UNIVERSITY OF EXETER
Centre for Water Systems

COLLEGE OF ENGINEERING, MATHEMATICS AND PHYSICAL SCIENCES

Near Real-Time Detection and Approximate Location of Pipe Bursts and Other Events in Water Distribution Systems

Submitted by Michele Romano to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Engineering
in November 2012

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: .....

ABSTRACT

The research work presented in this thesis describes the development and testing of a new data analysis methodology for the automated near real-time detection and approximate location of pipe bursts and other events which induce similar abnormal pressure/flow variations (e.g., unauthorised consumptions, equipment failures, etc.) in Water Distribution Systems (WDSs). This methodology makes synergistic use of several self-learning Artificial Intelligence (AI) and statistical/geostatistical techniques for the analysis of the stream of data (i.e., signals) collected and communicated on-line by the hydraulic sensors deployed in a WDS. These techniques include: (i) wavelets for the de-noising of the recorded pressure/flow signals, (ii) Artificial Neural Networks (ANNs) for the short-term forecasting of future pressure/flow signal values, (iii) Evolutionary Algorithms (EAs) for the selection of optimal ANN input structure and parameters sets, (iv) Statistical Process Control (SPC) techniques for the short and long term analysis of the burst/other event-induced pressure/flow variations, (v) Bayesian Inference Systems (BISs) for inferring the probability of a burst/other event occurrence and raising the detection alarms, and (vi) geostatistical techniques for determining the approximate location of a detected burst/other event.

The results of applying the new methodology to the pressure/flow data from several District Metered Areas (DMAs) in the United Kingdom (UK) with real-life bursts/other events and simulated (i.e., engineered) burst events are also reported in this thesis. The results obtained illustrate that the developed methodology allowed detecting the aforementioned events in a fast and reliable manner and also successfully determining their approximate location within a DMA. The results obtained additionally show the potential of the methodology presented here to yield substantial improvements to the state-of-the-art in near real-time WDS incident management by enabling the water companies to save water, energy, money, achieve higher levels of operational efficiency and improve their customer service.

The new data analysis methodology developed and tested as part of the research work presented in this thesis has been patented (International Application Number: PCT/GB2010/000961).

ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge all the people listed below for their contribution to this research, without whom this work would not have been accomplished.

Firstly, I would like to give very special thanks to my two Ph.D. supervisors, Professor Zoran Kapelan and Professor Dragan A. Savić. These special thanks are conveyed not only for their continuous support and guidance throughout this research, but also for their kindness, patience and cordial friendship. With their solid and innovative thinking in doing research, they have provided me with invaluable advice.

My further thanks go to Professor Orazio Giustolisi from the Technical University of Bari and Associate Professor Shie-Yui Liong from the National University of Singapore. They have been my mentors at the early stages of my research career and from them I learnt many invaluable skills.

I would like to acknowledge all of the incredibly dedicated and talented members of the Centre for Water Systems, with whom I had the pleasure of working, for creating an inspiring research environment.

I would like to thank Mr Ridwan Patel from Yorkshire Water Services for providing the data used for this research work and also Mr Kevin Woodward and Mr Derek Clucas from United Utilities for their kind support.

In addition, I wish to express my genuine gratitude to my wonderful parents and my two beloved sisters for the never-ending love, constant support and encouragement they have given me right from my childhood. They have always been a great motivation for me in every step of my life.

My gratitude extends to my fiancée Stacy for being a special person in my life. Her energy, support, encouragement, friendship and love made the journey towards this goal much easier.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under the Curve
BBA	Boundary Based Analysis
BN	Bayesian Network
BIS	Bayesian Inference System
BP	Back Propagation
CC	Customer Contact
CPD	Conditional Probability Distribution
CPT	Conditional Probability Table
CSV	Comma Separated Values
CWT	Continuous Wavelet Transform
DAG	Directed Acyclic Graph
DBA	Discrepancy Based Analysis
DMA	District Metered Area
DofW	Day of the Week
DWT	Discrete Wavelet Transform
EA	Evolutionary Algorithm
EANN	Evolutionary Artificial Neural Network
EE	Engineered Event
EM	Expectation Maximisation
ERS	Event Recognition System
ES	Evolutionary Strategy
FF	Feed Forward

FRF	Frequency Response Function
FT	Fourier Transform
GA	Genetic Algorithm
GPR	Ground Penetrating Radar
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
HM	Hydraulic Model
ICWT	Inverse Continuous Wavelet Transform
IDW	Inverse Distance Weighted
IDWT	Inverse Discrete Wavelet Transform
ITA	Inverse Transient Analysis
JPD	Joint Probability Distribution
LP	Local Polynomial
MAP	Maximum A Posteriori
ML	Maximum Likelihood
MLP	Multi-Layer Perceptron
MNF	Minimum Night Flow
MR	Main Repair
NAN	Not A Number
NOP	Normal Operating Pattern
OC	Ordinary Cokriging
ODBC	Open Database Connectivity
OFWAT	Office of Water Services (UK)
OK	Ordinary Kriging
PIG	Pipeline Inspection Gauge

RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristics
SCADA	Supervisory Control And Data Acquisition
SIM	Service Incentive Mechanism
SMS	Short Message Service
SPC	Statistical Process Control
SVM	Support Vector Machine
TBA	Trend Based Analysis
TofD	Time of the Day
UK	United Kingdom
WDR	Weight Decay Regularisation
WDS	Water Distribution System
WMS	Work Management System
WA	Wavelet Analysis
WT	Wavelet Transform
YWS	Yorkshire Water Services

CHAPTER 1 INTRODUCTION

1.1 Motivation

Fresh water can be considered as one of the most valuable of natural resources as it is essential not only to the human survival (safe drinking water and food production) but also important for the economic growth of our society. However, fresh water resources are scarce, unevenly distributed, and under increasing pressure. Thus, the security of supply is at risk.

It is estimated that less than 1% of the Earth's water is available for direct human consumption (USGS, 2011). This small fraction of fresh water is unevenly distributed in space and time (Gleick, 1995). Additionally, demographic, socioeconomic and environmental factors such as accelerated population growth, rapid urbanisation, unsustainable consumption patterns, pollution of underground and surface waters, and more extreme environmental fluctuations due to the climatic consequences of global warming put increasing pressure on the available fresh water resources.

Based on the above and bearing in mind that water crises are already evident in the competition for water for domestic, agriculture, and industrial use in many parts of the world (Barrett, 2000), it is clear that, if future generations are to enjoy a similar standard of living to the one we enjoy now, securing the supply of fresh water is a compelling and critical issue. Two potential responses can be identified: (i) meeting demand with new resources (i.e., supply-side), or (ii) managing the demand itself with the aim of postponing or avoiding the need to develop new resources (i.e., demand-side).

Desalination is currently expensive and will not secure the supply of fresh water in countries which cannot afford this technology. Also, the costs of constructing new water storages and/or water treatment plants have risen over the last twenty years. As a result, the emphasis is shifted towards managing water demand by better utilising the water that is already available. In this scenario, more intensive and effective water conservation efforts are needed. Although conserving water is everyone's ethical responsibility, the water industry plays a central role. Therefore, it has to stand out and lead the way.

Currently, the amount of water lost in Water Distribution Systems (WDSs) varies widely worldwide, with most countries falling into the range of 20% - 40%. In the light of the already outlined security of supply risk, these high levels of water loss cannot and should not be tolerated. Therefore, reducing these water loss levels is a high priority issue for the water companies. One way to achieve this is reducing the water lost due to the pipe burst events.

In the United Kingdom (UK), despite significant effort is put into the on-going rehabilitation and maintenance of the water supply infrastructure, the number of incidents caused by the pipe burst events is still significant. This is mainly because the UK has an ageing water supply infrastructure where some of the pipes were laid more than one hundred years ago and hence, today, many are in relatively poor condition. Additionally, it is important to stress that because of the stochastic nature of the pipe burst events, it is impossible to predict and completely eliminate them. Taking all this into consideration, the timely and reliable detection and location of these events (i.e., aim of the research work carried out in this thesis) may enable the water companies to repair them more quickly, thereby minimising their overall lifecycle. Considerable water savings may therefore be achieved.

Although conserving water is very important for the aforementioned reasons, it is only one of the drivers that motivate the water companies to seek to timely and reliably detect and locate the pipe burst events in their WDSs. Indeed, because of the outlined frequency and nature of these events, in their day-to-day operations, the water companies are not only tasked with operating their WDSs optimally to minimise the costs and to meet the required standards of service (e.g., in terms of water quality and providing adequate pressure at the customers' taps), but also with managing contingency situations when the pipe burst events occur. This additional burden is especially heavy because the water companies are mainly judged by the public and the regulatory agencies alike based on how well (or otherwise) they perform this task. Nowadays, the resulting potential unplanned interruptions to the water supply and the damaging consequences of the pipe burst events are tolerated to a lesser extent. It is in this scenario that the main driver for the research work carried out in this thesis arises. This is improving the water companies' operational performance and customer service. Indeed, the timely and reliable detection and location of pipe burst events in WDSs

provides opportunities for reducing the number and duration of the supply interruptions, enabling the water companies to allocate more time for planning and implementing mitigation measures (when the supply interruptions cannot be avoided), and allowing proactive and/or more informed communication with the customers. The fact that the required actions could be initiated before a pipe burst grows larger and causes serious damages or long interruptions in service is also important.

The aforementioned improved operational efficiency and customer service, in turn, may result in significant financial savings for the water companies. These include direct, indirect and social savings. For example, direct savings are achieved by reducing the production (i.e., power and chemicals) and distribution (i.e., pumping/energy) costs associated with the water that fails to reach the consumers, the costs associated with the compensation payments for the damages to the surrounding infrastructures and properties, the operational costs of dealing with the customer complaints, and the burst events management costs associated with focusing the detection and location efforts of the water company's personnel on the wrong parts of a WDS. Indirect savings are achieved by avoiding the UK Office of Water Services (i.e., OFWAT) penalties related to the Service Incentive Mechanism (SIM) and Serviceability measures. Social savings are achieved by reducing the costs associated with the interruption of supply and disruption of traffic and business, and the costs associated with the reduced fire-fighting capabilities and pollutant ingress through cracks.

Finally, it has to be stressed that the aforementioned cost savings, the improved operational performance and customer service, and the demonstration of commitment in conserving water and reducing their carbon footprint will all serve to favourably enhance the water companies' profile and increase the customers' trust in the water companies. The latter drivers are also very important for the research work carried out in this thesis. This is because they are fundamental for enabling the water companies successfully carrying out their statutory duty of promoting water efficiency. If the conservation messages the water companies are championing are heard and taken seriously by the public, the resulting benefits might have a positive and lasting impact that cannot be ignored since their value might be higher than the monetary benefits mentioned above.

1.2 Background

Despite the outlined importance of the timely and reliable detection and location of the pipe burst events in WDSs, this is one of the tasks faced by a water company that is becoming more and more difficult. The main reasons for this are the increasing complexity of the WDSs, the fact that water companies are starting routinely implementing pressure management programmes, and the installation of more and more poly-ethylene pipes (on which the conventional acoustic equipment-based techniques do not work that well). Therefore, new and more efficient techniques for the detection and location of pipe burst events in WDSs are required.

Currently, a wide range of pipe burst event detection and location techniques exists that are based on various principles (Puust et al., 2010). However, none is ideal and the number of techniques currently practised by the water companies is limited. In many cases, pipe bursts are brought to the attention of a water company only when someone calls in to report a visible event. Water companies embracing modern leakage management technologies devote considerable manpower and resources to proactive detecting and locating the pipe bursts by utilizing techniques that make use of highly specialised hardware equipment (e.g., leak noise correlators, acoustic sensors mounted on inline Pipeline Inspection Gauges - PIGs, Ground Penetrating Radars - GPRs, etc.). Despite some of these techniques being the most accurate ones used today (Puust et al., 2010), they are also costly, labour-intensive, and slow to run. Consequently, much research has been focussed on finding inexpensive techniques that can help the water companies significantly reducing the pipe bursts' lifecycle by making them aware of the occurrence of the pipe bursts much faster, and guiding the water company personnel straight to the problem areas.

Many techniques exist that promise low cost and faster solutions by endeavouring to solve the pipe burst detection and location problem by numerical analyses only (Puust et al., 2010). Among these, those that make use of statistical and Artificial Intelligence (AI) techniques for automatically processing the operational variables (e.g., pressure and flow) in an on-line fashion are of particular interest for the research work carried out in this thesis. These techniques have recently started to appear (e.g., Mounce et al., 2010a, 2011a; Fenner and Ye, 2011; Palau et al., 2012) mainly due to the latest developments in hydraulic sensor technology, and on-line data acquisition systems.

These developments are enabling the water companies to deploy a larger number of more accurate and increasingly cheaper pressure and flow devices and the data collected by these devices to be received in near real-time. As these data streams provide a potentially useful source of information for detecting and locating the pipe burst events both quickly and economically, the development of data analysis methodologies that allow their efficient and effective exploitation is a key factor for providing a rapid response to these events.

In the light of the above, statistical/AI-based techniques are very promising because they can automate mundane tasks involved in the data analysis process. Furthermore, they can efficiently deal with the vast amount of often imperfect sensor data collected by modern Supervisory Control And Data Acquisition (SCADA) systems, and extract information useful for making reliable operational decisions. These techniques also presents several advantages over other numerical techniques such as the steady state analysis-based (e.g., Pudar and Liggett, 1992; Wu et al., 2010), transient analysis-based (e.g., Liggett and Chen, 1994; Kapelan et al., 2003), and negative pressure wave-based (e.g., Misiunas et al., 2005c; Srirangarajan et al., 2012) techniques. For example, they have a requirement for pressure and/or flow measurements sampled much less frequently (e.g., every 15 minutes) than those required for transient analysis. Moreover, they rely on the empirical observation of the behaviour of a WDS. Therefore, precise knowledge of the WDS and instrumentation parameters is not required.

Despite their initial success, the aforementioned statistical/AI-based techniques can be improved in terms of both pipe burst event detection reliability and detection time (e.g., the time elapsed between a burst event occurrence and the generation of the corresponding alarm). Additionally, at the present, the available statistical/AI-based techniques allow the discovery of a pipe burst event in a particular section of a WDS (e.g., at the District Metered Area - DMA - level) without giving any information about its more precise location. Thus, they can be also improved in terms of pipe burst event location accuracy (i.e., to indicate more precisely the likely location of a pipe burst – to restrict the pipe burst search area).

The research work carried out in this thesis focuses on addressing the above issues. It establishes a novel data analysis methodology, based on sound theoretical principles,

with noticeable improvements over the aforementioned statistical/AI-based techniques. The main improvements involve: (i) using advanced techniques for more efficiently and effectively processing the hydraulic data gathered (e.g., wavelets – see Appendix A – for removing noise from the measured flow and especially pressure signals), (ii) using different ensembles of statistical and AI techniques each aimed at solving a different aspect of the pipe burst event detection problem (e.g., Statistical Process Control - SPC – see Appendix D, Artificial Neural Networks - ANNs – see Appendix B, and Evolutionary Algorithms - EAs – see Appendix C for the analysis of the short/long term pressure/flow variations and the recognition of the various types of evidence of an event occurrence), (iii) using Bayesian Inference Systems (BISs) – see Appendix E – for reasoning under uncertainty, and for performing classification/inference by simultaneously (synergistically) analysing multiple evidence/causes and multiple pressure and/or flow signals at the DMA level, and (iv) using geostatistical techniques – see Appendix F – for performing further diagnosis of the event occurring (i.e., determining the approximate location of a pipe burst event within the DMA). The more reliable and timely detection of the pipe burst events together with the successful identification of their approximate location, which could be achieved by using the proposed methodology, can facilitate prompt interventions and repairs. This, in turn, may yield substantial improvements to the state-of-the-art in near real-time WDS incident management and enable to proactively manage the water leakages in WDSs.

1.3 Research Scope and Objectives

The scope of the research work carried out in this thesis is to develop and test a data analysis methodology for the near real-time detection and approximate location of pipe burst events in WDSs that facilitates prompt interventions and repairs in response to these events, thereby enabling the water companies to save water, energy, money, achieve higher levels of operational efficiency and, most importantly, improve their customer service by minimising the supply interruption inconvenience. Such methodology had to be fast, reliable, non-invasive and cheap to employ.

The aforementioned research scope is achieved in this thesis through the following objectives:

1. To explore the feasibility of using the near real-time data collected by the hydraulic (pressure and flow) sensors as the only source of information about the normal and abnormal (e.g., when a pipe burst event occurs) operating conditions of a WDS;

2. To investigate the potential of using data-driven AI and statistical techniques for efficiently extracting and capturing information about the expected patterns of the signals coming from the hydraulic sensors during the normal operating conditions of a WDS (i.e., Normal Operating Patterns - NOPs);

3. To develop a suitable data analysis framework for identifying and estimating, in the signal patterns observed during the abnormal operating conditions of a WDS, the pipe burst event-induced deviations from the expected signal NOPs;

4. To investigate the potential of using Bayesian logic for combining multiple evidence of a pipe burst event occurrence (i.e., deviations mentioned in the previous bullet point) to enable more reliable and timely detection of these events and generation of corresponding alarms;

5. To explore the feasibility of using geostatistical techniques for allowing the approximate location of pipe burst events within a DMA;

6. To test, verify and demonstrate the effectiveness and efficiency of the novel methodology for detection and approximate location of pipe burst events on a number of real-life case studies from the UK.

1.4 Thesis Structure

This thesis is divided into six chapters including this introduction.

In Chapter 2, a review of the relevant literature is provided. The review gives a wide overview of the available techniques for the identification of the pipe burst events in WDSs, as well as techniques from research. It also highlights their individual advantages and disadvantages. The reviewed techniques range from simple sounding surveys aimed at identifying the pipe bursts already occurred in a WDS, to more sophisticated monitoring systems that allow the identification of the pipe burst events in near real-time.

In Chapter 3, the novel methodology for the near real-time detection of pipe burst events that makes use of data from the small number of sensors typically deployed nowadays in the UK DMAs is presented first, together with the associated implementation in a computer-based Event Recognition System (ERS). Next, the methodology customisation for enabling the efficient handling/processing of the data from a larger (than currently used in the UK practice) number of sensors is presented. Once this is done, the methodology further development for enabling the approximate location of a pipe burst event within a DMA by means of geostatistical techniques is described. This chapter also presents the theoretical background and the details of the (geo)statistical and AI based techniques used in the ERS.

In Chapter 4, the event detection capabilities of the novel methodology that makes use of data from the small number of sensors typically deployed in the UK DMAs are tested, verified and demonstrated on two case studies. This is done by using the historical pressure/flow data from several UK DMAs, and both real-life and engineered (i.e., simulated by opening fire hydrants) pipe burst events. Additionally, the results of a number of investigations, which evaluated the methodology's event detection capabilities for different choices of parameters and data analysis methods, are reported.

In Chapter 5, the results of several sensitivity type data analyses, that evaluated the event detection capabilities of the methodology that makes use of data from a larger number of sensors on a series of engineered pipe burst events, are reported first. Next, the details of the tests that focussed on evaluating the performances of different geostatistical techniques for the approximate location of the pipe burst events within a DMA are discussed.

In Chapter 6, the key findings of this thesis are summarised and the relevant conclusions are drawn. The novel aspects introduced in this thesis are highlighted, followed by the possible directions of future research work to enhance and extend the novel methodology presented in this thesis.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

Water supply pipes are the most important way of transporting water from one place to another. They provide a safe means of conduit and are one of the principal assets of a WDS. In a WDS, the water supply pipes can be broadly divided into two major categories, namely transmission and distribution pipes. Transmission pipes convey large amounts of water over great distances, typically between major facilities within the WDS. These pipes are connected to each other to form long *pipelines* whose topology is primarily branched. Distribution pipes further distribute the water to the residential areas and typically follow the alignment and topology of the residential areas' streets. They are connected to each other to form more or less complex *pipeline systems* whose topology is primarily looped. This provides an additional level of reliability of supply to the customers. Besides having different topologies, transmission pipelines and distribution pipeline systems have different hydraulic profiles, pipe sizes and materials. A variety of materials have been used in the production of water supply pipes, mainly depending on the pipe diameter and the year of installation. For large transmission pipelines (i.e., diameters over 300 millimetres), steel, mild steel cement-lined, or pre-stressed concrete cylindrical pipes are typically used. Older distribution pipeline systems are typically made of cast iron or asbestos cement. Mainly ductile iron and poly-vinyl chloride or poly-ethylene are used for newer distribution pipeline systems.

Whatever their function and material, the water supply pipes are designed and constructed to maintain their integrity over many years. However, it is difficult to avoid the occurrence of failures during their lifetime. This is because, water supply pipes are ubiquitous and may have to operate under adverse conditions (e.g., weather, corrosive environments, soil movement, poor construction standards, poor quality of fittings and workmanship, excessive traffic loads and vibration, fluctuation of water pressure, etc.). Furthermore, they deteriorate naturally with time and subsequently lose their initial water tightness. In many parts of the world, as well as in the UK, the majority of the water supply pipes were installed in the first part of the 20^{th} century and hence, today, are in poor condition. As water supply pipes continue to age and deteriorate it is likely that failures will occur more and more often.

A number of techniques have been developed and presented for the identification of failures in WDSs (see – e.g., AWWA, 2009; Colombo et al., 2009; Farley and Trow, 2003; Lambert et al., 1998; Puust et al., 2010; Thornton et al., 2008; Wang et al., 2001; Wu et al., 2011). This chapter present a review of these techniques. The reviewed techniques are classified here as either hardware-based or hydraulic (i.e., dependent on process parameters such as pressure and flow). On the one hand, hardware-based techniques make use of specialised devices such as leak noise correlators and gas sensors, among the others. On the other hand, hydraulic techniques numerically exploit the hydraulic behaviour of a WDS in the presence of a failure.

Before outlining the structure of this chapter, however, it is important to clarify some of the terms that will be used throughout it. Specifically, the following has to be noted:

- Different types of failures, depending on their size and character, exist. For most of the techniques described in the literature, however, the type of failure is not specified and the generic term 'leak' is predominantly used. Despite the fact that the term 'pipe burst event' would be more consistent with the terminology used in the rest of this thesis, in this chapter the term 'leak' will be used in order to conform with the mainstream jargon.
- The identification of leaks in WDSs involves leak assessment, detection and location. In this review, leak assessment will refer to the quantification of the water lost in a whole WDS or in a WDS section (e.g., particular section of a distribution pipeline system, particular section of a transmission pipeline, etc.). Leak detection will refer to the discovery of a leak in a specific WDS section without giving any information about its location. Leak location will refer to the determination of the position of a leak. With regard to the latter, however, the terms 'approximate leak location' and 'exact leak location' will be additionally used. Specifically, approximate leak location will refer to the determination of the likely (i.e., within a certain range) position of a leak. Exact leak location will refer to the determination of the precise position of a leak.
- Leak identification can be carried out by either inspecting or monitoring the whole or a section of a WDS. In this review, inspection will refer to a planned action that is performed at discrete time instances aiming at identifying the leaks that are already

present in a WDS/WDS section. Monitoring will refer to the continuous checking for the new leaks as they occur.

Note that although no specific division is made here to separate the reviewed techniques into assessment, detection and location groups, or into inspection and monitoring groups, what each technique allows performing and how is highlighted.

The chapter is organised as follows. After this introduction, an overview of the leakage management activities is given in Section 2.2. This section provides a picture of the broader context in which the techniques for the proactive identification of leaks find their application and highlights their fundamental role for the effective and efficient leakage management. Once this is done, the hardware-based techniques are reviewed in Section 2.3 and the hydraulic techniques are reviewed in Section 2.4. The review carried out in these latter sections covers the leak identification techniques that have been applied to both water transmission pipelines and distribution pipeline systems. Available techniques are revived, as well as techniques that are in the stage of research and development. These two sections provide a wide overview of the leak identification techniques. The aims are: (i) to highlight their capabilities and limitations, (ii) to identify the state-of-the-art in their development, and (iii) to evaluate their potential to help the water companies minimising the leaks' runtime. Finally, a summary of the chapter containing the main conclusions and considerations about the gaps in the current research is given in Section 2.5.

2.2 Leakage Management

The effective and efficient leakage management has been an issue associated with the WDSs' operation since some of the earliest WDSs were built. At the present, this issue still remains compelling and challenging. Two approaches to leakage management can be identified. These are: (i) the reactive approach, and (ii) the proactive approach.

The reactive approach has been the traditional approach to leakage management, whereby water companies wait for a leak to arise and repair it only when it becomes self-evident or is discovered through 'ad hoc' investigations following customer complaints. Note that, this type of leaks is often referred to as "reported leaks". This is because these leaks are typically reported to the water companies by the customers who have supply problems or by the public observing water escaping from the ground.

Although the reactive approach is simple and does not involve any systematic action, it presents inevitable problems for the customers (Ramos et al., 2001). Furthermore, it may easily lead to water losses up to 40% even in WDSs with a low number of leaks (Puust et al., 2010). This is because the greatest volume of water lost is often generated by the "unreported leaks". Indeed, as the water companies remain unaware of such leaks, their runtime is usually longer.

The proactive approach, on the other hand, is the approach that every modern water company should implement to achieve a significant reduction in the amount of water lost over the years. It not only involves proactive leak identification activities (i.e., active leakage control) that enable the water companies reducing the runtime of the unreported leaks but also other complementary activities such as pressure management and routine renewal/rehabilitation of assets (Thornton, 2002).

Pressure management activities allow reducing the number of new leaks which occur each year and the amount of water lost through the leaks that are already present in a WDS. Indeed, the pipe pressure affects the leakage levels in a number of ways (TWGWW, 1980). Firstly, it has an impact on the leak occurrence frequency (e.g., pressure transients can fracture the pipes and damage their joints, frequent pressure fluctuations may cause fatigue failures in the pipes, etc.). Secondly, it has an impact on the flow rate through a leak as theoretically the flow rate of a fluid through an opening is proportional to the square root of the pressure differential across the opening (provided the dimensions of the opening remain fixed). However, the leak's opening may enlarge because of the pressure and thus by reducing the pressure, much greater reduction in the amount of water lost through the leaks can be realised than that predicted by the square root relationship. This is especially true for small leaks from joints and fittings in most pipe types and large leaks in plastic pipes (Lambert, 2001). The above observations show that the leakage levels are tightly coupled with the WDS's pressure and that when the overall pressure is reduced the same happens to the leakage levels. However, it has to be stressed that despite the pressure management activities can save water they do not eliminate the leaks. Many techniques have been developed for dynamically performing the pressure reduction over a period of time. Examples of these techniques can be found in Jowitt and Xu (1990), Reis et al. (1997), Vairavamoorthy and Lumbers (1998), Tucciarelli et al. (1999), Reis and Chaudhry

(1999), and Ulanicka et al. (2001). All these techniques are based on the use of elements which provoke head-losses in a WDS section (e.g., pressure reducing valves).

Asset renewal/rehabilitation activities also allow reducing the number of new leaks that occur each year and the amount of water lost through the leaks that are already present in a WDS. Internationally, many water companies have tried to reduce their leakage levels by relying on a high pipe replacement rate. However, it has been observed that, if the proactive leak identification activities were not carried out, the leakage levels continued to increase (SOPAC, 2011). Furthermore, the cost of reducing the leakage levels through asset renewal/rehabilitation activities is much greater than the cost of reducing them through pressure management or proactive leak identification activities. Taking this into account, note that pipes do not last forever and targeting the pipes to be renewed/rehabilitated can ensure that the water companies get the maximum benefit from their investments. Various researchers have looked into approaches for optimising the pipe renewal/rehabilitation activities (e.g., Kleiner and Rajani, 2001; Kleiner et al., 2004, 2005; Dandy and Engelhardt, 2001; Boxall at al., 2005; Giustolisi et al., 2005; Giustolisi and Berardi, 2007; Berardi et al., 2007, 2009). Some of these approaches are based on the analysis of historical leak occurrence statistics and/or factors that influence the probability of a leak occurrence (e.g., pipe's age). Others take into account also the economic, technical and social criteria related to the renewal/rehabilitation strategy.

In view of the above, it is clear that the pressure management and the asset renewal/rehabilitation activities are very important for the effective and efficient leakage management and should always be part of a water company's water loss reduction strategy. However, it is also clear that leaks are dynamic and whilst initially, significant reductions can be made, the leakage levels over a period of time will tend to rise unless on-going proactive leak identification activities are carried out. Therefore, the proactive leak identification activities are fundamental for first reducing the leakage levels in a WDS and then maintaining these low levels.

2.3 Hardware-Based Techniques

2.3.1 Section overview

Hardware-based techniques mainly allow leak detection and location. They make use of specialised devices that range from simple listening rods, to more sophisticated

inspection gauges sending magnetic fields, electromagnetic waves, or ultrasounds through a pipeline's wall. Although some of these techniques have been around for years, because of the constant development, they are getting increasingly high tech and sophisticated (Puust et al., 2010). This section provides a brief review of literature dealing with several of the hardware-based techniques that have been used/proposed for the detection and location of leaks in WDSs. Furthermore, it provides a concise description of the principles on which each technique is based and the way it works. The aims are: (i) to identify their main capabilities and limitations, and (ii) to highlight some of the aforementioned latest developments.

The hardware-based techniques that are considered in this section include:

- Acoustic equipment-based techniques;
- Tracer gas-based techniques;
- Infrared thermographic techniques;
- GPR-based techniques;
- Inline PIG-based techniques.

Relevant literature dealing with each of the above techniques is reviewed in Sections 2.3.2 to 2.3.6. Once this is done, Section 2.3.7 provides a summary of the considered techniques' main characteristics, advantages and disadvantages and presents concluding remarks about their potential to help the water companies minimising the leaks' runtime.

2.3.2 Acoustic equipment-based techniques

Acoustic equipment-based techniques enable leak detection and location in WDSs by making use of specialised devices that detect the noise (i.e., acoustic signal) generated by a leaking pipe. Indeed, a pressurised pipe that is leaking emits an acoustic signal with different frequency ranges that depend on the fluid in the pipe, the pressure in the pipe, the pipe's material and size, the nature and conditions of the ground around the wall of the pipe, and the characteristics of the leak (type and size). The acoustic signal then travels through the fluid in the pipe, the pipe's wall and the ground surrounding the pipe in the immediate area of the leak.

Acoustic equipment include: (i) listening devices (i.e., listening rods and ground microphones), (ii) noise loggers, and (iii) leak noise correlators. Listening devices may be either mechanical or electronic. They make use of sensitive mechanisms or materials such as piezoelectric elements to sense a leak-induced acoustic signal. Modern equipment incorporates microphones, signal amplifiers and noise filters to enhance the leak-induced acoustic signal's resolution (Pilcher et al., 2007). Noise loggers are compact devices composed of an accelerometer (i.e., non-intrusive microphone) or a hydrophone (inserted directly into the water stream), and a programmable data logger. Leak noise correlators are also compact devices composed of accelerometers or hydrophones. These devices, however, are additionally fitted with a microprocessor that enables, based on the Cross-Correlation method (see – e.g., Duda and Hart, 1973), determining the location of a leak by using measurements of the leak-induced acoustic signal at two locations on either side of an inspected pipeline segment.

Generally speaking, initially, the leak inspectors detect and approximately locate (i.e., roughly bracket) the leaks by performing a 'sounding survey' using listening rods. This is achieved by walking along a WDS section and placing the listening rods onto (i.e., direct sounding) convenient pipe fittings such as fire hydrants and/or valves (i.e., general survey) or onto all the accessible pipe fittings (i.e., detailed surveys). Alternatively, a sounding survey can be carried out by using the noise loggers. According to Pilcher et al. (2007), this involves deploying the noise loggers in groups of 6 or more at adjacent pipe fittings. Typically, the noise loggers are left in place for a user defined period (e.g., 2 days) and the acoustic signals are recorded at one second intervals over a period of two hours during the night (usually between 2:00 and 4:00 a.m. - when the background noise is likely to be lower). By recording and analysing the intensity and consistency of the acoustic signals (e.g., frequency analysis of the noise levels), each logger indicates the likely presence (or absence) of a leak. The comparison of the leak-induced noise amplitude at the different logger locations then provides an indication of the likely leak position. Suspected leaks are then exactly located by using (a) ground microphones to listen for the leak-induced noise at the ground surface (i.e., indirect sounding) directly above the pipes at very small intervals (about 1 metre), or (b) leak noise correlators, which determine the position of the leak based on: (1) the time shift of the maximum correlation of the two leak-induced acoustic signals, (2) the

distance between the two measurement points, and (3) the propagation velocity of the leak-induced noise.

The main advantage of the techniques that make use of listening devices is that these pieces of equipment are easy to use. However, these techniques' leak detection and location efficiency strongly depends on the experience and skill of a leak inspector and the leak detection and location process is time consuming. Additionally, in spite of experience, the leak inspector's fatigue plays also an important role. Indeed, concentration decreases with time and weakly audible leak-induced noises may not be detected (Fuchs and Riehle, 1991). The degree of detail of a sounding survey carried out by using listening devices also affects the efficiency of these techniques (i.e., the general surveys mainly allow the detection of clearly audible leak-induced noises whilst the detailed surveys may allow the detection of weakly audible leak-induced noises also). Finally, the interference through noise from traffic, ground movement, wind and water use is another factor affecting the efficiency of these techniques.

The use of noise loggers for carrying out a sounding survey, on the other hand, can be beneficial in WDS sections that have high surrounding noise during the day. Additionally, it allows examining the sensors' response remotely. Although traditional noise loggers are not suitable for determining the exact location of a leak, the latest generation of noise loggers incorporates a multi-point correlation facility to provide a leak's exact location (Farley, 2005). Successful experiences of using such pieces of equipment for inspecting entire WDSs or WDS sections have been reported by several water companies in the UK (e.g., South East Water, 2005; Vision, 2005). However, their leak detection effectiveness is questionable (Hunaidi and Wang, 2006). For example, Van Der Kleij and Stephenson (2002) found that sounding surveys carried out by using the noise loggers are not a viable alternative to sounding surveys performed by skilled leak inspectors using listening devices. Taking this into consideration, it is worth noting that unlike other acoustic devices the noise loggers can also been used for monitoring purposes (i.e., detecting and locating new leaks as they occur). In this case, the noise loggers have to be permanently installed in a WDS section. Results from a 6 month pilot study whereby 75 kilometres of distribution pipes were monitored by using 460 noise loggers can be found in Sánchez et al. (2005). Of particular interest is the cost-benefit analysis that the authors performed in their study. This indicated that the

high installation cost of the noise loggers can only be justified in WDS sections with high deterioration rates (i.e., greater than 20 bursts/100 kilometres/year).

With regard to the leak noise correlators, it has to be noted that the first of these devices, developed in the late 1970s by the Water Research Centre in the UK, was found to be too time consuming, difficult to use and expensive for commercial use (Shaw Cole, 1979). However, leak noise correlators have been improved over the years quite considerably. Nowadays, quicker and easier to use leak noise correlators such as LOCAL (Fuchs and Riehle, 1991), LEAKTEC (Seaford, 1994), HEARLEAK (McNulty, 2001), and AQUASCAN (Gutermann, 2011) are available. Despite this, they are still quite expensive and remain beyond the means of many water and leak detection service companies (Hunaidi and Wang, 2006). The main advantages of using the leak noise correlators over the other acoustic devices are as follows. They allow: (i) suppressing the interference through noise from the environment and from the WDS section itself, and (ii) replacing the reliance on the leak inspector's experience with an objective location system. Existing leak noise correlators, however, require extensive training, are unable to deal with the presence of multiple leaks in an inspected pipeline segment, require accurate knowledge of the inspected pipeline segment's topology for locating the leaks (especially if the inspected pipeline segment is branched), and can be unreliable for quiet leaks in cast and ductile iron pipelines and for most leaks in large diameter and plastic pipelines (Hunaidi and Wang, 2006).

Regardless of the particular type of device employed, the acoustic equipment-based techniques are the techniques most commonly used by the water companies for detecting and locating leaks in their WDSs. Thus, not surprisingly, many papers have described the experiences and results obtained from their application in several real-life WDSs/WDSs sections or experimental facilities (e.g., Hunaidi et al., 2000; Fantozzi and Fontana, 2001; Bracken and Hunaidi, 2005). These experiences and results, while demonstrating that leaks may be successfully detected and located in most cases, also highlight that the effectiveness of the acoustic equipment-based techniques is influenced by a large variety of factors (e.g., pipeline type, size, and depth, soil type and water table level, leak type and size, pipeline pressure, interfering noise, sensitivity and frequency response of the acoustic equipment, etc.). Note that a more detailed description of how and to which extent some of these factors influence the leak

detection and location effectiveness of these techniques can be found in Hunaidi and Wang (2006). Having said this, the main drawback identified is as follows. As the acoustic equipment-based techniques have been primarily developed for metallic pipelines, they are not equally effective with every pipeline type. For example, Hunaidi and Chu (1999) and Hunaidi et al. (2000) described the difficulties they encountered with plastic pipelines. In particular, Hunaidi and Chu (1999) stated that, when using leak noise correlators, the high attenuation of the leak-induced noise may require the accelerometers to be placed at greatly reduced distances. This, in turn, entails a greater number of access points on plastic pipelines compared to the metallic ones and may render the use of these devices not practical. In this regard, note that the aforementioned requirement for the accelerometers to be placed at greatly reduced distances can be partly offset by using hydrophones. This is because, the water borne leak-induced noise is not damped to the same extent as the leak-induced noise travelling through the wall of a pipeline. However, the use of hydrophones has the following disadvantages: (i) it requires specialised insertion equipment and practices, and (ii) it may increase the risk of bacteria or pathogens entering the water stream. Finally, it is worth mentioning that the use of the acoustic equipment-based techniques on large diameter transmission pipelines has also been recognised as a difficult task (Farley, 2005; Pilcher et al., 2007). This is due not only to the poor sound propagation qualities of the transmission pipelines but also to the large distances between access points.

2.3.3 Tracer gas-based techniques

Leak detection and exact location by means of a tracer gas-based technique (Field and Ratcliffe, 1978; Hargesheimer, 1985; Hunaidi et al., 2000) consists of injecting an inert, harmless (i.e., non-toxic), water insoluble and lighter than air gas (e.g., helium, hydrogen, etc.) into an isolated WDS section. This is then followed by ground scanning with a portable and extremely sensitive gas sensor for detecting the gas as it escapes via the leaks and rises through the surrounding soil to the ground surface.

Although this technique may allow detecting and locating the leaks that are difficult to detect and locate by using any of the acoustic equipment-based techniques (i.e., leaks in non-metallic/large diameter pipelines, multiple leaks in an inspected pipeline segment, and leaks in a branched pipeline segment), Hunaidi et al. (2000) pointed out that it has several limitations. They observed that the time needed for the gas to surface is

relatively long and if the leaks are not near to the top of a pipe the gas may not escape. Given these observations this technique is impractical for routine leak inspection surveys. Furthermore, because the escaping gas can be detected only within a small distance from the leak location, a leak could be easily missed if the resolution of the survey is too coarse or if the scanning is not performed directly above the leaking pipe.

In general, in order to use a tracer gas-based technique the WDS section where a leak is suspected has to be isolated and dewatered. Furthermore, this technique often requires an expert contractor to carry out the work and is therefore very expensive. The effectiveness of this method strongly depends on the weather conditions and wind direction.

2.3.4 Infrared thermographic techniques

In a thermographic inspection survey a high resolution infrared camera is used to scan the ground above a pipeline for thermal anomalies. These thermal anomalies may be due to: (i) differences in the temperature between the leaking water and the overlying soil, and (ii) changes in the thermal characteristics (i.e., thermal conductance) of the soil close to the leak as it becomes saturated by the leaking water. The camera measures the energy pattern (i.e., thermal infrared radiation) at the ground surface and displays its visible images (i.e., areas of differing temperatures designated by differing grey tones on a black and white image, or by various colours on a colour image). As a result, the ground surface above a leak may appear cooler or warmer than the surface farther away from it. The survey can be carried out by foot, from moving vehicles or helicopters and has the potential to quickly cover large WDS sections.

Several application of the infrared thermographic technique applied to pipelines in chemical plants, and waste water and water pipelines/pipeline systems are reported in the literature (e.g., Weil, 1993; Weil et al., 1994; Wirahadikusumah et al., 1998). These case studies indicate the potential of the technique to detect and approximately locate not only leaks but erosion voids and poor backfill as well. Furthermore, they show that by performing computer analysis of the resulting thermal images, the accuracy and speed of the test interpretations can be greatly improved.

Despite the outlined advantages of the infrared thermographic technique, the major drawback of using this technique is that the inspection survey relies on the use of a

single sensor. This means that, even if computer analysis of the thermal images is performed, experience is crucial in interpreting the thermograms (Wirahadikusumah et al., 1998). Other issues to be considered are: (i) the most appropriate survey time, (ii) the effect of the season, (iii) the infiltration into adjacent sewer pipelines (as it prevents the moisture movement from reaching the ground surface), (iv) the ground water table, and (v) the effect of ambient conditions such as thermal noise (especially in urban settings), cloud cover, and relative humidity (Fahmy and Moselhi, 2009). The latter issue strongly undermines the application of this technique in urban areas.

2.3.5 Ground Penetrating Radar-based techniques

In a GPR inspection survey high frequency pulsed electromagnetic waves are sent into the ground. As the radar waves travel through different materials their velocity changes and reflections are produced. This is due to changes in the materials' permittivity and conductivity (Metje et al., 2007). The reflected radar waves are captured by a receiving antenna at the surface. The time lag between transmitted and reflected radar waves is then used to form a continuous cross-sectional profile or record (i.e., image) of the subsurface. In this context, a leak can be detected and approximately located by identifying: (i) the underground voids (created by the water as it circulates near the leaking pipe), or (ii) the anomalies in the depth of the leaking pipe (as soil that is saturated by leaking water slows down the radar waves and makes the leaking pipe appear deeper than it should be).

GPR-based techniques have mainly found their application in the mapping of the utility lines and buried tanks or pipelines (e.g., Annan et al., 1984; Caldecott et al., 1988; Zeng and McMechan, 1997). However, a large number of papers that deal with the detection and location of leaks in WDSs can be found in the literature. On the one hand, the vast majority of these papers report the results of tests conducted by mapping controlled releases of fluids in test pits or by using computer simulations (e.g., Graf, 1990; Hunaidi and Giamou, 1998; O'Brien et al., 2003; Lockwood et al., 2003; Hyun et al., 2003, 2007; Nakhkash and Mahmood-Zadeh, 2004). On the other hand, only a limited number of papers report results of experiments carried out under more realistic conditions (e.g., Charlton and Mulligan, 2001; Stampolidis et al., 2003). As an ensemble, these results show that GPR units can be an effective tool for detecting and approximately locating leaks as they are also suitable for leaks in large diameter or non-metallic pipelines.

Furthermore, Farley (2007) states that the use of GPR units is particularly cost effective when these devices are employed as a rapid reconnaissance survey tool on long lengths of transmission pipelines (e.g., survey speed of 15 - 30 kilometres per hour when the GPR units are mounted on a survey vehicle). However, the aforementioned results also show that despite its merits a number of drawbacks limits the use of the GPR-based techniques. Firstly, GPR units have difficulties in highly conductive clay and saturated soils. This is because the pulsed electromagnetic waves lose strength very quickly in conductive materials thereby affecting the depth of penetration. This, in turn, means that in situations where the pipes are laid deeper into the ground (e.g., more than 1.8 metres deep to avoid water freezing) the use of GPR units can prove totally unsuitable. Secondly, anomalies like metallic objects in the ground can lead to false conclusions. Thirdly, subtleties in processing and interpreting the images from a GPR inspection survey mean that less skilled investigators may fail to detect and locate the leaks.

Bearing in mind the above, it is worth highlighting some of the research studies into the GPR technologies that have focussed on overcoming such drawbacks. Hata et al. (1997) worked on antenna design and produced a deep GPR unit capable of surveying at depths of up to 5 metres in favourable conditions. Ciochetto and Polidoro (1998) proposed the use of an array of antennae operating simultaneously (instead of a single one). The multi-antenna system allows for a three-dimensional survey of the area under investigation. The bulk of the research, however, has focussed on the interpretation of the GPR inspection survey's images. Bernstein et al. (2000), for example, used digital signal processing software for suppressing images of discrete objects and enhancing the images of the pipelines. Their Ground Penetrating Imaging Radar system allowed creating sharp, three-dimensional images of the underground pipelines.

Despite the technical advancements outlined above, the complexity of interpreting the GPR inspection survey's images continues to deter the water companies from more extensive use of the GPR-based techniques (Farley, 2005). Furthermore, the properties of the ground surrounding the buried pipelines, in particular moisture content, still drastically affect the efficacy of these techniques for the detection and location of leaks (Thomas et al., 2006). Also, these techniques do not provide a reliable picture of the condition of a pipeline due to the reliance on a single mode of data collection.

2.3.6 Inline Pipeline Inspection Gauge-based techniques

Inline PIGs are computerised, self-contained devices inserted into a WDS section and propelled forward by the flowing liquid, recording information as they go. PIGs were originally developed to remove deposits which could obstruct or retard the flow through a pipeline (Trenchless Technology Network, 2002). Nowadays, PIGs carry a wide range of surveillance equipment and are used as an inspection tool for various purposes. These include leak detection and location, and pipeline's condition assessment (e.g., Licciardi, 1998; Anonymous, 1997).

Specialised robotic PIGs rely on various technologies for the detection and location of leaks, and/or for the characterisation of the features of a pipeline. Mergelas and Henrich (2005) evaluated the performance of a PIG for detecting and determining the exact location of a leak that carries microphones. As the PIG passes any leak it detects the sound generated and gives an indication to the operator. The PIG has a built in tracing device and is mounted on an umbilical cable which allows controlling its position along a pipeline precisely. Note that, some latest developments of this technology can be found in Fletcher (2008). On the other hand, pipeline corrosion, crack, and geometry detection PIGs commonly make use of magnetic fields, electromagnetic waves, or ultrasounds sending signals through a pipeline's wall. By analysing the reflected signals, changes in the thickness of the wall of a pipeline, cracks, and the pipeline's internal diameter can be identified/measured. Examples of using PIGs carrying permanent magnets or electro-magnets that generate magnetic flux in the pipeline can be found in Atherton (1989), Mukhopadahyay and Srivastava (2000), and Joshi et al. (2006). In McDonald and Makar (1996) results of several field trials using the Remote Field Eddy Current Hydroscope technology are reported. Willems and Barbian (1998) reported the results obtained from inspection of operational pipelines by using a PIG that is based on ultrasounds.

As the studies mentioned above show, the PIG-based techniques may be very effective for detecting and exactly locating leaks, and for assessing the integrity of a WDS section. However, access to the inside of a pipeline is needed and the application of PIGs is limited to large pipelines (i.e., diameters larger than 200 millimetres – see Anonymous, 1997). Furthermore, in distribution pipeline systems, the presence of valves, elbows, pipes with different diameters and pipe flanges may limit the PIG's

movement. In general these tools need clean pipelines. Therefore, it is difficult to apply this technique to old pipelines where there may be heavy corrosion. Before a survey is performed, attention must also be paid to the effect on the water quality (as the PIGs are in contact with the inner wall of a pipeline).

2.3.7 Section summary

Several hardware-based leak detection and location techniques have been reviewed in this section. These techniques have different limitations, capabilities, costs, and benefits. The performance and feasibility of each reviewed technique are mainly associated with its capability to inspect small or large WDS sections at once, its potential to be used for different pipeline types/sizes, its degree of reliance on the operators' experience and its applicability with or without the need of disrupting service. In this context, Table 1 summarises the main characteristics of the hardware-based leak detection and location techniques that have been reviewed in this section. Table 2 summarises their individual strengths and weaknesses.

Based on what reported in the aforementioned tables and on the review carried out in this section it is possible to state the following:

- Techniques that make use of acoustic equipment are, at the present, widely used by the water companies. Their overall effectiveness (under certain circumstances) is well established.
- Techniques that make use of infrared cameras, gas sensors, GPR units, and PIGs have found limited application so far. Although many of them proved effective in a number of cases, no definite answer regarding their overall effectiveness can be reliably established.

It is also possible to state that some of the aforementioned hardware-based techniques are the most accurate in today's leak detection and location surveys (Puust et al., 2010). This accuracy though, comes with the high costs of the devices they use. Furthermore, other factors such as the potential interruption to the water supply they may require or the potential water quality problems arising from the introduction of various devices into the pipelines have to be factored in as well.

Bearing in mind the above, it is important to note that, nowadays, the competition in the market and the increasingly cheaper manufacturing and electronics are bringing down the costs of the hardware devices reviewed in this section. Furthermore, the technological advancements and the continuous innovation are resulting in better and easier to use devices. This, in turn, is enabling the water companies to detect and locate the leaks that are present in their WDSs more efficiently and effectively than ever before, thereby avoiding expensive 'dry holes'. This is especially true when such challenges as the increasing use of plastic pipes and the implementation of pressure reduction programmes are taken into consideration. In this scenario, of particular relevance is the fact that the water companies are now able to adopt a "multi-sensor" approach whereby the whole range of devices is considered and an appropriate mix of equipment (for the specific WDS section characteristics, site locations, leak types, etc.) is selected.

Notwithstanding, the hardware-based techniques are mainly used for regular inspections rather than for continuous monitoring. Inspections of entire WDSs/WDS sections by using these techniques are often expensive in terms of man-hour needed. As a consequence, they are carried out infrequently. This implies that the detection time of a leak depends on the interval between inspections. In view of this, it is when the hardware-based techniques are observed from a "time to respond to a leak" perspective that their main limitation becomes apparent. This limitation is that they are, if used on their own, unsuitable for quickly responding to the new leaks. In this scenario, water companies have a need for techniques that would alert them as soon as a leak occurs and guide them to the problem area. Then, the hardware-based techniques could be used to quickly and accurately determining the leak position thereby contributing to minimise the overall leak runtime and improving the water companies' customer service by reducing the potential interruptions to the water supply.

To this end, numerical (i.e., less expensive) techniques that are used or have been proposed to overcome the aforementioned inability to quickly respond to the new leaks, and/or the other outlined limitations of the hardware-based techniques are reviewed in the following section.

Table 1. Main characteristics of the hardware-based techniques.

	Time consuming	**Invasive**	**Cost**	**Inspection/ Monitoring**	**Detection/ Location (Approx. – Exact)**	**Pipelines/ Pipeline systems**
Acoustic equipment-based - *Listening devices*	Very high	No	Medium	Inspection	Detection/ Location (Approx. – Exact)	Pipelines/ Pipeline systems
Acoustic equipment-based - *Noise loggers*	High	Depend	High	Inspection/ Monitoring	Detection/ Location (Approx. – Exact)	Pipelines/ Pipeline systems
Acoustic equipment-based - *Leak noise correlators*	High	Depend	Very high	Inspection	Detection/ Location (Approx. – Exact)	Pipelines/ Pipeline systems
Tracer gas-based techniques	Very high	Yes	Very high	Inspection	Detection/ Location (Exact)	Pipelines/ Pipeline systems
Infrared thermographic techniques	Medium	No	High	Inspection	Detection/ Location (Approx.)	Pipelines/ Pipeline systems
GPR-based techniques	Medium	No	High	Inspection	Detection/ Location (Approx.)	Pipelines/ Pipeline systems
Inline PIG-based techniques	High	Yes	High	Inspection	Detection/ Location (Exact)	Pipelines

Table 2. Main strengths and weaknesses of the hardware-based techniques.

	Pros	**Cons**
Acoustic equipment-based ***- Listening devices***	• Easy to use equipment	• Depend on the operator's experience • Affected by the operator's fatigue • Depend on the survey's degree of detail • Affected by the non-leak noise • Unreliable for non-metallic and large diameter pipelines
Acoustic equipment-based ***- Noise loggers***	• Beneficial in WDS sections that have high surrounding noise during the day • Enable examining the sensors' response remotely	• Limited leak detection effectiveness • Unreliable for non-metallic and large diameter pipelines • When used for monitoring purposes they are economically viable only for WDS sections with high deterioration rates
Acoustic equipment-based ***- Leak noise correlators***	• Enable suppressing the non-leak noise • Do not depend on the operator's experience	• Require extensive training • Unable to deal with multiple leaks • Leaks in branched pipeline segments require particular attention • Unreliable for non-metallic and large diameter pipelines
Tracer gas-based techniques	• Able to deal with multiple leaks and leaks in branched pipeline segments • Reliable for non-metallic and large diameter pipelines	• Impractical for routine surveys • Require isolation and dewatering of the inspected WDS section • Require expert contractors to carry out the work • Depend on the weather conditions and wind direction
Infrared thermographic techniques	• Can also detect and locate erosion voids and poor backfill	• Depend on the operator's experience • Affected by the ambient conditions such as thermal noise and relative humidity (i.e., difficult to use in urban areas)
GPR-based techniques	• Reliable for non-metallic and large diameter pipelines	• Depend on the operator's experience • Impractical in highly conductive clay and saturated soils • Affected by the presence of anomalies such as metallic objects in the ground
Inline PIG-based techniques	• Can also be used for condition assessment	• Affected by the presence of WDS elements (e.g., valves) and corrosion • Applicable to larger pipelines only • Entail contact with the inner surface of a pipeline

2.4 Hydraulic Techniques

2.4.1 Section overview

A leak is first of all a hydraulic phenomenon, thus it is not surprising that it can be identified by studying a WDS section hydraulic behaviour in its presence or when it occurs. Hydraulic techniques for leak identification in WDSs make use of process parameters such as pressure and flow. They range from techniques that only quantify the amount of water lost through leaks in a whole WDS or in a WDS section, to techniques that allow assessing, detecting and locating the leaks as soon as they occur. This section provides a review of literature dealing with some of the hydraulic leak identification techniques that are nowadays commonly used by the water companies or that are currently in the stage of research and development. Furthermore, for each reviewed technique, this section gives a brief description of the principles on which it is based and the way it works. The main aims of the review carried out here are: (i) to highlight the main capabilities and limitations of the hydraulic techniques, and (ii) to identify the state-of-the-art in their research and development.

The hydraulic techniques that are considered in this section include:

- Water audits;
- Step tests;
- Steady state analysis-based techniques;
- Transient analysis-based techniques;
- Negative pressure wave-based techniques;
- Statistical/AI-based techniques.

Relevant literature dealing with each of the above techniques is reviewed in Sections 2.4.2 to 2.4.7. Once this is done, Section 2.4.8 provides a summary of the considered techniques' main characteristics, advantages and disadvantages and presents concluding remarks about their potential to help water companies minimising the leaks' runtime.

2.4.2 Water audits

A water audit (known internationally as water balance) accounts for all water flowing into and out of a water company's WDS (i.e., system-wide audit) or a particular WDS

section with well-defined boundaries (i.e., district audit) over an arbitrary period of time (e.g., 12 months). It allows assessing the total amount of lost water and is based on measurements and/or estimations of the water produced, imported, exported, used and lost. Guidelines for conducting water audits have been published by the American Water Works Association (AWWA, 1999) and by the International Water Association (IWA, 2000). Table 3 shows the various components of consumption and losses according to the standard International Water Association water balance.

Table 3. International Water Association standard for the international water balance and terminology.

<table>
<tr><td rowspan="10">System Input Volume</td><td rowspan="3">Authorized Consumption</td><td rowspan="2">Billed Authorized Consumption</td><td>Billed Metered Consumption (including water exported)</td><td rowspan="2">Revenue Water</td></tr>
<tr><td>Billed Unmetered Consumption</td></tr>
<tr><td rowspan="2">Unbilled Authorized Consumption</td><td>Unbilled Metered Consumption</td><td rowspan="8">Non-Revenue Water</td></tr>
<tr><td>Unbilled Unmetered Consumption</td></tr>
<tr><td rowspan="7">Water Losses</td><td rowspan="3">Apparent Losses</td><td>Unauthorized Consumption</td></tr>
<tr><td>Customer Metering Inaccuracies</td></tr>
<tr><td>Systematic Data Handling Errors</td></tr>
<tr><td rowspan="4">Real Losses</td><td>Leakage on Transmission and Distribution Mains</td></tr>
<tr><td>Leakage and Overflows at Utility’s Storage Tanks</td></tr>
<tr><td>Leakage on Service Connections up to point of Customer Metering</td></tr>
</table>

It has to be stressed that whilst system-wide audits give an overall picture of the water losses in a WDS, district audits provide a much more detailed picture (i.e., water losses in different WDS sections). The latter type of audit may therefore enable the water companies identifying the sections of their WDSs that are experiencing the highest leakage levels and, as a consequence, allow them to target the resources for leak detection and location to the greatest effect.

In general, a water audit can be seen as a two part procedure. In the first part of the procedure, the Top-Down Real Losses Assessment method is used. According to this method the data about the various components of consumption and losses in Table 3 are extracted from a water company's existing records (e.g., bills, reports, etc.) without any field work. If some of these data are not available they will be based on estimates. In this regard, Liemberger and Farley (2004) describe what data have to be collected and how to estimate the unknown elements. Although the resulting real losses figure is only a rough estimate of its true value, the information obtained after this part of the procedure has been completed is valuable not only for providing a first understanding of the water company's water delivery efficiency (at the system-wide or district level) but also for raising awareness about the required actions in order to improve the water audit accuracy.

In the second part of the procedure (usually implemented after several Top-Down Real Losses Assessments have been completed and the water company has a thorough understanding of the findings from these "desktop exercises"), the Bottom-Up Real Losses Assessment (Liemberger and Farley, 2004) and the Real Losses Component Analysis (UKWIR, 1994; Lambert, 1994) methods are used in order to minimise the difference between the estimated figures of the consumption and losses components, and their real values. On the one hand, the Bottom-Up Real Losses Assessment method aims at more accurately estimating the "aggregated" real losses component. This is achieved by performing analysis of the night flow in a representative sample of WDS sections (i.e., for system-wide audits) or in a particular WDS section (i.e., for district audits). In this analysis the Net Night Flow is obtained by subtracting an assessed amount of legitimate night uses (i.e., exceptional night use, non-household night use, and household night use) from the measured Minimum Night Flow (MNF). The MNF in urban situations normally occurs during the early morning period (usually between 2:00 a.m. and 4:00 a.m.). During the early morning period, water demand reaches its minimum value while the opposite applies to the level of pressure. The combination of both these factors means that in the measured MNF the real losses are at the maximum percentage of the total flow. The daily level of real losses can then be estimated by: (i) multiplying the Net Night Flow by 20 hours (Stemberg, 1982), or (ii) applying the Fixed and Variable Area Discharge principle (May, 1994; Lambert, 2001). Note that the

latter approach allows taking into account the fact that pressure is not constant over the day. On the other hand, the Real Losses Component Analysis method aims at obtaining accurate figures for each of the real losses components in Table 3. This method was initially developed in the UK in the early 1990s and has meanwhile become the standard real losses analysis tool in many water companies around the world. In this analysis, data about the characteristics of a WDS section (e.g., total length of pipes, total number of service connections, etc.), the number of repairs carried out in the WDS section, the average pressure across the WDS section, and leak awareness, location and repair times are used to evaluate the water losses relative to each of the relevant real losses components.

By applying the two part procedure described above, more reliable figures of the various components of consumption and losses from the water balance shown in Table 3 are obtained, as well as pieces of information useful for identifying all the internal issues that are preventing a water company from achieving high water delivery efficiency. It has to be stressed that, in order facilitate the repeatability of this procedure, the boundaries of the considered WDS sections should be not only well-defined, but also permanent. In the UK, as well as in other parts of the world, this necessity motivated the sectorisation of the WDSs into DMAs (UKWIR, 1994, 1999a), which are WDS sections isolated from the rest of the WDS by closing appropriate valves and in which the quantities of water entering and leaving the section are metered.

When using a water audit, not only the amount of water lost through leaks but also other unaccounted-for water losses (e.g., water theft) can be estimated. However, the sensitivity of a water audit to leakages is low. It also depends on the accuracy of the flow meters, which typically have 15% inaccuracies.

The aforementioned low sensitivity of the standard water audits motivated researchers to develop methods for more accurately assessing the leakage levels in WDS sections. Notably, Buchberger and Nadimpalli (2004) demonstrated on several examples (with observed and simulated pipe flows) the good accuracy of a method based on sequential statistical analyses of continuous high frequency (i.e., 1 second sampling interval) flow measurements taken at the entry point of a DMA. This method was not field tested and, as a consequence, more work is needed to verify it. Despite this, such work highlights

the fact that the use of continuous high frequency measurements of discharge may yield more accurate leakage level estimates. Taking this into account, it is conversely important to note that this can be problematic since high frequency data are not used often in WDSs.

2.4.3 Step tests

Step tests allow detecting leaks within a particular subsection of a DMA. This is done by subdividing the DMA itself and then measuring flow rates while making temporary valve closures to cut off the different subsections in succession. The resulting reduction in flow rate following the closure of a particular valve indicates the total leakage rate plus legitimate consumption in that subsection. If the resultant reduction is greater than anticipated, taking into account the number and type of customers in the subsection being isolated, then an indication of the likely presence of leaks is obtained. Step tests are generally carried out during the period of MNF.

According to Pilcher et al. (2007), two methods for performing step tests can be identified. These are the Isolation, and the Close and Open methods. The Isolation method involves the successive closing of valves starting from the one farthest away from the flow meter and working back to the flow meter where the flow should drop to zero. The Close and Open method involves closing valves to isolate each individual subsection and once the reduction of flow has been recorded the valves are reopened.

In both methods outlined above the various subsections are de-pressurised for a certain amount of time due to their shutting down. This, in turn, can cause backsiphonage or the risk of infiltration of ground water. The shutting down of the various DMA subsections also causes supply interruption inconvenience. As these tests are usually carried out at night, however, supply problems to the majority of customers are avoided. Allowing for this, it has to be also noted that the fact that these tests are usually carried out at night is in itself a problem, as it requires ‘out-of-hours’ work.

2.4.4 Steady state analysis-based techniques

The techniques in this category enable identifying leaks in a WDS section by solving an inverse problem and using a steady state analysis approach. Generally speaking, in an inverse problem, the state of the WDS section is known/measured (e.g., pressures, some

demands, flows, etc.) but some parameters (e.g., roughness of the pipes, other demands, leaks, etc.) are unknown. Potential leaks (e.g., leaks of varying size modelled at the WDS section nodes) are tested in a numerical Hydraulic Model (HM) of the WDS section until the measurements satisfactorily match the corresponding simulated variables (e.g., pressure traces). The inverse problem is then solved for the WDS section parameters like pipe roughnesses and leak location/size.

Inverse analysis by using a steady state approach was first introduced by Pudar and Liggett (1992). Subsequently, many studies have been published that propose variations of the basic Inverse Steady State Analysis method (e.g., Mukherjee and Narasimhan, 1996; Holnicki-Szulc et al., 2005; Tabesh et al., 2005; Sage, 2005; Wu and Sage, 2006, 2007; Wu et al., 2008, 2010; Deagle et al., 2007; Puust et al., 2006; Borovik et al., 2009; Skworcow and Ulanicki, 2011; Misiunas et al., 2006). Here, a small selection of these studies is reviewed in more detail in order to highlight the main approaches used and the results obtained. This, in turn, enables identifying their main strengths and weaknesses.

Pudar and Liggett (1992) used pressure and flow "measurements" (i.e., pressure and flow simulated by using a steady state HM) at various nodal points in a simple pipeline system (i.e., 11 pipes and 7 nodes). Leaks at nodes were expressed in terms of pressure by an orifice formula. The objective function of the inverse problem was defined as the sum of squared differences between measured and simulated pressures. The minimisation of this objective function was carried out by means of a non-linear derivative-based optimisation method, namely the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963; Moré, 1977). The study highlights that even for small pipeline systems the Inverse Steady State Analysis method requires a huge amount of pressure and/or flow measurements and that the inverse problem is not well posed given the expected amount of available measured data. Furthermore, it shows that calculation of the leak location and size is sensitive not only to the quantity of the measurements but also to their quality (i.e., accuracy), the accuracy of the nodal demand and the accuracy of the pipe roughness estimates.

Sage (2005) proposed a method whereby the total leakage rate for a DMA was calculated as the difference between the metered inflow and the metered actual

consumption (or the modelled demand – based on a combination of metered and estimated consumption) and assumed to comprise of background leakage rate and burst leakage rate. The total leakage rate across each DMA was then attributed to the HM's nodes in direct proportion to the numbers of properties or mains length associated with each node. The background leakage rate for each node or pipe was calculated by an empirical formula, captured as the "Internal Condition Factor" which is based on the age of the pipe (i.e., the older a pipe, the greater the background leakage rate). The difference between the total leakage rate and the total background leakage rate was treated as the total burst leakage rate. A Genetic Algorithm (GA) optimisation technique (see – e.g., Holland, 1962; Holland, 1975) was then used to apply the total burst leakage rate across the HM's nodes. The solution fitness was evaluated by comparing the field observed pressures and the simulated pressures. Note that the availability of pressure measurements from sensors deployed with a density of one sensor per 200 houses was assumed. Hayuti et al. (2008) reported that a total of 59 DMAs were analysed by using this method to determine potential leakage "hotspots" (i.e., areas of water loss in the DMA – including leakages, unmetered and illegal consumptions, etc.). Leak inspectors were deployed to investigate the 26 "hotspots" identified. A total of 20 actual leaks were found on various asset types (i.e., mains, fittings, and services). Despite the encouraging results obtained, Wu et al. (2010) argued that this method presents the following two important limitations: (i) the leak is not modelled as a pressure dependent demand, and (ii) the leak identification problem is separated from the HM calibration process.

Wu and Sage (2006), despite modelling leaks as a volume based demand, worked on including the leak identification problem in the HM calibration process. The Darwin Calibrator (i.e., GA-based optimisation tool) within the WaterGEMS modelling package (Haestad Methods Inc., 2002; Bentley Systems Inc., 2006) was used for demand calibration. Subsequently, Wu and Sage (2007) used emitters to emulate the pressure dependent leaks at nodes. Note that, in contrast with the optimisation formulation that considers an emitter at each node of the HM (e.g., Pudar and Liggett, 1992), Wu and Sage (2007) proposed to prescribe the maximum number of emitters (i.e., possible leakage nodes). The GA-based optimisation was then used for simultaneously optimising the emitter locations and the corresponding emitter coefficients. In this way,

the proposed solution method avoided assigning one decision variable (i.e., emitter coefficient) to every node (for which the optimisation dimension increases exponentially with the number of nodes in the HM) and ensured scalability. The methods described in Wu and Sage (2006, 2007) were tested on a simple case study with simulation-generated field data and a relatively complex DMA with real field data. Some promising results were obtained whereby the deliberately allocated leakages in the simple case study were approximately located, and satisfactory indications of the leakage "hotspots" (checked against historical data on leak repairs) in the DMA case were provided. Building on these works Wu et al. (2008, 2010) combined the identification of leakage "hotspots" with the Extended Period Simulation HM calibration for improving the leak detection and location performances. In each of these papers the methodology was demonstrated on a real-life DMA in the UK. Similarly to Sage (2005), in Wu and Sage (2006, 2007) and in Wu et al. (2008, 2010) pressure measurements from sensors deployed with a density of one sensor per 200 houses were used. As an ensemble, the aforementioned studies emphasise the fact that to achieve reliable indication of the leakage "hotspots", the domestic and commercial demands have to be set up in the HM to an appropriate level of detail. If these are not representative of the real situation, the method may just highlight the deficiencies in demand allocation. A further concern emphasised relates to the accuracy of the pressure sensors used during field testing. Provided that equipment calibration is maintained, standard sensors used for modelling purposes are normally accurate to 0.1% of their full range (i.e., typically accurate to 0.1 metres). As the method identifies leakage "hotspots" based on the differences in total heads between different measurement locations in a DMA, sensor accuracy is an important issue, particularly when running simulations during the night period when head-losses across a DMA are often small.

Puust et al. (2006) proposed a method for solving the inverse problem that makes use of a stochastic (rather than deterministic) methodology. They used the Shuffled Complex Evolution Metropolis (Vrugt et al., 2003) optimisation algorithm. This algorithm allowed them to estimate the posterior probability density functions of the leak areas. The method was successfully demonstrated on two synthetic case studies (one of which with uncertain measurements). It has to be stressed that the use of a probabilistic methodology presents the following important advantage. As different kinds of errors

normally affect the calculations, obtaining a final discrete value for the leak size bounded with a certain probability (i.e., uncertainty) gives information about the result reliability. The main disadvantage of such methodology is the additional computational effort required.

The small selection of papers reviewed in this section clearly shows the potential of the techniques that make use of steady state HMs for assessing, detecting and approximately locating leaks within a DMA. Of particular interest is the fact that some of them aim at achieving this by making only use of a limited number of pressure sensors which are normally employed by the water companies for the calibration of their HMs. Nevertheless, at the present, the widespread application of these techniques in the water industry is seriously undermined by their total reliance on an accurate HM. Indeed, the roughness coefficients of the pipes and the demands have to be set up in the HM to a great level of detail. Additionally, assuming the above could be satisfactorily achieved, if even a single HM element is not correctly modelled (e.g., a valve set as opened in the HM is actually closed) these methods might give completely wrong results. In this regard, it has to be stressed that in many water companies the HMs (if at all available) are not constantly maintained, therefore the outlined situation often represents the rule not the exception. These techniques also require very accurate pressure measurements from the field tests. As a final note, it is worth mentioning that while the further development of these techniques for aiding the identification of the leaks that are already present in a WDS section is envisaged, it will be difficult to apply them for near real-time leak identification. This is mainly due to the long computation times they often require (e.g., long GA-based optimisation runs).

2.4.5 Transient analysis-based techniques

2.4.5.1 Overview

Transients are unsteady fluid flow. Transient flow occurs due to sudden changes in the WDS section conditions caused by rapid pressure/flow fluctuations. Some causes of transient flow include valve operations, pump start-up and shut-down and sudden leaks. These events generate transient waves that propagate away from the source. The speed of the transient wave depends on the characteristics of the WDS section (e.g., pipeline bulk modulus of elasticity) and density of the fluid. Since the transient wave speed can

be over 1000 metres per second, a high sampling frequency of measurements is required to measure the transient trace. It has to be noted that transient waves are shaped by their surroundings. Specifically, the shape of the waves is influenced by the physical (or propagation) structure of the WDS section (e.g., junctions, valves, elbows, blockages, leaks, changes in the conditions of the wall of the pipeline, etc.).

When a transient wave arrives at a leak a loss of energy (which manifests as a transient wave damping) occurs coupled with a reflected wave. The properties of the reflected wave and the transient wave damping can be exploited to achieve leak identification. For example, if the transient wave is sufficiently sharp and the reflected wave from the leak is large relative to reflections from other sources (e.g., junctions, valves, etc.), the latter may be usefully interpreted. However, if the transient wave is insufficiently sharp or the reflected wave from the leak is small, then the direct reflection information may be obscured by reflections from other sources and/or dispersion and damping. Nevertheless, energy is removed by the leak with each pass of the transient wave and this energy loss manifests as an increased damping. In this scenario, an understanding of the non-leak-related sources of dispersion and damping in a WDS section is required if the damping contribution from a relatively small leak, with indistinct direct reflections, is to be isolated.

The transient analysis-based techniques reviewed in this section make use of system elements (e.g., inline valves and pumps) or special devices (e.g., solenoid side discharge valves) for generating a customised transient event and allow identifying the leaks that are already present in a WDS section. The various transient analysis-based techniques in this category are classified here as follows:

- Inverse Transient Analysis (ITA) techniques;
- Frequency domain techniques;
- Direct transient techniques.

Sections 2.4.5.2 to 2.4.5.4 provide a review of each of the aforementioned techniques, respectively. This is then followed in Section 2.4.5.5 by concluding remarks about the practicality of applying these techniques for leak identification in real-life WDS sections.

2.4.5.2 Inverse Transient Analysis techniques

ITA techniques use the transient signal in the time domain in their analyses. They involve inverse calibration of a numerical HM of the pipeline/pipeline system section to the known/measured transient data. The main advantage of these techniques over the steady state analysis-based techniques described in Section 2.4.4 is that they make use of data from high speed pressure transducers. This way, each measurement station can provide an essentially unlimited number of data points in time, thus allowing calibrating the HM for the pipes' friction factor and the leak location/size simultaneously. By using more measurement data, a more accurate calculation of the friction factor of the pipes can be achieved. This, in turn, can lead to better determination of the leak location/size.

The first ITA technique was proposed by Liggett and Chen (1994). The authors applied their method to the simple pipeline system described in Pudar and Liggett (1992) using "measured" pressure heads obtained by first solving the forward problem in a transient simulator. As in Pudar and Liggett (1992) leaks were modelled at the pipeline system nodes. The objective function and the optimisation method used were as in Pudar and Liggett (1992), as well. Promising results were obtained as the procedure exhibited rapid convergence to the correct friction factor for all pipes and leak locations/sizes. However, it has to be stressed that the key to the theoretical success of the method lies in the fact that a perfect HM that accurately replicated the measured response was assumed.

The mathematical part of ITA technique proposed by Liggett and Chen (1994) has been refined in several subsequent studies. This has mainly involved: (i) finding more efficient methods for optimizing the objective function (e.g., Vítkovský and Simpson, 1997; Vítkovský et al., 1999, 2000, 2002; Kapelan et al., 2002, 2003), and (ii) modifying the classic Least-Squares objective function by adding to it a term to account for prior information (Kapelan et al., 2001, 2004).

The use of the ITA techniques for leak identification in WDSs has been demonstrated and tested in numerous studies reported in the literature. Besides the many numerical case studies and computer simulations (see – e.g., references provided above), experimental validations of the ITA techniques in transmission pipelines (e.g., Tang et al., 2000, 2001; Vítkovský et al., 2001, 2007; Covas et al., 2001, 2003) and in

distribution pipeline systems (e.g., Covas and Ramos 2001; Wang, 2002) have been performed. Furthermore, a number of field tests have been carried out by Stephens et al. (2004, 2005a, 2005b), Covas et al. (2004, 2005, 2006), Covas and Ramos (2010), Saldarriaga et al. (2006) and Karney et al. (2008), among the others.

As an ensemble, the aforementioned laboratory and field investigations highlight the seemingly forbidding complexities that can arise not only when the ITA technique is applied to real-life WDS sections but also to laboratory setups under controlled conditions. Vítkovský et al. (2007) stressed that these complexities arise because of the many sources of error that can creep into the ITA. These include noise in the measured variables, errors in the boundary conditions, and errors in the modelling of the transient (i.e., incorrectly-represented/unaccounted-for model processes, and numerical algorithm mistakes). They concluded that the latter errors are the most probable limiting factor for the successful field application of the ITA technique. To this end, experiences from the oil and gas industries show that, in single pipelines, a tuning process (i.e., calibration of the forward transient model) may normally address the aforementioned differences between measured and modelled transient response traces. According to this, several studies have attempted to calibrate model parameters such as the steady state friction factors, unsteady friction weighting functions and viscoelastic creep coefficients by using measured responses without any leak (e.g., Covas and Ramos, 2001; Covas et al., 2003). However, the possibility of applying similar procedures to real-life WDS sections, which have more uncertainties and are not known to be leak-free, is not clear. Also, note that having to calibrate the forward transient model is in contradiction with the fact that one of the most appealing theoretical advantages of the ITA techniques is to provide the location and size of the leaks with no calibration effort.

2.4.5.3 Frequency domain techniques

Frequency domain techniques use the transient signal in the frequency domain in their analyses. Generally speaking, they involve detecting and decoding the changes in the frequency domain response of a WDS section (i.e., Frequency Response Function – FRF – see e.g., Chaudhry, 1987) caused by the presence of the leak. That is to say, in order to determine the leak presence, location and size, the experimental/field-measured FRF from a WDS section with a leak is compared with that from an intact WDS section

which is itself obtained by numerically modelling from known WDS section geometry and parameters.

Applications of the frequency domain techniques to various synthetic cases including a single pipeline with one and two leaks, parallel and branched pipelines with one leak each and a simple reservoir-pipe-valve system with leaks of different sizes and at different positions can be found in Mpesha et al. (2001, 2002) and Lee et al. (2002, 2003, 2005a, 2005b). Among these, the works presented in Lee et al. (2005a) and Lee et al. (2005b) stand out. On the one hand, Lee et al. (2005a) introduced two novel frequency domain methods, namely the Inverse Resonant and the Peak Sequencing methods. The first method (like an ITA technique) entails inverse fitting. It seeks to minimise the sum of squared differences between measured and modelled FRFs by adjusting leak size and position within the numerical model. The second method is based on the uneven damping of the harmonic peaks across the frequency axis of the FRF caused by the presence of a leak. It allows determining the leak size and position within a WDS section through the matching of the observed FRF with the FRFs (for an identical WDS section) generated by leaks of different sizes and at various positions which have been previously archived in a lookup table. On the other hand, Lee et al. (2005b) derived an analytical expression for the leak-induced pattern on the FRF for a single pipeline. Using this analytical expression, a simple procedure was then developed for the identification of a leak from the experimental/field-measured FRF. This approach presents the great advantage of not relying on a comparison with the leak-free pipeline behaviour. Building on this work, Lee et al. (2006) presented the first experimental validation of a leak identification technique in the frequency domain. The technique was shown to be able to detect, locate and size a single leak under controlled conditions.

As an ensemble, the aforementioned studies show that the frequency domain techniques allow circumventing some of the problems that are faced when the ITA techniques are used. Indeed, instead of requiring the transient response at multiple locations throughout a WDS section, these techniques usually require measurements at only one location. This simplifies the process of data collection (as in many cases measurement data are not easily obtained) and also reduces the associated costs. Furthermore, since extensive measurements are not required, the possibility of obtaining questionable results due to

measurements with limited accuracy is reduced. Taking this into consideration, however, the aforementioned studies also show that the frequency domain techniques present several limitations that undermine their use in real-life WDS sections. Firstly they are strongly influenced by the location of the transient source and measurement station. Secondly, they require accurate modelling of the FRF from an intact WDS section. In this regard, it is important to note that, even if an analytical solution is used (i.e., Lee et al., 2005b), this is limited to single pipelines. As a result, only pipeline applications have been presented in the literature so far. Furthermore, tests on field pipelines are also lacking.

2.4.5.4 Direct transient techniques

Direct transient techniques attempt to isolate the leak-induced effects in the pressure trace, seeking to directly (i.e., without reproducing the transient trace in a simulator in the time or frequency domain) expose peculiarities/special features in the transient signal induced by the leak.

Various direct transient techniques have been proposed. For example, Jonsson and Larson (1992), Jonsson (1994, 2001), Brunone (1999), Brunone and Ferrante (1999, 2001), Covas and Ramos (1999), Covas et al. (2005a, 2005b), and Lee et al. (2007) presented approaches that aim at performing leak identification by analysing the pressure signals acquired during a transient in the time domain. Jonsson (1999), Covas et al. (2000), Ferrante et al. (2001), and Ferrante and Brunone (2003a), motivated by the observation that the aforementioned pressure signals usually exhibit a strong periodic component, made use of the Fourier analysis (see - e.g., Stein and Shakarchi, 2003). Jonsson (1999), Covas et al. (2000), Ferrante et al. (2001), and Ferrante and Brunone (2003a), in an attempt to preserve the information about the arrival time of the pressure waves when moving from the time to the frequency domain, made use of the Wavelet Analysis (WA) (see Appendix A). Aiming at better exposing the leak-induced peculiarities/special features in an acquired signal, Beck et al. (2005) made use of the Cross-Correlation analysis (see - e.g., Lange, 1987). Taghvaei et al. (2006) made use of the Cepstrum analysis (see - e.g., Bogert and Ossanna, 1966). Vítkovský et al. (2003), Kim (2005), and Lee et al. (2007) made use of the Impulse Response analysis (see - e.g., Suo and Wylie, 1989).

As an ensemble, the above studies show the potential of this form of analysis for leak identification in single pipelines and under controlled experimental conditions. Even in these cases, however, this form of analysis relies heavily on an accurate numerical model for obtaining the leak-free benchmark or on the existence of a measured leak-free response. Furthermore, it presents the following difficulties. Firstly, a sharp transient wave has to be usually used in order to allow the clear identification of a change in the measured pressure trace caused by the leak-induced transient reflection. This may be dangerous for a pipeline. Secondly, the identification of such change is usually carried out by visual inspection of the measured transient trace. This may give rise to ambiguous identifications. Thirdly, pipeline elements (e.g., elbows, valves, etc.), different pipe diameters, air pockets and blockages, among the others can also cause transient reflections. This may result in a change in the measured pressure trace that is not easily identifiable. Based on these observations, it is possible to state that, while the capabilities of the direct transient techniques have yet to be proven in field pipelines, the use of these techniques is unsuitable for more complex distribution pipeline systems.

