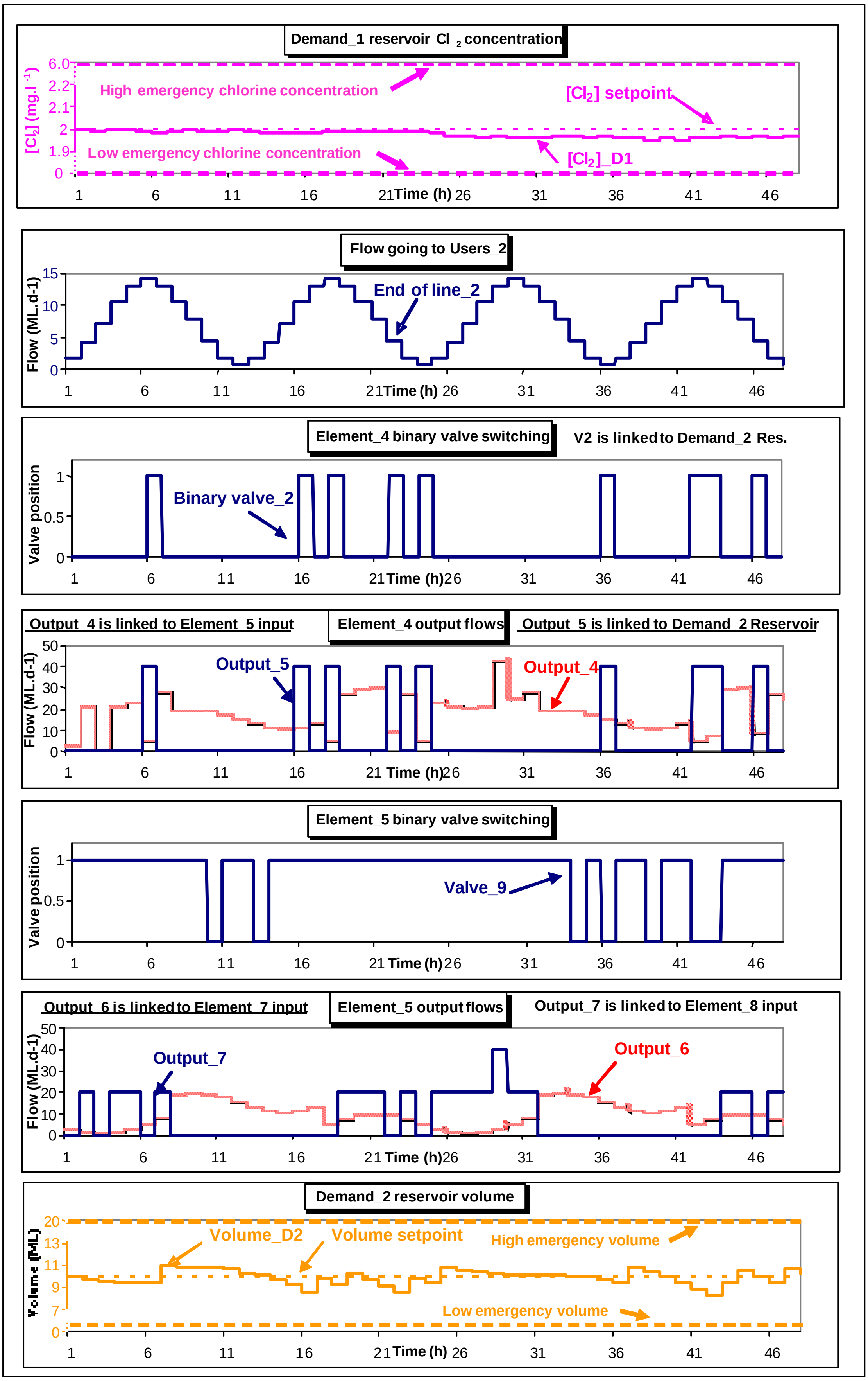


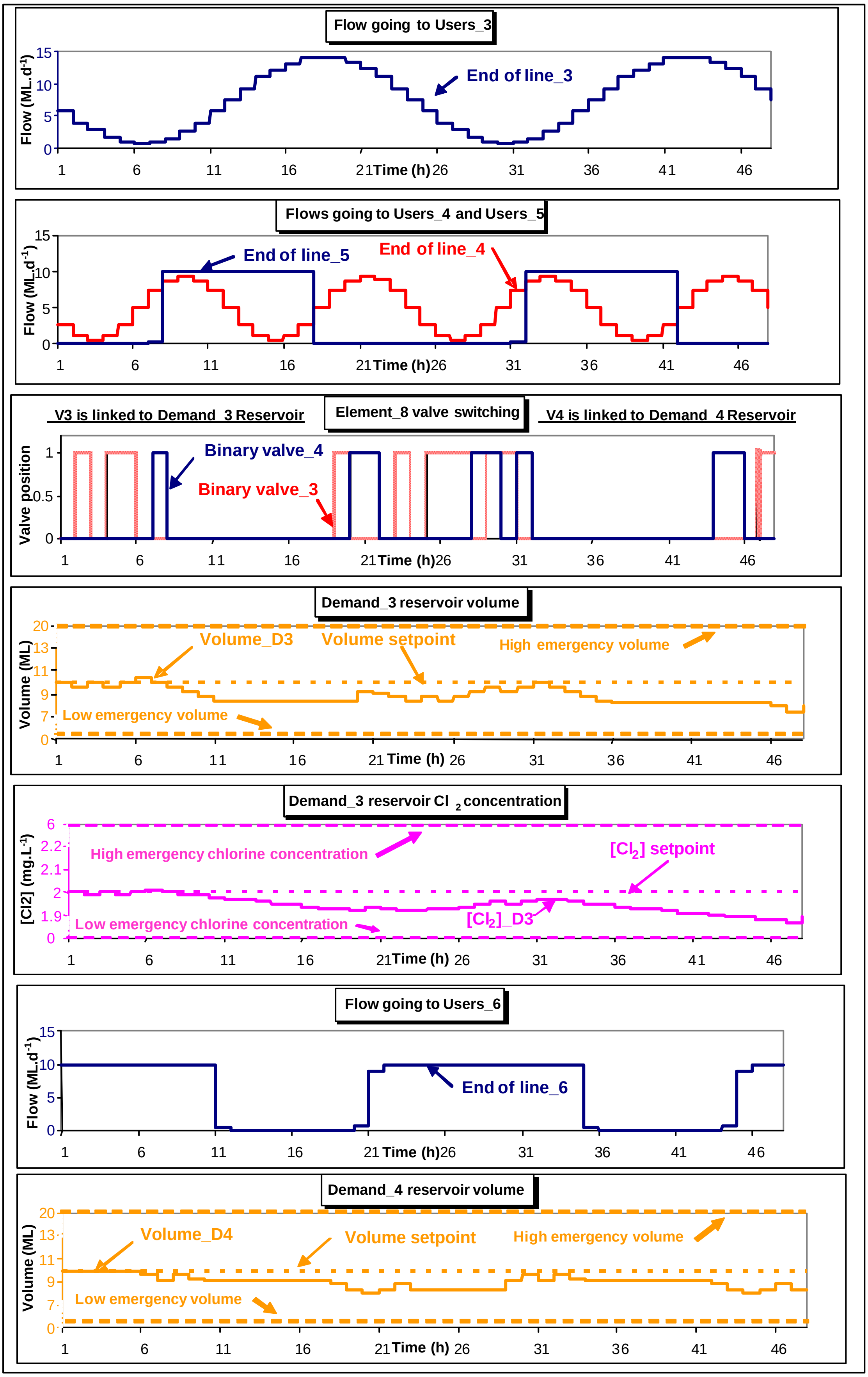
B.2 Schematic representation of the 10-element configuration

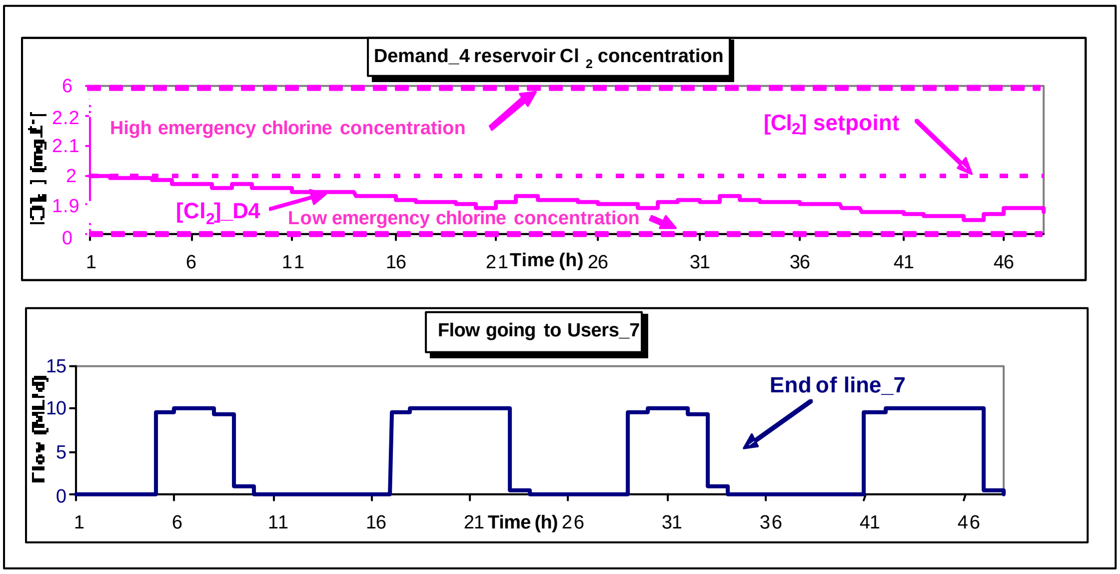


B.3 Results obtained with the MILP solver CPLEX



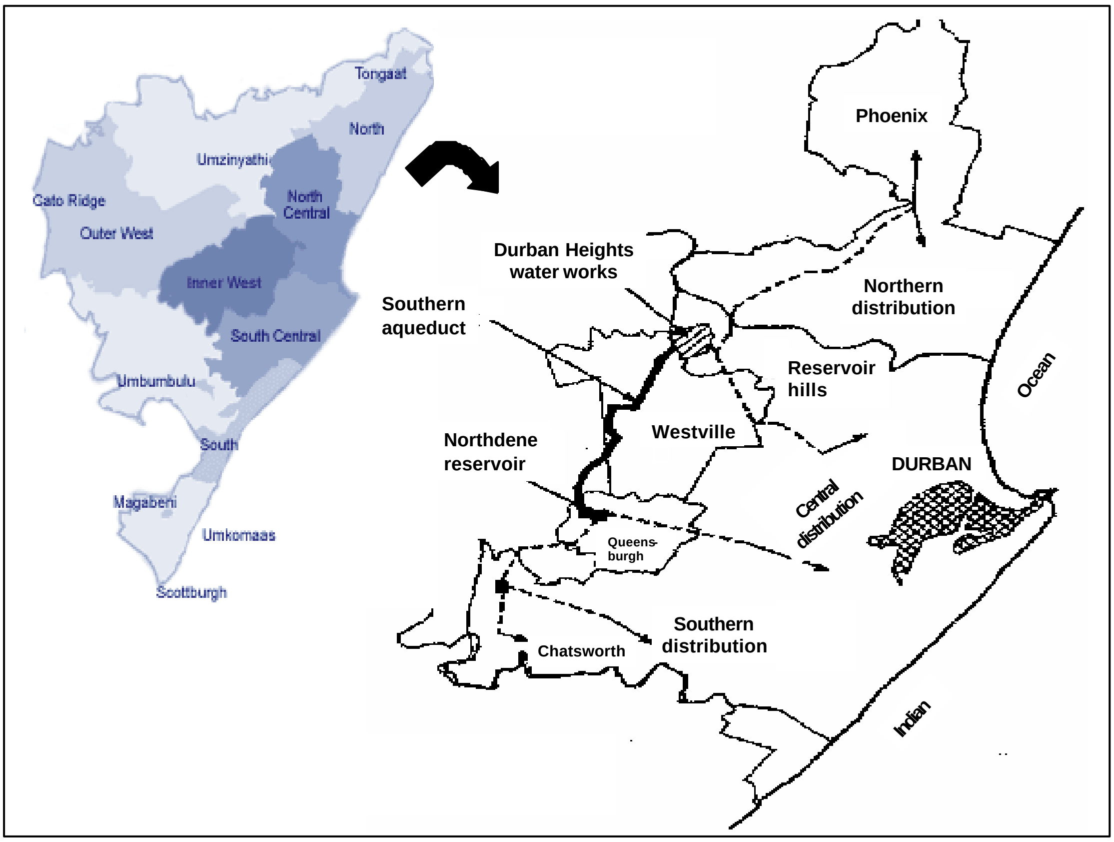

Demand_1 reservoir Cl $_2$ concentration
[Cl2] (mg.l-1)
High emergency chlorine concentration
[Cl2] setpoint
Low emergency chlorine concentration
[Cl2]_D1
Time (h)
Flow going to Users_2
Flow (ML.d-1)
End of line_2
Element_4 binary valve switching
V2 is linked to Demand_2 Res.
Valve position
Binary valve_2
Output_4 is linked to Element_5 input
Element_4 output flows
Output_5 is linked to Demand_2 Reservoir
Output_5
Output_4
Element_5 binary valve switching
Valve_9
Output_6 is linked to Element_7 input
Element_5 output flows
Output_7 is linked to Element_8 input
Output_7
Output_6
Demand_2 reservoir volume
Volume (ML)
Volume_D2
Volume setpoint
High emergency volume
Low emergency volume


Flow going to Users_3
Flow (ML.d-1)
End of line_3
Time (h)
Flows going to Users_4 and Users_5
End of line_5
End of line_4
V3 is linked to Demand_3 Reservoir
Element_8 valve switching
V4 is linked to Demand_4 Reservoir
Valve position
Binary valve_4
Binary valve_3
Demand_3 reservoir volume
Volume (ML)
Volume_D3
Volume setpoint
High emergency volume
Low emergency volume
Demand_3 reservoir Cl2 concentration
[Cl2] (mg.L-1)
High emergency chlorine concentration
[Cl2] setpoint
[Cl2]_D3
Low emergency chlorine concentration
Flow going to Users_6
End of line_6
Demand_4 reservoir volume
Volume_D4

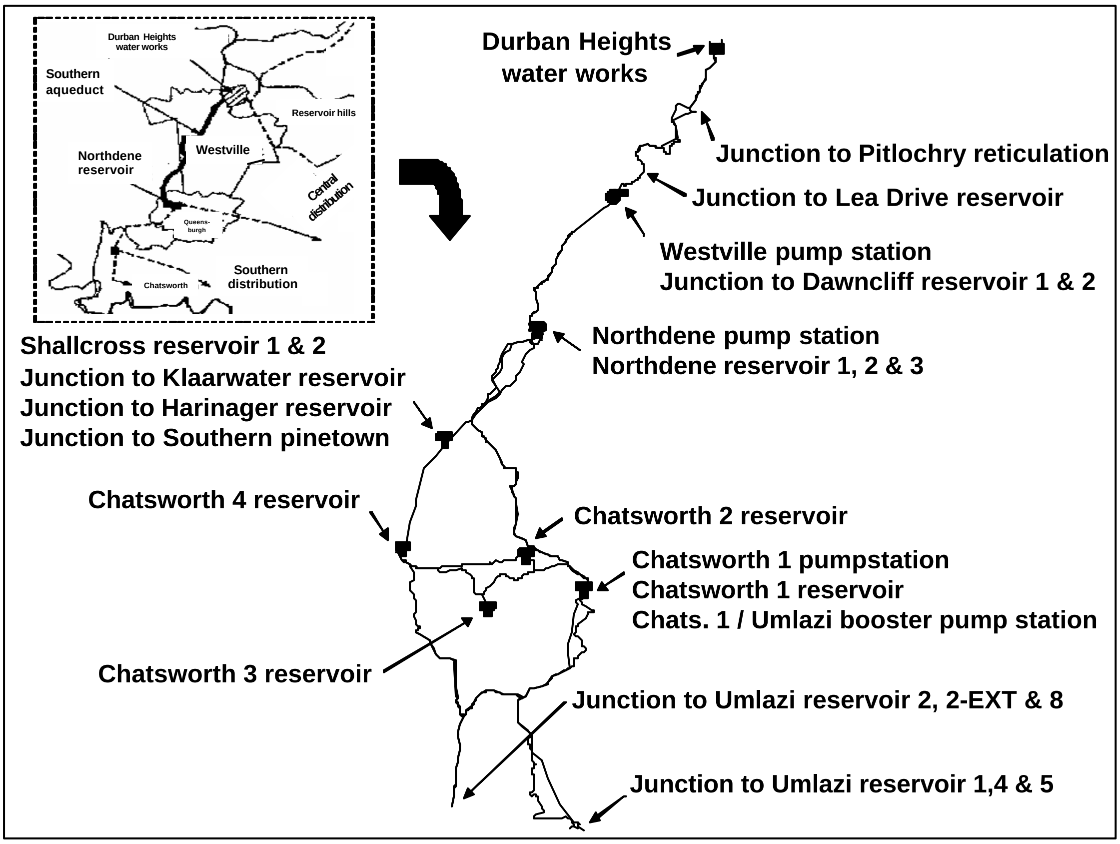

Demand_4 reservoir Cl 2 concentration
6
2.2
2.1
2
1.9
0
High emergency chlorine concentration
[Cl2] setpoint
[Cl2]_D4
Low emergency chlorine concentration
1 6 11 16 21 Time (h) 26 31 36 41 46
Flow going to Users_7
15
10
5
0
Flow [ML/d]
End of line_7
1 6 11 16 21 Time (h) 26 31 36 41 46