2.4.5.5 Concluding remarks

Literature dealing with the transient analysis-based techniques has been reviewed in this section. This review clearly highlights that the transient analysis-based techniques require a good understanding of the WDS section being studied. This understanding cannot be limited to factors such as unsteady friction, pipe roughness, precise geometry, WDS section features, and demands but has to take into consideration the other physical complexities that may be encountered in the field as well. These include air pockets and/or entrained air, pre-existing blockages, pipe's wall thickness or lining variations, soil/pipe interactions, unsteady phenomena (both in the context of friction and minor losses), and mechanical motion and vibration, to mention only a few. Such understanding constitutes a very high hurdle to overcome.

The above gives an important insight into the likely practicality of applying the frequency domain and the direct transient techniques for identifying leaks in real-life WDS sections. Indeed, the identification and interpretation of the non-leak related sources of reflections, dispersion and damping in the measured transient responses is critical if leaks are to be successfully identified. These techniques, however, cannot currently be adapted to the aforementioned physical complexities, and system

configuration restraints. The ITA techniques, on the other hand, have the potential to distinguish between leak and non-leak related sources of reflections, dispersion and damping. However, this would require an accurate forward transient model. Such a model, in turn, would need to include in the mathematical equations governing the WDS section behaviour as many of the aforementioned physical complexities/configuration restraints as possible. It is therefore envisaged that the transient simulation models need to be developed further before the ITA techniques can be utilised for leakage identification in real-life WDS sections.

Attempts to validate the transient analysis-based techniques have entailed numerical case studies, laboratory experiments and limited field testing. These have shown that, even under controlled conditions with the simplest WDS sections, the results do not yet achieve the level of validation required for a strategy that must work in complex WDS sections under a wide range of conditions. Although these techniques might find application for identifying leaks in single pipelines, at the present, they are in no position to supplant existing techniques (Colombo et al., 2009).

2.4.6 Negative pressure wave-based techniques

In contrast to the transient analysis-based techniques reviewed in Section 2.4.5, negative pressure wave-based techniques try to exploit the transient that naturally occur in a pressurised WDS section following a leak event. Generally speaking, the identification of the leak-induced transient wave allows leak detection. The measuring of its arrival time at the measurement station(s) and the knowledge of the wave speed allow leak location. The evaluation of the difference between the flow rate after and prior to the leak at a suitable measurement station or the estimation of the magnitude of the leak-induced transient wave allows estimating the size of the leak.

Several studies can be found in the literature that present examples of the successful application of the principles outlined above to laboratory and field transmission pipelines (e.g., Silva et al., 1996; Misiunas et al., 2003, 2005a, 2005b) and to real (but relatively small – e.g., 250 properties, WaterWiSe@SG test bed) distribution pipeline systems (e.g., Misiunas et al., 2004, 2005c; Srirangarajan et al., 2010a, 2010b, 2012). These studies show that, by effectively enabling the on-line monitoring of a WDS section, the negative pressure wave-based techniques have the potential to provide a fast

response to the leaks. However, they also show that the performances of these techniques are influenced by a large number of factors. These include: (i) leak size, (ii) speed of the leak opening, (iii) distance from the leak to the nearest transducer, (iv) WDS section topology, and (v) measurements accuracy, among the others. Furthermore, several other issues have to be considered as well. Firstly, in a WDS section there are always transient waves induced by the normal WDS operations (e.g., pump start-ups, valve closures, etc.) or by changes in demand. These make a leak-induced transient wave unobvious and difficult to be distinguished. In other words, a leak-induced transient wave can be easily 'drowned' by complicated background noise. This, coupled with the fact that the shape of a leak-induced transient wave is unknown, affects the capability of the negative pressure wave-based techniques to actually identify a leak-induced transient wave (and hence enable leak detection). It also affects the capability of these techniques to locate the change point (or inflexion) of a leak-induced transient wave precisely (and hence enable leak location). Secondly, a leak-induced transient wave attenuates due to energy dissipation as it propagates along a pipeline. This attenuation increases with pipeline bends, constrictions in the pipeline, valves, etc. Hence, the negative pressure wave-based techniques are also faced with the difficult task of having to recognise very weak leak signals.

In view of the above, it is evident that the near real-time leak identification task by using the negative pressure wave-based techniques is very complicated. Although this is true for transmission pipelines and distribution pipeline system alike, it can be safely concluded here that the existing negative pressure wave-based techniques are in no position yet to be successfully used in large distribution pipeline systems (e.g., average size of the UK DMAs) with complex configurations.

2.4.7 Statistical/Artificial Intelligence-based techniques

2.4.7.1 Overview

As discussed in Section 2.4.2, in the UK (as well as in the other parts of the world) a WDS is typically sectorised into DMAs to more effectively assess the leakage levels. As a consequence of the latest developments in hydraulic sensor technology, the water companies in the UK and in other parts of the world alike have been able to deploy a larger number of more accurate and increasingly cheaper pressure and flow devices.

Nowadays, the UK DMAs are usually observed by using: (i) pressure and flow sensors located at the DMA entry (and any water import/export) point and (ii) a pressure sensor located at the critical point in the DMA (i.e., the one located either at the point of highest elevation or alternatively at a location farthest away from the inlet). These sensors usually collect the data at low sampling frequencies (e.g., 15 minutes). Additionally, it has to be stressed that the latest developments in on-line data acquisition/telemetry systems have made possible to receive the data collected by the aforementioned devices at pretty much the same rate as they are collected. All this means that usually the UK DMAs are now being monitored in near real-time.

In the scenario outlined above, a number of techniques have been proposed (and/or are currently under development) that use the aforementioned near real-time telemetry data for the automated on-line leak identification. As it will be shown in this section, many of these techniques make use of statistical and AI data analysis tools including ANN, Fuzzy Logic (see – e.g., Zadeh et al., 1975), Support Vector Machines (SVM) (see – e.g., Vapnik, 1998), Kalman Filters (see – e.g., Grewal and Andrews, 2008), and Principal Component Analysis (see – e.g., Jolliffe, 2002).

The statistical/AI-based techniques considered in this section can be broadly grouped into two main categories. The techniques in the first category are classified in this review as *model-based techniques* as they make use of a process model or HM of the WDS section. Literature relevant to this first category is reviewed in Section 2.4.7.2. The techniques in the second category are classified here as *measurement-based techniques* as they make only use of the received pressure and/or flow measurements. Literature relevant to this second category is reviewed in Section 2.4.7.3.

The review of literature relevant to the two aforementioned groups of statistical/AI-based techniques is then followed in Section 2.4.7.4 by concluding remarks about the usefulness and practicality of applying these techniques for leak identification in real-life WDS sections.

2.4.7.2 Model-based techniques

Researchers have long attempted to evaluate the current state of a WDS section from the on-line pressure and/or flow measurements and demand predictions (i.e., pseudo-measurements). Evaluation of the current state of a WDS section is particularly relevant

to the diagnosis of leaks and other operational faults. In theory, this may be accomplished by formulating and solving the mass and/or energy conservation equations. The requirement is for as many independent equations as there are variables to be calculated. However, as the aforementioned set of measurements and pseudo-measurements does not usually permit to completely describe all the variables of a WDS section, state estimation techniques have been proposed (e.g., Sterling and Bargiela, 1984). These techniques allow the processing of additional equations which are added to the minimal set to build in a degree of redundancy and hence the ability to cope with the limited/inaccurate information. The drawback of this approach, however, is that it adds a considerable overhead in terms of computational effort. To alleviate this difficulty the use of sparsity exploiting techniques, numerically stable factorisation and parallel and distributed computations has been proposed (e.g., Bargiela and Hainsworth, 1989; Hartley and Bargiela, 1995). However, the use of ANNs for solving the systems of equations involved has been found to give better results (Gabrys and Bargiela, 1995, 1996). This is because of the ANNs' intrinsically parallel structure, high computational efficiency and robustness as global state estimators.

In the framework set above, the leak identification problem has been solved by making use of techniques based on the analysis of residuals (e.g., mass balance errors) at the nodes and along the paths linking them (e.g., Carpentier and Cohen, 1991, 1993; Powell, 1992). Gabrys and Bargiela (1999), on the other hand, proposed a Neuro-Fuzzy approach based on the examination of the patterns of the state estimates. They successfully applied their approach to a small, but realistic, distribution pipeline system. Generally speaking, these approaches have a requirement for a high number of measurements. In order to ease this requirement, an implicit formulation of the standard Weighted Least-Squares state estimation problem based on loop equations (instead of nodal) was proposed by Andersen and Powell (2000). This formulation is more suitable for WDS sections with a low measurements redundancy. The feasibility of identifying leaks was demonstrated by the authors on a simple distribution pipeline system without taking into account uncertainty and measurement errors. In practical applications, however, both metering and telemetry systems are subject to errors. Furthermore, a great level of uncertainty comes about from the predominance of pseudo-measurements needed to compensate for the lack of real measurements. These issues have a negative

impact on the accuracy of the state estimation calculation. Therefore, results may be very inaccurate when compared to the actual WDS section state.

Similarly to Gabrys and Bargiela (1999), Izquierdo et al. (2007) proposed a Neuro-Fuzzy approach for the identification of leaks and other failures/anomalies in a WDS section. Their method used a mathematical model of the WDS section to generate fuzzy estimated states (i.e., estimated states together with uncertainty levels) for different anomalous states of the WDS section. These were then used to train an ANN capable of estimating the WDS section's leaks/other-failures/anomalies associated with the particular sets of measurement received through telemetry and demand predictions. As for other state estimation approaches, a drawback of this method is that a large number of sensors are needed in order to obtain the required information. Furthermore, its applicability to real WDS sections cannot be reliably established as this methodology was applied to a synthetic case study only.

Rather than solving the over determined WDS section's set of equations, ANN/SVM have been used to directly correlate the state variables' values to the WDS section states (i.e., Caputo and Pelagagge, 2002, 2003; Shinozuka et al., 2005; Mashford et al., 2009). According to this simplified approach, the ANN/SVM is trained on different sets of input data, which characterise several states of the WDS section under the normal and abnormal operating conditions. Once this is done, the trained ANN/SVM acts as a classifier in order to estimate the actual WDS section status based on the available information. The ANN/SVM recognises patterns among measurement data from the pressure and flow sensors distributed on the WDS section, and classifies them to identify the leaks.

Caputo and Pelagagge (2002) proposed an approach based on the utilisation of ANNs. In this work, the authors tested on artificial data sets (obtained from a mathematical model) possible ANN architectures (i.e., Probabilistic, Radial Basis Function, and Multi-Layer Perceptron - MLP - see – e.g., Haykin, 1994). They concluded that the MLP ANN showed the best capability to correlate the state variables' values to the WDS section states. They then used MLP ANNs to build a system that analyses the data from the pressure and flow sensors in order to determine the location and size of a leak. The resulting system has a two level architecture composed of a main ANN at the first

level and several branch specific second level ANNs in cascade to the main one. The WDS section's branch in which the leak occurs is identified by the main ANN, while a specific second level ANN is activated to estimate the location and size of the leak in the selected branch. The authors followed this study with methodology tests on a real district heating system where no water out-take, except demonstrated leaks, occurred (Caputo and Pelagagge, 2003). These tests showed positive results as the system was always able to identify where the leaks occurred.

Shinozuka et al. (2005) proposed an approach to determine the location and extent of the damage of a leak caused by an earthquake. Their approach only makes use of pressure measurement at some selected positions in the WDS section. The ANN employed in their study was trained on data originating from a HM simulation of the WDS section (i.e., pressure variation at the monitoring stations as inputs and the location of the leak and "Damage Index" as outputs). The trained ANN was then used to provide the Euclidean distance (see – e.g., Deza and Deza, 2009) from a suspected leak location to a monitoring station and an indication of the extent of the damage. The approach was demonstrated on a simple synthetic case study having only one location of damage and three monitoring stations. It was found to be sufficiently effective for the purpose of damage identification. It has to be stressed, however, that the use of the Euclidean distance might have been suitable for the WDS section being studied, which was shaped as a rectangular grid with only two different pipe lengths, but could be inappropriate in more realistic situations.

Mashford et al. (2009) used SVMs for mining the data obtained by a collection of pressure sensors monitoring a WDS section and obtaining information about the location and size of leaks. They trained two SVMs on synthetic noise-free data obtained by using the EPANET2 software (Rossman, 2000). Performance assessment of their SVM method showed that the leak size and location were both predicted with a reasonable degree of accuracy. However, the assumptions of complete knowledge of the WDS section without uncertainties and availability of accurate pressure measurements make the study unfeasible for practical use.

In view of the above, it is worth mentioning that several issues have not been taken into consideration by the researchers that have proposed techniques that make use of the

simplified approach for estimating the current state of a WDS section. These include assessing the susceptibility to noise and measurement errors, evaluating the maximum size of a WDS section and its complexity to define the applicability limits of the particular technique, and determining the optimal number and placement of sensors to minimise the costs while maintaining a required level of performance, among the others. These issues will have to be investigated if such techniques are to be applied to real-life WDS sections.

2.4.7.3 Measurement-based techniques

The methods described in this section monitor and analyse in near real-time the process variable (e.g., pressure and flow) signals from a SCADA system to detect the leak-induced changes in these signals. Analysis can be simple, such as visually detecting abnormally low/high measurements or comparing a single measurement (i.e., flow or pressure at a single location in the WDS section) to a threshold limit. Alternatively, analysis might be extensive, involving the use of techniques from the statistical and AI fields to more reliably determine when the measurements are decreasing/increasing in a significant manner.

Generally speaking, a leak will cause an increase in the flow and a decrease in the pressure upstream of the leak, therefore it will affect the normal pressure/flow signal fluctuations that are observed at a sensor located upstream of the leak. Identification of these abnormal changes can therefore be used to indicate the occurrence of a leak. As mentioned above, in the simplest application of this principle (sometimes referred to as Pressure/Flow Trending), the current and recent historical pressure and/or flow trends are displayed in a graphical format on a monitor. These enable a controller to be cognisant of these parameters' fluctuations and infer the occurrence of a leak if significant abnormal changes are observed. Another application of such principle checks whether the pressure/flow measurements exceed a predetermined alarm threshold. This threshold is usually set out of the process variable's normal operating fluctuations range. If the threshold is exceeded an alarm is raised. Note that SPC techniques (e.g., Control Charts – see Appendix D) are sometimes used to define the alarm threshold from the historical data for a selected time window. Other similar methods include the Rate-of-Change monitoring, and the Pressure/Flow Point Analysis. Rate-of-Change monitoring calculates the variation in a process variable with respect to

a defined time interval. The rate at which the process variable changes is then used to identify the occurrence of a leak. Pressure/Flow Point Analysis, on the other hand, correlates the averaging of one or more process parameters over short and long time intervals in order to detect a leak. Specifically, a buffer of the most recent values of pressure/flow variable is divided into two parts and the statistical properties (e.g., mean and standard deviation) of the two parts are calculated and compared. If the mean of the newer portion of data is significantly smaller/bigger than the mean of the older portion, for example, a leak can be suspected. Note that to increase the sensitivity of the leak detection, a longer buffer of data can be used. Consequently, however, the response time to a leak will also increase. Note that the techniques described so far are commonly used to monitor oil and gas pipelines. A good review of these and other techniques that have been applied to pipelines in the oil and gas industries can be found in Whaley et al. (1992) and ADEC (2011). The primary advantage of these methods is their low complexity. However, although in a WDS they might be effective for single transmission pipelines (which are immune from pressure/flow variations due to demand fluctuations), they are unreliable for monitoring pressure and flow in DMAs.

Bearing in mind the above, the simplest technique for DMAs is the MNF monitoring. This technique involves measuring the MNF rate in a DMA at discrete time intervals (e.g., daily, weekly, etc.). Leaks are suspected if the MNF rate is greater than previously measured rates in the same DMA. The MNF monitoring is probably the most commonly used leak detection technique in practice. This is because it represents a simple approach to a complex problem, thus making it very attractive. Furthermore, it has proven its efficacy in the UK and in many other countries (e.g., UKWIR, 1999b; Pilcher, 2002; Valdés and Castelló, 2003; Sturm and Thornton, 2005). As a result, many water companies in the UK devote considerable manpower to analysing the MNF data. However, by using this technique, the changes in the MNF rates can only be detected the next day (in the best case scenario) and not in near real-time. This, coupled with the fact that the data analysis is often done manually, by using non-standard approaches, and at irregular time intervals, often results in a backlog of past events that need to be checked. Therefore, advanced statistical and AI techniques have been proposed for more efficiently and effectively analysing the data from the DMA sensors.

Palau et al. (2003, 2012) proposed the use of the Principal Component Analysis for analysing the hourly inflows into a DMA. In their work the Principal Component Analysis was used to simplify the original set of flow data recorded by a SCADA system, attempting to synthesise the most significant information into a statistical model. Control Charts based on suitable statistics (i.e., Distance to Model, and T^2 Hotelling – see e.g., Jackson, 1991) where then used for the detection of leaks and other anomalous behaviours regarding water use (e.g., illegal connections). On a single example with historical data from a real-life DMA, the authors showed that their technique offered higher detection sensitivity to leaks (or other unexpected consumptions) than conventional univariate statistical methods. It has to be stressed, however, that the authors only considered flow measurements collected during the night period (i.e., from midnight to 7:00 a.m.) in their analysis. This implies that, similarly to the MNF monitoring technique a leak can be detected (at best) the next day only. The authors stated that their approach could be extended to analyse data collected during the day. Due to the much greater demand fluctuations, however, the legitimacy of this statement cannot be established without further relevant tests.

Aksela et al. (2009) presented a method for leak detection based on the use of a Self-Organising Map ANN (see – e.g., Kohonen, 1990, 2001). The data used for training and validating the test results consisted of vectors of the flow meter readings and knowledge about reported leaks. Because of the great variance in behaviour between different days of the week, the authors used daily flow averages and carried out their analysis on a weekly basis. Despite the results of the experiments presented by the authors show that this method could be used for successfully detecting the leaks in a DMA, it is important to stress that the frequency at which the data analysis is performed makes it unsuitable for near real-time applications.

Mounce and Machell (2006) presented a study on using ANNs for classifying the signature of the flow and pressure signals under the normal and abnormal conditions of a DMA. Engineered Events (EEs), where leaks were simulated by using fire hydrants, were carried out to provide the one minute pressure and flow data used for the ANN training and testing. Two ANN architectures (i.e., Static and Time-Delay – see e.g., Haykin, 1994) were tested. The results obtained show that a Time-Delay ANN could learn the patterns of the simulated leaks. The 75% of the leaks simulated during the test

period were detected. Despite this study proved that ANNs can be used effectively for performing leak detection in a DMA, the authors noted that obtaining accurate targets for the classification task is difficult, unless EEs are carried out. Indeed, records of real leaks are not easy to obtain or may be inaccurate. Even where records do exist, the commencement time for a leak is not known (just the date on which the leak is reported or repaired) and the size is difficult to determine. It can be therefore concluded that the effectiveness of this methods and of the method proposed by Aksela et al. (2009) (although less remarkably being it based on the use of an unsupervised learning paradigm) is strongly dependent upon: (i) the availability of sufficient exemplars (i.e., leak records) which should also be representative of the entire feature space (i.e., all the possible leaks), and (ii) the quality of those exemplars.

Mounce et al. (2002), on the other hand, proposed the use of a Mixture Density ANN (Bishop, 1994) to predict future values of a flow signal coming from a DMA sensor. The predictions generated by the ANN were then used, together with the actual observed flow data, to form a comparative which, in turn, was analysed by a classification module. The classification module indicated abnormalities in its output (based on a modifiable error sensitivity level) by using a binary leak-no-leak indicator for analysis of windows of readings. Subsequently, Mounce et al. (2006) replaced their classification module with a Fuzzy Inference System such that confidence intervals could be assigned to the detected leak event. Mounce et al. (2007) further improved their leak detection methodology to provide accurate estimates of the average leak flow. In all these studies the methodology proposed by the authors was tested by using the historical flow data and data on repairs carried out in the DMAs and/or by using the data collected during several EEs. Later on, Mounce et al. (2008, 2010a) and Mounce and Boxall (2010) presented applications of their methodology in a near real-time environment and demonstrated its performance by checking the genuineness of the raised alarms against various sources of information (i.e., record of customer calls database, record of repairs database, ad hoc liaison with the water company personnel, etc.). Without doubt, the methodology developed by the authors represents an effective and viable means for on-line leak detection and assessment in DMAs as it provides an efficient and consistent means for the analysis of large data volumes. The main strength of this approach is that the ANNs only rely on the observed time series data. Therefore

it enables future flow values to be predicted solely on the basis of the observed past data without explicitly knowing the actual DMA dynamics. This, in turn, allows implementing a data driven approach that continuously and automatically learn the flow profiles and adapt to the DMA changes. However, the time window (i.e., 12 hours or 24 hours) used by the authors for leak detection is rather wide. This can be seen as a significant drawback to their approach. Indeed, late detection of a leak reduces the response time a water company might gain to locate and fix a leak before it affects the customers.

In order to reduce the leak detection time, Mounce et al. (2011a) proposed the use of SVMs. In this study, it was demonstrated by examples that the SVM methodology can provide faster alarm generation than the previously developed ANN/Fuzzy Inference System methodology (i.e., 8 hours earlier). The methodology was tested with historical pressure and flow data from five typical UK DMAs and the authors reported that the 22% of the raised alarms were false alarms (i.e., ghosts). This is a higher false alarms figure than the one reported in previous studies by the authors. For example, an 18% false alarms figure was reported in Mounce et al. (2010b). This fact raises concerns about the reliability of the SVMs methodology. This study, however, is of particular interest because data from the DMA pressure sensors were also used to perform leak detection. Although current research is investigating the use of pressure data collected from the DMA sensors for leak identification (e.g., Wu et al., 2010 - see Section 2.4.4), little work has been done in utilising these data in the context of the statistical/AI measurement-based techniques. The authors found that less leak detections resulted from the analysis of the historical pressure data. The authors also stated that the majority of ghosts resulted from the analysis of the pressure data, perhaps due to the less stationary nature of these data.

Observations about the usefulness of using pressure data have also been reported by Fenner and Ye (2011). They presented a method that makes use of an Adaptive Kalman Filter for analysing the DMA flow and pressure measurements. In their method, the Adaptive Kalman Filter was used to model the normal flow and pressure fluctuations, so that the residual of the filter (i.e., the difference between the predicted flow/pressure and the measured flow/pressure) provides an indication of the abnormal flow/pressure variations relating to the leaks. The methodology was tested by using EEs and historical

data from ten UK DMAs. Similarly to Mounce et al. (2011a), the authors stated that the results obtained suggest that the flow measurement data are more sensitive to a leak than the pressure measurement data. However, they also reported that the alarms generated by using the pressure data coincided well with the leak events.

With regard to the issue raised in Mounce et al. (2011a) and Fenner and Ye (2011) about the fewer detections resulted from the data analysis when the pressure signals were considered, the following has to be noted. It is well known that the effect of a leak will not always be observed in the measurements from a DMA pressure sensor (e.g., a service reservoir can support the pressure in a DMA even if a leak has occurred). Also, the sensitivity of a pressure signal to a leak depends on the location of the sensor. For example, a pressure sensor which is far from the leak location is unlikely to show a significant change from its normal profile. All this may explain the observations made by the authors. Despite its lesser sensitivity in terms of leak detection, however, the simultaneous analysis of multiple signals (i.e., from both flow and pressure sensors within a DMA) gives the opportunity to gain additional information about the occurrence of a leak. This, in turn, leads to more reliable and timely detections (as it will be shown in this thesis).

With regard to the issue raised in Mounce et al. (2011a) about the many false alarms resulted from the analysis of the pressure signals, the following has to be noted. Since the pressure measurements are usually instantaneous values, they tend to be noisy. Filtering/de-noising techniques could therefore be used, before training the ANNs/SVMs, to circumvent/alleviate the problem arising as a consequence of the less stationary nature of the pressure data (as it is done in the novel methodology presented in this thesis). In this regard, noteworthy is the work presented in Mounce et al. (2012). In that study, the authors investigated the benefit of averaging the pressure measurements over the sampling interval (similarly to what is commonly done for the flow measurements). They showed that this is a useful strategy for leak detection. However, averaging the pressure measurements may lose potentially useful information in the data. Therefore, it is necessary to explore alternative and more efficient de-noising techniques for improving the pressure measurements' "usability".

Given that the deployment of an increased number of pressure sensors in a single DMA is becoming more common, the analysis of multiple pressure signals can also provide useful information about the leak location as well. In this regard, it is worth mentioning that methodologies are being developed to identify the optimal placement of the pressure sensors to capture the leak effect no matter where in a DMA the leak occurs (Farley et al., 2008, 2010a, 2010b; Mounce et al., 2010c). These techniques have the potential to allow obtaining more reliable leak detections results and also enable performing approximate leak location within a DMA. Thus, their further validation and development are envisaged.

Bearing in mind the above, it has to be further noted that pressure and flow sensors are not the only source of information for performing near real-time leak detection by using statistical/AI measurement-based techniques. Turbidity sensors have been used too (Khan et al. 2002, 2005, 2006). Therefore, given that water quality sensors may further enhance the detection performance of these techniques and help performing approximate leak location within a DMA, advancements in their development are also envisaged.

2.4.7.4 Concluding remarks

A number of papers dealing with the statistical/AI-based techniques have been reviewed in this section. The techniques considered here have been grouped into two main categories, namely model-based and measurement-based techniques. Considerations about the practicality and usefulness of applying the techniques in each of these groups for leak identification in real-life WDS sections are reported below.

Using model-based techniques, leak identification is performed by correlating measured changes in hydraulic characteristics and changes in the analytical model of a WDS section. For single pipelines, this might not be a serious issue. Given the complexity of the problem for distribution pipeline systems, however, this procedure often suffers from the lack of available measurement data and measurement noise/errors. Therefore, many of the model-based approaches presented in the literature have been tested on single pipelines or on simple pipeline systems assuming ideal noise-free conditions. A further issue that would strongly affect the effectiveness of these techniques when applied to real-life WDS sections is the validity and correctness of the analytical model

(even if this is exclusively used to produce training data for AI data analysis tools such as ANNs and SVMs). Variations in the topography, WDS section geometry and roughness coefficients, among the others, would require accurate representation. Differences between the analytical model behaviour and the actual WDS section behaviour will affect the leak identification results.

The measurement-based techniques that adopt a pattern recognition methodology whereby a sensors output is treated as a signal and is analysed by techniques from the statistical and AI fields, on the other hand, are an effective and viable mean for enabling on-line leak detection and assessment in both real-life transmission pipelines and distribution pipeline systems. This is mainly because they only rely on the signals' data and hence on the empirical observation of the behaviour of the WDS section. Also, these techniques provide a consistent way of analysing the data by automating mundane tasks involved in the data analysis process. They can efficiently deal with the vast amount of often imperfect sensor data collected by modern SCADA systems, extract information useful for making reliable operational decisions, and present more "intelligent" information to the WDS operators. When distribution pipeline systems are considered, however, the following has to be noted. Despite all the aforementioned advantages, these techniques can be improved in terms of leak detection reliability and speed. Furthermore, they present an important limitation. Indeed, existing statistical/AI measurement-based techniques allow the discovery of a leak in a particular DMA without giving any information about its more precise location. Therefore, their further development in terms of enabling determining the likely location of a leak within a DMA is envisaged as well.

2.4.8 Section summary

Several hydraulic techniques have been described and reviewed in this section. These techniques have different limitations, capabilities, costs, and benefits. The performance and the feasibility of each considered technique are mainly associated with its capability to: (i) allow inspection or monitoring of a WDS section, (ii) perform leak identification in transmission pipelines only or in both transmission pipeline and distribution pipeline system, (iii) allow leak assessment, detection, location or relevant combinations (e.g., detection and location), and (iv) respond quickly to a leak. In this context, Table 4 summarises the main characteristics of the techniques that have been reviewed in this

section. Table 5, Table 6 and Table 7 summarise their individual strengths and weaknesses.

Based on what reported in the aforementioned tables and on the review carried out in this section it is clear that, at the present, hydraulic techniques do not achieve the high performance in terms of leak location accuracy of the hardware-based techniques. Indeed, among the reviewed hydraulic techniques, only the transient analysis-based and negative pressure wave-based techniques have the potential to allow exact leak location. In this context, the appeal of the transient analysis-based and negative pressure wave-based techniques arise a consequence of the greater energy and superior propagation distances of a low-frequency transient wave over an acoustic leak signal. Such characteristics make the propagation of pressure transients less sensitive to the characteristics of the pipes (e.g., material and diameter) and the feasibility of using such techniques less dependent on the sensor-to-sensor spacing. However, these techniques have had limited success so far and are in no position to supplant the techniques that make use of specialised hardware devices.

Bearing in mind the above, it has to be stressed that the key advantage of using the hydraulic techniques is that they may offer an inexpensive means to help the water companies significantly reducing the leaks' runtime. Indeed, bearing in mind that the exclusive use of the hardware-based techniques for inspecting entire WDSs or DMAs would require significant efforts in terms of man-hour needed, the hydraulic techniques can improve the situation by: (1) making the water company personnel aware of the presence/occurrence of a leak much faster, and (2) guiding the water company personnel straight to the problem area (i.e., narrow down the area to be searched). This, in turn, may enable reducing the time before starting the actual leak search by using the hardware devices and also reducing the time taken to find and repair the leak (i.e., improving the exact leak location efficiency and effectiveness).

In the scenario outlined above, it is clear that the hydraulic technique should be evaluated based on the extent they enable doing what stated in each of the two aforementioned points and hence improving the current situation. It is also clear that, among the reviewed hydraulic techniques, the statistical/AI measurement-based techniques that automatically analyse the hydraulic data transmitted by a SCADA

system to detect unusual changes in the process variable patterns are found to be the most promising in terms of enabling to proactively manage the leakage in WDSs. Indeed, despite the fact that the existing techniques in this group do not have any leak location capability, they are the only one holding the promise of an efficient and low time consuming means for the automated on-line leak identification in real-life WDS sections. Hydraulic techniques such as water audits, step tests, steady state analysis-based techniques and transient analysis-based techniques, on the one hand, entail intermittent inspections in the field, thereby making the leak detection time completely dependent on the inspection frequency. Negative pressure wave-based and statistical/AI model-based techniques, on the other hand, have shown little practical applicability to real-life situations so far.

Table 4. Main characteristics of the hydraulic techniques.

	Available/ R&D	**Time to react to a leak**	**Cost**	**Inspection/ Monitoring**	**Assessment/Detection/ Location (Approx. – Exact)**	**Pipelines/ Pipeline systems**
Water audits - ***System-wide***	Available	Very high	High	Inspection	Assessment	Pipelines/ Pipeline systems
Water audits - ***District***	Available	Very high	Very High	Inspection	Assessment	Pipelines/ Pipeline systems
Step tests	Available	High	High	Inspection	Assessment/Detection/ Location (Approx.)	Pipelines/ Pipeline systems
Steady state analysis-based techniques	R&D	High	Medium	Inspection	Assessment/Detection/ Location (Approx.)	Pipelines/ Pipeline systems
Transient analysis-based - ***ITA techniques***	R&D	High	High	Inspection	Assessment/Detection/ Location (Approx. - Exact)	Pipelines/ Pipeline systems
Transient analysis-based - ***Frequency domain techniques***	R&D	High	High	Inspection	Assessment/Detection/ Location (Approx. - Exact)	Pipelines
Transient analysis-based - ***Direct transient techniques***	R&D	High	High	Inspection	Assessment/Detection/ Location (Approx. - Exact)	Pipelines
Negative pressure wave-based techniques	R&D	Low	High	Monitoring	Assessment/Detection/ Location (Approx. – Exact)	Pipelines/ Pipeline systems
Statistical/AI techniques - ***model-based***	R&D	Low	Low	Monitoring	Assessment/Detection/ Location (Approx.)	Pipelines/ Pipeline systems
Statistical/AI techniques - ***measurement-based***	Available/ R&D	Medium/ Low	Medium/ Low	Monitoring	Assessment/Detection	Pipelines/ Pipeline systems

Table 5. Main strengths and weaknesses of the hydraulic techniques (Part 1 of 3).

	Pros	**Cons**
Water audits ***- System-wide***	• Provide an overall picture of the water losses in a WDS • Do not require any field work	• The resulting water losses figure is a rough estimate of its true value only • Carried out sporadically
Water audits ***- District***	• Enable identifying DMAs with the highest leakage levels • Enable identifying issues that prevent a water company from achieving high water delivery efficiency	• Rely on the accurate estimation of several night flow consumption components • Low sensitivity to leakages • Require 'out-of-hours' work • Carried out sporadically
Step tests	• Can narrow down the location of a leak within a DMA	• Rely on the accurate estimation of several night flow consumption components • The shutting down of various DMA subsections can cause backsiphonage and the risk of infiltration of ground water • Create supply interruption inconvenience to the customers • Require 'out-of-hours' work
Steady state analysis-based techniques	• Can narrow down the location of a leak within a DMA • Do not require high frequency measurements	• Rely on an accurate steady state HM • Require accurate measurements at multiple locations • Applicable to single pipelines or simple pipeline systems only • Involve manual and resource intensive processes for data logging, collection, and transfer to the point of use • Often require long computation times

Table 6. Main strengths and weaknesses of the hydraulic techniques (Part 2 of 3).

	Pros	**Cons**
Transient analysis-based - *ITA techniques*	• Can narrow down the location of a leak within a simple pipeline system or exactly locate a leak in a single pipeline • Have the potential to distinguish between leak and non-leak related sources of reflections, dispersion and damping	• Rely on an accurate transient HM • Require accurate high frequency measurements at multiple locations • Affected by errors in the boundary conditions • Involve manual and resource intensive processes for data logging, collection, and transfer to the point of use • Often require long computation times • Applicable to single pipelines or simple pipeline systems only
Transient analysis-based - *Frequency domain techniques*	• Can exactly locate a leak in a single pipeline • Require high frequency measurements at only one location (simpler and more economic data collection process, less affected by measurements with limited accuracy)	• Rely on an accurate transient HM • Require accurate high frequency measurements • Involve manual and resource intensive processes for data logging, collection, and transfer to the point of use • Influenced by the location of the transient source and measurement station • Require identification and interpretation of the non-leak related sources of reflections, dispersion and damping • Applicable to single pipelines only
Transient analysis-based - *Direct transient techniques*	• Can exactly locate a leak in a single pipeline • Require high frequency measurements at only one location (simpler and more economic data collection process, less affected by measurements with limited accuracy)	• Rely on an accurate transient HM or on the existence of a measured leak-free response • Require accurate high frequency measurements • Involve manual and resource intensive processes for data logging, collection, and transfer to the point of use • Often require a sharp transient wave that may be dangerous for a pipeline • Often rely on visual identification of the leak-induced transient reflections • Require identification and interpretation of the non-leak related sources of reflections, dispersion and damping • Applicable to single pipelines only

Table 7. Main strengths and weaknesses of the hydraulic techniques (Part 3 of 3).

	Pros	**Cons**
Negative pressure wave-based techniques	• Can effectively enable the on-line monitoring of a single pipeline or simple pipeline system • Can narrow down the location of a leak within a simple pipeline system or exactly locate a leak in a single pipeline	• Often rely on an accurate transient HM • Require accurate high frequency measurements at multiple locations • Have high data transmission costs • Have to identify very weak leak-induced transients with unknown shapes • Affected by the noise from the normal WDS operations • Applicable to single pipelines or simple pipeline systems only
Statistical/AI techniques **- *model-based***	• Can effectively enable the on-line monitoring of a single pipeline or simple pipeline system • Can narrow down the location of a leak within a DMA • Do not require high frequency measurements	• Rely on an accurate steady state HM • Require accurate measurements at multiple locations • Applicable to single pipelines or simple pipeline systems only
Statistical/AI techniques **- *measurement-based***	• Can effectively enable the on-line monitoring of a real-life pipeline or pipeline system • Do not require high frequency measurements • Only rely only on the empirical observation of the WDS section behaviour • Provide an efficient and consistent means for the analysis of large volumes of imperfect data • Can adapt to the changes in the WDS section's operating conditions	• Allow the discovery of a leak in a particular DMA without giving any information about its more precise location

2.5 Summary and Conclusions

In this chapter a review of literature focussing on leak identification in WDSs has been carried out. After the introduction in Section 2.1, a brief overview of the leakage management activities has been given in Section 2.2. Then, several hardware-based and hydraulic leak identification techniques have been reviewed in Section 2.3 and in Section 2.4, respectively. The considered hardware-based techniques have included acoustic equipment-based techniques, tracer gas-based techniques, infrared thermographic techniques, GPR-based techniques, and PIG-based techniques. The considered hydraulic techniques have included water audits, step tests, steady state analysis-based techniques, transient analysis-based techniques, negative pressure wave-based techniques, and statistical/AI-based techniques. For each of the reviewed techniques, a brief description of the principles on which it is based and the way it works has been provided, its main capabilities and limitations have been identified, the state-of-the-art in its development has been highlighted, and its potential for enabling the water companies minimising the leaks' runtime has been evaluated. This, in turn, enables us making several considerations about the gaps in the current research. These considerations together with the main conclusions that can be drawn from the literature review carried out in this chapter are detailed below.

None of the existing leak identification techniques is perfect and hence can be used on its own. In this scenario, an appropriate combination of techniques has to be used. It is well known that when selecting the most appropriate combination of leak identification techniques for a particular water company or scenario, a number of economical, operational, environmental, social and legal factors have to be considered (Hunaidi, 2005). The specific characteristics of a WDS section (e.g., age of the pipes, leaks frequency, etc.) and the water company operations (e.g., cost of water, availability of trained and skilled staff, pressure management strategies, leak identification budget, water quality concerns, renewal strategies, etc.), along with other social and environmental factors typically determine the mix of techniques' applicability and feasibility. Independently from all this, however, it is clear that to reduce the water losses from a WDS and to improve a water company's operational efficiency and customer service, an appropriate combination of techniques that allows a fast response leak identification protocol is required.

Currently, the best practical protocol for leak identification in WDSs is the combined application of water audits/MNF monitoring and step tests/hardware-based techniques. Water audits enable the water companies identifying the DMAs that are experiencing the highest leakage levels and hence allow them to efficiently target the resources for leak detection and location. They, however, rely on the accurate estimation of several consumption components that contribute to the night flow. As more cities and towns move towards 24-hour activity, the estimation of these components will become significantly more difficult. Furthermore, water audits are expensive in terms of man-hour needed and therefore they are carried out sporadically. MNF monitoring, on the other hand, has the potential to assess the DMAs' leakage level at much shorter time intervals. The data analysis required, however, tends to be a manual process. Human analysis of the data and interpretation of the results are unfeasible for ever larger volumes of imperfect data. As a result, the MNF analysis is often carried out using non-standard approaches and irregularly. Whether water audits or MNF monitoring is used, if it is established that a DMA warrant attention, the next stage involves manual leak detection and location activities which are carried out by using step tests and/or hardware-based techniques. Step tests require night work and the shutting down of DMA subsections. For these reasons, in the UK during the 1990s, step tests began to be replaced by the use of acoustic equipment-based and other hardware-based techniques. When it comes to the hardware-based techniques, significant technological advancements have been made in the past decades. These advancements have made them achieve high performance in terms of leak detection and location accuracy. The availability of more diverse and improved hardware devices has also enabled the water companies to adopt a "multi-sensor" approach whereby the whole range of hardware devices is considered and an appropriate mix of equipment for the specific WDS section characteristics, site locations and types of leak is selected. Bearing this in mind, further developments of these devices are envisaged, with special regard to key technologies such as the correlating noise loggers and the GPRs. Despite competition in the market, however, the use of the hardware-based techniques still remains expensive in terms of equipment renting or owning. Furthermore, it is time consuming and relatively expensive in terms of man-hour needed to carry out the survey of entire DMAs.

Given the current leak identification protocol described above, it is clear that there is the need for a numerical (i.e., inexpensive) technique that enables: (1) making a water company aware of the occurrence of a leak as soon as it occurs, thereby minimising the time before starting the actual leak search using the hardware devices, and (2) restricting the area to be searched, thereby allowing to not only optimise the use of the hardware devices but also give the leak inspectors higher motivation and hence improve the exact leak location efficiency and effectiveness. To this end, transient analysis-based, steady state analysis-based, negative pressure wave-based and statistical/AI-based techniques have been proposed.

Transient analysis-based and steady state analysis-based techniques have the potential to significantly restrict the leak search area. The former techniques, however, often rely on complex and inaccurate transient HMs and require precise knowledge of the WDS section parameters along with precise measurements. As a result, they had limited success so far, generally on single pipelines only. Similarly, the latter techniques are reliant on the availability of accurate steady state HMs and require an extensive number of precise measurements. This significantly limits their widespread use by the water companies. In addition to this, it has to be noted that both transient analysis-based and steady state analysis-based techniques entail intermittent WDS section inspections in the field. This implies that the leak detection time will be completely dependent on the frequency of the inspections. Furthermore, they also involve manual and resource intensive processes for data logging, collection, and transfer to the point of use. These processes together with the actual time necessary for performing the required data analyses may take weeks.

Opposite of the above, negative pressure wave-based and statistical/AI-based techniques can provide a rapid (i.e., in near real-time) response to a leak occurring. This is because they automatically analyse the data from permanently installed sensors. The sensors required by the former techniques, however, are quite expensive to buy and maintain. Furthermore, as the data are recorded at very high frequencies, the transmission costs are significant as well. Economic issues aside, it has to be stressed that the negative pressure wave-based techniques present several of the limitations of the transient analysis-based techniques. Additionally, they have to cope with the unknown shape of a leak-induced transient, noise from the normal WDS section operations and often very

weak leak signals. All this makes the negative pressure wave-based techniques difficult to apply to large WDS sections with complicated configurations. When it comes to the statistical/AI-based techniques, a distinction has to be made between model-based and measurement-based techniques. Similarly to the other techniques that make use of a steady state or transient HM of a WDS section, the statistical/AI model-based techniques have shown little practical applicability to real-life situation so far. Statistical/AI measurement-based techniques that automatically detect unusual changes in the process variable patterns, on the other hand, are found to be the most promising techniques in the context of on-line leak identification in real-life WDSs. Indeed, these techniques enable extracting useful information (required for making reliable operational decisions) from the vast and often imperfect sensor data collected by modern SCADA systems. They do not have a requirement for expensive high frequency measurements and, most importantly, they only rely on the empirical observation of the behaviour of a WDS section. Existing statistical/AI measurement-based techniques, however, do not have leak location capabilities. Furthermore, they can be improved in terms of reducing the leak detection time while improving the leak detection reliability.

Having recognised that the statistical/AI measurement-based techniques hold the promise of an efficient and low time consuming means for the automated on-line leak identification in real-life WDSs, it is possible to state that there is scope for introducing a novel data analysis methodology that improves the existing techniques in this group by enabling: (i) the near real-time detection of leaks in a more reliable and timely fashion, and (ii) the successful approximate location of leaks within a DMA. The more reliable and timely detection of leaks together with the successful identification of their approximate location, which can be achieved by using this novel data analysis methodology, may facilitate prompt interventions and repairs. This, in turn, may enable the water companies to save water, improve their operational efficiency and, most importantly, their customer service by reducing the supply interruption inconvenience.

In view of the above, the development and testing of the aforementioned novel data analysis methodology are presented in the following chapters. In particular, the next chapter will presents the methodology's theoretical background and implementation details.

CHAPTER 3 EVENT DETECTION & LOCATION METHODOLOGY

3.1 Introduction

As a consequence of the latest developments in hydraulic sensor technology and on-line data acquisition systems, the cost of pressure and flow monitoring devices has reached the level that enables their large scale deployment in a WDS (Kapelan et al., 2005). As discussed in Chapter 2, the UK WDSs are typically subdivided into DMAs to more effectively manage the water leakages. Nowadays, the UK DMAs are usually observed by using pressure and flow sensors located at the DMA entry/import/export points and a pressure sensor located at the critical DMA point. The data streams (i.e., signals) from these sensors provide a potentially useful source of information for identifying pipe burst events both quickly and economically. As water companies recognise this fact more and more, and that several other important benefits are yielded by the monitoring of their DMAs in near real-time (e.g., improved network visibility and management, higher compliance with regulatory targets, etc.), an increase in the density of coverage of pressure/flow monitoring locations is expected in the near future.

Given the outlined current status and trend in availability of pressure and/or flow data collected and communicated on-line by the sensors deployed in the UK DMAs, this chapter presents a novel pipe burst event identification methodology that makes use of these data. Specifically, if the DMA being studied is observed by using a number of sensors as the one currently used in the UK practice (i.e., normally-sampled DMA), the methodology presented here enables: (i) detecting the pipe burst events at the DMA level and as they occur (i.e., in near real-time), and (ii) assessing their size. If the DMA being studied is observed by using a larger (than currently used in the UK practice) number of sensors (i.e., over-sampled DMA), then the methodology also enables determining the approximate location of the pipe burst events within the DMA. Note that the methodology presented here also allows the identification of other events that induce abnormal pressure/flow variations similar to those induced by a pipe burst event, and that are often of interest for the water companies (e.g., illegal connections, unexpected water usages, etc.). In the light of this and for the sake of simplicity,

hereafter, "event" will be used as a generic term to indicate both pipe burst events and other aforementioned events. Accordingly, ERS will be used to denote the computer-based early warning system that implements the novel data analysis methodology presented here. Taking this into account, it can be noted that for some of the event types that can be identified by using the methodology presented here it does not make sense to talk about assessing their size (e.g., sensor failures). Therefore, the use of the term "event assessment" will be generally avoided (although the methodology's capability to estimate the size of pipe bursts and other relevant events will be implicitly assumed when discussing event detection and location).

Bearing in mind the above, this chapter is organised as follows. After this introduction, the philosophy behind the development of the novel event detection and location methodology is discussed in Section 3.2. Then a high level overview of the data processing route for performing near real-time event detection and location by using the novel methodology presented here is given in Section 3.3. Once this is done, Section 3.4 provides the theoretical background and the methodological details of the data analyses that enable performing event detection in normally-sampled DMAs. Section 3.5 provides the theoretical background and the methodological details of the data analyses that enable performing event detection and location in over-sampled DMAs. Finally, a summary of the chapter and the main conclusions are presented in Section 3.6.

3.2 Event Detection & Location Philosophy

The pressure and flow signals from the sensors deployed in a DMA show daily, weekly and seasonal variations and are also influenced by socioeconomic and meteorological factors such as population characteristics or number of industrial establishment and air temperature or precipitation. Furthermore, they show a different behaviour under the DMA normal and abnormal (e.g., when a pipe burst event has occurred) operating conditions. However, their pattern is specific to a particular location. During the DMA normal operating conditions, this means that a signal may show a predictable fingerprint. It should be therefore quite feasible, during the DMA abnormal operating conditions, to identify the event-induced pressure/flow deviations from this predictable fingerprint. Note that, hereafter, the acronym NOP will be used to refer to the pattern of a signal expected during the normal operating conditions of a DMA.

In the light of the above, it is possible to state that a qualified and experienced human operator can (in general) decide if an event has occurred in a DMA based on the available data from the hydraulic sensors deployed in it. This is mainly because the human operator is able to "capture" (i.e., learn/estimate) the NOPs of the DMA signals and hence identify, in the signal patterns observed during the DMA abnormal operating conditions, the event-induced deviations from the "captured" DMA signal NOPs. In addition to these abilities, however, other "human qualities" may help him/her achieve the above goal. For example, thanks to his/her ability to ponder, the operator will be more confident that an event has occurred if the scale of an identified deviation is significant and/or if multiple deviations are identified in a DMA signal. Similarly, thanks to his/her ability to form relationships, the operator's confidence that an event has occurred will be higher if deviations are simultaneously identified in more than one DMA signal. The persistence of the identified deviations (i.e., deviations identified in successive time steps) is also an important factor that the operator will take into consideration in order to make a decision about the actual occurrence of an event. In addition to this, the human operator is also able to (continuously) improve the reliability of his/her own event occurrence decisions. Indeed, by observing how past events of different type and size, which occurred in different areas of the DMA, affected the DMA signals, he/she (will be) is able to review and fine-tune his/her own event occurrence decision mechanisms. Finally, as the human operator is aware that the response of a particular sensor to an event is dependent on the event location, he/she can use this information to determine the area within the DMA where the event most likely occurred.

The idea (i.e., philosophy) behind the development of the novel event detection and location methodology, which (as it will be shown in this chapter) is implemented in a computer-based ERS, is to mimic the behaviour of a qualified and experienced human operator. This is motivated by the fact that, despite the above stated human operator's ability to decide if an event has occurred in a DMA and to determine its approximate location, the volume and complexity of the received SCADA data often exceeds the human capability to analyse, interpret, and extract useful information for making reliable operational decisions. Thus, the advantages of an automated data analysis

methodology that simultaneously and synergistically analyses all the signals coming on-line from the sensors deployed in a DMA are evident.

In order to pursue the above idea, the novel event detection and location methodology presented in this chapter makes use of a number of advanced data-driven (geo)statistical and AI techniques which include, among the others, a BIS for inferring the probability of an event occurrence at the DMA level (i.e., DMA level BIS). As an ensemble, these techniques enable the novel event detection and location methodology performing the following five basic actions:

1. To "capture" the DMA signal NOPs;
2. To identify and estimate, in the DMA signal patterns observed when an event has occurred, the event-induced deviations from the "captured" DMA signal NOPs;
3. To infer the probability that an event has actually occurred in the DMA based on the identified deviations;
4. To (re)calibrate the parameters of the DMA level BIS based on the information about past events occurred in the DMA being studied;
5. To determine the approximate location of an event within the DMA based on the analysis of the extent to which each DMA sensors has been affected by the event occurring.

The first three basic actions aim at mimicking the process that leads the human operator to make a reliable decision about the occurrence of an event. The fourth basic action aims at mimicking the process that leads the human operator to (continuously) improve the reliability of his/her own event occurrence decisions. The fifth basic action aims at mimicking the process that leads the human operator to determine where an event most likely occurred.

In the light of the automated data analysis framework outlined above, it is important to stress that the fourth basic action is not 'strictly' necessary for the actual event detection and location. Furthermore, it has a requirement for information (i.e., start time and duration) about the past events occurred in the DMA being studied, and it also involves a certain degree of human interaction (as the past events information has to be checked for reliability before being used for the DMA level BIS calibration/recalibration). For

these reasons the fourth basic action may not be performed if the past events information is unavailable, or not taken into consideration (e.g., to avoid any human interaction with the ERS).

3.3 Event Detection & Location Data Processing Route

As mentioned in Section 3.2, the novel event detection and location methodology presented in this chapter is based on the simultaneous analysis of all the (pressure/flow) signals coming on-line from the sensors deployed in a DMA. This methodology is implemented in an automated ERS in the form of a piece of software which can be installed on a computer located in the control room of a water company. Bearing this in mind, this section gives an overview of the data processing route that could/should be followed for effectively enabling performing near real-time event detection and location by using the novel methodology behind the ERS. Before this, however, an assumption about the DMA sensors that have to be used to satisfy the methodology's on-line data requirement is made below.

The aforementioned assumption is that the DMA sensors are equipped with a Global System for Mobile communications (GSM) modem capable of individual General Packet Radio Service (GPRS) data connectivity (i.e., newest UK best practice for leading water companies). The use of GPRS allows the data collected at regular time intervals (e.g., 15 minute) to be communicated to the water companies without incurring a per-call charge that is a characteristic of other communication technologies (e.g., GSM calls, Short Message Service - SMS - messages). The data from each sensor may be communicated at longer intervals (e.g., every 30 minutes) to increase the sensor battery life. Despite this assumption, it has to be stressed that the novel event detection and location methodology is able to work even if the sensor data are communicated at much longer (e.g., every 24 hours) intervals. This said, however, the above assumption serves the purpose of showing that the communication technology already employed as best practice by the leading water utilities in the UK, may be beneficially used together with the methodology presented here for shortening the detection time of the events occurring. That is to say, for enabling their detection shortly after they occur and, possibly, to intervene before the event impacts are noticed and reported by the customers.

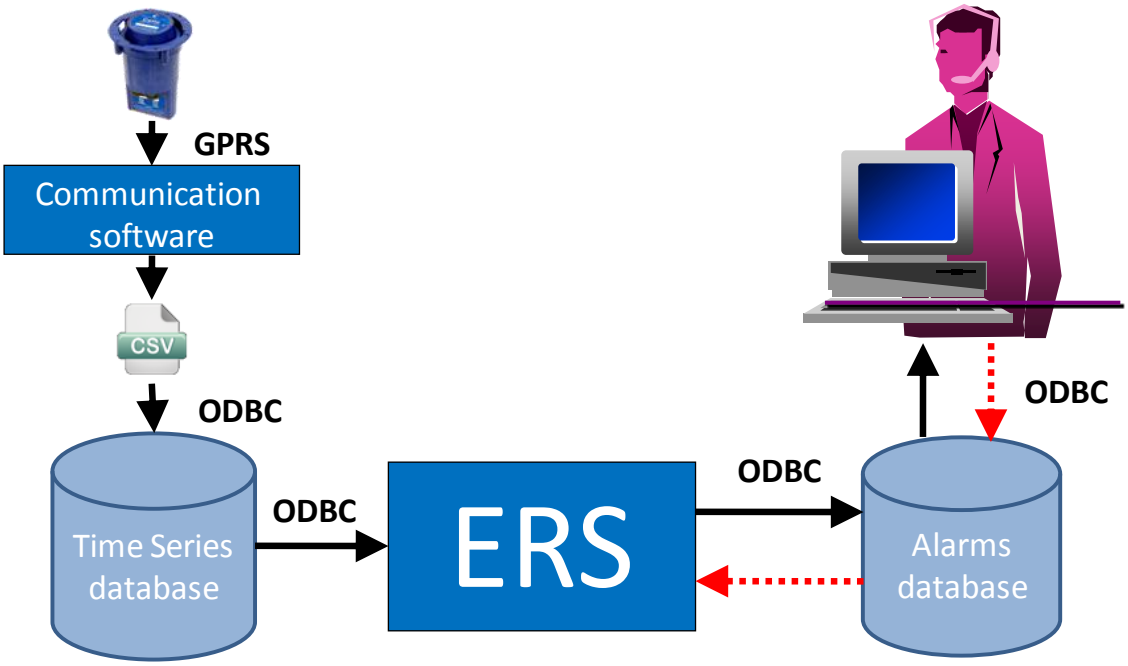


Figure 1. Schematic representation of the data processing route for performing near real-time event detection and location by using the novel methodology behind the Event Recognition System.

Bearing in mind the above, Figure 1 provides a schematic representation of the data processing route for performing near real-time event detection and location by using the novel methodology behind the ERS. On the one hand, the solid black arrows show the data processing route for performing the first, second, third and fifth basic actions described in Section 3.2. It starts from the sensor units which initiate calls to the communication software at regular time intervals. For each DMA signal and at each communication interval *u* readings are obtained (e.g., 2 readings – assuming 15 minute sampled data, and data communicated every 30 minutes). Export functionality in the communication software is used to automatically update a set of Comma Separated Values (CSV) files. An Open Database Connectivity (ODBC) text driver is used to interface to the Time Series database storing the time series history of the DMA signals. Once the data from all the DMA signals are fully processed as described in Sections 3.4 and 3.5, the resulting *u* probability values that an event has occurred in the DMA and the additional ERS output information useful for enabling a diagnosis of the event occurring (i.e., to determine the approximate event location and the likely cause of an alarm) are stored in the Alarms database. If any of the *u* probability values exceed a user defined detection threshold an alarm is generated and the user is notified. Following the

generation of an alarm, the additional output information provided by the ERS is processed further to determine the approximate location of the event occurring. On the other hand, the dashed red arrows show the data processing route for performing the fourth basic action described in Section 3.2. It forms a feedback loop which gives the user the possibility to store, in the Alarms database, information about past events occurred in the DMA being studied. This information could be obtained from the water company's historical records. Alternatively, it could be obtained from a periodic review of the alarms raised by the ERS. This review has the potential to minimise the user interaction with the ERS by enabling him/her to simply flag the genuine alarms (which have an associated event start time) as confirmed and check the associated event durations. The stored past events information may then be used to (re)calibrate the parameters of the DMA level BIS, thereby enabling (continuously) improving the reliability of the raised alarms.

3.4 Event Detection in Normally-Sampled District Metered Areas

3.4.1 Section overview

Given the current status in availability of pressure and/or flow data from the sensors deployed in the UK DMAs (see Section 3.1), the readily available on-line data can be used to develop a novel methodology for the automated detection of events at the DMA level and as they occur. This section presents such a methodology. That is to say, it presents a novel methodology for the near real-time detection of events at the DMA level that is based on the simultaneous and synergistic analysis of all the (pressure/flow) signals coming on-line from the small number of sensors that are nowadays usually deployed in a UK DMA. This methodology makes use of different AI and statistical techniques and does not require a hydraulic or any other physically based simulation model of the DMA.

This section is organised as follows. An overview of the computer-based ERS is presented first in Section 3.4.2. In that section, it is shown that the novel methodology is implemented in the ERS by making use of six main subsystems. In view of this, Sections 3.4.3 to 3.4.8 describe the theoretical background and the methodological details of the data analyses performed in each of the six ERS subsystems. Specifically,

Section 3.4.3 presents the data analyses for the first ERS subsystem, namely the *Setup* subsystem. Section 3.4.4 presents the data analyses for the second ERS subsystem, namely the *Discrepancy Based Analysis (DBA)* subsystem. Section 3.4.5 presents the data analyses for the third ERS subsystem, namely the *Boundary Based Analysis (BBA)* subsystem. Section 3.4.6 presents the data analyses for the fourth ERS subsystem, namely the *Trend Based Analysis (TBA)* subsystem. Section 3.4.7 presents the data analyses for the fifth ERS subsystem, namely the *Inference* subsystem. Section 3.4.8 presents the data analyses for the sixth ERS subsystem, namely the *BIS parameters learning* subsystem. Once this is done, a brief discussion about the main limitations of the ERS methodology is presented in Section 3.4.9. Finally, a summary of the section is given in Section 3.4.10.

Bearing in mind the above, it has to be stressed that the description of the relevant data analysis routines developed is organised in terms of the ERS subsystems (and associated modules) to make it easier for the reader to identify the function and "location" of each of these data analysis routines in the complex methodology developed and presented here.

3.4.2 Event Recognition System overview

Figure 2 shows a diagrammatic representation of the ERS. As it can be observed from this figure, the first four basis actions described in Section 3.2 (i.e., shown as dotted dashed rectangles) are performed in the ERS by making use of six subsystems (i.e., shown as solid snipped corner rectangles) each containing a number of different modules (i.e., shown as solid rectangles). The six ERS subsystems are as follows: (1) the *Setup* subsystem, (2) the *DBA* subsystem, (3) the *BBA* subsystem, (4) the *TBA* subsystem, (5) the *Inference* subsystem, and (6) the *BIS parameters learning* subsystem.

The first ERS subsystem is used to perform the first basic action mentioned is Section 3.2 (i.e., signal pattern capturing). The second, third and fourth ERS subsystems are used together (i.e., synergistically) to perform the second basic action mentioned in Section 3.2 (i.e., deviations identification/estimation). The reason for using three analysis subsystems is that, by making use of different ensembles of statistical/AI techniques, each of them focuses on recognising a specific type of evidence that an event has occurred. Furthermore, since they perform tasks in parallel they allow

simultaneously looking at how an event affects the pressure/flow measurements from different perspectives (e.g., short-term and long-term effects). All this enables the three analysis subsystems to complement each other by: (a) providing more conclusive evidence of an event occurrence - different analysis subsystems may independently identify the deviations caused by a particular event, and (b) enabling the identification of the deviations caused by different event types - each analysis subsystem looks at a particular deviation indicator that is more suitable than the others for identifying the deviations caused by a particular event type (e.g., the *DBA* and *BBA* subsystems are more suitable for sudden pipe burst events while the *TBA* subsystem is more suitable for gradually developing events such as pipe bursts developing gradually from background leaks). The fifth ERS subsystem is used to perform the third basic action mentioned in Section 3.2 (i.e., event occurrence probability inference). Finally, the sixth ERS subsystem is used to perform the fourth basic action mentioned in Section 3.2 (i.e., calibration/recalibration of the DMA level BIS).

Figure 2 also shows that the ERS has three main modes of operation: (1) the "Assemble" mode, (2) the "Execute" mode, and (3) the "Learn" mode. These modes of operation define if and when the data analyses in each ERS subsystem are performed. Specifically, the "Assemble" mode is used for 'tuning' the data-driven ERS when it is initialised (i.e., used for the first time in a DMA). Later on, it is used: (i) regularly (e.g., weekly) when the ERS is updated (to capture the latest normal operating conditions of a DMA) thereby providing a continuously adaptive ERS, and (ii) periodically (e.g., every three months) when the ERS is reinitialised (to account for the seasonal variations in the DMA's pressure/flow regime, growing demand over time, etc.; or following occasional operational/other DMA changes - e.g., re-valving). The "Execute" mode is the normal operating mode used at every communication interval to detect the events occurring and raise the alarms. Finally, the "Learn" mode may be used for the initial calibration and/or for the follow-on periodic recalibration of the DMA level BIS. Note however that, as explained in Section 3.2, its actual utilisation depends on whether or not the past events information is available/considered.

Details about the ERS subsystems and the associated modules can be found in the following six sections.

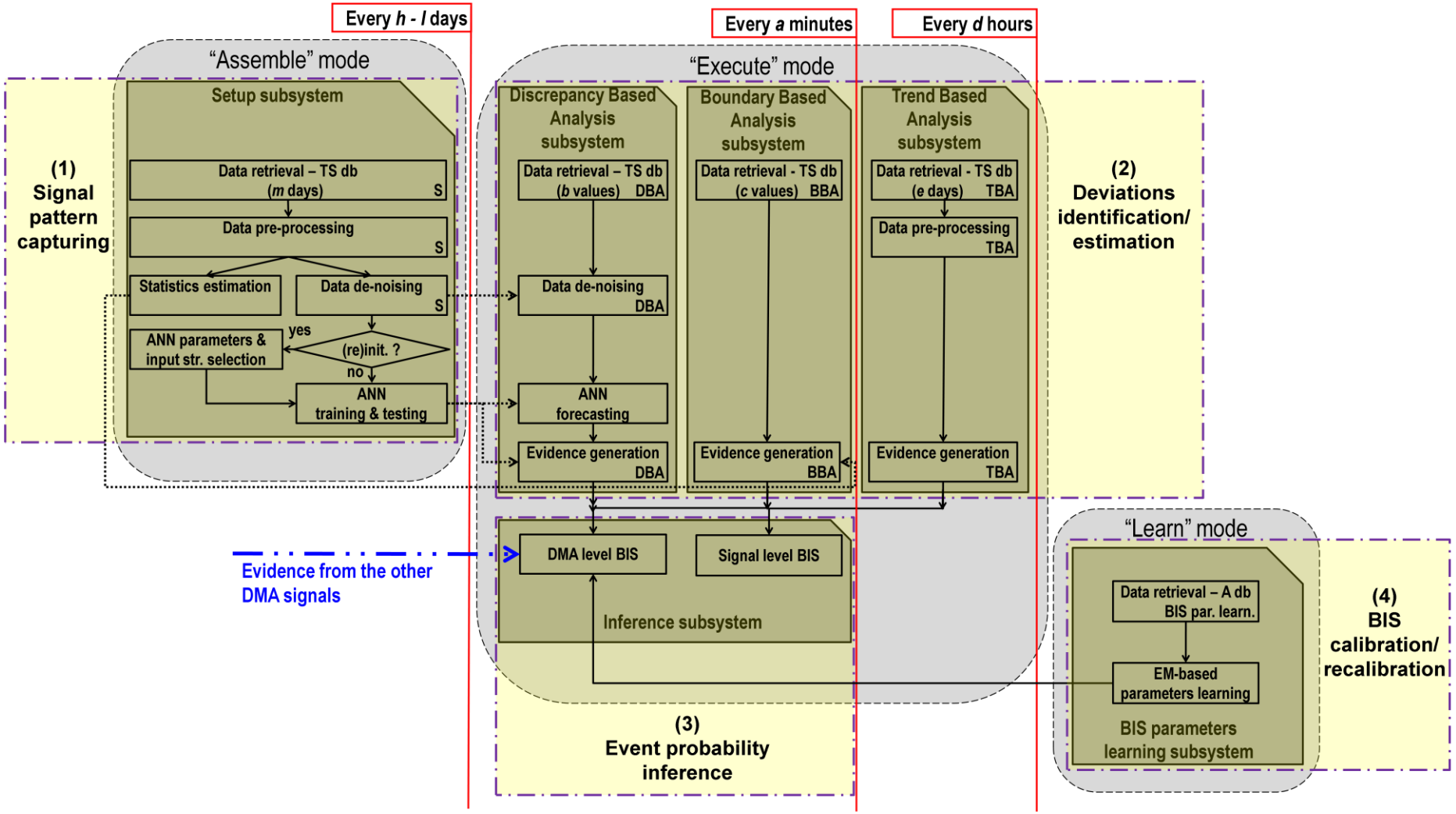


Figure 2. Diagrammatic representation of the Event Recognition System.

3.4.3 Setup subsystem

3.4.3.1 Subsystem overview

The *Setup* subsystem constitutes the initial processing stage of the novel methodology that enables performing event detection in normally-sampled DMAs. In this ERS subsystem the historical data for the specific DMA signal being analysed undergo a series of analyses to allow the ERS to "capture" different types of information (i.e., statistically-based and model-based) about the NOP of that signal.

Figure 3 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are organised into six modules. These six modules are as follows: (1) data retrieval module, (2) data pre-processing module, (3) statistics estimation module, (4) data de-noising module, (5) ANN parameters & input structure selection module, and (6) ANN training & testing module. The first module is used for retrieving from the Time Series database the historical data for the specific DMA signal being analysed. The second module is used for assembling a set of pressure/flow data that best represents the most recent NOP of that DMA signal (i.e., 'NOP data set'). Once this is done, the third module is used for estimating (starting from the 'NOP data set') several vectors of descriptive statistics (i.e., averages and standard deviations). These vectors provide basic statistical information about the DMA signal NOP. The remaining modules are used for first removing noise from the 'NOP data set' and then for: (i) selecting the optimal (i.e., that yield the best forecasting performance) ANN input structure and parameters set, (ii) training and testing a signal-specific ANN prediction model (by using the optimised ANN input structure and parameters set), and (iii) estimating the ANN model prediction error's variability. Note that since the resulting ANN model assumes that no event occurred in the DMA, it provides a model-based type of information about the DMA signal NOP.

As it can be observed from Figure 2, this subsystem runs in the "Assemble" mode when the ERS is initialised, reinitialised (i.e., every h days - e.g., every 90 days) and updated (i.e., every l days - e.g., every 7 days).

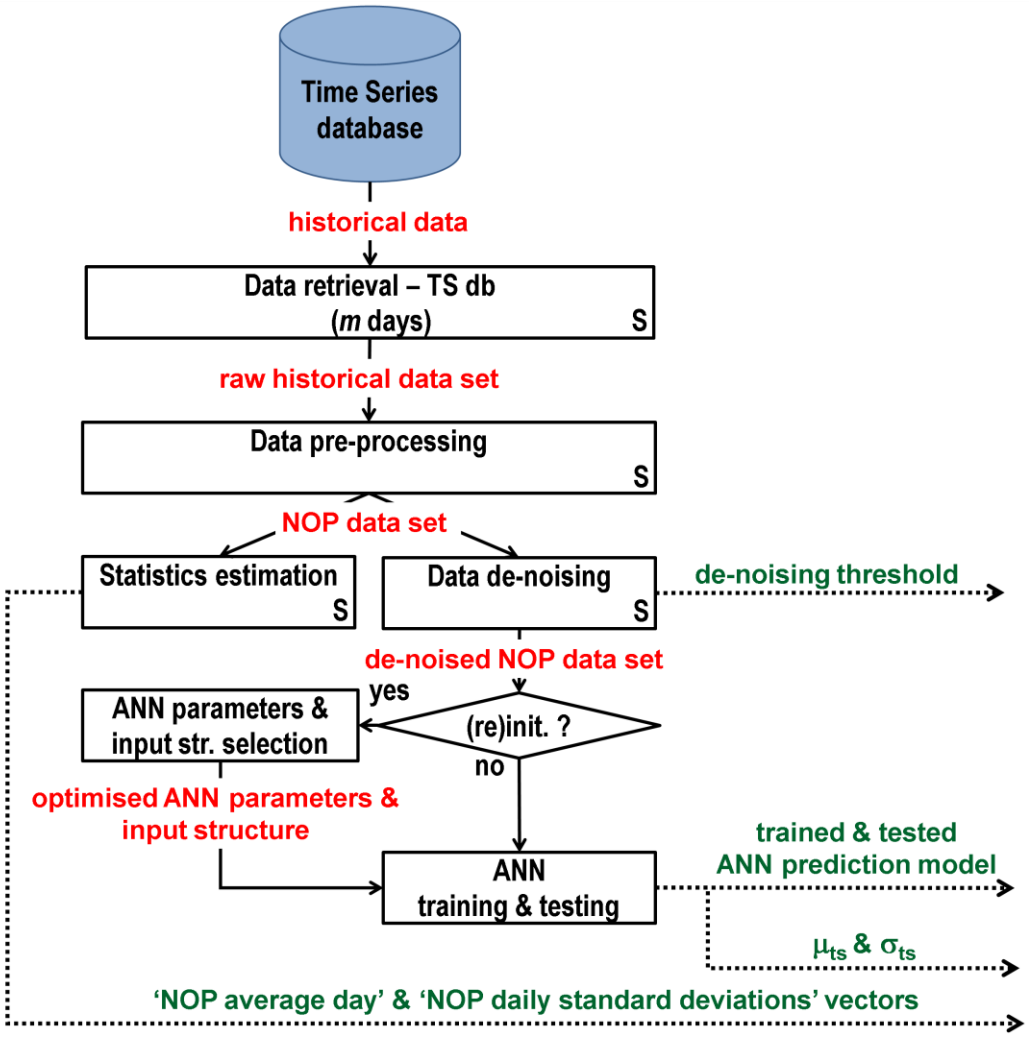


Figure 3. Processing of data in the *Setup* subsystem.

3.4.3.2 Data retrieval module

The objective of this module is to retrieve, for the DMA signal being analysed, *m* days (e.g., 14 days) of past raw data (i.e., 'raw historical data set') from the Time Series database. Note that in order to obtain *m* 'full' days of past raw data (e.g., from midnight to 23:45 – if a 15 minute sampling rate is considered) the ERS is bounded to retrieve past raw data starting from the DMA signal value measured at the time step immediately before midnight on that day. The *m*-day 'raw historical data set' is then passed on to this subsystem's data pre-processing module for further processing (see Figure 2 and Figure 3).

3.4.3.3 Data pre-processing module

The objectives of this module are: (a) to assemble (starting from the 'raw historical data set') a uniformly spaced time series without missing values (i.e., 'repaired historical data set'), and (b) to assemble (starting from the 'repaired historical data set') a set of pressure/flow data that best represents the most recent NOP of the DMA signal being analysed (i.e., 'NOP data set'), These objectives are achieved by: (i) identifying and repairing erroneous time stamps and applying a heuristics-based procedure for dealing with the missing values, and (ii) performing in sequence three statistical tests aimed at gradually filtering out outliers and measurements that are not consistent with the expected DMA signal pressure/flow variations assuming that no event occurred in the DMA. Note that, while the 'repaired historical data set' is used in this module only, the assembled 'NOP data set' is passed on to the statistics estimation module and to the data de-noising module of this subsystem for further processing (see Figure 2 and Figure 3).

Assembling the 'repaired historical data set'

According to Fayyad et al. (1996) about 80% of the resources in statistical and AI applications are spent on cleaning and pre-processing the data. In the light of this and given the "dirty" nature of the pressure/flow data from the sensors deployed in a DMA (i.e., erroneous timestamps, large chunks of missing data, etc.), effectively cleaning and pre-processing the retrieved raw data is a key issue for achieving efficient event detection in the ERS methodology presented here.

Starting from the *m*-day 'raw historical data set', the ERS checks and corrects erroneous timestamps, creates a uniformly spaced time series, replaces blank entries with missing value indicators (i.e., NAN – Not A Number) and associates to each measured value its Time of the Day (TofD) (i.e., a value between 1 and *g*, where 1 corresponds to midnight and *g* is the number of samples in one day – e.g., g=96 if a 15 minute sampling rate is considered) and its Day of the Week (DofW) (i.e., a value between 1 and 7, where 1 corresponds to Monday). The resulting *m*-day time series is then rearranged into *m* vectors (i.e., one vector for each day with *g* pressure/flow/NANs values in each vector). Once this is done, the ERS uses a heuristics-based procedure to discard vectors containing large chunks of missing data (i.e., the linear interpolation would result in unrealistic pressure/flow daily trends) and uses the linear interpolation to fill in the

missing values in each of the remaining vectors (i.e., 'valid' days). As a result, an *n*-day 'repaired historical data set' is obtained.

The heuristics-based procedure used here discards a vector based on the number v_1 of continuous periods of missing data values (i.e., two or more missing values in consecutive time steps) in it, and the number w_1 of missing data values in each identified continuous period of missing data values. This procedure was identified after relevant sensitivity analysis (not shown here). The sensitivity analysis conducted showed that the absence of up to about 25% of data did not affect the shape of the pressure/flow daily trend significantly when this heuristics-based procedure was applied. Note that the term significantly is used here to indicate that the Mean Absolute Percentage Error (see e.g., Hyndman and Koehler, 2006) between an observed pressure/flow daily trend and the same daily trend with interpolated data values, after fictitiously removing a certain percentage of data, is less than 1%. The sensitivity analysis conducted also suggested that, although the actual v_1 and w_1 that are used by the ERS depend on the sampling rate of the measurements, the following heuristic rule could be used as a reference: maximum 6 continuous periods of missing data of maximum 1 hour (e.g., v_1=6 and w_1=4 – if a 15 minute sampling rate is considered).

Assembling the 'NOP data set'

After filling in the missing data, the 'repaired historical data set' is processed further by means of three statistical tests.

The *first statistical test* involves two steps. In the first step, considering that each vector (i.e., one day of data) in the *n*-day 'repaired historical data set' can be represented as $(x_{1j}, x_{2j},\ldots, x_{ij})$ where $i = 1{:}\,g$ and $j = 1{:}\,n$, a vector of 'daily' averages $(\bar{x}_1, \bar{x}_2,\ldots, \bar{x}_i)$ is computed using the following equation:

$$\bar{x}_i = \frac{\sum_{j=1}^{n} x_{ij}}{n} \tag{3.1}$$

Similarly, a vector of 'daily' standard deviations $(\sigma_1, \sigma_2,\ldots, \sigma_i)$ is computed using the following equation:

$$\sigma_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2} \tag{3.2}$$

In the second step, the ERS checks that every x_{ij} value falls inside the interval $(\bar{x}_i - N_{l1^{st}}\sigma_i,\ \bar{x}_i + N_{u1^{st}}\sigma_i)$, where $N_{l1^{st}}$ and $N_{u1^{st}}$ are user defined multipliers for the 1^{st} statistical test, denoting the acceptable lower and upper control limits (e.g., $N_{l1^{st}} = N_{u1^{st}} = 3$ – i.e., those values outside the 99.7% confidence interval assuming a normal distribution). Note that these user defined multipliers, as well as the other multipliers that will be described below and in the following sections, can take different values depending on the particular DMA signal being analysed (i.e., pressure vs. flow). This fact represents the main reason for treating the pressure and flow measurements as independent variables. Indeed, by doing this the most appropriate parameters for the particular signal being analysed can be used during the various data analysis stages. If x_{ij} is outside these limits the entire j vector is discarded. As a result the number of 'valid' days may be reduced from *n* to *p*.

The *second statistical test* aims to assess how the mean of the remaining p 'valid' days changes over time. It involves SPC-based control charting (see – e.g., Shewhart, 1931; Deming, 1950, 1975). A Control Chart is a graphical representation of descriptive statistics that can be used to study how a process variable changes over time and hence to track unusual variations in a process. Note that further information about Control Charts is given in Appendix D Section D.2. A mean μ_k, where $k = 1{:}\,p$, is calculated for each of the remaining p 'valid' days of data. The Control Chart plots the p values μ_k in time order, a centre line at the average of the means $\bar{\mu}_k$, and upper and lower control limits at a user defined number $N_{l2^{nd}}$ and $N_{u2^{nd}}$ of standard deviations from the centre line. The standard deviation is estimated by taking the average of the p 'valid' days standard deviations. If μ_k is an outlier of the range defined by the upper and lower control limits, the k^{th} day is defined as 'out of control' (Breyfogle, 1999) and the k^{th} vector removed. After applying this statistical test, the number of 'valid' days may be reduced further from *p* to *q*.

The *third statistical test* reapplies the first statistical test, but this time considering the remaining *q* 'valid' days and new user defined multipliers for the acceptable lower and

upper control limits $N_{l3^{rd}}$ and $N_{u3^{rd}}$, respectively. Once the third statistical test has been performed, the ERS is potentially left with a dataset consisting of *r* 'valid' days (i.e., 'NOP data set'), which should now best represent the NOP of the DMA signal being analysed.

Alternatively, if a 'repaired historical data set' containing a significant number of 'valid' days (i.e., $n \geq f$ – where *f* is a user defined ERS parameter – e.g., $n \geq 90$ days) can be assembled, the ERS groups the *n* 'valid' days according to their relevant DofW before performing the three statistical tests described above. In this case, the statistical tests are applied separately to each DofW-group. The *r*-day 'NOP data set' is then put together by using the 'valid' days left in each DofW-group. Note that this procedure allows better accountability (from a statistical point of view) of the differences between the various DofWs (e.g., a weekday is likely to show a different pattern from a day during the weekend).

The reason for using the two alternative approaches described above is that each of those approaches has its own advantages and disadvantages. For example, with regard to the latter approach, the number of 'valid' days *n* in the 'repaired historical data set' has to be significant in order to provide the ERS with a reasonable (i.e., statistically relevant) number of days representing a particular DofW. This implies that the number of days of past raw data to be retrieved from the Time Series database for analysis (i.e., *m* days) has to be significant as well. This requirement is not ideal when (re)initialisation of the ERS is required (e.g., when the ERS is used to analyse a DMA where the pressure/flow sensors have just been installed, following a DMA reconfiguration, etc.) because the ERS will have to remain inoperative until the necessary number of days of past raw data is stored in the Time Series database.

Figure 4 shows an example of how the sequential use of the three statistical tests described above allows removing days containing measurements that are not consistent with the expected DMA signal pressure/flow variations assuming that no event occurred in the DMA being studied (i.e., red dotted, green dashed, and purple dotted dashed daily trends). The figure refers to the case where $n < f$ and the user defined multipliers for the acceptable lower and upper control limits were set as equal to 3 for all the statistical tests.

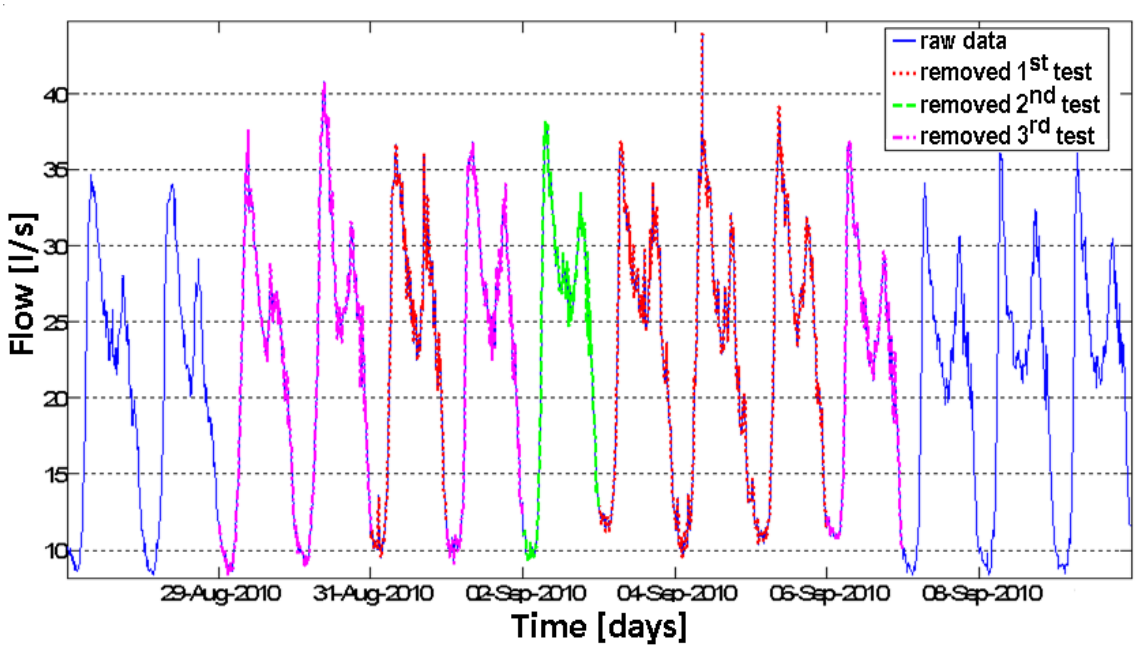


Figure 4. Example of the results obtained after applying the statistical tests for assembling the 'Normal Operating Pattern data set'.

3.4.3.4 Statistics estimation module

The objective of this module is to "capture" basic statistical information about the NOP of the DMA signal being analysed. By using Equations (3.1) and (3.2) and depending on the number of 'valid' days, *n*, in the 'repaired historical data set': (i) one 'NOP average day' vector and one 'NOP daily standard deviations' vector are calculated based on the *r*-day 'NOP data set' (i.e., $n < f$), or (ii) seven (i.e., one for each DofW-group) 'NOP average day' vectors and seven 'NOP daily standard deviations' vectors are calculated based on the 'valid' days left in each DofW-group after the three statistical tests have been performed (i.e., $n \geq f$).

The estimated 'NOP average day' and 'NOP daily standard deviations' vectors are then passed on to the evidence generation module of the *BBA* subsystem (see the fourth dotted arrow from the top in Figure 2). These vectors will be repeatedly used in the "Execute" mode for establishing the boundaries within which the DMA signal data coming on-line from the relevant pressure/flow sensor should lie assuming that no event has occurred in the DMA (see Section 3.4.5.3).

3.4.3.5 Data de-noising module

The objective of this module is to create a 'de-noised NOP data set'. This is achieved by removing noise from the assembled 'NOP data set' using a wavelet-based de-noising procedure (Donoho, 1995; Donoho and Johnstone, 1994, 1995). The main reason for this is that, as shown in Romano et al. (2010a) and in Chapter 5, it improves the prediction accuracy of the ANN prediction model (especially for the pressure signals). The 'de-noised NOP data set' is then passed on to the ANN parameters & input structure selection module (i.e., if the *Setup* subsystem is used when the ERS is initialised/reinitialised) or directly on to the ANN training & testing module (i.e., if the *Setup* subsystem is used when the ERS is updated) - (see Figure 2 and Figure 3).

A wavelet-based de-noising procedure is used in the ERS methodology because in the de-noising of non-stationary signals (such as those coming from a DMA), it has been found to be very successful (e.g., Nason and Sachs, 1999; Percival 2000). Also, the de-noising of a DMA signal can be performed without compromising its non-stationary or transitory characteristics, which are of particular importance in anomaly detection (especially when the de-noising of the data coming on-line is considered – see Section 3.4.4.3). This is because, WA allows signals to be studied using a dual frequency-time representation which, in turn, allows noise to be removed from signal frequencies that are likely to contain important information whilst isolating and preserving the characteristics of the events occurring (which might be smothered or hidden away in the noisy pressure/flow signals resulting from the legitimate consumer demand). Note that further details about the general wavelet theory and the wavelet-based de-noising procedure used in this module and in the data de-noising module of the *DBA* subsystem are given in Appendix A.

The wavelet-based de-noising procedure used here is as follows:

1. Use the Daubechies wavelet of the fourth order (Daubechies, 1988) for performing a one level Discrete Wavelet Transform (DWT) of the original signal (Mallat, 1989). This results in two sub-signals: (i) approximation, which refers to the high-scale (i.e., low-frequency) component of the original signal, and (ii) detail, which refers to the low-scale (i.e., high-frequency) component of the original signal;

2. Calculate the Universal Threshold (Donoho and Johnstone, 1994, 1995), and apply it to the detail sub-signal's coefficients by using the Soft Thresholding method (Donoho, 1995);
3. Perform Inverse Discrete Wavelet Transform (IDWT) using the original approximation sub-signal's coefficients and the modified detail sub-signal's coefficients.

Figure 5 shows an example of the results obtained after applying the first step of the procedure outlined above for the de-noising of a DMA flow signal. It is possible to observe from this figure how the original signal (Figure 5a) is decomposed into an approximation sub-signal (Figure 5b) and a detail sub-signal (Figure 5c), which contains most of the high-frequency noise.

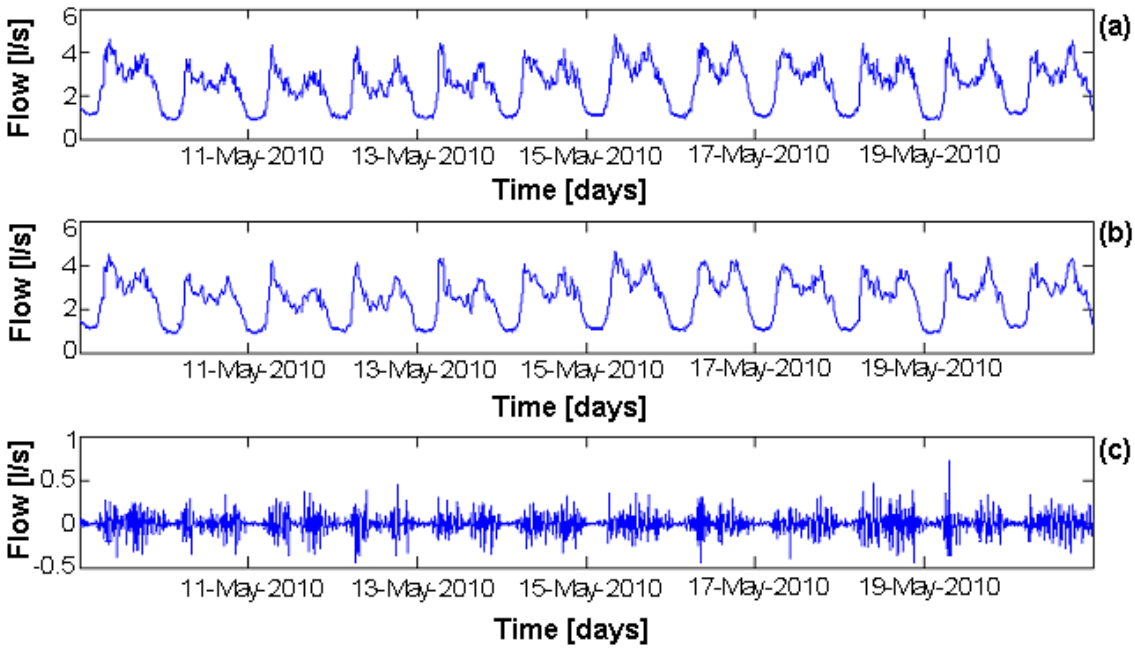


Figure 5. One level Discrete Wavelet Transform of a District Metered Area flow signal. Original signal (a), approximation sub-signal (b), and detail sub-signal (c).

It is important to stress that although the wavelet-based de-noising of signals is theoretically effective and widely applied in various engineering activities and applied research fields (see – e.g., Ruskai, 1992; Jaffard et al., 2001), in practice the procedure

used for implementing it is influenced by several key factors. These include the choice of: (i) the mother wavelet, (ii) the decomposition level, (iii) the thresholding method, and (iv) the method for estimating the threshold value. Many studies have been conducted in various fields for addressing these issues (e.g., Torrence and Compo, 1998; Schaefli et al., 2007; Sang et al., 2009, 2010a, 2010b; Jansen, 2006). However, most of the quantitative methods proposed are not applicable to the de-noising of DMA signals. This is because it is not possible to identify the 'real signals' in observed DMA signals. Thus, heuristic judgement is required for selecting a suitable de-noising procedure. Taking this into consideration, the choice of the particular wavelet-based de-noising procedure outlined above has been supported by the results obtained after tests on several DMA signals (results not shown here) and the following theoretical considerations (see also Appendix A).

With regard to the mother wavelet, in order to have a unique reconstructed signal from IDWT, orthogonal wavelets have to be selected. Choosing a mother wavelet that has compact support in both time and frequency is also important for the complexity of the computation. Additionally, the number of vanishing moments increases the computational complexity of the DWT. Hence, for on-line applications, a wavelet with a high number of vanishing moments cannot be suggested. However, a wavelet with more vanishing moments provide better reconstruction quality, introduce less distortion into the processed signals, and concentrate more signal energy in few coefficients. In view of these considerations, the selection of a wavelet that belongs to the Daubechies family represents a good choice. This is because the Daubechies wavelets not only are orthogonal but have also the property of being the most compactly supported wavelets with a certain number of vanishing moments (Daubechies, 1988). When wavelets in this family are considered, the Daubechies wavelet of the fourth order, which has four vanishing moments and is supported over eight time points, represents a good compromise between computational simplicity and reconstruction quality. Furthermore, it has been suggested that the lower orders of the Daubechies family (such as the fourth order) are suitable for non-stationary signals while the higher orders (such as the twentieth order) are suitable for stationary signals (Abdullah et al., 2008).

With regard to the decomposition level, many levels of decomposition brings with it complex computation. Furthermore, if the number of levels is more than required, the

event information from the signal might be lost, whilst, if it is less than required, the noise might not be eliminated completely. Taking this into account and bearing in mind that the main purpose of the de-noising procedure used here is to improve the forecasting performance of the ANN prediction model, the choice of only one level of decomposition is justified by the following fact. It allows removing "enough" noise from a DMA signal (for achieving the stated goal) without the risk of losing the important event information. The faster computation is also an important factor for enabling near real-time event detection.

With regard to the thresholding method, it must be suitable for the characteristics of an analysed signal. In view of this, for the pressure/flow signals from a DMA, the Soft Thresholding method produces better results than the Hard Thresholding method, as it does not introduce discontinuities in the reconstructed signal (Donoho and Johnstone, 1994, 1995).

Finally, with regard to the threshold, the Universal Threshold has many advantages - as reported in Donoho and Johnstone (1994, 1995). Furthermore, its value is based on an estimate of the noise level from the pressure/flow data in the 'NOP data set' (see Appendix A). This allows taking into consideration the most recent DMA normal operating conditions at every ERS initialisation/updating/reinitialisation.

The Universal Threshold computed and used in this module is then passed on to the data de-noising module of the *DBA* subsystem (see the first dotted arrow from the top in Figure 2) where it will be repeatedly used in the "Execute" mode for removing noise from the DMA signal data coming on-line from the relevant pressure/flow sensor (see Section 3.4.4.3).

3.4.3.6 Artificial Neural Network parameters & input structure selection module

The objective of this module is as follows. For each DMA signal being analysed, to automatically select the ANN input structure (e.g., number of past pressure/flow values) and set of parameters (e.g., number of hidden units, number of training cycles), that, when used for developing the relevant short-term ANN prediction model, enables it to yield the best forecasting performance. The main reason for this is that, with a view to the application of the novel ERS methodology to entire WDSs, signals from different DMA types (e.g., rural, residential, etc.) and different DMA signal types (e.g., pressure

vs. flow) will have to be analysed. As these signals may show extremely varying patterns, the use of a pre-defined ANN input structure and parameters set (e.g., the same for all the analysed DMA signals) may lead to developing ANN prediction models that exhibit sub-optimal forecasting performance. Taking this into consideration, the potential benefits resulting from the use of the approach proposed here are two-fold. On the one hand, the quality of the ANN models' predictions improves. On the other hand, as the resulting ANN models are signal-specific, the ERS becomes "tailored" to the particular DMA to which it is applied, whilst more generally applicable to different DMAs. In view of this, the aforementioned module's objective is achieved here by using an EA-based optimisation strategy. The selected ANN input structure and parameters set is then passed on to the ANN training & testing module (see Figure 2 and Figure 3) where it will be repeatedly (i.e., at every ERS updating) used for training and testing the signal-specific ANN prediction model, until being replaced by a newly selected one when the ERS is reinitialised.

Bearing in mind the above, in the module's description that follows the various issues that have been considered for building ANN models that exhibit good forecasting performance for different DMA signals are discussed first. Next, the employed EA-based optimisation strategy is described. Note that additional information about the ANN models and the EA-based optimisation strategy is provided in Appendix B and Appendix C, respectively.

ANN model building issues

The rationale for using ANN models in the novel ERS methodology is the inherent complexity of the WDSs. The main aim is to exploit their ability to model any function without explicit knowledge of the parameters involved. Allowing for this and despite the fact that the efficiency of the ANNs in modelling and forecasting water consumption has been demonstrated in a large number of studies (e.g., Adamowski, 2008), several issues have to be considered in order to build ANN models that exhibit good forecasting performance for different DMA signals. These issues include the choice of: (i) the ANN structure, (ii) the transfer function, (iii) the training algorithm, (iv) the ANN parameters, and (v) the ANN input structure. A number of preliminary sensitivity analysis type tests (not shown here) were performed in order to investigate these issues for the problem at

hand (i.e., short-term forecasting of future DMA signal values). A brief overview of the tests performed and of the main findings from these tests is given below.

With regard to the first three issues, the aforementioned tests investigated their influence on the forecasting performance and training speed. The investigated ANN structures included: (i) the Feed Forward (FF) MLP ANNs (Bishop, 1995) with one and two hidden layers, (ii) the Jordan ANN (Jordan, 1986a, 1986b), and (iii) the Elman ANN (Elman, 1990) - (see Appendix B Section B.2.2). The investigated transfer functions (for the FF ANN models only) included the Logistic and the Hyperbolic Tangent transfer functions for the neurons in the hidden layer/layers, and the Logistic, Hyperbolic Tangent and Linear transfer functions for the neuron in the output layer - (see Appendix B Section B.2.1). The investigated training algorithms (for the FF ANN models only) included the Back Propagation (BP) method (Rumelhart et al., 1986) - (see Appendix B Section B.3), and the Conjugate Gradient (Masters, 1995) and Levenberg-Marquardt algorithms - (see Appendix B Section B.2.2). FF ANNs with a Hyperbolic Tangent transfer function for the neurons in the single hidden layer and a Linear transfer function for the neuron in the output layer, trained by using the BP method were identified as the most suitable candidates (i.e., faster training and better predictive accuracy). Furthermore, an approach whereby the novel ERS methodology makes use of the same ANN structure, transfer function, and training algorithm for every ANN model (i.e., for all the DMA signals and for any DMA) was found not to affect the forecasting performance significantly. In the light of results obtained, the aforementioned ANN structure, transfer function, training algorithm, and approach were selected. Bearing this in mind, it is important to stress that the following theoretical considerations also supported this choice: (i) the effectiveness of FF-BP ANNs for the modelling and forecasting of water resources variables has been proven by many researchers (see – e.g., Maier and Dandy, 2000), (ii) it has been proven that a one hidden layer FF-BP ANN is capable of arbitrary non-linear function approximation (see – e.g., Hornick et al., 1989; Leshno et al., 1993), and (iii) it has been shown that using a Hyperbolic Tangent transfer function for the neurons in the single hidden layer has a positive impact on the predictive accuracy and also results in faster training (see – e.g., Brown et al., 1993; Kalman and Kwasny, 1993).

With regard to the selection of the ANN parameters, the aforementioned tests investigated the influence of the number of hidden neurons (i.e., elementary processing elements in the hidden layer) on the forecasting performance. On the one hand, the tests revealed that training an ANN model using too few hidden neurons leads to poor performance. On the other hand, using an arbitrary large number of hidden neurons leads to overfitting the data (i.e., such ANN model is very flexible and closely approximates the training set but it lacks the power to generalise). These findings (rather obviously) confirmed what observed in many ANN studies (e.g., Lawrence et al, 1996; Geman et al., 1992). In view of this, the approach selected for use in the novel ERS methodology involves the use of the Early Stopping (Weigend, 1994) and the Weight Decay Regularisation (WDR) (Bishop, 1995) techniques (see Appendix B Section B.4). These techniques have been successfully used (e.g., Nelson and Illingworth, 1990; Moody, 1992) to allow striking a balance between ANN learning and generalisation (i.e., achieving good generalisation performance while controlling the flexibility/complexity of an ANN model). Early Stopping involves controlling the number of training cycles while WDR involves applying a penalisation coefficient α (i.e., coefficient of WDR) to the weights of the ANN model. In this scenario, for each ANN model the "right" number of hidden units, the "right" number of training cycles, and the "right" value of the coefficient of WDR have to be accurately chosen in order to achieve the best forecasting performance.

With regard to the input structure of an ANN prediction model, the aforementioned tests investigated the influence of different input structures on the forecasting performance. The investigated ANN input structures included combinations of the following information: (i) a certain number of past pressure/flow values (i.e., *LagSize*); (ii) the TofD index associated with the forecasting horizon (i.e., the next time step) converted into a field type form (i.e., ones and zeros); and (iii) the DofW index associated with the forecasting horizon converted into a field type form. This is shown in Figure 6. Note that the choice of using a field type form for encoding the TofD and DofW indices is motivated by data representation issues when using ANN models. These indices, in fact, can be seen as a finite set of categorical data. Despite they have a natural ordering (e.g., Tuesday follows Monday), they do not have an intrinsic 'value' associated with them (e.g., Tuesday is more important than Monday). Thus to prevent the ANN model from

learning these artificial relationships a 'binary' flag is created for each possible value (e.g., Monday-flag=0000001, Tuesday-flag=0000010, etc.). The results obtained revealed that not only the *LagSize* but also the use (or omission) of the other considered explanatory variables strongly influence the ANN forecasting performance. In particular, they showed that if only the *LagSize* is varied, choosing a small *LagSize* (e.g., 4 past pressure/flow values) improves the accuracy of the predictions when large sampling intervals are considered (e.g., 10 and 15 minute intervals) and does not negatively affect the ANN prediction performance when smaller sampling intervals are considered (see Romano et al., 2010a and Chapter 5). However, if the other explanatory variables are considered (i.e., TofD and/or DofW) no general rule for the selection of the input structure can be applied. Thus, similarly to what found for the ANN parameters, for each ANN prediction model the "right" *LagSize* and the "right" combination of the other considered explanatory variables have to be accurately chosen in order to achieve the best forecasting performance.

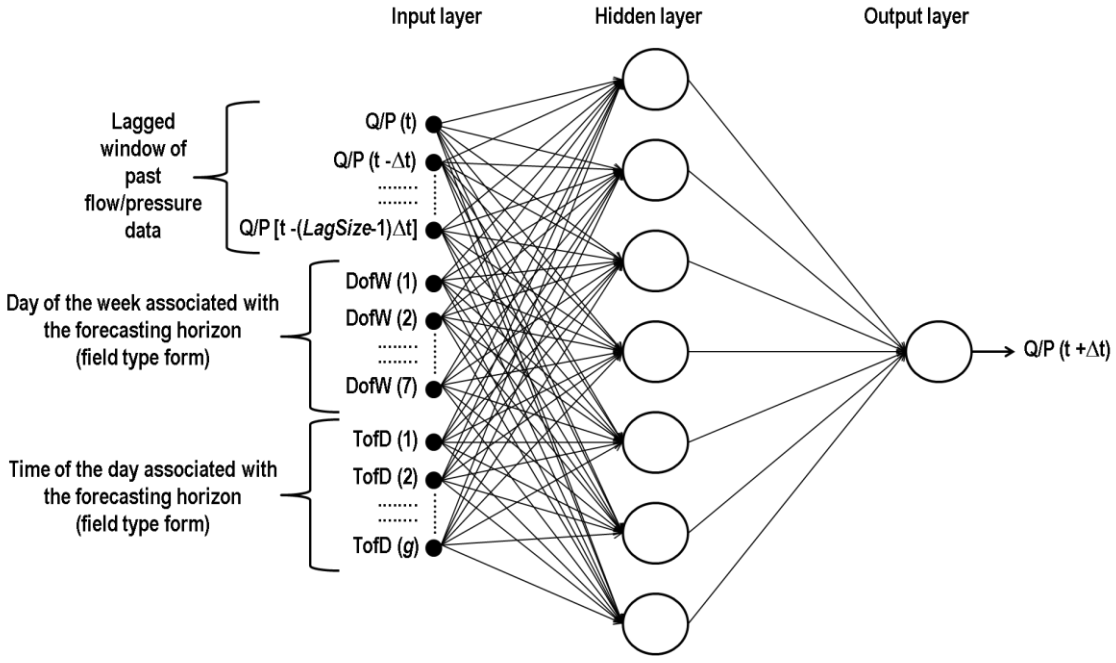


Figure 6. Artificial Neural Network for the one step ahead prediction of flow/pressure values showing a generic example of the investigated input structures.

EA-based optimisation strategy

In the light of the findings outlined above, it is possible to conclude that when the selection of the ANN parameters and input structure is considered, an approach similar to that used in the novel ERS methodology for the selection of the ANN structure, transfer function and training algorithm (i.e., using the same for every ANN model) cannot be recommended. That is to say, an optimal set of parameters and an optimal input structure are required to build an ANN model able to accurately model/forecast the particular DMA signal being analysed.

It is clear, however, that the above is a combinatorial problem. In this scenario and bearing in mind that the ERS has to potentially deal with many hundreds of different signals, the use of a manual trial and error procedure would not be feasible. Similarly, the use of an automated full enumeration procedure would be far too computational expensive. Therefore, an EA-based optimisation strategy was selected for use in the novel ERS methodology. The main reason for this is that EAs do well in large search spaces by working only with a sample available population and have the power to discover good solutions rapidly for difficult high-dimensional problems. Thus, they enable circumventing the computational limitations of the "brute-force" methods that use full search space enumeration.

The term EA is used for a broad spectrum of heuristic approaches that simulate evolution (see Appendix C Section C.1). Here, an Evolutionary Strategy (ES) algorithm (Rechenberg, 1973; Schwefel, 1981, 1995), which represents a particular instance of the several EAs that have been developed so far, is used (see Appendix C Section C.2). The parameters of this algorithm are: (i) the number of parents per generation – μ (Appendix C Section C.2.2), (ii) the number of offspring per generation – λ (Appendix C Section C.2.2), (iii) the total number of fitness function evaluations – $N_{f.f.e.}$ (i.e., termination condition) (Appendix C Section C.2.3), (iv) the probability of a parameter being perturbed – $P_{mut.}$ (Appendix C Section C.2.4), and (v) the standard deviation of normal (Gaussian) perturbation – σ (i.e., Mutation Strength) (Appendix C Section C.2.4). The "+" or "," selection operators can be selected for use in the algorithm (Appendix C Section C.2.5).

Considering the ANN parameters and the variables that define the ANN input structure shown in Table 8 as the decision variables for the problem at hand, the aim of the ES used here can be stated as follows: automatically finding the set of decision variables that minimises the ANN model prediction error on the test set (i.e., a randomly chosen subset - e.g., 30% - of the 'de-noised NOP data set' – see Section 3.4.3.7). Having said this, note that Table 8 also shows the range of values used in optimisation for each of the ANN parameters and variables that define the ANN input structure. These ranges were selected after carrying out a number of preliminary tests (not shown here) aimed at defining the size of the search space that is likely to enable finding an optimal solution for the problem at hand.

Table 8. Decision variables and associated ranges of variability.

Decision variable	**Range of values used in optimisation**
Number of hidden neurons	10 - 100
Number of training cycles	50 - 500
Value of the coefficient of Weight Decay Regularisation	10^{-5} - 10^{3}
LagSize (i.e., number of past pressure/flow values in input to the ANN prediction model)	4 - 72
Time of the Day	use/do-not-use
Day of the Week	use/do-not-use

For each generation (i.e., cycle of the ES algorithm), the ANN model prediction error on the test set is computed by using the Nash-Sutcliffe index (Nash and Sutcliffe, 1970):

$$NS_{index} = 1 - \frac{\sum_{i=1}^{n_{ts}}(M_i - S_i)^2}{\sum_{i=1}^{n_{ts}}(M_i - \bar{M})^2} \tag{3.3}$$

where M_i is an observed pressure/flow value, $\overline{M}$ is the mean of M_i, S_i is a predicted value and n_{ts} is the number of pressure/flow values in the test set.

The Nash-Sutcliffe index is a normalised statistic that determines the relative magnitude of the residual variance compared to the measured data variance. It ranges between $-\infty$ and 1, with 1 being the optimal value. $NS_{index} = 0$ indicates that the model predictions are as accurate as the mean of the observed data, whereas a $NS_{index} < 0$ occurs when

the observed mean is a better predictor than the model. The Nash-Sutcliffe index is commonly used in the literature and recommended by many (e.g., ASCE, 1993; Legates and McCabe, 1999; Moriasi et al., 2007). It is dimensionless, robust in terms of applicability to various models and overcomes some of the limitations of other widely used quantitative statistics for model evaluation.

On the one hand, error indices such as the Mean Absolute Error, the Mean Square Error and the Root Mean Square Error (RMSE) – (see – e.g., Hyndman and Koehler, 2006) – are valuable because they indicate error in the units (or squared units) of the constituent of interest. Low values are better. However, there is no commonly accepted guideline to quantify what is considered as a low value for these indices. On the other hand, standard regression indices such as the Pearson's Correlation Coefficient and the Coefficient of Determination – (see – e.g., Everitt, 2006) – are oversensitive to high extreme values (outliers) and insensitive to additive and proportional differences between model predictions and measured data (Legates and McCabe, 1999). Furthermore, Murphy (1995) shows that some of the above indices are not independent. For example, the following relevant relationship exists, $NS_{index} \leq$ Coefficient of Determination.

3.4.3.7 Artificial Neural Network training & testing module

The objectives of this module are: (a) to develop an ANN model for the short-term prediction of the DMA signal values assuming that no event occurred in the DMA being studied, and (b) to compute two key descriptive statistics to serve as a measure of the ANN model prediction error's variability. All this allows the ERS to "capture" a model-based type of information about the NOP of the DMA signal being analysed. These objectives are achieved: (i) by training (i.e., calibrating) and testing a one step ahead ANN prediction model using the 'de-noised NOP data set' and the optimised ANN input structure and parameters set selected in the ANN parameters & input structure selection module, and (ii) by calculating μ_{ts} and σ_{ts} – i.e., the values of the mean and standard deviation of the discrepancies between the ANN prediction model training set's observed and predicted values.

Note that, as it can be observed from Figure 2 and Figure 3, when this module runs at ERS updating, it continues to make use of the ANN input structure and parameters set automatically selected at ERS (re)initialisation. The rationale is that, in the absence of

operational/other DMA changes, it is expected that the DMA signal NOP will be affected only by relatively minor changes in the interval between two ERS reinitialisations. Thus, continuing to make use of the optimised ANN input structure and parameters set is likely not to affect the ANN model forecasting performance significantly. Bearing this in mind, it is possible to state that, in principle, the added computational burden (which may cause a problem if the computing power is scarce) of using the EA-based optimisation strategy at short regular time intervals (e.g., weekly) is not justified. Using the EA-based optimisation strategy periodically (e.g., every three months - when the ERS is reinitialised), on the other hand, enables the ERS to take into account factors such as the seasonal variations in the DMA's pressure/flow regime and growing demand over time. Also, it would hardly pose a computational problem even if hundreds of DMAs are monitored and the computing power is scarce. This is because the ERS reinitialisations can be scheduled to run at different times for different DMAs (i.e., in a sequential fashion) during, for example, a three month period.

Once the ANN prediction model has been trained and tested, it is passed on to the ANN forecasting module of the *DBA* subsystem (see the second dotted arrow from the top in Figure 2). This ANN prediction model will be repeatedly used in the "Execute" mode together with the incoming DMA signal data for performing one step ahead predictions of the signal values assuming that no event has occurred in the DMA (see Section 3.4.4.4). On the other hand, μ_{ts} and σ_{ts} are passed on to the evidence generation module of the *DBA* subsystem (see the third dotted arrow from the top in Figure 2). These statistics will be repeatedly used for establishing the limits within which the discrepancies between the predicted and observed incoming DMA signal values should lie assuming that no event has occurred in the DMA being studied (see Section 3.4.4.5).

Developing a short-term ANN prediction model

The ANN prediction model developed here takes as an input a number (i.e., *LagSize*) of past DMA signal values. Furthermore, depending on the input structure selected in the ANN parameters & input structure selection module, it may have the following additional inputs: (1) the TofD index associated with the forecasting horizon, and (2) the DofW index associated with the forecasting horizon. The output of the ANN prediction model is the predicted DMA signal value one step ahead in time (see Figure 6). Note

that, results obtained when analysing the ANN prediction models performance for different values of the *LagSize* are reported in Romano et al. (2010a) and in Chapter 5.

The ANN prediction model training set consists of a subset (i.e., *Train%* - e.g. 80%) of the 'de-noised NOP data set'. The remaining data (i.e., *Test%*) form the test set which is used to evaluate the ANN prediction model performance. The goodness-of-fit measure, used to compare the performance of the predicted values with their observed counterpart, is the Nash-Sutcliffe index – i.e., Equation (3.3) (see Section 3.4.3.6).

It is worth mentioning that two important checks are performed in this module. The first check makes sure that the size of the assembled training set is "satisfactory". Note that, depending on the number of 'valid' days n in the 'repaired historical data set' (see Section 3.4.3.3), the term "satisfactory" indicates that: (i) the number of 'valid' days r in the 'de-noised NOP data set' is greater than a user defined r_{min} (i.e., if $n < f$), or (ii) the number of 'valid' days left in each DofW-group after the three statistical tests described in Section 3.4.3.3 have been performed is greater than a user defined q_{min} (i.e., if $n \geq f$). The second check makes sure that the trained ANN prediction model allows achieving good forecasting performance. That is, it makes sure that the value of the calculated Nash-Sutcliffe index is greater than a minimum user defined NS_{min} (e.g., 0.9). On the one hand, if this module is running at ERS (re)initialisation and any of the above conditions is not satisfied, the problem is notified to the user of the ERS. The user can then take corrective measures (e.g., initialise/reinitialise the ERS by using more days of past data). On the other hand, if this module is running at ERS updating and any of the above conditions is not satisfied, the ERS automatically discards the newly trained ANN prediction model and continues using (in the ANN forecasting module of the *DBA* subsystem) the latest trained and successfully tested model (e.g., ANN model from the previous week, if the "Assemble" mode is set to run weekly). The latter allows the ERS to automatically take into consideration the occurrence of different situations that could affect the training of the ANN prediction model.

As an example, consider the not desirable but realistic scenario (given the current level of employed technologies) where the lifecycle of an occurred event (i.e., awareness, location, and repair times) is as long as the interval between two scheduled ERS updates (e.g., one week). In this case, the presence of too many abnormal measurements in the

past data retrieved from the Time Series database may render the size of the assembled training set not satisfactory, or the ANN model unable to accurately "capture" the NOP of the DMA signal being analysed. This, in turn, may lead to poor event detection performance. However, by always using the latest "good" ANN prediction model, the ERS is able to continue working without interruptions and producing reliable results. Note that the same considerations apply to the case of measurements from a faulty sensor and to the case of too many missing values in the historical data retrieved from the Time Series database, among the others. Further note that, based on the outcome of the above checks, the ERS behaves in a similar way with regard to: (i) the 'NOP average day' and 'NOP daily standard deviations' vectors calculated in the statistics estimation module (see Section 3.4.3.4), and (ii) the Universal Threshold calculated in this subsystem's data de-noising module (see Section 3.4.3.5).

Computing the ANN model prediction error's variability statistics

After the ANN prediction model has been trained and successfully tested, the discrepancies between the training set's observed and predicted values are used to calculate the values of the mean μ_{ts} and standard deviation σ_{ts}.

3.4.4 Discrepancy Based Analysis subsystem

3.4.4.1 Subsystem overview

The *DBA* subsystem is the first of the three ERS subsystems that aim at performing deviations identification/estimation. It provides a first way of doing this, which is based on the information about the NOP of the DMA signal being analysed "captured" by means of the ANN prediction model trained and tested in the *Setup* subsystem. Specifically, it checks that the discrepancies between the incoming observed DMA signal values and their ANN predicted counterparts do not exceed pre-defined limits based on the estimated measures of the ANN model prediction error's variability (i.e., μ_{ts} and σ_{ts} – see Section 3.4.3.7). This way, it focuses on the identification of the pressure/flow deviations induced by sudden events and can identify small to large deviations. As a matter of fact, the size of the minimum identifiable deviation is theoretically of the same order of magnitude as σ_{ts} (see Section 3.4.4.5). Because of all this, the *DBA* subsystem is particularly well suited for identifying the beginning and end of an event.

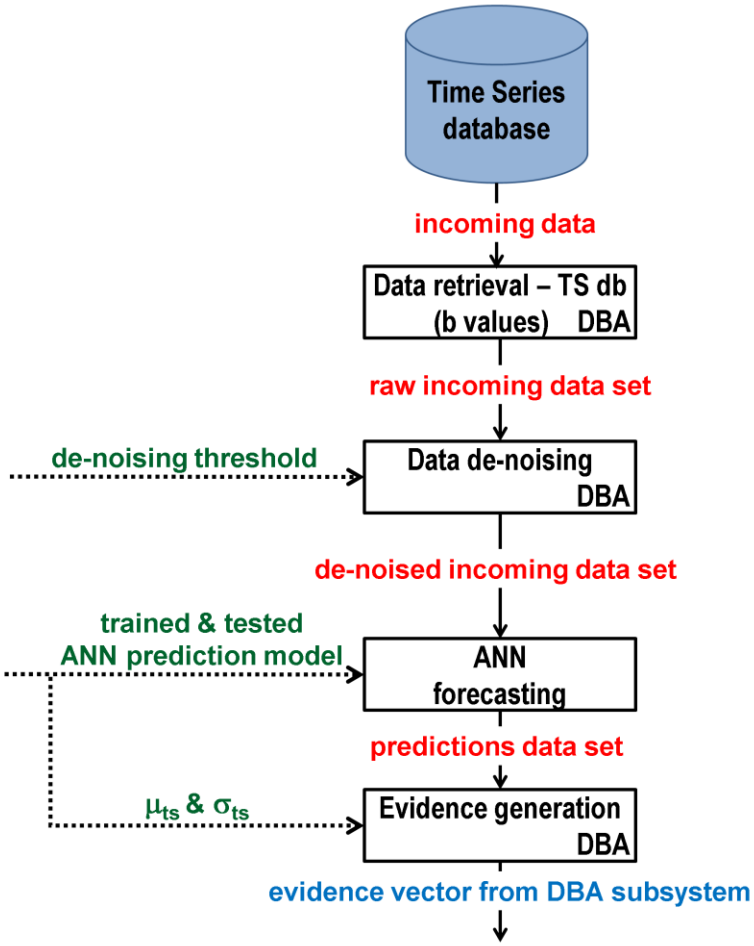


Figure 7. Processing of data in the *Discrepancy Based Analysis* subsystem.

Figure 7 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are organised into four modules. These modules are as follows: (1) data retrieval module, (2) data de-noising module, (3) ANN forecasting module, and (4) evidence generation module. The first two modules are used for retrieving from the Time Series database and de-noising the DMA signal data coming on-line. Then, the third module is used for performing one step ahead prediction of the DMA signal values. Finally, the fourth module is used for: (i) comparing the incoming observed DMA signal values to the signal values predicted by the ANN model, (ii) identifying/estimating significant (i.e.,

indicative of an event occurrence) discrepancies between those values, and (iii) further processing the identified discrepancies to provide reliable evidence of an event occurrence.

As it can be observed from Figure 2, this subsystem runs in the "Execute" mode every *a* minutes (e.g., every 30 minutes – depending on the data communication frequency).

3.4.4.2 Data retrieval module

The objective of this module is to retrieve (from the Time Series database) the latest *b* raw pressure/flow values for the DMA signal being analysed and associating to each measured value its TofD and DofW. Note that *b* is the number of past pressure/flow values that enables: (1) the discrepancies between the incoming observed DMA signal values and their ANN predicted counterparts to be analysed over consecutive time steps (see Section 3.4.4.5), and (2) the resulting pieces of event occurrence evidence (which are relative to multiple successive time steps) to be synergistically processed by the Signal level BIS and the DMA level BIS modules in the *Inference* subsystem (see Section 3.4.7). Bearing this in mind, the value of *b* is determined automatically by the ERS by taking into consideration the following: (i) the *LagSize* used for training and testing the ANN prediction model (see Sections 3.4.3.6 and 3.4.3.7), (ii) the maximum number of consecutive time steps analysed in this subsystem's evidence generation module, and (iii) the number of time steps considered by the Signal level BIS and the DMA level BIS modules in the *Inference* subsystem. For example, given that a maximum of 8 consecutive time steps are analysed in this subsystem's evidence generation module, and assuming that the *LagSize* is as equal to 4 and that 3 time steps are considered in the Signal level BIS and in the DMA level BIS modules, the value of *b* is as equal to 15 (= 4 + 8 + 3). This set of values forms the 'raw incoming data set', which is then passed on to this subsystem's data de-noising module for further processing (see Figure 2 and Figure 7).

3.4.4.3 Data de-noising module

The objective of this module is to remove noise from the 'raw incoming data set' while preserving and isolating the characteristics of an event occurring. This is achieved by using the same procedure and de-noising threshold described/computed in the data de-noising module of the *Setup* subsystem (see Section 3.4.3.5). The reason for using that

de-noising threshold is that it is based on the 'NOP data set', therefore it is "tailored" to the noise level assuming that no event occurred in the DMA being studied. The resulting 'de-noised incoming data set' is then passed on to the ANN forecasting module (see Figure 2 and Figure 7).

3.4.4.4 Artificial Neural Network forecasting module

The objective of this module is to obtain the predicted DMA signal values corresponding to the latest *b-LagSize* values in the 'de-noised incoming data set'. This is achieved by using the pressure/flow values in the 'de-noised incoming data set' as input to the ANN prediction model developed in the ANN training & testing module (see Section 3.4.3.7). The resulting 'predictions data set' is then passed on to this subsystem's evidence generation module (see Figure 2 and Figure 7).

3.4.4.5 Evidence generation module

The objectives of this module are: (a) to reliably identify the event-induced deviations from the NOP of the analysed DMA signal, and (b) to generate, at each relevant time step, an evidence that reflects the event-induced deviation presence (or absence) and its severity (if present). These objectives are achieved by: (i) analysing the discrepancies between the DMA signal values in the 'predictions data set' and their observed counterparts in successive time steps using SPC-based Control Rules (Shewhart, 1931; Deming, 1950, 1975) (see Appendix D Section D.3), and (ii) determining and classifying, at each relevant time step, the number of Control Rule violations. The resulting vector of evidence is then forwarded as an input into the relevant BIS modules (i.e., Signal level and DMA level) of the *Inference* subsystem (see Figure 2).

In this module, a starting point for the identification of the event-induced deviations is postulated to be the continuous comparison (at any particular time step) of an ANN predicted pressure/flow value and its observed counterpart. The comparative variable (or just comparative) considered here is the discrepancy. The discrepancy is defined as $x_{obs,t} - x_t$ where $x_{obs,t}$ is a pressure/flow measurement at time t and x_t is the corresponding value predicted by the ANN model. Figure 8 shows a qualitative example of the continuous comparison between observed and predicted pressure/flow values and a graphical representation of the discrepancy comparative.

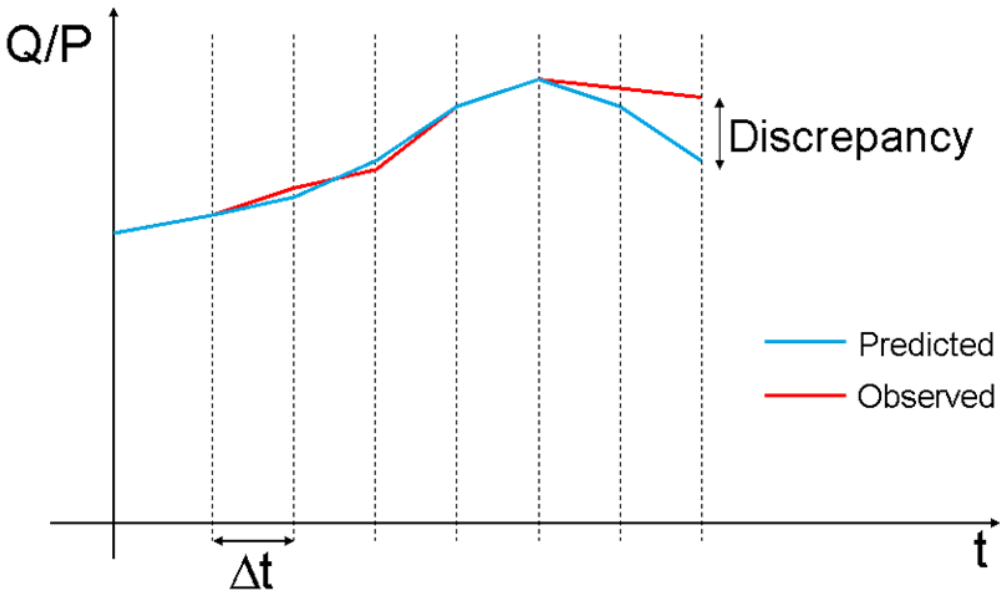


Figure 8. Qualitative example of the continuous comparison between observed and predicted pressure/flow values and graphical representation of the discrepancy comparative.

A simple Control Rule can be used to test if a discrepancy falls outside a confidence interval ($\mu_{ts} - N_l\sigma_{ts}$, $\mu_{ts} + N_u\sigma_{ts}$) where μ_{ts} and σ_{ts} are the mean and the standard deviation of the discrepancies between the training set's observed and predicted values (see Section 3.4.3.7) and N_l and N_u are user defined multipliers denoting the acceptable lower and upper control limits, respectively. If, at any time step, a discrepancy falls outside these limits this indicates the presence of an event-induced deviation from the NOP of the DMA signal being analysed. However, if the same observation is made in consecutive time steps, it increases the chance of an event actually occurring. In view of this, the use of a Control Rule that focuses the attention on the recent history of the discrepancy comparative may enable further increasing the effectiveness and reliability of the deviations identification. Similarly, as different Control Rules provide different criteria for deciding if the observed pressure/flow variations are due to the occurrence of an event, the use of a set of Control Rules may be beneficial. Allowing for this, note that a more exhaustive discussion about SPC-based Control Rules is reported in Appendix D Section D.3.

For the reasons mentioned above, the following modified subset of the Western Electric Control Rules (Western Electric Company, 1958) for detecting 'out of control' situations is used here:

- any discrepancy falls outside the $\pm 4\sigma_{ts}$ limits;
- two out of three consecutive discrepancies fall above the $+3\sigma_{ts}$ limit or below the $-3\sigma_{ts}$ limit;
- four out of five consecutive discrepancies fall above the $+2\sigma_{ts}$ limit or below the $-2\sigma_{ts}$ limit;
- eight consecutive discrepancies fall above the $+1\sigma_{ts}$ limit or below the $-1\sigma_{ts}$ limit.

This modified set of Control Rules (i.e., N_l and N_u chosen as shown above) was identified after relevant sensitivity analysis, as shown in Romano et al. (2010b) and in Chapter 5. The same analysis also tested different types of comparative variables and concluded that the discrepancy shown above generated the best event detection results.

The above Control Rules are applied to each discrepancy and a number of Control Rule violations is determined. This number is then classified in a number of classes, $n_{classes}$, that reflect the severity of an event-induced deviation. For example, if $n_{classes}$ is chosen as equal to 3, the number of Control Rule violations can be classified as follows: (i) high (3 or 4 Control Rule violations), (ii) moderate (1 or 2 Control Rule violations), or (iii) none (0 Control Rule violations).

3.4.5 Boundary Based Analysis subsystem

3.4.5.1 Subsystem overview

The *BBA* subsystem is the second of the three ERS subsystems that aim at performing deviations identification/estimation. It provides another, second way of doing this based on the basic statistical information about the NOP of the DMA signal being analysed "captured" in the *Setup* subsystem. Specifically, it checks that the incoming observed signal values lie inside a "data envelope" whose boundaries are defined by using the vectors of descriptive statistics estimated from the 'NOP data set' (i.e., 'NOP average day' and 'NOP daily standard deviations' vectors – see Section 3.4.3.4). Similarly to the *DBA* subsystem, this subsystem focuses on the identification of the pressure/flow

deviations induced by sudden events. However, it can only identify medium to large deviations. Indeed, at each time step (i.e., a particular TofD), the magnitude of the minimum identifiable deviation is theoretically of the same order of magnitude as three times the corresponding standard deviation value in the relevant 'NOP daily standard deviations' vector (see Section 3.4.5.3). Despite this, it complements the *DBA* subsystem by providing further evidence of an event occurrence. This is because, by analysing the incoming observed DMA signal values as described above, this subsystem has the potential to enable identifying deviations throughout the duration of an event.

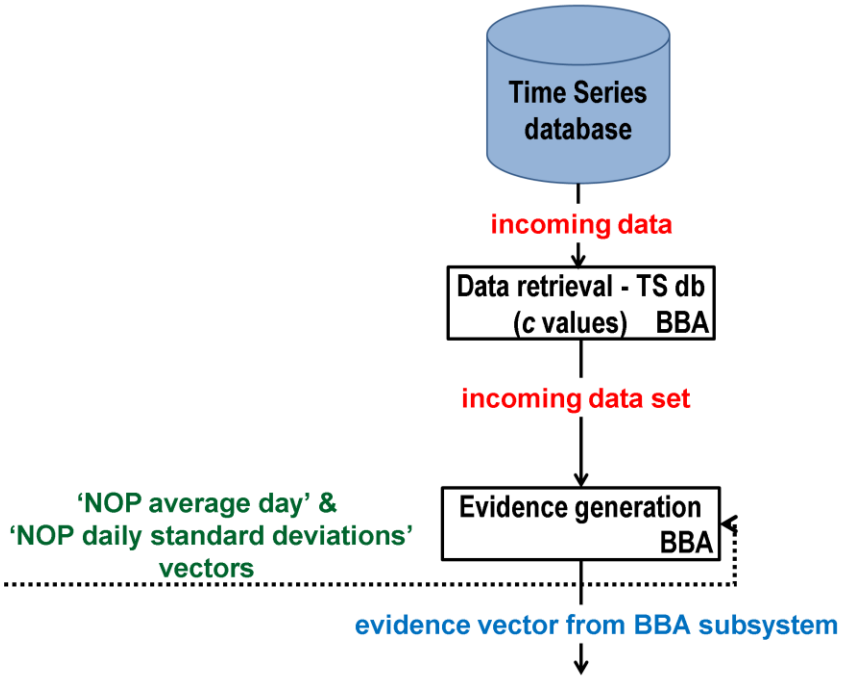


Figure 9. Processing of data in the *Boundary Based Analysis* subsystem.

Figure 9 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are organised into the following two modules: (1) data retrieval module, and (2) evidence generation module. The first module is used for retrieving the incoming DMA signal data from the Time Series database. Then, the second module is used for: (i) checking whether the incoming observed DMA signal values lie inside the aforementioned

statistical boundaries, (ii) identifying/estimating the event-induced excursions outside those boundaries, and (iii) further processing the identified excursions to provide reliable evidence of an event occurrence.

As it can be observed from Figure 2, this subsystem also runs in the "Execute" mode every *a* minutes.

3.4.5.2 Data retrieval module

The objective of this module is to retrieve from the Time Series database the latest *c* raw pressure/flow values for the DMA signal being analysed and associating to each measured value its TofD and DofW. For the same reason given in Section 3.4.4.2 with regard to the value of *b*, the value of *c* is determined automatically by the ERS. However, it only depends on: (i) the maximum number of consecutive time steps analysed in this subsystem's evidence generation module (see Section 3.4.5.3), and (ii) the number of time steps considered by the Signal level BIS and the DMA level BIS modules in the *Inference* subsystem (see Section 3.4.7). This set of values forms the 'incoming data set', which is then passed on to this subsystem's evidence generation module for further processing (see Figure 2 and Figure 9).

3.4.5.3 Evidence generation module

The objectives of this module are identical to the objectives of the evidence generation module of the *DBA* subsystem (see Section 3.4.4.5). However, the way these are achieved is as follows.

Firstly, data analyses are performed to check whether each DMA signal value in the 'incoming data set' exceeds the statistical boundaries. As each DMA signal value has an associated TofD and DofW, these boundaries are computed by using the corresponding (same TofD or same TofD and DofW - depending on the number of 'valid' days *n* in the 'repaired historical data set'– see Section 3.4.3.3) values of the average μ_i, and of the standard deviation σ_i in the relevant 'NOP average day' and 'NOP daily standard deviations' vectors. Specifically, the boundaries are defined as $\mu_i \pm N\sigma_i$, where N is a user defined multiplier. Once this is done, the identified excursions outside the statistical boundaries are identified/estimated and analysed in successive time steps using SPC-based Control Rules. Next, the number of Control Rule violations is classified into $n_{classes}$ categories and the resulting vector of evidence forwarded as an

input into the relevant BIS modules (Signal level and DMA level) of the *Inference* subsystem (see Figure 2).

Note however that, based on the results obtained after relevant sensitivity analysis (not shown here), the following set of Control Rules is used in this module:

- any single observed value falls above the $+3\sigma_i$ boundary or below the $-3\sigma_i$ boundary;
- two out of three consecutive observed values fall above the $+3\sigma_i$ boundary or below the $-3\sigma_i$ boundary;
- four out of five consecutive observed values fall above the $+3\sigma_i$ boundary or below the $-3\sigma_i$ boundary;
- eight consecutive observed values fall above the $+3\sigma_i$ boundary or below the $-3\sigma_i$ boundary.

Based on the above, a comparison can be made between the analysis carried out here and the 'flat line' (i.e., pre-defined min/max pressure/flow value) event detection methodology which is sometimes used by the water companies. The main differences are that the min/max value used here varies with the TofD/DofW and that Control Rules are used to increase the reliability of the deviations identification. However, similarly to that methodology this module also allows the user to set a pre-defined lower/upper bound should a statistically estimated min/max value be unreasonably low/high.

Figure 10 shows an example of the aforementioned statistical boundaries. The figure refers to the case where $n < f$ (see Section 3.4.3.3). Thus, only one 'NOP average day' vector and one 'NOP daily standard deviations' vector were calculated by using the *r*-day (i.e., 11-day) 'NOP dataset' (i.e., remaining 'valid' days in the 14-day period from the 16th of May to the 29th of May 2010 – see legend in Figure 10). These vectors were then used to build the upper and lower boundaries shown in the figure (i.e., data points with squares and diamonds, respectively). The figure also shows the 'NOP average day' vector (i.e., data points with circles) and the *optional* user defined upper bound set at 40 litres per second (i.e., dotted green line).

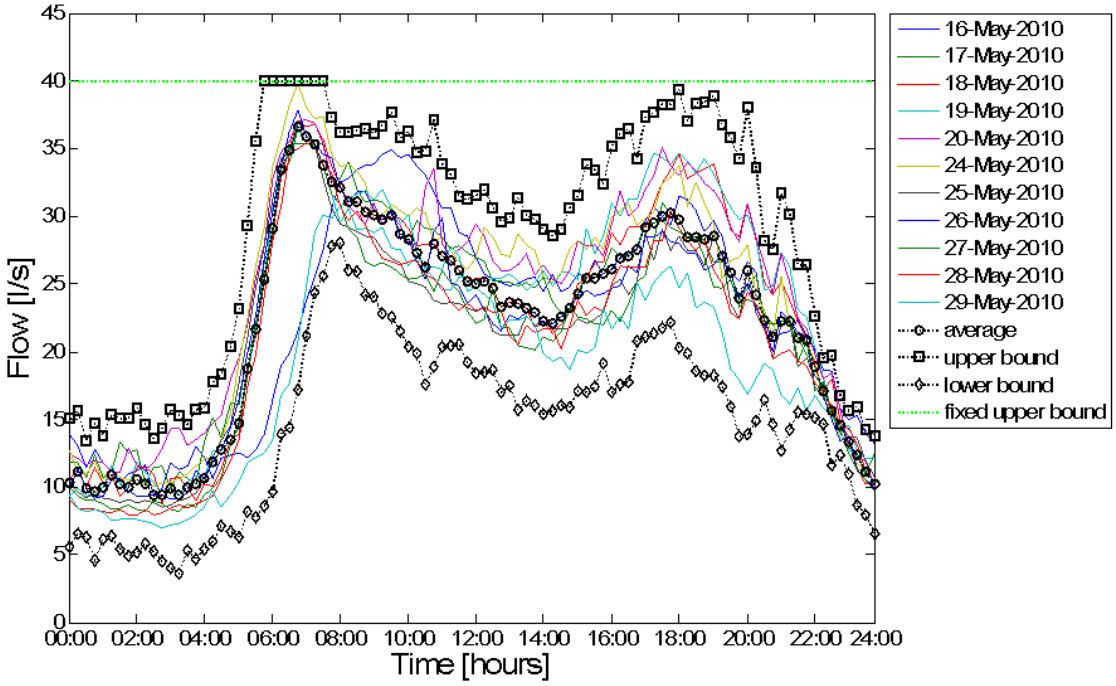


Figure 10. Example of the statistically determined boundaries, 'Normal Operating Pattern average day', and *optional* upper bound used in the *Boundary Based Analysis* subsystem.

3.4.6 Trend Based Analysis subsystem

3.4.6.1 Subsystem overview

The *TBA* subsystem is the last of the three ERS subsystems that aim at performing deviations identification/estimation. It provides a further way of doing this, which is neither dependent on the model-based nor the basic statistical information "captured" in the *Setup* subsystem. This subsystem focuses on the identification of the pressure/flow deviations induced by gradually developing events (e.g., pipe bursts developing gradually from background leaks, cumulative effect of background leaks, etc.). However, it also supports the other ERS analysis subsystems in the identification of the pressure/flow deviations induced by sudden events.

Figure 11 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are

organised into the following three modules: (1) data retrieval module, (2) data pre-processing module, and (3) evidence generation module. The first two modules are used for: (i) retrieving from the Time Series database the latest historical DMA signal data measured during a particular time window during the day, (ii) dealing with the erroneous timestamps and the missing data, and (iii) filtering out those time windows containing outliers and/or measurements that are not consistent with the expected DMA signal values assuming that no event has occurred in the DMA being studied. At the end, the third module is used for: (i) monitoring on a Control Chart the trend of the average of the measurements in each of the remaining time windows, and (ii) identifying 'out of control' situations indicative of the longer term event-induced pressure/flow variations.

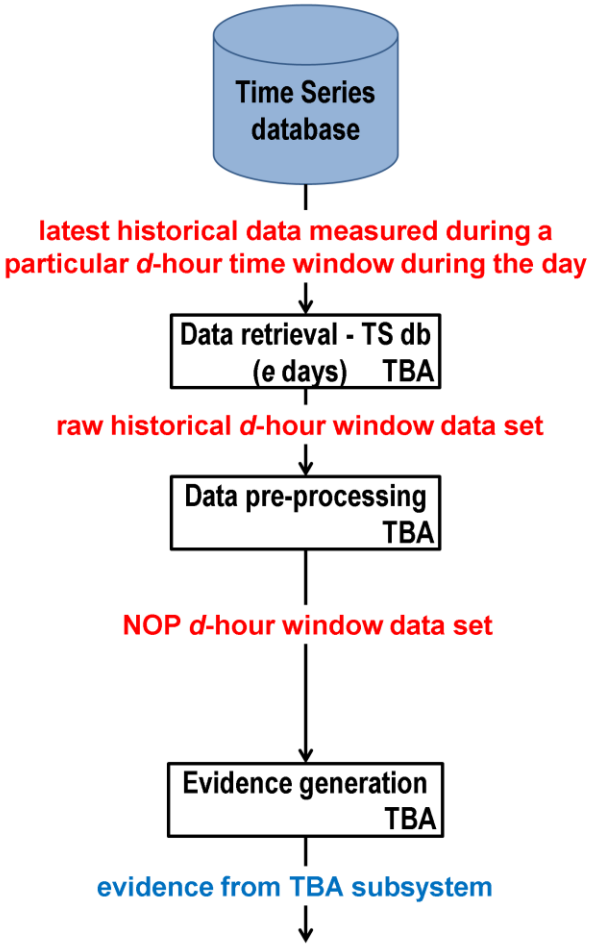


Figure 11. Processing of data in the *Trend Based Analysis* subsystem.

As it can be observed from Figure 2, this subsystem runs in the "Execute" mode every *d* hours (e.g., 4 hours).

3.4.6.2 Data retrieval module

The objective of this module is to retrieve from the Time Series database the latest *e* days (e.g., 14 days) of raw past DMA signal data measured during the *d*-hour (e.g., 4-hour) window to be analysed (e.g., latest 14 days of pressure/flow measurements recorded between midnight and 4:00 a.m., 4:00 a.m. and 8:00 a.m., etc.). The retrieved data form the 'raw historical *d*-hour window data set' which is then passed on to this subsystem's data pre-processing module for further processing (see Figure 2 and Figure 11).

3.4.6.3 Data pre-processing module

The objective of this module is to assemble (starting from the 'raw historical *d*-hour window data set') a set of data that, assuming that no event has occurred in the DMA being studied, best represents how the pressure/flow measurements recorded during a particular *d*-hour window vary over time (i.e., 'NOP *d*-hour window data set'). This is achieved by using a data pre-processing procedure similar to that used in the data pre-processing module of the *Setup* subsystem (see Section 3.4.3.3). The main differences are as follows: (i) after correcting the erroneous timestamps and replacing the blank entries with missing value indicators, the data are rearranged to form *e* vectors (i.e., one for each *d*-hour window) each containing $d \times v_{ph}$ pressure/flow values, where v_{ph} (i.e., values per hour) is the number of pressure/flow values recorded in one hour (e.g., $v_{ph} = 4$ – if a 15 minute sampling rate is considered), (ii) v_2 (i.e., number of continuous periods of missing data) and w_2 (i.e., number of missing data in each of the identified continuous period of missing data) are used in the heuristics-based procedure for dealing with the missing values, and (iii) only one statistical test (similar to the first and third statistical tests described in Section 3.4.3.3) is performed by using $N_{l4^{th}}$ and $N_{u4^{th}}$ as user defined multipliers. The rationale for using only one statistical test is allowing more variability in the assembled data set. The resulting 'NOP *d*-hour window data set' is then passed on to this subsystem's evidence generation module (see Figure 2 and Figure 11).

Note however that the above procedure for dealing with the missing values and the statistical test are applied to all but the e^{th} vector (i.e., the most recent *d*-hour window). This is because the focus of the analysis performed in the following evidence generation module (Section 3.4.6.4) is to identify the event-induced pressure/flow variations that may have been recorded during this particular time window.

3.4.6.4 Evidence generation module

The objectives of this module are similar to the objectives of the evidence generation modules of the other ERS analysis subsystems (see Sections 3.4.4.5 and 3.4.5.3). The main differences are as follows: (a) the evidence is generated every *d* hours, and (b) the evidence only reflects the presence (or absence) of an event-induced pressure/flow variation. These objectives are achieved by: (i) using a Control Chart to monitor how the mean of the pressure/flow measurements relative to the particular *d*-hour window being analysed changes over time, and (ii) determining and classifying the occurrence of an 'out of control' situation. Similarly to the other evidence generation modules, the evidence is then forwarded as an input into the two BIS modules (Signal level and DMA level) of the *Inference* subsystem (see Figure 2).

In the analysis carried out here, a mean is calculated for each of the remaining *d*-hour vectors in the 'NOP *d*-hour window data set'. The Control Chart plots these means in time order, a centre line at the average of the means, and upper and lower control limits at three standard deviations from the centre line. The standard deviation is estimated by taking the average of the remaining *d*-hour vectors' standard deviations. If the mean of the measurements recorded during the latest *d*-hour time window is an outlier of the range defined by the upper and lower control limits, those pressure/flow measurements are considered as indicators of an 'out of control' situation. This information is then classified as either 'out of control' or 'in control'.

Note that a loose comparison can be made between the analysis carried out here and the monitoring of the variations in the night flow (i.e., values usually recorded between midnight and 4:00 a.m. to capture the MNF) which is sometimes carried out by the water companies. Indeed, the data analysis method implemented in this module can be seen as a generalisation of such approach whereby, through the analysis of a *d*-hour

"moving" window during the day, the event-induced pressure/flow variations may be detected hours before they become apparent in the night trend.

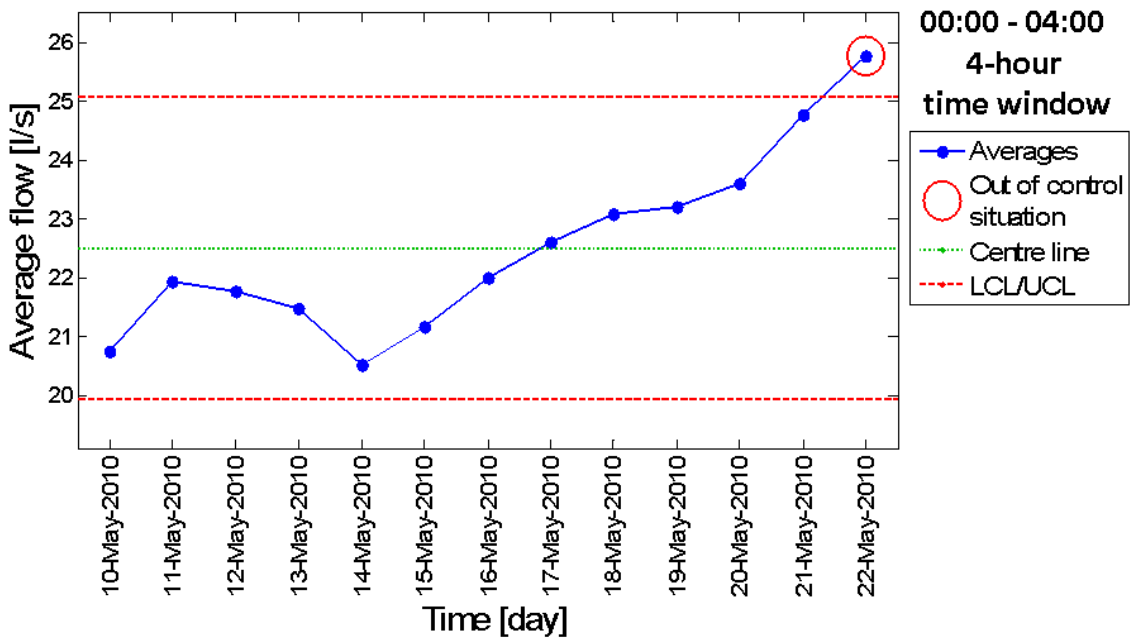


Figure 12. Example of a Control Chart used in the *Trend Based Analysis* subsystem.

Figure 12 shows an example of a Control Chart used in the *TBA* subsystem. The figure refers to the data from a flow signal recorded during the 4-hour time window between midnight and 4:00 a.m. The 14-day period of past raw data analysed goes from the 10th of May 2010 to the 22nd of May 2010. It is possible to observe from this figure that the means of the 4-hour vectors are plotted in time order (i.e., data points represented as blue filled circles), a centre line is plotted at the average of these means (i.e., dotted green line), and upper and lower control limits are plotted at three standard deviations from the centre line (i.e., LCL/UCL – dashed red lines). The figure shows that the average night flow gradually increases during the considered 14-day period and that the mean of the measurements recorded during the latest 4-hour time window is an outlier of the range defined by the upper and lower control limits (i.e., 'Out of control situation' – red circle).

3.4.7 Inference subsystem

3.4.7.1 Subsystem overview

The *Inference* subsystem constitutes the final event detection processing stage of the novel methodology implemented in the computer-based ERS. In fact, it enables raising the detection alarms.

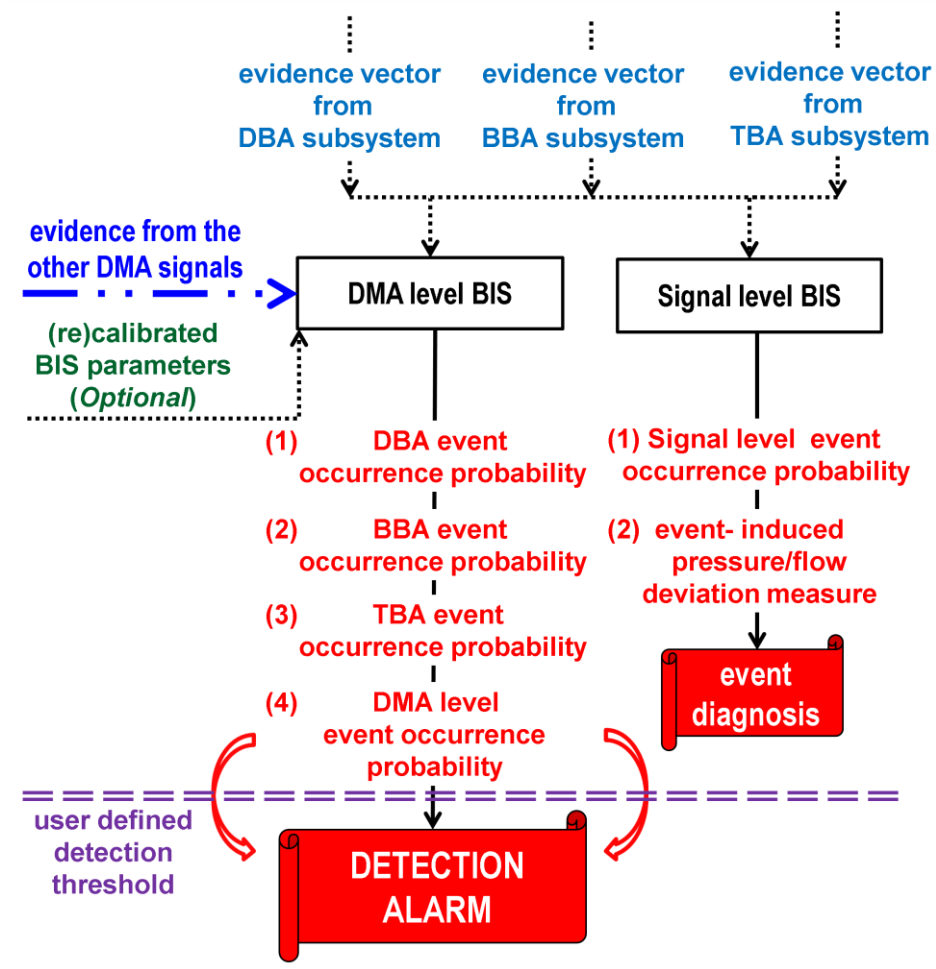


Figure 13. Processing of data in the *Inference* subsystem.

Figure 13 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are organised into the following two modules: (1) Signal level BIS module, and (2) DMA level BIS module. In these two modules all the generated evidence of an event occurrence at hand are combined to infer the probability that an event has occurred in

the DMA, raise detection alarms, and provide additional information that allows performing a diagnosis of the event occurring.

As it can be observed from Figure 2, this subsystem runs in the "Execute" mode every *a* minutes.

3.4.7.2 Signal level Bayesian Inference System module

The objective of this module are: (a) to calculate, at each time step during the data communication interval, the probability that an event has occurred (i.e., Signal level event occurrence probability) based on the evidence resulting from the three ERS analysis subsystems (see Figure 2 and Figure 13), and (b) to provide, at each time step during the data communication interval, a measure of the pressure/flow deviation from the NOP of the DMA signal being analysed "captured" by means of the basic descriptive statistics in the *Setup* subsystem. These objectives are achieved by: (i) using a BIS that consists of a discrete input Bayesian Network (BN) (Edwards, 2000; Jensen, 2001) for performing inference, and (ii) estimating the difference between an observed DMA signal value and its corresponding value in the relevant 'NOP average day' vector. Note that the latter measure of deviation may be used (after a detection alarm is raised – see Section 3.4.7.3) to evaluate the likely event magnitude (e.g., assessing the burst flow rate). The module's output is then stored in the Alarms database.

A discrete input BN is a Directed Acyclic Graph (DAG) consisting of nodes and arcs. Each node represents a random variable or a group of random variables (e.g., event occurrence evidence from a particular ERS analysis subsystem at a specific time) and is discretised into states (e.g., high, moderate, or none). The arcs express probabilistic relationship between these variables. A Conditional Probability Table (CPT) is associated with each node and contains the parameters (i.e., probability values) that are used for performing inference (see Appendix E Section E.2.2). Explicitly, in a CPT associated with an input node (i.e., root node with no predecessors) a prior probability value is assigned to each of the node's state. In a CPT associated with a non-input node a conditional probability value is assigned to each of the possible combinations of the node's states and the states of the node's direct predecessors. Note that additional information about the BNs is given in Appendix E.

A BIS that consists of a discrete input BN is used for achieving this module's first objective because BNs allow reasoning under uncertainty. A BN is ideally suited for situations that require combining multiple inputs in order to infer a meaningful (probabilistic) output for decision-making. Furthermore, it allows updating the probability of an outcome (e.g., event occurrence) as evidence accumulates. With enough evidence, it should become very high or very low.

In view of the above, the structure of the BIS used in this module (i.e., Signal level BIS) is designed in such a way that it is able to process synergistically the evidence from the different ERS analysis subsystems (i.e., *DBA*, *BBA*, and *TBA*) and to encode the temporal sequence of the incoming data. The temporal sequence of the incoming data is encoded by considering the event occurrence evidence relative to a number of time steps, *z*. This allows reasoning over time by updating probabilities of outcomes over consecutive time steps. Figure 14 shows the structure of the Signal level BIS for z=3 (i.e., t, t-Δt and t-$2\Delta t$).

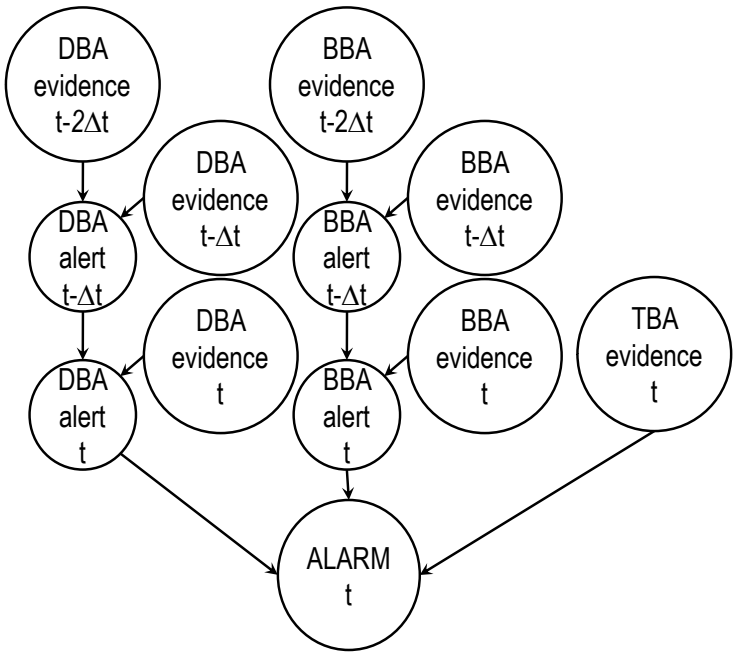


Figure 14. Structure of the Signal level Bayesian Inference System considering three consecutive time steps.

Although this module is not used to raise the detection alarms, it provides information that may be used in the diagnosis of an event occurring. In fact, it allows the ERS to assess the extent to which the particular signal has contributed to generate a detection alarm. As it will be shown in the remainder of this chapter and in Chapter 4 and Chapter 5, the latter is particularly useful because it may allow: (1) to identify the area of the DMA that has been affected the most by the event (i.e., the likely event location), and (2) to determine the likely root cause of the event (e.g., faulty sensor, failure of a pressure modulating valve, etc.).

3.4.7.3 District Metered Area level Bayesian Inference System module

The main objective of this module is to raise detection alarms at the DMA level if and when necessary. Similarly to the Signal level BIS module, a BIS that consists of a discrete input BN encoding the temporal sequence of the incoming data is used here for inferring, at each time step during the data communication interval, the probability that an event has occurred (i.e., DMA level event occurrence probability). The main difference is that the BIS used in this module (i.e., DMA level BIS) takes as an input the evidence resulting from the different ERS analysis subsystems coming simultaneously from all the DMA signals (see Figure 2 and Figure 13). It has to be stressed that, by enabling the processing of all the event occurrence evidence available in a synergistic way it is expected that the ERS will be able to increase its efficiency and reliability of detecting the occurrence of an event in the studied DMA. This hypothesis will be proved on three case studies (see Chapter 4 and Chapter 5). Note that the increased reliability is also achieved because the detection alarms will continue to be raised even when one or multiple DMA signals are not being analysed by the ERS (e.g., due to sensors that stopped transmitting the data). Indeed, as long as at least one DMA signal is available for analysis, the ERS will continue to perform inference, although without exploiting its full potential (i.e., synergistically analysing the evidence from different and multiple sensors). Finally, a detection alarm is generated when the DMA level event occurrence probability goes above a user defined detection threshold (i.e., λ).

In addition to the aforementioned DMA level event occurrence probability, the DMA level BIS also outputs the event occurrence probabilities based on: (i) the evidence from the *DBA* subsystems only, (ii) the evidence from the *BBA* subsystems only, and (iii) the evidence from the *TBA* subsystems only. These three additional event occurrence

probabilities can help evaluate the extent to which a particular type of analysis has contributed to generate a detection alarm and, for example, identify the event as a growing rather than a sudden pipe burst.

Once a detection alarm is raised, it is stored in the Alarms database and the following information is saved: (i) the DMA location, (ii) the DMA level event occurrence probability, (iii) the event occurrence probabilities for each type of analysis performed (i.e., short-term *DBA* and *BBA*, and long-term *TBA*), and (iv) the detection alarm start time. Note that, in order to avoid raising unnecessary alarms for the same event in the future, the ERS simply suppresses any further detection alarm for a user specified *'alarm inactivity time'* (e.g., 1 week).

3.4.8 Bayesian Inference System parameters learning subsystem

3.4.8.1 Subsystem overview

The *BIS parameters learning* subsystem constitutes a further processing stage of the novel methodology implemented in the computer-based ERS. It enables performing the fourth basic action mentioned in Section 3.2 (i.e., BIS calibration/recalibration) aiming at further (continuously) improving the classification performance of the DMA level BIS and hence the event detection performance of the ERS.

Figure 15 shows a flowchart describing the processing of data in this subsystem and the main inputs and outputs of each of the data analyses performed. As it can be observed from this figure and from Figure 2, the data analyses carried out in this subsystem are organised into the following two modules: (1) data retrieval module, and (2) Expectation Maximisation (EM) based parameters learning module. The first module is used for retrieving, from the Alarms database, information (i.e., start time and duration) about past events occurred in the DMA being studied. This information is then used in the second module to automatically (re)calibrate the parameters of the DMA level BIS.

As it can be observed from Figure 2, this subsystem runs in the "Learn" mode. Therefore, as mentioned in Section 3.4.2, this subsystem is not used by the ERS if the past events information is unavailable or not considered. Taking this into account, however, it is important to stress that the data analyses performed in this subsystem

constitute an important part of the novel methodology developed as part of the research work carried out in this thesis and they are fully integrated into the ERS presented here.

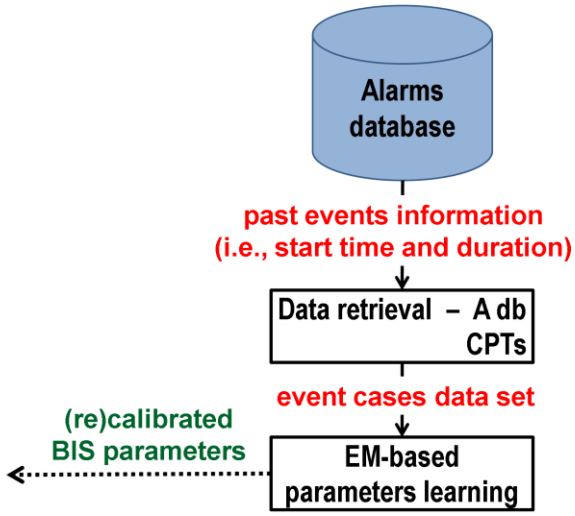


Figure 15. Processing of data in the *Bayesian Inference System parameters learning* subsystem.

3.4.8.2 Data retrieval module

The objective of this module is to retrieve from the Alarms database a certain amount of data about the past events occurred in the analysed DMA. Taking into consideration the way the past events information can be obtained (see Section 3.3), the amount of data to be retrieved depends on: (i) the length of the time window for which the water company's historical records are available, or (ii) the frequency at which the review of the alarms raised by the ERS is performed. The retrieved data form the 'event cases data set', which is then passed on to the EM-based parameters learning module (see Figure 2 and Figure 15).

3.4.8.3 Expectation Maximisation based parameters learning module

The objective of this module is to (re)calibrate the parameters of the DMA level BIS. This is achieved (as shown below) by using the 'event cases data set' and the EM strategy (Dempster et al., 1977; Lauritzen, 1995). The (re)calibrated DMA level BIS parameters are then passed on to the *Inference* subsystem (see Section 3.4.7.3) for use in

the DMA level BIS module. Note that, in the description that follows, it is also shown how a 'cumulative learning' procedure could be used in a semi-automatic fashion for enabling the (re)calibration of the DMA level BIS parameters as knowledge about confirmed past events becomes available and, as a consequence, continuously improving the ERS detection performance.

As mentioned in Section 3.4.7, the nodes of the two BISs used in the *Inference* subsystem (i.e. Signal level BIS and DMA level BIS) represent a random variable/group of random variables. These variables are discrete and a CPT is associated with each node. The CPTs contain the prior probabilities of all the root nodes and the conditional probabilities of all the non-root nodes. These prior/conditional probabilities encode the strengths of the dependencies among the nodes. According to Jensen (2009), there are two knowledge sources for selecting the parameters (i.e., probabilities) in the CPTs of a BIS. These are: (i) domain experts, and (ii) databases.

The CPT parameters 'normally' used (i.e., when no knowledge about past events is available/considered) by the ERS in its Signal level BISs are the same for each signal and for every DMA. The CPT parameters 'normally' used by the ERS in its DMA level BIS are the same for every DMA. These CPT parameters have been selected according to the former knowledge source (i.e., domain experts). Specifically, this has been done by carrying out a number of preliminary tests (results not shown here) and by incorporating domain knowledge obtained by the theoretical research framework, literature, and observational experience.

Figure 16 shows an example of the aforementioned domain experts' knowledge based parameters. Note that this figure shows the structure of the DMA level BIS assuming that three consecutive time steps (i.e., t, t-Δt and t-$2\Delta t$) are considered (see Section 3.4.7.2), $n_{classes}$ is as equal to 3 (see Section 3.4.4.5), and that this DMA level BIS is used for simultaneously analysing one flow and one pressure signal only. Further note that, in the figure, nodes with the same symbols (e.g., stars, diamonds, etc.) have the same parameters in their CPTs.