APPENDIX C – LARGE-SCALE APPLICATION: SCADA REPRESENTATION OF THE SOUTHERN AQUEDUCT

C.1 Localisation of the Southern Aqueduct within the Durban Metropolitan Area



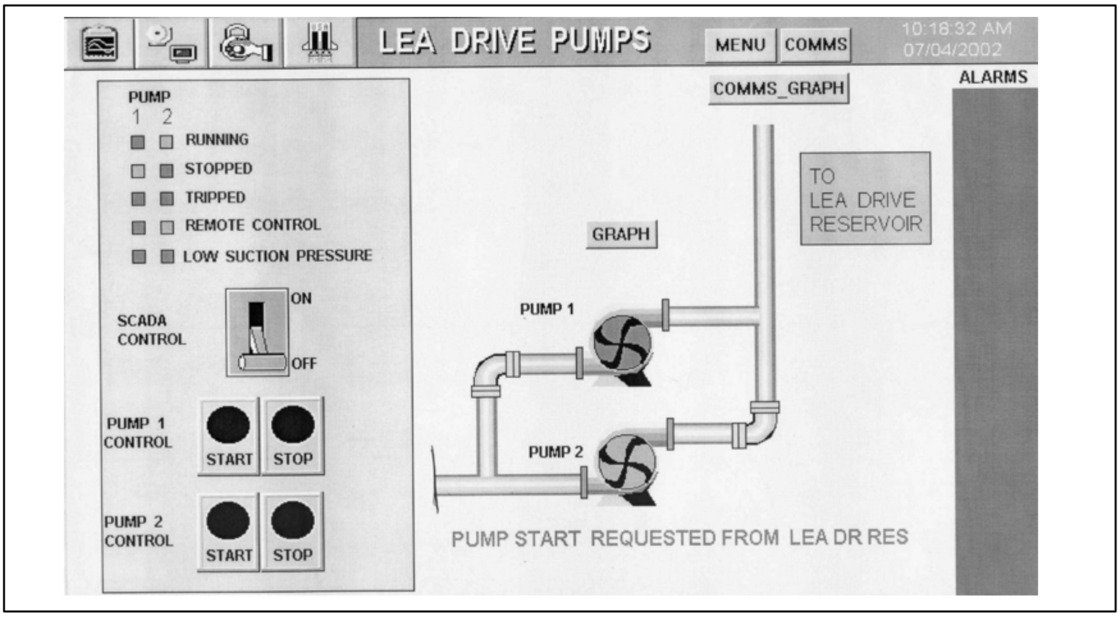
C.2 Schematic representation of the Southern Aqueduct



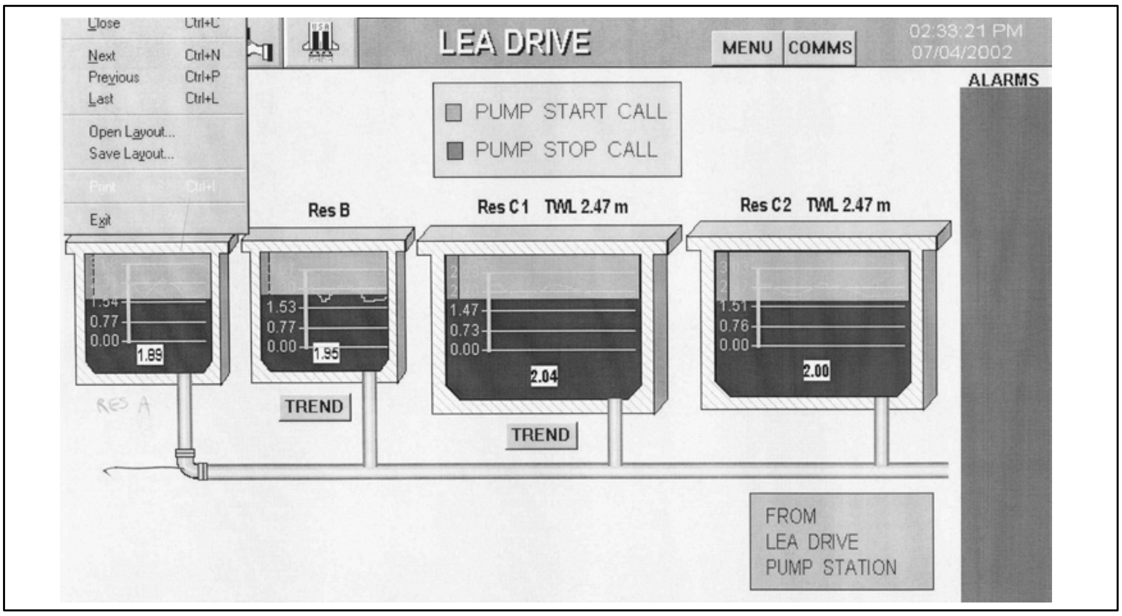
C.3 SCADA representation of the sub-systems of the Southern Aqueduct

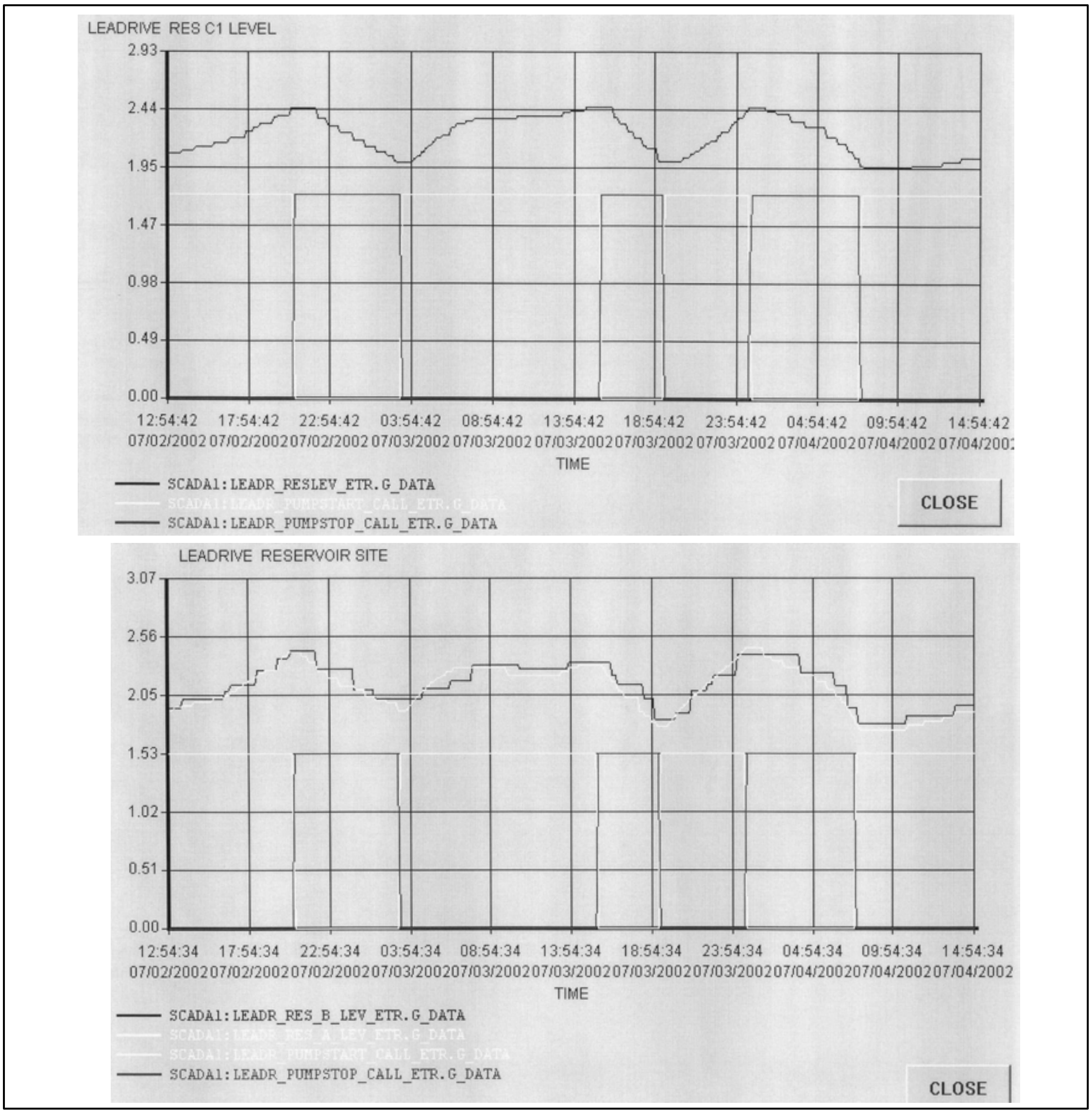
All the snapshots presented in this section are courtesy of eThekwini Water Services.

C.3.1 Lea Drive pump station

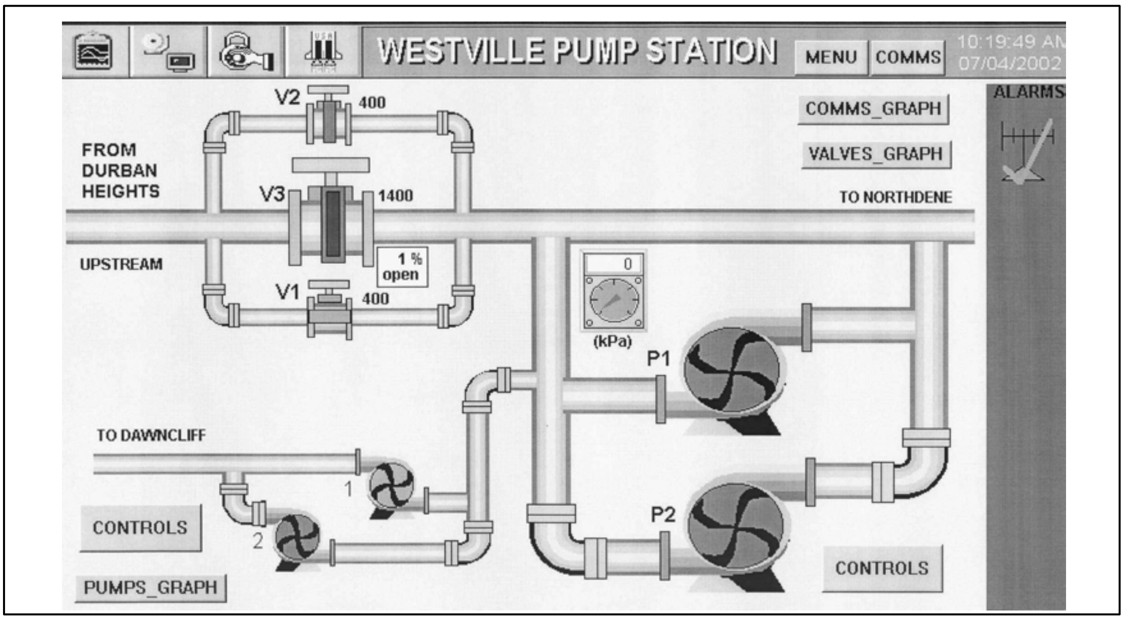


C.3.2 Lea Drive reservoirs and reticulation system

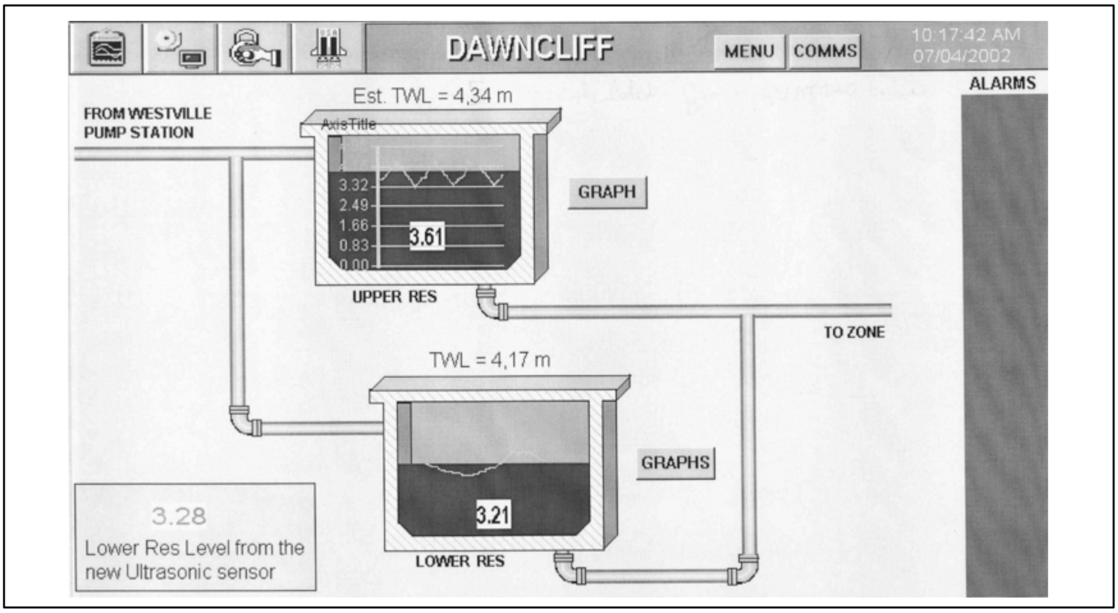


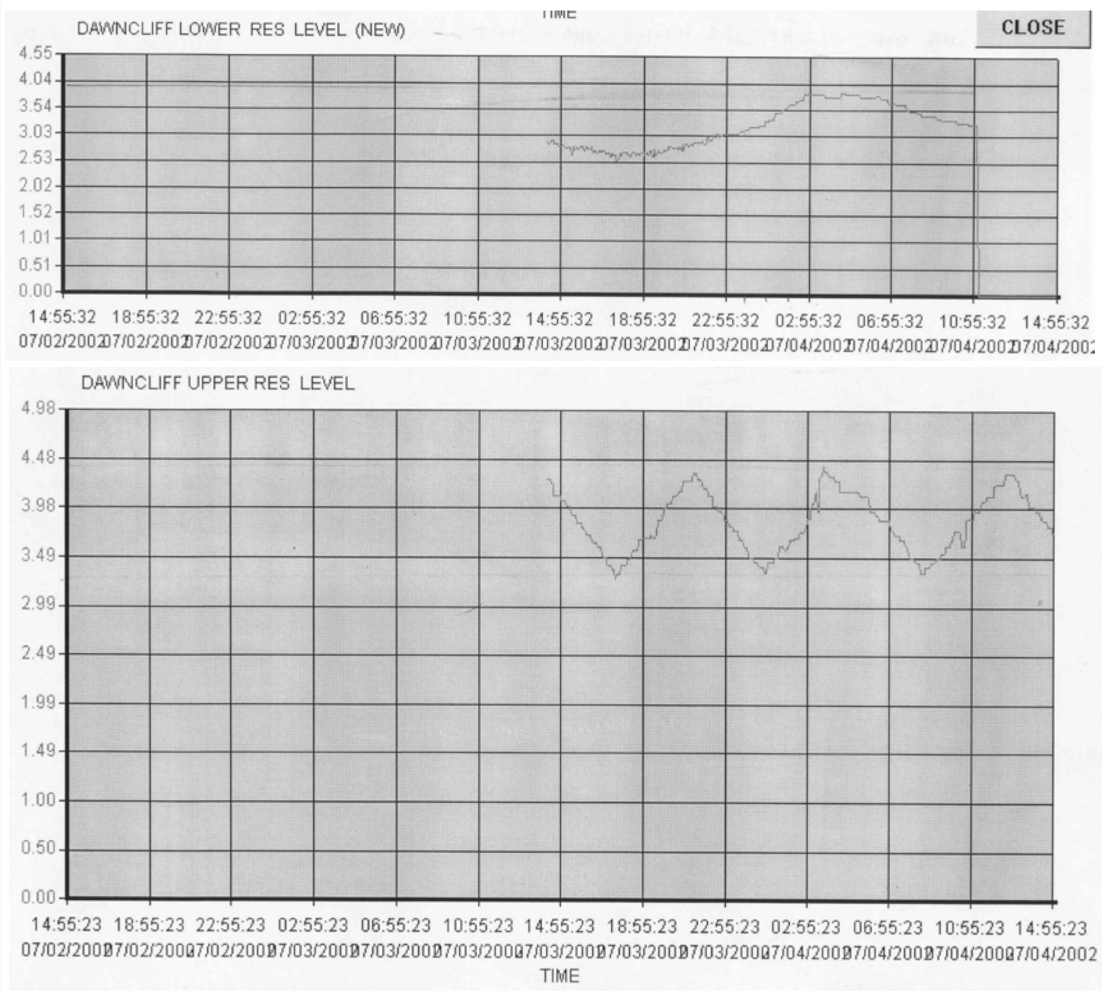


C.3.3 Westville pump station

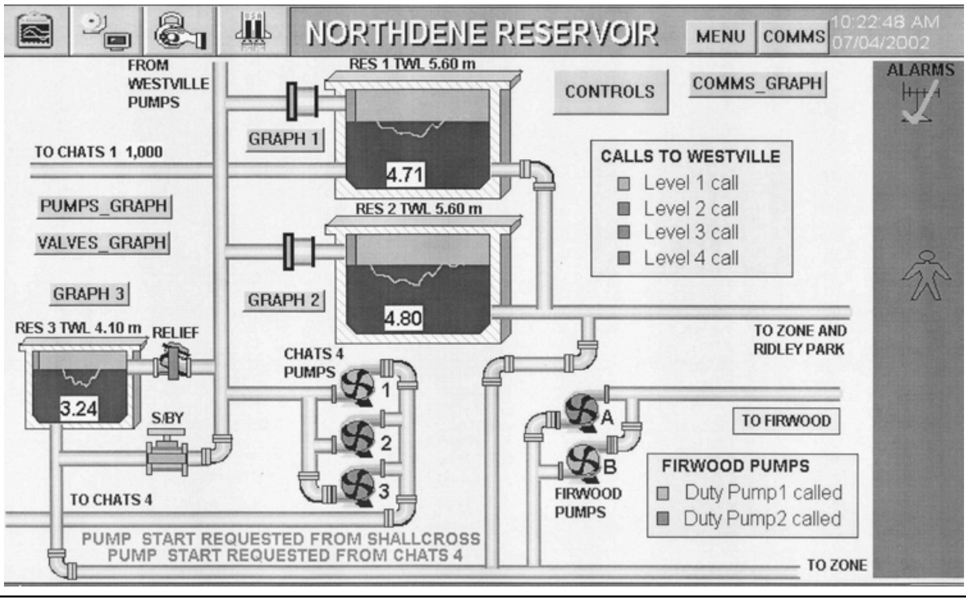


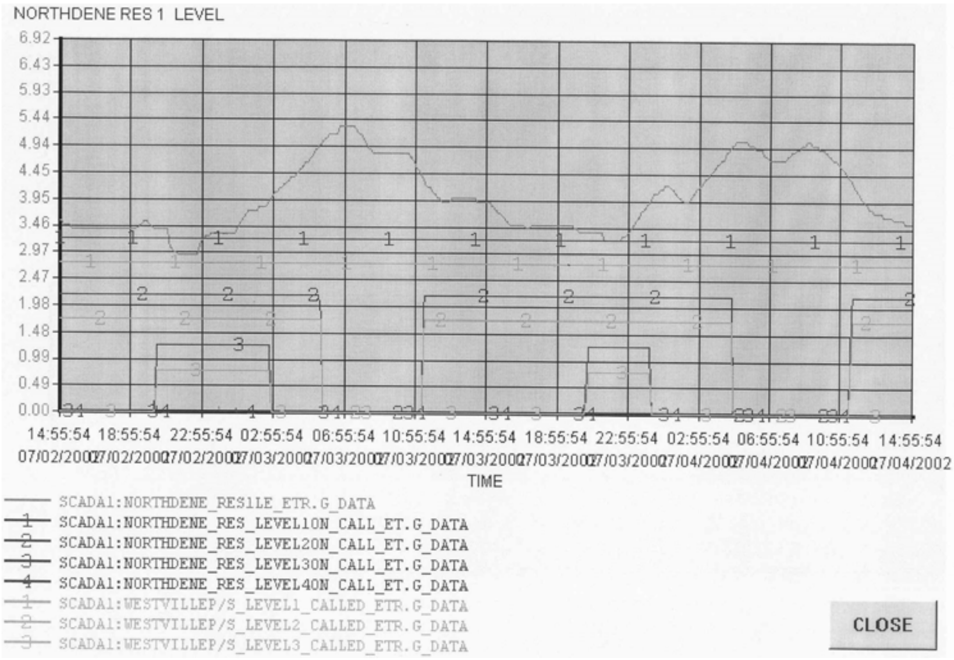
C.3.4 Dawncliff reservoirs and reticulation system

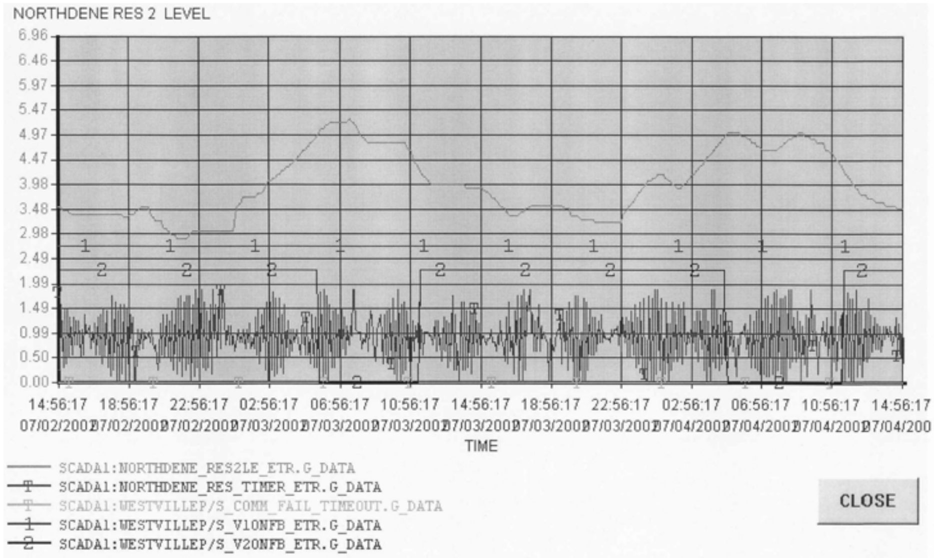


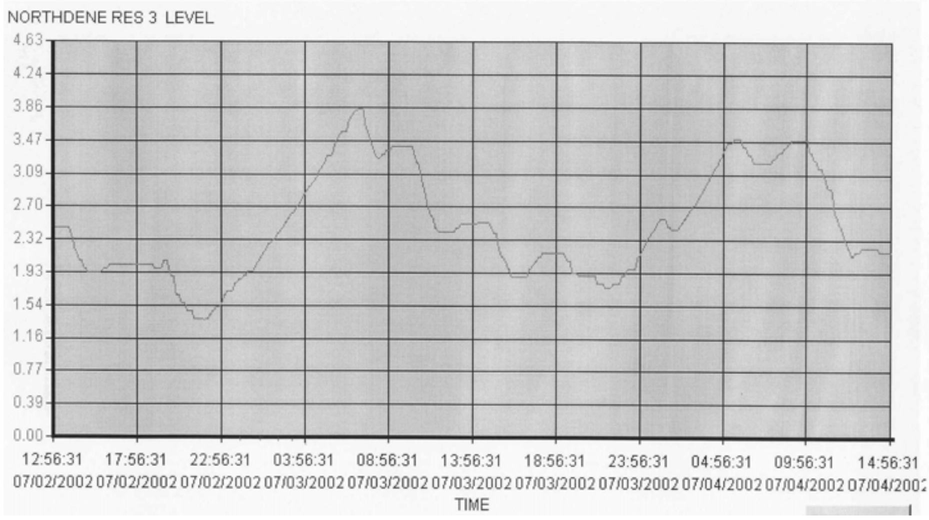


C.3.5 Northdene reservoirs and reticulation system

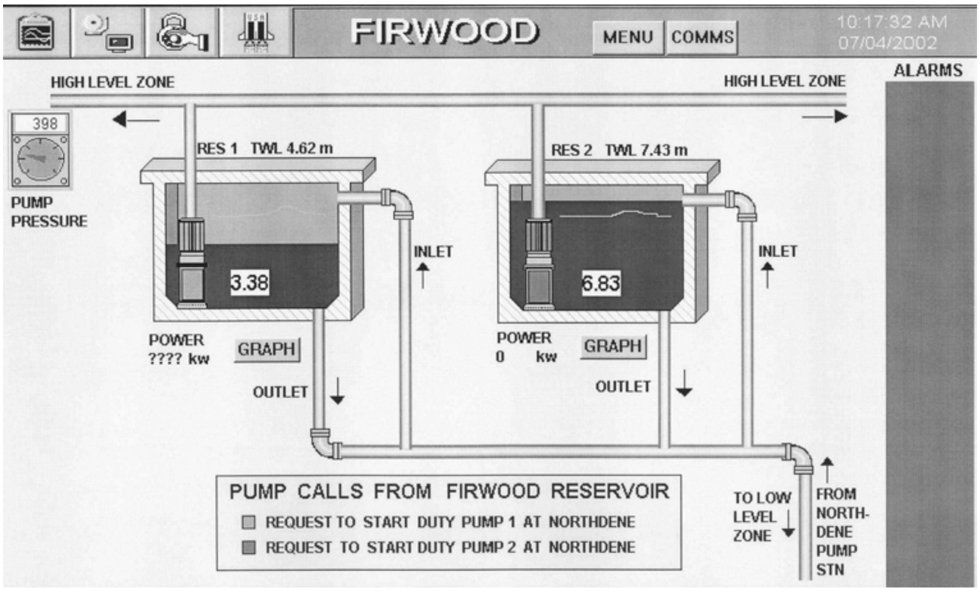


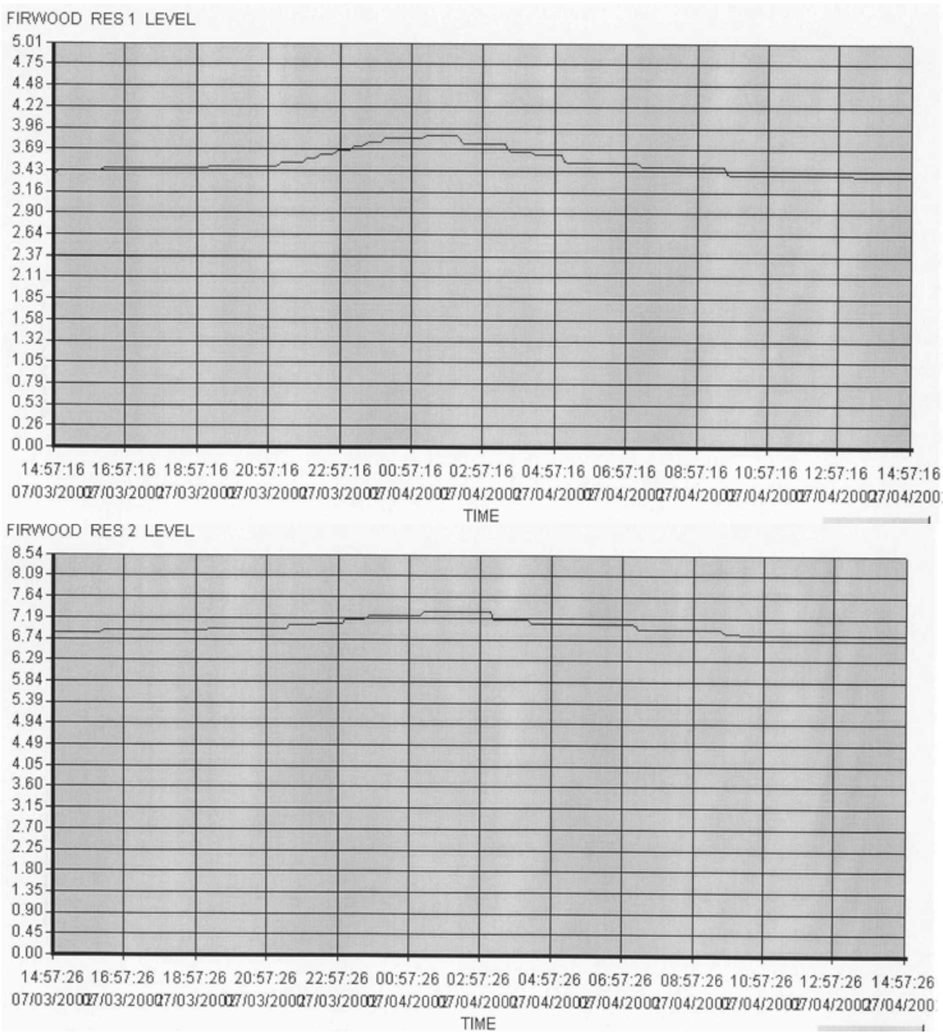

NORTHDENE RES 1 LEVEL
6.92
6.43
5.93
5.44
4.94
4.45
3.95
3.46
2.97
2.47
1.98
1.48
0.99
0.49
0.00
14:55:54 18:55:54 22:55:54 02:55:54 06:55:54 10:55:54 14:55:54 18:55:54 22:55:54 02:55:54 06:55:54 10:55:54 14:55:54
TIME
SCADA1:NORTHDENE_RES1LE_ETR.G_DATA
SCADA1:NORTHDENE_RES_LEVEL1ON_CALL_ET.G_DATA
SCADA1:NORTHDENE_RES_LEVEL2ON_CALL_ET.G_DATA
SCADA1:NORTHDENE_RES_LEVEL3ON_CALL_ET.G_DATA
SCADA1:NORTHDENE_RES_LEVEL4ON_CALL_ET.G_DATA
SCADA1:WESTVILLEP/S_LEVEL1_CALLED_ETR.G_DATA
SCADA1:WESTVILLEP/S_LEVEL2_CALLED_ETR.G_DATA
SCADA1:WESTVILLEP/S_LEVEL3_CALLED_ETR.G_DATA
CLOSE


NORTHDENE RES 2 LEVEL
6.96
6.46
5.97
5.47
4.97
4.47
3.98
3.48
2.98
2.49
1.99
1.49
0.99
0.50
0.00
14:56:17 18:56:17 22:56:17 02:56:17 06:56:17 10:56:17 14:56:17 18:56:17 22:56:17 02:56:17 06:56:17 10:56:17 14:56:17
TIME
SCADA1:NORTHDENE_RES2LE_ETR.G_DATA
SCADA1:NORTHDENE_RES_TIMER_ETR.G_DATA
SCADA1:WESTVILLEP/S_COMM_FAIL_TIMEOUT.G_DATA
SCADA1:WESTVILLEP/S_V1ONFB_ETR.G_DATA
SCADA1:WESTVILLEP/S_V2ONFB_ETR.G_DATA
CLOSE

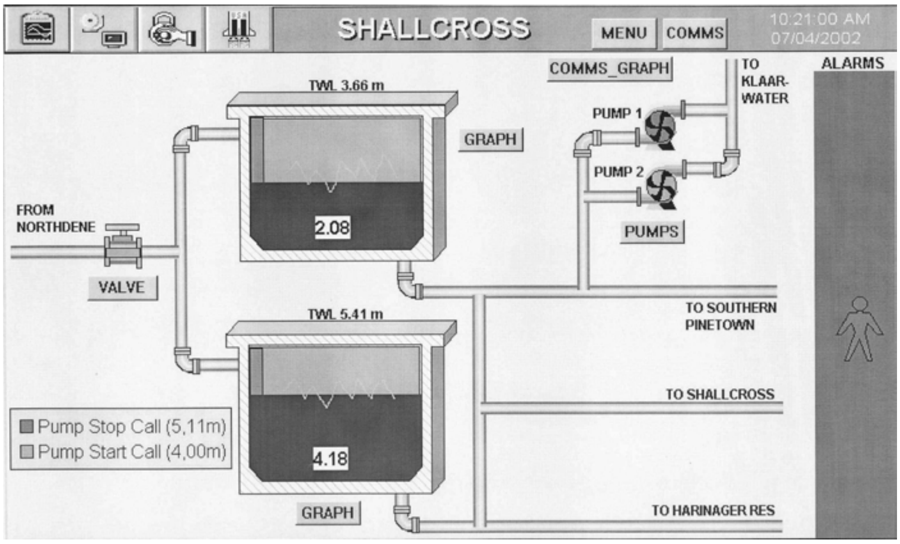

NORTHDENE RES 3 LEVEL
4.63
4.24
3.86
3.47
3.09
2.70
2.32
1.93
1.54
1.16
0.77
0.39
0.00
12:56:31 17:56:31 22:56:31 03:56:31 08:56:31 13:56:31 18:56:31 23:56:31 04:56:31 09:56:31 14:56:31
07/02/2002 07/02/2002 07/02/2002 07/03/2002 07/03/2002 07/03/2002 07/03/2002 07/03/2002 07/04/2002 07/04/2002 07/04/2002
TIME