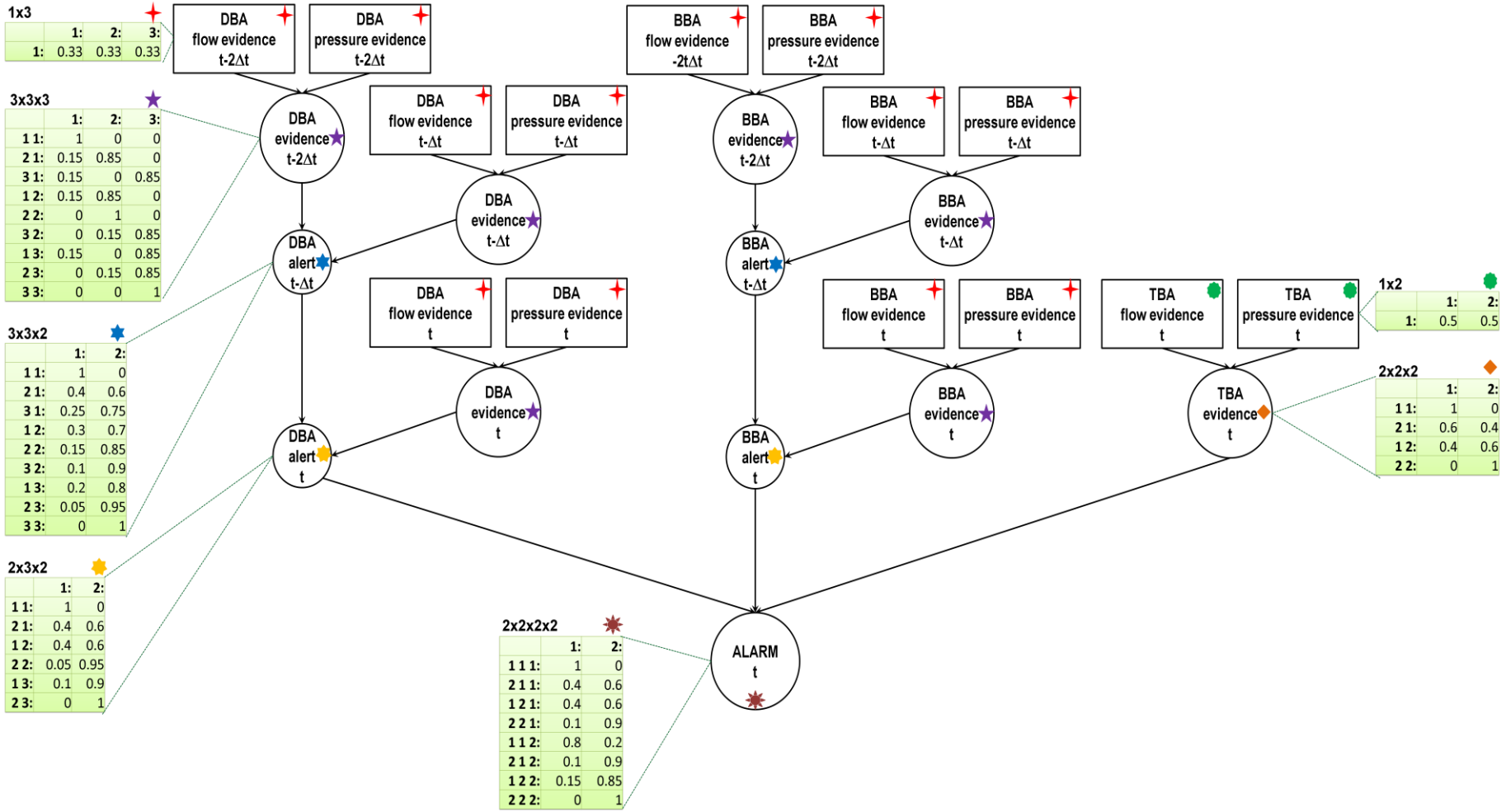


Figure 16. Structure of the District Metered Area level Bayesian Inference System with examples of the domain experts' knowledge based parameters used in the Conditional Probability Tables associated with its nodes.

Alternatively, when information about past events occurred in the DMA being studied is available, the CPT parameters in the Signal level BISs and in the DMA level BIS could be learned (i.e., calibrated/recalibrated) directly from these data (i.e., 'event cases data set'). In this scenario, if the structures of these BISs had no hidden nodes (i.e., those for which there is no observed data), the estimation of the parameters would be simple and could be done just by calculating (counting and dividing) the prior or conditional probabilities. However, as the Signal level BISs and the DMA level BIS include such nodes (see for example the 'alert' nodes in Figure 14 and Figure 16), they are only partially observable. Thus, in order to estimate the CPT parameters an algorithm capable of estimating them from incomplete data must be used.

The EM algorithm is the most commonly employed algorithm for estimating the CPT parameters from incomplete data (Jensen, 2009). Furthermore, it has the advantage of being simple, robust and easy to implement (Do and Batzoglou, 2008). For these reasons, the EM algorithm has been selected for use in the ERS methodology. The EM algorithm is an iterative algorithm that, given a set of training data (i.e., 'event cases'), determines estimates of the CPT parameters that are optimal within a neighbouring set of solutions. It does this by assuming the pattern of missing data is uninformative (i.e., missing at random or missing completely at random). Maximum A Posteriori (MAP) is estimated in the situation where initial knowledge about the parameters is assigned. Maximum Likelihood (ML) is estimated in the situation where non-informative prior beliefs are used. No initial knowledge about the parameters is assigned in the developed ERS. This is because, with the increasing availability of 'event cases' (see the 'cumulative learning' procedure described below), the effect of the prior becomes small (Heckerman, 1995). Furthermore, the ML allows simplifying the calculation. Taking this into consideration, the EM algorithm works as follows. It starts with initial values for all the CPT parameters (e.g., chosen at random) and then iteratively refines them to produce a locally optimal ML fit (i.e., when convergence is reached). The iteration process consists of two steps, namely the Expectation and Maximisation steps. The Expectation step fills-in the hidden parameters with the distribution given by the current model. The Maximisation step then treats these filled-in values as if they had been observed and re-estimates the model parameters. Note that further details about the EM algorithm and other methods for estimating the BIS parameters from a set of training data (i.e., learning) are provided in Appendix E.

Bearing in mind the above, it is important to say that both the CPT parameters of the various Signal level BISs and the CPT parameters of the DMA level BIS could be (re)calibrated by using information about confirmed past events. However, in the ERS such information is used for (re)calibrating the CPT parameters of the DMA level BIS only. This is because, in the ERS, only the DMA level BIS is used for raising the detection alarms (see Section 3.4.7). Furthermore, as a particular event may have occurred anywhere in the DMA, it may have not affected the measurements from a sensor located farther away from it. Therefore, "forcing" changes (through learning) to the CPT parameters of that sensor's Signal level BIS(s) in order to obtain high Signal level event occurrence probabilities could be counterproductive.

In the framework outlined above, a 'cumulative learning' procedure could be used to efficiently exploit the past events information. Note that, although independent from the way this information is obtained (see Section 3.3), it is assumed here that this procedure takes advantage of the ERS feedback mechanism (shown in Figure 1 as dashed red arrows) which may allow the user to periodically review the ERS alarms and easily populate the Alarms database with new confirmed 'event cases'. Taking this into consideration, the proposed 'cumulative learning' procedure works as follows. Once a review of the raised alarms has been carried out for the first time, the defined 'event cases' can be used for calibrating the CPT parameters. Subsequently, when a second review process is carried out, the newly defined 'event cases' (i.e., event occurred in the time window between the two reviews) can be used together with ('cumulative learning') the 'event cases' defined in the previous review for recalibrating the CPT parameters. Progressing in a like fashion, over time, the Alarms database will be populated with 'cases' of events of different type and size that occurred in different areas of the DMA being studied. Thus, it is hypothesised here that the ERS may be able to learn recognising the features of a large variety of events, thereby potentially continuously improving its generalisation and event detection capabilities. This hypothesis will be proven on a case study in Chapter 5 whereby the use of the aforementioned procedure, in a semi-automatic fashion, for (re)calibrating the DMA level BIS parameters as knowledge about the events occurred in the DMA being studied becomes available, is simulated.

3.4.9 Discussion

In the data pre-processing module of the *Setup* subsystem (see Section 3.4.3.3), outliers that are not related to bursts or other similar events and that should really be part of the 'NOP data

sets' may be eliminated. This could render the ERS detection framework more sensitive to the presence of such outliers (hereafter referred to as 'normal' outliers) in the observed data streams. This is because those pre-processed data series are then employed to train the ANN models used in the *DBA* subsystem and to compute the relevant statistics used in the *BBA* subsystem. However, the 'normal' outliers usually occur at random times and affect a single data point in time, as opposed to burst and other events induced outliers that usually affect multiple, consecutive data points in time (and may exhibit an identifiable or quantifiable pattern). Bearing this in mind, in the ERS, the potential false alarms resulting from single outliers are reduced by using SPC-based Control Rules to analyse the incoming data over consecutive time steps. Indeed, by using these rules, persistent outliers have a greater chance of triggering an alarm than single outliers. Additionally, in the ERS, the potential false positive alarms resulting from single outliers are further reduced by using the DMA level BIS, via event occurrence evidence processing and the use of a threshold on estimated event occurrence probabilities.

In addition to the above, it could also be argued that the persistent outliers themselves may not be related to events that are of interest for water companies (e.g., bursts, sensor faults, illegal consumptions, etc.) and that should really be part of the 'NOP data sets'. These may be due to increased/decreased consumption during a very hot day, a holyday or a special occasion such as a big sport event in town, for example. At the present, the ERS methodology does not account for these situations (although this would be possible and could be quite straightforward - e.g., information about holydays might be considered when assembling the 'NOP data sets', and/or as an extra input to the ANN prediction model, and/or as an extra input to the DMA level BIS). As a consequence, such outliers might trigger detection alarms. Allowing for this, it is worth mentioning that the aforementioned situations would likely result in a number of alarms being simultaneously raised in multiple DMAs that cannot however be affected by the same local anomaly (e.g., a burst). Therefore, following aposteriory analysis of the raised alarms (e.g., by means of an additional ERS module or manually by the operator), it could be possible to attribute the cause of the alarms to some global anomaly (e.g., a very hot day).

The minimum size of the events detectable by using the presented methodology depends on the DMA size (i.e., the larger the size of the DMA the larger an event has to be in order to be detected), and on the DMA head loss characteristics. This is because in larger or high head

loss DMAs the legitimate variability in pressure/flow may overwhelm the changes in pressure/flow due to an event. Furthermore, in the case of a large DMA, the pressure signals may not be affected by an event if such event is small and far away from the sensors. Taking all this into consideration, it could be noted that with more (than one or two) pressure sensors within a DMA, the accuracy could improve and the DMA size may not be an issue. Also, the aforementioned legitimate variability in pressure/flow is less accentuated during the night hours and, it could be noted that events such as pipe burst often occur at night when the pressure in the DMA is at its highest.

3.4.10 Section summary

The current UK practice normally involves deploying pressure and flow sensors at a number of locations in each DMA, typically at the entry/import/export points and a pressure sensor at the critical point in the DMA. The DMA sensors typically collect and transmit the data at regular time intervals (e.g., every 15 or 30 minutes). A novel data analysis methodology that makes use of these pressure/flow data streams (i.e., signals) to detect pipe bursts and other similar events in a timely and reliable manner has been presented in this section.

The novel methodology is implemented in a computer-based ERS. The ERS simultaneously processes the DMA signals and provides a range of outputs as follows: (a) probability of an event occurring in the DMA being studied (used to raise detection alarms by comparing it to a pre-specified threshold value), (b) specific event occurrence probabilities for each DMA signal analysed and for each type of analysis performed (i.e., *DBA*, *BBA*, and *TBA*), (c) estimates of the pressure/flow magnitude variations (e.g., likely burst flow), and (d) event timing. The information generated by the ERS may enable water companies to: (i) react more quickly after an event occurrence in order to mitigate its negative impact, (ii) determine the likely cause of the event occurring, and (iii) rank alarms and hence prioritise responses.

In this section, bearing in mind that the description of the developed data analysis methods' theoretical background and implementation details has been organised in terms of the ERS subsystems and associated modules, an overview of the ERS has been given first in Section 3.4.2. Then, the six main ERS subsystems, namely the *Setup* subsystem, the *DBA* subsystem, the *BBA* subsystem, the *TBA* subsystem, the *Inference* subsystem and the *BIS parameters learning* subsystem have been described in detail in Sections 3.4.3 to 3.4.8, respectively. Finally, a brief discussion about the main limitations of the ERS methodology has been presented in Section 3.4.9.

As an ensemble, Sections 3.4.3 to 3.4.9 have shown the following:

- By making use of a range of AI and statistical techniques, the ERS enables efficiently extracting and capturing information about the NOPs of the DMA signals being analysed;
- As the AI/statistical techniques used are data-driven and have the self-learning capability, the ERS is able to work in an on-line context and generically applicable to different DMAs. Indeed, the ERS is not only able to adapt to changes in the DMA operating conditions but also to "tailor" itself to the particular DMA being studied;
- The ERS enables identifying and estimating the event-induced deviations from the expected signal NOPs by simultaneously looking at: (i) different types of event occurrence indicators, and (ii) both the short and the long term effects of an event on the pressure/flow measurements;
- In order to generate reliable and timely detection alarms, the ERS enables synergistically processing the results from the complementary subsystems that analyse each DMA signal from all the different and multiple DMA signals;
- The ERS enables the (re)calibration of its parameters as knowledge about past events occurred in the DMA being studied is acquired aiming at further (continuously) improving the reliability and speed of the raised detection alarm.

The methodology presented in this section enables detecting pipe bursts and other events by making use of the pressure and flow data from the small number of sensors typically deployed nowadays in the UK DMAs. In the following section it is shown how, when a DMA is observed by means of a larger than currently used in the UK practice number of sensors, the capabilities of the developed methodology can be further enhanced and extended to efficiently deal with the increased data availability and to determine the approximate location of an event within a DMA.

3.5 Event Detection & Location in Over-Sampled District Metered Areas

3.5.1 Section overview

As mentioned in Section 3.1, the latest developments in hydraulic sensor technology and on-line data acquisition systems are enabling the UK water companies to deploy a larger number of more accurate and increasingly cheaper pressure and flow sensors and the data collected

by these devices to be transmitted to the water companies in near real-time. Given this trend and the fact that the UK water companies have started recognising the benefits of monitoring their WDSs in near real-time, it is possible to envisage that in the near future a larger number of these devices will be deployed in the UK DMAs.

In view of the above, this section focuses on enhancing and extending the capabilities of the methodology that enables performing event detection in normally-sampled DMAs (see Section 3.4). The main aim is to enable it to: (i) efficiently deal with the increased data availability, and (ii) exploit the event information resulting from the analysis of the larger number of DMA signals for determining the approximate location of an event within the DMA (in addition to detecting it in a fast and reliable manner). Given this objective, in this section the work done in order to customise and further develop the ERS described in Section 3.4 is presented. The ERS's customisation mainly involved replacing the discrete input BN used in the DMA level BIS module (see Section 3.4.7.3) with a continuous input multivariate Gaussian mixtures-based graphical model (see – e.g., Duda and Hart, 1973). This was done to overcome the BN limitations when dealing with the availability of an increased number of event occurrence evidence. The ERS's further development involved introducing an additional ERS subsystem that processes on-line the output information generated by the various Signal level BISs (see Section 3.4.7.2) by using the Geostatistical Analysis (see – e.g., Isaaks and Srivastava, 1989). This analysis enables associating a probability value of an event occurrence to each DMA pipe. The intention was to provide a means to identify the group of DMA pipes where the event most likely occurred.

The section is organised as follows. A brief overview of the customised and further developed ERS focussing on the main differences with the ERS described in Section 3.4 is given in Section 3.5.2. Once this is done, in Sections 3.5.3 and 3.5.4 the theoretical background and implementation details of the techniques used for the ERS's customisation and further development are respectively presented. Finally, a summary of the section is given in Section 3.5.5.

3.5.2 Customised-further developed Event Recognition System overview

A diagrammatic representation of the customised and further developed ERS is shown in Figure 17. By comparing this figure with Figure 2, it is possible to observe two key differences. The first difference is in the *Inference* subsystem and the second difference is in the introduction of a new subsystem, namely the *Location* subsystem.

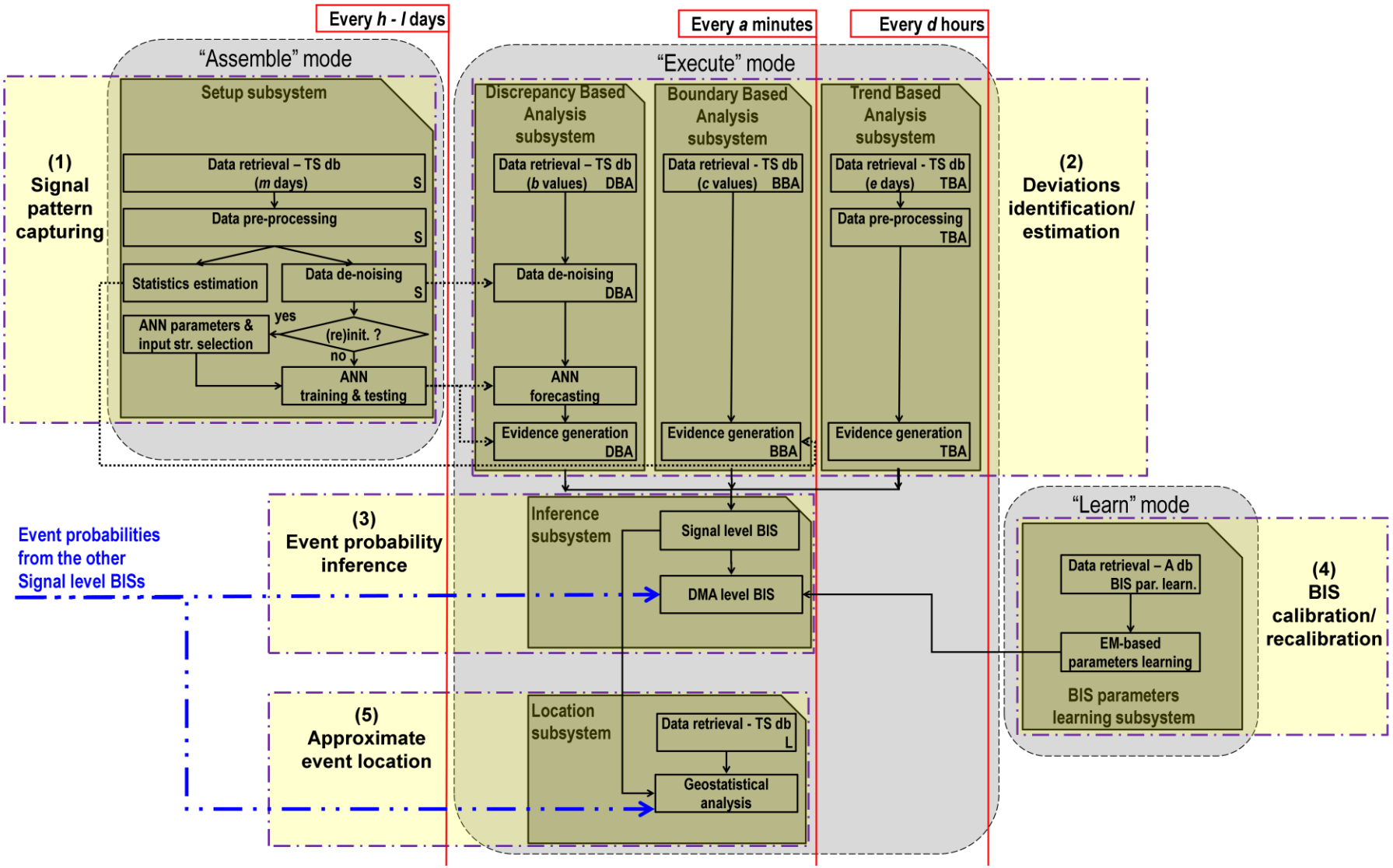


Figure 17. Diagrammatic representation of the customised and further developed Event Recognition System.

With regard to the first difference, it is possible to observe from Figure 17 that the DMA level BIS module takes as an input the event occurrence probabilities estimated at the sensor locations (i.e., the output of the various Signal level BISs) rather than, directly, the event occurrence evidence resulting from the different ERS analysis subsystems and coming simultaneously from all the DMA signals. As it will be explained in more detail in Section 3.5.3, this data processing route diversion accompanies the replacement, in the DMA level BIS module, of the discrete input BN described in Section 3.4.7.3 with a continuous input multivariate Gaussian mixtures-based graphical model. This new DMA level BIS enables inferring the probability of an event occurrence at the DMA level while circumventing several of the difficulties/limitations that would arise if the discrete input BN had to continue to be used for processing the event information resulting from the analysis of the larger number of DMA signals.

With regard to the second difference, Figure 17 shows that the newly introduced *Location* subsystem is used to perform the fifth basic action mentioned is Section 3.2 (i.e., approximate event location). In this subsystem, following the generation of a detection alarm, the output information provided by the various Signal level BISs is processed further by using the Geostatistical Analysis. Based on the locations of the deployed sensors, this analysis enables performing spatial interpolation of the event occurrence probability variable. As a result, a model (i.e., interpolation surface) that predicts the probability value of an event occurrence associated with each DMA pipe is built. The predicted DMA pipe event occurrence probabilities are then used to identify the group of DMA pipes where the event most likely occurred.

3.5.3 Inference subsystem customisation

As described in Section 3.4.7.3, the main objective of the DMA level BIS module is to infer, at each time step during a data communication interval, the probability that an event has occurred in the DMA being studied. This then enables (by means of a user defined detection threshold) raising the detection alarms if and when necessary.

In the ERS described in Section 3.4, the DMA level BIS consists of a BN which takes as inputs the evidence of an event occurrence resulting from the different ERS analysis subsystems (i.e., *DBA*, *BBA*, and *TBA*) and coming simultaneously from all the DMA

signals. Consequently, if an increase in the number of DMA signals that are simultaneously analysed in the ERS is assumed, the number of BN inputs and hence the overall number of BN nodes and parameters in the CPTs increases drastically. This may affect the computational efficiency and therefore effectiveness of the inference process. Furthermore, manual specification of all the required CPT parameters is tedious, and necessitates not only domain knowledge but also understanding of the probabilistic calculus and of the probabilistic graphical models. Even when the automatic learning of the CPT parameters from the past events is considered (see Section 3.4.8 and Appendix E Section E.3), the increased number of 'missing' parameters (i.e., those in the hidden BN nodes) together with the large multidimensional data structure pose serious challenges to the efficiency of the algorithm for learning from incomplete data that has to be used (e.g., the EM algorithm).

To avoid the potential computational inefficiency during the inference process and to circumvent the other difficulties/limitations outlined above, in the ERS customised and further developed here, the DMA level BIS consists of a two-class (i.e., alarm on, alarm off), two-component multivariate Gaussian mixtures-based graphical model. This DMA level BIS works in an *n*-dimensional feature space by inferring, at any time step, the probability that an event has occurred in the DMA based on the continuous output (i.e., event occurrence probability value) of the *n* Signal level BISs. This implies that, unlike the discrete input BN-based DMA level BIS used previously (see Section 3.4.7.3), the various types of event occurrence evidence resulting from the different data analyses performed on each DMA signal are not used as input directly. Indeed, here, they are first processed in the Signal level BISs. The outputs of these BISs are then fed into the multivariate Gaussian mixtures-based graphical model. Taking this into account, note that, by processing the event occurrence evidence in such a way, the new multivariate Gaussian mixtures-based DMA level BIS still enables synergistically combining the event information from all the analysed DMA signals aiming at increasing the reliability of the detection alarms.

It is worth stressing that the above results in an ERS with a 'hierarchical architecture' (i.e., Signal level BISs - DMA level BIS). This has the potential to enable (if desired and/or beneficial) embedding the local detection intelligence (i.e., data analyses performed on the single DMA signal) into the sensor device itself ("smart device" in the

AI jargon). The central detection intelligence (i.e., the data analysis performed in the DMA level BIS module), on the other hand, can become part of the overall decision support type system used in the water company's control room.

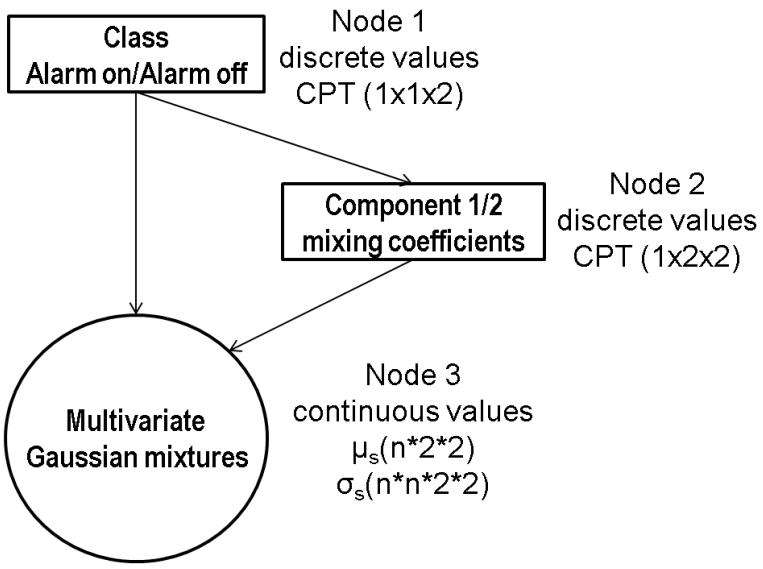


Figure 18. Diagrammatic representation of the District Metered Area level Bayesian Inference System in the customised and further developed Event Recognition System.

Figure 18 shows the structure of the new DMA level BIS. It can be observed from this figure that this graphical model has three nodes, two of which have discrete parameters for each of the allowed states of the relevant variables. In the third node (i.e., input node), on the other hand, the probability distribution is described by using the multivariate Gaussian mixture function formula (see Appendix E Section E.3.2).

The parameters of the multivariate Gaussian mixtures (i.e., means and covariances) together with the other model's parameters are determined by using the EM algorithm (see Appendix E Section E.3)

3.5.4 Location subsystem development

As mentioned in Section 3.5.1, the ERS further development involved introducing an additional ERS subsystem (i.e., the *Location* subsystem) which enables determining the approximate location of an event within the DMA. The methodology behind this ERS

subsystem makes use of the Geostatistical Analysis for the on-line processing of the output information generated by the various Signal level BISs (i.e., Signal level event occurrence probabilities).

Geostatistical Analysis focuses on the relationship between the value of a variable at a given location (e.g., event occurrence probability at a generic point within the DMA) and the values of the same and (possibly) other variables (e.g., pressure, flow, etc.) at a set of measured/analysed locations at a certain distance from it (e.g., sensor locations). Its basic goal is to interpolate the value of a variable at locations which have not been measured/analysed, using data from the surrounding measured/analysed locations. Ultimately, it allows creating a model (i.e., interpolation surface) of how the variable's values are distributed across the entire domain of interest (e.g., DMA). Note that further details about the general Geostatistical Analysis theory and several techniques (i.e., geostatistical techniques) that are used for performing it are given in Appendix F.

As it will be shown in Chapter 5, a number of geostatistical techniques were tested for enabling estimating the probability values of an event associated with the DMA pipes and, thus, provide a means for identifying the group of DMA pipes where the event most likely occurred. These included the Inverse Distance Weighted (IDW) (Shepard, 1968), Local Polynomial (LP) (Gandin, 1963; Cleveland and Devlin, 1988), Ordinary Kriging (OK) (Isaaks and Srivastava, 1989), and Ordinary Cokriging (OC) (Myers, 1982) geostatistical techniques (see Appendix F Sections F.2 to F.5). These techniques were selected because they are among the most well-known and widely used geostatistical techniques. Furthermore, as an ensemble, they are representative of the six main classes the geostatistical techniques can be divided into. Indeed, it has to be noted that the geostatistical techniques can be classified according to three different criteria as follows: (1) depending on the nature of the function that is used to interpolate the variable's values, they can be classified as deterministic (i.e., IDW and LP) or stochastic (i.e., OK and OC), (2) depending on which measured/analysed locations are considered when interpolating the variable's values at the locations which have not been measured/analysed, they can be classified as global (i.e., IDW) or local (i.e., LP, OK and OC), and (3) depending on whether the values of a single or multiple variables at the measured/analysed locations are used for estimating the variable's values at locations which have not been measured/analysed, they can be classified as univariate

geostatistical techniques (i.e., IDW, LP, and OK) or multivariate geostatistical techniques (i.e., OC).

Following the aforementioned tests, the OC technique was eventually selected for use in the customised and further developed ERS. This is because it was found to improve the prediction performance and reduce the variance of the estimation error. Indeed, OC is a geostatistical technique developed to improve the estimation of a variable (i.e., primary variable or variable of interest) by drawing on the additional information from other spatially correlated variables (i.e., secondary/auxiliary variables). It exploits the auxiliary information by directly incorporating the values of the secondary variables and measuring the degree of spatial association with the primary variable through the cross-semivariogram. Note that further information about this technique is provided in Appendix F Section F.5).

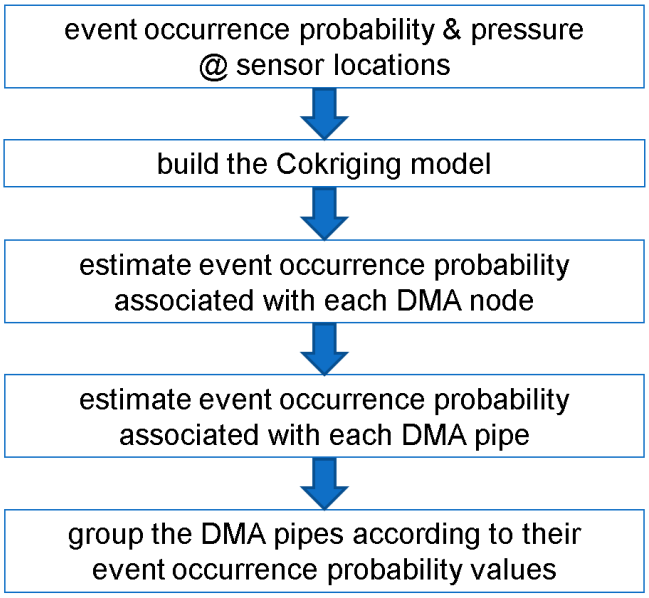


Figure 19. Flow chart showing the main procedural steps for performing approximate event location by means of the Ordinary Cokriging geostatistical technique.

Bearing in mind the above and considering the event occurrence probability value as the variable of interest and the measured pressure at the various sensor locations as the

secondary variable, approximate event location by means of the OC geostatistical technique is achieved in the *Location* subsystem according to the procedure shown in Figure 19.

As soon as an alarm is raised, the available values of the variable of interest estimated at the sensor locations (i.e., output of the various Signal level BISs), and the values of the secondary variable measured at the sensor locations (which are retrieved from the Time Series database) are used to build the OC interpolation surface. This interpolation surface, in turn, is used for estimating the probability values of an event occurrence at every network junction (nodes - i.e., prediction locations). A probability value of an event occurrence is then calculated for each pipe (link) of the network by taking the average between the event occurrence probability values estimated at the pipe's start and end nodes. In this way, an event occurrence probability value is associated with each DMA pipe. Note that the rationale for doing this is that an event (e.g., a burst) usually occurs at a point along the length of a pipe rather than at a junction. Bearing this in mind, note that an average OC standard deviation of interpolated junction values is also calculated for each DMA pipe in order to assess the uncertainty of the prediction made. Once this is done, the DMA pipes are grouped based on their event occurrence probability value in order to identify DMA pipes with similar event occurrence probability values. Here, this grouping is carried out by using the Jenks Natural Breaks Classification method (Jenks, 1967). This method allows determining the best arrangement of the event occurrence probability values into a user defined number of classes (i.e., $Jenks_{classes}$) by seeking to minimise the variance within classes and maximise the variance between classes. The rationale for using this particular classification method is that it aims to present a series of break values that best represent the actual breaks observed in the data as opposed to an arbitrary classificatory scheme (e.g., equal interval). In this way the actual clustering of the event occurrence probability values is preserved (subject to the arbitrary specification of $Jenks_{classes}$). As a result of the application of this procedure, the group of DMA pipes with the highest predicted event occurrence probability values indicates the DMA area where the event most likely occurred.

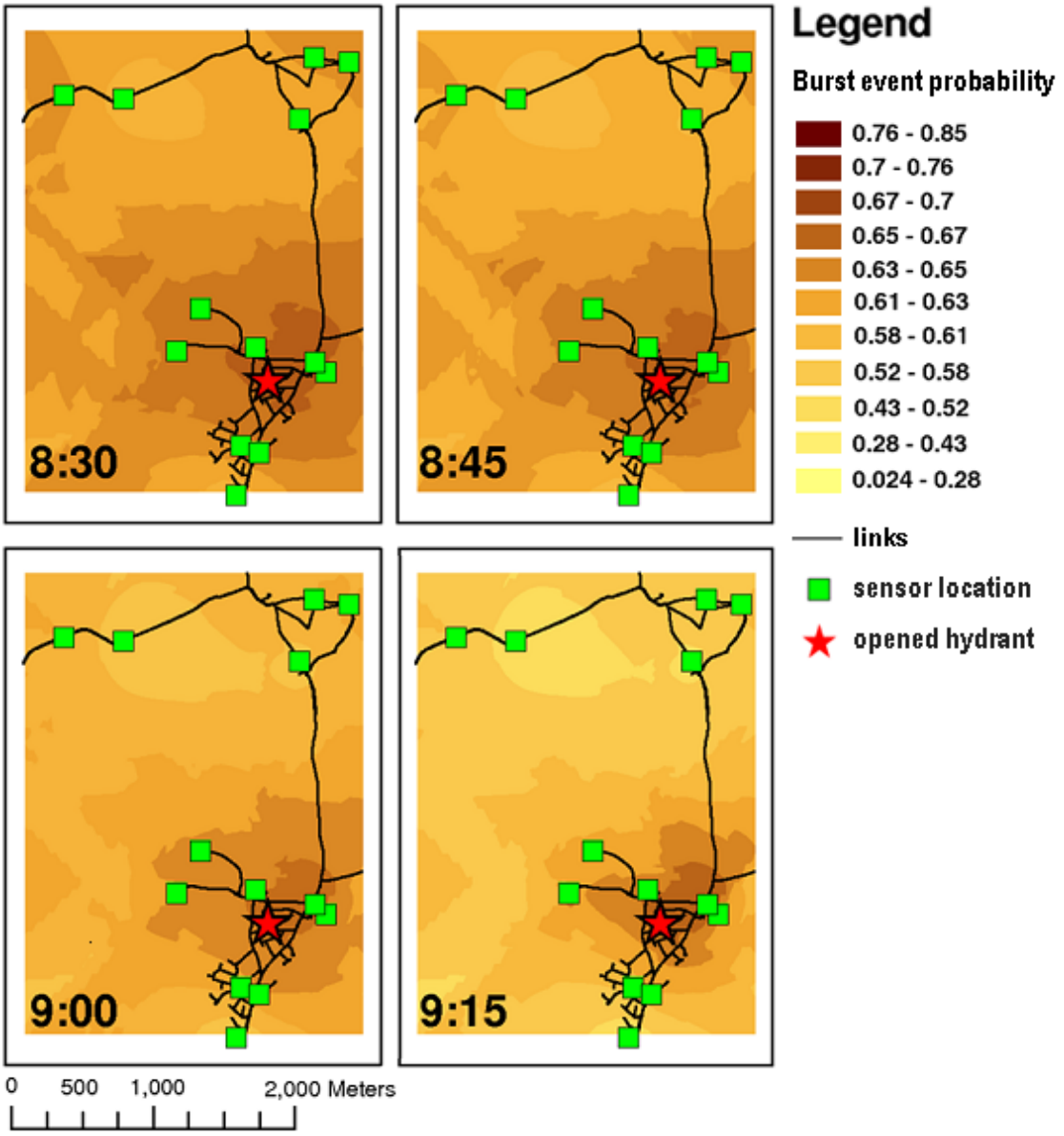


Figure 20. Example of how the interpolation surface changes over time.

The procedure outlined above is repeated at every time step after an alarm is raised (e.g., every 15 minutes). However, due to the dynamics of a distribution pipeline system, an event might affect the pressure/flow measurements from different sensors at different times. Thus, in order to obtain a more precise indication of the likely event location as time progresses, the ERS uses the Signal level BISs' output at the time steps following detection in a cumulative fashion. That is to say, at the time step an event is

detected, only the output of the various Signal level BISs at that particular time step is used. At the following time step, the output of the various Signal level BISs at the time step the event was detected (i.e., previous time step) and the output of the various Signal level BISs at the current time step are used together, and so on for the following time steps (while the ERS remains in an alarm state). As an example, Figure 20 shows how, following a burst event simulated by opening a hydrant (location indicated with a red star symbol), the interpolation surface obtained by using the OC technique changes with time. Taking this into account, note that detailed approximate event location results, obtained by using the OC and the other considered geostatistical techniques, can be found in Section 5.6.

3.5.5 Section summary

The wider availability of more accurate and lower cost pressure/flow sensors allows us to envisage that in the near future a larger number of these devices will be deployed in the UK DMAs. Assuming the availability of pressure/flow data from a DMA which is observed by means of a larger than currently used in the UK practice number of sensors (i.e., over-sampled DMA), this section has focussed on enhancing and extending the capabilities of the event detection methodology described in Section 3.4 aiming at efficiently deal with the increased data availability and determining the approximate location of an event within the DMA.

Data analysis methods have been developed for achieving the above aims. As the implementation of these methods in the computer-based ERS has involved customising and further developing it, this section has been organised as follows. An overview of the customised and further developed ERS has been given first in Section 3.5.2. Then, the theoretical background and procedural details of the methods whose implementation has involved the ERS's customisation have been presented in Section 3.5.3. The theoretical background and procedural details of the methods whose implementation has involved the ERS's further development have been presented in Section 3.5.4.

As an ensemble, these sections have shown the following:

- By using a multivariate Gaussian mixtures-based DMA level BIS for inferring the probability that an event has occurred in an over-sampled DMA, the ERS can efficiently deal with the increased number of event occurrence evidence;

- By using the OC geostatistical technique for processing in near real-time the output information from the various Signal level BISs, the ERS is able to associate an event occurrence probability value to each DMA pipe thereby providing a means to identify the group of DMA pipes where the event most likely occurred.

3.6 Summary and Conclusions

The on-line pressure and flow data from the sensors deployed in the UK DMAs provide a potentially useful source of information for detecting and locating pipe bursts and other similar events both quickly and economically. A novel detection and location methodology that makes use of these data has been presented in this chapter.

The methodology presented here is implemented in a computer based ERS. If only the pressure and flow data from the small number of sensors typically deployed nowadays in the UK DMAs is available, the ERS enables the detection of pipe bursts and other events at the DMA level in a timely and reliable manner. It also provides a range of outputs that may allow determining the likely cause of the events occurring and ranking/prioritising the alarms raised. If data from a larger than current UK practice number of sensors deployed in a DMA is available, then the ERS also enables determining the approximate location of these events within the DMA. The reliable and timely detection of events together with the identification of their likely cause and approximate location, which could be achieved by using the novel methodology developed and presented here, can facilitate prompt interventions and repairs. This, in turn, may reduce the potential damages to the infrastructures and to third parties and improve the water company's operational performance and customer service thereby yielding substantial improvements to the state-of-the-art in near real-time WDS incident management.

The chapter has been organised as follows. After the introduction in Section 3.1, the philosophy behind the development of the novel event detection and location methodology has been discussed in Section 3.2. Then, a high level overview of the data processing route for performing near real-time event detection and location by using the novel methodology presented here has been given in Section 3.3. Next, Section 3.4 has provided the theoretical background and the methodological details of the data analyses that enable the ERS performing event detection in normally-sampled DMAs. Section

3.5 has provided the theoretical background and the methodological details of the data analyses that enable the ERS performing event detection and location in over-sampled DMAs.

This chapter has shown the following. The ERS makes use of a range of AI and (geo)statistical techniques that enable effectively and efficiently mimicking the behaviour of a qualified and experienced human operator. It does not make use of hydraulic or any other physically based simulation model of the WDS. The AI and (geo)statistical techniques used are data-driven and have the self-learning capability. Hence, they give the ERS the ability to (re)initialise, update, and (re)calibrate its parameters as conditions in the DMA being studied change and knowledge about the past events occurred in that DMA is acquired. This, in turn, enables the ERS to work in an on-line context and to be generically applicable to different DMAs whilst being "tailored" to the specific DMA to which it is applied. The ERS not only enables analysing a DMA signal by looking at different aspects of the event detection problem (i.e., different event occurrence indicators, short-long term effect of an event on the pressure/flow measurements) but also simultaneously analysing the multiple and different DMA signals in a synergistic manner. When data from an over-sampled DMA are used, the ERS can efficiently deal with the increased number of event occurrence evidence for inferring the probability that an event has occurred. Finally, the ERS is able to further analyse its output spatially and identify where in the DMA the event most likely occurred. Additional comment made here is that the developed ERS is readily transferable to practice.

In the next two chapters the event detection capabilities of the ERS presented in Section 3.4 and the event detection and location capabilities of the customised and further developed ERS presented in Section 3.5 are respectively tested, verified and demonstrated on a number of case studies. Additionally, the results of several sensitivity type data analyses that focussed on evaluating the detection and location performances of the novel methodology presented in this chapter for different choices of parameters and data analysis methods are also reported.

CHAPTER 4 CASE STUDIES FOR EVENT DETECTION IN NORMALLY-SAMPLED DISTRICT METERED AREAS

4.1 Introduction

This chapter reports the results of a number of data analyses carried out on two case studies. The main aim of these analyses was to test, evaluate and illustrate the capabilities of the novel methodology that enables performing event detection in normally-sampled DMAs - i.e., as implemented in the ERS described in Section 3.4. In the case studies presented in this chapter the historical data from the pressure and flow sensors deployed in several Yorkshire Water Services (YWS) DMAs were used. These data were collected as part of the NEPTUNE project (Savić et al., 2008).

The chapter is organised as follows. After this introduction, Sections 4.2 and 4.3 report the results obtained on the first and on the second case study, respectively. In each of these sections a general description of the case study including details of: (i) the area(s) being studied, (ii) the data used for the analyses, and (iii) the practical implementation of the methodology in the specific case study being analysed is given first. This general description is then followed by the details of the various data analyses performed and their outcomes. Once this is done, a summary of the chapter and the main conclusions are given in Section 4.4.

4.2 Case Study #1

4.2.1 Section overview

In this section, after a general description of the case study, five analyses that focussed on testing and evaluating the capabilities of the methodology that enables performing event detection in normally-sampled DMAs are presented. The main aims of the analyses performed and presented here were as follows.

The first analysis aimed at testing and evaluating the methodology's capabilities on real-life events such as pipe bursts, sensor failures and other pressure/flow anomalies in the recorded DMA signal trends.

The second analysis aimed at testing and evaluating the methodology's capabilities on burst events simulated by opening fire hydrants.

The third analysis aimed at investigating the issue concerning the optimal selection of the user defined detection threshold value to be used for raising the alarms and evaluating the methodology's capabilities for different values of this parameter.

The fourth analysis aimed at evaluating the benefits of the EM strategy for calibrating the parameters in the CPTs of the DMA level BIS.

The fifth analysis aimed at performing a comparison of the methodology's capabilities in the following two cases: (1) a BN is used in the DMA level BIS module of the ERS for determining the occurrence of an event, and (2) an ANN is used instead.

Details of how these analyses were carried out and of the results obtained are reported in Sections 4.2.3 to 4.2.7.

4.2.2 Case study #1 description

The Harrogate and Dales region in the Yorkshire county in the north of the UK consists of nearly 200 DMAs and includes approximately 122,000 properties. Prior to the start of the NEPTUNE project, YWS implemented a pilot area in this region. Specifically, YWS installed about 450 Cello Loggers (Technolog, 2011) at the inlet/outlet and critical points of the DMAs in the pilot area. The Cello Loggers were equipped with a GSM modem capable of individual GPRS data connectivity. The aim was to dramatically improve data transfer for both pressure and flow data. It is worth noting that these sensors are unique in the UK water industry in that they communicate via GPRS to YWS's mobile telephony provider Orange's Data Centre, and from there the data are relayed to YWS via a high speed broadband connection known as the Orange Pipeline. Data from these sensors are communicated every 30 minutes and two readings are obtained (15 minute sampled data). These data are then stored in a YWS database.

The case study examined and presented in this section made use of the historical data from the pressure and flow sensors deployed in one of the YWS pilot area's DMAs. The DMA being studied is predominantly urban. It has a total of 2,640 domestic properties and 500 commercial users, of which 48 have a demand greater than 400 cubic metres

per year. There is one major metered consumer within this DMA which has an average annual metered consumption of 12,072 cubic metres. The total mains length is 24 kilometres. Figure 21 shows a satellite view of the DMA being studied with overlaid pipe network.

Figure 21. Case study #1 District Metered Area.

The historical data provided by the YWS consisted of 15 minute readings from the flow and pressure sensors located at the DMA inlet as well as 15 minute readings from the pressure sensor located at the critical DMA point. The flow data were averaged values during the sampling interval whilst the pressure data were 15 minute instantaneous values. These data referred to an eleven month period, between July 2009 and June 2010. For the same time period, Work Management System (WMS) data (i.e., data on

repairs carried out in the network) and Customer Contacts (CCs) data (i.e., customer complaints about problems related to water supply) were also provided.

In this case study, information about the real-life events occurred in the DMA being studied was used to achieve the aim of the first analysis mentioned in Section 4.2.1. Information about a series of EEs carried out in the DMA being studied was used to achieve the aims of the second, third, fourth and fifth analyses mentioned in Section 4.2.1. The EEs were carried out by fitting a standpipe to fire hydrants at different locations within the DMA being studied. The fire hydrant valve (with an in-line flow meter connected) was then slowly opened until the required flow rate was reached. This enabled simulating different types of burst occurring in different areas of the DMA being studied. The rationale for using information about the EEs is that they allowed achieving the relevant analyses' aims without needing to wait for actual pipe bursts to occur and without the associated uncertainty about the duration and magnitude of these events.

The historical data collected between September 2009 and February 2010 and April-May 2010 (i.e., a total of eight months) were used for the analysis that made use of the real-life events information. The remaining data (i.e., July-August 2009 and March 2010) were used for the analyses that made use of the EEs information (because the EEs were conducted in these three months).

Table 9. Event Recognition System user defined parameters used in the case study.

Schedule & data retrieval	l=7 days; m=14 days; a=30min; b=15values; c=11values; d=4 hours; e=14 days; f=90 days; g=96; v_{ph}=4.
Statistical tests	$N_{l1^{st}} = N_{u1^{st}} = N_{l2^{nd}} = N_{u2^{nd}} = 3$. $N_{l3^{rd}} = N_{u3^{rd}} = N_{l4^{th}} = N_{u4^{th}} = 3$.
Procedure for dealing with the outliers	v_1=6; w_1=4. v_2=1; w_2=4.
ANN prediction models	$LagSize$=4 values; α=0.01; training cycles=400; $Train\%$=80%; $Test\%$=20%; NS_{min}=0.9; r_{min}= 7 days.
BISs	$n_{classes}$=3; z=3.

With regard to the practical implementation of the methodology, Table 9 summarises the values of the user defined ERS parameters used for the five analyses carried out on this case study (see Section 4.2.1). Note, however, that Table 9 does not show the values of all the user defined ERS parameters. This is because the values of parameters such as the *alarm inactivity time* and the detection threshold were varied according to the needs of the particular analysis performed. For this reason, the 'analysis-specific' values of these user defined ERS parameters are given in the relevant sections.

Bearing in mind the above, it is important to stress that although in the practical implementation of the methodology historical time series were used, the pressure and flow data were fed to the ERS in a simulated on-line fashion (i.e., as the ERS would have been used in real-life). According to this, the ERS was initialised with data recorded during the first two weeks (i.e., m=14 days) of July 2009 and then updated weekly (i.e., l=7 days). For the purposes of the analyses carried out on this case study (see Section 4.2.1), however, the periodic ERS reinitialisation was not performed and the ANN parameters & input structure selection module (see Figure 2 and Section 3.4.3.6) was not used. Here, for each of the three signals coming from the sensors deployed in the DMA being studied, the parameters and input structure of the relevant ANN prediction model were chosen as follows. A rule of thumb was used for selecting the number of hidden neurons (Neuroshell2 manual, 1996). This rule of thumb selected the number of hidden neurons as equal to half the sum of the number of ANN inputs and outputs plus the square root of the number of measurements in the training set. The coefficient of WDR was chosen as equal to 0.01. The number of training cycles was chosen as equal to 400. The ANN input structure included 4 past pressure/flow values (i.e., *LagSize*=4) and the TofD and DofW indices associated with the forecasting horizon. Note that the use of the aforementioned rule of thumb, coefficient of WDR, number of training cycles and input structure was found to ensure (after a series of preliminary tests – not shown here) that, for all the analysed signals, the resulting ANN prediction models were able to closely approximate the training sets whilst allowing good generalisation performances. Finally, note that the *BIS parameters learning* subsystem (see Figure 2 and Section 3.4.8) was only used in the last two analyses carried out on this case study.

4.2.3 Testing of detection capabilities on real-life events

In the analysis carried out in this section, the ERS was used for detecting the real-life events occurred in the DMA being studied. Here, the value of the *alarm inactivity time* parameter was set as equal to 1 week. The value of the user defined detection threshold for raising the alarms (i.e., λ) was set as equal to 0.5 (see Section 3.4.7.3). With these settings the ERS raised a total of 14 detection alarms in the considered eight month test period.

In order to test and evaluate the methodology's event detection capabilities, the WMS and CCs data provided by the YWS were used as source of information about the real-life events occurred in the DMA being studied. The WMS records taken into consideration were the ones tagged as 'Main Repair' (MR), whilst the CC records were the ones tagged as 'Pressure/Flow', 'Burst/Leak' and 'No Water'.

Bearing in mind the above, 2 MRs and 2 clusters of CCs (i.e., several CCs received during a particular day) that were not accompanied by any MR were recorded during the test period. The former were considered as an indication of the occurred burst events. The latter were considered as an indication of the occurred pressure/flow anomalies whose exact cause is uncertain (e.g., illegal water usage, unusual system activity, operational DMA changes, etc.) but for which a record existed. However, it is important to stress that the likely start and end date-time (and hence duration) of these events could not be determined based on the available information. For example, the MR records only had a creation date (without time) attached to them. Thus, they did not provide information about either the likely start date-time of the relevant burst events or the actual completion date-time of the MRs (i.e., likely end date-time of the relevant burst events). For this reason, a careful visual inspection of the data was carried out. This inspection involved looking at the trends of the pressure and flow signals coming from the DMA being studied and attempting to visually recognise the likely start and end date-time of the occurred burst events and pressure/flow anomalies. Consequently, it was possible to establish the following. The first burst event lasted about 5 days. The second burst event lasted about 4 weeks. Each of the two pressure/flow anomalies lasted about one day.

The visual inspection of the data carried out as outlined above, however, not only provided the timing information of the aforementioned events but also allowed the identification of other real-life events occurred in the considered eight month test period. These additional real-life events included 3 sensor failure events (i.e., characterised by a distinctive trend signature), and 2 other visible pressure/flow anomalies (i.e., large and short-lived abnormal increases in pressure) that were, however, not accompanied by any CC or MR record (i.e., did not impact the customers). With regard to the duration of these additional real-life events, it was observed that two of the sensor failure events lasted less than 1 week, whilst the third one lasted about 12 days. The first visible abnormal increase in pressure lasted about 1 day, whilst the second only a few hours.

All the above events (a total of 9) formed the set of real-life events against which the 14 detection alarms raised by the ERS were compared in order to see if a correlation existed (i.e., genuine alarms) or not (i.e., false alarms). The following results were obtained from this comparison:

- 6 alarms were correlated with the 2 recorded MRs (i.e., burst events). Note that, since the *alarm inactivity time* parameter was set as equal to 1 week, five of these alarms were related to the second burst event;
- 2 alarms were correlated with the 2 clusters of recorded CCs (i.e., pressure/flow anomalies for which a record existed);
- 4 alarms were correlated with the 3 sensor failure events. Two of these alarms were related to the same sensor failure event (bearing in mind the *alarm inactivity time* parameter used);
- 2 alarms were correlated with the 2 other visible pressure/flow anomalies that did not impact the customers.

In view of the above results, the following has to be stressed in order to avoid potential confusion. The ERS cannot automatically distinguish between different types of events. The aforementioned comparison only involved looking for matches between the start date-time and duration of the 9 real-life events (established after visual inspection of the data) and the 14 detection alarms raised by the ERS.

The results obtained show that the ERS detected all the real-life events occurred during the considered eight month test period in the DMA being studied. Furthermore, it did not generate any false alarm. All this provides evidence that the ERS methodology can detect real-life pipe bursts and other events both efficiently and reliably. Allowing for this, two examples that refer to the detection of different types of real-life events are provided below.

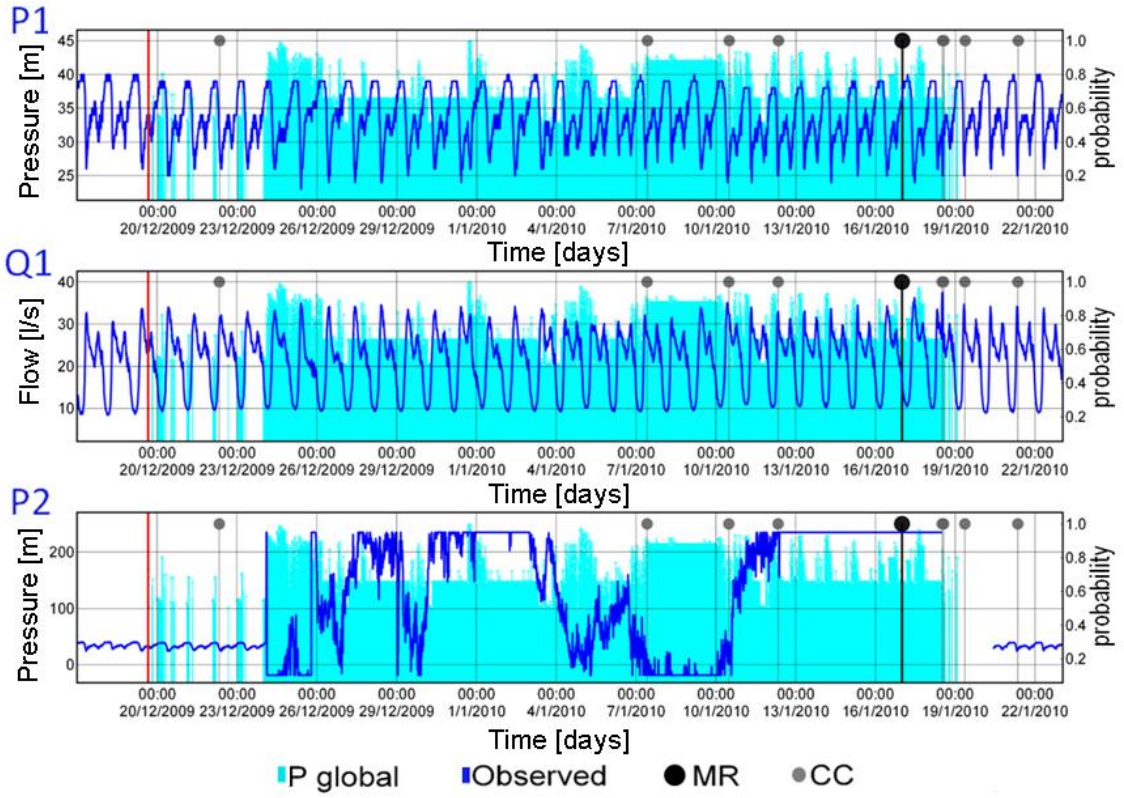


Figure 22. Results obtained by the Event Recognition System when a real-life burst event was experienced.

Figure 22 shows the results obtained by the ERS when the burst event that lasted about 4 weeks was experienced. The figure is divided into three parts each representing one of the three signals coming from the sensors deployed in the DMA being studied (one pressure signals P1 and one flow signal Q1 from the sensors located at the DMA inlet, and one pressure signals P2 from the sensor located at the critical DMA point). For each signal, the vertical axis on the right refers to the DMA level event occurrence probability (i.e., a value between 0 and 1). The vertical axis on the left refer to the

observed pressure or flow. The leftmost vertical red line indicates the time the first alarm was raised (i.e., at 15:45 on the 19th of December 2009). Note that, out of the five related alarms raised for this event, only the first alarm is shown in the figure. The vertical light blue bars (i.e., P global) indicate the DMA level event occurrence probability (i.e., only if greater than the user defined detection threshold - i.e., greater than 0.5) at every time step (i.e., every 15 minutes) following the alarm. In the figure, the vertical black line with a big circle and the vertical grey lines with small circles indicate the recorded MR and CCs, respectively.

As it can be observed from Figure 22, a detection alarm was generated two days before the first CC was received, while the MR was carried out about one month later (probably only after receiving three further CCs). The flow rate estimated by the ERS for this burst event was 2.3 litres per second. Note that the DMA level event occurrence probability changed with time and it drastically increased when the pressure sensor at the critical DMA point (i.e., P2) started recording abnormal values that were a long way outside its usual range (probably due to a fault). Also note that the ERS continued working even after this sensor stopped recording data. This fact highlights the ERS robustness (ability to continue to function even under less than ideal conditions), which is a consequence of the simultaneous analysis of multiple (pressure and flow) DMA signals.

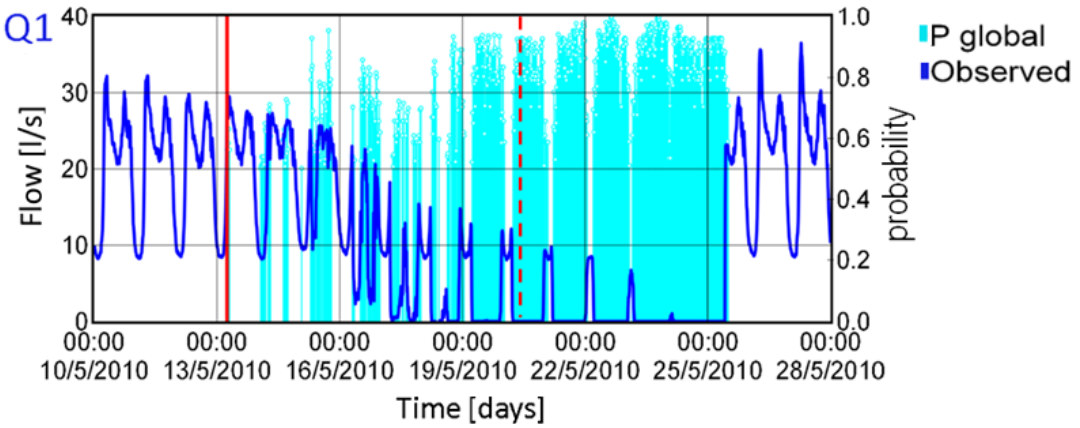


Figure 23. Results obtained by the Event Recognition System when a real-life sensor failure event was experienced.

Figure 23 shows the results obtained when a second type of real-life event (i.e., a sensor failure) was experienced. Note that only the DMA signal responsible for generating the alarm (as it will be explained later) is shown in this figure (i.e., Q1). In contrast to Figure 22, this figure displays the two related alarms raised for this event. The leftmost solid vertical red line represents the first alarm (i.e., at 06:00 on the 13th of May 2010). The dashed vertical red line represents the second alarm (i.e., at 06:00 on the 20th of May 2010 – exactly after one week). For both the displayed alarms (different time snapshots), the corresponding tabular results (i.e., alarm start time, DMA level event occurrence probability, Sensor level event occurrence probabilities, and estimates of the deviations from the statistically-based NOPs of the pressure/flow signals being analysed) are provided in Table 10.

Table 10. Tabular results obtained by the Event Recognition System when a real-life sensor failure event was experienced.

Alarm start time	DMA level event occurrence probability	Signal	Signal level event occurrence probability	Deviation from the statistically-based NOPs	Deviation from the statistically-based NOPs as % of the average pressure/flow
13th May 2010 06:00	0.56	P1	0.1	0.5 m	1.4
		Q1	0.56	-1.2 l/s	6.5
		P2	0	0.0 m	0
20th May 2010 06:00	0.92	P1	0	0.2 m	0.6
		Q1	0.99	-18.3 l/s	99
		P2	0	0.1 m	0.3

From Figure 23 and Table 10 it is possible to observe that the ERS raised a detection alarm with a DMA level event occurrence probability as equal to 0.56 at the very early stage of the sensor failure event (i.e., variation in flow of just 1.2 litres per second). Furthermore, it is possible to observe that the DMA level event occurrence probability increased with time as the sensor started drifting more and more until it failed completely. These observations confirm that the ERS methodology has the potential to detect precursor features of anomalies and can be used for monitoring their development over time.

The above detection example also highlights another important feature of the ERS methodology. This is as follows. Although the ERS cannot automatically distinguish

between different types of events, its output information may help determining the likely root cause of a raised alarm (i.e., performing a first diagnosis of the incident occurring). Indeed, by taking into consideration the information reported in Table 10 it is possible to infer, even without observing the trends of the DMA signals, that the most likely cause of the raised alarms was the failure of the flow sensor. This is because Table 10 shows that a drastic (and somehow unrealistic) flow decrease was recorded over the two relevant time snapshots. Additionally, Table 10 also shows that the Signal level event occurrence probabilities at both relevant time snapshots and for both pressure signals were very low, whilst those for the flow signal were high. This fact suggests that the flow sensor only was responsible for raising the alarm at the DMA level. Having said this, it is important to stress that the fact that the ERS output information helps determining the likely root cause of an alarm can, in turn, support alarm and other (e.g., sensor) management decisions.

4.2.4 Testing of detection capabilities on Engineered Events

In the analysis carried out in this section, the ERS methodology was tested and verified on the EEs where pipe burst events were simulated by opening a number of fire hydrants in the DMA being studied. This way, all "burst" related information (i.e., timing, and approximate flow rate) was known which, in turn, enabled the evaluation of the methodology's capabilities based on comparisons between the ERS detection times and the corresponding actual hydrant opening times. In addition to this, the Receiver Operating Characteristics (ROC) graph (Egan, 1975) and related performance metrics – i.e., Area Under the Curve (AUC) (Hanley and McNeil, 1982) were used too.

ROC graphs have long been used in signal detection theory to depict the performance of a classifier. They are useful for visualizing, organizing and selecting classifiers based on their performance. The AUC of a classifier is equal to the value of the Wilcoxon test of rank (Hanley and McNeil, 1982). It represents the probability that the classifier will score a randomly drawn positive sample higher than a randomly drawn negative sample. The closer the AUC value to 1 the better.

Generally speaking, given the DMA level BIS (i.e., the classifier) and an event, there are four possible outcomes at any particular time step. These can be categorised as follows: (i) True Positive - i.e., the event is occurring and the DMA level BIS classifies

it as occurring, (ii) False Negative - i.e., the event is occurring and the DMA level BIS classifies it as not occurring, (iii) True Negative - i.e., the event is not occurring and the DMA level BIS classifies it as not occurring, and (iv) False Positive - i.e., the event is not occurring and the DMA level BIS classifies it as occurring. Bearing this in mind, given the DMA level BIS and a set of events, a two-by-two confusion matrix (also known as a contingency table), representing the disposition of the possible outcomes, can be constructed. This matrix forms the basis for building a ROC graph and calculating related performance metrics. Figure 24 shows a confusion matrix and the equations of four common performance metrics that can be calculated from it.

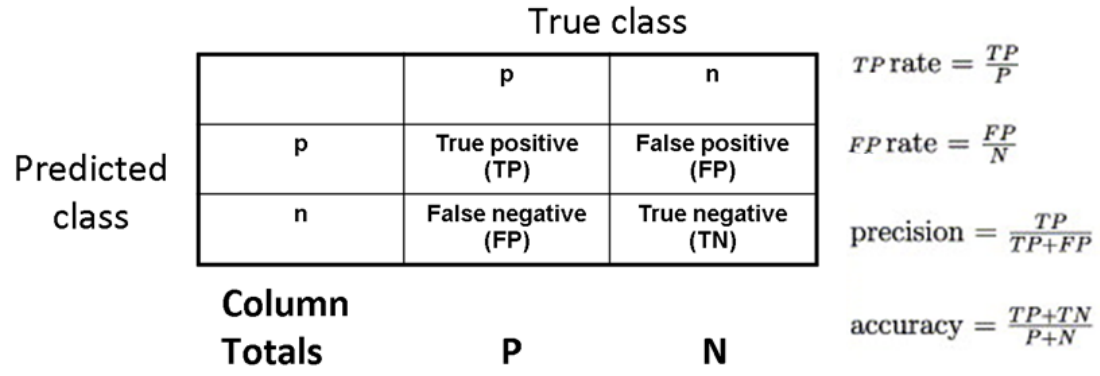


Figure 24. Confusion matrix and common performance metrics.

Evaluation of detection capabilities based on comparison between Event Recognition System detection and actual hydrant opening times

A total of 9 EEs were carried out during the considered three month test period: 3 events in July 2009, 3 events in August 2009, and 3 events in March 2010. Note that any single EE lasted one day maximum and that different EEs were sometimes carried out during the same week. Because of this, the *alarm inactivity time* parameter had to be chosen carefully. Choosing it as one week, for example, would have implied that only one alarm could have been raised for multiple EEs carried out during the same week. Thus, for the purposes of the analysis that evaluated the methodology's event detection capabilities based on comparisons between the ERS detection times and the corresponding actual hydrant opening times, it was set as equal to 1 day. The value of the user defined detection threshold for raising the alarms was set as equal to 0.5.

Table 11. Engineered Events schedule and detection results.

	Alarm start time	Hydrant opened time	Hydrant closed time	Actual flow rate [l/s]	% DMA average inflow	Estimated flow rate [l/s]
July 2009	20/07/2009 - 08:00	20/07/2009 - 08:00	21/07/2009 - 08:00	2	11	1.9
	22/07/2009 - 06:15	22/07/2009 - 06:05	23/07/2009 - 08:00	2	11	2
	24/07/2009 - 16:15	24/07/2009 - 16:00	25/07/2009 - 16:00	2	11	2
August 2009	17/08/2009 - 08:30	17/08/2009 - 08:20	18/08/2009 - 07:05	3	16	2.7
	21/08/2009 - 08:00	21/08/2009 - 07:20	22/08/2009 - 08:05	1	5	0.8
	26/08/2009 - 08:00	26/08/2009 - 07:55	27/08/2009 - 07:00	1 to 3	5-16	2.2
March 2010	**01/03/2010 - 21:00**	01/03/2010 - 09:10	02/03/2010 - 07:50	2	11	2.1
	02/03/2010 - 21:00	02/03/2010 - 08:00	03/03/2010 - 07:10	2	11	2
	15/03/2010 - 07:45	15/03/2010 - 07:20	16/03/2010 - 07:15	2	11	2

Table 11 shows the ERS detection times and the corresponding actual hydrant opening and closing times for all the EEs. The actual flow rates and their corresponding values estimated by the ERS are also given together with the values of the actual flow rate expressed as percentage of the average DMA inflow. It can be observed from this table that all the EEs were successfully detected. The underlined alarm start times refer to those events that were detected at the best possible time (within the 15 minute sampling rate). Alarm start times in normal text font refer to those events that were detected with a delay not greater than 1 hour, whilst alarm start times in bold refer to those events that were detected with delays longer than 1 hour. The main reason for the comparatively long delay in detection for the EE carried out on the 1st of March 2010 is that the DMA pressure signals were only mildly affected by this event (most likely due to the location of the hydrant opened being far away from the pressure sensors). As a result, until late evening that day (when the legitimate variability in pressure/flow became less accentuated), the pieces of event occurrence evidence generated by the three ERS analysis subsystems were not sufficient to produce a DMA level event occurrence probability greater than the user defined detection threshold (i.e., greater than 0.5) and hence trigger an alarm. With regard to the comparatively long delay in detection for the EE carried out on the 2nd of March 2010, on the other hand, note that only 10 minutes separate this EE from the EE carried out on the 1st of March 2010. Therefore, with the relevant value for the *alarm inactivity time* parameter, it was not possible to correctly identify its start time, although the ERS generated a DMA level event occurrence

probability of 0.68 at 08:00 that day (i.e., detection at the best possible time). The results obtained confirm that the ERS methodology can efficiently detect burst events. Furthermore, they show that timely detections were achieved (i.e., within one hour for all but one of the simulated burst events).

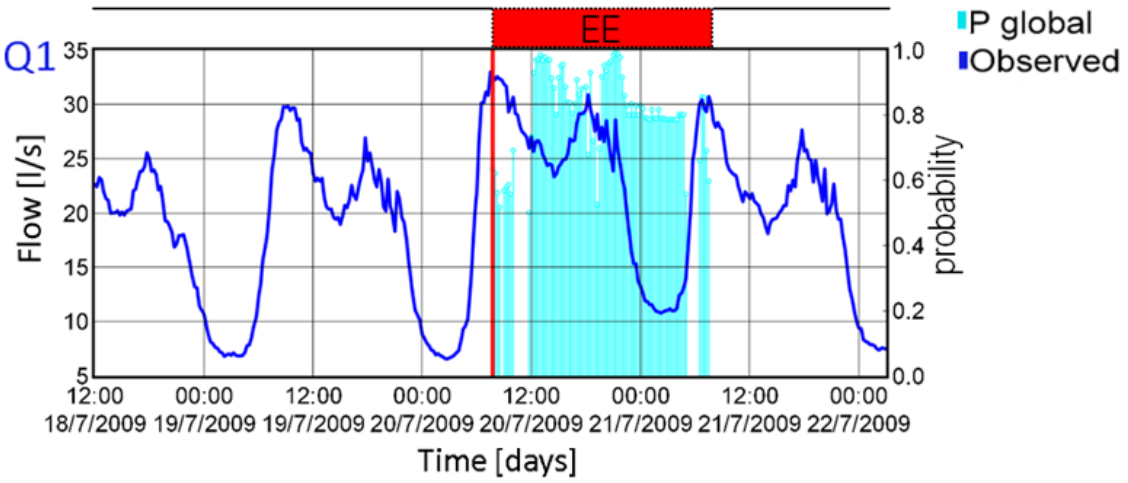


Figure 25. Results obtained by the Event Recognition System when an Engineered Event was carried out.

As an example, Figure 25 shows the results obtained by the ERS when the first EE was carried out. Note that only the trend of the flow signal (i.e., Q1) is shown in the figure. Similarly to Figure 22, the leftmost vertical red line indicates the time the alarm was raised and the vertical light blue bars indicate the DMA level event occurrence probability (i.e., only if greater than 0.5) at every time step following the alarm. Here, however, the red filled rectangle in the top part of the figure indicates the EE duration.

Evaluation of detection capabilities based on Receiver Operating Characteristics graphs and Area Under the Curve values

For the purposes of the analysis that evaluated the capabilities of the ERS methodology by using the ROC graph and the AUC values, the value of the user defined detection threshold for raising the alarms was varied in the interval between 0 and 1. The value of

the *alarm inactivity time* was set as equal to 15 minutes (i.e., all the DMA level global detection probabilities greater than the user defined detection threshold raised detection alarms). This is because the true and false positive rates are obtained by comparison, at every time step, between the status of the hydrant (i.e., opened/closed) and the output of the DMA level BIS (i.e., DMA level event occurrence probability greater/smaller than the user defined detection threshold).

Here, each vector containing the DMA level event occurrence probabilities returned by the ERS, at every time step, during one of the three months analysed was considered as a single set of data (i.e., reference set). Then, because the exact opening and closing times of the hydrants were known, a vector consisting of a binary value (i.e., hydrant opened/hydrant closed) for each time step (i.e., target vector) was attached to each of the three reference sets. Once this was done, a ROC curve for each reference set (i.e., July 2009, August 2009, and March 2010) was built. The resulting curves are shown in Figure 26.

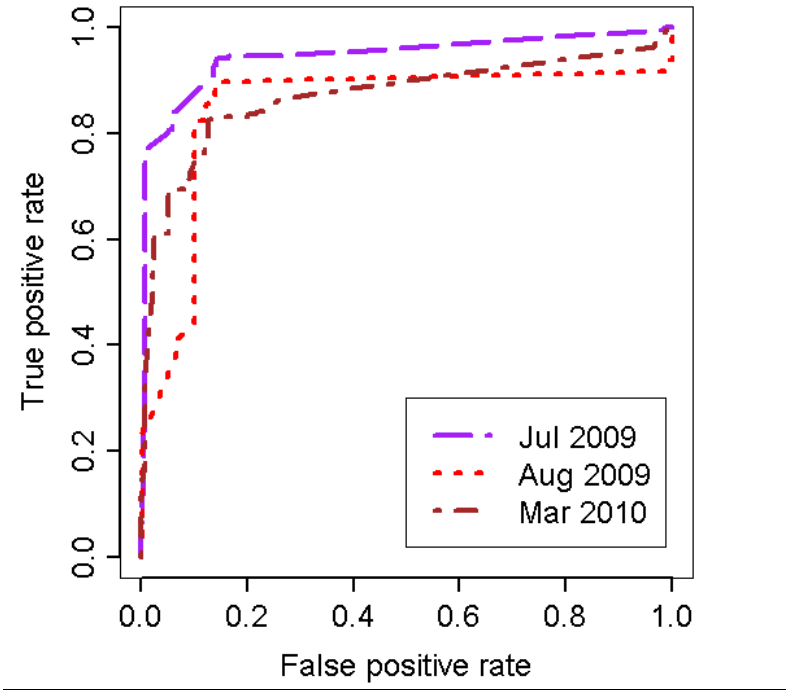


Figure 26. Receiver Operating Characteristics curves for the July 2009, August 2009 and March 2010 reference sets.

Similarly to the above, the ROC curve for the single three month data set obtained by merging the three reference sets together was also built. This curve is shown in Figure 27. Note that Figure 27 also shows the values of the detection threshold printed at the corresponding curve positions. Furthermore the parameterisation is visualised by colouring the curve according to the values of the detection threshold.

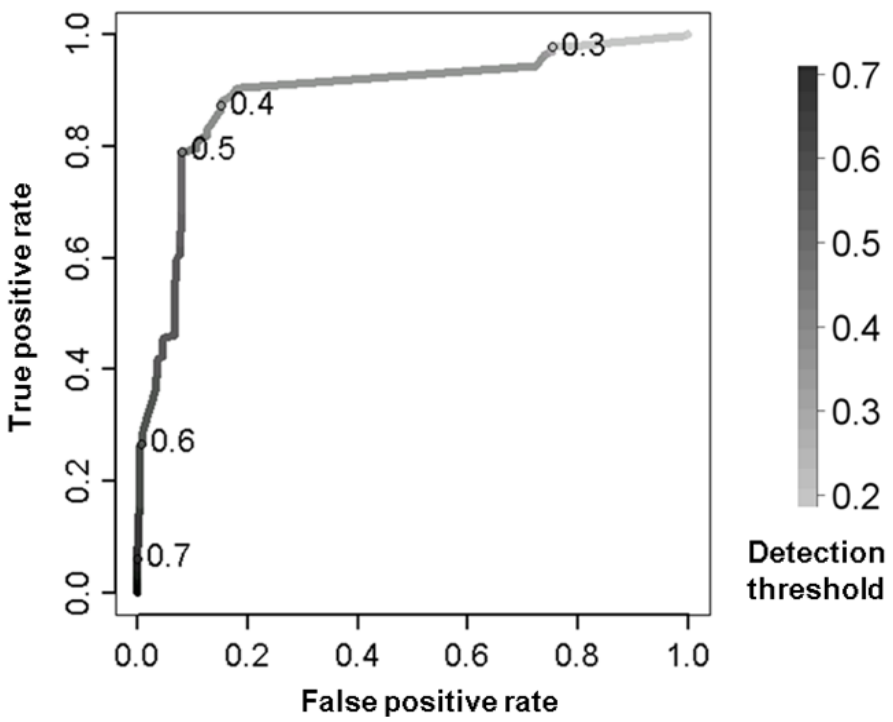


Figure 27. Receiver Operating Characteristics curve for the three month merged data set and detection threshold values.

The ROC curves shown in Figure 26 can be seen as two-dimensional depictions of the DMA level BIS classification performance (i.e., trade-off between true and false positive rates) on each reference set. The ROC curve shown in Figure 27, on the other hand, can be seen as a two-dimensional depiction of the DMA level BIS "average" classification performance over the whole test period. Allowing for this, it is possible to reduce the DMA level BIS classification performance to a single scalar value representing expected performance by calculating the area under the ROC curve (i.e., AUC). According to this, the calculated AUC values for the three reference sets are: (July 2009) 0.94, (August 2009) 0.84, and (March 2010) 0.87. The calculated AUC

value for the three month merged data set is 0.88. These values together with the ROC curves shown in Figure 26 and Figure 27 provide strong evidence of the good classification performance of the DMA level BIS.

4.2.5 Investigating the optimal detection threshold selection issue

In the analysis carried out in this section, the issue concerning the optimal selection of the user defined detection threshold value to be used for raising the alarms was investigated.

Here, the three month merged data set obtained as described in Section 4.2.4 was used to compute the Accuracy and Precision performance metrics. As it is shown in Figure 24, Accuracy is computed as the ratio between: (i) the sum of true positives and true negatives, and (ii) the sum of total positives and total negatives. Precision is computed as the ratio between: (i) true positives, and (ii) the sum of true positives and true negatives. Figure 28 shows a graphical representations of the resulting Accuracy (Figure 28a) and Precision (Figure 28b) performance metrics according to the detection threshold values.

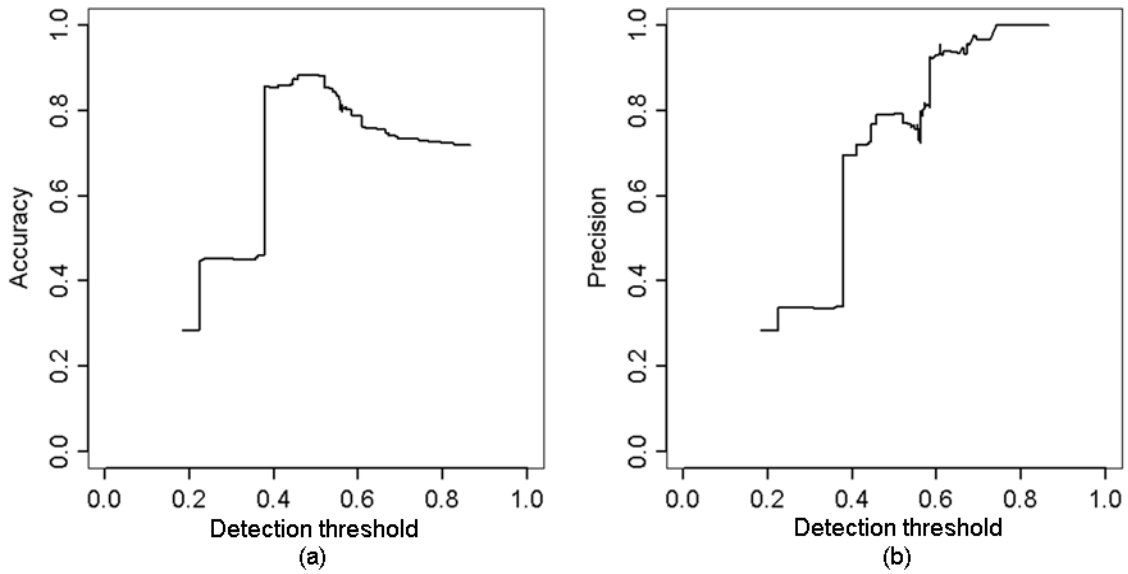


Figure 28. Accuracy (a), and Precision (b) according to the detection threshold values for the three month merged data set.

By observing Figure 27 and Figure 28, a few considerations about the choice of the best detection threshold value to be used for raising the alarms can be made. On the one hand, it can be observed from Figure 27 that choosing a low value for the detection threshold (e.g., below 0.4) has the potential to improve the methodology's sensitivity to the events (i.e., true positive rate close to one). However, this choice may lead to raising too many alarms, including the false ones. On the other hand, choosing a high value for the detection threshold (e.g., above 0.6) has the potential to improve the methodology's reliability (i.e., false positive rate close to zero). However, this choice may lead to detecting fewer events only. Bearing this in mind, for the particular case study analysed here, it is possible to state that a reasonable choice for the value of the detection threshold would have been 0.4. This is because Figure 28 shows that both Accuracy and Precision greatly improve at this value. Furthermore, Figure 27 shows that the true positive ratio grows faster than the false positive ratio between the detection threshold value as equal to 0.5 and the detection threshold value as equal to 0.4. To support these observations, note that, when the first of the March 2010 EEs was carried out (i.e., the one detected with an eleven hour and forty five minute delay – see Table 11), a detection threshold value as equal to 0.4 would have led to the detection of that simulated burst event at the best possible time (not shown here). This example highlights the fact that a more suitable selection of the detection threshold value could have further improved the performances of the ERS methodology.

A further consideration to be made here is as follows. When the parameters in the CPTs of the DMA level BIS are based on domain experts' knowledge (i.e., the *BIS parameters learning* subsystem is not used - see Section 3.4.8), choosing a value for the detection threshold as equal to 0.5 does not imply that an alarm is raised when there is more than a 50% chance that an event is actually occurring. This is because the 0.5 value does not represent a calibrated probability but only a score for discriminating between negative and positive instances of an event occurring.