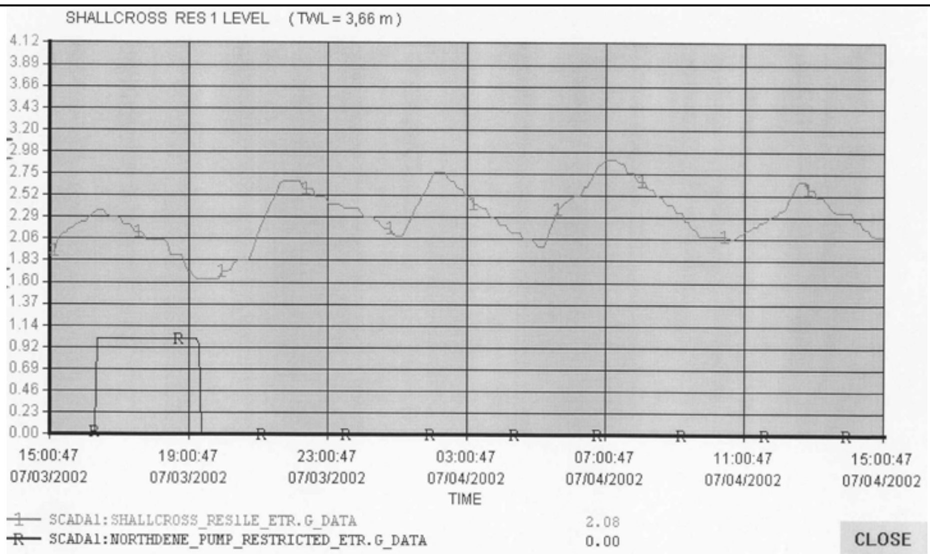
C.3.6 Firwood reservoirs and reticulation system

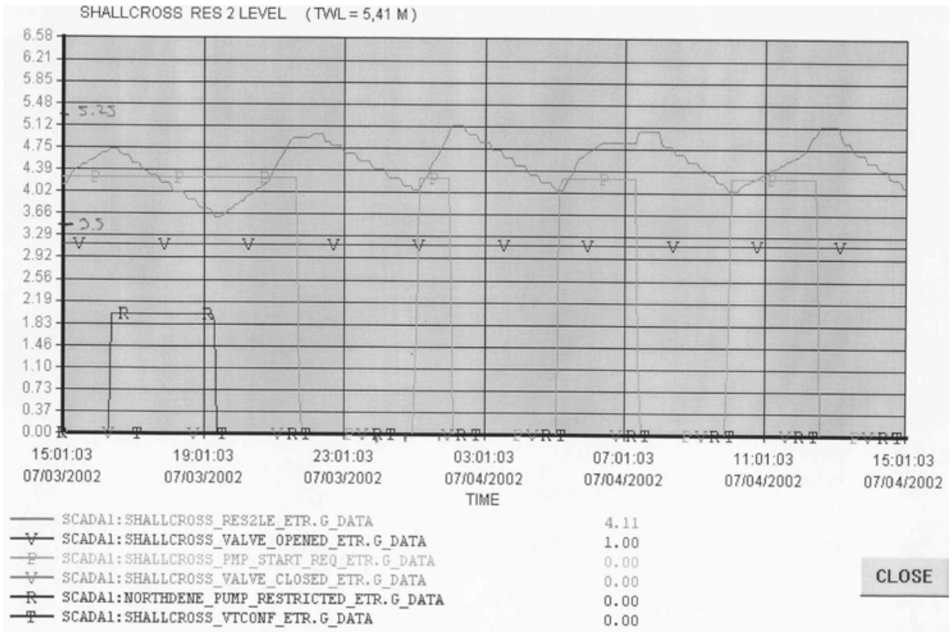




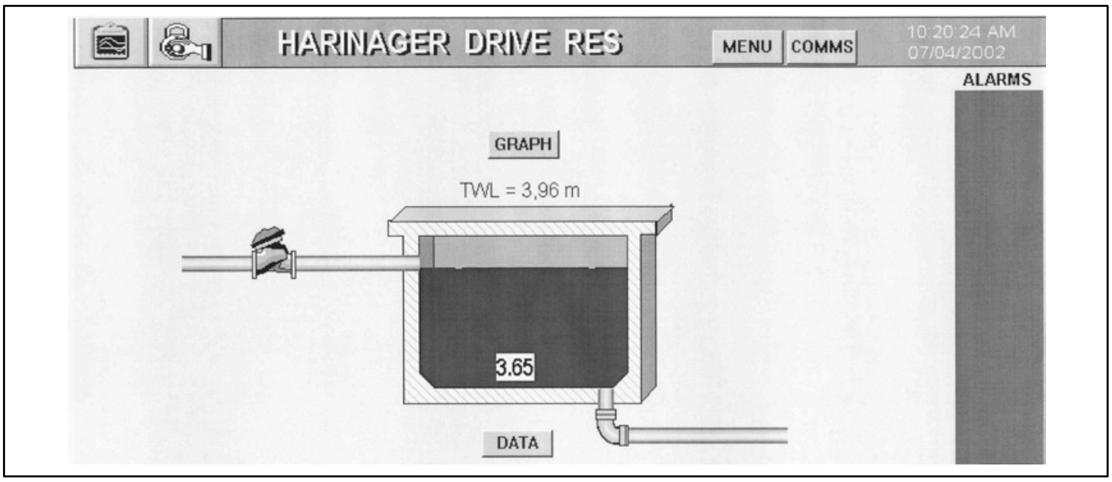
C.3.7 Shallcross reservoirs and reticulation system

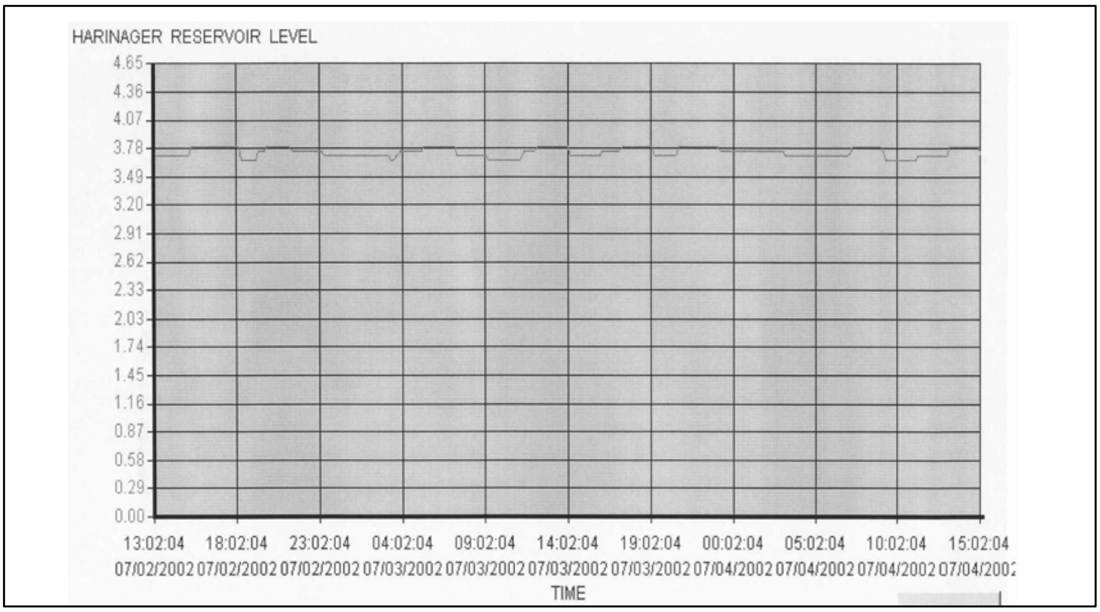




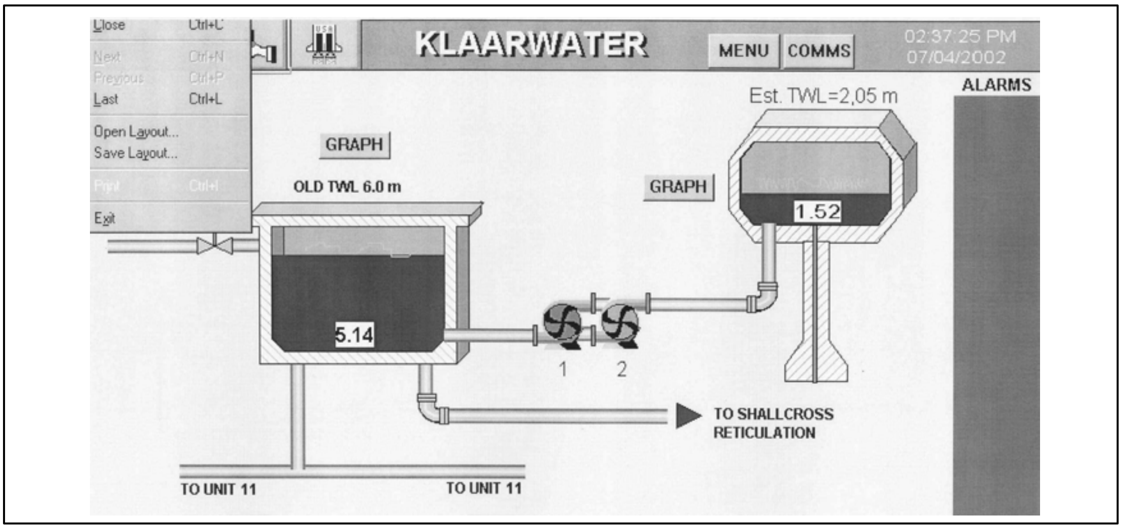


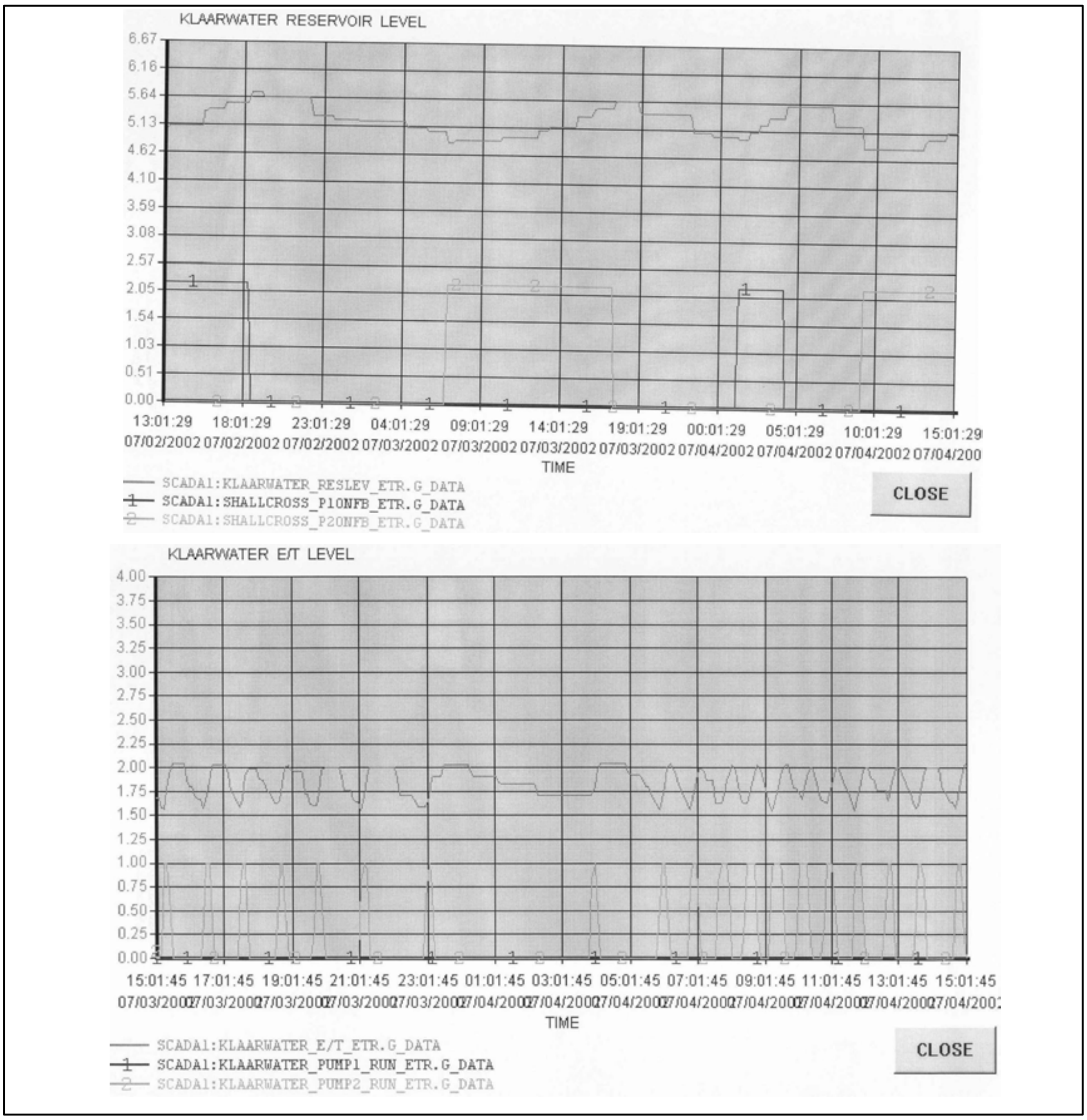
C.3.8 Harinager drive reservoir and reticulation system



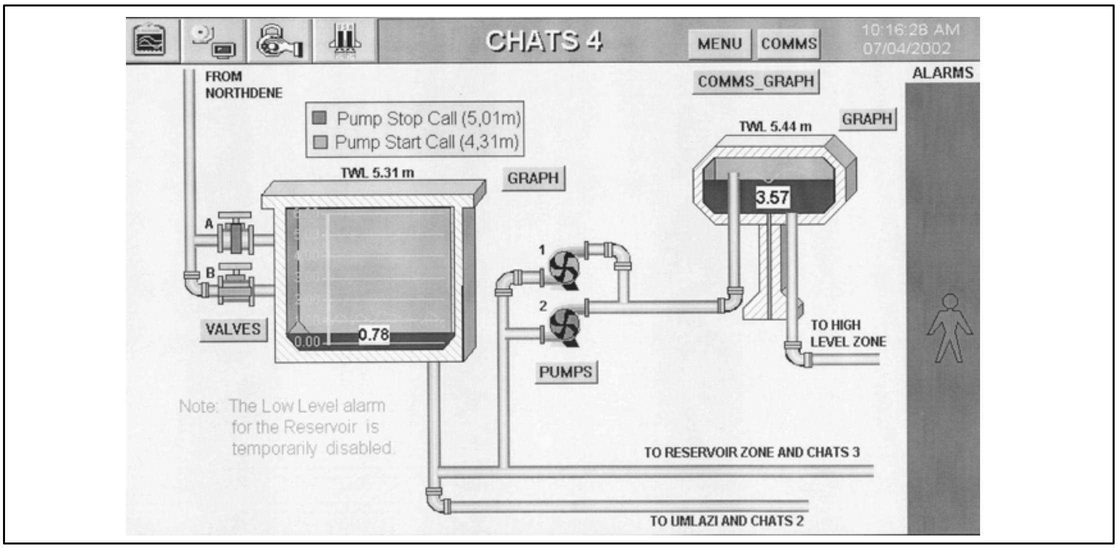


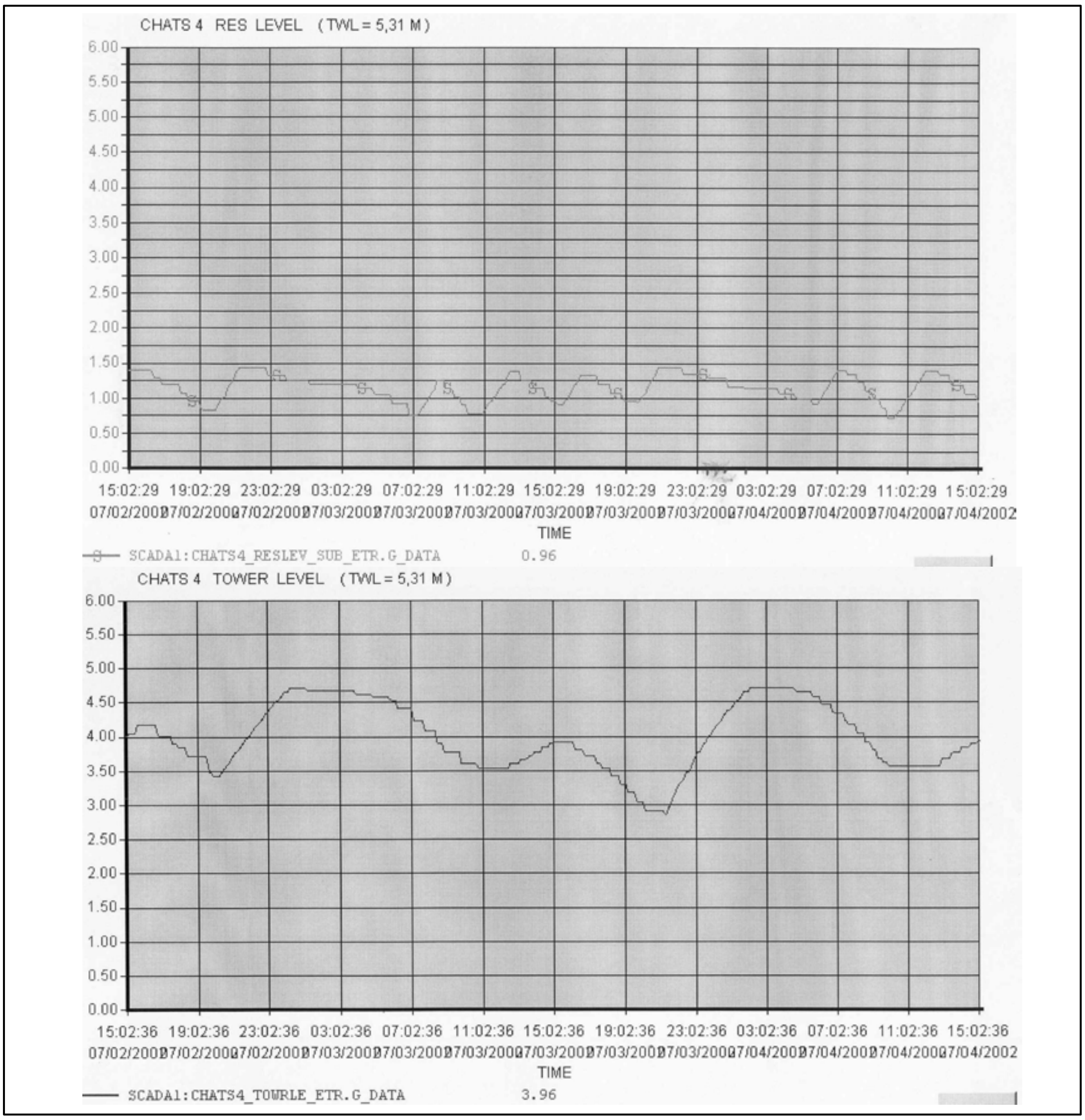
C.3.9 Klaarwater reservoir and reticulation system



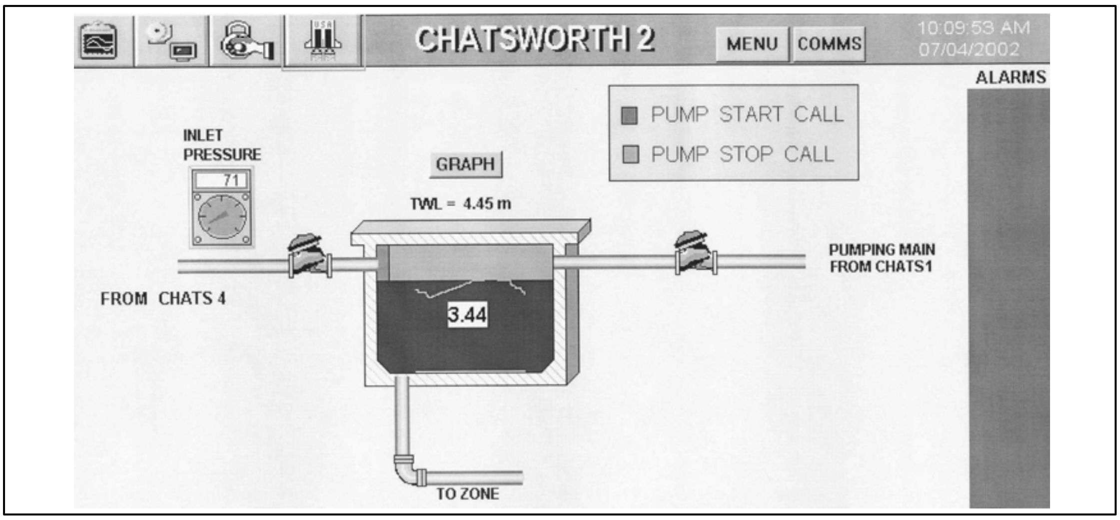


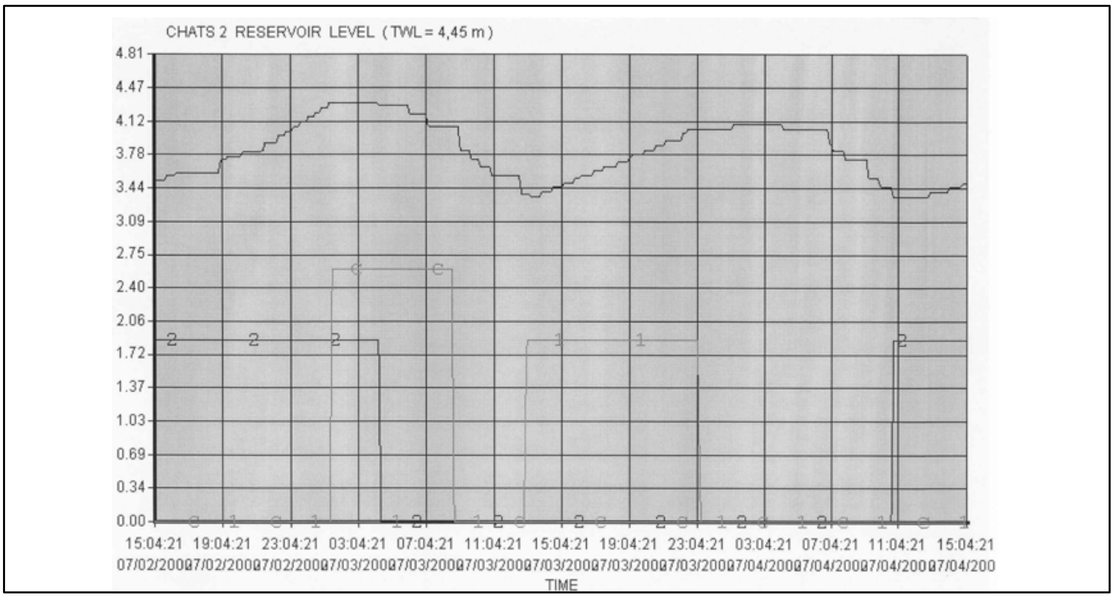
C.3.10 Chatsworth 4 reservoir and reticulation system



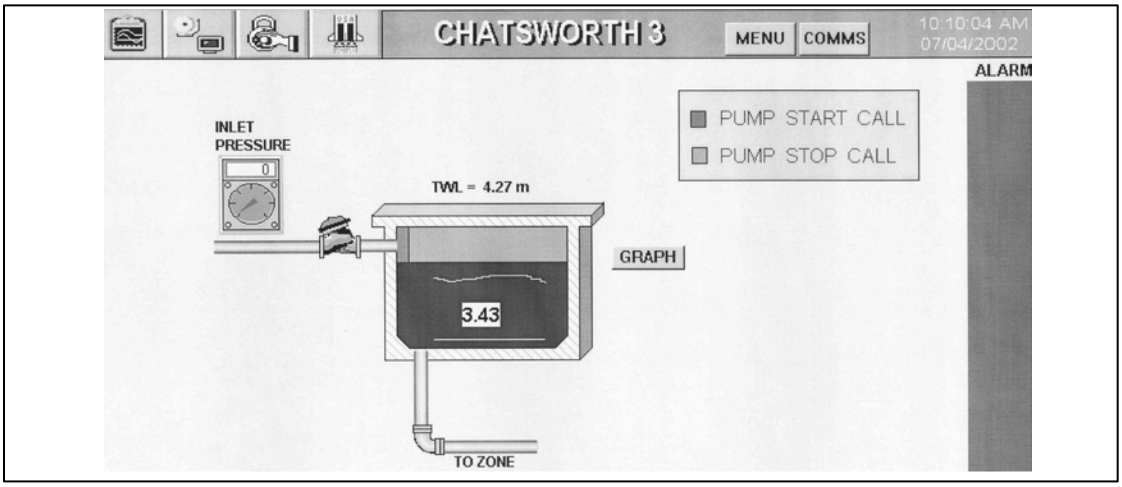


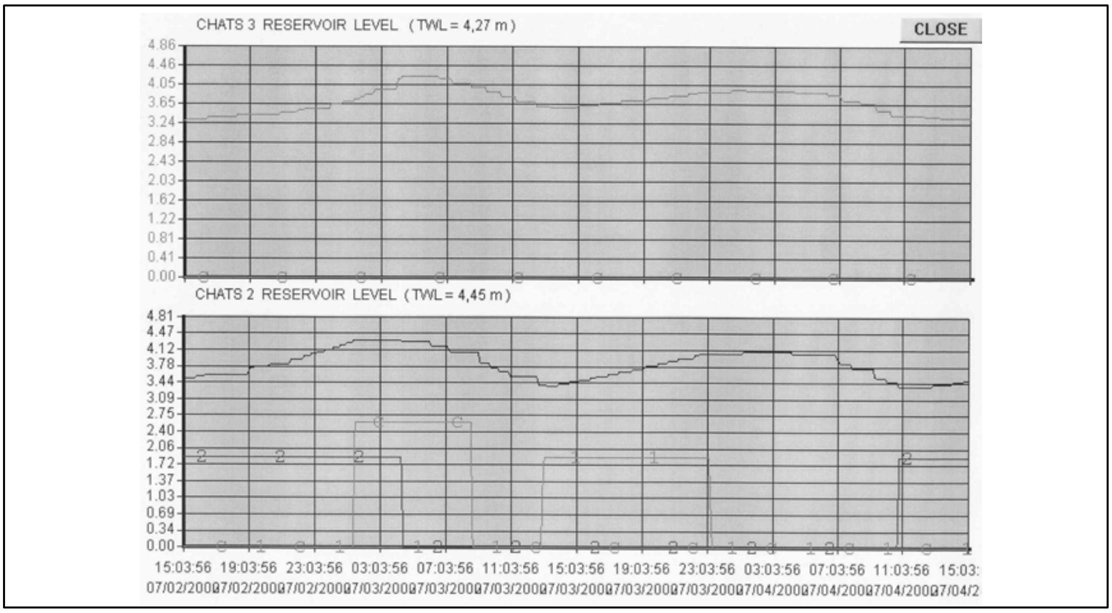
C.3.11 Chatsworth 2 reservoir and reticulation system



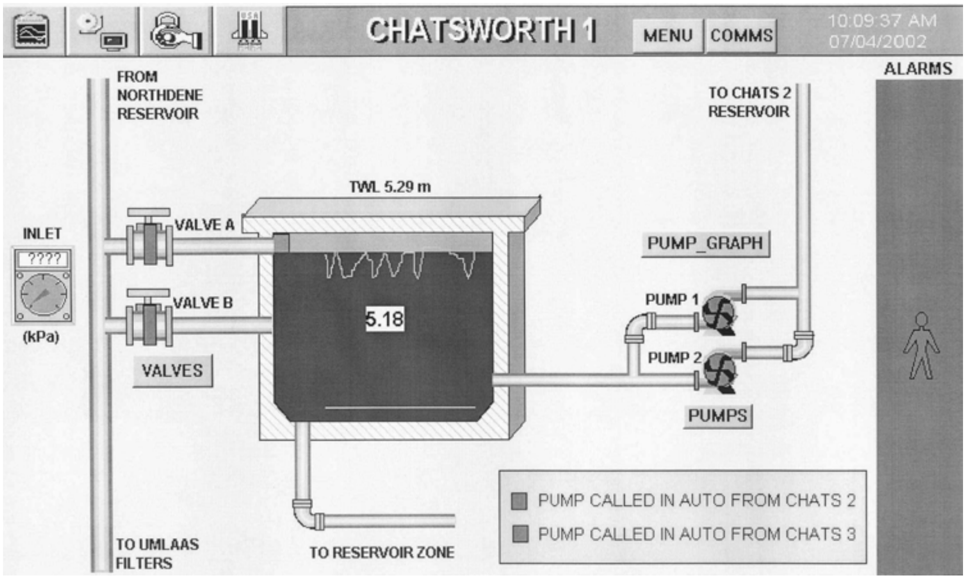


C.3.12 Chatsworth 3 reservoir and reticulation system

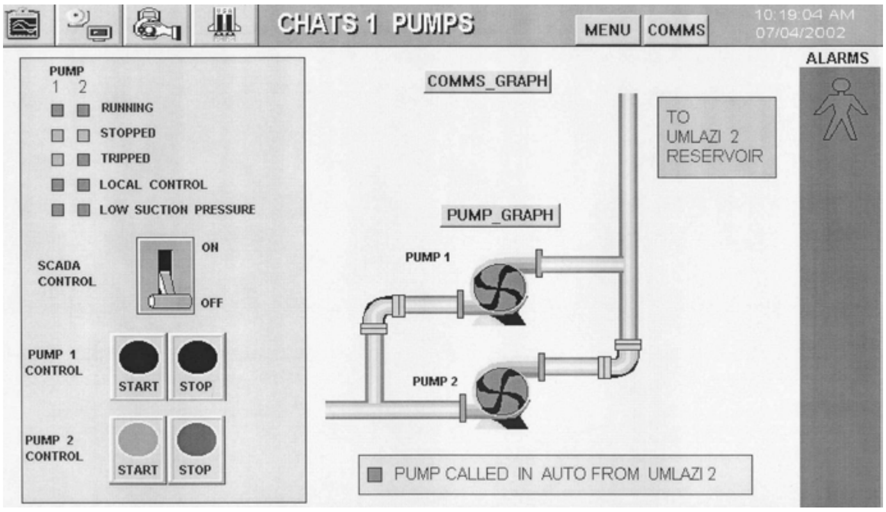


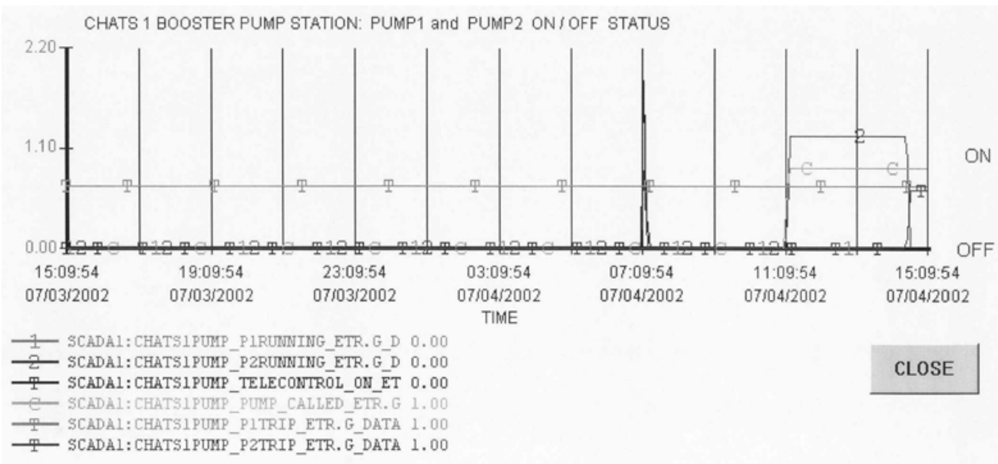


C.3.13 Chatsworth 1 reservoir and pump station



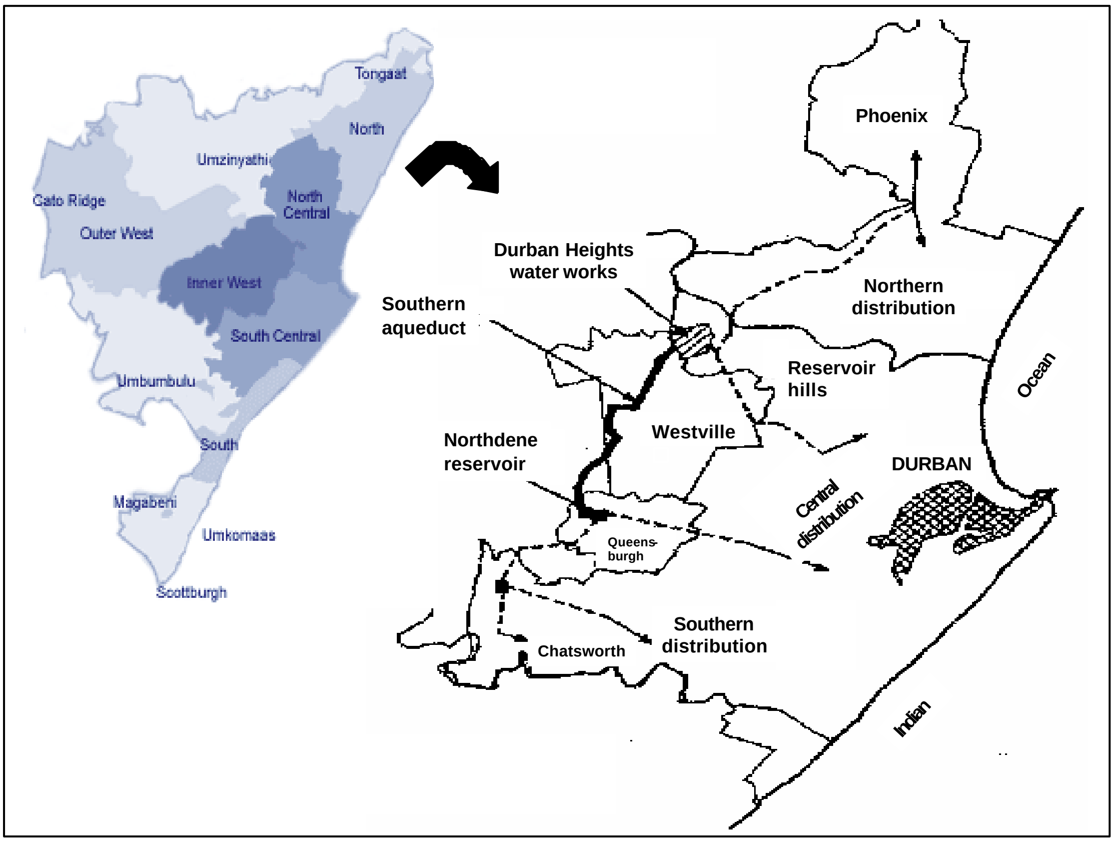
C.3.14 Chatsworth 1 / Umlazi booster pump station to Umlazi area