Based on the above, it can be concluded that, although a detection threshold value as equal to 0.5 allows achieving good detection performances (see also the results of the tests on the real-life events presented in Sections 4.2.3 and 4.3.3), it does not represent the optimal choice. Therefore the selection of the most suitable detection threshold value has to be ultimately based on the operator's decision and remains subject to the

operator's confidence and expertise. In this scenario, the analysis carried out in this section can be used to help the operator to make more informed decisions.

4.2.6 Evaluating the benefits of the Expectation Maximisation strategy

In the analysis carried out in this section, the benefits of the EM strategy for calibrating the parameters in the CPTs of the DMA level BIS were evaluated (see also Section 3.4.8 and Appendix E). In order to achieve this, the capabilities of the ERS with the parameters in the CPTs of the DMA level BIS based on domain experts' knowledge were compared with those of the ERS with the parameters in the CPTs of the DMA level BIS calibrated by using the EM strategy and information about some of the EEs carried out. Similarly to what done in Section 4.2.4, here, the capabilities of the ERS methodology were evaluated by: (i) performing, for the two relevant cases considered, a comparison between the ERS detection times and the corresponding actual hydrant opening times, and (ii) using ROC graphs.

Evaluation of detection capabilities based on comparison between Event Recognition System detection and actual hydrant opening times

For the purposes of the analysis carried out here, the CPT parameters were calibrated by making use of the information about the start and end times of the 6 EEs carried out in July 2009 and August 2009. The *alarm inactivity time* parameter was set as equal to 1 day. The value of the detection threshold was set as equal to 0.5.

Table 12 shows the detection times obtained for the two relevant cases considered and the corresponding actual hydrant opening and closing times for all the EEs. Note that the detection times obtained by the ERS with the parameters in the CPTs of the DMA level BIS based on domain experts' knowledge and the actual hydrant opening and closing times are as in Table 11 (see Section 4.2.4). However, they are repeated here for convenience. Further note that, similarly to Table 11, in Table 12 the underlined alarm start times refer to those events that were detected at the best possible time. Alarm start times in normal text font refer to those events that were detected with a delay not greater

than 1 hour. Alarm start times in bold refer to those events that were detected with delays longer than 1 hour.

As it can be seen from Table 12, in both relevant cases considered, all the EEs were successfully detected. However, when the calibration procedure was used, the detection speed improved significantly. Given that the CPT parameters were calibrated by using the July-August 2009 EEs start and end times information, it is not unexpected that the EEs carried out during these two months were timely detected (i.e., the third of the July 2009 EEs with a 15 minute delay only, and the other EEs at the best possible time). This said, the significant detection speed improvement is evident when the EEs carried out in March 2010 are considered (i.e., set of EEs not used for calibration). Indeed, the first of these EEs was detected 11 hours and 30 minutes earlier than in the case where the calibration procedure was not used. Additionally, the third of these EEs was detected 15 minutes earlier (resulting in detecting it at the best possible time). Finally, a "detection speed improvement" can also be observed for the second of the March 2010 EEs. With regard to this particular EE, however, the following has to be noted. On the 2nd of March 2010 at 08:00 the ERS generated a detection probability of 0.68 and 0.74 in the case of CPT parameters based on domain experts' knowledge and calibrated by using the EEs information, respectively (see also Section 4.2.4). This means that in both cases a detection at the best possible time could have been achieved if this EE had not been carried out the day immediately after the first of the March 2010 EEs, or if a different (shorter) value for the *alarm inactivity time* parameter had been used.

Table 12. Engineered Events time schedule and Event Recognition System detection times for the two relevant cases considered.

	Alarm start time CPT parameters based on domain experts' knowledge	**Alarm start time calibrated CPT parameters**	**Hydrant opened time**	**Hydrant closed time**
July 2009	20/07/2009 - 08:00	20/07/2009 - 08:00	20/07/2009 - 08:00	21/07/2009 - 08:00
	22/07/2009 - 06:15	22/07/2009 - 06:15	22/07/2009 - 06:05	23/07/2009 - 08:00
	24/07/2009 - 16:15	24/07/2009 - 16:15	24/07/2009 - 16:00	25/07/2009 - 16:00
August 2009	17/08/2009 - 08:30	17/08/2009 - 08:30	17/08/2009 - 08:20	18/08/2009 - 07:05
	21/08/2009 - 08:00	21/08/2009 - 07:30	21/08/2009 - 07:20	22/08/2009 - 08:05
	26/08/2009 - 08:00	26/08/2009 - 08:00	26/08/2009 - 07:55	27/08/2009 - 07:00
March 2010	**01/03/2010 - 21:00**	01/03/2010 - 9:30	01/03/2010 - 09:10	02/03/2010 - 07:50
	02/03/2010 - 21:00	**02/03/2010 - 9:30**	02/03/2010 - 08:00	03/03/2010 - 07:10
	15/03/2010 - 07:45	15/03/2010 - 07:30	15/03/2010 - 07:20	16/03/2010 - 07:15

Evaluation of detection capabilities based on Receiver Operating Characteristics graphs

For the purposes of the analysis carried out here, the value of the *alarm inactivity time* was set as equal to 15 minutes. The value of the detection threshold was varied in the interval between 0 and 1.

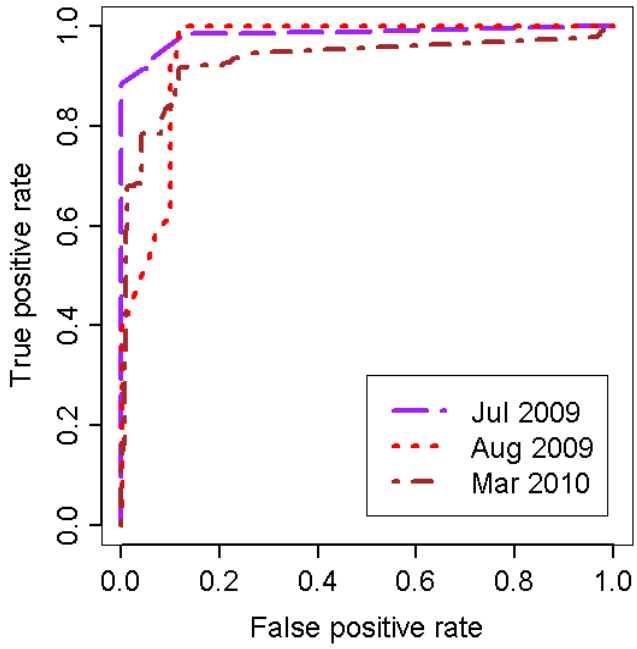


Figure 29. Receiver Operating Characteristics curves for the July 2009, August 2009 and March 2010 reference sets when the Conditional Probability Tables parameters were calibrated by using information about the Engineered Events.

Here, the classification performances of the DMA level BIS with CPT parameters based on domain experts' knowledge were compared with those of the DMA level BISs with CPT parameters calibrated by using the EEs information. Because information from a small number of EEs was available, the Three-Fold Cross-Validation technique (see – e.g., Devijver and Kittler, 1982) was used here to circumvent the limited data (i.e., 'event cases') availability. This involved calibrating the CPT parameters by using, in turn, information about the EEs carried out during two of the three months considered (i.e., August 2009 and March 2010, July 2009 and March 2010, and July 2009 and

August 2009) and evaluating the resulting DMA level BIS classification performance on the remaining month (i.e., July 2009, August 2009, and March2010).

Following the above, similarly to what done in Section 4.2.4, a target vector was attached to each of the resulting reference sets and the corresponding ROC curves were built. These are shown in Figure 29 The calculated AUC values for the three ROC curves shown in Figure 29 are as follows: (July 2009) 0.98, (August 2009) 0.95, and (March 2010) 0.93. These AUC values are significantly higher than the corresponding values obtained when the DMA level BIS with CPT parameters based on domain experts' knowledge was used (see Section 4.2.4).

The above ROC curves and AUC values alone, however, do not allow making conclusion about the calibrated DMA level BISs superiority. This is because, in order to compare classifiers, it is necessary to (Fawcett, 2006): (1) use a suitable method for averaging the ROC curves, and (2) derive a measure of variance. With regard to the former, the Vertical Averaging technique (see – e.g., Fawcett, 2006) was used here. It involved taking vertical samples of the ROC curves for fixed false positive rates and averaging the corresponding true positive rates. With regard to the latter, Box plots (see – e.g., Tukey, 1977) were used to visualise the classification performance variability across three months studied (i.e., Three-Fold Cross-Validation runs).

Figure 30 shows the results of applying the above procedure to the ROC curves shown in Figure 26 (i.e., obtained by using CPT parameters based on domain experts' knowledge) and Figure 29 (i.e., obtained by using calibrated CPT parameters). The resulting ROC curves represent the average classification performance (across the three months studied) of the relevant DMA level BISs. The Box plots show the classification performance variability. It is possible to observe from this figure that performance was similar for low false positive rates (i.e., 0.05). That is, in both cases, reliable true positive classifications (i.e., an alarm is raised and the event is actually occurring) were made with strong evidence. However, for higher false positive rates the DMA level BISs with calibrated CPT parameters performed better than the DMA level BIS with CPT parameters based on domain experts' knowledge and also showed less variability. The results of this analysis show clearly that the detection reliability and effectiveness

of the ERS methodology can be improved if information about past events is used for calibrating the parameters in the CPTs of the DMA level BIS.

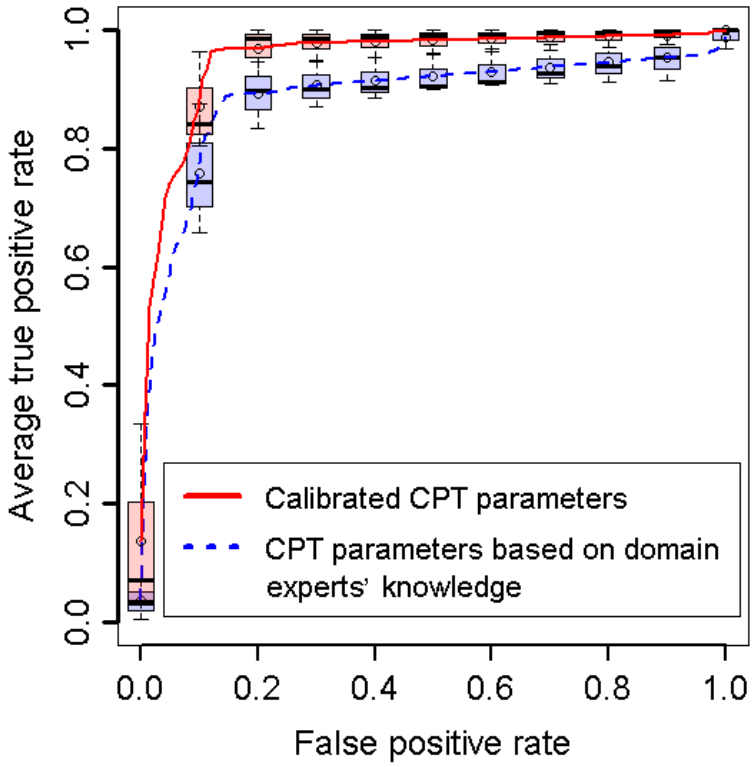


Figure 30. Performance comparison between the District Metered Area level Bayesian Inference System with Conditional Probability Table parameters based on domain experts' knowledge, and with calibrated Conditional Probability Table parameters.

The improved detection reliability and effectiveness, however, are not the only benefit yielded by the use of the EM strategy. In fact, the calibration of the CPT parameters has the further benefit of rendering the selection of the most suitable detection threshold value less reliant on the operator's confidence and expertise (see Section 4.2.5). This is because the CPT parameters are 'adjusted' in such a way that the DMA level BIS output (i.e., DMA level event occurrence probability) better approximate an accurate probability estimate (i.e., calibrated probability). Thus, generally speaking, choosing a value for the detection threshold as equal to 0.5 should guarantee that an alarm is raised when there is more than a 50% chance that an event is actually occurring.

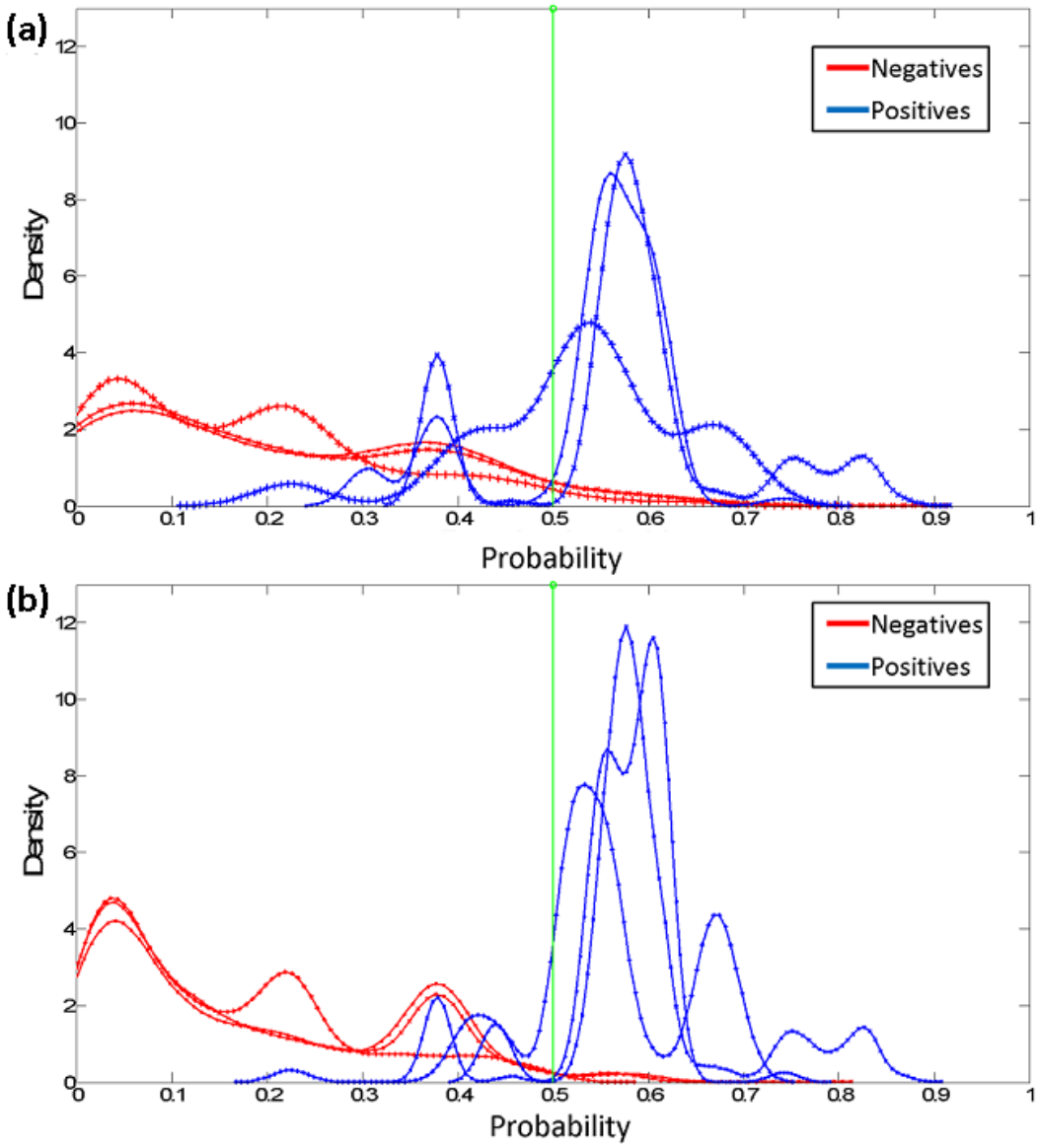


Figure 31. Density estimates for each relevant reference set and for each class. Conditional Probability Table parameters based on domain experts' knowledge (a), and calibrated Conditional Probability Table parameters (b).

To support what stated above, Figure 31 shows how well, for the two relevant cases considered, the 0.5 detection threshold value enabled discriminating between the two possible classes/scenarios (i.e., negative = the event was not occurring; positive = the event was occurring). The figure is divided into two parts. Figure 31a refers to the DMA level BIS with CPT parameters based on domain experts' knowledge. Figure 31b refers to the DMA level BIS with CPT parameters calibrated by using the EEs information. In

each part of the figure the density estimates for each relevant reference set and for the two possible classes/scenarios are shown. It can be observed from this figure that when the calibrated CPT parameters were used, the 0.5 detection threshold value enabled to better discriminate between the negative and positive event occurrence instances.

Note that the density estimates shown in Figure 31 were obtained by using a Kernel Density Estimation method (see – e.g., Rosenblatt, 1956; Parzen, 1962). Generally speaking, such methods involve solving a nonparametric data smoothing problem where inferences about the population are made based on a finite data sample. In this particular application: (i) the density estimate for each relevant reference set and for each possible class/scenario (i.e., reference-class set) was based on a normal (Gaussian) kernel function, (ii) a window parameter (i.e., width), function of the number of points in the reference-class set, was used, and (iii) the density was evaluated at a sufficiently high number of equally spaced points covering the range of the data in the reference-class set.

4.2.7 Evaluating the performance of different classifiers

In this section the issue concerning the use of different classifiers was investigated. Specifically, this was done by comparing the classification performances of the BN-based DMA level BIS with those of classifiers based on an ANN. The aim was to explore the possibility of using, in the DMA level BIS module, an ANN-based classifier to replace or complement the BN-based classifier.

Here, similarly to what done in Section 4.2.6, the ANN-based classifiers were trained and evaluated by using the Three-Fold Cross-Validation technique and information about the EEs carried out during the considered three month test period. Target vectors were attached to the resulting reference sets. The corresponding ROC curves were built. The Vertical Averaging technique was used for averaging the ROC curves obtained. Box plots were used to visualise the classification performance variability.

Figure 32a shows a comparison between the classification performance of the DMA level BIS classifier (BN-based) with CPT parameters based on domain experts' knowledge and the performance of the ANN-based classifiers. Figure 32b shows a comparison between the detection performance of the DMA level BIS classifiers with calibrated CPT parameters and the performance of the ANN-based classifiers. It is

possible to observe from these figures that performances were similar for low false positive rates. However, regardless of the CPT parameters used in the DMA level BIS, the BN-based classifiers performed better than the ANN-based classifiers for higher false positive rates. They also showed less variability. These results demonstrate clearly that the BN-based classifiers outperform their ANN-based counterpart, thereby supporting the choice of using BNs for effectively inferring the probability of an event occurrence.

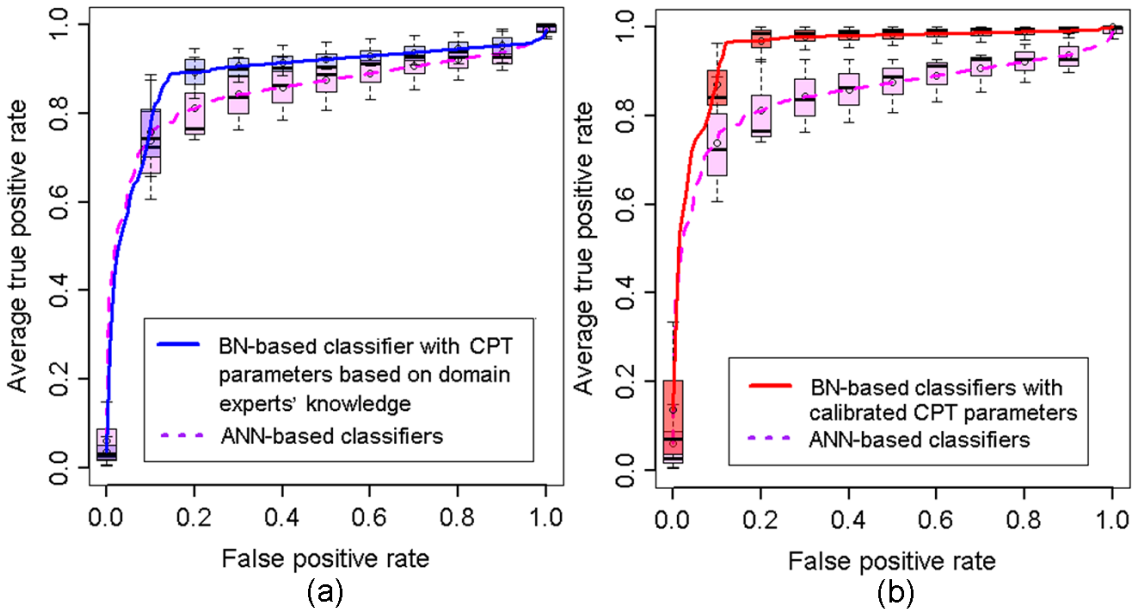


Figure 32. Comparison of classification performance. District Metered Area level Bayesian Inference System with Conditional Probability Table parameters based on domain experts' knowledge and Artificial Neural Network -based classifiers (a), and District Metered Area level Bayesian Inference System with calibrated Conditional Probability Table parameters and Artificial Neural Network -based classifiers (b).

4.2.8 Section summary

In this section, the capabilities of the methodology that enables performing event detection in normally-sampled DMAs have been tested and evaluated on real-life pipe bursts/other events and simulated pipe burst events. Additionally, the following issues have been investigated: (i) the optimal selection of the detection threshold value to be

used for raising the alarms, (ii) the benefits of the EM strategy for calibrating the parameters in the CPTs of the DMA level BIS, and (iii) the potential for using ANN based classifiers in the DMA level BIS module. The main findings from these tests/investigations are briefly summarised below.

The ERS successfully identified all the real-life events and the simulated pipe burst events in a timely and reliable manner. Furthermore, it showed the ability to identify precursor features of anomalies and to provide information that could be used for determining the likely root cause of an event.

Although a detection threshold value as equal to 0.5 allowed achieving good detection performances when the parameters in the CPTs of the DMA level BIS were based on domain experts' knowledge, it did not represent the optimal choice. In the light of the analysis performed, it was therefore concluded that the selection of the most suitable detection threshold value has to be ultimately based on the operator's decision and remains subject to the operator's confidence and expertise. Bearing this in mind, when the CPT parameters were calibrated by using information about confirmed past events occurred in the DMA being studied, the 0.5 detection threshold value allowed to better discriminate between positive and negative instances of an event occurrence. This enabled to conclude that the use of the EM strategy diminishes the aforementioned reliance on the operator's confidence and expertise.

The results obtained also showed that the use of the EM strategy improved the ERS performance strongly. The BN-based classifiers outperformed their ANN-based counterpart.

4.3 Case Study #2

4.3.1 Section overview

In this section, after a general description of the case study, one analysis that aimed at evaluating the benefits of the EA optimisation strategy for selecting optimal ANN input structure and parameters sets is presented. However, it has to be noted that the ERS was used here for detecting the real-life events occurred in several DMAs within the YWS pilot area (see Section 4.2.2). Therefore, the analysis presented in this section also

served the purpose of further testing the capabilities of the ERS methodology and providing additional evidence of its good detection performance.

Details of how the above analysis was carried out and of the results obtained are reported in Section 4.3.3.

4.3.2 Case study #2 description

The case study examined and reported in this section made use of the historical data from the pressure and flow sensors deployed in five DMAs located in the YWS pilot area. These DMAs were selected because of their different characteristics (e.g., rural/industrial/urban) and varying sizes. The details of the DMAs that were selected are as outlined below.

The first DMA contains urban, light industrial, semi-rural and rural areas. It has a total of 2,415 domestic properties and 230 commercial users, of which 55 have a demand greater than 400 cubic metres per year. There is one major metered consumer within this DMA which has an average annual metered consumption of 14,530 cubic metres. The total mains length is 30 kilometres. The DMA is equipped with one inlet flow sensor and one pressure sensor at the critical point in the DMA.

The second DMA is predominantly urban and includes several hotels. It has a total of 3,493 domestic properties and 193 commercial users with an annual demand in excess of 400 cubic metres. There are three major metered consumers within this DMA with an average annual demand in excess of 10,000 cubic metres. The total mains length is 25 kilometres. The DMA is equipped with one inlet flow sensor and one pressure sensor at the critical point in the DMA.

The third DMA is predominantly urban. It has a total of 2,441 domestic properties and 231 commercial users, of which 75 have a demand greater than 400 cubic metres per year. The total mains length is 27 kilometres. There are two major metered consumers within this DMA with an average annual metered consumption of 13,072 cubic metres and 15,488 cubic metres, respectively. The DMA is equipped with one flow and pressure sensor at the DMA inlet.

The fourth DMA contains urban and light industrial areas. It has a total of 409 domestic properties and 11 commercial users with an annual demand in excess of 400 cubic metres. There are no major metered consumers within this DMA. The total mains length is 6.3 kilometres. The DMA is equipped with one inlet flow sensor and one pressure sensor at the critical point in the DMA.

The fifth DMA contains urban and light industrial areas. It has a total of 1,906 domestic properties and 141 commercial users, of which 30 have a demand greater than 400 cubic metres per year. There are no major metered consumers within this DMA. The total mains length is 16.5 kilometres. The DMA is equipped with one inlet flow sensor and a pressure sensor at the critical point in the DMA.

Similarly to the first case study described in this chapter, the historical data provided by the YWS referred to an eleven month period, between July 2009 and June 2010. The flow data were averaged values during the 15 minute sampling interval, whilst the pressure data were 15 minute instantaneous values. The WMS and CCs data (for the same time period) were also provided.

In order to achieve the main aim of the analysis performed in this case study, the novel methodology was (practically) implemented in the ERS as follows. The historical pressure and flow data were processed, in turn, with and without making use of the ANN parameters & input structure selection module (see Figure 2 and Section 3.4.3.6). When this ERS module was used, the parameters of the ES algorithm (see also Appendix C Section C.2) were selected as shown in Table 13. When this ERS module was not used, the number of hidden neurons, the coefficient of WDR, the number of training cycles and the ANN input structure (for each of the pressure/flow signals coming from the sensors deployed in the DMAs being studied) were chosen as described in Section 4.2.2. In both aforementioned cases: (1) the ERS user defined parameters not related to the ANN input structure and parameters were as those reported in Table 9, (2) the *alarm inactivity time* parameter was set as equal to 1 week, (3) the value of the detection threshold for raising the alarms was set as equal to 0.5, (4) the historical pressure and flow data were fed to the ERS in a simulated on-line fashion but without performing the periodic ERS reinitialisation, and (5) the *BIS parameters learning* subsystem was not used.

It has to be stressed that the ES algorithm parameters used for this case study (see Table 13) were identified after a sensitivity type analysis. The details of the results obtained are not shown here. However, it is worth reporting that the main finding from this analysis confirmed the observation from research on meta-EAs that most EAs are fairly insensitive to exact parameter settings (see – e.g., De Jong, 2007). Indeed, for a wide range of parameters, the ES algorithm allowed finding, within a reasonable time (i.e., max 15 minutes per DMA sensor on a standard dual core personal computer with 3.48 Gigabytes of Random Access Memory), ANN input structure and parameters sets that led to the development of ANN prediction models with very good forecasting performances (i.e., $NS_{index} > 0.95$ – see Section 3.4.3.6).

Table 13. Evolutionary Strategy algorithm parameters used for the case study.

Parameter	**Value**
Number of parents per generation – μ	15
Number of offspring per generation – λ	100
Number of fitness function evaluations – $N_{f.f.e.}$	515
Probability of a parameter being perturbed – $P_{mut.}$	0.6
Mutation Strength – σ	0.4
Selection operator	+

4.3.3 Evaluating the benefits of the Evolutionary Algorithm strategy

In the analysis presented here, the capabilities of the ERS methodology were tested and evaluated on real-life events with and without making use of the EA optimisation strategy. When the EA optimisation strategy was used, the ERS raised a total of 37 alarms. In the opposite case, it raised a total of 38 alarms.

Similarly to what done in Section 4.2.3, a careful visual inspection of the signal trends (supported by the WMS and CC data) was carried out here. This visual inspection enabled to define the set of real-life events against which to compare the alarms raised by the ERS in both relevant cases considered and, thus, to evaluate the ERS detection performance (i.e., success rate, and reliability). The defined set of real-life events

included: (i) 17 burst events (i.e., related to the MR records), (ii) 7 pressure/flow anomalies whose exact cause is uncertain but for which a record existed (i.e., clusters of CC records), (iii) 1 sensor failure event, and (iv) 5 other visible pressure/flow anomalies which were not accompanied by any CC or MR records (i.e., did not impact the customers).

In the case of ANN optimisation, the following results were obtained from the comparison between ERS alarms and set of real-life events. Of the 37 alarms that were raised:

- 21 alarms were correlated with the 17 recorded MRs (i.e., burst events). Note that because of the *alarm inactivity time* parameter used, multiple alarms were related to the same events. Furthermore, some event caused alarms in different DMAs;
- 7 alarms were correlated with the 7 clusters of recorded CCs (i.e., pressure/flow anomalies not followed by any MR record but for which a record existed);
- 3 alarms were correlated with the identified sensor failure event;
- 5 alarms were correlated with the 5 other visible pressure/flow anomalies that did not impact the customers;
- 1 alarm was a false alarm (i.e., it could not be correlated with any of the real-life events defined after visual inspection of the signal trends).

On the other hand, of the 38 alarms that were raised when the ANN optimisation was not carried out:

- 21 alarms were correlated with the 17 recorded MRs;
- 6 alarms were correlated with the clusters of recorded CCs;
- 3 alarms were correlated with the identified sensor failure event;
- 5 alarms were correlated with the 5 other visible pressure/flow anomalies that did not impact the customers;
- 3 alarms were false alarms.

The above results are summarised in Figure 33. In the light of these results, it is possible to state that the use of the EA optimisation strategy improved both the methodology's

reliability and effectiveness. Indeed, the number of false alarms was reduced three-fold (from 3 to 1), and one additional event that resulted in several CCs being received was detected.

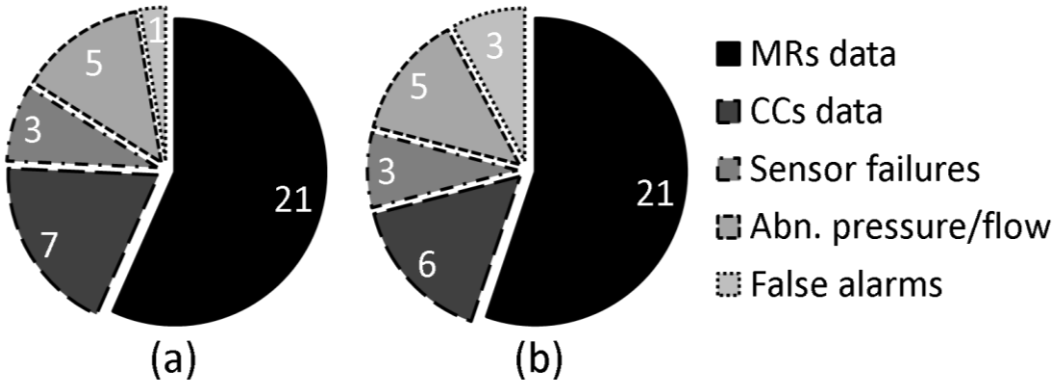


Figure 33. Number of alarms correlated with the real-life events defined by visual inspection of the data with (a), and without (b) using the Evolutionary Algorithm optimisation strategy.

The improved event detection reliability and efficiency obtained on this case study, however, are not the only benefits yielded by the use of the EA optimisation strategy. In actual fact, the methodology's detection speed was also enhanced. As an example, Figure 34 shows the results obtained when a large (i.e., about 30 litres per second) burst event occurred. The figure is divided into two parts. Figure 34a refers to the results obtained when the ANN optimisation was performed. Figure 34b refers to the results obtained when the ANN optimisation was not performed. In each part of the figure, the vertical red line indicates the time the detection alarm was raised. The vertical light blue bars indicate the DMA level event occurrence probability (if greater than 0.5) at every time step following the alarm. The vertical grey lines with small circles indicate the 6 recorded CCs. The vertical black line with a big circle indicates the recorded MR. Note that, as the MR did not have a time associated with it, it is plotted at the 00:00 on the 12th of January 2010. However, it is very likely that this MR was carried out some time after the burst.

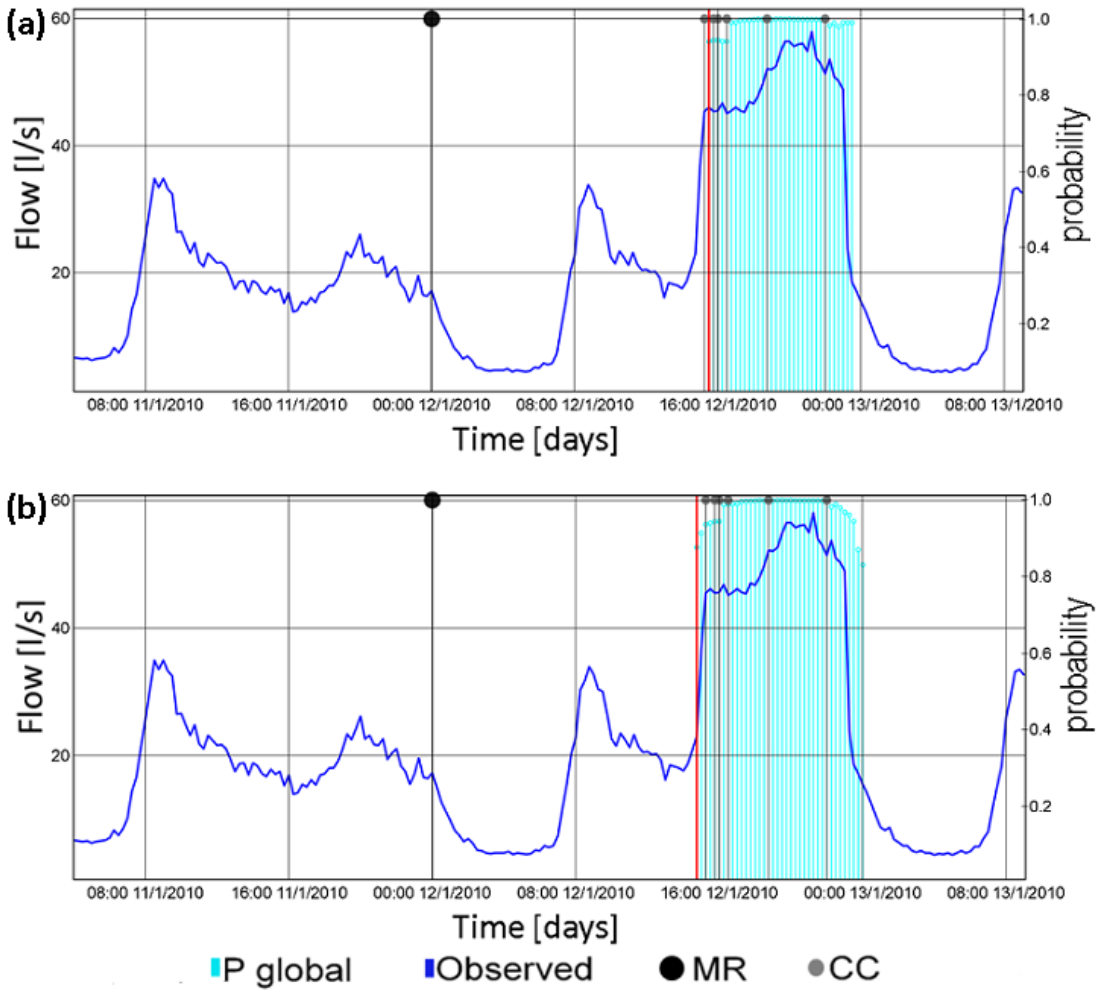


Figure 34. Detection of a large burst event with (a), and without (b) using the Evolutionary Algorithm optimisation strategy.

It can be observed from Figure 34 that when the ANN optimisation was used, the ERS raised an alarm 30 minutes before the first CC was received (i.e., at 15:30 on the 12th of January 2010). On the other hand, when it was not used, the ERS raised an alarm 45 minutes later (i.e., at 14:45 on the 12th of January 2010). That is to say, 15 minutes after the first CC. The fact that by making use of the EA optimisation strategy the ERS could have detected this burst event ahead of the customers is very important. This is because the early information may have enabled the water company to react more quickly and decrease the potential damages to the infrastructures and to third parties. Furthermore, it may have also helped the company to improve its customer service by: (i) reducing the time of the supply interruption, (ii) enabling to allocate more time for planning and

implementing mitigation measures (if the supply interruption could have not been avoided), and (iii) allowing proactive and/or more informed communications with the customers.

4.3.4 Section summary

The investigation of the benefits that the EA optimisation strategy for selecting optimal ANN input structure and parameters sets yield to the ERS methodology has been the focus of the analysis carried out in the case study presented in this section. This investigation has involved using the ERS for analysing the historical pressure and flow data from five YWS DMAs with different characteristics and varying sizes. The ERS analysed the historical data in a simulated on-line fashion with and without making use of the EA optimisation strategy. The main finding from this investigation is that by using the EA optimisation strategy the methodology's detection reliability and speed were enhanced.

4.4 Summary and Conclusions

In this chapter, the capabilities of the novel methodology that enables performing event detection in normally-sampled DMAs have been tested, evaluated and illustrated on real-life pipe bursts/other events and simulated pipe burst events. This has been done by using the ERS for processing the historical pressure and flow data from a number of UK DMAs. Furthermore, several investigations that evaluated the methodology's performances for different choices of parameters and data analysis methods have also been carried out.

After the introduction in Section 4.1, two case studies have been presented in Section 4.2 and Section 4.3, respectively. The results obtained from the analyses carried out in these case studies have shown the following: (i) the novel ERS methodology allowed the detection of both real-life pipe burst/other events and simulated pipe burst events in a timely and reliable manner, (ii) the novel ERS methodology showed the ability to identify precursor features of anomalies and to provide information that could be used for determining in a semi-automatic fashion the likely root cause of the events occurring, (iii) although the novel ERS methodology was found already fast and reliable, the use of the EM strategy for calibrating the parameters in the CPTs of the DMA level BIS further improved the methodology's performance. It also enabled

reducing the reliance on the operator's confidence and expertise when selecting the most suitable detection threshold value to be used for raising the alarms, (iv) BN-based classifiers were found to perform better than their ANN-based counterpart, and (v) the use of the EA optimisation strategy for selecting optimal parameters and input structure for each ANN prediction model further enhanced the methodology's event detection reliability and speed.

All of the above demonstrated that the ERS presented in Section 3.4 can be effectively used for detecting pipe bursts and other events in a reliable and timely manner. The term 'reliable' refers here to the fact that the obtained false alarm rates in all aforementioned tests are very low, typically below 10% - i.e., lower than the corresponding values reported in the literature, e.g., 15-18% (Mounce et al., 2010a; Mounce and Boxall, 2010) and 22% (Mounce et al., 2011). Similarly, the term 'timely' refers to the fact that most events are detected and alarms generated in 15-30 minutes after the event which is much faster than the values reported in the literature - e.g., 12 hours (Mounce et al., 2010a; Mounce and Boxall, 2010) and 4 hours (Mounce et al., 2011). However, it has to be stressed that a direct comparison of the above results obtained is not possible as the tests carried out in the aforementioned studies involved different DMAs and events. Furthermore, it should be noted that tests on live data/events (e.g., Mounce et al., 2010a; Mounce and Boxall, 2010) are more challenging than tests performed by using historical data/events. Reliable and timely event detections are important as they enable the water companies to gain confidence in the raised alarms and, in turn, react quickly to minimise the negative impacts of the burst/other events leading to an improvement of the water companies' operational efficiency and customer service.

In the following chapter the capabilities of the methodology that enables performing event detection and location in over-sampled DMAs are tested, evaluated and demonstrated on a case study involving the use of pressure measurements recorded during a series of EEs.

CHAPTER 5 CASE STUDY FOR EVENT DETECTION & LOCATION IN OVER-SAMPLED DISTRICT METERED AREAS

5.1 Introduction

This chapter reports the results of a number of sensitivity type data analyses carried out on a single case study. The main aim of these analyses was to test, evaluate and illustrate the capabilities of the methodology that enables performing event detection and location in over-sampled DMAs – i.e., as implemented in the customised and further developed ERS described in Section 3.5. Similarly to the two case studies presented in Chapter 4, the historical data used in this case study were collected as part of the NEPTUNE project. They consisted of pressure measurements recorded during a series of EEs carried out in one of the YWS pilot area's DMAs (see Section 4.2.2).

The chapter is organised as follows. After this introduction, the aims and the rationale behind the analyses performed and presented here are given in Section 5.2. Then, the case study is introduced in Section 5.3. In this section, a description of the DMA being studied, the EEs carried out, and the available data is given. Once this is done, Section 5.4 presents the details concerning the practical implementation of the methodology in the specific case study being analysed. Section 5.5 reports the results of the data analyses that focussed on testing and evaluating the methodology's event detection capabilities. Section 5.6 reports the results of the data analyses that focussed on testing and evaluating the methodology's event location capabilities. Finally, a summary of the chapter and the main conclusions are given in Section 5.7.

5.2 Case Study #3 Aims

As mentioned above, the analyses carried out in this chapter mainly aimed at testing, evaluating and illustrating the capabilities of the methodology that enables performing event detection and location in over-sampled DMAs. However, they also aimed at achieving the following goals:

1. Evaluating the performance of the ANN prediction models for different *LagSize* values;
2. Evaluating the benefits of the wavelet-based de-noising procedure;
3. Evaluating the methodology's event detection capabilities for different measurement sampling rates;
4. Evaluating the methodology's event detection capabilities for different 'comparative/Control Rule set' combinations;
5. Evaluating the benefits of the 'cumulative learning' procedure;
6. Evaluating the performances of different geostatistical techniques for the approximate event location within a DMA;
7. Evaluating the usefulness of the pressure measurements for enabling the timely and reliable near real-time detection and location of events in WDSs.

Note that the first, fourth, fifth, and sixth goals stated above can be seen as targeted to the evaluation of the methodology's event detection and location capabilities when different parameters and/or data analysis methods are used. However, the other goals also provide an insight on issues which have implications that go beyond the developed methodology itself. These issues are highlighted in the discussion that follows below. The reason for this is that these issues are deemed to be important for the effective management of incidents in WDSs and it is envisaged that they will become more and more compelling in the near future.

As stated in Section 3.5.1, the wider availability of more accurate and cheaper pressure/flow sensors and the consequent envisaged deployment (in the near future) of a larger number of these devices in a single DMA has motivated the customisation and further development of the ERS. Bearing this in mind, it is expected that the deployment of an increased number of pressure sensors would be the water companies' preferred choice. The rationale behind this is that these sensors can be installed at significantly lower costs than flow sensors. Furthermore, their calibration and maintenance requirements are also far less onerous.

Using pressure measurements in a data analysis methodology that aims at performing event detection and location, however, is more challenging than using flow measurements. This is mainly because pressure signals from a DMA sensor typically do not exhibit the same degree of stationarity over extended periods of time as flow signals. Indeed, the less a DMA signal exhibit stationarity over extended periods of time the more difficult it is to model and forecast its NOP and hence identify when an event occurs. Taking this into consideration, it is worth mentioning that the weak stationarity nature of the pressure signals is partially due to the fact that, unlike flow measurements that are sensitive to 'downstream' changes, pressure measurements are sensitive to changes in head-loss along prescribed 'up-stream' flow routes. This fact makes: (i) the pressure sensors more sensitive to local changes (e.g., opening of cross connections to adjacent DMAs), and (ii) the sensors' response to a particular event more dependent on location (e.g., a signal from a pressure sensor at the inlet of a DMA which is fed by a service reservoir is unlikely to display a significant drop in reaction to a burst within the DMA). Because of all this, pressure is considered a less reliable parameter than flow (Mounce et al., 2010c). As a result, at the present, pressure sensors are almost exclusively used for assessing whether the pressure measured at the critical point in a DMA has fallen below some pre-defined threshold level or not. Given the trends in availability of pressure measurements from an increased number of sensors deployed in a DMA, however, this is no longer acceptable. Supported by the development of suitable data analysis methodologies, it is therefore important to evaluate if the wider availability of on-line pressure data can play an important role in the context of near real-time event detection and location.

A second issue arising from the use of pressure sensors is the measurement noise. Indeed, noise in the measurements from a DMA sensor has a negative impact on the performances of an event detection and location data analysis methodology. In the UK water industry, flow measurements are routinely averaged over a fifteen minute interval. This is mainly due to the way flow measurements are obtained. The output of most flow sensors is, in fact, an electronic pulse per unit of flow. The number of these pulses is counted over and then divided by the measurement sampling interval (i.e., fifteen minutes). As a result, a 'smoothed' flow time series, which is immediately suitable for use in an event detection and location data analysis methodology, is obtained. On the

other hand, irrespective of the basis of operation, pressure measurements tend to be instantaneous values. This can generate data analysis problems as an instantaneous reading could be at a "level" not representative of the general trend. As mentioned in Section 2.4.7.3, Mounce et al. (2012) recently investigated the benefit of averaging the pressure measurements from a WDS and showed that this is a useful strategy not only for event detection but also for other WDSs management purposes (e.g., HM calibration, provision of better regulatory compliance, etc.). Although the averaging strategy is simple and achievable by a simple low cost firmware upgrade without any additional overheads on communications (Mounce et al., 2012), it may lose potentially useful information in the data. Therefore, exploring alternative and more efficient de-noising strategies may prove fundamental for improving the "usability" of the pressure measurements.

The last issue highlighted here concerns the sampling frequency of both pressure and flow measurements. As mentioned above, in the UK a fifteen minute sampling interval is commonly used in the water companies' SCADA systems. The main reason for this practice is that most pressure and flow changes in a WDS are gradual. Hence, this particular sampling rate is seen to provide a reasonable trade-off between amount of data to be handled/analysed and definition of the daily pressure/flow dynamics within a WDS. Also, since the battery life of a sensor strongly decreases as the sampling rate increases, frequent recharge is required when the data are more frequently sampled. This creates an additional burden from an operational and managerial point of view. However, pressure/flow measurements recorded at sub-fifteen minute intervals have the potential to provide a more detailed picture of the dynamics of a WDS. Furthermore, since the current instrumentation's fast technical improvements (e.g., logging memory, battery life, communication infrastructures) may render collection and on-line communication of sub-fifteen minute data much more feasible than has previously been the case, the availability of such data may also have the potential of significantly shortening the time needed to detect an incident occurring.

5.3 Case Study #3 Description

5.3.1 Study area

The DMA being studied is predominantly rural. It has a total of 897 domestic properties and 28 commercial users with an annual demand in excess of 400 cubic metres. There are no major metered consumers (i.e., with an annual demand in excess of 5% of the DMA flow or greater than 10,000 cubic metres per year). This DMA has a total of 17.8 kilometres of pipes with diameters from 50 to 300 millimetres. Its configuration is a combination of loops and branches. It has one import and two export points. The elevation in the DMA ranges from 43 to 70 metres (above ordinance datum).

Figure 35 shows a satellite view of the study area with overlaid DMA boundaries and pipe network. This DMA is deemed representative of a typical UK DMA with wide range of conditions.

Figure 35. Case study #3 District Metered Area.

5.3.2 Engineered Events

The EEs considered here were carried out on the 7th, 8th, and 12th of August 2008. Several pipe bursts were simulated at different locations within the DMA by opening fire hydrants to create additional network flows to waste (to street gulley or to land to soak away without causing flooding or disruption). Specifically, on the 7th of August 2008, 5 hydrants in different locations were opened, at different times, for approximately 1 hour each. On the 8th of August 2008, 2 hydrants at different locations were opened, such that each ran individually for approximately an hour and both ran simultaneously for approximately an hour. This was then repeated using 2 different hydrants. Finally, on the 12th of August 2008, 1 hydrant was opened in six stages starting with a very small flow and progressing to a large flow, thereby simulating a growing burst. Table 14 reports the time schedule of the EEs carried out. Figure 36 shows a comparison between the total inflow into the DMA being studied for the day immediately before the EEs execution (i.e., the 6th of August 2008), and the DMA total inflow for the days during which the EEs were carried out (i.e., the 7th, 8th, and 12th of August 2008).

Table 14. Engineered Events time schedule.

7th August	Time	Flow [l/s]
Hydrant 1	08:25	0.00
	08:26	5.00
	08:34	6.67
	09:29	0.00
Hydrant 2	09:33	0.00
	09:34	5.17
	09:35	5.75
	09:36	6.25
	10:24	18.83
	10:37	0.00
Hydrant 3	10:56	0.00
	10:57	5.00
	10:58	6.67
	11:56	18.33
	11:58	0.00
Hydrant 4	12:06	0.00
	12:07	6.17
	12:08	7.33
	13:05	13.00
	13:09	0.00
Hydrant 5	13:17	0.00
	13:18	5.00
	13:19	7.00
	13:30	7.50
	13:32	5.00
	14:18	5.00
	14:20	0.00

8th August	Time	Flow [l/s]
Hydrant 5	08:41	0 to 5
	08:42	2.00
Hydrant 4	09:44	0.00
	09:45	2.00
Hydrant 5	10:44	2.00
	10:44	0.00
Hydrant 4	11:45	2.00
	11:46	0.00
Hydrant 1	11:53	0.00
	11:54	2.00
Hydrant 2	12:54	0.00
	12:55	2.00
Hydrant 1	13:52	2.00
	13:53	0.00
Hydrant 2	14:53	2.00
	14:54	0.00

12th August	Time	Flow [l/s]
Hydrant 5	11:57	0.20
	13:06	0.42
	14:02	0.78
	15:01	1.00
	16:04	1.23
	17:10	5.00
	18:12	-
	18:14	0.00

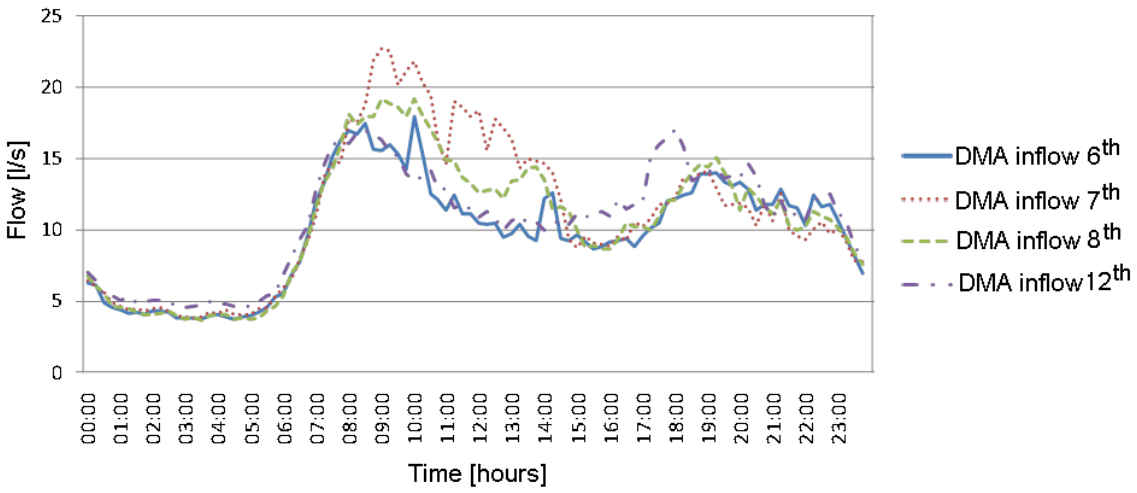


Figure 36. Total inflow into the District Metered Area being studied for the 6th, 7th, 8th, and 12th of August 2008.

5.3.3 Available data

The DMA being studied had existing flow and pressure sensors installed at its inlet and export points. These sensors recorded data at fifteen minute intervals. Instrumentation within the DMA consisted of a pressure sensor recording data at fifteen minute intervals. For the purpose of the EEs detailed in Section 5.3.2, however, the DMA was temporarily equipped with a larger number of pressure sensors. Specifically, 13 extra sensors were deployed. These sensors recorded data at 5 second intervals during the 6th, 7th, 8th, and 12th of August 2008.

For the purposes of the data analyses presented in this chapter, only the 5 second pressure data from the 13 extra sensors were used. Allowing for this, for each sensor, 6 additional pressure time series were created by re-sampling the recorded data at 10 seconds, 1 minute, 3 minutes, 5 minutes, 10 minutes, and 15 minutes. This way, the availability of pressure time series recorded at decreasing sampling rates was simulated.

5.4 Customised and Further Developed Event Recognition System Modifications

5.4.1 Overview

In the case study presented here, several changes to the way the methodology that enables performing event detection and location in over-sampled DMAs normally works (see Section 3.5) had to be implemented. These changes were necessary for the following two main reasons: (1) circumventing the lack of historical data from the 13 pressure sensors considered (as they were *temporarily* installed for the purpose of carrying out the EEs only), and (ii) enabling the selection of different parameters/data analysis methods and thus achieving the goals listed in Section 5.2. The changes implemented here resulted in a customised and further developed ERS as the one shown in Figure 37.

By comparing Figure 37 with Figure 17 it is immediately evident that, in the resulting customised and further developed ERS, the *BBA* and *TBA* subsystems are missing. The reason for this is that the data analyses in these subsystems have a requirement for several weeks of historical data. Thus, given the aforementioned limited historical data availability, they could not be performed. Taking this into account, it is important to stress that the omission of these subsystems was easily achieved due to the fact that the ERS is fully modular. Additionally, note that it did not affect the validity of the results obtained. Indeed, the use of the historical data from sensors *permanently* installed in the DMA could have only further improved the methodology's performance.

Although not as immediately evident as the omission of the *BBA* and *TBA* subsystems, by comparing Figure 37 with Figure 17, it is also possible to note that other modules of the customised and further developed ERS were affected by the changes implemented here. These are: (i) the data pre-processing module, (ii) the data de-noising module, (iii) the ANN parameters & input structure selection module, (iv) the evidence generation module, and (v) the geostatistical analysis module. Details of the changes implemented in each of these modules together with other details relevant to the practical implementation of the customised and further developed ERS methodology in the case study presented here are given in the following sections.

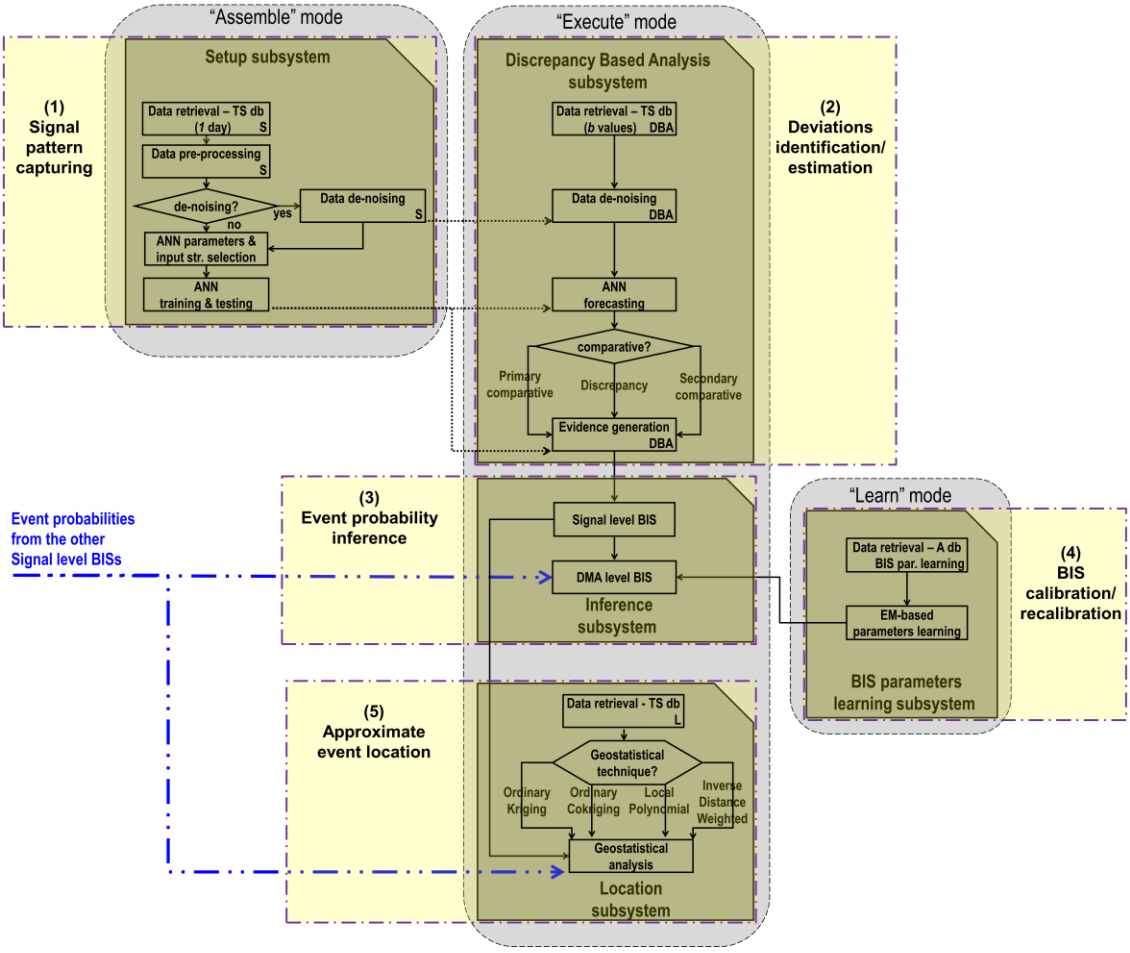


Figure 37. Practical implementation of the event detection and location methodology in the customised and further developed Event Recognition System.

5.4.2 Data pre-processing module

For each signal being analysed, the goal of this module is to assemble a 'NOP data set' (see Section 3.4.3.3). In this case study, since only one day of data during which the pipe bursts were not simulated was available (i.e., the 6th of August 2008), this goal was achieved according to the following procedure. The raw pressure data recorded during the 6th of August 2008 were first checked for erroneous time stamps/missing values, repaired accordingly, and then taken as representative of the signal daily variations assuming that no burst occurred in the DMA.

5.4.3 Data de-noising module

For each signal being analysed, the goal of this module is to remove noise from the 'NOP data set' (see Section 3.4.3.5). In this case study, in order to allow evaluating the benefits of the wavelet-based de-noising procedure, the customised and further developed ERS was given the capability of skipping this data analysis step (see Figure 37).

5.4.4 Artificial Neural Network parameters & input structure selection module

For each signal being analysed, the goal of this module is to select the optimal ANN input structure and parameters set to be used for building the relevant ANN prediction model (see Section 3.4.3.6). In this case study, due to the desire to evaluate the performance of the ANN prediction models for different values of the *LagSize* parameter and because of the aforementioned limited historical data availability, the use of the EA optimisation strategy was deemed inappropriate. Consequently, here, the ANN parameters and input structures were chosen as follows. The customised and further developed ERS was given the capability of selecting the value of the *LagSize* as equal to 4, 6, 8, 10, or 12. The Neuroshell2 manual (1996) rule of thumb was used for selecting the number of hidden neurons. The coefficient of WDR was set as equal to 0.1. The number of training cycles was set as equal to 250. Note that, for all the pressure signals being analysed, these ANN parameters and input structures combinations were found (after a series of preliminary tests – not shown here) to ensure the resulting ANN prediction models were able to closely approximate the training sets whilst allowing good generalisation performances.

5.4.5 Artificial Neural Network training & testing module

For each signal being analysed, the goal of this module is to train and test an ANN model for the one step ahead prediction of pressure values assuming that no burst is occurring in the DMA being studied (see Section 3.4.4.4). In this case study, the first 80% (i.e., *Train%*) of data in the 'de-noised NOP data set' or in the 'NOP data set' (depending on whether or not the wavelet-based de-nosing procedure was carried out) was used as the training set. The remaining 20% (i.e., *Test%*) was used as the test set.

5.4.6 Evidence generation module

For each signal being analysed and at each time step, the goals of this module are: (a) to reliably identify the event-induced deviations from the signal NOP, and (b) to generate a piece of evidence that reflects the event-induced deviation presence (or absence) and its severity (if present) (see Section 3.4.4.5).

Generally speaking, the above goals can be achieved according to the following procedure. Firstly, a comparative variable (or simply comparative) has to be selected and used to represent the comparison, at each time step, between ANN predicted and actually observed signal values. Secondly, Control Rules (see Appendix D Section D.3) have to be applied to the comparatives in order to reliably identify the event-induced deviations from the signal NOP. Finally, the number of Control Rule violations has to be determined and classified. In view of this and due to the desire to evaluate the methodology's event detection capabilities for different 'comparative/Control Rule set' combinations, in this case study, the aforementioned procedure was implemented as outlined below.

Firstly, the customised and further developed ERS was given the capability of selecting and using one of the following comparatives (see Figure 37):

- Primary comparative = observed (t)/predicted (t);
- Secondary comparative = [observed (t)/predicted (t)] – [observed (t-Δt)/predicted (t-Δt)];
- Discrepancy = observed (t) – predicted (t).

Control Rules were then applied to the selected comparative. Note that a single Control Rule tests whether a comparative falls outside a confidence region that is plus and minus a user defined number of standard deviation limits from the mean or:

$$\mu_{comp.,t} - N_l\sigma_{comp.,t} \leq x_{comp.,t} \leq \mu_{comp.,t} + N_u\sigma_{comp.,t} \quad (5.1)$$

where $x_{comp.,t}$ is the comparative's value at time t, $\mu_{comp.,t}$ is the mean of the comparative's values, $\sigma_{comp.,t}$ is the standard deviation of the comparative's values, and N_l and N_u are user defined values denoting the acceptable lower and upper control limits. Here, the mean and standard deviation of the selected comparative's values were

computed based on the data resulting from the ANN prediction models' training. The values of N_l and N_u were chosen to be the same and as equal to 1, 2, 3 and 4.

As shown in Figure 38, the above choice of N_l and N_u values resulted in a situation that can be represented graphically by means of a Control Chart. It is possible to observe from this figure that five zones (i.e., A, B, C, D and E) can be identified. In this framework, at any time step, the value of the selected comparative had to fall in one of the zones. This allowed establishing a direct correspondence between the zones and the discrete input of the Signal level BIS, whereby the evidence relating to an event occurrence is classified according to: (A) high, (B) moderate, (C) low, (D) very low, or (E) none.

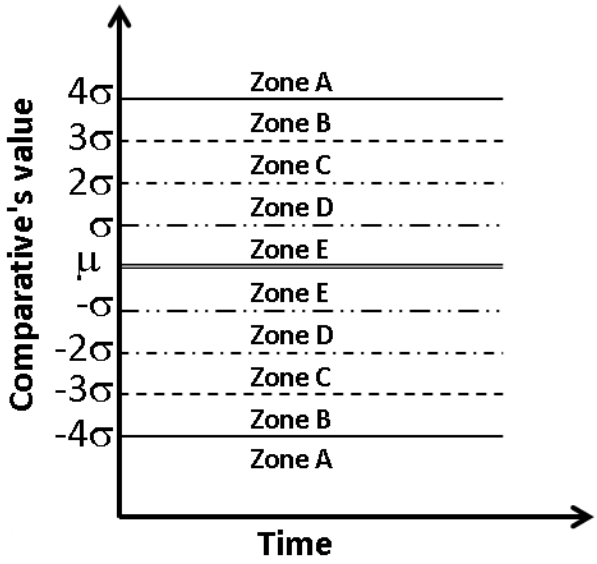


Figure 38. Typical Control Chart divided into zones for different control criteria

Bearing in mind the above, a set of Control Rules that, at every time step, simply checks which zone the comparative fell in, was applied to the primary and secondary comparatives. On the other hand, the modified subset of the Western Electric Control Rules described in Section 3.4.4.5 was applied to the discrepancy. The rationale for this is that such 'comparative/Control Rule set' combinations allowed evaluating the benefit of analysing a comparative over multiple time steps. In fact, it is possible to note that

only a single time step (i.e., *t*) is considered if the primary comparative is used. Two consecutive time steps (i.e., *t* and *t*-Δ*t*) are considered if the secondary comparative is used. Up to eight consecutive time steps (i.e., *t*, *t*-Δ*t*,..., *t*-7Δ*t*) are considered if the discrepancy is used.

5.4.7 Signal level Bayesian Inference System module

For each signal being analysed, the goal of this module is to output the probability that an event has occurred in the DMA being studied (see Section 3.4.7.2). In this case study, since the evidence resulted from the *DBA* subsystem only, the structure of the BN used for achieving the aforementioned goal had to be modified. The modified BN structure used here is shown in Figure 39.

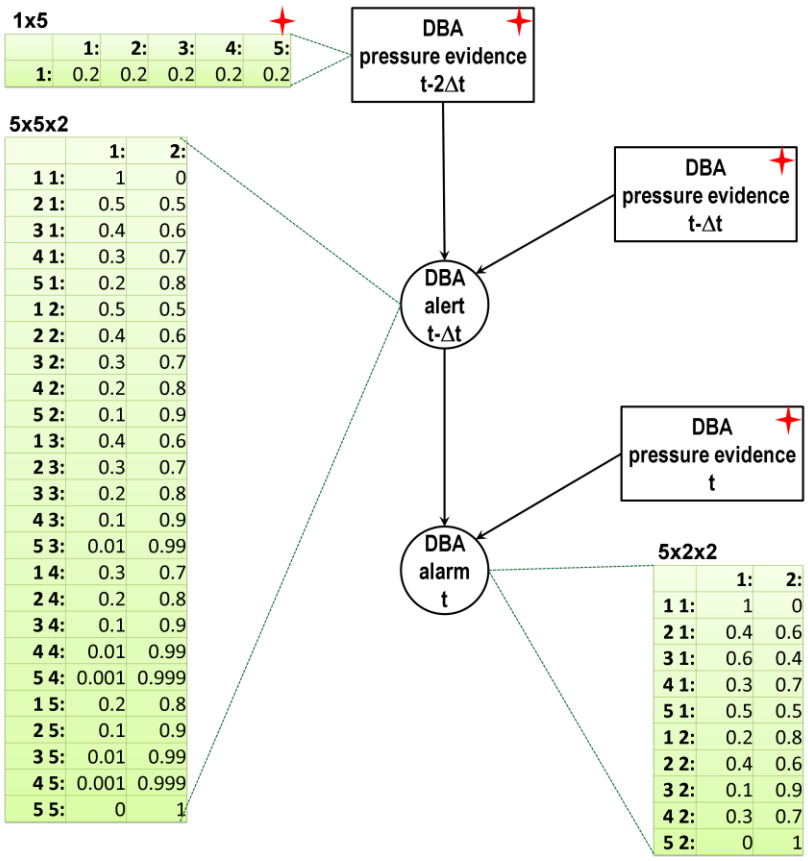


Figure 39. Modified structure of the Signal level Bayesian Inference System.

Figure 39 shows that each input node has five classes (see also Section 5.4.6), whilst the alert and alarm nodes have two classes (i.e., alarm on/alarm off). The figure also shows the values of the parameters used in the CPTs associated with each node and that three time steps were considered (i.e., t, t-Δt and t-$2\Delta t$) to allow updating the event occurrence evidence probability over consecutive time steps.

5.4.8 District Metered Area level Bayesian Inference System module

The goal of this module is to raise the detection alarms if and when necessary. In this case study, this was achieved by: (i) calibrating the parameters of the multivariate Gaussian mixtures-based DMA level BIS described in Section 3.5.3 using the EM strategy (see also Section 3.4.8 and Appendix E Section E.3.2) and information (i.e., start time and duration) about the first two EEs carried out on the 7th of August 2008, (ii) using a detection threshold (see Section 3.4.7.3) set as equal to 0.5, and (iii) using an *alarm inactivity time* (see Section 3.4.7.3) set as equal to the relevant measurement sampling rate.

5.4.9 Geostatistical analysis module

The goal of this module is to determine the approximate location of the pipe burst events within the DMA (see Section 3.5.4). In this case study, in order to evaluate the performances of different geostatistical techniques, the customised and further developed ERS was given the capability of selecting and using the OC, the OK, the LP, or the IDW technique (see Figure 37 and Appendix F).

5.5 Testing of Event Detection Capabilities

5.5.1 Section overview

In this section four analyses that focussed on testing and evaluating the event detection capabilities of the customised and further developed ERS methodology are presented.

The first is a sensitivity type analysis that mainly aimed at evaluating the benefits of the wavelet-based de-noising procedure. However, the effects on the ANN models performance of the different *LagSize* values as the measurement sampling rate increases were also investigated.

The second analysis is a sensitivity type analysis that aimed at evaluating the methodology's event detection capabilities when pressure measurements recorded at sub-fifteen minute sampling rates are used and the potential benefit that measurements at such sampling rates could yield.

The third analysis is also a sensitivity type analysis. It focussed on the role played by the use of Control Rules for achieving more reliable and efficient event detection. In particular, this analysis aimed at evaluating the benefits of analysing a comparative over multiple time steps.

The fourth analysis aimed at evaluating the benefits of integrating the EM strategy into a 'cumulative learning' procedure (i.e., dynamic re-calibration of the multivariate Gaussian mixtures-based DMA level BIS parameters).

Details of how these analyses were carried out and of the results obtained are reported in the following sections.

5.5.2 Evaluating the benefits of the wavelet-based de-noising procedure and investigating the LagSize selection issue

In the analysis presented here, the customised and further developed ERS shown in Figure 37 was applied to the case study with and without making use of the data de-noising module (see Figure 37). When the data de-noising module was not used (i.e., the 'NOP data sets' was used directly to train and test the ANN prediction models), the 13 pressure signals available, the 7 different sampling rates considered, and the 5 different *LagSize* values selectable by the customised and further developed ERS led to a total number of trained and tested ANN prediction models as equal to 455. Similarly, 455 ANN prediction models were trained and tested when the data de-noising module was used (i.e., the 'de-noised NOP data set' was used to train and test the ANN prediction models).

Figure 40 shows the average Nash-Sutcliffe indices (over 13 analysed signals) for the training and test data sets, for the two principal cases studied (with/without noise) and all for 5 different *LagSize* values and 7 different sampling rates. This figure emphasises the importance of the wavelet-based de-noising procedure for the effective modelling/forecasting of pressure signals. Indeed, as it can be seen from this figure, the

performances of the ANN prediction models were poor when the 'NOP data sets' (i.e., original signals) were used directly. However, they were greatly improved when the wavelet-based de-noising was performed. It can be concluded that, because of the non-stationary nature of the pressure signals, even a powerful modelling tool such as an ANN will not be able to accurately capture and predict their variations if some smart 'smoothing' technique (e.g., de-noising) is not applied to the data.

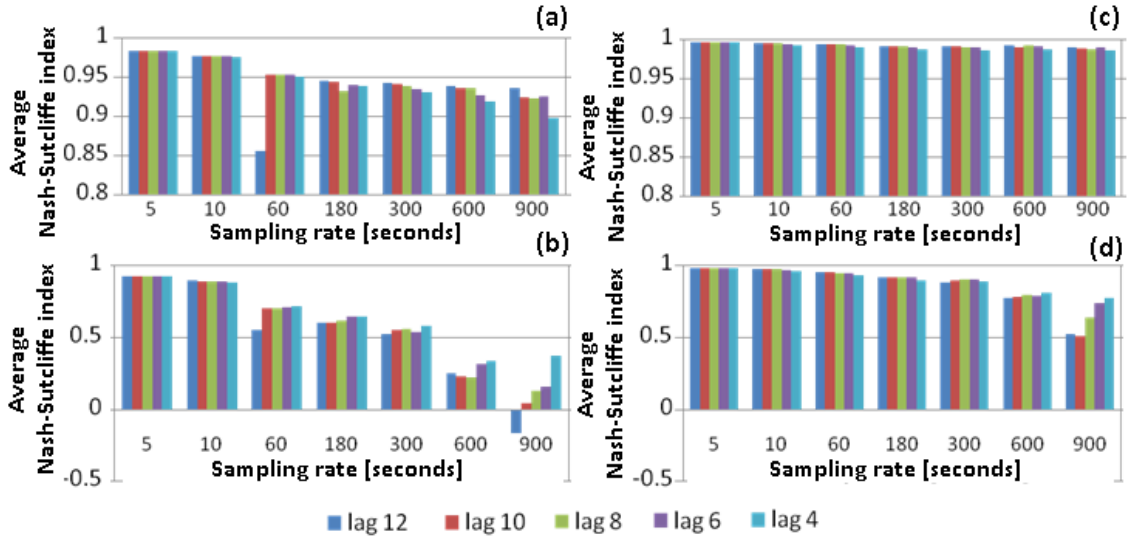


Figure 40. Artificial Neural Networks performances. For the training (a) and test (b) sets using the original signals, and for the training (c) and test (d) sets using the de-noised signals.

Secondly, it can be observed from Figure 40 that even though the performances of the ANN prediction models decreased with the decreasing sampling frequency (an expected consequence of the decreasing size of the training and test sets - only 96 data points in 24 hours for the 15 minute sampling interval), choosing a small *LagSize* value (e.g., 4) improved the accuracy of the predictions when large sampling intervals were considered (e.g., 10 and 15 minute intervals). At the same time, the choice of a small *LagSize* value did not negatively affect the performances when small sampling intervals were considered. This finding is important for near real-time event detection because choosing a small *LagSize* value has the advantage of including only a small number of

abnormal values as ANN input which, in turn, increases the reliability of the ANN predictions. However, it is to be stressed that this finding resulted from training the ANN prediction models based on a single day of data. Therefore, as stated in Section 3.4.3.6, it does not necessarily apply to the more general case whereby several days of data are considered (including both weekdays and weekends) or when modelling/forecasting more complex signals (e.g., pressure measured downstream a pressure modulating valve).

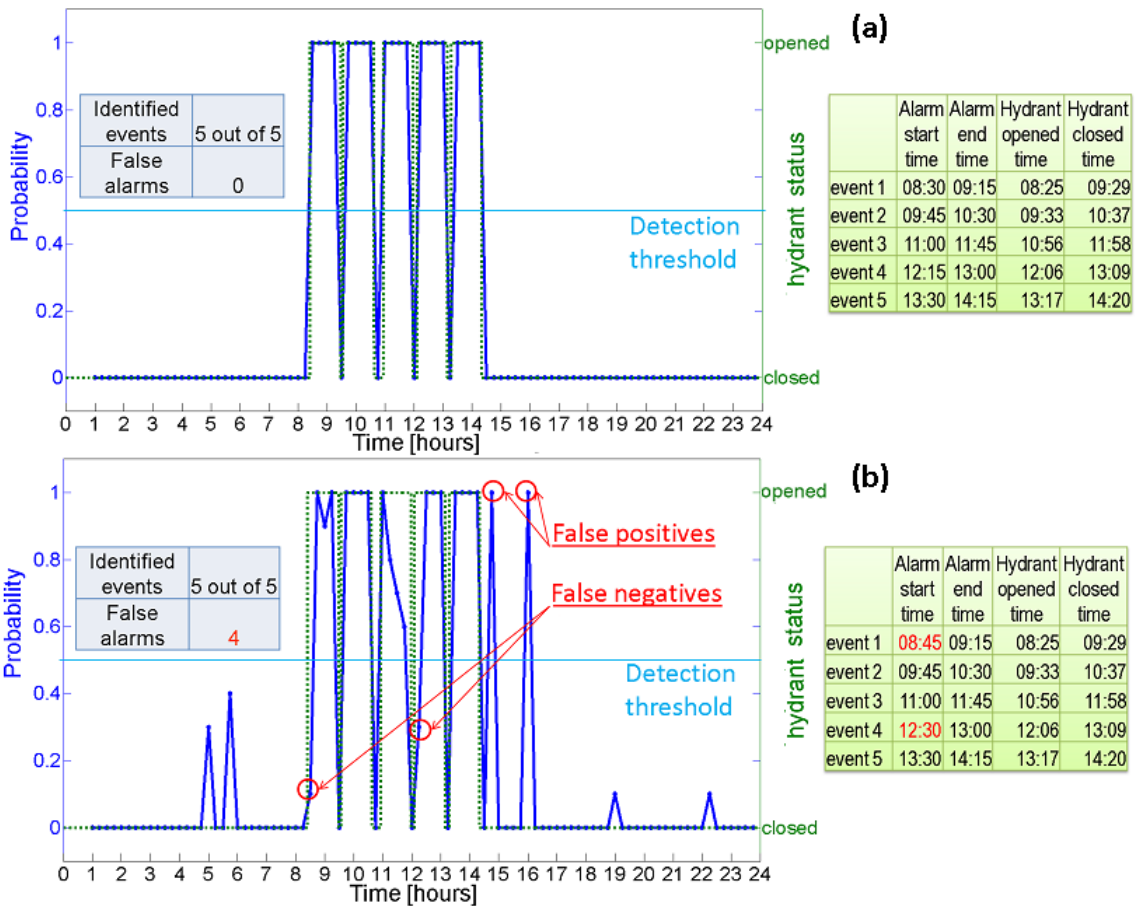


Figure 41. Detection results obtained when the customised and further developed Event Recognition System was used for processing the pressure time series recorded on the 7th of August 2008 and re-sampled at 15 minutes with (a), and without (b) the wavelet-based de-noising.

Figure 41 shows the detection results obtained when the customised and further developed ERS was used for processing the pressure data collected on the 7th of August 2008 (i.e., first day of EEs). In particular, it shows the results obtained when the pressure time series re-sampled at 15 minutes and a *LagSize* value as equal to 4 were used. Note that the results obtained by using the pressure time series re-sampled at 15 minutes are shown here because this represents the worst case scenario (i.e., very limited information about the simulated burst events embedded in the re-sampled pressure time series and longest detection time). The *LagSize* value as equal to 4 was chosen for the reasons mentioned above. Note further that a *LagSize* value as equal to 4 was also used for the other data analysis presented in this chapter. Therefore, hereafter and where not mentioned otherwise, the above value can be assumed. Figure 41 is divided into two parts. Figure 41a refers to the case where the data de-noising module was used. Figure 41b refers to the case where the data de-noising module was not used. Each part of the figure has two vertical axes - the left one referring to the DMA level event occurrence probability (i.e., solid blue line) and the right one referring to the status (opened/closed) of the five hydrants (i.e., dotted green line). Note the actual detection times and the corresponding hydrant opening and closing times provided in the small table on the right and the number of detected events and false alarms in the small table on the left. False alarms are divided into two categories: (i) false positives (i.e., the customised and further developed ERS raised an alarm when the hydrants were closed) and (ii) false negatives (i.e., the customised and further developed ERS did not raise an alarm when a hydrant was opened). As it can be observed from Figure 41, when the customised and further developed ERS was used for processing the pressure signals without performing de-noising (Figure 41b), both its reliability (i.e., two false positives) and efficiency (i.e., the two false negatives led to detecting events 1 and 4 with a 15 minute delay) were worse than in the case where de-noising was performed (Figure 41a), resulting in all five simulated burst events detected at best possible time (bearing in mind the measurement sampling rate used).

5.5.3 Evaluating the benefits of sub-fifteen minute pressure measurements

In the analysis presented here, the detection results obtained when the customised and further developed ERS shown in Figure 37 was used for processing the pressure time

series recorded on the 7th of August 2008 and re-sampled, in turn, at 5 seconds, 10 seconds, 30 seconds, 1 minute, 3 minutes, 5 minutes, and 15 minutes were analysed.

Table 15. Alarm start and end times for different measurement sampling rates.

Sampling rate [seconds]	Alarm start/end time	event 1	event 2	event 3	event 4	event 5
5	**Alarm start time**	08:26:55	09:35:20	10:58:45	12:08:10	13:18:25
	Alarm end time	09:31:30	10:40:10	11:59:45	13:10:40	14:21:50
10	**Alarm start time**	08:27:00	09:35:20	10:58:40	12:08:00	13:18:30
	Alarm end time	09:31:30	10:40:00	11:59:40	13:10:40	14:21:50
30	**Alarm start time**	08:27:00	09:35:30	10:59:00	12:08:30	13:18:30
	Alarm end time	09:31:30	10:40:00	11:59:30	13:10:30	14:21:30
60	**Alarm start time**	08:27:00	09:36:00	10:59:00	12:09:00	13:19:00
	Alarm end time	09:31:00	10:40:00	11:59:00	13:10:00	14:21:00
180	**Alarm start time**	08:27:00	09:36:00	11:00:00	12:09:00	13:21:00
	Alarm end time	09:30:00	10:39:00	11:57:00	13:09:00	14:21:00
300	**Alarm start time**	08:30:00	09:40:00	11:00:00	12:10:00	13:20:00
	Alarm end time	09:25:00	10:40:00	11:55:00	13:10:00	14:20:00
900	**Alarm start time**	08:30:00	09:45:00	11:00:00	12:15:00	13:30:00
	Alarm end time	09:15:00	10:30:00	11:45:00	13:00:00	14:15:00

Table 16. Number of identified events and false alarms for different measurement sampling rates.