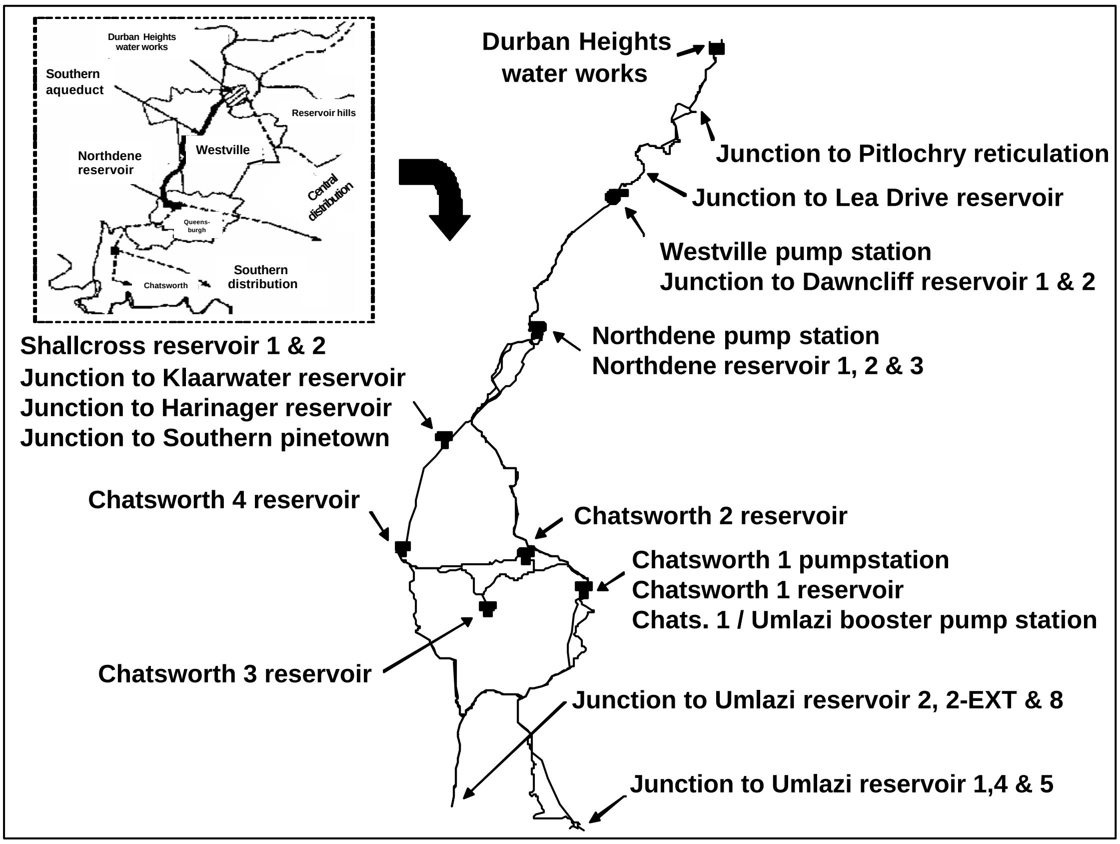


APPENDIX D – LARGE-SCALE APPLICATION: ePANET REPRESENTATION OF THE SOUTHERN AQUEDUCT

D.1 Localisation of the Southern Aqueduct within the Durban Metropolitan Area



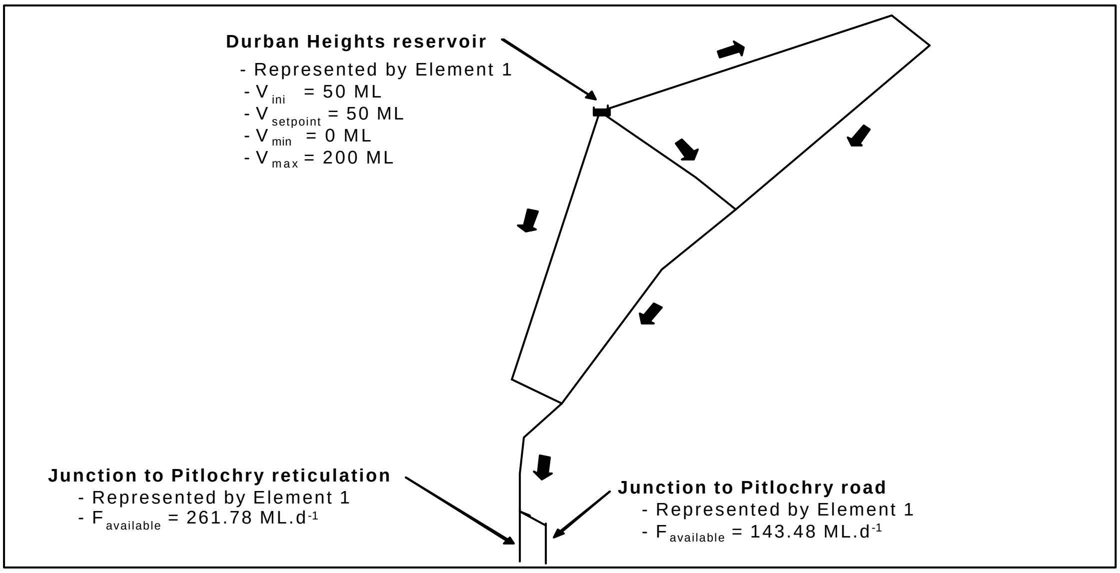
D.2 Schematic representation of the Southern Aqueduct



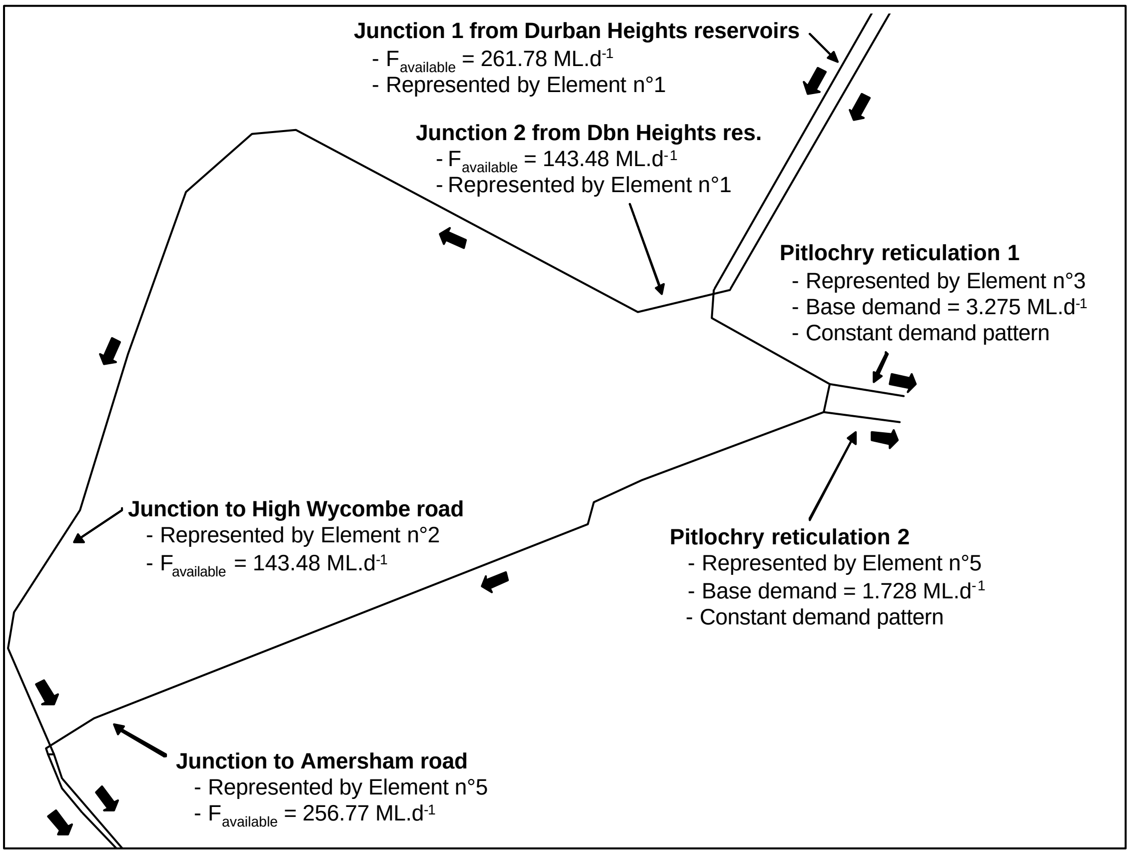
D.3 ePANET representation of the sub-systems of the Southern Aqueduct

ALL of the graphs presented in the section are courtesy of eThekwini Water Services and are relative to the first scenario of operations of the Southern Aqueduct that is detailed in Chapter 7, at the paragraph 7.1.5.1.

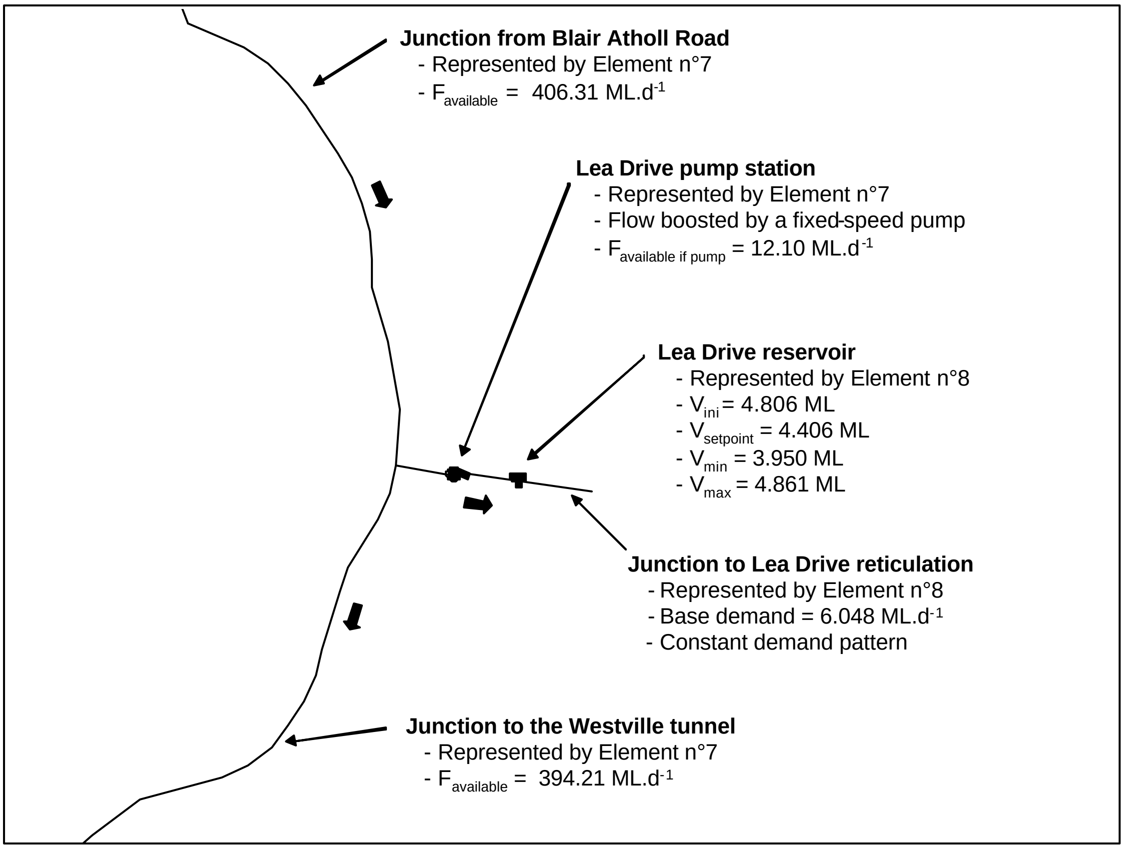
D.3.1 Durban Heights water works output



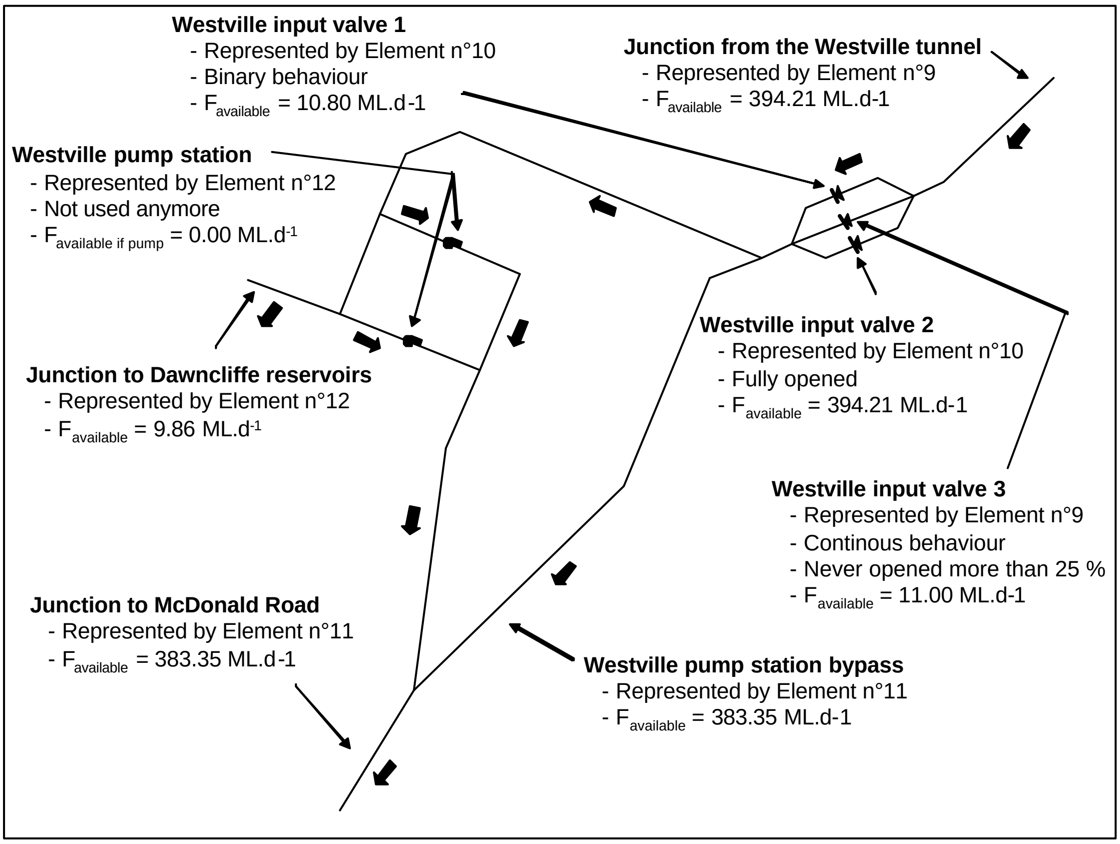
D.3.2 Pitlochry Road reticulation system



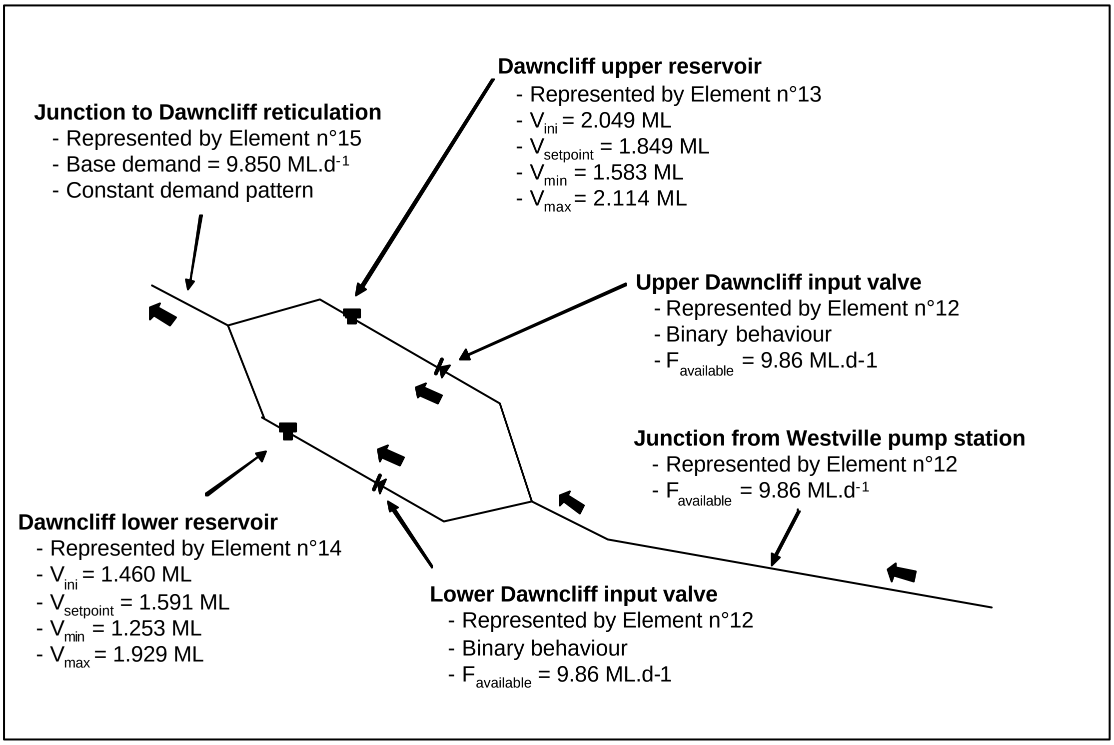
D.3.3 Lea drive reservoir and reticulation system



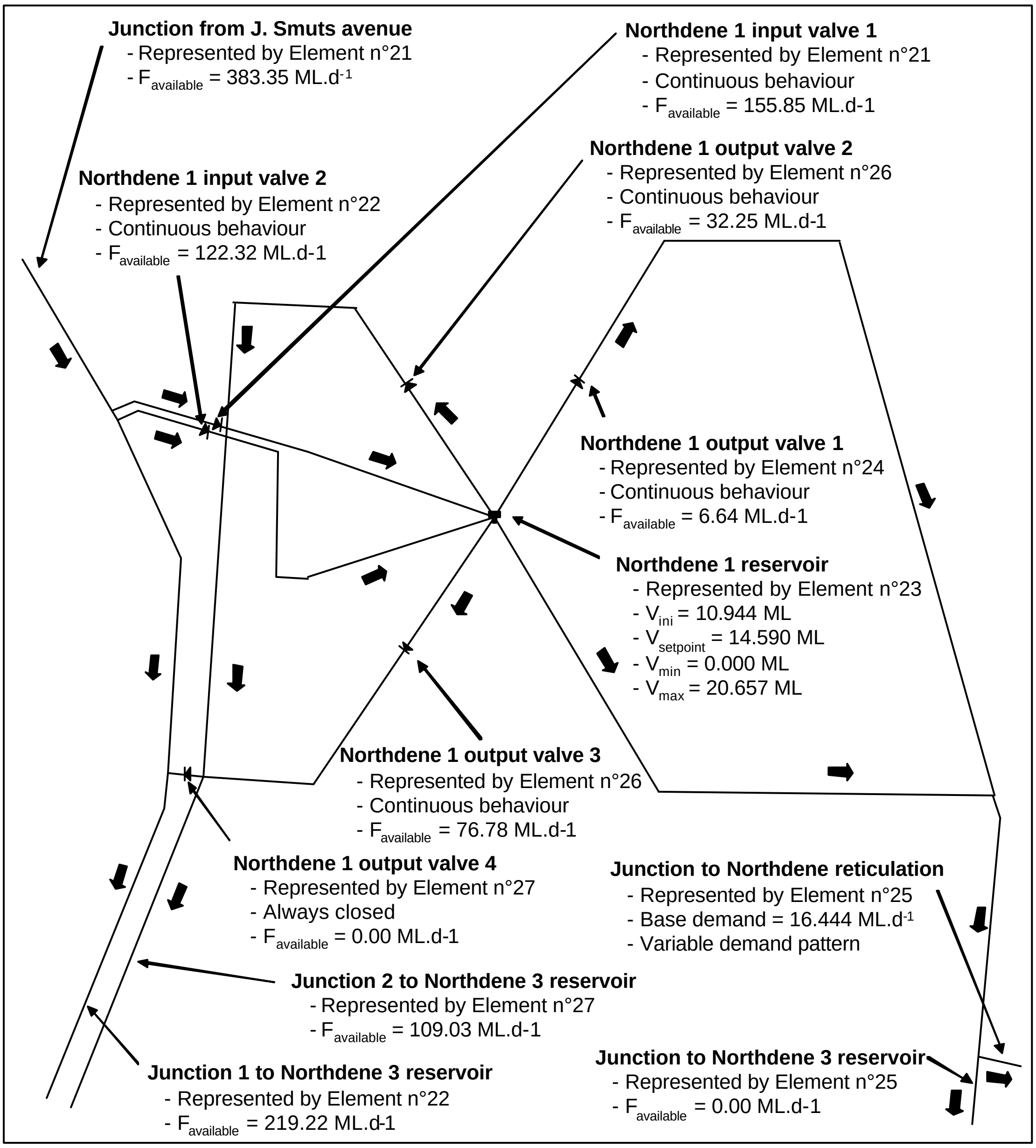
D.3.4 Westville pump station and bypass



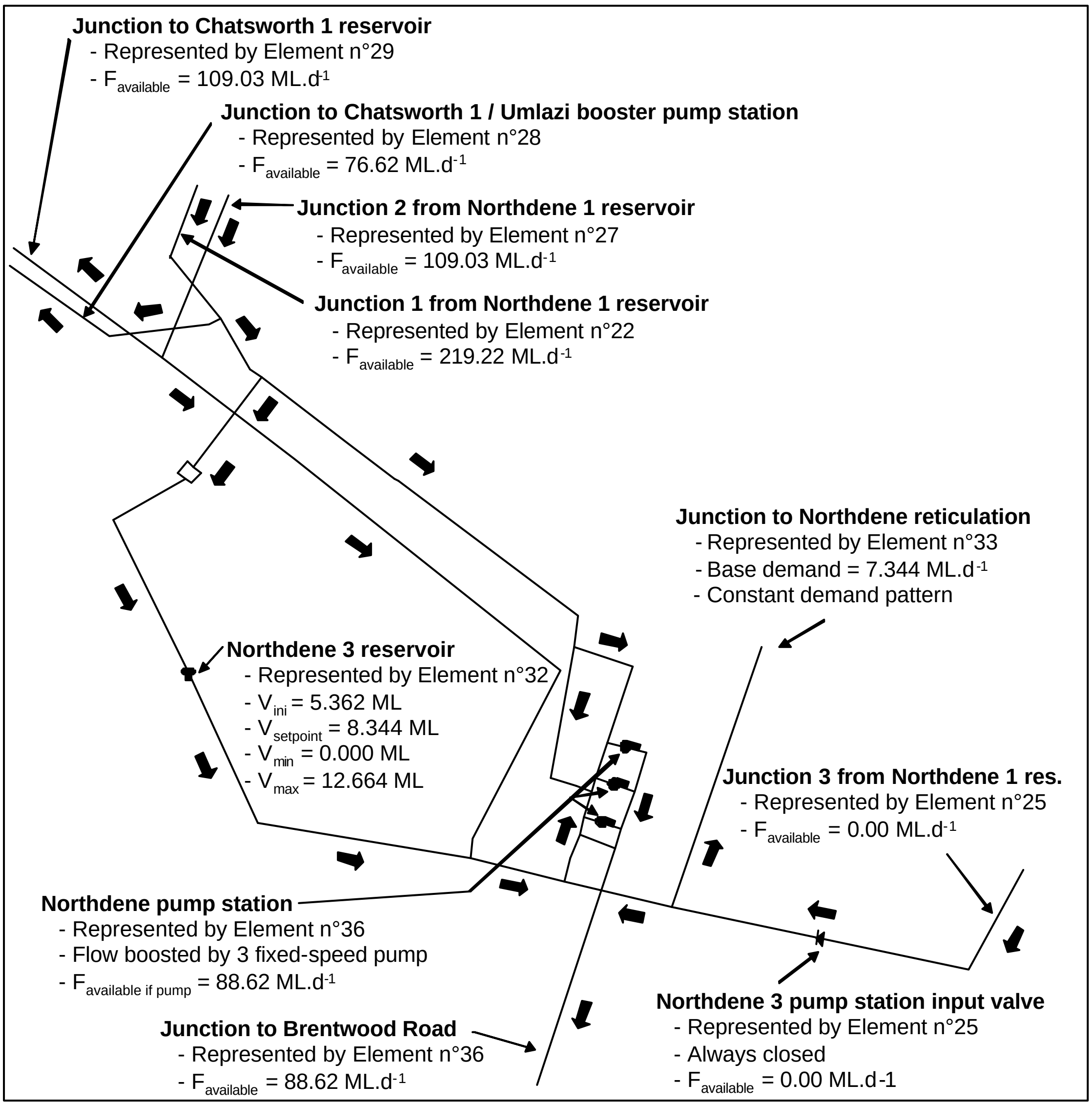
D.3.5 Dawncliff reservoir and reticulation system



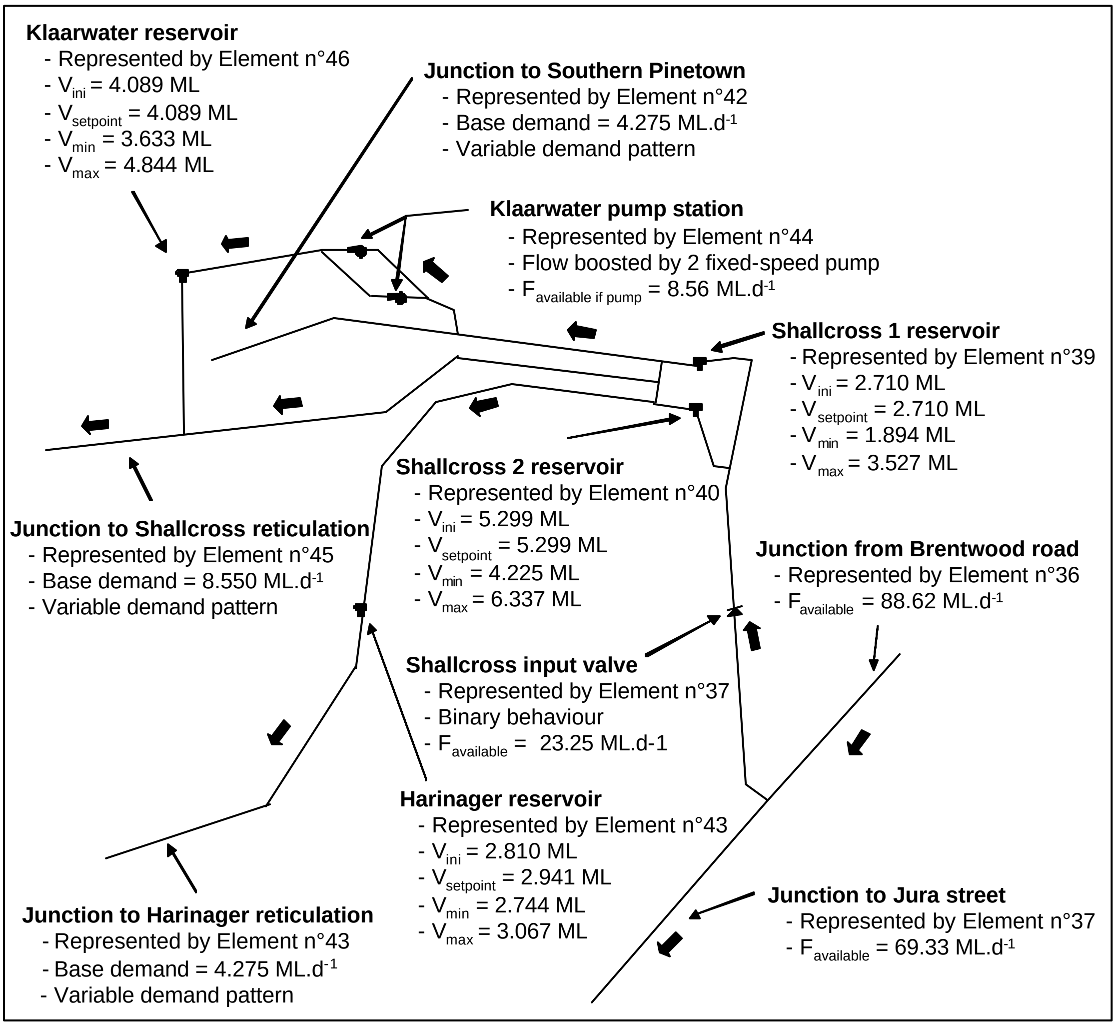
D.3.6 Northdene 1 reservoir and reticulation system



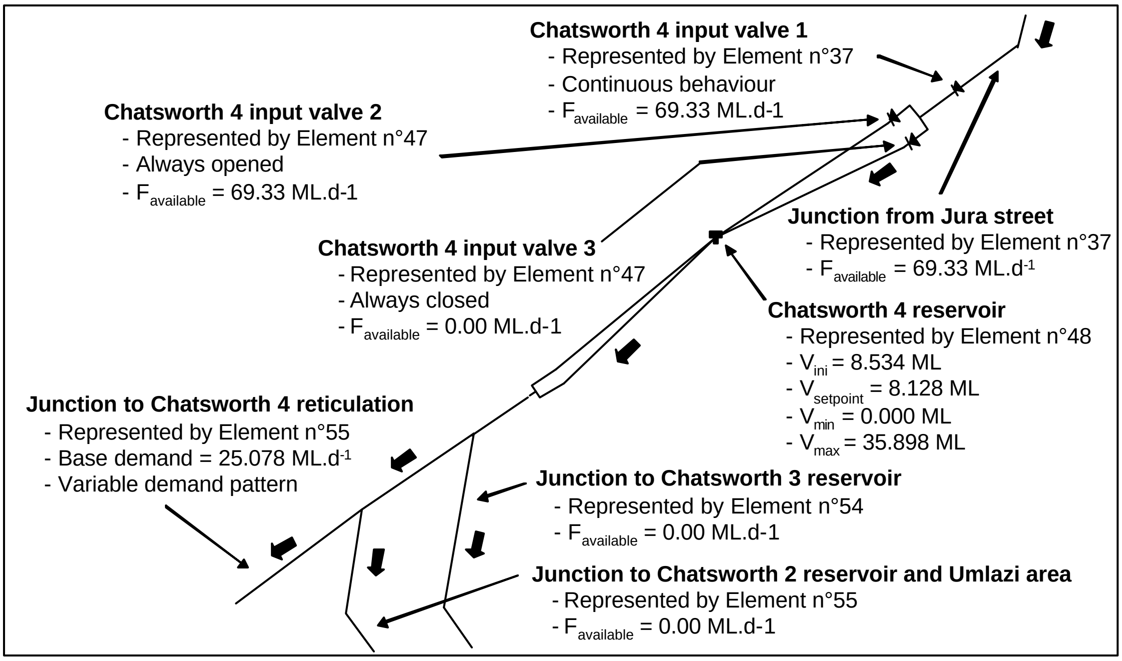
D.3.7 Northdene 3 reservoir and pump station & Firwood reticulation