Sampling rate [seconds]	Identified events	Number of false alarms	False alarm rate [%]
5	5 out of 5	258	1.5
10	5 out of 5	127	1.5
30	5 out of 5	43	1.5
60	5 out of 5	21	1.5
180	5 out of 5	10	2.1
300	5 out of 5	4	1.4
900	5 out of 5	0	0.0

Table 15 reports the detection times (i.e., alarm start and end times) for the 5 burst events simulated on the 7th of August 2008 and for the 7 measurement sampling rates considered. In this table, the values of the alarm start and end times in normal text font indicate a mismatch between customised and further developed ERS detection times and actual hydrant opening and closing times. The underlined values indicate that the simulated burst events were detected at the best possible time (bearing in mind the relevant measurement sampling rate used). Table 16 summarises the number of identified events and false alarms raised. Figure 42, as an example, shows the detection results obtained when the 5 second pressure time series were used.

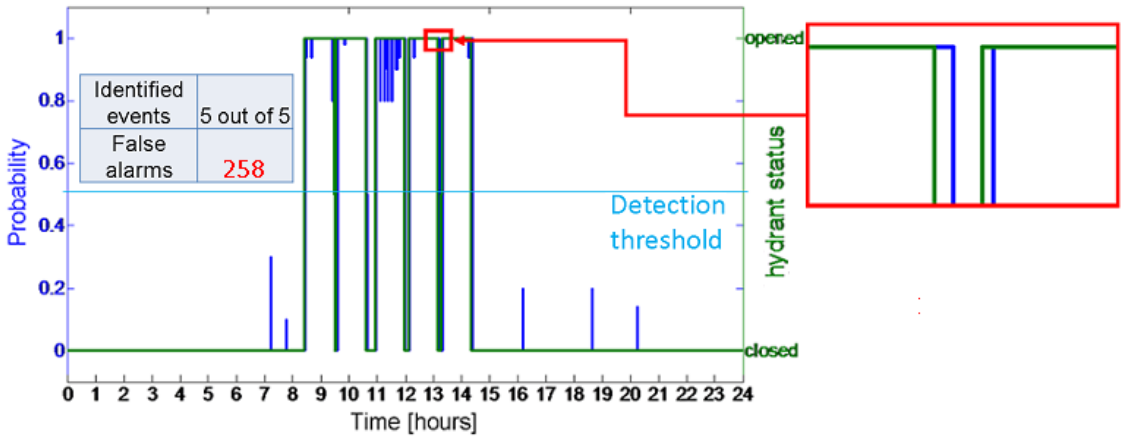


Figure 42. Detection results obtained when the customised and further developed Event Recognition System was used for processing the 5 second pressure time series recorded on the 7th of August 2008.

By analysing the results in Table 15 and Table 16 and by observing Figure 42 it is possible to state that, although all the simulated burst events were always detected, the 'absolute' number of false alarms drastically increased when sub-fifteen minute sampling rates were used. However, it is also important to stress that the false alarms raised by the customised and further developed ERS were, in every case, no more than a few minutes away from an hydrant manoeuvre (i.e., they could be logically related to the simulated burst events). This circumstance seems to suggest that the resulting large

number of false alarms may have been due to the fact that the logged hydrant opening and closing times did not exactly match the signals' time stamps, or to the fact that the pressure sensors did not react instantaneously to the manoeuvres on the hydrants. To support this hypothesis, note that the "zoom in" window shown in Figure 42 provides evidence that, when compared to the logged hydrant opening and closing times, the customised and further developed ERS alarms were slightly shifted in time (i.e., delayed).

With regard to the high 'absolute' numbers of false alarms raised when sub-fifteen minute sampling rates were used, the following has to be also noted. In the analysis carried out here, because of the way the false alarms were declared, a small mismatch between customised and further developed ERS detection times and actual hydrant opening and closing times (e.g., 1 minute) resulted in a large number of false negatives at the beginning of a simulated pipe burst event (e.g., 12 false negatives – for a 5 second sampling rate) and a corresponding large number of false positives at the end of the simulated pipe burst event.

In the light of the above observations and bearing in mind the false alarm percentages shown in Table 16, it can be concluded that the customised and further developed ERS showed good event detection performances for all the measurement sampling rates considered.

Figure 43 shows the average (over the 5 simulated pipe burst events) delay in detection for each of the 7 measurement sampling rate considered. This figure shows that the detection times were significantly reduced by using pressure measurements recorded at sub-fifteen minute sampling rates. However, it also suggests that no appreciable further improvements were achieved by using pressure measurements sampled more frequently than 1 minute. Therefore, it can be concluded that while increasing the measurement sampling rates to 5, 3 or 1 minute may allow shortening the detection times considerably, the use of measurements recorded at even higher sampling rates may not warrant the increased operational and managerial burden (e.g., frequent sensors' battery replacements and data management overheads).

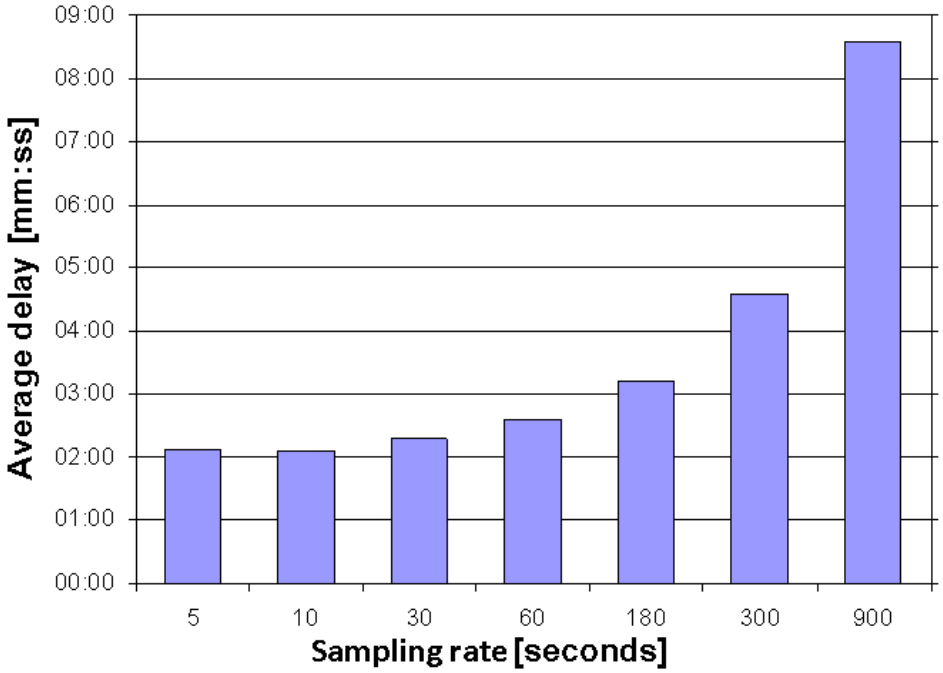
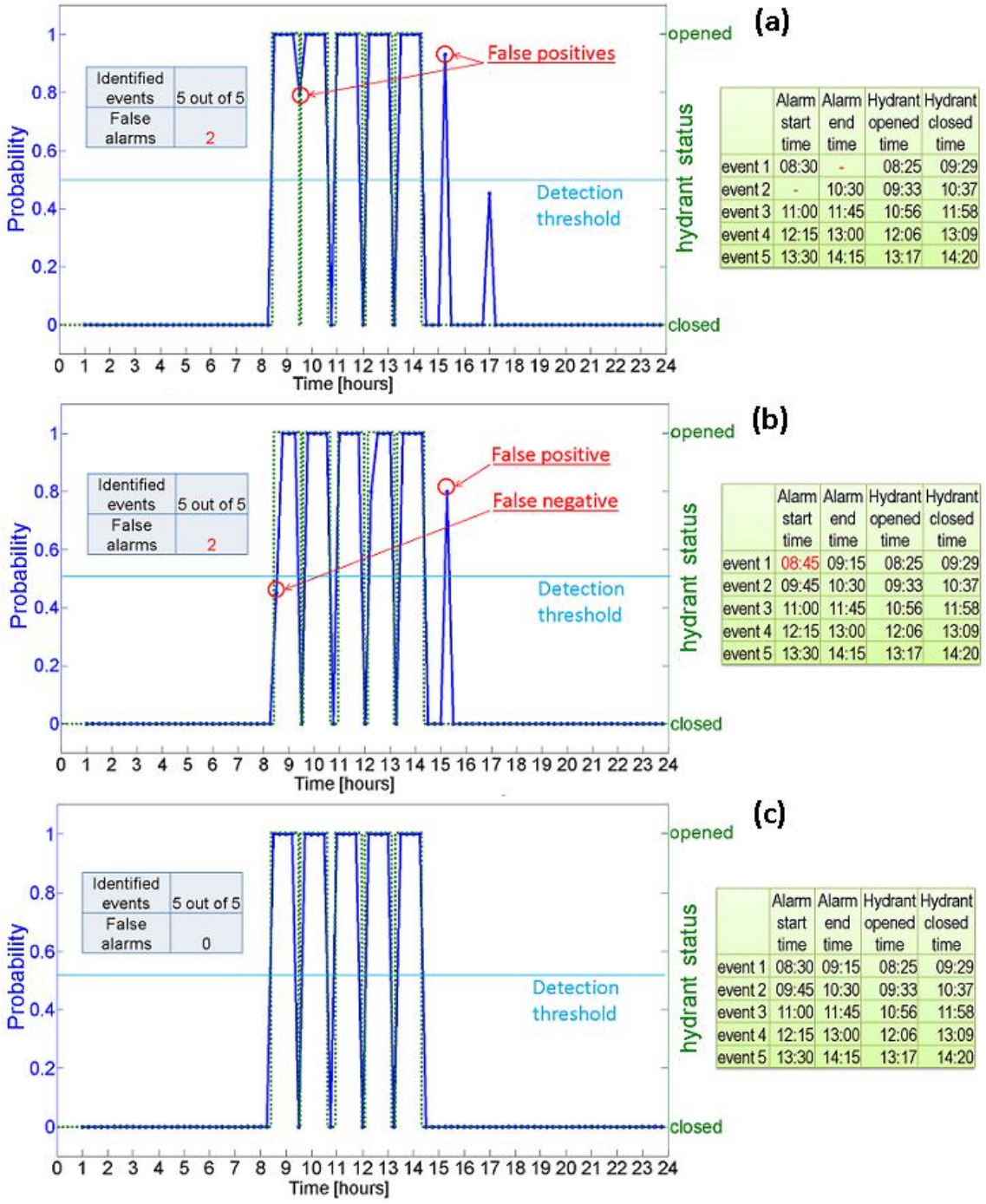


Figure 43. Average delay in detection for different measurement sampling rates.

5.5.4 Evaluating the benefits of different comparatives and anomaly detection rules

In the analysis presented here, the customised and further developed ERS shown in Figure 37 was applied to the case study using, in turn, the three 'comparative/Control Rule set' combinations described in Section 5.4.6.

Similarly to Figure 41, Figure 44 shows the detection results obtained when the customised and further developed ERS was used for processing the pressure time series recorded on the 7th of August 2008 and re-sampled at 15 minutes. In particular, this figure reports the results of applying the customised and further developed ERS to the case study by using, in turn, the Control Rule set associated with: (i) the primary comparative, (ii) the secondary comparative, and (iii) the discrepancy. According to the analyses performed, the figure is divided into three parts (i.e., a, b and c). Again, the left and right vertical axes refer to the DMA level event occurrence probability and to the status of the five hydrants, respectively. The actual detection times and the corresponding hydrant opening and closing times are in the table on the right. The number of detected events and false alarms are in the table on the left.



	Alarm start time	Alarm end time	Hydrant opened time	Hydrant closed time
event 1	08:30	-	08:25	09:29
event 2	-	10:30	09:33	10:37
event 3	11:00	11:45	10:56	11:58
event 4	12:15	13:00	12:06	13:09
event 5	13:30	14:15	13:17	14:20

	Alarm start time	Alarm end time	Hydrant opened time	Hydrant closed time
event 1	08:45	09:15	08:25	09:29
event 2	09:45	10:30	09:33	10:37
event 3	11:00	11:45	10:56	11:58
event 4	12:15	13:00	12:06	13:09
event 5	13:30	14:15	13:17	14:20

	Alarm start time	Alarm end time	Hydrant opened time	Hydrant closed time
event 1	08:30	09:15	08:25	09:29
event 2	09:45	10:30	09:33	10:37
event 3	11:00	11:45	10:56	11:58
event 4	12:15	13:00	12:06	13:09
event 5	13:30	14:15	13:17	14:20

Figure 44. Detection results obtained when the customised and further developed Event Recognition System was used for processing the pressure time series recorded on the 7th of August 2008 and re-sampled at 15 minutes using the Control Rule set relative to: the primary comparative (a), the secondary comparative (b), and the discrepancy (c).

As it can be observed from Figure 44, when the customised and further developed ERS was used for processing the pressure measurements by means of the Control Rule set relative to the primary comparative (Figure 44a), two false alarms (i.e., false positives) were generated and because the first of these false alarms was raised at 9:30 a.m., event 1 and event 2 could not be distinguished. When the customised and further developed ERS was used for processing the pressure measurements by means of the Control Rule set relative to the secondary comparative (Figure 44b), one false positive was generated at 15:15 p.m. and additionally, because of a false negative at 8:30 a.m., event 1 was detected with a 20 minute delay from the actual hydrant opening time. On the other hand, when the customised and further developed ERS was used for processing the pressure measurements by means of the Control Rule set relative to the discrepancy (Figure 44c), both its reliability and its efficiency were improved. Indeed, as already noted in section 5.5.2, all the simulated pipe burst events were detected at the best possible times and no false alarms were raised.

Based on the above results, it can be concluded that the analysis of the discrepancy comparative over multiple time steps enabled more reliable detection of the simulated pipe burst events.

5.5.5 Evaluating the benefits of the 'cumulative learning' procedure

In the analysis presented here, the customised and further developed ERS shown in Figure 37 was applied to the case study using the 'cumulative learning' procedure described in Section 3.4.8.3.

Here, information about the hydrant opening and closing times of the five pipe burst events simulated on the 7th of August 2008 was used to calibrate the parameters of the multivariate Gaussian mixtures-based DMA level BIS. The customised and further developed ERS with calibrated DMA level BIS parameters was then used for processing the pressure time series recorded during the 8th of August 2008 (i.e., EEs second day). Progressing in a like fashion, information about the hydrant opening and closing times for the EEs carried out on the 8th of August 2008 was used together with ('cumulative learning') the previously used information about the EEs carried out on the 7th of August 2008 to recalibrate the DMA level BIS parameters. Finally, the customised and further developed ERS with the DMA level BIS parameters recalibrated by using the

'cumulative learning' procedure was used for processing the pressure time series recorded during the 12th of August 2008.

The above was done in order to simulate the real-life on-line scenario whereby the DMA level BIS parameters are (re)calibrated as new information about past events occurred in the DMA being studied becomes available. The simulated scenario is as follows: (1) on the 7th of August 2008, alarms were raised by the customised and further developed ERS following the occurrence of a number of pipe burst events (e.g., as shown in previous sections), (2) later that day, an operator confirmed (i.e., tagged as genuine) these alarms and stored information about the start and end times of the events detected in the Alarms database, (3) upon completion of this alarms review process, the stored information about the events occurred on the 7th of August 2008 was used to calibrate the DMA level BIS parameters, (4) once this was done, the customised and further developed ERS with the DMA level BIS parameters calibrated as in point 3 was used for processing the data coming on-line, (5) on the 8th of August 2008, alarms were raised following the occurrence of new pipe burst events, (6) later that day, the operator performed a new alarms review, (7) upon completion, the stored information about the events occurred on the 8th of August 2008 together with information about the events occurred on the 7th of August 2008 was used to recalibrate the DMA level BIS parameters, (8) the customised and further developed ERS with the DMA level BIS parameters recalibrated as in point 7 was then used for processing the data coming on-line, and (9) on the 12th of August 2008, alarms were raised following the occurrence of a new pipe burst event.

Figure 45 shows the detection results obtained when the customised and further developed ERS, with the DMA level BIS parameters calibrated as described above, was used for processing the pressure time series recorded on the 8th of August 2008 and re-sampled at 3 minutes. In this figure, the vertical axis on the left refers to the DMA level event occurrence probability (i.e., solid blue line). The vertical axis on the right refers to the status of the four hydrants opened and closed according to the schedule reported in Table 14 (i.e. dashed and dotted coloured lines). The actual detection times and the corresponding hydrant opening and closing times are shown in the table on the right. The number of detected events and false alarms are shown in the table on the left. It can

be observed from this figure that reliable (i.e., only two false alarms) and quick (i.e., within 4 minutes) event detections were obtained.

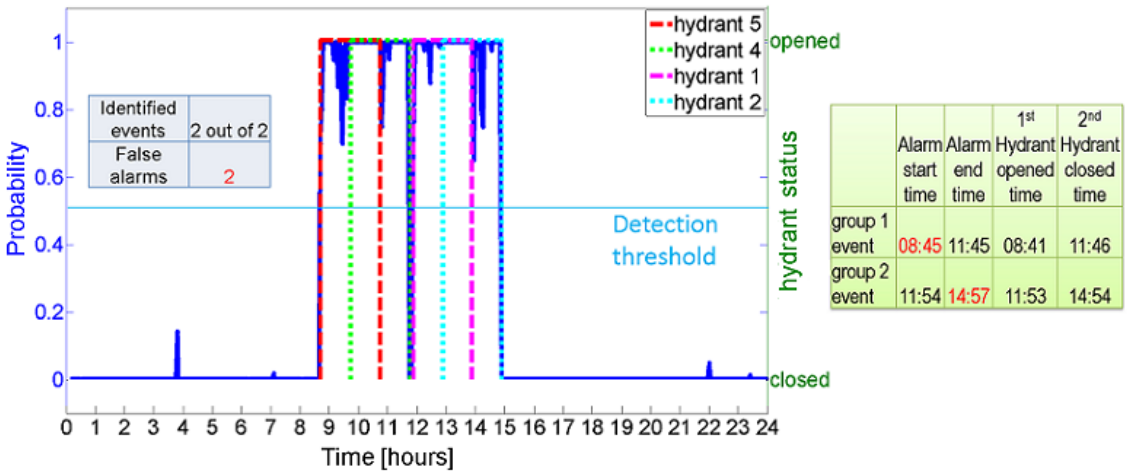


Figure 45. Detection results obtained when the customised and further developed Event Recognition System, with the District Metered Area level Bayesian Inference System parameters calibrated using information about all the events simulated on the 7th of August 2008, was used for processing the pressure time series recorded on the 8th of August 2008 and re-sampled at 3 minutes.

Despite the good detection results reported above, it has to be stressed that they were obtained when processing the pressure time series re-sampled at 3 minutes. When the pressure time series re-sampled at 15 minutes were processed, on the other hand, the customised and further developed ERS was not able to separate the two events. This is because of the very short time interval between the closing time of the second hydrant in the first group and the opening time of the first hydrant in the second group (i.e., 6 minutes only). In the light of these findings, a further benefit yielded by the use of measurements sampled at a sub-fifteen minute intervals is highlighted (see also Section 5.5.3). This is the ability of discriminating between events very near in time to each other and occurring in the same DMA. However, it has to be noted here that such a circumstance is unlikely in a real-life scenario.

The setup of the EEs carried out on the 8th of August 2008 also highlighted the inability of the customised and further developed ERS to detect in isolation events occurring "simultaneously" (meaning that overlap in time) at different locations within the DMA (i.e., the two hydrants that ran simultaneously in each group). However, it is possible to observe from Figure 45 how the variations in the DMA level event occurrence probability can provide useful information to further diagnose the events occurring and overcome this limitation. In fact, when the hydrants were simultaneously opened the DMA level event occurrence probabilities were larger than the case where each hydrant ran independently.

Figure 46 shows the detection results obtained when the customised and further developed ERS was used for processing the pressure time series recorded during the 12th of August 2008 and re-sampled at 15 minutes for the following two cases: (i) DMA level BIS parameters calibrated using information about the events simulated on the 7th of August 2008 (i.e., no 'cumulative learning'), and (ii) DMA level BIS parameters recalibrated using information about the events simulated on the 7th and on the 8th of August 2008 (i.e., 'cumulative learning'). In this figure, the vertical axis on the left refers to the DMA level event occurrence probability (i.e., solid blue line). The vertical axis on the right refers to the status (opened/closed) of the hydrant that was opened in six stages according to the schedule reported in Table 14 (i.e., solid grey line). The vertical dotted coloured lines indicate the timing of the five flow increments. The actual detection times and the corresponding hydrant opening and closing times are shown in the table on the right. The number of detected events and false alarms are shown in the table on the left.

As it can be observed from Figure 46, when the 'cumulative learning' procedure was not used (Figure 46b), both efficiency (the simulated pipe burst event was detected with a 1 hour and 15 minute delay) and reliability (i.e., the simulated pipe burst event was detected after the first flow increment only - the minimum detectable burst rate was 0.42 litres per second) of the customised and further developed ERS were worse than in the case where the 'cumulative learning' procedure was used (Figure 46a). Indeed, when the 'cumulative learning' procedure was used, the simulated burst was detected with a 15 minute delay from the best possible detection time (bearing in mind the sampling interval used) and at its minimum flow rate (i.e., 0.2 litres per second).

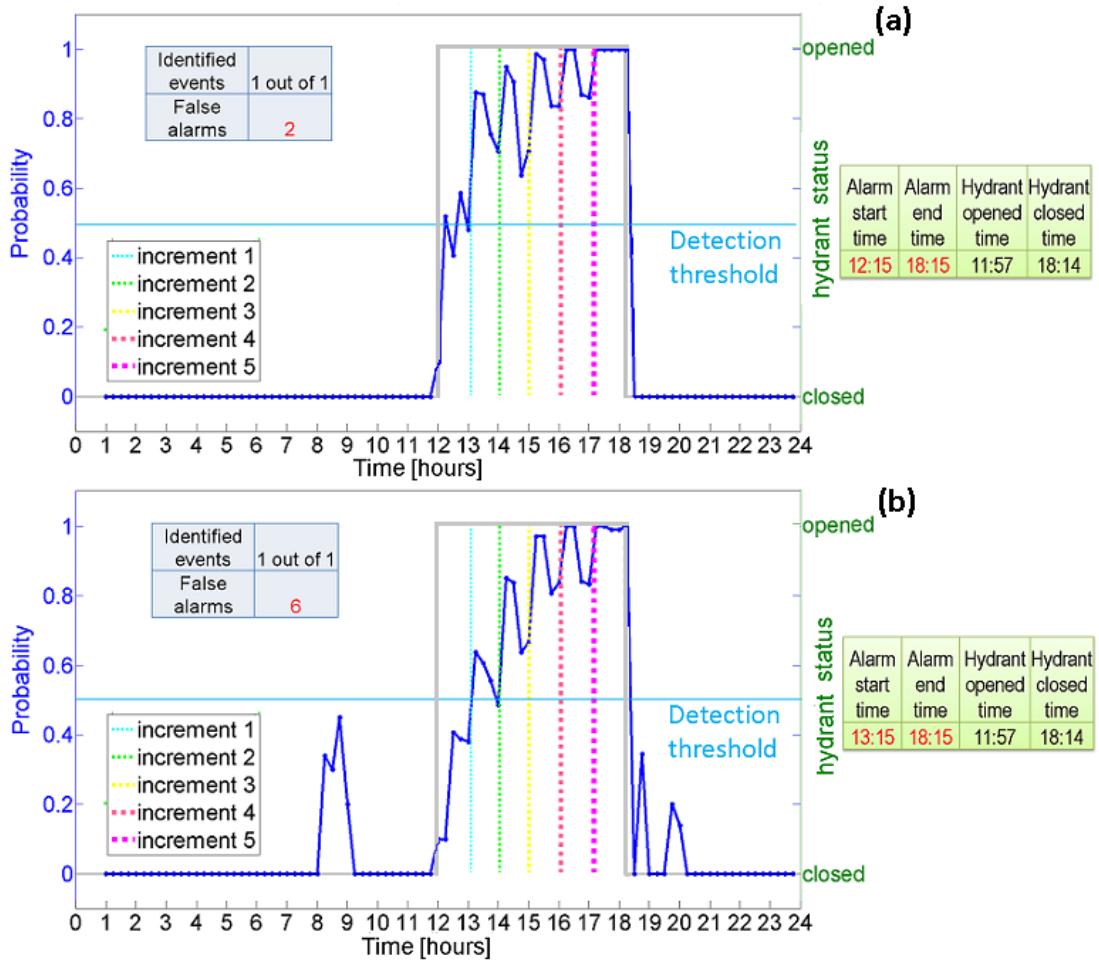


Figure 46. Detection results obtained when the customised and further developed Event Recognition System was used for processing the pressure time series recorded during the 12th of August 2008 and re-sampled at 15 minutes with (a), and without (b) using the 'cumulative learning' procedure.

Based on the above, it can be concluded that, by using the 'cumulative learning' procedure, the customised and further developed ERS was able to learn recognising the features of simulated pipe burst events of different type and size (i.e., the size of the pipe burst simulated on the 8th of August 2008 was significantly smaller than the size of the pipe bursts simulated on the 7th of August 2008). This led to improved generalisation capabilities which, in turn, allowed the pipe burst with a very small flow rate simulated on the 12th of August 2008 to be detected in a good time. In view of this,

it is possible to infer that the methodology's event detection performances could be continuously improved if the 'cumulative learning' procedure described here is used in an 'on-line' fashion for (re)calibrating the parameters of the DMA level BIS.

Finally, note that in order to detect the pipe burst simulated on the 12th of August 2008 at its minimum flow rate, when the DMA level BIS parameters were calibrated using only information about the events simulated on the 7th of August 2008, the value of the user defined detection threshold could have been set as equal to 0.4. This, however, would have had the unwanted effect of raising a false positive alarm at 8:45 a.m. (see Figure 46).

5.5.6 Section summary

The methodology's event detection capabilities for different parameters/data analysis methods have been tested and evaluated in this section. Issues such as the use of measurements sampled at sub-fifteen minute intervals and the value of the on-line pressure data for near real-time event detection have also been investigated. The main findings from the analyses performed in this section are briefly summarised below.

The careful selection of the number of input data for the ANN prediction models and the use of a wavelet-based procedure for de-noising the pressure signals coming from the sensors deployed in the DMA being studied significantly improved the detection reliability and efficiency of the methodology.

The methodology showed reliable results and faster detection times when pressure measurements recorded at sub-fifteen minute sampling rates were used. However, no further appreciable improvements to the detection speed were achieved by using pressure measurements sampled more frequently than 1 minute. It was also noted that the use of measurements sampled at a sub-fifteen minute intervals enabled the customised and further developed ERS discriminating between events very near in time and occurring in the same DMA.

The analysis of the discrepancy comparative by using a set of Control Rules applied to consecutive time steps improved the detection reliability and efficiency of the methodology.

The methodology showed to be already fast and reliable when the parameters of the DMA level BIS were calibrated by using information about a small number of simulated pipe burst events. However, further improvements to its accuracy, reliability and efficiency were achieved by using information about a larger number of simulated pipe burst events. In the light of this finding, it was inferred that continuous detection performance improvements could be achieved if the past events information is used in a 'cumulative learning' procedure for the 'on-line' (re)calibration of the DMA level BIS parameters.

By using pressure measurements only, the developed methodology reliably and rapidly detected different types of simulated pipe burst events.

5.6 Testing of Event Location Capabilities

5.6.1 Section overview

In this section two data analysis that focussed on testing and evaluating the capabilities of the customised and further developed ERS methodology for determining the approximate location of an event within an over-sampled DMA are presented.

The first data analysis aimed at testing and evaluating the methodology's event location capabilities when the OC technique was used for developing a model that predicts the probability value of a burst event associated with each DMA pipe.

The second analysis aimed at comparing the methodology's event location capabilities when models developed by using different geostatistical techniques were employed.

Details of how these analyses were carried out and of the results obtained are reported in the following sections.

5.6.2 Evaluating the methodology's event location capabilities when using Ordinary Cokriging

In the analysis presented here, the customised and further developed ERS shown in Figure 37 was applied to the case study using the OC technique (see Appendix F Section F.5) according to the procedure described in Section 3.5.4.

Figure 47 to Figure 51 show the location results obtained when the first, second, third, fourth and fifth pipe burst events simulated on the 7th of August 2008 were considered, respectively. Note that here the cumulative procedure described in Section 3.5.4 was used. In this regard, note that the results shown in these figures refer to the third time step after the relevant simulated pipe burst event was detected (e.g., 9:15 a.m. for the first EE, 10:30 a.m. for the second EE, etc.). However, similar results were obtained for the time step each of the events was detected and for the two time steps immediately after that (not shown here).

Figure 47 to Figure 51 are all divided into two parts. The left part (i.e., burst event probability map) shows the burst event probability values for each DMA pipe. The right part (i.e., OC standard deviation map) shows the corresponding OC standard deviations (see also Appendix F Section F.5). In both parts of each figure, the locations of the deployed pressure sensors are indicated by using green squares and the real location of the hydrant opened in order to simulate the relevant pipe burst event is indicated by a red star symbol. In the left part of each figure, the DMA pipes are grouped by using the Jenks Natural Breaks Classification method (see Section 3.5.4) and coloured according to their burst event probability value. In the right part of each figure, pipes are grouped by using the Jenks Natural Breaks Classification method and coloured according to their OC standard deviation value.

Bearing in mind that the higher the value of the burst event probability for a pipe the more likely that pipe is the ‘failed’ pipe, Figure 47 to Figure 51 show that, for each of the five simulated pipe burst events, the customised and further developed ERS successfully identified the group of pipes in the proximity of the ‘failed’ pipe. Furthermore, bearing in mind that the lower the value of the OC standard deviation value for a pipe the higher the confidence that the predicted burst event probability value for that pipe is close to its ‘real’ value (i.e., good estimate), it is possible to observe how the predicted burst event probability values are more accurate the closer they are to the locations of the deployed sensors.

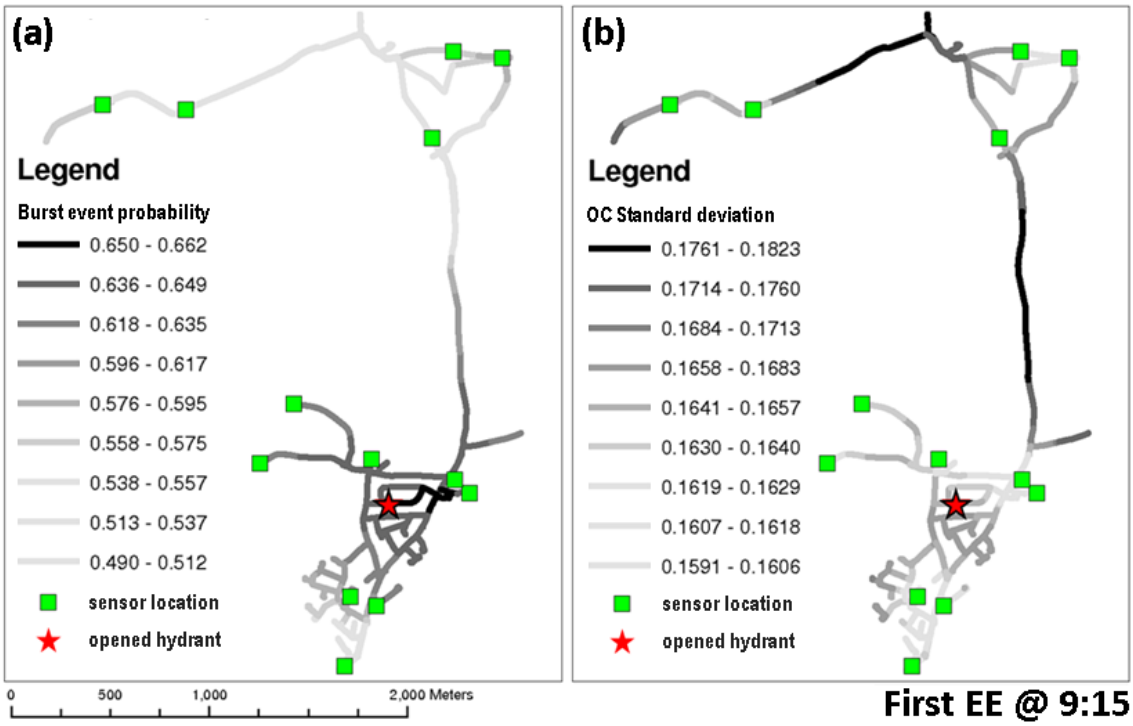


Figure 47. Approximate location results for the first Engineered Event. Burst event probability map (a), and Ordinary Cokriging standard deviation map (b).

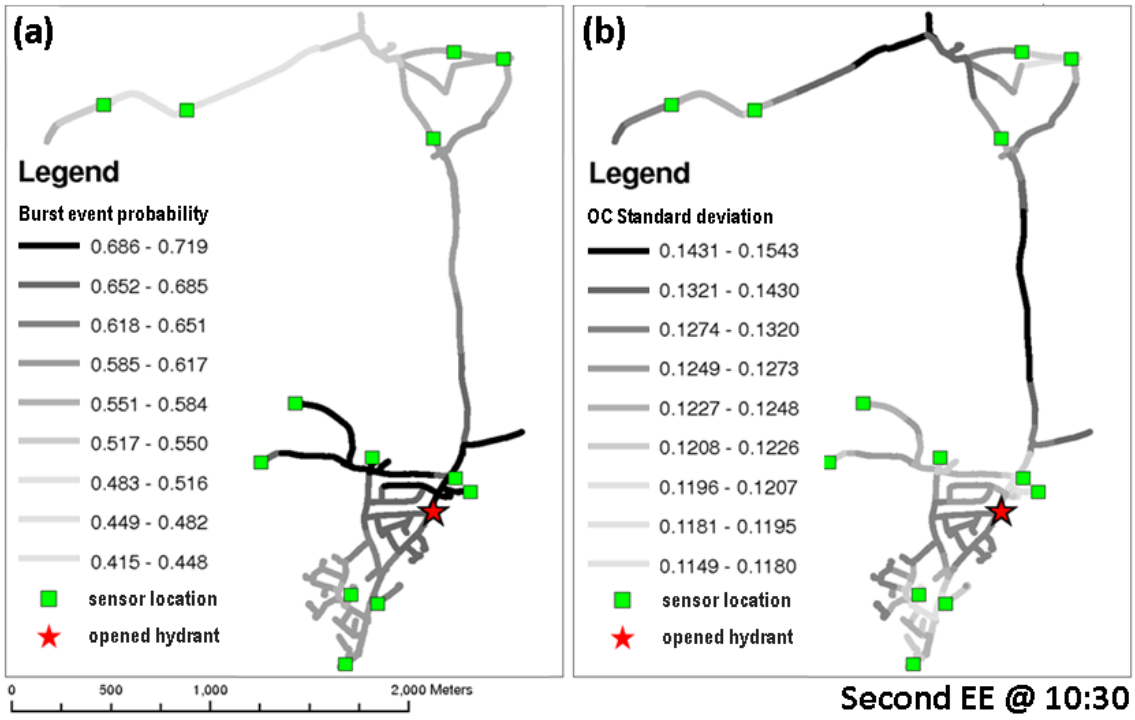


Figure 48. Approximate location results for the second Engineered Event. Burst event probability map (a), and Ordinary Cokriging standard deviation map (b).

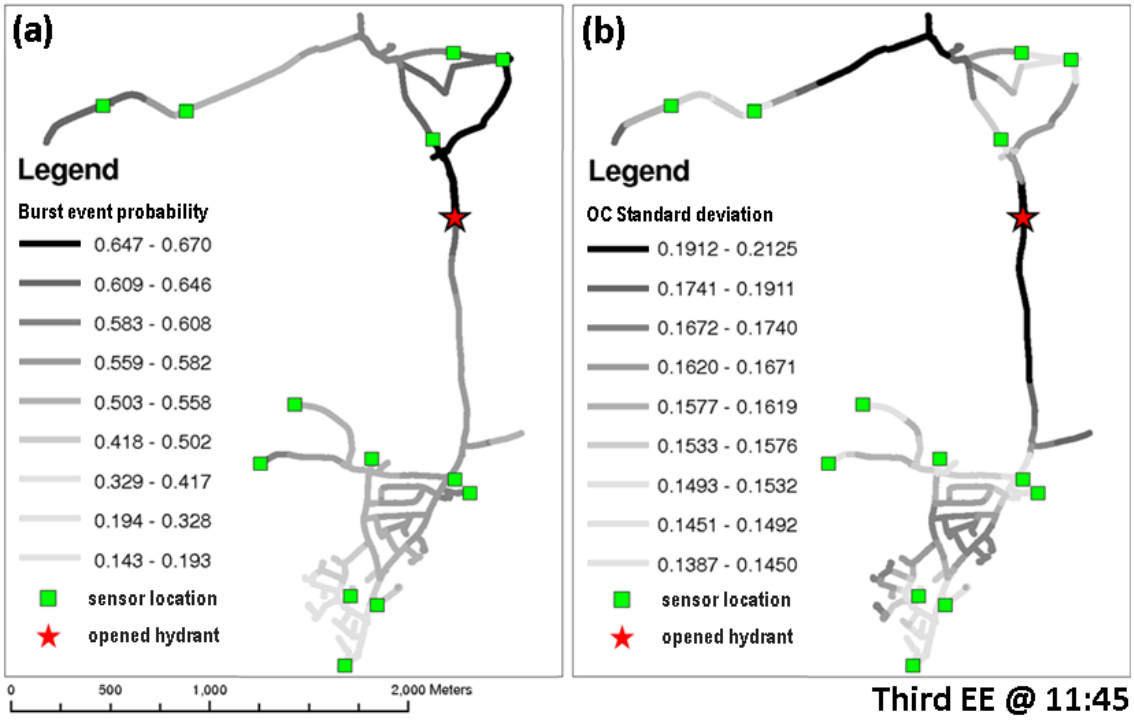


Figure 49. Approximate location results for the third Engineered Event. Burst event probability map (a), and Ordinary Cokriging standard deviation map (b).

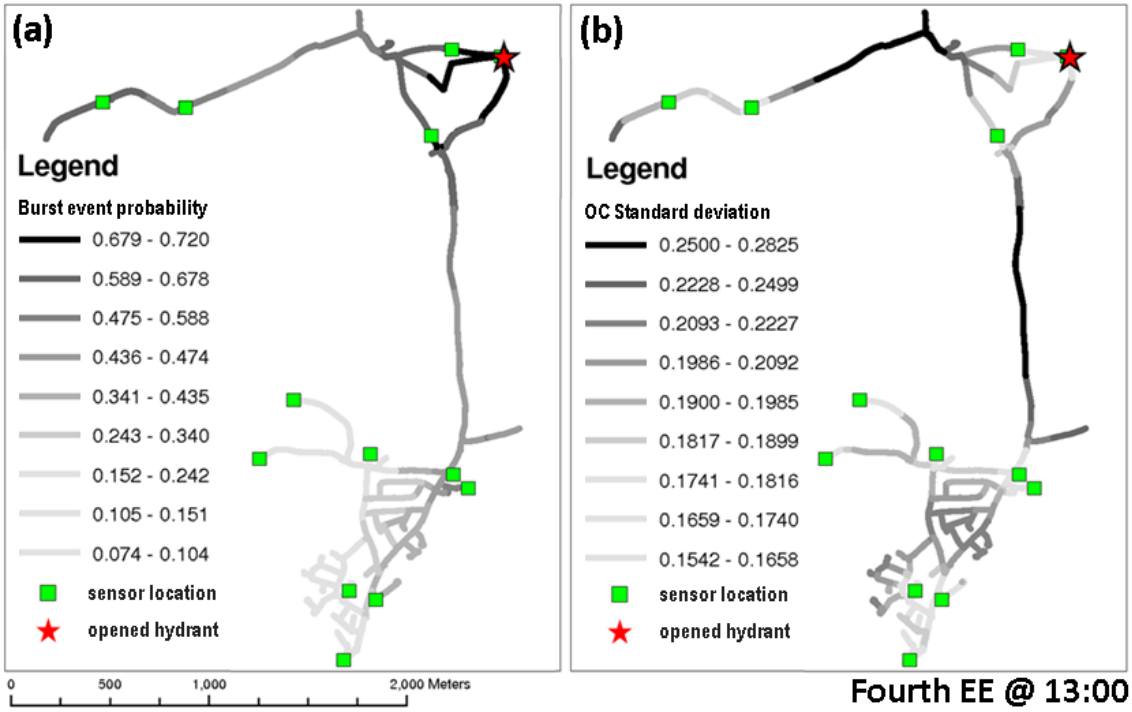


Figure 50. Approximate location results for the fourth Engineered Event. Burst event probability map (a), and Ordinary Cokriging standard deviation map (b).

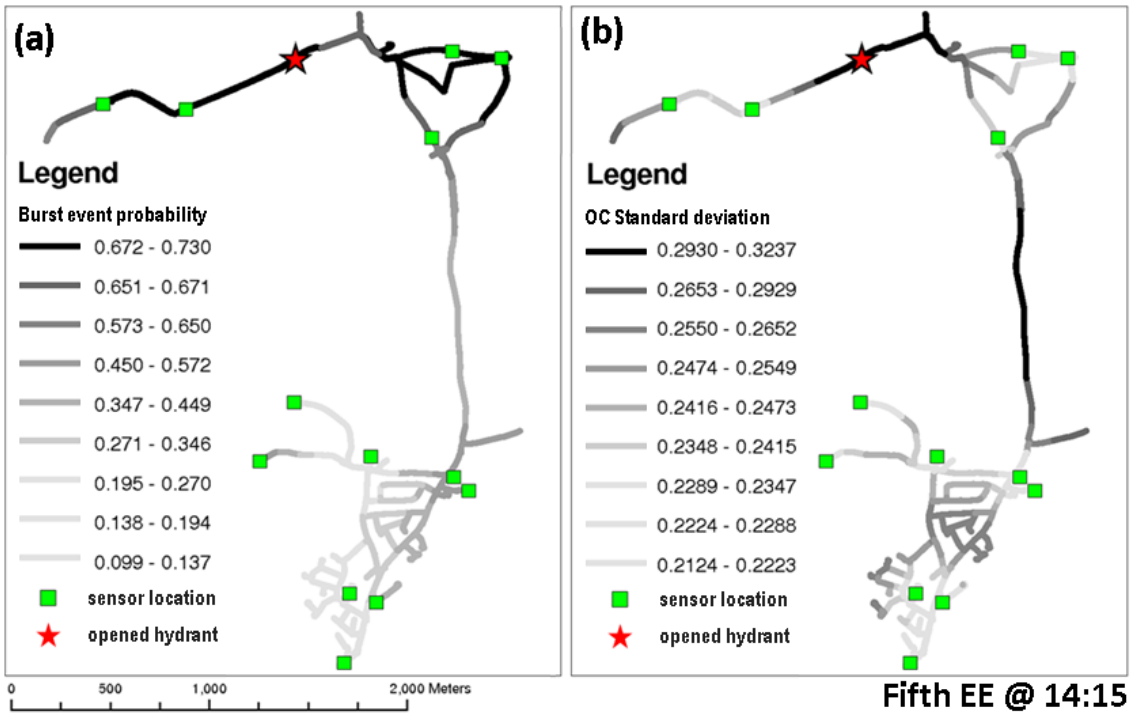


Figure 51. Approximate location results for the fifth Engineered Event. Burst event probability map (a), and Ordinary Cokriging standard deviation map (b).

It is important to note that the approximate pipe burst event location methodology effectiveness/accuracy depends on the number of pressure (and/or flow) devices deployed in the DMA – the more the better. Furthermore, it depends on the spatial layout of these devices within the DMA. To support these statements, four scenarios analysed are shown in Figure 52. Similarly to Figure 47, Figure 52 shows the location results obtained when the first pipe burst event simulated on the 7th of August 2008 was considered. Here, however, the figure is divided into four quadrants that refer to the results obtained by using pressure data from: (Figure 52a) 3 sensors located as in Scenario 1, (Figure 52b) 6 sensors located as in Scenario 2, (Figure 52c) 9 sensors located as in Scenario 3, and (Figure 52d) 6 sensors located as in Scenario 4. Similarly to Figure 47a, in each quadrant, the DMA pipes are grouped by using the Jenks Natural Breaks Classification method and coloured according to their burst event probability value. The locations of the deployed pressure sensors are indicated by using green squares and the real location of the hydrant opened in order to simulate the pipe burst event is indicated by a red star symbol.

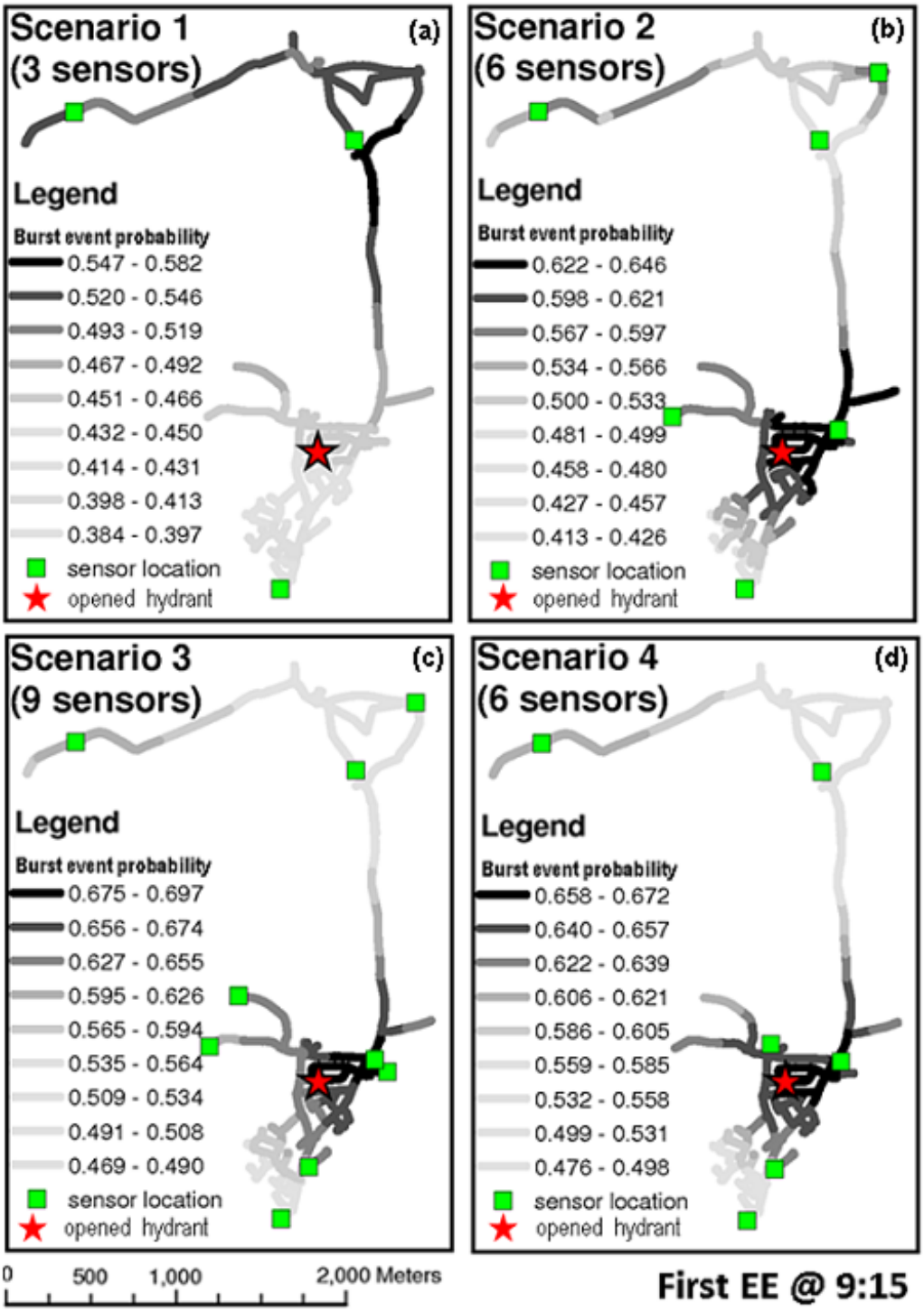


Figure 52. Approximate location results for the first Engineered Event. Pressure data from 3 sensors located as in scenario 1 (a), pressure data from 6 sensors located as in scenario 2 (b), pressure data from 9 sensors located as in scenario 3 (c), and pressure data from 6 sensors located as in scenario 4 (d).

By observing Figure 52, it can be seen that, in Scenario 1, the customised and further developed ERS was not able to identify the group of DMA pipes in the proximity of the 'failed' pipe. In the remaining three scenarios, on the other hand, the group of DMA pipes in the proximity of the 'failed' pipe was successfully identified. However, in Scenario 3, the customised and further developed ERS provided a more precise indication of the likely location of the 'failed' pipe than the one provided in either Scenario 2 or Scenario 4.

Furthermore, by comparing Scenarios 2 and 4, it is possible to observe that, although the same number of sensors was used, the spatial arrangement of the deployed sensors considered in Scenario 4 enabled the customised and further developed ERS to indicate the likely location of the 'failed' pipe more precisely. In view of this, it is worth highlighting that the use of the novel approximate pipe burst event location methodology could/should be supported by the development and use of a methodology for optimising the number and spatial arrangement of the sensors to be employed in order to achieve a required degree of location accuracy no matter where in the DMA the pipe burst event occurs. Bearing this in mind, note that this topic is beyond the scope of the work presented here and hence will not be discussed in greater detail.

5.6.3 Evaluating the methodology's event location capabilities when using different geostatistical techniques

In the analysis presented here, the customised and further developed ERS shown in Figure 37 was applied to the case study using, in turn, the four geostatistical techniques mentioned in Section 5.4.9.

Similarly to Figure 52, Figure 53 shows the location results obtained when the first, pipe burst event simulated on the 7th of August 2008 was considered. Here, however, the four quadrants refer to the results obtained by using: (Figure 53a) the OC, (Figure 53b) the OK, (Figure 53c) the LP, and (Figure 53d) the IDW geostatistical interpolation techniques. Again, similarly to Figure 52, each quadrant shows the DMA pipes grouped by using the Jenks Natural Breaks Classification method and coloured according to their burst event probability value. In each of these quadrants, the locations of the deployed sensors are indicated by using green squares and the real location of the hydrant opened in order to simulate the pipe burst event is indicated by a red star symbol.

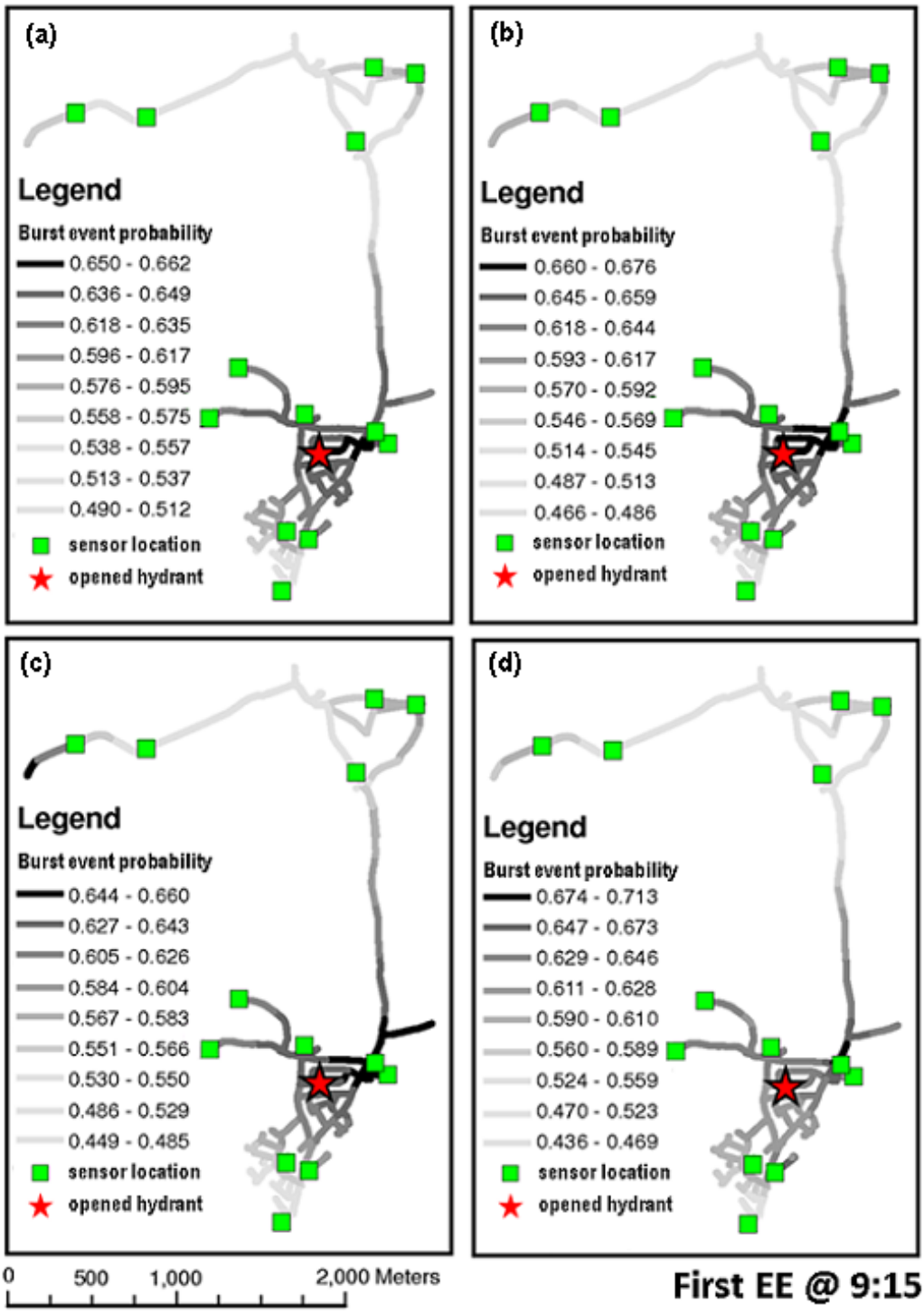


Figure 53. Comparison of the results obtained for the first Engineered Event when the Ordinary Cokriging (a), the Ordinary Kriging (b), the Local Polynomial (c), and the Inverse Distance Weighted (d) geostatistical interpolation techniques were used.

As it can be seen from Figure 53, when the OC or the OK geostatistical techniques were used, the group of DMA pipes in the proximity of the 'failed' pipe was successfully identified. This was not the case when the LP and IDW geostatistical techniques were used. Furthermore, by comparing the results shown in Figure 53a and Figure 53b it is possible to observe that the OC technique allowed obtaining a more precise indication of the likely location of the 'failed' pipe (see also Table 17 below). Indeed, the number of DMA pipes in the group of DMA pipes with the highest burst event probability values was reduced. As a consequence, the burst event search area was reduced as well. Note that similar results were obtained for the other four pipe burst events simulated on the 7th of August 2008 (not shown here).

Table 17. Root Mean Square Errors.

	Ordinary Cokriging	**Ordinary Kriging**	**Local Polynomial**	**Inverse Distance Weighted**
event 1	0.175	0.177	0.178	0.204
event 2	0.135	0.135	0.136	0.140
event 3	0.126	0.129	0.129	0.132
event 4	0.102	0.147	0.152	0.154
event 5	0.094	0.100	0.104	0.154

To support the above findings, Table 17 reports the values of the RMSE (i.e., performance indicator – see Appendix F Section F.6) calculated for each of the four geostatistical interpolation techniques tested and for all the pipe burst events simulated on the 7th of August 2008. This table shows clearly that the models (i.e., interpolation surfaces) obtained by using the OC technique outperformed the models obtained by using the other geostatistical techniques. Thus, it provides strong evidence that better models to predict the probability value of a burst event associated with each DMA pipe can be obtained by simultaneously considering the information from the correlated pipe burst event probability and pressure variables. This, in turn, leads to more accurate location of pipe bursts within a DMA.

5.6.4 Section summary

The methodology's event location capabilities have been tested and evaluated in this section. The use of different geostatistical interpolation techniques has also been

investigated. The main findings from the analyses performed in this section are briefly summarised below.

The use of the OC geostatistical interpolation technique enabled to successfully determine the approximate location of a number of simulated pipe burst events at different location within the DMA being studied.

The OC technique outperformed the other geostatistical interpolation techniques tested. Indeed, even though the use of the OK technique already allowed to successfully determine the approximate location of each considered simulated pipe burst event, the use of the OC technique allowed to determine it more accurately (i.e., reducing the number of potential 'failed' pipes).

By using pressure measurements only, the developed methodology successfully determined the approximate location of a number of simulated pipe burst events at different location within the DMA being studied.

5.7 Summary and Conclusions

The wider availability of more accurate and increasingly cheaper pressure/flow sensors allows us to envisage that in the near future a larger number of these devices (particularly pressure sensors) will be deployed in UK DMAs. The capabilities of the methodology that enables performing event detection and location based on the analysis of data from a larger (than currently used in the UK practice) number of sensors have been tested, evaluated and illustrated in this chapter. This has been done by using the customised and further developed ERS for processing the pressure data recorded during a series of EEs in an over-sampled UK DMA. Several sensitivity type data analyses that evaluated the methodology's performances for different choices of parameters and data analysis methods have been performed. Furthermore, issues such as the use of measurements sampled at sub-fifteen minute intervals and the value of the on-line pressure data for near real-time event detection and location have been investigated.

After the introduction in Section 5.1, the main aims and the rationale behind the analyses performed and presented in this chapter have been discussed in Section 5.2 and a description of the case study has been given in Section 5.3. The details of the methodology's practical implementation in the specific case study being analysed have

then been presented in Section 5.4. Finally, the results of the aforementioned investigations and sensitivity type data analyses have been reported in Sections 5.5 and 5.6. These results have shown the following: (i) the careful ANN *LagSize* selection and the use of the wavelet-based pressure signals de-noising improved the methodology's detection reliability and efficiency, (ii) the use of pressure measurements recorded at sub-fifteen minute sampling rates (e.g., 5, 3 or 1 minute) increased the methodology's detection speed considerably and did not negatively affect its reliability, however, little further detection speed improvement resulted from the use of pressure measurements sampled more frequently than 1 minute, (iii) the use of a set of Control Rules that analyses the discrepancy comparative over consecutive time steps improved the methodology's detection reliability and efficiency, (iv) the use a 'cumulative learning' procedure for the (re)calibration of the DMA level BIS parameters further improved the methodology's accuracy, reliability and efficiency, (v) the use of the OC geostatistical interpolation technique enabled to successfully determine the approximate location of a number of simulated pipe burst events at different location within the DMA being studied and outperformed the other geostatistical techniques tested in terms of event location accuracy (i.e., restricting the pipe burst event search area), and (vi) by using pressure measurements only, the developed methodology not only detected different types of simulated pipe burst event reliably and rapidly but also successfully determined their approximate location. This, in turn, provides evidence that the pressure data can play an important role in the context of near real-time event detection and approximate location.

This chapter has demonstrated that the customised and further developed ERS described in Section 3.5 enhances and further extends the capabilities of the ERS described in Section 3.4. Not only is the customised and further developed ERS able to reliably and rapidly detect the occurrence of a burst event but also to successfully determine its approximate location within the DMA. This has the potential to enable the water company personnel to more quickly locate bursts in a DMA which, in turn, may improve their customer service and reduce potential damages to third parties.

In the following chapter a summary of the research work carried out and presented in this thesis is given together with the thesis main conclusions and a discussion about the directions of the future research work.

CHAPTER 6 SUMMARY, CONCLUSIONS AND FUTURE WORK RECOMMENDATIONS

6.1 Thesis Summary

WDSs are critical infrastructures that increasingly large numbers of people rely on daily. As population grows, the demand on these critical infrastructures also grows and the need for their near real-time monitoring becomes vital to ensure timely responses to events such as pipe bursts and, thus, efficient and reliable operations. Sensing technology has advanced to the point that the deployment of dense networks of low-cost hydraulic sensors is now feasible. When combined with an appropriate data analysis methodology, the increased density and availability of measurements from these devices can provide the capability to detect and locate the pipe bursts/other events quickly and economically, thereby enabling to proactively manage the WDSs.

The research presented in this thesis resulted in the development of that appropriate data analysis methodology mentioned above. That is to say, it resulted in the development of a novel data analysis methodology capable of exploiting the increased density and availability of measurements from the hydraulic sensors for detecting the pipe bursts and similar events in a timely and reliable manner, and for successfully identifying their approximate location. The novel methodology developed works by simultaneously analysing the on-line pressure and flow signals from all the sensors deployed in a DMA in an efficient and automated manner. It does this by using a range of AI and (geo)statistical techniques which are data-driven and have the self-learning capability. The approach has been to first process the DMA signals in order to capture their patterns during the normal operating conditions of the DMA being studied. Then, to identify and estimate, in the observed DMA signals patterns, the deviations from the captured DMA signals patterns occurring in response to events such as pipe bursts. Finally, to infer the probability that an event has occurred based on the identified deviations, and to determine the approximate location of the event within the DMA based on the analysis of the extent to which each DMA signals pattern has been affected by the event occurring. The key concept is that the AI and (geo)statistical techniques are used here to effectively mimic the behaviour of a qualified and experienced human

operator when faced with the task of deciding if an abnormal event has occurred based on the available data from the hydraulic sensors deployed in a DMA and if yes, determining its approximate location.

The novel data analysis methodology developed has been implemented into a dataflow for on-line operation. This resulted in a computer-based ERS that can accept and store flow and pressure data from the DMA sensors and automatically convert it into useful operational information for the management of WDSs. The operational information provided by the ERS not only enables detecting the pipe bursts and other similar events at the DMA level and determining their approximate location within the DMA in near real-time but also determining the likely root cause of the events occurring and ranking/prioritising the alarms raised.

The novel methodology has been applied to the pressure/flow data from several DMAs in the UK with real-life bursts/other events and simulated (i.e., engineered) burst events. The results obtained have shown that it enables detecting the pipe bursts/other events more quickly and reliably than other existing statistical/AI-based techniques and it also improves over these techniques by providing a means to successfully determine the approximate location of these events within a DMA. The results obtained therefore demonstrate that the novel methodology developed and presented in this thesis has the potential to yield substantial improvements to the state-of-the-art in near real-time WDS incident management by facilitating prompt interventions and repairs.

6.2 Summary of the Contributions

The work carried out in this thesis forms a useful contribution to the research field. The main contributions of the work presented in this thesis are as follows:

- The development of a new methodology for on-line analysis of DMA pressure/flow signals that enables efficiently extracting and capturing information about the signals' expected patterns during the DMA normal operating conditions (i.e., NOPs). Specifically, this has involved the development and use of the following new methods:
  - A method for effective data pre-processing, including a novel heuristics-based method for dealing with the missing data;

- A method that makes use of a series of statistical tests for effectively filtering out outliers and pressure/flow measurements that are not consistent with the expected signal NOPs;
- A method for determining the statistical boundaries within which the observed signal patterns should lie assuming that no event has occurred in the DMA;
- A wavelet-based method for removing noise from the analysed pressure and flow signals;
- An ANN-based model for capturing the NOP of a pressure/flow signal and forecasting future signal values;
- An EA-based method for automatically selecting the optimal ANN input structure and related parameters.

- The development of a novel data analysis framework that enables identifying and estimating the deviations from a pressure/flow signal's NOP induced by an event (i.e., event occurrence evidence). Specifically, this has involved the development and use of three new complementary data analysis procedures that employ:
  - Control Rules for analysing, over consecutive time steps, the discrepancies between the observed DMA signal values and the signal values predicted by the relevant ANN model;
  - Control Rules for analysing, over consecutive time steps, the excursions of the observed DMA signal values outside the relevant, pre-defined, statistical boundaries;
  - Control Charts to monitor how the mean of the observed DMA signal values recorded during a particular time window during the day changes over time.

- The investigation and application of Bayesian logic for reasoning under uncertainty and combining multiple event occurrence evidence aiming at developing a new methodology for reliable and timely event detection, and for providing additional information useful for performing event diagnosis. Specifically, this has involved the development and use of:
  - A new Signal level BIS capable of: (i) synergistically processing the output information (i.e., event occurrence evidence) resulting from the various statistical/AI-based data analyses performed on a DMA pressure/flow signal, and (ii) encoding the temporal sequence of the incoming data, thereby effectively

enabling to update the probability of an event occurrence as evidence accumulates;

- A new DMA level BIS capable of synergistically processing the evidence resulting from the various statistical/AI-based data analyses coming simultaneously from all the DMA pressure/flow signals.

- The development of an EM-based new method for (re)calibrating the parameters of the DMA level BIS.
- The investigation and application of geostatistical techniques for approximate location of pipe burst events within a DMA. Specifically, this has involved the development of a novel data analysis method that makes use of a stochastic spatial interpolation technique for estimating the probability of a burst event occurrence associated with each DMA pipe.
- The encapsulation of the above methods into a novel computer-based ERS for on-line event detection and location. The novel ERS was tested and evaluated on three real-life UK case studies by considering both real-life pipe bursts/other events and simulated pipe burst events.

The research work presented in this thesis is based on the application of existing advanced technologies to solve a complex engineering problem. Bearing this in mind, these technologies, some of which have brought a real breakthrough in the AI and cognitive sciences fields, are synergistically combined in a novel methodological framework to efficiently and effectively detect and approximately locate events in the WDSs. Note also that the research work presented in this thesis has an interdisciplinary nature covering AI/(geo)statistical techniques and WDSs management processes such as data collection, manipulation and analysis for event detection and location. It should therefore be seen as an attempt to find links/synergies between the two fields and to provide an approach for integrating them, with the results demonstrating a viable way forward.

6.3 Conclusions

The main conclusion that summarises the research work presented in this thesis is as follows:

The novel ERS developed and presented in this thesis can be used in near real-time to reliably and in a timely and cost efficient manner detect pipe bursts and other events at the DMA level and, in cases where enough observed information is available, also approximately locate these events within the DMA. This in turn, enables the water company to respond more promptly to these events and if necessary, undertake necessary repairs so that the impact on the customer is minimised.

Other observations include:

- None of the existing techniques for the detection and location of pipe bursts and associated leakages seems ideal. The current best practical approach is the combined application of water audits/MNF monitoring and hardware-based techniques. This approach, however, is not suitable for quickly responding to the burst events. Conversely, the combined use of the ERS presented in this thesis and techniques that make use of specialised hardware devices would provide a much faster response to these events. Indeed, the ERS presented in this thesis could be used to enable a water company to become aware of the occurrence of a burst as soon as it occurs and, eventually, guide the personnel to the likely burst location. Then, the hardware-based techniques could be used to microlocate the burst.

- Many water companies are beginning to amass large volumes of pressure and flow data by means of modern SCADA systems. At the present, this information is mainly used for simple monitoring type activities and/or regulatory reporting. Manual analyses of these data are unfeasible. As a result, the available data are analysed by using non-standard approaches and irregularly. Furthermore, the data analysts do not generally look at individual time series for short-term events such as sudden pipe bursts. They are usually interested in pressure/flow values averaged over some time interval (e.g., one day or one month). In this scenario, the use of the technique presented in this thesis would be beneficial for enabling automating the mundane tasks involved in the analytical processing of the data as well as for extracting useful information (required for making reliable operational decisions) from the vast amount of often imperfect sensor data collected. When integrated with near real-time communications, the ERS presented in this thesis can provide the capability to automatically detect unusual (sudden or more gradual) changes in the process

variable patterns thereby effectively enabling a water company to proactively manage its WDS.

- The ERS presented in this thesis is particularly suited for working in an on-line environment. To this end, advantages of this system are as follows: (i) it is based on a data-driven paradigm, (ii) it has the self-learning capability, and (iii) it is robust to noisy or incomplete data, among the others. The first two advantages listed above are of particular relevance. This is because they enable the ERS developed to adapt to changes in the DMA operating conditions, to "tailor" itself to the particular DMA being studied and to learn from the past events occurred, thereby leading to a continuous improvement of its performance. Furthermore, the novel ERS is able to interface in a straightforward manner with existing monitoring systems. It makes use of the already available pressure and flow data or data that could be obtained by incrementally deploying more sensor devices in a DMA.

- The ERS presented in this thesis does not require a hydraulic or any other physically based simulation model of the WDS. It only relies on the empirical observation of the patterns and relationships which exist in the observed pressure/flow data without explicitly knowing the actual WDS dynamics. This represents a further advantage of this system because, at the present, the aforementioned simulation models often do not exist and/or are expensive to build/calibrate/maintain as they require more detailed WDS data.

- Unlike some other existing methods, the ERS presented in this thesis works well with both flow and pressure data, despite more noise being present in the latter than the former. The key reason for this is the presence of the novel wavelet-based de-noising method. This is very important as most of field measurements are pressure data (due to the much lower cost of pressure sensors when compared to flow sensors). Despite all the smart technologies built into the ERS, however, its performance is ultimately dependent on both quality and quantity of the sensor data available. Taking this into consideration, the good news is that an increased number of more accurate sensors is being installed in real-life WDSs as we speak.

6.4 Future Work Recommendations

The main goal for future work should involve further testing and validation of the proposed data analysis methodology. This should be done by applying it to a large number of DMAs presenting a wide range of operational conditions. The testing and validation of the methodology in an on-line environment should also be performed.

In addition to the above, some components of the work presented in this thesis could be extended or improved upon. These include:

1. Development of the methodology capability to handle cascading DMAs and identify events on the trunk mains;
2. Improvement of the methodology capabilities to perform approximate event location within a DMA;
3. Improvement of the methodology capabilities to rank the raised alarms;
4. Improvement of the methodology capabilities to identify the likely root cause of an event occurring;
5. Transfer of the methodology developed here to other fields – e.g., detection of (un)intentional contamination events in water distribution systems or toxic events in wastewater networks.

With regard to point 1, it has to be noted that the existing methodology returns multiple alarms if an event occurs in cascading DMAs or on a trunk main, hence cannot automatically tell where the event occurred. This could be improved, for example, by using a set of heuristic rules based on engineering knowledge, the pipeline system schematics and the ERS output. All this may offer improved interpretation of results reducing the potential of false alarms and the overall number of alarms.

With regard to point 2, it has to be noted that the existing methodology can approximately locate an event within a DMA based exclusively on the spatial analysis of the output of the various Signal level BISs and the measured pressure at the sensors locations. This could be improved, for example, by making use of simple HMs of the pipeline system and burst prediction models based on the characteristics of the pipes (e.g., age, material, diameter, etc.) along with historical data of burst occurrences.

With regard to point 3, it has to be noted that the existing methodology offers limited alarm ranking capabilities. This could be improved, for example, by adopting a risk-based approach whereby, for each event being detected, a suitable risk score is calculated as a function of the following two risk components: (a) the DMA level event occurrence probability (i.e., the probability of an event occurrence), and (b) the estimate of the event magnitude/severity (i.e., the event consequence). Factors such as the total number of users in a DMA, the number of vulnerable users in a DMA, and the time the event occurs could be taken into consideration as well. The calculated risk scores could then be used to rank the alarms (i.e., detected events occurring in different DMAs) and thus enable the water company personnel to better prioritise responses.

With regard to point 4, it has to be noted that the existing ERS offers limited capabilities to discriminate between event types. This could be improved, for example, by using a set of heuristic rules based on engineering knowledge of the pipeline system and the ERS output. All this may offer improved interpretation of results enabling the water company operators to make more informed intervention decisions.

With regard to point 5, it has to be noted that the methods developed and presented here have a generic nature. Provided domain-specific knowledge is taken into consideration, the ERS methodology could be applied not only to WDSs but to any type of distribution or collection system having sensors at a number of locations in the network measuring parameters such as flow, pressure, quality or the flow of data itself.

APPENDIX A WAVELET ANALYSIS

A.1 Introduction

A wavelet is a mathematical function which, when plotted, resembles a 'little wave' with amplitude that starts out at zero, increases, and then decreases back to zero. Wavelets are useful for decomposing data into 'wavelet coefficients' (see below), which can then be processed in a way that depends on the aim of the analysis. Wavelets have been widely applied in the last two decades in many different fields, such as signal and image processing, data compression, communication, computer graphics, astronomy, quantum mechanics, and turbulence. A discussion of these and other areas of application can be found in Ruskai (1992) and Jaffard et al. (2001). Other important fields of application are numerical analysis (see – e.g., Cohen, 2003) and statistics (see – e.g., Vidakovic, 1999). In the fields of hydrology and water resources research, wavelets have been applied to study the temporal variability of the rainfall-runoff relationship (Labat et al., 2000), heat fluxes under unstable atmospheric conditions (Katul and Parlange, 1995), large and small scale features of spatial rainfall fields (Kumar and Foufoula-Georgiou, 1993), and pressure signal analysis for leak detection and location (Ferrante and Brunone, 2001, 2003b; Brunone and Ferrante, 2004; Ferrante et al., 2005, 2007), among the others.

In the following two sections of this appendix, an introduction about the general wavelet theory is given first. Note that this introduction focuses only on the application of the WA to the processing of non-stationary signals (such as those coming from a DMA). More comprehensive mathematical treatment of this subject can be found in several good books (e.g., Daubechies, 1992; Meyer and Ryan, 1996; Meyer and Roques, 1993). This is then followed, in Section A.3, by a description of the procedure for performing the wavelet-based de-noising of signals, which is of particular interest for the research work presented in this thesis. In this latter section, the implementation details of the de-noising procedure used in the developed ERS are presented.

A.2 Wavelet Analysis

Generally speaking, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. They are compactly or almost compactly supported, and they integrate to zero. This is in contrast to 'big waves' – i.e., sines and

cosines in Fourier analysis (see – e.g., Stein and Shakarchi, 2003), which also oscillate, but the amplitude of their oscillation never changes. Figure A.1 graphically shows this difference.

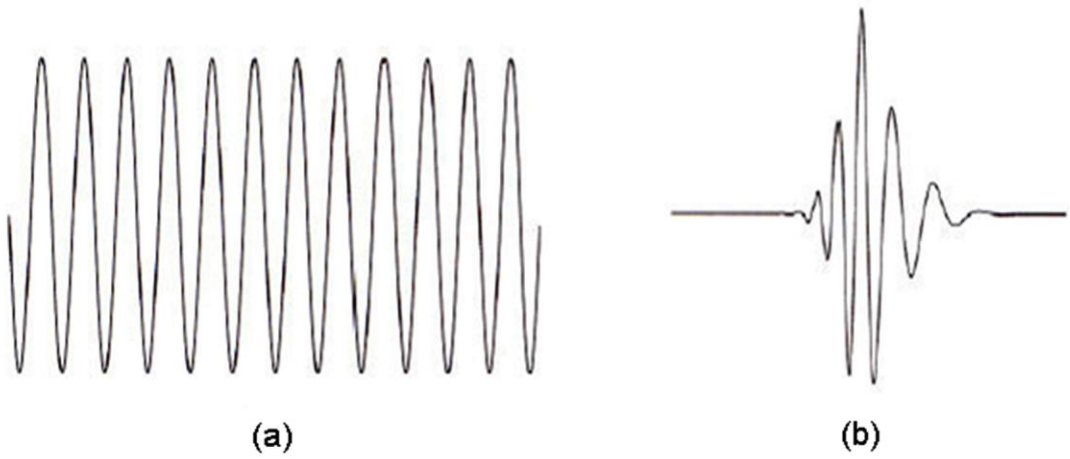


Figure A.1. Difference between a wave (a) and a wavelet (b).

Signal processing by using the WA provides a way of analysing the frequency content of a signal locally in time. This means that the properties of a signal can be studied simultaneously in the frequency and time domains (i.e., dual frequency-time representation). In the processing of non-stationary signals, the application of the WA has been found to be very successful (e.g., Nason and Sachs, 1999; Percival, 2000), presenting better results than traditional Fourier and Short-Time (i.e., Windowed) Fourier analyses (Gabor, 1946).

Fourier analysis is used here as a starting point to introduce the Wavelet Transform (WT), and as a benchmark to demonstrate cases where the WA provides a more useful characterisation of signals.

Fourier analysis breaks down a signal into constituent sinusoids of different frequencies. Mathematically, the process of Fourier analysis is represented by the Fourier Transform (FT):

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} f(t)\, dt \qquad \text{(A.1)}$$

which is the integral over all values of time of the signal $f(t)$ multiplied by a complex exponential (which can be broken down into real and imaginary sinusoidal components). For each angular frequency ω, the integral over all values of time produces a scalar, $\hat{f}(\omega)$. These complex-valued scalars are the Fourier coefficients. Conceptually, multiplying each Fourier coefficient $\hat{f}(\omega)$ by a complex exponential (sinusoid) of frequency ω yields the constituent sinusoidal components of the original signal. If a signal contains significant oscillations at an angular frequency of ω_0 the absolute value of $\hat{f}(\omega_0)$ will be large. By examining a plot of $|\hat{f}(\omega)|$ as a function of the angular frequency, it is possible to determine what frequencies contribute most to the variability of $f(t)$. Note that the FT maps a function of a single variable into another function of a single variable. Taking this into account, another way to think of Fourier analysis is as a mathematical technique for transforming the view of the signal from the time domain to the frequency domain. For many signals, this is extremely useful because the signal's frequency content is of great importance. However, Fourier analysis has a serious drawback. In transforming to the frequency domain (i.e., integrating over all time), the time information is lost. That is, when looking at the FT of a signal, it is impossible to tell when a particular event took place. Most signals contain numerous non-stationary or transitory characteristics (i.e., drift, trends, abrupt changes, and beginnings and ends of events). These characteristics are often the most important part of the signal and Fourier analysis is not suited to detect them.

To overcome this drawback the Short-Time FT has been proposed by Gabor (1946). The author adapted the FT to analyse only a small section of the signal at a time. The technique works by choosing a time function that is nonzero only on a finite interval. As an example, consider a compactly supported function $g(t)$ centred around t=0. Note that a very common choice for g is the Gaussian function. Shifting this function by τ centres it around τ. Multiplying a signal by $g(t-\tau)$ selects a portion of the signal centred at τ. Taking the FT of these windowed segments for different values of τ, produces the Short-Time FT. Mathematically, this is:

$$\bar{f}(\omega,\tau) = \int_{-\infty}^{+\infty} e^{-i\omega s} f(t)\, g(t-\tau) dt \qquad \text{(A.2)}$$

The Short-Time FT maps a function of one variable into a function of two variables, ω and τ. Therefore, the Short-Time FT represents a sort of compromise between time and frequency -based views of a signal. It provides some information about both when and at what frequencies an event occurs. However, this information can only be obtained with limited precision. This precision is determined by the size of the window. Once a particular size for the time window has been chosen that window is the same for all the frequencies. Most signals require a more flexible approach whereby the window size can be varied to determine more accurately either time or frequency.

Based on the above, WA represents the next logical step. Indeed, WA can be seen as a windowing technique with variable-sized regions. The window is shifted along the signal and, for every position, the spectrum is calculated. This process is then repeated many times with a slightly shorter (or longer) window for every new cycle. In the end, the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations, it is possible to speak of a multiresolution analysis. However, in the case of wavelets it is normal to speak about time-scale representations not time-frequency representations, scale being (in a way) the opposite of frequency. In conclusion, WA allows the use of long time intervals where more precise low-frequency information is required, and shorter time intervals where high-frequency information is required.

Mathematically, the process of WA is represented by the Continuous Wavelet Transform (CWT):

$$\check{f}(s,\tau) = \int_{-\infty}^{+\infty} f(t)\ \psi^*_{s,\tau}(t)\, dt \qquad \text{(A.3)}$$

where $\check{f}(s,\tau)$ is the wavelet coefficient and '*' denotes the complex conjugation. This equation shows how a signal $f(t)$ is decomposed into a set of basis functions $\psi_{s,\tau}(t)$ called the wavelets. The variables s and τ are the new dimensions, scale and translation, after the CWT. The wavelets are generated, by scaling and translation, from a single basic wavelet $\psi(t)$ (i.e., the mother wavelet):

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}}\ \psi\left(\frac{t-\tau}{s}\right) \qquad \text{(A.4)}$$

where the factor $\frac{1}{\sqrt{|s|}}$ is for energy normalisation across the different scales s. Similarly to the FT, the CWT seeks level of similarity between $f(t)$ and the wavelet function at different scales and positions and generates a wavelet coefficients contour map (also known as a scalogram). Unlike FT which (as previously mentioned) breaks up a signal into sine waves of various frequencies, however, the WA breaks up a signal into scaled (i.e., stretched/compressed) and shifted (i.e., delayed/hastened) versions of the mother wavelet. Figure A.2 shows a graphical representation of the difference between FT and WT.