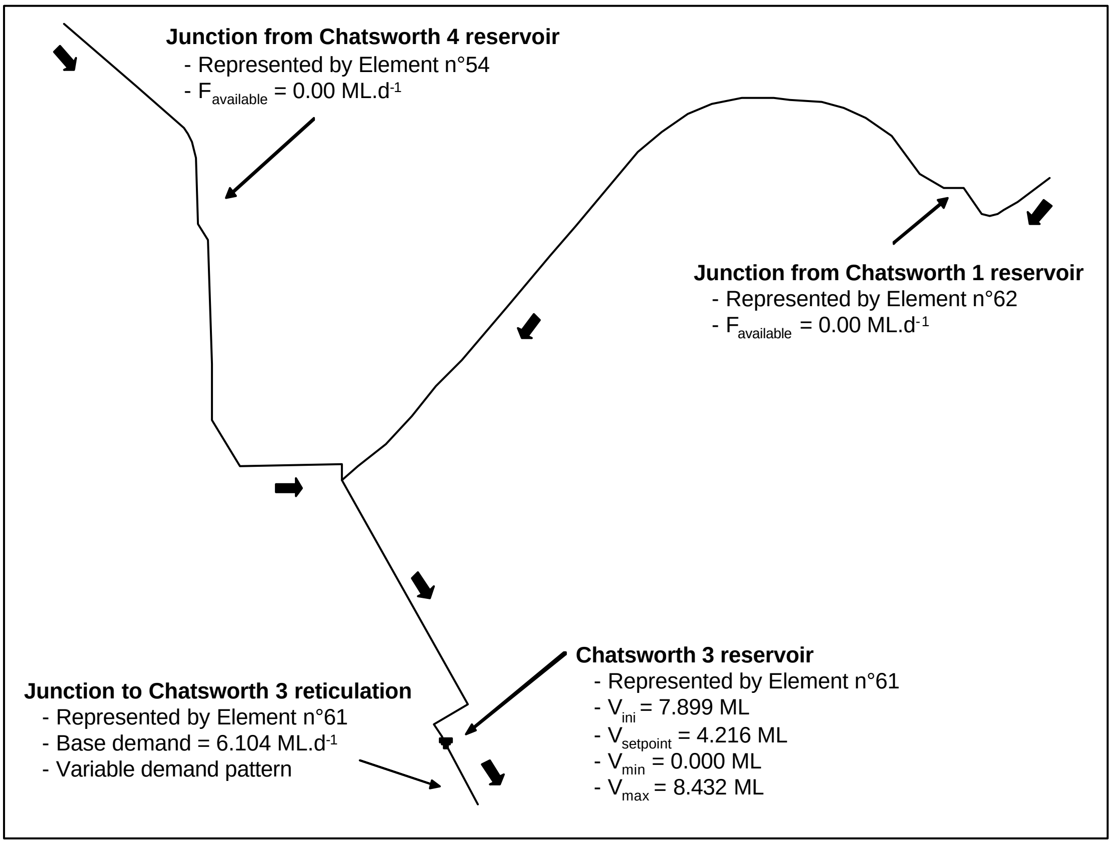
D.3.8 Shallcross area reservoirs and reticulation system



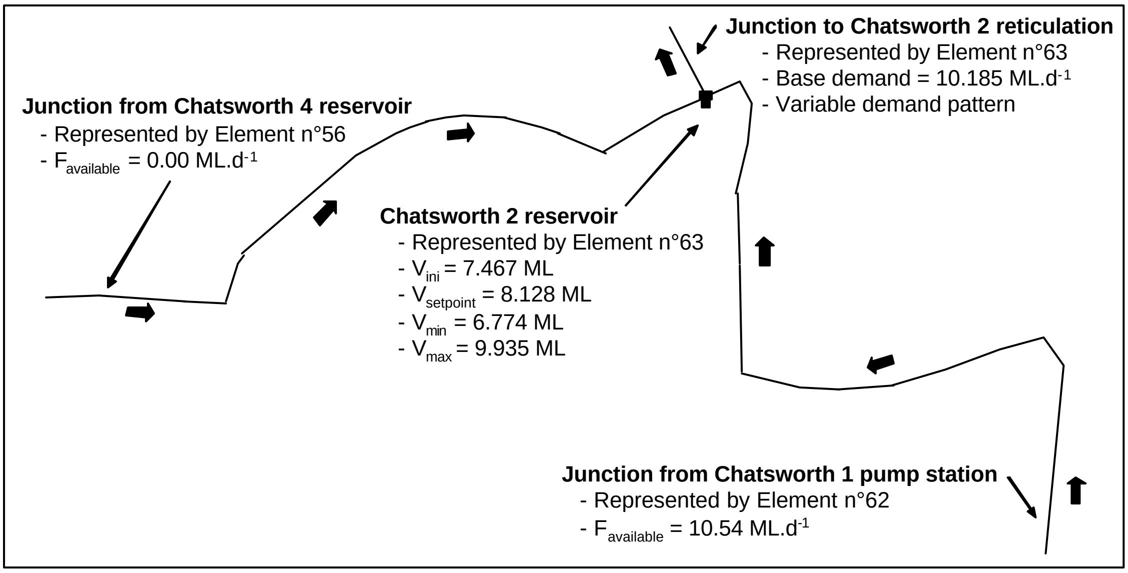
D.3.9 Chatsworth 4 reservoir and reticulation system



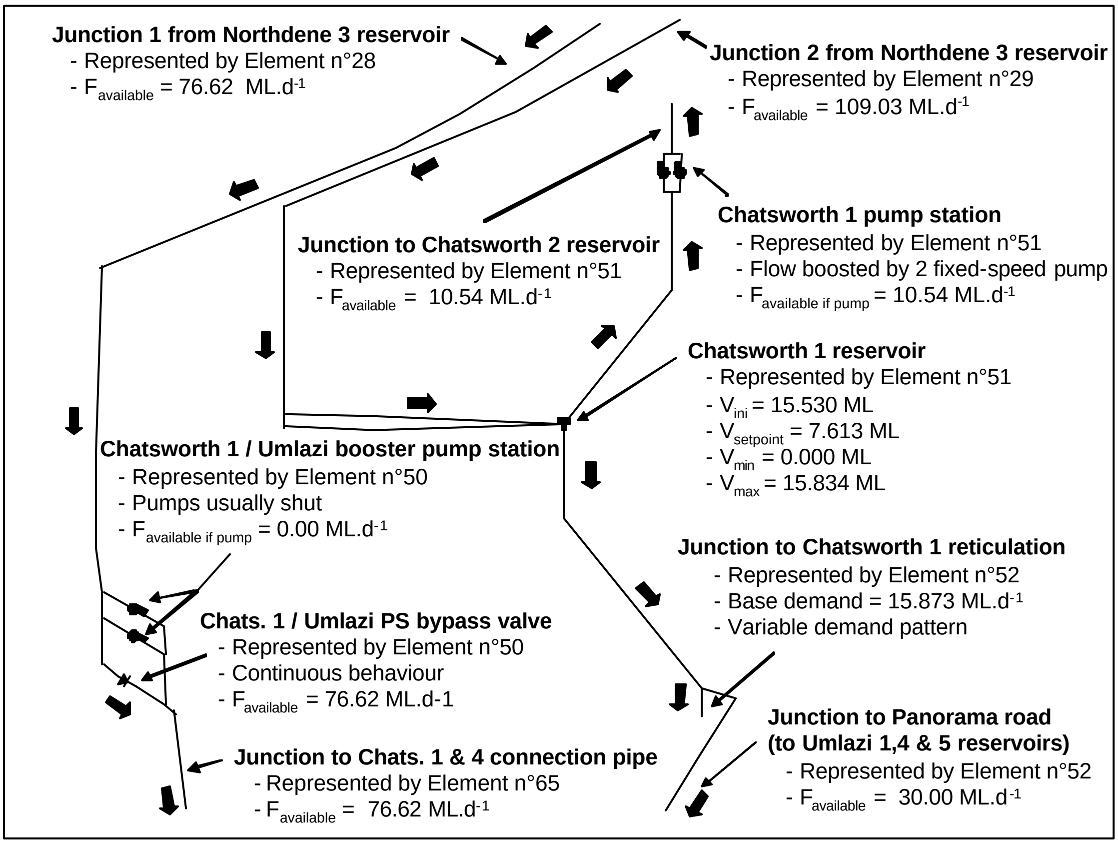
D.3.10 Chatsworth 3 reservoir and reticulation area



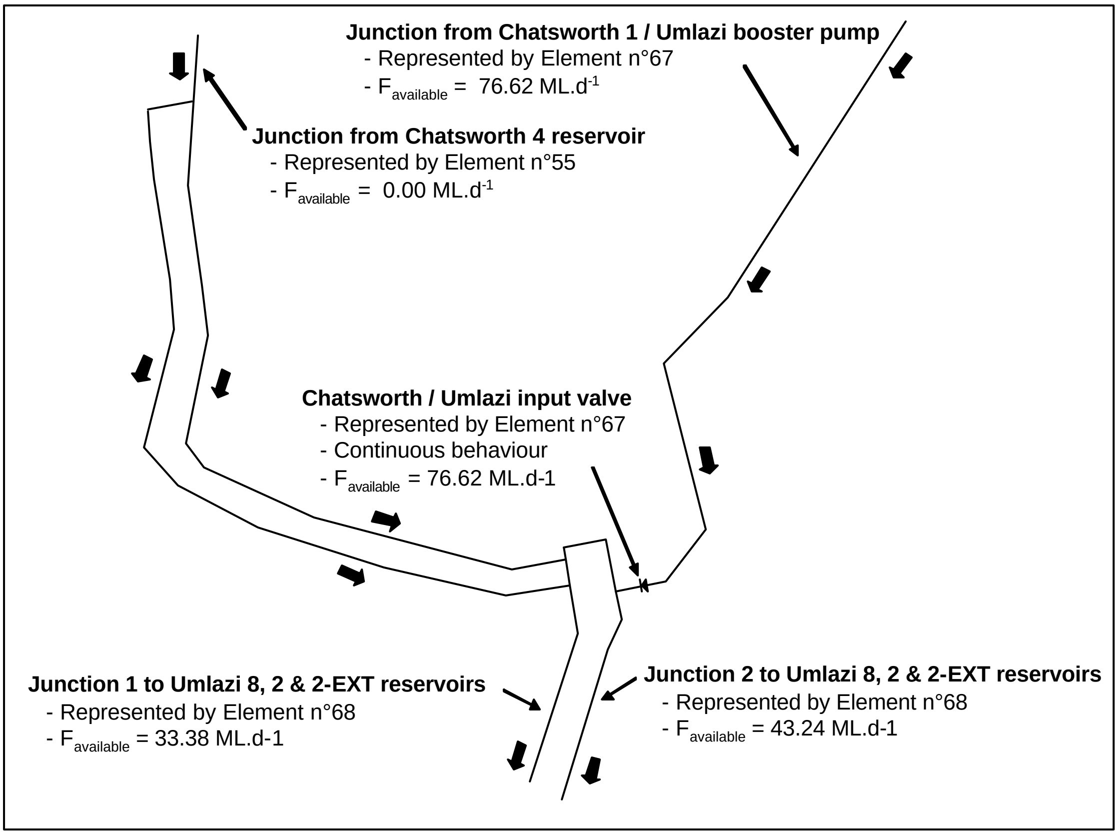
D.3.11 Chatsworth 2 reservoir and reticulation area



D.3.12 Chatsworth 1 reservoir, pump stations and reticulation area



D.3.13 Linking of Chats. 1 / Umlazi booster pump station with Chatsworth 4 reservoir output



APPENDIX E – LARGE-SCALE APPLICATION: MODELLING CHOICES MADE BEFORE PERFORMING THE OPERATIONAL OPTIMISATION OF THE SOUTHERN AQUEDUCT

THE modelling choices that needed to be done before performing the operational optimisation of the Southern Aqueduct are presented by areas. All of the figures included in this Appendix are extracted from the EXCEL spreadsheet entitled "Southern Aqueduct data.xls" and stored on the CD attached to the thesis at the following location:

/Large-scale application\Available flows strategy

E.1 Durban Heights waterworks

THE existing three output reservoirs have been combined into one unique reservoir. However, the capacity of these three reservoirs being unknown, it was decided for the data tables included in the MATLAB file to create a large reservoir in which the volume would not change much, even for high water consumption. A total capacity of 200 ML (initialised at 50 ML) has been therefore fixed on this reservoir.

E.2 Lea Drive reservoirs and reticulation system

IN the ePANET version of the Southern Aqueduct, the Lea Drive area is represented by a single connection to reticulation. However, on the SCADA system, it appears that the Lea Drive reservoir system is represented by four different compartments following each other. Unfortunately, data were not available for all of these reservoirs. It has been decided therefore to create a single reservoir with a total capacity of 5.002 ML (initialised at 4.806 ML).

Furthermore, the SCADA system shows the presence of two pumps at the entrance of the reservoir system. However, no data were available for these items. Therefore, these pumps have been considered as fixed-speed pumps and identical to the pumps operated at the Northdene pump station (SULZER SM 402 – 720, 1490 RPM), for which a characteristic is available (downloaded from the constructer website). From the operational point of view, only one pump can work at a time, which explains why only one pump is represented in the main data table of the MATLAB file representing the Southern Aqueduct. The pump needs to be switched on to see a flow entering the reservoir system.

E.3 Dawncliff reservoir and reticulation system

SIMILARLY to the Lea Drive reservoir and reticulation system, Dawncliff is represented by a single connection to reticulation in the ePANET version of the Southern Aqueduct. Here it was decided to rather follow the representation of the supervisory system, which shows two reservoirs in parallel, preceded each by one valve. These valves have been represented with a binary behaviour. The reservoir properties have been extracted from the SCADA system. Hence, the total capacity of Dawncliff upper reservoir is taken equal to 2.187 ML (with an initial capacity equal to 2.049 ML). Similarly, the total capacity of Dawncliff lower reservoir is equal to 2.064 ML, while its volume is initialised to 1.460 ML.

Furthermore, the SCADA system shows the presence of two pumps at the entrance of the reservoir system. However, it appears that these pumps are not operated any more. This explains why no pump appears at the input of the Dawncliff area in the main data table of the MATLAB document.

E.4 Northdene reservoirs and reticulation system

NORTHDENE 2 reservoir does not appear in the ePANET file, while it is present on the graphic mimic of the SCADA system. A closer look at the ePANET version of the Southern Aqueduct shows that this reservoir was combined with Northdene 1 reservoir. Therefore, in the MATLAB file used to represent the Southern Aqueduct within the optimisation tool, a unique reservoir with a combined capacity of 20.678 ML was created (the volume is initialised at 10.944 ML).

Since only one pump at a time can be operated in the facility of the Northdene pump station, the MATLAB file only models one pump in the main data table relative to the Southern Aqueduct. In this version of the file, the pump needs to be on to see a flow leaving the reservoir system.

E.5 Firwood reservoirs and reticulation system

IN the ePANET version of the Southern Aqueduct, the Firwood area is represented by a single connection to reticulation. However, on the SCADA system, it appears that the Firwood reservoir system is represented by two different compartments in parallel. Unfortunately, data were not available for all of these reservoirs. It has been therefore decided to create a unique reservoir with a total capacity of 12.664 ML (initialised at 5.362 ML).

E.6 Shallcross reservoirs and reticulation system

UNDER ePANET, the Shallcross area is represented by a single connection to reticulation. However, on the SCADA system, it appears that the Shallcross reservoir system is represented by two different compartments in parallel. This representation was the one adopted in the MATLAB file used to construct the Southern Aqueduct within the optimisation tool. The reservoir properties were directly extracted from the SCADA system. Hence, Shallcross 1 reservoir total capacity is taken equal to 4.335 ML, while its initial volume is equal to 2.710 ML. Similarly, for Shallcross 2 reservoir, the total capacity is equal to 6.537 ML and the initial volume to 5.299 ML.

Furthermore, within ePANET, no mention is made of the Klaarwater and Harinager drive reservoirs, of the pump station located at the entrance of the Klaarwater reservoir and of the multiple connections to the reticulation occurring in this area. Therefore, the SCADA representation has been adopted and all of the missing figures were extracted from the supervisory control system. For instance, Klaarwater reservoir total capacity is fixed at 4.718 ML and its initial volume at 4.089 ML. Similarly, Harinager drive reservoir total capacity was taken equal to 3.114 ML, while its initial capacity has been chosen equal to 2.810 ML. However, since data were not available for the user demand flows of this area, all of the values had to be guessed. Similarly, the pumps of the Klaarwater pump station were considered identical to the ones of the Northdene pump station, for which a characteristic is available. Since it was estimated that only one pump can work at a time, the Klaarwater pump station is represented by a unique pump in the MATLAB file. In this version of the file, the pump needs to be switched on to have a flow entering the reservoir system.

E.7 Chatsworth 2 reservoir and reticulation system

IN the ePANET version of the Southern Aqueduct, not all of the information required for the representation of the Chatsworth 2 reservoir, according to the principles of the available flows strategy, was available. Therefore, all of the missing data have been estimated. For instance, the reservoir initial volume was fixed at 7.467 ML, while its total capacity was chosen equal to 9.944 ML.

E.8 Chatsworth 1 pump station

THE properties of the pumps in this area being unavailable, they were taken identical to those located in the facility of the Northdene pump station.

Furthermore, from an operational point of view, it was estimated that only one pump could work at a time. This explains why only one pump was included in the main data table of the MATLAB file representing the Southern Aqueduct. In this version of the file, the pump needs to be switched on to have a flow entering the reservoir system.

E.9 Chatsworth 1 / Umlazi booster pump station

THE comments are similar to those of paragraph E.9. Since the operational optimisation of the Southern Aqueduct is conducted on the first scenario of operations presented in Chapter 7, no pumps were modelled at this particular location of the distribution scheme. It is recalled here that this pump station is not used in this particular scenario.