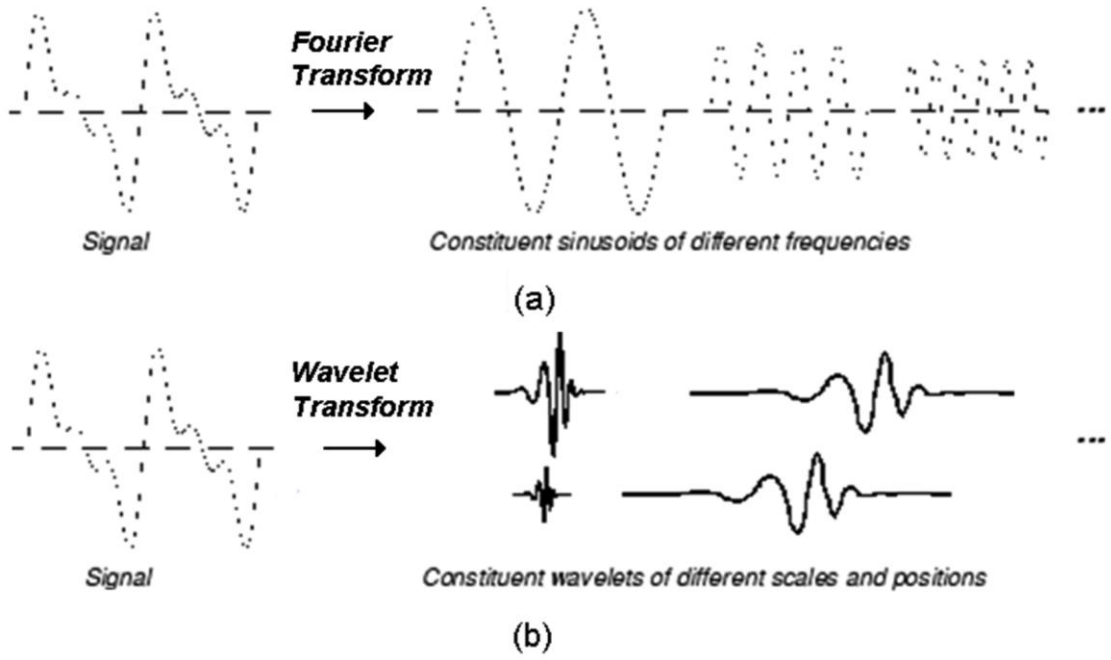


Figure A.2. Difference between Fourier Transform (a) and Wavelet Transform (b).

It can be shown (see – e.g., Sheng, 1996) that a mother wavelet satisfying the admissibility condition:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|}\, d\omega < +\infty \qquad \text{(A.5)}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$ and C_ψ is the admissibility coefficient, can be used to first analyse and then reconstruct a signal without loss of information. The admissibility condition implies in particular, that:

$$\int_{-\infty}^{+\infty} \psi(t)\, dt = 0 \qquad \text{(A.6)}$$

Equation (A.5) means that $\psi(t)$ should be localised in frequency. However, Equation (A.6) means that $\psi(t)$ is localised in time and also oscillatory. Hence, the name 'wavelet'.

Note that a wavelet is said to have *n* vanishing moments if:

$$\int_{-\infty}^{+\infty} t^k \psi(t)\, dt = 0 \qquad \text{(A.7)}$$

for $k \in \{0, 1, \ldots, n\}$.

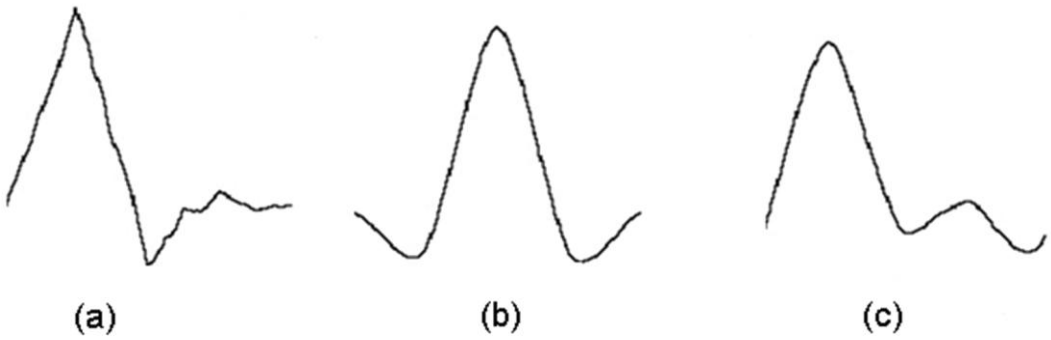


Figure A.3. Examples of commonly used orthogonal wavelets in the Daubechies (a), Coiflet (b), and Symlet (c) families.

Different families of wavelets, whose qualities vary according to several criteria, exist. The main criteria are: (1) the speed of convergence to zero of these functions when the time or the frequency goes to infinity - which quantifies both time and frequency localisations, (2) the symmetry, (3) the number of vanishing moments, and (4) the regularity. The most commonly used wavelets are the orthogonal ones, such as the Daubechies, Coiflet or Symlet wavelets (see – e.g., Addison, 2002). Examples of

wavelets in these families are shown in Figure A.3. Further details on these and other families of wavelets can be found in – e.g., Vidakovic (1999).

If the admissibility condition is satisfied, then the following inverse formula (i.e., 'resolution of identity') holds:

$$f(t) = \frac{1}{C_\psi} \iint_{-\infty}^{+\infty} \check{f}(s,\tau)\ \psi_{s,\tau}(t)\ \frac{1}{s^2}\, ds\, d\tau \tag{A.8}$$

which represents the Inverse Continuous Wavelet Transform (ICWT).

CWT generates large amount of data for all s and τ (i.e., redundant information). Discrete wavelets can be used to reduce this amount of data, thereby resulting in more efficient data analysis. Discrete wavelets are not continuously scalable and translatable but can only be scaled and translated in discrete steps. This is achieved by modifying the wavelet representation in Equation (A.4) to create (see – e.g., Daubechies, 1992):

$$\psi_{j,i}(t) = \frac{1}{\sqrt{|s_0^j|}} \psi\left(\frac{t - i\tau_0 s_0^j}{s_0^j}\right) \tag{A.9}$$

where j and i are integers that control the wavelet scale/dilation and translation, respectively; s_0 is a specified fine scale step greater than 1; and τ_0 is the location parameter, which must be greater than zero and depends on the dilation step. The most common choice for these parameters is: s_0=2 and τ_0=1. This power-of-two logarithmic scaling of the dilations and translations is known as dyadic grid arrangement. It is the simplest discretisation choice and it is particularly convenient as it enables a fast implementation of the DWT and of the IDWT.

A framework for this fast decomposition/reconstruction of signals is the discrete wavelet multiresolution analysis introduced by Mallat (1989). Basically, the discrete wavelet multiresolution analysis (commonly based on the Daubechies orthogonal wavelets – Daubechies, 1992) allows decomposing a signal into a number of sub-signals, namely approximations (i.e., high-scale, low-frequency components) and details (i.e., low-scale, high-frequency components), by using a filter bank composed of both Low-Pass (L-P) and High-Pass (H-P) filters. The approximation sub-signals represent the background information of the data and the detail sub-signals capture the small

features of interpretational value in the data. This process can be repeated *m* times, producing *m* levels of decomposition, but decomposing the approximations only. Indeed, the details are neglected. In this way, a signal can be first decomposed into an approximation A_1 and a detail D_1 (that is the level 1 of the decomposition). Then A_1 can be decomposed into an approximation A_2 and a detail D_2 (that is the level 2 of the decomposition) and so on. Considering *m* levels of decomposition, the reconstruction process allows recovering the initial signal by summing the *m* details and the approximation A_m of level *m*. As an example, Figure A.4 shows a 3-level decomposition of a signal and its reconstruction.

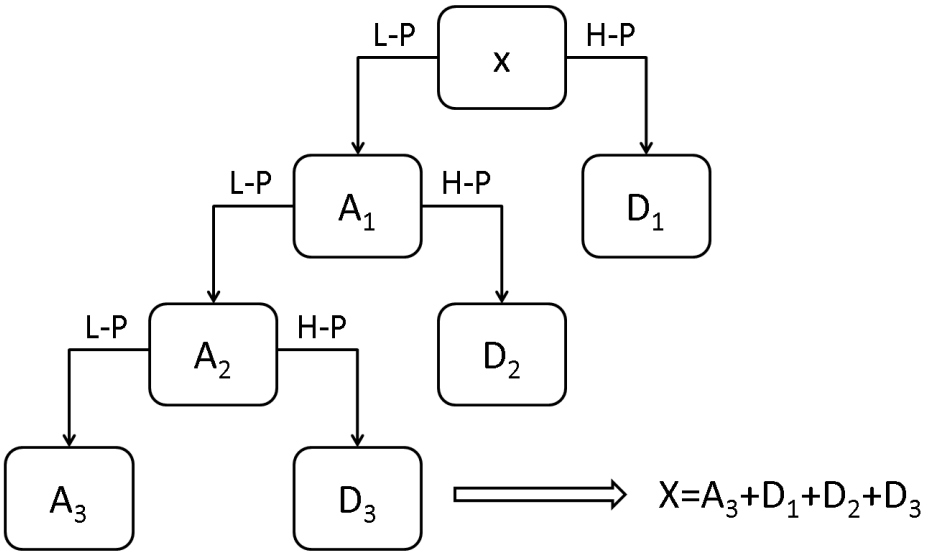


Figure A.4. Multiresolution analysis leading to a 3-level decomposition of a signal x.

A.3 Wavelet-based De-noising

Removing noise from a signal by using wavelets is an approach that shows superior performance when compared to other signal de-noising algorithms that have been proposed in the literature (e.g., Least-Square Spline Fitting method, Kalman Filters, and Adaptive Filtering based on ANNs). Indeed, it has been found to be asymptotically near optimal for a wide class of signals, especially if corrupted by white Gaussian noise (Donoho and Johnstone, 1994, 1995). The main reason for its success is that in the wavelet domain, signal energy is dominated by a few large-amplitude wavelet

coefficients, whereas noise energy is distributed uniformly over all wavelet coefficients, especially in the detail components.

In general a wavelet-based de-noising procedure is carried out in three steps as follows:

1. Select a mother wavelet and a decomposition level and perform DWT of the original signal at that level (Mallat, 1989). Note that, as described in Section A.2, this results in several approximation and detail sub-signals;
2. Select a threshold and a thresholding method and apply the selected threshold to the detail sub-signals' coefficients;
3. Perform IDWT using the original approximation sub-signal's coefficients (at the chosen decomposition level) and the modified detail sub-signals' coefficients.

Assuming additive white Gaussian noise, any signal $y(t)$ can be represented by the summation of the 'real signal' $x(t)$ and the noise $e(t)$:

$$y(t) = x(t) + e(t) \tag{A.10}$$

where time t is equally spaced. The objective of this model is to remove noise by suppressing the noise part of the signal y and to recover x.

The traditional approach to noise removal models noise as a high frequency signal added to the 'real signal'. FT could be used to identify this high frequency and remove it by adequate filtering. However, when the 'real signal' has important information associated with the same frequency as the noise, filtering out this frequency is likely to induce a noticeable loss of information. Since pressure and flow time series data from a DMA contain numerous non-stationary or transitory characteristics (e.g., drifts, seasonal trends, abrupt changes, etc.), the wavelets approach is more appropriate due to the fact that (as also mentioned in Section 3.4.3.5) a pressure/flow signal is studied using a dual frequency-time representation. This allows noise to be removed from signal frequencies that are likely to contain important information.

After applying DWT to $y(t)$ the following is obtained:

$$y_{s,k} = x_{s,k} + e_{s,k} \tag{A.11}$$

where $y_{s,k}$ is the k^{th} wavelet coefficient in the scale s. Once this is done, a suitable strategy for noise removal would consists in making the coefficients associated with the noise as equal to zero. Strategies for doing so differ in the way these coefficients are tracked and taken out from the representation. The most popular of these strategies is the one that makes use of the Universal Threshold and a Soft or Hard Thresholding method (Donoho and Johnstone, 1994, 1995; Donoho, 1995).

The Universal Threshold is based on the statistical properties of the white Gaussian noise and is calculated as follows:

$$\lambda_{Universal} = \sigma\sqrt{2log(d)} \tag{A.12}$$

where d is the length of the signal and σ is the standard deviation of the noise. As the actual noise standard deviation is typically unknown, a measure of the noise level is estimated from the data by using the following formula:

$$\sigma = \frac{MAD}{0.6745} \tag{A.13}$$

where *MAD* is the Median Absolute Deviation (see – e.g., Everitt, 2006) of the absolute value of the detail coefficients of the first level of decomposition. Note that σ is obtained by dividing the *MAD* by 0.6745 so that, for white Gaussian noise, the correct standard deviation would be returned.

Once the Universal Threshold has been calculated, it is necessary to decide how to apply it to the detail sub-signals' coefficients. The so-called Hard Thresholding method sets to zero the coefficients whose absolute values are smaller than the threshold and leaves the other ones unchanged. In contrast, the Soft Thresholding method first sets to zero the coefficients whose absolute values are smaller than the threshold, and then shrinks (by a quantity as equal to the threshold value) the non-zero coefficients towards zero. Mathematically, let λ denote the threshold. The "Hard Thresholded" signal is y if $|y| > \lambda$, and is 0 if $|y| \leq \lambda$. The "Soft Thresholded" signal is [sign $(y)(|y| - \lambda)$] if $|y| > \lambda$ and is 0 if $|y| \leq \lambda$. An example showing how these methods work is shown in Figure A.5. As it can be observed from this figure, the Hard Thresholding procedure creates discontinuities at $y = \pm\lambda$, while the Soft Thresholding procedure does not.

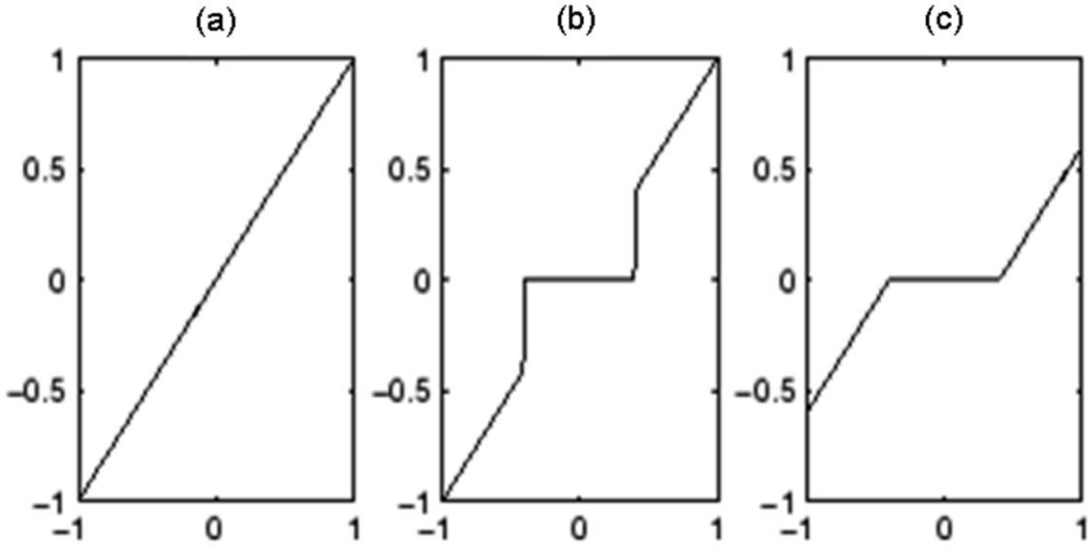


Figure A.5. Hard and Soft Thresholding methods applied to a simple signal. Input data scaled in the interval [-1, 1] and λ=0.5. Original signal (a), Hard Thresholded signal (b) and Soft Thresholded signal (c).

Finally, IDWT is performed as detailed in the third step of the general wavelet-based de-noising procedure outlined above.

APPENDIX B ARTIFICIAL NEURAL NETWORKS

B.1 Introduction

ANNs are systems based on the operation of the biological neural networks, in other words, they are an emulation of biological neural systems. ANNs are networks of many simple processor units (i.e., neurons). These units are connected by communication channels (i.e., connections). Each connection is associated with a real number (i.e., weight), which may be determined from a set of training examples and represents the connection strength between neurons. ANNs may allow the modelling of complex relationships between inputs and outputs and learning these relationships directly from the data being modelled (Haykin, 1994).

Over the past decades, ANNs have been successfully applied in a wide range of problem domains. These include pattern classification and clustering (e.g., Bishop, 1995), forecasting (e.g., Weigend and Rumelhart, 1992), and non-linear modelling (e.g., Haykin, 1994). ANNs have also been successfully applied to many water resources modelling and forecasting problems. For example, Romano et al. (2009) used ANNs for forecasting the arrival time and high of tsunami waves. Bowden et al. (2002) applied ANNs for forecasting salinity in a river. Liong et al. (2000) used ANNs for river stage forecasting. Note that a comprehensive review of papers dealing with the use of ANNs for the modelling and forecasting of water resources variables can be found in Maier and Dandy (2000).

As an ensemble, the above studies show that, in the majority of cases, the ANN approach outperformed the traditional statistical approaches to the given problem. For example, with regard to the ANNs' efficiency in modelling and forecasting water demand, Jain et al. (2001), Jain and Ormsbee (2002), and Bougadis et al. (2005) observed that ANN models outperform regression and time series analysis. Similarly, Adamowski (2008) developed and compared relative performance of: (i) 39 Multiple Linear Regression models, (ii) 9 Auto Regressive Integrated Moving Average models, and (iii) 39 ANN models. The study concluded that the ANN models perform the best.

When the specific problem of pipe bursts and associated leaks in WDSs is considered, several researchers have used ANNs for predicting the pipe burst rate (e.g., Skipworth et

al., 2000) and for assessing the probability of the pipe burst risk (e.g., Babovic and Drecourt, 2000). Furthermore, as discussed in Chapter 2, several authors have explored the possibility of using ANNs for the detection and location of these events (e.g., Gabrys and Bargiela, 1999; Mounce et al., 2002, 2007, 2010a; Mounce and Machell, 2006, 2010; Caputo and Pelagagge, 2002, 2003; Aksela et al., 2009).

In this appendix several ANN topics relevant to the research work carried out in this thesis are discussed. The ANNs theoretical background is given in Section B.2. In this section, an analogy between biological and artificial neurons is made first in order to explain how an artificial neuron model works and introduce the most commonly used activation functions. Then the different ANN structures and learning paradigms are examined. The focus being on the Jordan (Jordan, 1986a, 1986b) and Elman (Elman, 1990) ANNs, and on the different algorithms that are used in supervised learning. In Section B.3, the details about the FF ANNs and the BP algorithm (Rumelhart et al., 1986) are given. Finally, in Section B.4 the overfitting and generalisation issues are considered.

B.2 Theoretical Background

B.2.1 Artificial neuron model

As mentioned in Section B.1, ANNs emulate biological neural systems. A biological neuron consists of a dendrite, a cell body, and an axon. This is shown in Figure B.1a. The connections between the dendrite and the axons of other biological neurons are called synapses. Electric pulses coming from the other biological neurons are translated into chemical information at each synapse. The cell body inputs these pieces of information and fires an electric pulse if the sum of the inputs exceeds a certain threshold. Bearing this in mind, it is possible to state that the main job of a biological neuron is to output electric pulses according to the sum of multiple signals from the other neurons with the characteristics of a Pseudo-Step function. A second job involves changing the transmission rate at the synapses to optimise the whole neural network (i.e., the network consisting of the various biological neurons, which is the most essential part of the brain activity).

As it can be observed from Figure B.1b, an artificial neuron model simulates multiple inputs and one output, the switching (i.e., activation) function of the input-output

relationship, and the adaptive synaptic weights. Each unit performs a relatively simple job, it receives the inputs from its neighbours or external sources and uses these to compute (i.e., weighted sum of the inputs) an output signal which is propagated to the other units. As it will be shown below, apart from this processing, a second task is the adjustment of the weights.

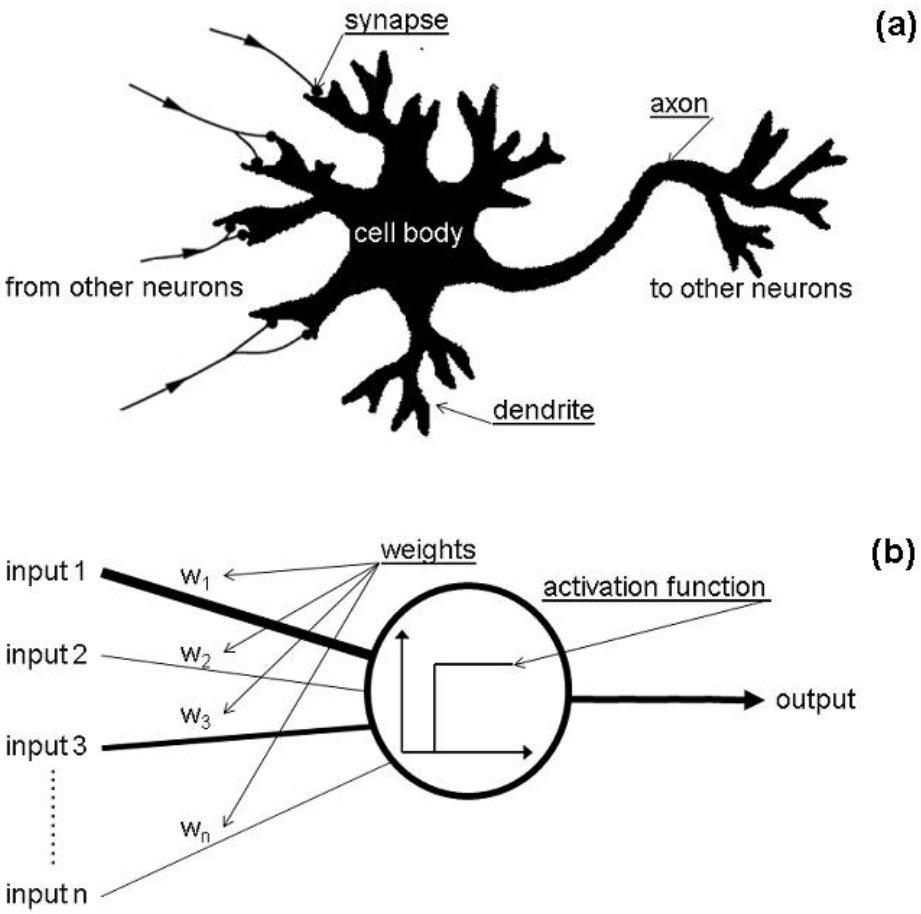


Figure B.1. A biological neuron (a), and an artificial neuron model (b).

The first artificial neuron model proposed by McCulloch and Pitts (1943) used a Step function for the activation function. However, due to the constraints of the binary on/off signals, the perceptron (Rosenblatt, 1958) - which is an ANN consisting of this type of neurons - has limited capability. Figure B.2 shows several continuous functions that are commonly used, nowadays, as activation functions together with their derivatives. The use of these functions results in higher performance of the ANNs.

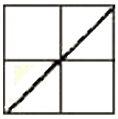
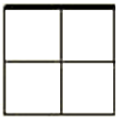
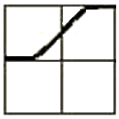
Name	Function $y = f(x)$	Derivative $\frac{\partial y}{\partial x}$
Linear	x 	1 
Logistic	$\frac{1}{1+e^{-x}}$ 	$y(1-y)$ 
Hyperbolic Tangent	$Tanh(x)$ 	$1-y^2$ 

Figure B.2. Various differentiable activation functions and their derivatives.

B.2.2 Types of Artificial Neural Networks

Many ANN structures and learning paradigms have been proposed. Typical ANN structures include FF and Recurrent (or Feed Back) ANNs. Learning paradigms can be categorised into supervised and unsupervised learning. This section provides an overview of these structures and paradigms.

Artificial Neural Network structures

Two major kinds of ANN structures (or network topologies) can be identified based on the pattern of the connections between the units and the direction of the data propagation. These are the FF and the Recurrent ANNs. Figure B.3 shows a representation of these two ANN structures.

In a FF ANN, the data flow in one direction only and the connections between units do not form cycles (i.e., connections extending from outputs of units to inputs of units). A FF ANN acts as a filter which outputs the processed input signal. FF ANNs usually produce a response to an input quickly. Note that FF ANNs will be discussed further in Section B.3.

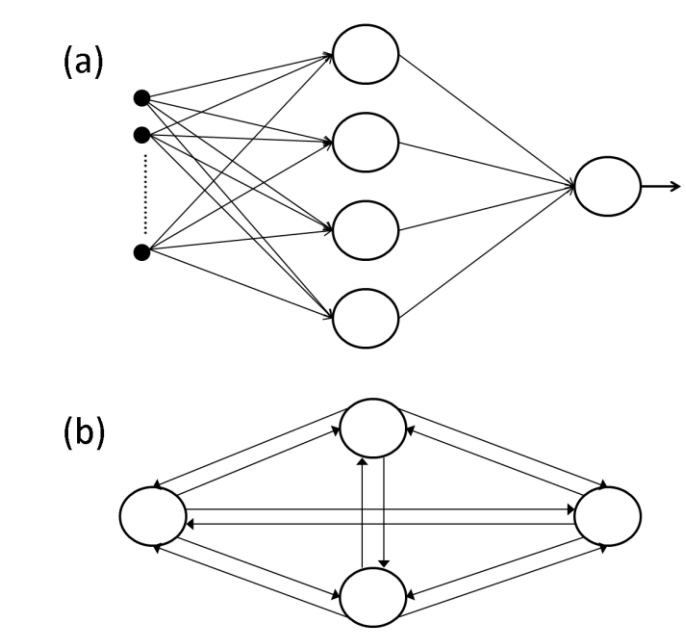


Figure B.3. A Feed Forward Artificial Neural Network (a), and a Recurrent Artificial Neural Network (b).

The Recurrent ANNs have connections between network outputs and some or all the other neuron units. Certain unit outputs are used as activated inputs to the network, and other unit outputs are used as network outputs. In some Recurrent ANNs, each time an input is presented, the ANN must iterate for a potentially long time before it produces a response. Also, they are usually more difficult to train than FF ANNs and, according to Takagi (Takagi, 1997), the following has to be noted. Due to the feed-back, there is no guarantee that the networks become stable. Some networks converge to one stable point, and other networks become limit-cycle, chaotic or divergent. These characteristics are common to all non-linear systems which have feed-back. To guarantee stability, constraints on synaptic weights are introduced so that the dynamics of a network are expressed by the Lyapunov function (see – e.g., Keller-Ressel, 2012). Concretely, a constraint of equivalent mutual connection weights of two units is implemented.

The Jordan ANN (Jordan, 1986a, 1986b) and the Elman ANN (Elman, 1990) are two examples of the Recurrent ANNs. Figure B.4 shows exemplars of these two ANNs.

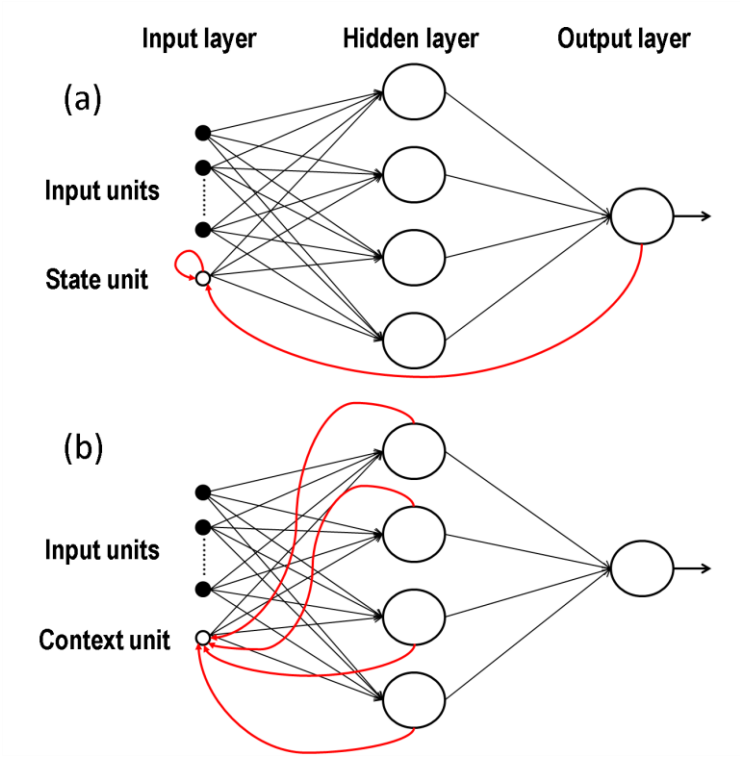


Figure B.4. A Jordan Artificial Neural Network (a), and an Elman Artificial Neural Network (b).

In a Jordan ANN, the activation values of the output units are fed back into the input layer through a set of extra input units called the state units. There are as many state units as there are output units in the network.

In an Elman ANN, a set of context units are introduced, which are extra input units whose activation values are fed back from the hidden units. This ANN is very similar to the Jordan ANN, except that: (i) the hidden units instead of the output units are fed back, and (ii) the extra input units have no self-connections.

The two Recurrent ANNs described above have been successfully used in predicting financial markets because of their ability to learn sequences. When time series data are used, as they learn the influence of the previous time steps themselves, they eliminate the need to present lagged data (i.e., windows of past data) as input to the ANN. All this, in turn, results in a reduction of the ANN complexity (e.g., network size, etc.).

Learning paradigms

The most well-known learning paradigms are the supervised learning and the unsupervised learning.

In supervised learning, the ANN is provided with a set of correct results. Supervised learning algorithms adjust the synaptic weights using the input-output data to match the input-output characteristics of the ANN to the desired (correct) characteristics. This process is known as training. Through training, an ANN can learn a mapping between input and output without an explicit model for the mapping being provided. The BP algorithm (Rumelhart et al., 1986), the Conjugate Gradient (Masters, 1995), and the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963; Moré, 1977) are three representatives of the supervised learning paradigm's algorithms.

It is important to stress that training an ANN involves, in most cases, numerical optimisation of a usually non-linear 'objective function'. Note that the term 'objective function' is used here rather than 'error function' (i.e., difference between input-output characteristics of the ANN and desired characteristics). This is because the former term is more general, in that it may include other quantities such as penalties for weight decay (see Section B.4). Methods of non-linear optimisation have been studied for hundreds of years, and there is a huge literature on the subject in fields such as numerical analysis, operations research, and statistical computing (e.g., Fletcher, 1987; Gill et al., 1982).

Bearing in mind the above, because of its relevance for the research work carried out in this thesis, the BP algorithm is explained in detail in Section B.3. With regard to the Conjugate Gradient and the Levenberg-Marquardt algorithms, on the other hand, only the following is noted here. As the memory required by the Levenberg-Marquardt algorithm is proportional to the square of the number of weights, this algorithm is efficient for simpler ANNs only (i.e., small number of weights). For larger number of weights, the Conjugate Gradient algorithm is more suitable as the memory required is proportional to the number of weights. However, the Conjugate Gradient algorithm is quite complicated to implement. Note that a detailed discussion of the Conjugate Gradient and the Levenberg-Marquardt algorithms in the context of the ANNs can be found in Masters (1995).

In unsupervised learning, the ANN is not provided with a set of correct results during training. Unsupervised learning algorithms use a mechanism that changes synaptic weight values according to the input values to the ANN (unlike supervised learning which changes the weights according to supervised data for the output of the ANN). Since the output characteristics are determined by the ANN itself, this mechanism is called self-organisation. Unsupervised ANNs usually perform some kind of data compression, such as dimensionality reduction or clustering. Hebbian (Hebb, 1949) and Competitive Learning (Fritzke, 1997) are two representatives of the unsupervised learning paradigm's algorithms.

B.3 Feed Forward Artificial Neural Networks & Back Propagation Algorithm

The data flow of a FF ANN is unidirectional from the input to the output units. Figure B.5 shows a numerical example of the data flow of a three layer (i.e., input, hidden, and output layers) FF ANN (or MLP).

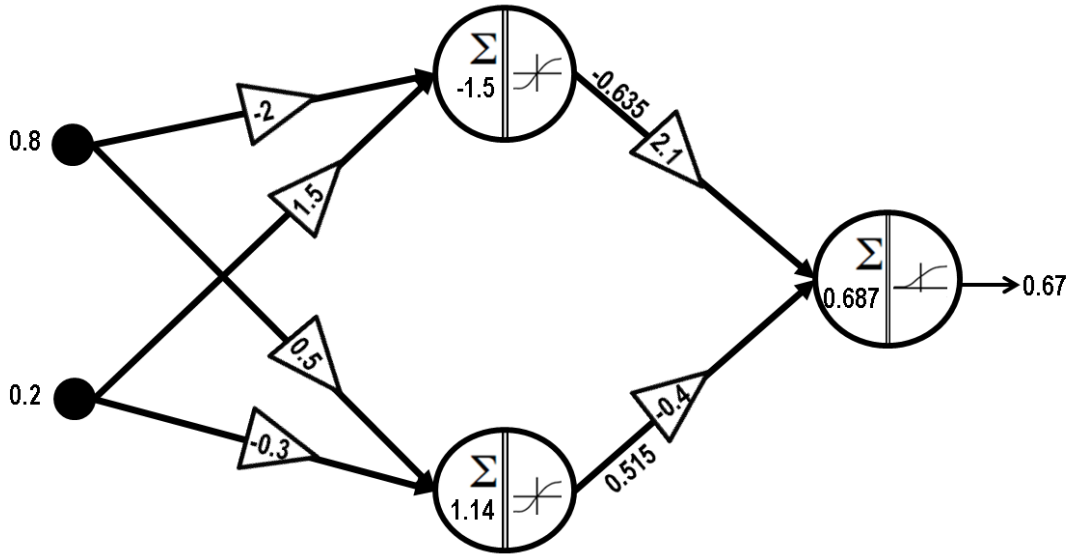


Figure B.5. Example of the data flow in a simple 1-hidden layer Feed Forward Artificial Neural Network.

As it can be seen from Figure B.5, the inputs are fed into the input layer and are multiplied by interconnection weights as they are passed from the input layer to the hidden layer. Within the hidden layer, they are summed, and then processed by a selected non-linear activation function. As the processed data leaves the hidden layer the data are multiplied by interconnection weights and processed one last time within the output layer to produce the ANN output.

With BP (see Figure B.6), which is one of the most popular learning algorithms for iteratively determining the weights of the FF ANNs, the input data are presented repeatedly to the ANN. With each presentation, the ANN output is compared to the desired output and an error is computed. This error is then fed back (back propagated) to the ANN and used to adjust the weights such that the error decreases with each iteration and the ANN becomes closer to producing the desired output.

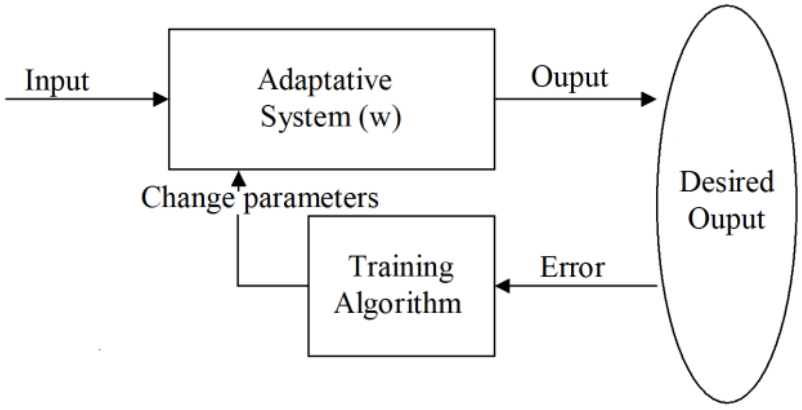


Figure B.6. Schematic representation of the Back Propagation data flow.

The BP algorithm is an extension of the Delta Rule (see – e.g., Rumelhart et al., 1986; McClelland et al., 1986), which is a simple learning algorithm that modifies the weights between the output and the hidden layers. Specifically, the BP algorithm enables modifying not only the weights between output and hidden layers but also between hidden and input layers.

Mathematically, let E be the error between the ANN output v^3 and the supervised data y, for the exemplar FF ANN in Figure B.5 (where the number at superposition - i.e., 3 - means the layer number). Since the ANN output is changed when the synaptic weights w are modified, the E must be a function of the synaptic weights. This implies that the searching direction of the smaller error point is obtained by calculating a partial differential. This technique is called the Gradient method, and the Steepest Descent method is the base of the BP algorithm (see – e.g., Rumelhart et al., 1986; McClelland et al., 1986). In this framework, the searching direction is $g = -\partial E(w)/\partial w$. The modification of weights is $\Delta w = \varepsilon g$, where ε is the learning rate.

From the above, it is possible to obtain the following equations that define the BP algorithm:

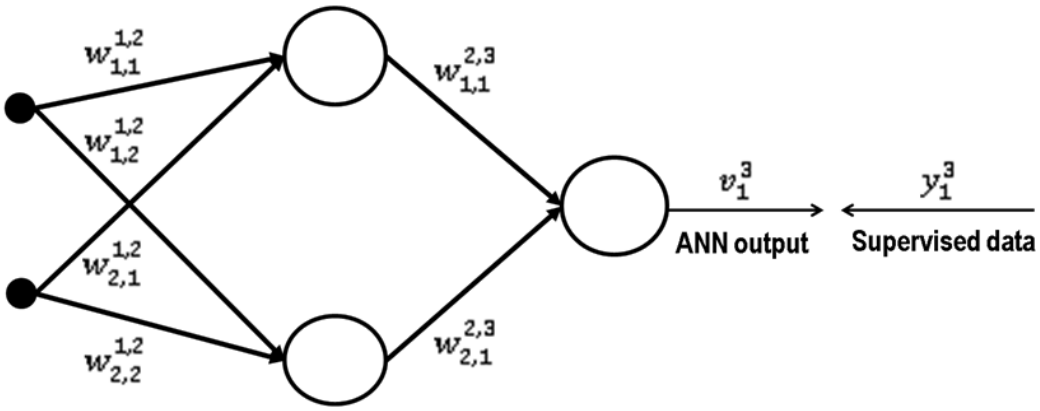
$$\Delta w_{i,j}^{k-1,k} = -\varepsilon d_j^k v_i^{k-1} \tag{B.1}$$

and

$$d_j^k = \begin{cases} \left(v_j^3 - y_j\right)\dfrac{\partial f\left(U_j^k\right)}{\partial U_j^k} & \text{For output layer} \\ \displaystyle\sum_{h=1}^{N_{k+1}} d_j^{k+1} w_{i,h}^{k,k+1} \frac{\partial f\left(U_j^k\right)}{\partial U_j^k} & \text{For hidden layer(s)} \end{cases} \tag{B.2}$$

where $w_{i,h}^{k,k+1}$ is the connection weight between the i^{th} unit in the $(k-1)^{\text{th}}$ layer and the j^{th} unit in the k^{th} layer, and U_j^k is the total amount of input to the j^{th} unit at the k^{th} layer. To calculate d_j^k, d_j^{k+1} must be previously calculated. This implies that the calculation must be conducted in the order of the direction from the output layer to the input layer. Thus, the name BP algorithm.

When a Logistic function (see Section B.2.1) is used for the activation function $f()$ of the neuron units, the calculation of the BP algorithm becomes simple. This is illustrated in Figure B.7.



$$\begin{cases} d_j^3 = (v_j^3 - y_j)(1 - v_j^3)\, v_j^3 \\ \Delta w_{i,j}^{2,3} = -\varepsilon d_j^3\, v_i^2 \end{cases} \quad \text{For output layer}$$

$$\begin{cases} d_j^2 = \sum_{h=1} d_h^3 w_{j,h}^{2,3} (1 - v_j^2)\, v_j^2 \\ \Delta w_{i,j}^{1,2} = -\varepsilon d_j^2\, v_i^1 \end{cases} \quad \text{For hidden layer(s)}$$

Figure B.7. Back Propagation algorithm for neurons with a Logistic activation function.

B.4 Overfitting & Generalisation

FF ANNs with non-linear activation functions and "enough" hidden neurons are very "flexible". That is to say, given a sufficient amount of training data and sufficient training time, they are capable of approximating any function to a desired degree of accuracy (see – e.g., Hornick et al., 1989; Leshno et al., 1993). However, it is not always desirable to have very "flexible" ANNs. In some cases, it is better to have ANNs that do not perform perfectly on the training data set. This is because, ultimately, the performance of an ANN is measured on "new" data that are not part of the training data set.

The above basic concept is illustrated through an example in Figure B.8. This example refers to the case where two ANNs (i.e., a simple perceptron and a MLP) are used for performing a classification task. In Figure B.8a the data set has been separated by a simple ANN and, as the data set is not linearly separable, a few errors remain. On the

other hand, the more "flexible" ANN in Figure B.8b is capable of a perfect separation. If the point indicated by the arrow is a "new" data point (i.e., not part of the training data set) the ANN in Figure B.8b would classify it as a cross, the one in Figure B.8a as a circle. The problem is that a priori it is not known which one of the two classifications is correct. For example, it may be that the cross which caused the tongue in the separation boundary is the results of a noisy measurement and that the likelihood of the circles is much higher in that region than the one for the crosses. Then, the prediction of the simple ANN would indeed be better than the one of the more "flexible" ANN. It follows that (in particular for noisy data) the "flexibility" of an ANN has to be controlled. Otherwise, the ANN may learn the noise in the data instead of the underlying structure of the input-output mapping. This problem is known as overfitting.

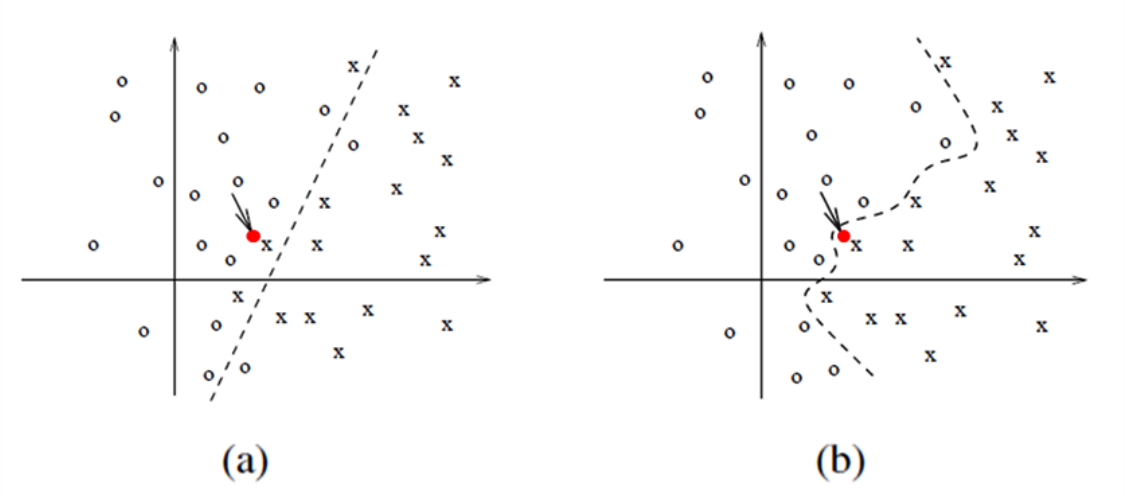


Figure B.8. The separation of a data set by a simple perceptron (a), and by a Multi-Layer Perceptron (b).

Overfitting occurs if the effective number of free parameters in an ANN is too large. In the ANNs, the number of free parameter increases with the number of neurons in the hidden layer. Taking this into consideration, it is also important to stress that overfitting may additionally occur if an ANN is trained for too long. The reason for this is that, at the beginning of the training process, an ANN does not yet use all the "flexibility" provided by its free parameters. Indeed, at the beginning of the training process, several neurons in the hidden layer learn approximately the same weight vector. Only after some time, they specialise to detect different features.

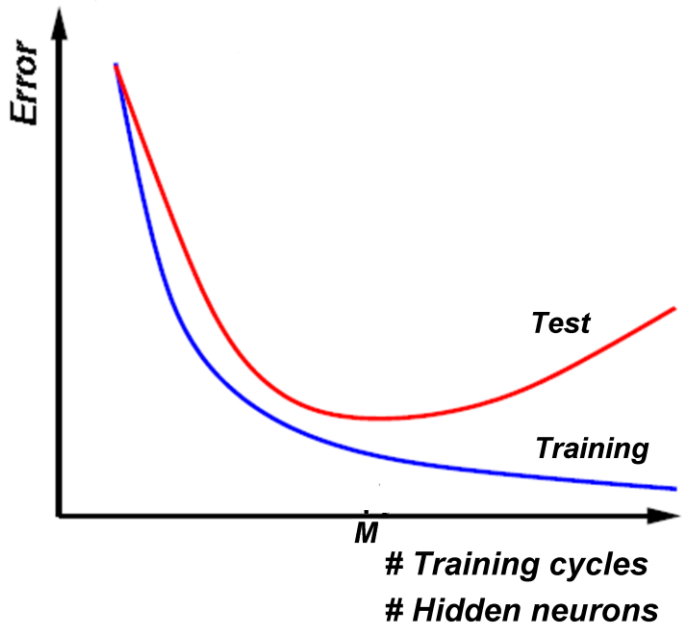


Figure B.9. Error on the training and test data sets as a function of the number of training cycles and hidden neurons.

Figure B.9 shows a qualitative example of how the error on the training and test (i.e., unseen) data sets changes with the number of training cycles (i.e., training time) and with the number of hidden neurons. It is possible to observe from this figure that the error on the training set becomes smaller as the number of training cycles or the number of hidden neurons increases (i.e., overfit the data). However, the error on the test set will encounter a minimum (i.e., optimum) and then increase.

Bearing in mind the above, several approaches for improving generalisation have been proposed. The most commonly used are the Early Stopping (e.g., Weigend, 1994) and the WDR (e.g., Bishop, 1995) techniques.

B.4.1 Early Stopping

Early Stopping is a simple method of flexibility control. It only involves controlling the number of training cycles and does not formally reduce the number of free parameters. It works by stopping the training process at different times (i.e., different number of training cycles) and testing the ANN performance on the independent test set. Every

time the training is stopped the weights of the ANN are saved. The weights that lead to the lowest error on the test set (i.e., point M in Figure B.9) are then used in the final ANN.

B.4.2 Weight Decay Regularisation

WDR is probably the most systematic way to control the flexibility of an ANN. It works by ensuring that the function computed by an ANN is no more curved than necessary (i.e., it encourages smoother ANN mapping). WDR involves applying a penalty term to the error function, giving:

$$E_{WDR} = E + \alpha\, \Omega \tag{B.3}$$

where α is the coefficient of WDR which controls the trade-off between reducing the error E and increasing the smoothing. As an overfitted mapping with regions of large curvature requires large weight, WDR penalise large weights by choosing the Regularisation function as:

$$\Omega = \frac{1}{2}\sum_{i} w_i^2 \tag{B.4}$$

APPENDIX C EVOLUTIONARY ALGORITHMS

C.1 Introduction

The term EA is used for a broad spectrum of heuristic approaches that apply the biological search process of natural evolution to artificial systems. In general, EAs can be seen as heuristic search algorithms that are useful for finding near-optimal solutions to difficult problems within an acceptable time. As these algorithms provide a framework for effectively sampling large search spaces, in fact, they are particularly suitable for problems for which deterministic search techniques incur difficulties (or fail completely), and for which exhaustive search is impractical. Primary examples of EAs include ESs (Rechenberg, 1973; Schwefel, 1981, 1995), GAs (Holland, 1962, 1975), Evolutionary Programming (Fogel et al., 1966), and Genetic Programming (Koza, 1992). Each of these algorithms has been proved capable of yielding approximately optimal solutions given complex, multimodal, non-differential, discontinuous, noisy and time-dependent search spaces (Jones, 1998). Furthermore, these algorithms offer conceptual and computational simplicity, good applicability to a broad class of problems, capability of self-optimisation, and have the potential of being used in conjunction (i.e., hybridisation) with other methods (Fogel, 1997).

EAs have been successfully applied to numerous problems from different domains. These include optimisation, automatic programming, machine learning, and population genetics (Mitchell, 1996). Book length reviews of the many EA applications in these and other fields can be found in Goldberg (1989), Davis (1991), and Sipper (2002), to mention only a few. With specific regard to the water resources planning and management field, Nicklow et al. (2010) provide a comprehensive review of recent EA applications. These include WDSs and closely related applications (e.g., Savić and Walters, 1997; Wu and Walski, 2005), urban drainage and sewer system applications (e.g., Walters and Smith, 1995; Liang et al., 2004), and water supply and wastewater treatment applications (e.g., Murthy and Vengal, 2006; Burn and Yulianti, 2001), among the others.

Because of its relevance for the work carried out in this thesis, it is essential to stress here that an important field of research in evolutionary computation concerns Evolutionary Artificial Neural Networks (EANNs). EANNs are biologically-inspired

computational models that use EAs in conjunction with ANNs. In this framework, EAs are often used for "designing" ANNs. Common approaches involve performing tasks such as connection weight optimisation (e.g., Montana and Davis, 1989; Keesing and Stork, 1991), architecture optimisation (e.g., Harp et al., 1991; Miller et al., 1989), parameter optimisation (e.g., Merelo Guervós et al., 1993; Castillo et al., 2000), and input data selection (e.g., Brill et al., 1992; Reeves and Taylor, 1998). Note that a comprehensive (but not so recent) review of the different interactions/combinations between EAs and ANNs that have been proposed in the literature is given in Yao (1999). With particular regard to the water resources management field, EANN applications are scarce. Noteworthy is however the work presented in Giustolisi and Simeone (2006) whereby a multi-objective EA was used for the selection of the input data and the number of hidden neurons of ANN models for groundwater level prediction. The key advantages of the EANNs identified in many of the aforementioned studies include: (i) their adaptability to dynamic environments (i.e., EANNs can adapt to an environment as well as to changes in the environment), and (ii) the fact that they dramatically reduce the effort required from a human expert to "design" an ANN model for a given problem whilst enabling replicating or outperforming the quality of the results achievable through human expert intervention.

Bearing in mind the above, this appendix is organised as follows. After this introduction, Section C.2 focuses on the particular family of EAs employed in the research work presented in this thesis for the optimal selection of the ANN input structure and parameters, namely the ESs family. In this section, the general functioning of an EA is described first in order to explain why different families of EAs exist. Subsequently, the main characteristics and different components of an ES algorithm are analysed. Note that this appendix only aims at presenting a brief introduction to the EAs. Further details can be found in the reference provided above and in several good books (e.g., Back, 1996; De Jong, 2007).

C.2 Evolutionary Strategy

C.2.1 Overview

EAs are computer programs that attempt to solve complex problems by mimicking the processes of Darwinian evolution (i.e., Darwin, 1973). All EAs have basic properties which distinguish them from other search algorithms. Firstly, EAs are population-based.

They make use of the collective learning process within a population of individuals, each of which represents a search point in the space of the potential solutions to a given problem (i.e., search space). These individuals compete continually with each other to evolve towards areas of the search space that give globally optimal solutions to the problem (much in the same way as in nature populations of organisms tend to adapt to their surrounding environment). Note that this process is equivalent to carrying out parallel explorations of the overall search space in a problem. Secondly, a population is evolved by using randomised processes of mutation, recombination (which may be completely omitted in some algorithmic realisations), and selection. Mutation corresponds to an "erroneous" self-replication of individuals and introduces innovation into the population. Recombination allows the mixing of information between some individuals (commonly known as parents) and passing it to other individuals (commonly known as offsprings or descendants). Finally, as the environment delivers quality information about each individual (i.e., fitness value), selection performs a comparison of individual fitness and favours the evolution of those individuals with higher fitness. Taking all this into account, it is important to stress that the aforementioned properties are general. Indeed, at the algorithmic level, different representations of individuals, schemes for implementing fitness evaluation, and different mutation/recombination/selection operators are adopted in different EAs.

In view of the above, this section focuses on the ES algorithm, which represents a particular instance of the several EAs that have been developed so far. According to De Jong (2009), ES algorithms were developed to solve difficult parameter optimisation problems arising from engineering design problems. In that scenario, the general framework in which they were applied was one in which a parameterised set of designs was developed along with a measure of effectiveness (i.e., a fitness function). These algorithms were then used to find a set of design parameter values that optimised the given fitness function. Bearing this in mind, the main characteristics of a state-of-the-art ES algorithm are presented here as follows. In section C.2.2 the standard ES notation is introduced. In Section C.2.3 the basic ES algorithm, as it is used in this thesis, is outlined. Finally, in Sections C.2.4 and C.2.5, the mutation and selection operators are respectively briefly described. Note that further details about this topic can be found in – e.g., Beyer and Schwefel (2002).

C.2.2 Evolutionary Strategy standard notation

ESs aim at optimising a given fitness function f with respect to a set of decision variables $Y = (y_1, y_2, \ldots)$, where $Y \in \upsilon$, and υ is a set of data structures of finite but not necessarily fixed length (e.g., the real-valued N-dimensional search space $\mathbb{R}^N$, the integer search space $\mathbb{Z}^N$, etc.).

ESs operate on populations P of individuals a. An individual a_k comprises the specific set of decision variables Y_k and its fitness function value $f_k = f(Y_k)$. Within one ES generation step, λ offspring individuals a_l are generated from the set of μ parent individuals a_m. Note that, as it can be easily inferred, the size λ of the offspring population P_o is usually not equal to the size μ of the parent population P_p. Further note that the ES-specific parameters λ and μ as well as the mixing number ρ (see below) are kept constant during all the generation steps.

The way the offspring population is generated can be expressed by using the $(\mu/\rho + \lambda)$ or $(\mu/\rho, \lambda)$ notation, to be read as "mu slash rho plus lambda" or "mu slash rho comma lambda", respectively (Rechenberg, 1978). Where ρ refers to the number of parent individuals involved in the procreation of one offspring. The "+" and "," refer to the kind of selection operator used (see Section C.2.5).

C.2.3 Basic Evolutionary Strategy algorithm

Figure C.1 shows the pseudo-code of the basic ES algorithm as it is employed in this thesis. It is possible to observe from this figure that at generation $g = 0$ the parent population $P_p^{(0)} = (a_1, \ldots, a_\mu)$ is (randomly) initialised. Then, the repeat-until loop is entered. From the parent population $P_p^{(g)}$ at generation g a new offspring population $P_o^{(g)}$ is produced by running the for-end loop λ times. Here, by selecting $\rho = 1$ (i.e., the recombination only involves producing a copy of a parent - asexual reproduction), the μ parents produce λ offspring by cloning themselves and mutating the values in the decision variable sets using a Gaussian mutation mechanism (see Section C.2.4). After a complete offspring population is obtained, selection is performed. This results in a new parent population $P_p^{(g+1)}$. The outlined process repeats itself until a certain number of generations (i.e., termination condition) is reached, at which time the population

member with the best fitness is returned as the best estimate of the optimal decision variable set.

Procedure

Begin

$g = 0$

$Initialise\left(P_p^{(0)} = \left\{\left[Y_m^{(0)}, f\left(Y_m^{(0)}\right)\right], \quad m = 1, \ldots., \mu\right\}\right)$

Repeat

For $\boldsymbol{l = 1}$ **to** $\boldsymbol{\lambda}$

$\xi_l = marriage\left(P_p^{(0)}, \ \rho\right);$

$Y_l = recombination\ (\xi_l)$

$\tilde{Y}_l = mutation\ (Y_l)$

$\tilde{f}_l = f\left(\tilde{Y}_l\right)$

End

$P_o^{(g)} = \left\{(\tilde{Y}_l, \tilde{f}_l), \quad l = 1, \ldots., \lambda\right\}$

Case selection type

(μ, λ): $P_p^{(g+1)} = selection\left(P_o^{(g)}, \mu\right)$

$(\mu + \lambda)$: $P_p^{(g+1)} = selection\left(P_o^{(g)}, P_p^{(g)}, \mu\right)$

End

$g = g + 1$

Until termination condition

End

Figure C.1. Pseudo-code of a basic Evolutionary Strategy algorithm.

C.2.4 Mutation operator

The mutation operator is the primary means for introducing variation in the ES algorithm. Its purpose is to simulate the effect of an "erroneous" self-replication of individuals that can occur with a certain probability P_{mut}. The design of a mutation operator is problem-dependent, thus various mutation operators have been proposed in the literature (see – e.g., Beyer, 2001). In this thesis, a well-known and commonly used standard operator is employed, namely the Isotropic Gaussian mutation operator. This is described below.

By Isotropic Gaussian mutation a set of decision variables $\tilde{Y}$ is perturbed according to the following formula:

$$\tilde{Y} = Y + Z \quad where \quad Z = \sigma[N_1(0,1), \dots, N_N(0,1)] \tag{C.1}$$

That is to say, each component $\tilde{y}_i$ of the decision variable set $\tilde{Y}$ is perturbed independently with a random number from a Gaussian distribution with zero mean and standard deviation σ, which is called the Mutation Strength. Consequently, it is clear that the Isotropic Gaussian mutation operator defined in Equation (C.1) only needs the Mutation Strength parameter for its control. This characteristic represents the main strength of this operator.

Noteworthy is the fact that the Mutation Strength parameter can be either kept constant during all the generation steps (as in the particular ES algorithmic realisation used in the work presented in this thesis), or adapted according to the local search space characteristics in order to improve convergence. In the case of a high number of offspring having a better fitness of their parents (as it is often observed at the beginning of a search), for example, a large Mutation Strength parameter is advantageous in order to promote exploration of the search space. At later phases of a search, on the other hand, a small Mutation Strength parameter is often more appropriate for promoting exploitation of the parent information. An example of an adaptive control of the Mutation Strength parameter is the 1/5-th rule proposed by Rechenberg (1973). Having said this, note that the topic of self-adaptation is beyond the scope of this appendix. A detailed discussion about this topic can be found in – e.g., Fogel et al. (1996) and Schwefel (1981).

C.2.5 Selection operator

In an ES algorithm, the selection operator is used to guide the search into promising regions of the decision variable space. It guarantees that only the μ best individuals (i.e., those with the highest fitness function value) from the selection pool of generation (g) are transferred into $P_p^{(g+1)}$.

As mentioned in Section C.2.2, two selection operators may be used (i.e., the "+" and "," operators). These define whether or not the parent population at generation (g) is included in the process outlined above, respectively. In the case of $(\mu + \lambda)$ selection both the parents and the offspring individuals are copied into the selection pool, which is therefore of size $\mu + \lambda$. In the case of (μ, λ) selection, only the λ newly generated offspring individuals (i.e., the $P_o^{(g)}$population) define the selection pool. Note that each of these variants has its specific strengths and weaknesses. However, the former guarantees the survival of the best individual (i.e., elitist technique) thereby guaranteeing a monotonic improvement of the population. Furthermore, it has been recommended for combinatorial optimisation problems (see – e.g., Herdy, 1990; Beyer, 1992).

APPENDIX D STATISTICAL PROCESS CONTROL

D.1 Introduction

SPC involves using statistical techniques to measure and analyse the variations in a process. The concepts of SPC were initially developed by Walter A. Shewhart in the early 1920s (Shewhart, 1931) and were expanded upon by Edwards W. Deming (Deming, 1950, 1975), who introduced SPC to the Japanese industry after the Second World War. Nowadays, SPC is used by manufacturing organisations around the world and has been successfully applied to non-manufacturing processes as well (see – e.g., Humphrey, 1989; Roberts, 2005). In the water resources management field, Smeti et al. (2007) used SPC for quality improvement of treated water and Sumer et al. (2007) used SPC for the detection of sanitary sewer overflows. SPC concepts have also been used for burst detection in pipelines (e.g., Al-Rafai and Barnes, 1999).

SPC's general idea is that while every process displays variations, some processes display controlled variations that are 'natural' to the process (i.e., common causes of variations), while others display uncontrolled variations that are not present in the process causal system at all times (i.e., special causes of variations). A process that features only common causes of variations (in a state of statistical control - Shewhart, 1931) is called a stable process. This implies that its variations are predictable within statistically established limits. A process that features both common and special causes of variations is called an unstable process. Note that this does not implies that its variations are large but only that the magnitude of the variation from one time period to the next is unpredictable.

Complex pipeline systems do not respond in a linear fashion to changes in their inputs and non-linear changes make even simple systems impossible to predict. Therefore, pressure/flow SCADA measurements can be seen as the output of an unstable process. In a pipeline system, common causes of variations include measurement uncertainties, seasonal pressure/flow variations, growing demand over time, etc. Special causes of variations include sensor failures/drifts and the occurrence of a burst, among the others.

It has to be stressed that SPC refers to a number of different methods to detect any special cause of variations in a process. In this thesis, SPC-based Control Charts and

Control Rules are used for determining whether the pressure/flow variations that are observed when operating a complex pipeline system are indicative of the pipe burst events/other (similar) events occurrence or consistent with the variations due to common causes. A brief description of these two methods is given in the following sections. Detailed descriptions of these two and other SPC-based methods can be found in – e.g., Hart and Hart (1989), Pyzdek (1989), Buffa (1972), and Grant and Leavenworth (1980).

D.2 Control Charts

A Control Chart is a graphical representation of certain descriptive statistics for specific quantitative measurements of the process. It is used to study how a process changes over time and is very useful to track unusual variations and trends. A Control Chart always has a centre line for the average, an upper line for the upper control limit and a lower line for the lower control limit. By comparing current measurements to these limits, conclusions can be drawn about whether the process is in control or affected by special causes of variations (i.e., 'out of control' – Breyfogle, 1999).

Even though it is possible to arbitrarily establish the aforementioned lines, it is common practice to apply statistical principles to determine them from the historical measurements. Usually, in a Shewhart Control Chart:

- Points representing a statistic of a sample of measurements (e.g., mean, range, etc.) are plotted in time order;
- The centre line is drawn at the value of the mean of the statistic (e.g., mean of the means, mean of the ranges, etc.);
- Upper and lower control limits (that indicate the threshold at which the process output is considered statistically "unlikely") are drawn at 3 (i.e., a user defined multiplier) standard deviations (e.g., calculated as the mean of standard deviations of the means, mean of the standard deviations of the ranges, etc.) from the centre line.

Note that, as it is shown in Figure D.1, a Control Chart may also have other optional features such as upper and lower warning limits. These are drawn as separate lines at 2 (or 1) standard deviations above and below the centre line.

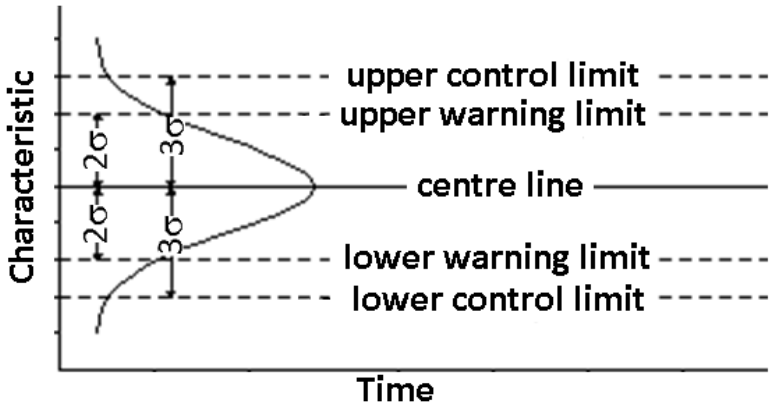


Figure D.1. Shewhart Control Chart with control and warning limits.

If a statistic is normally distributed, according to the Central Limit theorem (see – e.g., Fischer, 2010), the user defined multipliers typically used in a Shewhart Control Chart have the following theoretical interpretation (see – e.g., Swift, 1995):

- 68.27% of the points should be within the $\pm 1\sigma$ limits from the centre line;
- 27.18% of the points should be between the $\pm 1\sigma$ and $\pm 2\sigma$ limits from the centre line;
- 4.28% of the points should be between the $\pm 2\sigma$ and $\pm 3\sigma$ limits from the centre line;
- No more than 0.27% of the points should exceed the $\pm 3\sigma$ limits from the centre line.

Bearing this in mind, it is important to stress the following. Control Charts can be used also if the data are not normally distributed. However, the theoretical interpretation given above does not hold true.

D.3 Control Rules

When using Control Charts for deciding about the occurrence of an 'out of control' situation, collecting the measurements, calculating the control limits, and plotting the data points is only the first step of the process. In actual fact, once a Control Chart is constructed and the data points are plotted, the Control Chart needs to be evaluated and a decision about whether or not the process is operating in statistical control has to be made. There are several guidelines for evaluating Control Charts, all with a prevailing

philosophy. However, the general idea is that the plotted data points should only exhibit random variations that cannot be fitted with an identifiable or quantifiable pattern. It is therefore clear that, in order to effectively evaluate a Control Chart, it is necessary to focus the attention not only on the latest data point but also on the recent history of the process. The importance of focussing on the recent history of the process is well highlighted by the following example. When a large number of consecutive data points lies above (or below) the centre line on a Control Chart for means, it may indicate that the mean is off target. If consecutive data points are taken into consideration during evaluation, this situation could be identified even though no single data point has exceeded the control limits.

In the scenario outlined above, several sets of Control Rules have been proposed for helping making more reliable decisions about the occurrence of 'out of control' situations (e.g., Page, 1955; Western Electric Company, 1958; Bissell, 1978; Wheeler, 1983). All these sets of Control Rules have a common aim. This is highlighting when something "unlikely" happens in a process. The most widely used Control Rules are the Western Electric Company Control Rules.

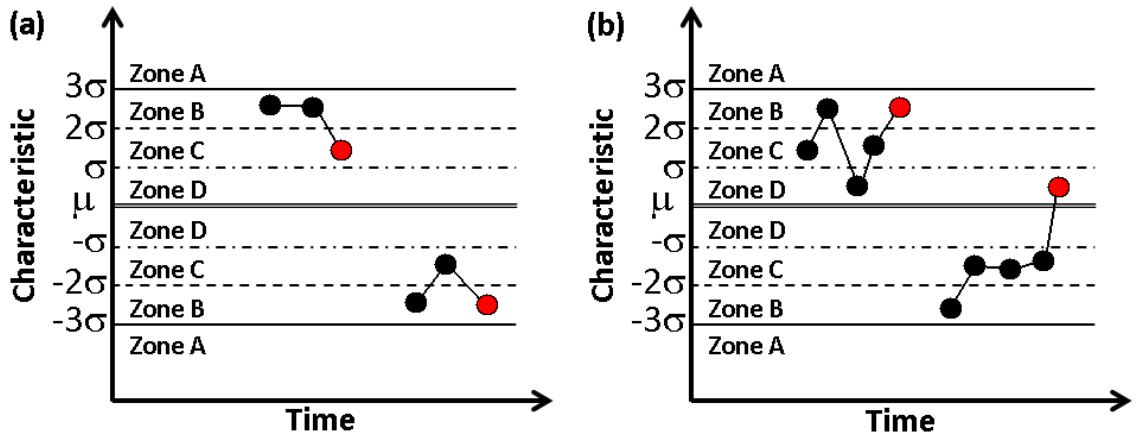


Figure D.2. Examples of Shewhart Control Charts evaluated by using two of the Western Electric Company Control Rules. Two out of three points in zone B or beyond on the same side of the centre line (a), and four out of five points in zone C or beyond on the same side of the centre line (b).

Figure D.2 shows, as an example, an illustration of how Control Charts are evaluated by using two of the aforementioned Western Electric Company Control Rules. Note that the Control Charts shown in Figure D.2 are divided into zones. If these zones are measured in units of standard deviation (as it is shown in the figure), the theoretical interpretation of the considered Control Rules (for normally distributed data) is as follows: (i) the probability that 2 out of 3 points in a row fall outside the $\pm 2\sigma$ limits on the same side of the centre line is about 1%, and (ii) the probability that 4 out of 5 points fall outside the $\pm 1\sigma$ limits on the same side of the centre line is about 3%. Note that for situations where the data are not normally distributed the same consideration made in Section D.2 applies. Control Rules can be used in isolation or together. They provide different criteria for deciding if the observed variations in a process are "unlikely" due to common causes. When the Control Rules flag the observed variations as "unlikely" due to common causes, the decision that these variations are due to special causes can then be made.

APPENDIX E BAYESIAN NETWORKS

E.1 Introduction

BNs are used to represent knowledge about an uncertain domain and enable reasoning using probabilities. They have the ability of capturing both qualitative and quantitative knowledge and unifying them into a single form of representation. BNs combine principles from graph theory and probability theory and make use of statistical and computational methods. It is this combination which makes them intuitively understandable whilst mathematically rigorous (Ben-Gal, 2008). As BNs enable simplifying the manipulations in Bayesian inference, they are seen as a powerful data analysis tool within both statistical and AI societies.

BNs have been used widely from problems in medical diagnosis (e.g., Heckerman, 1990; Spiegelhalter et al., 1989) to manufacturing control (e.g., Nadi et al., 1991), from language understanding (e.g., Goldman, 1990) to image recognition (e.g., Booker and Hota, 1986), from forecasting (e.g., Abramson, 1994) to heuristic search (e.g., Hansson and Mayer, 1989), and so on. Note that a book length review of recent BNs applications in these and other fields can be found in Pourret et al. (2008). With specific regard to the water resources management field, BNs have been used in applications such as groundwater remediation (Stiber et al., 1999, 2004), groundwater quality modelling (Shihab, 2005), and prediction of estuarine water quality (Stow et al., 2003), among the others. However, few applications of BNs can be found in the drinking water research literature. These include the works presented in Dawsey et al. (2007), Perelman and Ostfeld (2010), and Murray et al. (2011), which mainly focussed on water quality monitoring in WDSs.

In this appendix, several BN topics relevant to the research work carried out in this thesis are discussed. After this introduction, the fundamentals of BNs are given in Section E.2 as follows. Firstly, the relevant BN terminology is introduced and the BNs are formally defined. Secondly, the probabilistic inference process via BNs and the methods for performing belief updating are introduced. Section E.3 then deals with the task of building BNs. In particular, the problems of learning the structure and parameters of a BN given full and partial observability are considered first. Once this is done, the EM algorithm is introduced and its application to: (i) learning the parameters

of a BN with discrete variables, and (ii) learning the parameters of a mixture of multivariate Gaussian densities is described. Note that although a number of algorithms and techniques are mentioned here, it is beyond the scope of this appendix to discuss each of them in great detail. Allowing for this, references to several good introductory and more advanced publications dealing with the general BN theory and the BN topics discussed here are provided in the following sections.

E.2 Fundamentals of Bayesian Networks

E.2.1 Definitions

BNs correspond to a class of probabilistic graphical models known as DAGs. They enable the effective representation and computation of the Joint Probability Distribution (JPD) over a set of random variables (Pearl, 1988).

BNs represent domain knowledge qualitatively by the use of a DAG with nodes and edges. A node represents a random variable (or a set of random variables). An edge from node X_i to node X_j represents a statistical dependence between the corresponding random variables. That is to say, an edge indicates that the value taken by variable X_j depends on the value taken by variable X_i. Bearing this in mind, node X_i is often referred to as a parent of X_j and, similarly, X_j is referred to as the child of X_i. Furthermore, by continuing to use this "genealogical jargon", the terms descendants (i.e., the set of nodes that can be reached on a direct path from the node), and ancestors (i.e., the set of nodes from which the node can be reached on a direct path) are used too (Griffiths and Yuille, 2006). Taking this into account, it important to highlight that, in a DAG, there is no node that can be its own ancestor or its own descendant.

In addition to the DAG, BNs represent domain knowledge quantitatively by encoding the statistical dependencies between the random variables. For BNs which make use of discrete variables (i.e., taking a fixed number of states) the relationships between variables are often encoded by using CPTs. A CPT is associated with each node and contains the probabilities that the variable will be in each of its possible (i.e., allowed) states given the states of its parents. For BNs which make use of continuous variables (i.e., continuous state set) the relationships between variables are encoded by using Conditional Probability Distributions (CPDs) (e.g., Gaussian distributions).

The above can be formally expressed as follows (Friedman et al., 1997). A BN B, is defined by a pair $\langle G, \Theta \rangle$, where G is the DAG with nodes $X_1, X_2, \ldots, X_n$ that represent the set of random variables V. The graph G encodes the independence assumptions, by which each variable X_i is independent of its non-descendants given its parents in G. The second component Θ denotes the set of parameters of the BN. This set contains the parameters $\Theta_{ijk} = P_B(X_i = j|\pi_i = k)$, where j is a value of variable X_i and k is a value of π_i, the set of parents of X_i in G. With this notation, the unique JPD over the set of random variables V is expressed as follows:

$$P_B(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P_B(X_i|\pi_i) \quad \text{(E.1)}$$

If X_i has no parents, its local probability distribution is said to be unconditional, otherwise it is conditional. If the variable represented by a node is observed, then the node is said to be an evidence node, otherwise the node is said to be hidden or latent (Ben-Gal, 2008).

E.2.2 Inference via Bayesian Networks

Given a BN that specifies the JPD as in Equation (E.1) (i.e., complete model for the variables and their relationships) and evidence of a set of variables e, it is then possible to find out updated knowledge about the values of another set of variables $U \subset V$. This process of updating (i.e., modifying) beliefs in light of the evidence is called probabilistic inference. It involves the task of calculating $P(U|e)$ - i.e., the probability of the subset U given the evidence. Unfortunately, although conceptually easy, in many circumstances this problem is nondeterministic polynomial time hard (see – e.g., Cooper, 1990; Russell and Norvig, 2010). That is to say, the computation needed to calculate the probabilities scales badly with the number of nodes.

Exact methods to solve the above problem in an acceptable time exist. The most popular exact methods are the Message Passing algorithm (Pearl, 1988; Lauritzen and Spiegelhalter, 1988), and the Variable Elimination algorithm (Zhang and Poole, 1996). In some classes of BNs (e.g., tree structured BNs), these methods attempt to exploit the structure of the DAG to derive efficient algorithms. For many BNs, however, the structure of the DAG prohibits exact solutions. As a consequence, a number of

approximation methods have been proposed in the literature. These include the Monte Carlo sampling (e.g., Pearl, 1987), a variety of standard Markov Chain Monte Carlo methods (e.g., Griffiths and Yuille, 2006; Gilks et al., 1995), the Loopy Belief Propagation (e.g., Kim and Pearl, 1983) and a number of Variational methods (e.g., Jordan et al., 1998).

Beside the references provided above, note that detailed descriptions of the aforementioned methods for probabilistic inference via BNs can also be found in several books (e.g., Cowell et al., 1999; Korb and Nicholson, 2004; Castillo et al., 1997). Finally, it is worth mentioning that a number of software tools, such as GeNIe and SMILE (GeNIe and SMILE, 2012), HUGIN (Jensen et al., 2002), and BNT (Murphy, 2001) provide libraries of belief updating algorithms.

E.3 Bayesian Networks Learning

E.3.1 Overview

In many practical applications the structure (variables with states and causal relationship between variables) and the parameters (conditional probabilities) of a BN are unknown. Although domain experts may provide both structure and parameters of a BN, experience shows that humans are good at providing the structure of a BN but not at estimating its parameters (Jensen, 2009).

Bearing in mind the above, both structure and parameters of a BN may be learned directly from a set of training data. This is known as the BN learning problem, which has been one of the most active areas of research within the statistical and AI communities in recent years. In general, it is possible to look at this problem along two dimensions: (1) the structure of the BN can be known or unknown, and (2) the observability of the BN can be full (i.e., when there are no hidden nodes and the set of training data is complete) or partial (i.e., when some of the nodes are hidden or when data are missing). In the remainder of this overview section a brief description of the methods used for learning the structure of a BN, assuming full and partial observability, is given first. Then a brief description of the methods used for learning the BN parameters in the case of known structure and full observability, and in the case of known structure and partial observability is presented. Note that a particular method used in the latter case, namely the EM algorithm, is discussed in more detail in Section

E.3.2. Further details about other aforementioned methods can be found in – e.g., Jordan (1998), Neapolitan (1990), and Heckerman (1996).

Learning the Bayesian Network structure

As the problem of learning a BN structure (as well as the associated parameters) given full observability is nondeterministic polynomial time hard (see – e.g., Chickering and Heckerman, 2004), the main algorithms used for solving it are iterative. In general, they start with a simple DAG, and incrementally modify this DAG, until reaching some termination condition. Two main classes of algorithms can be identified. These are: (1) *independence-based* algorithms (e.g., Glymour et al., 1987; Pearl, 2000), and (2) *score-based* algorithms (e.g., Cooper and Herskovits, 1992; Kirkpatrick et al., 1983). *Independence-based* algorithms work by performing a set of independence tests between every pair of unconnected nodes in the current DAG to assess if the data set supports the claim that they are independent given the rest of the graph structure. *Score-based* algorithms typically employ a local search to evaluate the current DAG, as well as every DAG obtained by implementing some slight modification (e.g., adding/deleting/reversing an edge), and "climb" to the new DAG with the highest score. In this framework, several scoring functions have been proposed, including the Minimum Description Length score (e.g., Lam and Bacchus, 1994; Suzuki, 1978), and the Akaike Information Criterion score (e.g., Bozdogan, 1987). Note that, often, the aforementioned scoring functions explicitly include a penalty term for complex structures which enables avoiding overfitting.

The problem of learning a BN structure (as well as the associated parameters) in the case of partial observability, on the other hand, is much more complicated. This is because there are two nested optimisation problems to be solved: (i) choosing the DAG, and (ii) choosing the parameters for a given DAG. Techniques such as the Structural EM (e.g., Friedman, 1998) have been proposed to solve the aforementioned problem. In particular, Structural EM works by making use of specific data structures that allow the computational results to be used across the different iterations of the algorithm.

Learning the Bayesian Network parameters

The problem of learning the BN parameters given full observability is the easiest problem to solve. It can indeed be seen as another example of Bayesian inference,

whereby the value of the BN parameters are inferred by using a whole set of training data, which are all understood to be potentially generated from a BN with one set of parameters. It involves finding the parameters set that maximises the likelihood that the set of training data came from the given BN. Specifically, assuming that the set of training data D, consists of m examples – i.e., $D = d_{1,} d_{2,} \dots , d_m$, it is possible to consider the likelihood of the model $L(\Theta)$, as the likelihood of seeing the data, given the model:

$$L(\Theta) = P(D|\Theta) = \prod_{l=1}^{m} P\,(d_l|\Theta) \qquad \text{(E.2)}$$

The *ML* learning paradigm can then be used to maximise $L(\Theta)$. It approximates the probability of a new example d, given the training data D, as $P(d|D) \approx P(d|\Theta_{ML})$, where Θ_{ML} is the maximum loglikelihood model which aims to maximise $ln\ P(D|\Theta)$ - i.e., $\Theta_{ML} = arg\,max_{\Theta}\ ln\,P(D|\Theta)$. It is important to stress that this learning paradigm does not assume any prior.

The *MAP* learning paradigm, on the other hand, enables a "full" Bayesian approach. Indeed, it approximates the probability of a new example d, given the training data D, as $P(d|D) \approx P(d|\Theta_{MAP})$, where Θ_{MAP} is the MAP probability (likelihood of the ''model given the data'') which aims to maximise $ln\ P(D|\Theta)$, i.e., $\Theta_{MAP} = arg\,max_{\Theta}\,ln\ P(\Theta|D)$. Thus, since through Bayes' theorem $P(\Theta|D) = P(D|\Theta)P(\Theta)/P(D)$, the prior is taken into account. This prior could be a broad and unrestrictive distribution, expressing little knowledge in the values that the conditional probabilities take.

The problem of learning the BN parameters given partial observability is more difficult than the previous outlined case. This is because the presence of missing training data implies that there is no tractably computable closed form solution. However, effective learning algorithms to solve this problem exist. Primary examples are the EM (Dempster et al., 1977) and the Gibbs sampling (Geman and Geman, 1984) algorithms. Such algorithms are often iterative, and work by making inferences about the data to complete the missing values, which, in turn, are used to update the model parameters. Noteworthy is the fact that, despite the merits of the EM and Gibbs sampling

algorithms, other algorithms have been proposed to overcome some of their limitations (e.g., sensitivity to the initial starting points). For example, the Information-Bottleneck EM algorithm (Elidan and Friedman, 2003) and the Data Perturbation method (Elidan et al., 2002) have been proposed to escape local maxima. The Generalised Conjugate Gradient algorithm (Thiesson, 1995) and the Online Updating Rules (Bauer et al., 1997) have been proposed to speed up the learning procedure.

E.3.2 Expectation Maximisation algorithm

The EM algorithm represents a general method for estimating likelihood functions and is a standard tool for statisticians. It is the most commonly employed algorithm for learning from incomplete data (Jensen, 2009). This is mainly because the EM is useful in situations where other optimisation methods fail and has the advantage of being simple, robust and easy to implement (Do and Batzoglou, 2008). This algorithm was developed in the statistics community by Dempster et al. (1977) and adapted for the use with the BNs by Lauritzen (1995). The EM algorithm is an algorithm that, given a database of 'cases' (i.e., training data set), determines estimates of the BN parameters that are optimal within a neighbouring set of solutions. It starts with initial values (e.g., chosen at random) for all the BN parameters, and then iteratively refines them. Each iteration ensures that the likelihood function increases and eventually converges to a local maximum. The iteration process consists of two steps, namely the E-step (i.e., expectation) and the M-step (i.e., maximisation), which are performed in alternating manner until convergence or a user defined number of iterations is reached.

In the remainder of this section, because of their relevance to the research work carried out in this thesis, the EM applications to learning the parameters of a BN with discrete variables, and to learning the parameters of a mixture of multivariate Gaussian densities are described.

Expectation Maximisation for Bayesian Networks with discrete variables

Mathematically, the EM algorithm applied to the task of learning the CPT parameters can be expressed as follows (Zhang, 1996). Considering that the loglikelihood function of Θ given data D is defined by:

$$L(\Theta|D) = ln\, P(D|\Theta) = \sum_{l} ln\, P(d_l|\Theta) \qquad \text{(E.3)}$$

where d_l represent a data case. The algorithm starts with an initial estimate $\Theta^{(0)}$ of Θ and improves it iteratively. From the current estimate $\Theta^{(t)}$, the next estimate $\Theta^{(t+1)}$ is obtained by firstly (i.e., E-step) computing the current expected loglikelihood function of Θ given data D:

$$Q\left(\Theta\middle|\Theta^{(t)}\right) = \sum_{l}\sum_{X_l} ln\, P(d_l, X_l\,|\Theta)\, P\left(X_l\,\middle|d_l, \Theta^{(t)}\right) \tag{E.4}$$

where X_l is the set of variables whose values are missing from data case d_l. Then, the M-step chooses the next estimate $\Theta^{(t+1)}$ by maximising the current expected loglikelihood: $Q\left(\Theta^{(t+1)}\middle|\Theta^{(t)}\right) \geq Q\left(\Theta\middle|\Theta^{(t)}\right)$ for all Θ. By making use of Equation (E.1), the following equation is obtained:

$$Q\left(\Theta\middle|\Theta^{(t)}\right) = \sum_{i,j,k} f_t\,(X_i = j, \pi_i = k) ln\, \Theta_{ijk} \tag{E.5}$$

where

$$f_t(X_i, \pi_i) = \sum_{l} P\ (X_i, \pi_i | d_l, \Theta^{(t)}) \tag{E.6}$$

Hence, the E-step reduces to computing the function $f_t(X_i, \pi_i)$ for each variable X_i, while the M-step reduces to computing the next estimate of Θ by setting:

$$\Theta_{ijk}^{(t+1)} = \frac{f_t(X_i = j, \pi_i = k)}{\sum_j f_t\ (X_i = j, \pi_i = k)} \quad \text{for all } i, j, \text{and } k. \tag{E.7}$$

Expectation Maximisation for mixture of multivariate Gaussian densities

A multivariate mixture function can be defined as the weighted sum of multivariate component densities:

$$f(y) = \sum_{i=1}^{c} p_i\, g(y; \theta_i) \tag{E.8}$$

where p_i represents the weight (or mixing coefficient) for the i^{th} term (i.e., component), c is the number of components, and $g(y; \theta_i)$ denotes a probability density, with

parameters represented by the vector θ_i. Note that to ensure that $f(y)$ is a bona fide density, the following conditions hold: $p_1 + \cdots + p_c = 1$, and $p_i > 0$.

When a multivariate Gaussian is used as the component density, the following equation represents an estimate of the multivariate mixture function:

$$\hat{f}(y) = \sum_{i=1}^{c} \hat{p}_i \, \emptyset(y; \hat{\mu}_i, \hat{\Sigma}_i) \tag{E.9}$$

where y is a n-dimensional vector, $\hat{\mu}_i$ is a n-dimensional vector of means, and $\hat{\Sigma}_i$ is a $n \times n$ covariance matrix.

Bearing in mind the above, the EM algorithm is described here as it can be applied to estimating the parameters of a mixture of multivariate Gaussian densities (Redner and Walker, 1984).

The first step is to determine the posterior probabilities given by:

$$\hat{\tau}_{ij} = \frac{\hat{p}_i \emptyset(y_j; \hat{\mu}_i, \hat{\Sigma}_i)}{\hat{f}(y_j)} \qquad i = 1, \ldots, c; \;\; j = 1, \ldots, n. \tag{E.10}$$

where $\hat{\tau}_{ij}$ represents the estimated posterior probability that point y_j belongs to the i^{th} component, $\emptyset(y_j; \hat{\mu}_i, \hat{\Sigma}_i)$ is the multivariate normal density for the i^{th} component evaluated at y_j, and

$$\hat{f}(y_j) = \sum_{k=1}^{c} \hat{p}_k \, \emptyset(y_j; \hat{\mu}_k, \hat{\Sigma}_k) \tag{E.11}$$

is the multivariate Gaussian mixture estimate at point y_j.

The posterior probability indicates the likelihood that a point belongs to each of the separate component densities. This estimated posterior probability can then be used to obtain a weighted update of the parameters for each component. This yields the iterative EM update equations for the mixing coefficients:

$$\hat{p}_i = \frac{1}{n} \sum_{j=1}^{n} \hat{\tau}_{ij} \tag{E.12}$$

the means:

$$\hat{\mu}_i = \frac{1}{n}\sum_{j=1}^{n}\frac{\hat{\tau}_{ij}y_j}{\hat{p}_i} \tag{E.13}$$

and the covariance matrices:

$$\hat{\Sigma}_i = \frac{1}{n}\sum_{j=1}^{n}\frac{\hat{\tau}_{ij}\left(y_j - \hat{\mu}_i\right)(y_j - \hat{\mu}_i)^T}{\hat{p}_i} \tag{E.14}$$

APPENDIX F GEOSTATISTICAL TECHNIQUES

F.1 Introduction

Geostatistics is a branch of statistics concentrating on spatial datasets. Geostatistical techniques focus on the relationship between the value of a variable at a given location and the values of the same and (possibly) other variables at locations some distance from it (i.e., at a set of measured locations). Their basic goal is to interpolate the value of a variable at locations which have not been measured, using data from the surrounding measured locations. Ultimately, they allow creating a model (i.e., interpolation surface) of how the variable's values are distributed across the entire domain of interest. Geostatistical techniques were originally developed for the analysis of ore reserves in mining engineering (Matheron, 1963; 1971), but have since been applied to a number of problems in many other fields (see – e.g., Isaaks and Srivastava, 1989; Cressie, 1993; Banerjee et al., 2004). According to the bibliographic research of Zhou et al. (2007) and Hengl et al. (2009), the top ten application fields of the geostatistical techniques are: (1) geosciences, (2) water resources, (3) environmental sciences, (4) agriculture and/or soil sciences, (5/6) mathematics and statistics, (7) ecology, (8) civil engineering, (9) petroleum engineering, and (10) meteorology. An application of these techniques for WDS management is presented in Magini et al. (2007). In their work the authors used geostatistical techniques in order to overcome the lack of hydraulic measurements when calibrating a steady-state HM aiming at performing leak detection and location.

In spatial statistics, geostatistical techniques have been often synonymous with Kriging methods (Krige, 1951, 1966) which are a statistical version of interpolation. Nevertheless, the current widened definition includes both stochastic and deterministic interpolation techniques. Both these types of techniques allow performing interpolation by relying on the fundamental geographic principle which states that things that are closer together tend to be more alike than things that are farther apart (i.e., Tobler's principle – Tobler, 1970). However, deterministic techniques do not use concepts of probability theory, thus they provide no indication of the extent of possible interpolation errors. Conversely, stochastic techniques incorporate the concept of randomness, meaning that the resulting interpolation surface is conceptualised as one of many that might have been observed and all of which could have produced the measured

variable's values. Stochastic techniques allow the computation of the statistical significance of the interpolation surface and of the predicted variable's values uncertainty. This difference is fundamental for assessing the "quality" of the predictions.

In this appendix two deterministic and two stochastic geostatistical techniques are described first. Specifically, the IDW (see – e.g., Shepard, 1968), and the LP (see – e.g., Gandin, 1963; Cleveland and Devlin, 1988) deterministic interpolation techniques are respectively described in Sections F.2 and F.3. The OK (see – e.g., Isaaks and Srivastava, 1989), and the OC (see – e.g., Myers, 1982) stochastic interpolation techniques are respectively described in Sections F.4 and F.5. Note that, since in-depth discussions about these and other geostatistical techniques for spatial interpolation can be found in the above references and in – e.g., Cressie (1993), Deutsch and Journel (1998), and Banerjee et al. (2004), only an outline of their main characteristics is given here. Once this is done, the method and summary statistic used in the research work presented in this thesis for testing the prediction performance of the considered geostatistical techniques are presented in Section F.6.

F.2 Inverse Distance Weighted Interpolation

IDW interpolation is a deterministic technique that estimates the variable's value at a prediction (i.e., unmeasured) location using a linear combination of the measured variable's values surrounding the prediction location weighted by an inverse function of the distance from the prediction location to the measured locations. It explicitly implements Tobler's principle by assuming that the measured variable's values closest to the prediction location have the greatest influence on the predicted variable's value.

The variable's value at a prediction location s_0, is computed as a weighted average according to the following formula:

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i \, Z(s_i) \qquad \text{(F.1)}$$

where n is the number of measured variable's values surrounding the prediction location, λ_i are the weights assigned to each measured variable's value, and $Z(s_i)$ is the measured variable's value at the location s_i.

The weights are usually determined by using the following weighting function:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^{n} d_{i0}^{-p}} \qquad \text{where} \qquad \sum_{i=1}^{n} \lambda_i = 1 \tag{F.2}$$

In the latter equation, the weights are scaled so that their sum is as equal to 1. This is to ensure an unbiased interpolation. The quantity d_{i0} is the distance between the prediction location and each of the measured locations. As d_{i0} increases, a weight approaches zero. The factor p is a power parameter that influences how fast the weighs decrease as d_{i0} increases. If the p value is very high, only the immediate few surrounding points influence the prediction. The choice of the p value is arbitrary (Webster and Oliver, 2001). If $p = 0$, there is no decrease with distance and, because each weight is the same, the prediction is the mean of all the measured values. If $p = 1$, the predictions are obtained by simple linear interpolation. This said, the most popular choice of p is 2 (Li and Heap, 2008) and the resulting method is often called Inverse Square Distance or Inverse Distance Squared. The power parameter can also be chosen on the basis of an error measurement (e.g., minimum Mean Absolute Error) and the resulting method is often called the Optimal IDW (Collins and Bolstad, 1996). Note that the comparative merits of other weighting functions are discussed in detail by Lancaster and Salkauskas (1986).

IDW interpolation is one of the simplest geostatistical techniques. However, it does not take advantage of the spatial correlation structure of the data explicitly. Furthermore, IDW is an exact interpolation technique. That is, it predicts a variable's value identical to the measured variable's value at a measured location (i.e., it forces the resulting surface to pass through all the measured variable's values). This implies that this technique may generate surfaces with sharp peaks or valleys. Finally, the interpolated values at any location within the domain of interest are bounded by the maximum and minimum of the measured variable's values. This is considered to be an important shortcoming because, in order to be useful, an interpolation surface should predict accurately certain important features of the "true" surface. For example, the locations and magnitudes of maxima and minima, even when they are not included in the set of measured variable's values (Lam, 1983).

F.3 Local Polynomial Interpolation

LP interpolation is a deterministic technique that finds its roots in the Trend Surface Analysis theory (see – e.g., Cressie, 1993). Trend Surface Analysis assumes that the variable's values in the domain of interest are a function of the geographic coordinates. Each measured variable's value $Z(s)$ is considered to be the sum of a deterministic polynomial function of the geographic coordinates $f(x, y)$ (i.e., trend surface) plus a random error term ε (Webster and Oliver, 2001):

$$Z(s) = f(x, y) + \varepsilon \tag{F.3}$$

The polynomial function can be expanded to any desired degree. The coefficients of the polynomial function are found by the method of Least-Squares which makes sure that the sum of the squared deviations from the trend surface is minimised. The variable's values at the prediction locations are then estimated by substituting the coordinates of the prediction locations into the polynomial function (i.e., the predicted variable's values are approximated by the fitted trend surface).

In the framework outlined above, LP interpolation uses the variable's values at the measured locations within localised (and overlapping) windows rather than using the variable's values at all the measured locations. This implies that a set of "local" trend surfaces are fitted to the measured variable's values and then patched together to construct the final interpolation surface. The window can be moved around and the surface value at the centre of the window $\mu_0(x, y)$ is estimated at each point. Weighted Least-Squares is used by minimizing:

$$\sum_{i=1}^{n} w_i \left[Z(x_i, y_i) - \mu_0(x_i, y_i)\right]^2 \tag{F.4}$$

where n is the number of points within the window, $Z(x_i, y_i)$ is the measured variable's value at location (x_i, y_i), $\mu_0(x_i, y_i)$ is the value of the polynomial at location (x_i, y_i), and w_i is a weight which can be computed as follows:

$$w_i = \exp\left(\frac{-3d_{i0}}{a}\right) \tag{F.5}$$

where d_{i0} is the distance between the location (x_i, y_i) and the centre of the window, and a is a parameter that controls how fast the weights decay with distance.

For a first-order polynomial $\mu_0(x_i, y_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i$ meaning that a single plane is fitted through the variable's values measured within the considered moving window. For a second-order polynomial $\mu_0(x_i, y_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i + \beta_3 x_i^2 + \beta_4 y_i^2 + \beta_5 x_i y_i$ meaning that a surface with a bend in it is fitted through the measured variable's values. Progressing in a like fashion, a similar expression (including the sums of powers and cross products of the x and y coordinates) can be written for a third-order polynomial, which implies that a surface with two bends is fitted through the measured variable's values; and so forth for a higher order polynomial. The minimisation occurs for the coefficients $\{\beta_i\}$. These coefficients are re-estimated whenever the centre point and, consequently, the window moves (Gandin, 1963; Johnston et al., 2001).

Similarly to the IDW interpolation, LP interpolation does not take advantage of the spatial correlation structure of the data explicitly. Indeed, it only uses the geographical coordinates to predict the variable's values. Differently from the IDW interpolation, however, LP is an inexact interpolation technique. When the input dataset exhibits short-range variations, LP interpolation can be a good method to capture the finer details (Akima, 1970).

F.4 Ordinary Kriging

OK is the most widely used version of Kriging. It is similar to the IDW interpolation technique, in that it uses a weighted linear combination of the measured variable's values surrounding the prediction location to estimate the variable's value at the prediction location. Because OK is a stochastic technique, however, two statistical data analysis steps have to be followed for generating an interpolation surface. These steps are: (1) quantifying the spatial correlation structure of the measured variable's values (also known as variography), and (2) estimating the variable's values at the prediction locations. During the first step, a spatial dependence model is fit to the variable's values at the measured locations. During the second step, the fitted model from variography, the spatial data configuration of the prediction locations and the measured variable's values surrounding the prediction locations are used to perform prediction. By following this procedure, OK provides a solution to the problem of estimation of the interpolation surface that takes into account the spatial correlation structure of the data.

In OK the predictions are based on the following model:

$$Z(s) = \mu + \varepsilon(s) \tag{F.6}$$

where $Z(s)$ is the variable's value at location s, μ is an unknown constant mean, and $\varepsilon(s)$ is the spatially correlated part of variation. The predictions are made according to Equation (F.1). Here, however, the weights are based not only on the distance between measured and prediction locations, but also on the overall spatial arrangement of the measured locations, and on their values.

The spatial correlation between the measured variable's values can be quantified by means of the following semivariance function:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} [Z(s_k + h) - Z(s_k)]^2 \tag{F.7}$$

where $N(h)$ is the number of pairs of measurement locations with distance h apart. If the two locations, s_k and $s_k + h$, are close to each other, they are expected to be similar, and so the difference in their values, $Z(s_k + h) - Z(s_k)$, is expected to be small. As s_k and $s_k + h$ get farther apart, they are expected to become less similar and so the difference in their values is expected to become larger. A plot of the calculated $\gamma(h)$ against the distance h is referred to as an experimental semivariogram.

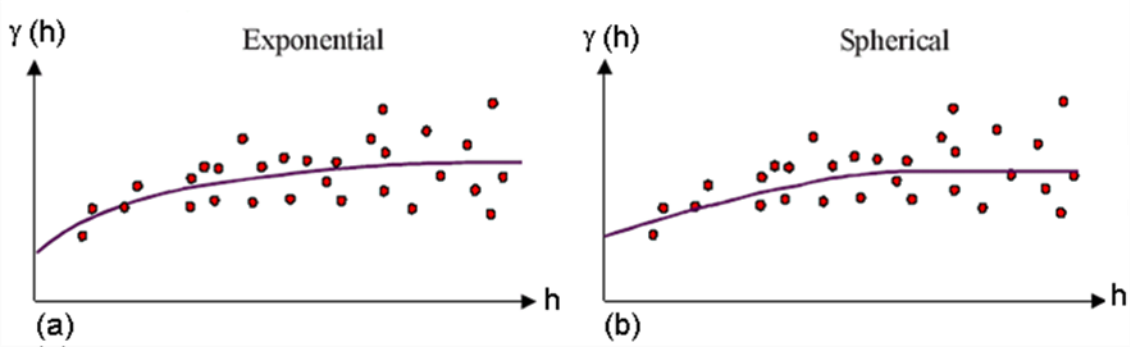


Figure F.1. Parametric functions for experimental semivariogram fitting. Exponential (a), and Spherical (b).

A parametric function, such as Linear, Spherical, Exponential, Circular, Gaussian, Bessel, Power and similar (see – e.g., Isaaks and Srivastava, 1989) is then fitted to the experimental semivariogram. Figure F.1 shows, as an example, two of the aforementioned parametric functions. The experimental semivariogram is commonly fitted by Iterative Reweighted Least-Squares estimation, where the weights are determined based on the number of point pairs or based on the distance (Hengl, 2009).

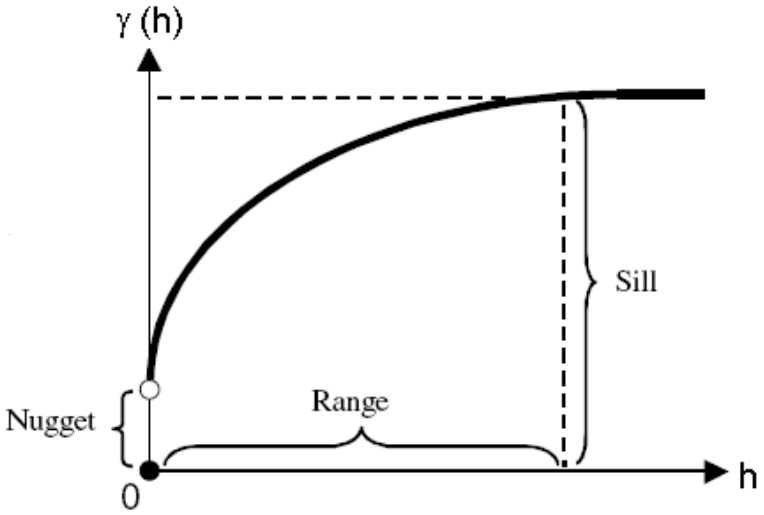


Figure F.2. Parameters of a typical parametric function for experimental semivariogram fitting.

As shown in Figure F.2, a fitted parametric function has several parameters which have to be estimated. The distance where the parametric function first flattens out is known as the Range. The value at which the parametric function attains the Range is called the Sill. The value at which the parametric function intercepts the *y* axis is called the Nugget. The Nugget is simply the sum of measurement error and microscale variation. Note that the choice among the various parametric functions depends on the shape of the experimental semivariogram. However, the Spherical parametric function is the most widely used and often preferred when the Nugget variance is important and there is a clear Range and Sill effect (Cressie, 1993; Burrough and McDonnell, 1998). Once the parametric function's parameters have been estimated (i.e., the semivariogram has

been modelled), the fitted parametric function can be used to derive the semivariances at all locations and compute the OK weights.

The unknown weights are calculated by minimising the estimation variance $Var\left[\hat{Z}(s_0) - Z(s_0)\right]$, subject to the constraint that the estimate is unbiased – i.e., $E\left[\hat{Z}(s_0) - Z(s_0)\right] = 0$. This yields the following OK system of equations:

$$\begin{aligned} &\sum_{i=1}^{n} \lambda_i \gamma(s_m, s_i) + \varphi = \gamma(s_m, s_0); \;\; m = 1, \dots, n \\ &\sum_{i=1}^{n} \lambda_i = 1 \end{aligned} \tag{F.8}$$

where φ is the Lagrange Multiplier (introduced for minimisation of the estimation variance). Once the above system of equations has been solved for the weights and the Lagrange Multiplier, the variable's values at the prediction locations are calculated by inputting the weights into Equation (F.1).

The estimation variance (also known as the OK variance, $\hat{\sigma}_e^2$) for each prediction location can also be computed by using the following formula:

$$\hat{\sigma}_e^2(s_0) = \sum_{i=1}^{n} \lambda_i \gamma(s_i, s_0) + \varphi \tag{F.9}$$

The OK variance (or more usually the OK standard deviation $\hat{\sigma}_e$) can be mapped. This provides an indication of the reliability of the estimate at each prediction location.

OK presents several advantages over the IDW and LP interpolation techniques. Firstly, OK provides the best linear unbiased estimate as it attempts to optimise the weights assigned to the variable's values at the neighbouring measured locations. Secondly, OK also provides a measure of the error or uncertainty at the prediction locations. Furthermore, it does not produce edge-effects resulting from trying to force a polynomial to fit the data.

F.5 Ordinary Cokriging

OC is a stochastic technique that can be seen as an extension of OK to the case of more than one variable (Journel and Huijbregts, 1978). Similarly to OK, OC computes the

interpolated values of the variable of interest by optimising the weights assigned to the variable's values at the neighbouring measured locations based on the spatial correlation between the measured variable's values. However, OC also relies on the relationship between the variable of interest and other variables (secondary/auxiliary variables) and uses the information from other variables in an attempt to create a better prediction model.

In the case of two variables U and V which are spatially correlated, the OC predictions are made according to the following formula:

$$\hat{Z}_U(s_0) = \sum_{i=1}^{n_U} \lambda_{Ui} Z_U(s_i) + \sum_{j=1}^{n_V} \lambda_{Vj} Z_V(s_j) \quad \text{(F.10)}$$

where $\hat{Z}_U(s_0)$ is the estimate for U at the prediction location s_0, n_U and n_V are the number of measured values (used for the prediction) of the primary variable U, and of the secondary variable V, respectively, $Z_U(s_i)$ is the measured value of the primary variable at the location s_i, $Z_V(s_j)$ is the measured value of the secondary variable at the location s_j, and λ_{Ui} and λ_{Vj} are the associated weights.

In OC, besides the experimental semivariograms for both U and V, information on the joint spatial co-variation (i.e., interaction) of both variables is taken into consideration as well. For any pair of variables U and V the cross-semivariance $\gamma_{UV}(h)$, at lag h is defined as:

$$\gamma_{UV}(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} [Z_U(s_k + h) - Z_U(s_k)] * [Z_V(s_k + h) - Z_V(s_k)] \quad \text{(F.11)}$$

A plot of the calculated $\gamma_{UV}(h)$ against the distance h is referred to as an experimental cross-semivariogram. A parametric function is then fitted to the experimental cross-semivariogram and its parameters estimated. Once this is done, the estimated parameters of the parametric function fitted to the experimental cross-semivariogram are used together with the estimated parameters of the parametric functions fitted to the U and V experimental semivariograms for computing the OC weights.

In OC, similarly to the OK procedure, the unknown weights are calculated by minimising the estimation variance subject to the constraint that the estimate is unbiased, resulting in the following OC system of equations:

$$\begin{aligned}
&\sum_{i=1}^{n_U} \lambda_{Ui}\,\gamma_{UU}(s_m, s_i) + \sum_{j=1}^{n_V} \lambda_{Vj}\,\gamma_{UV}(s_m, s_j) + \varphi_U = \gamma_{UU}(s_m, s_0); m = 1, \dots, n_U \\
&\sum_{i=1}^{n_U} \lambda_{Ui}\,\gamma_{UV}(s_m, s_i) + \sum_{j=1}^{n_V} \lambda_{Vj}\,\gamma_{VV}(s_m, s_j) + \varphi_V = \gamma_{UV}(s_m, s_0); m = 1, \dots, n_V \\
&\sum_{i=1}^{n_U} \lambda_{Ui} = 1 \\
&\sum_{j=1}^{n_V} \lambda_{Vj} = 0
\end{aligned} \tag{F.12}$$

Solving this system of equations for the weights λ_{Ui} and λ_{Vj}, and for the Lagrange Multipliers φ_U and φ_V, allows the calculation of the values of the variable of interest at the prediction location by using Equation (F.10), and the estimation of the OC variance:

$$\hat{\sigma}_e^2(s_0) = \sum_{i=1}^{n_U} \lambda_{Ui}\,\gamma_{UU}(s_i, s_0) + \sum_{j=1}^{n_V} \lambda_{Vj}\,\gamma_{UV}(s_j, s_0) + \varphi_U \tag{F.13}$$

OC presents the same advantages of OK over the IDW and LP interpolation techniques. It however may improve the predictions and reduce the variance of the estimation error by drawing on the additional information from the other spatially correlated variables to help with interpolation.

F.6 Test of Prediction Performance

The evaluation of the performance of the different geostatistical techniques described in the previous sections can be performed by using the Leave-One-Out Cross-Validation technique (see – e.g., Devijver and Kittler, 1982). This technique removes each of the n measured locations s_i, one at a time, and estimates the associated variable's value $\hat{Z}(s_i)$, using the remaining measured locations. The estimated and actual - i.e., $Z(s_i)$ - variable's values are then compared and a summary statistic is computed.

The summary statistic used in the research work presented in this thesis is the RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\hat{Z}(s_i) - Z\,(s_i)]^2} \qquad \text{(F.14)}$$

The RMSE is the square root of the sum of the squared residuals. This statistic not only allows seeing how closely a resulting interpolation surface estimates the actual variable's values, but may also be used to compare the performance of the different interpolation surfaces. The smaller the RMSE the better.

BIBLIOGRAPHY

Papers Presented by the Candidate

Romano, M., Kapelan, Z., and Savić, D. A. (2012a). "Geostatistical techniques for approximate location of pipe burst events in water distribution systems." *Journal of Hydroinformatics* (Submitted).

Romano, M., Kapelan, Z., and Savić, D. A. (2012b). "Evolutionary algorithm and expectation maximisation strategies for improved detection of pipe bursts and other events in water distribution systems." *Journal of Water Resources Planning and Management* (Submitted).

Romano, M., Kapelan, Z., and Savić, D. A. (2012c). "Automated detection of pipe bursts and other events in water distribution systems." *Journal of Water Resources Planning and Management* (Submitted).

Romano, M., Kapelan, Z., and Savić, D. A. (2012d). "Testing of the system for detection of pipe bursts and other events in a UK water distribution system." *Proceedings of the 14th Water Distribution Systems Analysis Conference*, Adelaide, Australia.

Romano, M., Kapelan, Z., and Savić, D. A. (2012e). "Evolutionary algorithm strategy for improved detection of bursts in water distribution systems." *Proceedings of the 10th International Conference on Hydroinformatics*, Hamburg, Germany.

Romano, M., Kapelan, Z., and Savić, D. A. (2011a). "Geostatistical techniques for approximate location of bursts in water distribution systems." *Proceedings of the 11th International Conference on Computing and Control for the Water Industry*, Exeter, UK.

Romano, M., Kapelan, Z., and Savić, D. A. (2011b). "Burst detection and location in water distribution systems." *Proceedings of the World Environmental and Water Resources Congress*, Palm Springs, USA.

Romano, M., Kapelan, Z., and Savić, D. A. (2010a). "Pressure signal de-noising for improved real-time leak detection." *Proceedings of the 9th International Conference on Hydroinformatics*, Tianjin, China.

Romano, M., Kapelan, Z., and Savić, D. A. (2010b). "Real-time leak detection in water distribution systems." *Proceedings of the 12th Annual International Symposium on Water Distribution Systems Analysis*, Tucson, USA.

Romano, M., Kapelan, Z., and Savić, D. A. (2010c). "Bayesian inference system for the fast and reliable detection of burst related leaks in real-time." *Proceedings of the 9th Water Loss Conference*, Sao Paulo, Brazil.

Romano, M., Kapelan, Z., and Savić, D. A. (2009a). "Bayesian-based online burst detection in water distribution systems." *Proceedings of the 10th International Conference on Computing and Control for the Water Industry*, Sheffield, UK.

Romano, M., Liong, S. Y., Vu, M. T., Zemskyy, P., Doan, C. D., Dao, M. H., and Tkalich, P. (2009b). "Artificial neural network for tsunami forecasting." *Journal of Asian Earth Sciences*, vol. 36, no. 1, pp. 29–37.

Patent

Romano, M., Kapelan, Z., and Savić, D. A. (2009). "Anomaly detection based in Bayesian inference." *International Application Number:* PCT/GB2010/000961.

List of References

Abdullah, S., Sahadan, S. N., Nuawi, M. Z., and Nopiah, Z. M. (2008). "Fatigue road signal denoising process using the 4th order of Daubechies wavelet transforms." *Journal of Applied Sciences,* vol. 8, pp. 2496–2509.

Abramson, B. (1994). "The design of belief network-based systems for price forecasting." *Computers and Electrical Engineering*, vol. 20, no. 2, pp. 163–180.

Adamowski, J. F. (2008). "Peak daily water demand forecast modeling using artificial neural networks". *Journal of Water Resource Planning and Management*, vol. 134, no. 2, pp. 119–128.

Addison, P. S. (2002). "*The illustrated wavelet transform handbook*." Institute of Physics Publishing Ltd., Bristol, UK.

ADEC (2011). "*Technical review of leak detection technologies*." Alaska Department of Environmental Conservation, http://dec.state.ak.us/spar/ipp/docs/ldetect1.pdf (Accessed 28 Nov 2011).

Akima, H. (1970). "A new method of interpolation and smooth curve fitting based on local procedures." *Journal of Association for Computing Machinery*, vol. 17, no. 4, pp. 589–602.

Aksela, K., Aksela, M., and Vahala, R. (2009). "Leakage detection in a real distribution network using a SOM." *Urban Water Journal*, vol. 6, no. 4, pp. 279–289.

Al-Rafai, W., and Barnes, R. J. (1999). "Underlying the performance of real-time software-based pipeline leak-detection systems." *Pipes and Pipelines International*, vol. 44, no. 6, pp. 44–51.

Andersen, J. H., and Powell, R. S. (2000). "Implicit state-estimation technique for water network monitoring." *Urban Water*, vol. 2, no. 2, pp. 123–130.

Annan, A. P., Davis, J. L., and Vaughan, C. J. (1984). "*Radar mapping of buried pipes and cables*." Technical Note no. 1, A-Cubed Inc., Mississauga, Canada.

Anonymous (1997). "Intelligent pigs now inspect pipelines down to 200 mm diameter." *NonDestructive Testing and Evaluation International*, vol. 30, no. 1, pp. 41.

ASCE (1993). "Criteria for evaluation of watershed models." American Society of Civil Engineers Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, Irrigation and Drainage Division, *Journal of Irrigation and Drainage Engineering*, vol. 119, no. 3, pp. 429–442.

Atherton, D. L. (1989). "Magnetic inspection is key to ensuring safe pipelines." *Oil and Gas Journal*, vol. 8, no. 32, pp. 52–61.

AWWA (1999). "*Water audits and leak detection*." American Water Works Association, Manual of Water Supply Practices M36, Denver, USA.

AWWA (2009). "*Water audits and loss control programs*." 3rd edition, American Water Works Association, Manual of Water Supply Practices M36, Denver, USA.

Babovic, V., and Drecourt, J. (2000). "Data mining approach to pipe burst risk assessment." *Proceedings of Water Network Modelling for Optimal Design and Management*, Exeter, UK.

Back, T. (1996). "*Evolutionary algorithms in theory and practice*." Oxford University Press, New York, USA.

Banerjee, S., Carlin, C. P., and Gelfand, A. E. (2004). "*Hierarchical modeling and analysis for spatial data*." Monographs on Statistics and Applied Probability, Chapman and Hall/CRC, Boca Raton, USA.

Bargiela, A., and Hainsworth, G. (1989). "Pressure and flow uncertainty in water systems." *Journal of Water Resource Planning and Management*, vol. 115, no. 2, pp. 212–229.

Barrett, C. B. (2000). "World water crisis seeks engineering solutions." *Journal of Water Resources Planning and Management*, vol. 126, no. 5, pp. 268–269.

Bauer, D., Koller, E., and Singer, Y. (1997). "Update rules for parameter estimation in Bayesian networks." *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, Providence, Rhode Island, USA.

Beck, S. B. M., Curren, M. D., Sims, N. D., and Stanway, R. (2005). "Pipeline network features and leak detection by cross-correlation analysis of reflected waves." *Journal of Hydraulic Engineering*, vol. 131, no. 8, pp. 715–723.

Ben-Gal, I. (2008). "Bayesian networks." in "*Encyclopaedia of statistics in quality and reliability*." John Wiley and Sons, Chichester, UK.

Bentley Systems Inc. (2006). "*WaterGEMS v8 user's manual*." Watertown, USA.

Berardi, L., Giustolisi, O., and Primativo, F. (2007). "Exploiting multi-objective strategies for optimal rehabilitation planning." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Berardi, L., Giustolisi, O., and Savić, D. A. (2009). "An operative approach to water distribution system rehabilitation." *Proceedings of the World Environmental and Water Resources Congress*, Kansas City, USA.

Bernstein, R., Oristaglio, M., Miller, D. E., and Haldorsen, J. (2000). "Imaging radar maps underground objects." *Institute of Electrical and Electronics Engineers Transactions in Computer Applications in Power*, vol. 13, no. 3, pp. 20–24.

Beyer, H. -G. (1992). "Some aspects of the 'evolution strategy' for solving tsp-like optimization problems." in "*Parallel problem solving from nature*." Elsevier, Oxford, USA.

Beyer, H. -G. (2001). "*The theory of evolution strategies*." Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, Berlin, Germany.

Beyer, H. -G., and Schwefel, H. -P. (2002). "Evolution strategies: a comprehensive introduction." *Journal of Natural Computing*, vol. 1, no. 1, pp. 3–52.

Bishop, C. M. (1994). "*Mixture density networks*." Technical Report, Department of Computer Science and Applied Mathematics, Aston University, Birmingham, UK.

Bishop, C. M. (1995). "*Neural networks for pattern recognition*." Oxford University Press, New York, USA.

Bissell, A. F. (1978). "An attempt to unify the theory of quality control procedures." *Bulletin in Applied Statistics*, vol. 5, no.2, pp. 113–128.

Bogert, B. P., and Ossanna, J. F. (1966). "Computer experimentation on echo detection using cepstrum and pseudo auto covariance." *Journal of the Acoustical Society of America*, vol. 39, no. 6, pp. 1258–1265.

Booker, L. B., and Hota, N. (1986). "Probabilistic reasoning about ship images." *Proceedings of the 2nd Annual Conference on Uncertainty in Artificial Intelligence*, Philadelphia, USA.

Bougadis, J., Adamowski, K., and Diduch, R. (2005). "Short-term municipal water demand forecasting." *Hydrological Processes*, vol. 19, no. 1, pp. 137–148.

Bowden, G. J., Maier, H. R., and Dandy, G. C. (2002). "Optimal division of data for neural network models in water resources applications." *Water Resources Research*, vol. 38, no. 2, pp. 1–11.

Boxall, J. B., O'Hagan, A., Pooladsaz, S., Saul, A., and Unwin, D. M. (2005). "Pipe level estimation of burst rates in water distribution mains." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Bozdogan, H. (1987). "Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions." *Psychometrika*, vol. 52, no. 3, pp. 345–370.

Bracken, M., and Hunaidi, O. (2005). "Practical aspects of acoustical leak location on plastic and large diameter pipe." *Proceedings of the International Water Association Conference: Leakage 2005*, Halifax, Canada.

Breyfogle, F. W. (1999). "*Implementing six sigma: smarter solutions using statistical methods*." John Wiley and Sons, New York, USA.

Brill, F. Z., Brown, D. E., and Martin, W. N. (1992). "Fast genetic selection of features for neural network classifiers." *Institute of Electrical and Electronics Engineers Transactions on Neural Networks*, vol. 3, no. 2, pp. 324–328.

Brown, M., An, P. C., Harris, C. J., and Wang, H. (1993). "How biased is your multi-layer perceptron?" *Proceedings of the World Congress on Neural Networks*, Portland, USA.

Brunone, B. (1999). "Transient test-based technique for leak detection in outfall pipes." *Journal of Water Resources Planning and Management*, vol. 125, no. 5, pp. 302–306.

Brunone, B., and Ferrante, M. (1999). "On leak detection in single pipes using unsteady-state tests." *Proceedings of the International Conference on Modelling and Simulation*, Philadelphia, USA.

Brunone, B., and Ferrante, M. (2001). "Detecting leak in pressurized pipes by means of transient." *Journal of Hydraulic Research*, vol. 39, no. 5, pp. 539–547.

Brunone, B., and Ferrante, M. (2004). "Pressure waves as a tool for leak detection in closed conduits." *Urban Water Journal*, vol. 1, no. 2, pp. 145–155.

Buchberger, S. G., and Nadimpalli, G. (2004). "Leak estimation in water distribution systems by statistical analysis of flow readings." *Journal of Water Resources Planning and Management*, vol. 130, no. 4, pp. 321–329.

Buffa, E. S. (1972). "*Operations management: problems and models*." 3rd edition, John Wiley and Sons, New York, USA.

Burn, D. H., and Yulianti, J. S. (2001). "Waste-load allocation using genetic algorithms." *Journal of Water Resources Planning and Management*, vol. 127, no. 2, pp. 121–129.

Burrough, P. A., and McDonnell, R. A. (1998). "*Principles of geographical information systems*." Oxford University Press, New York, USA.

Caldecott, R., Poirier, M., Scofea, D., Svoboda, D. E., and Terzuoli, A. J. (1988). "Underground mapping of utility lines using impulse radar." *Institution of Electrical Engineers Proceedings F, Radar and Signal Processing*, vol. 135, no. 4, pp. 343–353.

Caputo, A. C., and Pelagagge, P. M. (2002). "An inverse approach for piping networks monitoring." *Journal of Loss Prevention in the Process Industries*, vol.15, no. 6, pp. 497–505.

Caputo, A. C., and Pelagagge, P. M. (2003). "Using neural networks to monitor piping systems." *Process Safety Progress*, vol. 22, no. 2, pp. 119–127.

Carpentier, P., and Cohen, G. (1991). "State estimation and leak detection in water distribution networks." *Civil Engineering Systems*, vol. 8, no.4, pp. 247–257.

Carpentier, P., and Cohen, G. (1993). "Applied mathematics in water supply network management." *Automatica*, vol. 29, no. 5, pp. 1215–1250.

Castillo, E., Gutierrez, J. M., and Hadi, A. S. (1997). "*Expert systems and probabilistic network models*." Springer-Verlag, New York, USA.

Castillo, P. A., Carpio, J., Merelo Guervós, J. J., Rivas, V., Romero, G., and Prieto, A. (2000). "Evolving multilayer perceptrons." *Neural Processing Letters*, vol. 12, no. 2, pp. 115–127.

Charlton, M. B., and Mulligan, M. (2001). "Efficient detection of mains water leaks using ground-penetrating radar." *Proceedings of Subsurface and Sensing Technologies and Applications III*, San Diego, USA.

Chaudhry, H. M. (1987). "*Applied hydraulic transients*." Van Nostrand Reinhold, New York, USA.

Chickering, D. M., and Heckerman, D. (2004). "Large-sample learning of Bayesian networks is NP-hard." *Journal of Machine Learning Research*, vol. 5, pp. 1287–1330.

Ciochetto, G., and Polidoro, R. (1998). "Site investigation and output of utilities map using GPR." *Centro Studi E Laboratori Telecomunicazioni Technical Reports*, vol. 26, no. 2, pp. 177–184.

Cleveland, W. S., and Devlin, S. J. (1988). "Locally weighted regression: an approach to regression analysis by local fitting." *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 596–610.

Cohen, A. (2003). "*Numerical analysis of wavelet methods*." Elsevier, Oxford, UK.

Collins, F. C., and Bolstad, P. V. (1996). "A comparison of spatial interpolation techniques in temperature estimation." *Proceedings of the 3rd International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, USA.

Colombo, A. F., Lee, P. J., and Karney, B. W. (2009). "A selective literature review of transient-based leak detection methods." *Journal of Hydro-environment Research*, vol. 2, no. 4, pp. 212–227.

Cooper, G. (1990). "The computational complexity of probabilistic inference using Bayesian belief networks." *Artificial Intelligence*, vol. 42, no. 2-3, pp. 393–405.

Cooper, G., and Herskovits, E. (1992). "A Bayesian method for the induction of probabilistic networks from data." *Machine Learning*, vol. 9, no. 4, pp. 309–347.

Covas, D., and Ramos, H. (1999). "Leakage detection in single pipelines using pressure wave behaviour." *Proceedings of the 5th International Conference on Computing and Control in the Water Industry*, London, UK.

Covas, D., and Ramos, H. (2001). "Hydraulic transients used for leak detection in water distribution systems." *Proceedings of the 4th International Conference on Water Pipeline Systems*, York, UK.

Covas, D., and Ramos, H. (2010). "Case studies of leak detection and location in water pipe systems by inverse transient analysis." *Journal of Water Resources Planning and Management*, vol. 136, no. 2, pp. 248–257.

Covas, D., Graham, N., Maksimovic, C., Ramos, H., Kapelan, Z., Savić, D. A., and Walters, G. A. (2003). "An assessment of the application of inverse transient analysis for leak detection. II: Collection and application of experimental data." *Proceedings of the 7th International Conference on Computing and Control in the Water Industry*, London, UK.

Covas, D., Ramos, H., and Almeida, A. B. (2000). "Leak location in pipe systems using pressure surges." *Proceedings of the 8th International Conference on Pressure Surges*, The Hague, Netherlands.

Covas, D., Ramos, H., Brunone, B., and Young, A. (2004). "Leak detection in water trunk mains using transient pressure signals: field tests in Scottish Water." *Proceedings of the 9th International Conference on Pressure Surges*, Chester, UK.

Covas, D., Ramos, H., Graham, N., and Maksimovic, C. (2005a). "Application of hydraulic transients for leak detection in water supply systems." *Water Supply*, vol. 4, no. 5-6, pp. 365–374.

Covas, D., Ramos, H., Lopes, N., and Almeida, A. B. (2006). "Water pipe system diagnosis by transient pressure signals." *Proceedings of the 8th Annual International Symposium on Water Distribution Systems Analysis*, Cincinnati, USA.

Covas, D., Ramos, H., Young, A., Graham, I., and Maksimovic, C. (2005b). "Uncertainties of leak detection by means of hydraulic transients from the lab to the field." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Covas, D., Stoianov, I., Butler, D., Maksimovic, C., Graham, N., and Ramos, H. (2001). "Leak detection in pipeline systems by inverse transient analysis-from theory to practice." *Proceedings of the 6th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Cowell, R., Dawid, A., Lauritzen, S., and Spiegelhalter, D. (1999). "*Probabilistic networks and expert systems*." Springer-Verlag, New York, USA.

Cressie, N. (1993). "*Statistics for spatial data*." John Wiley and Sons, New York, USA.

Dandy, G. C., and Engelhardt, M. (2001). "Optimal scheduling of water pipe replacement using genetic algorithms." *Journal of Water Resources Planning and Management*, vol. 127, no. 4, pp. 214–223.

Darwin, C. (1973). "*The origin of species*." Dent Gordon, London, UK.

Daubechies, I. (1988). "Orthonormal bases of compactly supported wavelets." *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996.

Daubechies, I. (1992). "*Ten lectures on wavelets*." Society for Industrial and Applied Mathematics, New York, USA.

Davis, L. (1991). "*Handbook of genetic algorithms*." Van Nostrand Reinhold, New York, USA.

Dawsey, W., Minsker, B., and Amir, E. (2007). "Real time assessment of drinking water systems using a dynamic Bayesian network." *Proceedings of the World Water and Environmental Resources Congress*, Tampa, USA.

De Jong, K. A. (2007). "*Evolutionary computation: a unified approach*." Massachusetts Institute of Technology Press, Cambridge, USA.

De Jong, K. A. (2009). "Evolutionary computation." *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no.1, pp. 52–56.

Deagle, G., Green, A., and Scrivener, J. (2007). "Advanced tools for burst location." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Deming, W. E. (1950). "*Lectures on statistical control of quality*." Nippon Kagaku Gijutsu Remmei, Tokyo, Japan.

Deming, W. E. (1975). "On probability as a basis for action." *The American Statistician*, vol. 29, no. 4, pp. 146–152.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38.

Deutsch, C. V., and Journel, A. G. (1998). "*GSLIB: geostatistical software and user's guide*." 2nd edition, Oxford University Press, New York, USA.

Devijver, P. A., and Kittler, J. (1982). "*Pattern recognition: a statistical approach*." Prentice-Hall, Englewood Cliffs, USA.

Deza, E., and Deza, M. M. (2009). "*Encyclopedia of distances.*" Springer-Verlag, Berlin, Germany.

Do, C. B., and Batzoglou, S. (2008). "*What is the expectation maximization algorithm?*" Nature Publishing Group, http://www.nature.com/naturebiotechnology (Accessed 21 Mar 2011).

Donoho, D. L., and Johnstone, I. M. (1994). "Ideal spatial adaptation via wavelet shrinkage." *Biometrika*, vol. 81, no. 3, pp. 425–455.

Donoho, D. L., and Johnstone, I. M. (1995). "Adapting to unknown smoothness via wavelet shrinkage." *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224.

Donoho, D. L. (1995). "De-noising by soft-thresholding." *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 41, no. 3, pp. 613–617.

Duda, R., and Hart, P. (1973). "*Pattern classification and scene analysis*." John Wiley and Sons, New York, USA.

Edwards, D. (2000). "*Introduction to graphical modelling*." 2nd edition, Springer-Verlag, New York, USA.

Egan, J. P. (1975). "*Signal detection theory and ROC analysis*." Academic Press, New York, USA.

Elidan, G., and Friedman, N. (2003). "The information bottleneck EM algorithm." *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence*, Acapulco, Mexico.

Elidan, G., Ninio, M., Friedman, N., and Schuurmans, D. (2002). "Data perturbation for escaping local maxima in learning." *Proceedings of the 18th National Conference on Artificial Intelligence*, Alberta, Canada.

Elman, J. L. (1990). "Finding structure in time." *Cognitive Science*, vol. 14, no. 2, pp. 179–211.

Everitt, B. S. (2006). "*The Cambridge dictionary of statistics*." 3rd edition, Cambridge University Press, Cambridge, UK.

Fahmy, M., and Moselhi, O. (2009). "Detecting and locating leaks in underground water mains using thermography." *Proceedings of the 26th International Symposium on Automation and Robotics in Construction*, Austin, USA.

Fantozzi, M., and Fontana, E. (2001). "Acoustic emission techniques: the optimum solution for leakage detection and location in water pipelines." *Insight: Non-Destructive Testing and Condition Monitoring*, vol. 43, no. 2, pp. 105–107.

Farley, B., Boxall, J. B., and Mounce, S. R. (2008). "Optimal locations of pressure meter for burst detection." *Proceedings of the 10th Annual International Symposium on Water Distribution Systems Analysis*, Kruger National Park, South Africa.

Farley, B., Mounce, S. R., and Boxall, J. B. (2010a). "Field testing of an optimal sensor placement methodology for event detection in an urban water distribution network." *Urban Water Journal*, vol. 7, no. 6, pp. 345–356.

Farley, B., Mounce, S. R., and Boxall, J. B. (2010b). "Field validation of 'optimal' instrumentation methodology for burst/leak detection and location." *Proceedings*

of the 12th Annual International Symposium on Water Distribution Systems Analysis, Tucson, USA.

Farley, M. (2005). "Technology and equipment for water loss management – what's new?" *Proceedings of the International Water Association Conference: Leakage 2005*, Halifax, Canada.

Farley, M. (2007). "Finding the 'difficult' leaks." *Water 21*, issue 9.6, pp. 24–25.

Farley, M., and Trow, S. (2003). "*Losses in water distribution networks: a practitioners' guide to assessment, monitoring and control*." International Water Association Publishing, London, UK.

Fawcett, T. (2006). "An introduction to ROC analysis." *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874.

Fayyad, U. M., Piatetsky-Shapiro, G., and Smyth, P. (1996). "From data mining to knowledge discovery: an overview." in "*Advances in knowledge discovery and data mining*." Massachusetts Institute of Technology Press, Cambridge, USA.

Fenner, R. A., and Ye, G. (2011). "Kalman filtering of hydraulic measurements for burst detection in water distribution systems." *Journal of Pipeline Systems Engineering and Practice*, vol. 2, no. 1, pp. 14–22.

Ferrante, M., and Brunone, B. (2001). "Leak detection in pressurised pipes by means of wavelet analysis." *Proceedings of the 4th International Conference on Water Pipeline Systems*, York, UK.

Ferrante, M., and Brunone, B. (2003a). "Pipe system diagnosis and leak detection by unsteady-state tests. 1: Harmonic analysis." *Advances in Water Resources*, vol. 26, no. 1, pp. 95–105.

Ferrante, M., and Brunone, B. (2003b). "Pipe system diagnosis and leak detection by unsteady-state tests. 2. Wavelet analysis." *Advances in Water Resources*, vol. 26, no. 1, pp. 107–116.

Ferrante, M., Brunone, B., and Meniconi, S. (2007). "Wavelets for the analysis of transient pressure signals for leak detection." *Journal of Hydraulic Engineering*, vol. 133, no. 11, pp. 1274–1282.

Ferrante, M., Brunone, B., and Rossetti, A. G. (2001). "Harmonic analysis of pressure signal during transients for leak detection in pressurised pipes." *Proceedings of the 4th International Conference on Water Pipeline Systems*, York, UK.

Ferrante, M., Brunone, B., Meniconi, S., and Almadori, C. (2005). "Wavelet analysis of numerical pressure signals for leak monitoring." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Field, D. B., and Ratcliffe, B. (1978). "*Location of leaks in pressurised pipelines using sulphur hexafluoride as a tracer*." Technical Report no. 80, Water Research Centre, UK.

Fischer, H. (2010). "*A history of the central limit theorem.*" Springer-Verlag, New York, USA.

Fletcher, R. (1987). "*Practical methods of optimization*." 2nd edition, John Wiley and Sons, Chichester, UK.

Fletcher, R. (2008). "SmartBall™ – a new approach in pipeline leak detection." *Proceedings of the International Pipeline Conference*, Calgary, Canada.

Fogel, D. B. (1997). "The advantages of evolutionary computation." *Proceedings of the International Conference on Biocomputing and Emergent Computation*, Skövde, Sweden.

Fogel, L. J., Owens, A. J., and Walsh, M. J. (1966). "*Artificial intelligence through simulated evolution*." John Wiley and Sons, New York, USA.

Friedman, N. (1998). "The Bayesian structural EM algorithm." *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, Madison, USA.

Friedman, N., Geiger, D., and Goldszmidt, M. (1997). "Bayesian network classifiers." *Machine Learning*, vol. 29, no. 2-3, pp. 131–163.

Fritzke, B. (1997). "*Some competitive learning methods*." Institute for Neural Computation, Ruhr-Universität Bochum, Bochum, Germany, ftp://ftp.neuroinformatik.ruhr-uni-bochum.de/pub/software/NN/DemoGNG/sclm.ps.gz (Accessed 5 Jan 2012).

Fuchs, H. V., and Riehle, R. (1991). "Ten years of experience with leak detection by acoustic signal analysis." *Applied Acoustic*, vol. 33, no. 1, pp. 1–19.

Gabor, D. (1946). "Theory of communication." *Journal of the Institute of Electrical Engineering*, vol. 93, no. 26, pp. 429–457.

Gabrys, B., and Bargiela, A. (1995). "Neural simulation of water systems for efficient state estimation." *Proceedings of the Modelling and Simulation Conference*, Prague, Czech Republic.

Gabrys, B., and Bargiela, A. (1996). "Integrated neural based system for state estimation and confidence limit analysis in water networks." *Proceedings of the European Simulation Symposium*, Genova, Italy.

Gabrys, B., and Bargiela, A. (1999). "Neural networks based decision support in presence of uncertainties." *Journal of Water Resource Planning and Management*, vol. 125, no. 5, pp. 272–280.

Gandin, L. S. (1963). "*Objective analysis of meteorological fields*." GIMIZ, Leningrad, Russia, English translation from the Russian, Israel Program for Scientific Translation, 1965, Jerusalem, Israel.

Geman, S., and Geman, D. (1984). "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images." *Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741.

Geman, S., Bienenstock, E., and Doursat, R. (1992). "Neural networks and the bias/variance dilemma." *Neural Computation*, vol. 4, no. 1, pp. 1–58.

GeNIe and SMILE (2012). "*Graphical Network Interface and Structural Modeling, Inference, and Learning Engine*." http://genie.sis.pitt.edu (Accessed 10 Jan 2012).

Gilks, W. R., Richardson, S., and Spiegelhalter, D. (1995). "*Markov chain monte carlo in practice*." Chapman and Hall/CRC, Boca Raton, USA.

Gill, P. E., Murray, W., and Wright, M. H. (1982). "*Practical optimization*." Academic Press, Bingley, USA.

Giustolisi, O., and Berardi, L. (2007). "Pipe level burst prediction using EPR and MCS-EPR." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Giustolisi, O., and Simeone, V. (2006). "Optimal design of artificial neural networks by a multi-objective strategy: groundwater level predictions." *Hydrological Sciences Journal*, vol. 51, no. 3, pp. 502–523.

Giustolisi, O., Laucelli, D., and Savić, D. A. (2005). "A decision support framework for short-time rehabilitation planning in water distribution systems." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Gleick, P. H. (1995). "*Human population and water: to the limits in the 21st century*." http://www.aaas.org/international/ehn/fisheries/gleick.htm (Accessed 6 Nov 2011).

Glymour, C., Scheines, R., Spirtes, P., and Kelly, K. (1987). "*Discovering causal structure*." Academic Press Inc., London, UK.

Goldberg, D. E. (1989). "*Genetic algorithms in search, optimization, and machine learning*." Addison-Wesley, Reading, USA.

Goldman, R. (1990). "*A probabilistic approach to language understanding*." Technical Report, Department of Computer Science, Brown University, Providence, USA.

Graf, F. L. (1990). "Using ground-penetrating radar to pinpoint pipeline leaks." *Materials Performance*, vol. 29, no. 4, pp. 27–29.

Grant, E. L., and Leavenworth, R. S. (1980). "*Statistical quality control*." 5th edition, McGraw-Hill, New York, USA.

Grewal, M. S., and Andrews, A. P. (2008). "*Kalman filtering: theory and practice using MATLAB*." 3rd edition, John Wiley and Sons, Chichester, UK.

Griffiths, T. L., and Yuille, A. (2006). "A primer on probabilistic inference." *Trends in Cognitive Sciences, Supplement to Special Issue on Probabilistic Models of Cognition*, vol. 10, no. 7, pp. 1–11.

Gutermann (2011). "*Gutermann Australia leak noise correlators*." http://www.gutermann-au.com (Accessed 18 Apr 2011).

Haestad Methods Inc. (2002). "*WaterCAD v5 user's manual*." Waterbury, USA.

Hanley, J. A., and McNeil, B. J. (1982). "The meaning and use of the area under a receiver operating characteristic (ROC) curve." *Radiology*, vol. 143, pp. 29–36.

Hansson, O., and Mayer, A. (1989). "Heuristic search as evidential reasoning." *Proceedings of the 5th Workshop on Uncertainty in Artificial Intelligence*, Mountain View, USA.

Hargesheimer, E. E. (1985). "Identifying water main leaks with trihalomethane tracers." *Journal of American Water Works Association*, vol. 77, no. 11, pp. 71–75.

Harp, S., Samad, T., and Guha, A. (1991). "Towards the genetic synthesis of neural networks." *Proceedings of the 4th International Conference on Genetic Algorithms*, San Diego, USA.

Hart, K. M., and Hart, R. F. (1989). "*Quantitative methods for quality improvement*." American Society for Quality: Quality Press, Milwaukee, USA.

Hartley, J. K., and Bargiela, A. (1995). "Parallel simulation of large scale water distribution system." *Proceedings of the Modelling and Simulation Conference*, Prague, Czech Republic.

Hata, N., Masuda, J., and Takatsuka, T. (1997). "Deep ground-penetrating radar technology for surveying buried objects." *Proceedings of the 15th International Conference No-Dig*, Taipei, Taiwan.

Haykin, S. (1994). "*Neural networks: a comprehensive foundation*." Macmillan, New York, USA.

Hayuti, M., Wheeler, M., Harford, A., and Sage, P. (2008). "Leakage hotspot prediction and water network models." *Proceedings of Water Loss Seminar and Workshop*, Marbella, Spain.

Hebb, D. O. (1949). "*The organization of behavior*." John Wiley and Sons, New York, USA.

Heckerman, D. (1990). "*Probabilistic similarity networks*." Technical Report, Departments of Computer Science and Medicine, Stanford University, Stanford, USA.

Heckerman, D. (1996). "*A tutorial on learning with Bayesian networks*." Technical Report no. MSR-TR-95-06, Microsoft Research, ftp://ftp.research.microsoft.com/pub/tr/tr-95-06.pdf (Accessed 12 May 2011).

Hengl, T. (2009). "*A practical guide to geostatistical mapping*." http://spatial-analyst.net/book/system/files/Hengl_2009_GEOSTATe2c1w.pdf (Accessed 11 Dec 2011)

Hengl, T., Sierdsema, H., Radovic, A., and Dilo, A. (2009). "Spatial prediction of species' distributions from occurrence-only records: combining point pattern analysis, ENFA and regression-kriging." *Ecological Modelling*, vol. 220, no. 24, pp. 3499–3511.

Herdy, M. (1990). "Application of the 'evolutionsstrategie' to discrete optimization problems." in "*Parallel problem solving from nature*." Springer-Verlag, Berlin, Germany.

Holland, J. H. (1962). "Outline for a logical theory of adaptive systems." *Journal of the Association for Computing Machinery*, vol. 9, no. 3, pp. 297–314.

Holland, J. H. (1975). "*Adaptation in natural and artificial systems*." University of Michigan Press, Ann Arbor, USA.

Holnicki-Szulc, J., Kolakowski, P., and Nasher, N. (2005). "Leakage detection in water networks." *Journal of Intelligent Material Systems and Structures*, vol. 16, no. 3, pp. 207–219.

Hornick, K., Maxwell, S., and Halbert, W. (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks*, vol. 2, no. 5, pp. 359–366.

Humphrey, W. S. (1989). "*Managing the software process*." Addison-Wesley, Reading, USA.

Hunaidi, O., and Chu, W. T. (1999). "Acoustical characteristics of leak signals in plastic water distribution pipes." *Applied Acoustics*, vol. 58, no. 3, pp. 235–254.

Hunaidi, O., and Giamou, P. (1998). "GPR for detection of leaks in buried plastic water distribution pipes." *Proceedings of the 7th International Conference on Ground-Penetrating Radar*, Lawrence, USA.

Hunaidi, O., and Wang, A. (2006). "A new system for locating leaks in urban water distribution pipes." *Management of Environmental Quality: An International Journal*, vol. 17, no. 4, pp. 450–466.

Hunaidi, O., Chu, W., Wang, A., and Guan, W. (2000). "Detecting leaks in plastic pipes." *Journal American Water Works Association*, vol. 92, no. 2, pp. 82–94.

Hyndman, R. J., and Koehler, A. B. (2006). "Another look at measures of forecast accuracy." *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688.

Hyun, S. -Y., Jo, Y. -S., Oh, H. -C., and Kim, S. -Y. (2003). "An experimental study on a ground-penetrating radar for detecting water-leaks in buried water transfer pipes." *Proceedings of the 6th International Symposium on Antennas, Propagation and Electromagnetic Theory*, Beijing, China.

Hyun, S. -Y., Jo, Y. -S., Oh, H. -C., Kim, S. -Y., and Kim, Y. -S. (2007). "The laboratory scaled-down model of a ground-penetrating radar for leak detection of water pipes." *Measurement Science and Technology*, vol. 18, no. 9, pp. 2791–2799.

Isaaks, E. H., and Srivastava, R. M. (1989). "*An introduction to applied geostatistics*." Oxford University Press, New York, USA.

IWA (2000). "*Manual of best practice: performance indicators for water supply services*." International Water Association, London, UK.

Izquierdo, J., López, P. A., Martínez, F. J., and Pérez, R. (2007). "Fault detection in water supply systems using hybrid (theory and data-driven) modelling." *Mathematical and Computer Modelling*, vol. 46, no. 3-4, pp. 341–350.

Jackson, J. E. (1991). "*A user's guide to principal components*." John Wiley and Sons, New York, USA.

Jaffard, S., Meyer, Y., and Ryan, R. D. (2001). "*Wavelets: tools for science & technology*." Society for Industrial and Applied Mathematics, New York, USA.

Jain, A., and Ormsbee, L. E. (2002). "Short-term water demand forecast modeling techniques – conventional methods versus AI." *Journal American Water Works Association*, vol. 94, no. 7, pp. 64–72.

Jain, A., Varshney, A. K., and Joshi, U. C. (2001). "Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks." *Water Resources Management*, vol. 15, no. 5, pp. 299–321.

Jansen, M. (2006). "Minimum risk thresholds for data with heavy noise." *Institute of Electrical and Electronics Engineers Signal Processing Letters*, vol. 13, no. 5, pp. 296–299.

Jenks, G. F. (1967). "The data model concept in statistical mapping." *International Yearbook of Cartography*, vol. 7, pp. 186–190.

Jensen, F. V. (2001). "*Bayesian networks and decision graphs*." Springer-Verlag, New York, USA.

Jensen, F. V. (2009). "Bayesian networks." *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 3, pp. 307–315.

Jensen, F. V., Kjaerulff, U. B., Lang, M., and Madsen, A. L. (2002). "HUGIN - the tool for Bayesian networks and influence diagrams." *Proceedings of the 1st European Workshop on Probabilistic Graphical Models*, Cuenca, Spain.

Johnston, K., Ver Hoef, J. M., Krivoruchko, K., and Lucas, N. (2001). "*Using ArcGIS geostatistical analyst*." ESRI, Redlands, USA.

Jolliffe, I. T. (2002). "*Principal component analysis*." 2nd edition, Springer-Verlag, New York, USA.

Jones, G. (1998). "Genetic and evolutionary algorithms." in "*Encyclopaedia of computational chemistry, volume III - databases and expert systems*." John Wiley and Sons, Chichester, UK.

Jonsson, L. (1994). "*Leak detection in pipelines using hydraulic transients - laboratory measurements*." Report to the Carl Trygger Foundation, Department of Water Resources Engineering, University of Lund, Lund, Sweden.

Jonsson, L. (1999). "Hydraulic transients as a monitoring device." *Proceedings of the 28th International Association for Hydro-Environment Engineering and Research Congress*, Graz, Austria.

Jonsson, L. (2001). "Interaction of a hydraulic transient with a leak in a pipe flow." *Proceedings of the 14th Australasian Fluid Mechanics Conference*, Adelaide, Australia.

Jonsson, L., and Larson, M. (1992). "Leak detection through hydraulic transient analysis." *Proceedings of the International Conference on Water Pipeline Systems*, Manchester, UK.

Jordan, M. I. (1986a). "Attractor dynamics and parallelism in a connectionist sequential machine." *Proceedings of the 8th Annual Conference of the Cognitive Science Society*, Hillsdale, USA.

Jordan, M. I. (1986b). "*Serial order: a parallel distributed processing approach*." Technical Report no. 8604, Institute for Cognitive Science, University of California, USA.

Jordan, M. I. (1998). "*Learning in graphical models*." Massachusetts Institute of Technology Press, Cambridge, USA.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., and Saul, L. K. (1998). "An introduction to variational methods for graphical models." in "*Learning in graphical models*." Kluwer Academic Publishers, Dordrecht, Netherlands.

Joshi, A., Udpa, L., Udpa, S., and Tamburrino, A. (2006). "Adaptive wavelets for characterizing magnetic flux leakage signals from pipeline inspection." *Institute of Electrical and Electronics Engineers Transactions on Magnetics*, vol. 42, no. 10, pp. 3168–3170.

Journel, A. G., and Huijbregts, C. J. (1978). "*Mining geostatistics*." Academic Press Inc., London, UK.

Jowitt, P. W., and Xu, C. (1990). "Optimal valve control in water distribution networks." *Journal of Water Resources Planning and Management*, vol. 116, no. 4, pp. 455–472.

Kalman, B. L., and Kwasny, S. C. (1993). "*TRAINREC: a system for training feedforward & simple recurrent networks efficiently and correctly*." http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.5463 (Accessed 21 Mar 2011).

Kapelan, Z., Savić, D. A., and Walters, G. A. (2001). "Use of prior information on parameters in inverse transient analysis for leak detection and roughness calibration." *Proceedings of the World Water and Environmental Resources Congress*, Orlando, USA.

Kapelan, Z., Savić, D. A., and Walters, G. A. (2002). "Hybrid GA for calibration of water distribution models." *Proceedings of the Water Resources Planning and Management Conference*, Roanoke, USA.

Kapelan, Z., Savić, D. A., and Walters, G. A. (2003). "A hybrid inverse transient model for leakage detection and roughness calibration in pipe networks." *Journal of Hydraulic Research*, vol. 41, no. 5, pp. 481–492.

Kapelan, Z., Savić, D. A., and Walters, G. A. (2004). "Incorporation of prior information on parameters in inverse transient analysis for leak detection and roughness calibration." *Urban Water Journal*, vol. 1, no. 2, pp. 129–143.

Kapelan, Z., Savić, D. A., and Walters, G. A. (2005). "Optimal sampling design methodologies for water distribution model calibration." *Journal of Hydraulic Engineering*, vol. 131, no. 3, pp. 190–200.

Karney, B., Khani, D., Halfawy, M. R., and Hunaidi, O. (2008). "A simulation study on using inverse transient analysis for leak detection in water distribution networks." *Proceedings of the International Stormwater and Urban Water Systems Modeling Conference*, Toronto, Canada.

Katul, G. G., and Parlange, M. B. (1995). "Analysis of land surface heat fluxes using the orthonormal wavelet approach." *Water Resources Research*, vol. 31, no. 11, pp. 2743–2749.

Keesing, R., and Stork, D. G. (1991). "Evolution and learning in neural networks: the number and distribution of learning trials affect the rate of evolution." *Advanced in Neural Information Processing Systems*, vol. 3, pp. 805–810.

Keller-Ressel, M. (2012). "*Lyapunov function*." MathWorld - A Wolfram Web Resource, created by Eric W. Weisstein, http://mathworld.wolfram.com/LyapunovFunction.html (Accessed 7 Aug 2012).

Khan, A, Widdop, P. D., Day, A. J., Wood, A. S., Mounce, S. R., and Machell, J. (2002). "Low-cost failure sensor design and development for water pipeline distribution systems." *Water Science and Technology: A Journal of the International Association on Water Pollution Research*, vol. 45, no. 4-5, pp. 207–215.

Khan, A., Widdop, P. D., Day, A. J., Wood, A. S., Mounce, S. R., and Machell, J. (2005). "Performance assessment of leak detection failure sensors used in a water distribution system." *Journal of Water Supply: Research and Technology – AQUA*, vol. 54, no. 1, pp. 25–36.

Khan, A., Widdop, P. D., Day, A. J., Wood, A. S., Mounce, S. R., and Machell, J. (2006). "Artificial neural network model for a low cost failure sensor: performance assessment in pipeline distribution." *Transactions on Engineering, Computing and Technology, ENFORMATIKA*, vol. 15, pp. 195–201.

Kim, J. H., and Pearl, J. (1983). "A computational model for causal and diagnostic reasoning in inference systems." *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Los Altos, USA.

Kim, S. H. (2005). "Extensive development of leak detection algorithm by impulse response method." *Journal of Hydraulic Engineering*, vol. 131, no. 3, pp. 201–208.

Kirkpatrick, S., Gelatt, D., and Vecchi, M. (1983). "Optimization by simulated annealing." *Science*, vol. 220, no. 4598, pp. 671–680.

Kleiner, Y., Rajani, B., and Sadiq, R. (2004). "Modelling failure risk in buried pipes using fuzzy Markov deterioration process." *American Society of Civil Engineers International Conference of Pipeline Engineering and Construction*, San Diego, USA.

Kleiner, Y., Rajani, B., and Sadiq, R. (2005). "Application of a fuzzy Markov model to plan the renewal of large-diameter buried pipes: a case study." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Kleiner, Y., and Rajani, B. (2001). "Comprehensive review of structural deterioration of water mains: statistical models." *Urban Water Journal*, vol. 3, no. 3, pp. 131–150.

Kohonen, T. (1990). "The self-organizing map." *Proceedings of the Institute of Electrical and Electronics Engineers*, vol. 78, no. 9, pp. 1464–1480.

Kohonen, T. (2001). "*Self-organizing maps*." 3rd edition, Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, Berlin, Germany.

Korb, A., and Nicholson, A. (2004). "*Bayesian artificial intelligence*." Chapman and Hall/CRC, Boca Raton, USA.

Koza, J. R. (1992). "*Genetic programming*." Massachusetts Institute of Technology Press, Cambridge, USA.

Krige, D. G. (1951). "A statistical approach to some mine valuations problems at the Witwatersrand." *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, vol. 52, no. 6, pp. 119–139.

Krige, D. G. (1966). "Two-dimensional weighted moving average trend surfaces for ore evaluation." *Journal of South African Institute of Mining and Metallurgy*, vol. 66, no. 1, pp. 13–38.

Kumar, P., and Foufoula-Georgiou, E. A. (1993). "Multicomponent decomposition of spatial rainfall fields 1. Segregation of large and small scale features using wavelet transform." *Water Resources Research*, vol. 29, no. 8, pp. 2515–2532.

Labat, D. (2005). "Recent advances in wavelet analyses: part 1. A review of concepts." *Journal of Hydrology*, vol. 314, no. 1-4, pp. 275–288.

Labat, D., Ababou, R., and Mangin, A. (2000). "Rainfall-runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses." *Journal of Hydrology*, vol. 238, no. 3-4, pp. 149–178.

Lam, N. S. (1983). "Spatial interpolation method: a review." *The American Cartographer*, vol. 10, no. 2, pp. 129–135.

Lam, W., and Bacchus, F. (1994). "Learning Bayesian belief networks: an approach based on the MDL principle." *Computation Intelligence*, vol. 10, no. 4, pp. 269–293.

Lambert, A. (1994). "Accounting for losses – the bursts and background estimates concepts." *Journal of the Institution of Water and Environmental Management*, vol. 8, no. 2, pp. 205–214.

Lambert, A. (2001). "What do we know about pressure/leakage relationships in distribution systems?" *Proceedings of the International Water Association Conference: System Approach to Leakage Control and Water Distribution Systems Management*, Brno, Czech Republic.

Lambert, A., Myers, S. D., and Trow, S. (1998). "*Managing water leakage - economic and technical issues*." Financial Times Energy, London, UK.

Lancaster, P., and Salkauskas, K. (1986). "*Curve and surface fitting: an introduction*." Academic Press Inc., London, UK.

Lange, F. H. (1987). "*Correlation techniques*." Van Nostrand, Englewood Cliffs, USA.

Lauritzen, S. L. (1995). "The EM algorithm for graphical association models with missing data." *Computational Statistics and Data Analysis*, vol. 19, no. 2, pp. 191–201.

Lauritzen, S. L., and Spiegelhalter, D. J. (1988). "Local computations with probabilities on graphical structures and their application to expert systems (with discussion)." *Journal of the Royal Statistical Society*, vol. 50, no. 2, pp. 157–224.

Lawrence, S., Giles, C. L., and Tsoi, A. C. (1996). "*What size neural network gives optimal generalization? Convergence properties of backpropagation*." Technical Report, UMIACS-TR-96-22 and CS-TR-3617, Institute for Advanced Computer Studies, University of Maryland, USA, http://www.neci.nj.nec.com/homepages/lawrence/papers/minima-tr96/minima-tr96.html (Accessed 25 Mar 2011).

Lee, P. J., Lambert, M. F., Simpson, A. R., and Vítkovský, J. P. (2006). "Experimental verification of the frequency response method of leak detection." *Journal of Hydraulic Research*, vol. 44, no. 5, pp. 451–468.

Lee, P. J., Lambert, M. F., Simpson, A. R., Vítkovský, J. P., and Misiunas, D. (2007). "Leak location in single pipelines using transient reflections." *Australian Journal of Water Resources*, vol. 11, no. 1, pp. 53–65.

Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2003). "Frequency response coding for the location of leaks in single pipeline systems." *Proceedings of the International Conference Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, Valencia, Spain.

Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2005a). "Frequency domain analysis for detecting pipeline leaks." *Journal of Hydraulic Engineering*, vol. 131, no. 7, pp. 596–604.

Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2005b). "Leak location using the pattern of the frequency response diagram in pipelines: a numerical study." *Journal of Sound and Vibration*, vol. 284, no. 3-5, pp. 1051–1073.

Lee, P. J., Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2007). "Leak location in pipelines using the impulse response function." *Journal of Hydraulic Research*, vol. 45, no. 5, pp. 643–652.

Lee, P. J., Vítkovský, J. P., Simpson, A. R., Lambert, M. F., and Liggett, J. A. (2002). Discussion of ''Leak detection in pipes by frequency response method using a step excitation." Mpesha, W., Chaudhry, M. H., and Gassman, S. L. (2002). *Journal of Hydraulic Research*, vol. 40, no. 1, pp. 55–62, in: *Journal of Hydraulic Research*, vol. 41, no. 2, pp. 221–223.

Legates, D. R., and McCabe, G. J. (1999). "Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation." *Water Resources Research*, vol. 35, no. 1, pp. 233–241.

Leshno, M., Ya Lin, V., Pinkus, A., and Schocken, S. (1993). "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function." *Neural Networks*, vol. 6, no. 6, pp. 861–867.

Levenberg, K. (1944). "A method for the solution of certain problems in least squares." *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168.

Li, J., and Heap, A. D. (2008). "*A review of spatial interpolation methods for environmental scientists*." Record 2008/23, Geoscience Australia, Canberra, Australia.

Liang, L. Y., Thompson, R. G., and Young, D. M. (2004). "Optimising the design of sewer networks using genetic algorithms and tabu search." *Engineering, Construction and Architectural Management*, vol. 11, no. 2, pp. 101–112.

Licciardi, S. V. (1998). "Quantification and location of leaks by internal inspection." *Tunnelling and Underground Space Technology*, vol. 13, no. 2, pp. 5–15.

Liemberger, R., and Farley, M. (2004). "Developing a non-revenue water reduction strategy: investigating and assessing water losses." *Proceedings of the International Water Association 4th World Water Congress*, Marrakech, Morocco.

Liggett, J. A., and Chen, L. -C. (1994). "Inverse transient analysis in pipe networks." *Journal of Hydraulic Engineering*, vol. 120, no. 8, pp. 934–955.

Liong, S. Y., Lim, W. H., and Paudyal, G. N. (2000). "River stage forecasting in Bangladesh: neural network approach." *Journal of Computations in Civil Engineering*, vol. 8, no. 2, pp. 201–220.

Lockwood, A., Murray, T., Stuart, G., and Scudder, L. (2003). "A study of geophysical methods for water leak location." *Proceedings of Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, Valencia, Spain.

Magini, R., Pallavicini, I., and Verde, D. (2007). "Estimation of leakages in water distribution systems with measurements at few nodes." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." *Environmental Modelling and Software*, vol. 15, no. 1, pp. 101–124.

Mallat, S. (1989). "A theory for multiresolution signal decomposition: the wavelet representation." *Institute of Electrical and Electronics Engineers Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693.

Marquardt, D. (1963). "An algorithm for least-squares estimation of nonlinear parameters." *Society for Industrial and Applied Mathematics Journal on Applied Mathematics*, vol. 11, no. 2, pp. 431–441.

Mashford, J., De Silva, D., Marney, D., and Burn, S. (2009). "An approach to leak detection in pipe networks using analysis of monitored pressure values by support vector machine." *Proceedings of the 3rd International Conference on Network and System Security*, Gold Coast, Australia.

Masters, T. (1995). "*Advanced algorithms for neural networks: a C++ sourcebook*." John Wiley and Sons, New York, USA.

Matheron, G. (1963). "Principles of geostatistics." *Economic Geology*, vol. 58, no. 8, pp. 1246–1266.

Matheron, G. (1971). "*The theory of regionalized variables and its applications*." Cahiers du Centre de Morphologie Mathematique, Ecole des Mines, Fountainebleau, France.

May, J. (1994). "Pressure dependent leakage." *World Water and Environmental Engineering*.

McClelland, J. L., Rumelhart, D. E. and the Parallel Distributed Processing Research Group (1986). "*Parallel distributed processing: explorations in the microstructure of cognition, vol. 2: psychological and biological models*." Massachusetts Institute of Technology Press, Cambridge, USA.

McCulloch, W. S., and Pitts, W. H. (1943). "A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biology*, vol. 5, no. 4, pp. 115–133.

McDonald, S., and Makar, J. (1996). "*Assessment of the Hydroscope 201TM condition index evaluation of grey cast iron pipe from Gatineau, Quebec*." National Research Council, Report no. A-7015.3, Ottawa, Canada.

McNulty, J. G. (2001). "An acoustic-based system for detecting, locating and sizing leaks in water pipelines." *Proceedings of the 4th International Conference on Water Pipeline Systems*, York, UK.

Merelo Guervós, J. J., Patón, M., Cañas, A., Prieto, A., and Morán, F. (1993). "Optimization of a competitive learning neural network by genetic algorithms." in "*New trends in neural computation*." Springer-Verlag Berlin and Heidelberg GmbH & Co. K, Berlin, Germany.

Mergelas, B., and Henrich, G. (2005). "Leak locating method for precommissioned transmission pipelines: North American case studies." *Proceedings of the International Water Association Conference: Leakage 2005*, Halifax, Canada.

Metje, N., Atkins, P., Brennan, M. J., Chapman, D., Lim, H., Machell, J., Muggleton, J. M., Pennock, S., Ratcliffe, J., Redfern, M., Rogers, C., Saul, A., Shan, Q., Swingler, S., and Thomas, A. (2007). "Mapping the underworld - state-of-the-art review." *Tunnelling and Underground Space Technology*, vol. 22, no.5-6, pp. 568–586.

Meyer, Y., and Roques, S. (1993). "*Progress in wavelet analysis and applications*." Frontières editions, Gif-sur-Yvette, France.

Meyer, Y., and Ryan, R. D. (1996). "*Wavelets: algorithms and applications*." Society for Industrial and Applied Mathematics, New York, USA.

Miller, G. F., Todd, P. M., and Hegde, S. U. (1989). "Designing neural networks using genetic algorithms." *Proceedings of the 3rd International Conference on Genetic Algorithms*, Fairfax, USA.

Misiunas, D., Lambert, M. F., Simpson, A. R., and Olsson, G. (2005a). "Burst detection and location in water transmission pipelines." *Proceedings of the World Water and Environmental Resources Congress*, Anchorage, USA.

Misiunas, D., Lambert, M. F., Simpson, A. R., and Olsson, G. (2005c). "Burst detection and location in water distribution networks." *Water Science and Technology: Water Supply*, vol. 5, no. 3-4, pp. 71–80.

Misiunas, D., Vítkovský, J. P., Olsson, G., Lambert, M. F., and Simpson, A. R. (2006). "Failure monitoring in water distribution networks." *Water Science and Technology: A Journal of the International Association on Water Pollution Research*, vol. 53, no. 4-5, pp. 503–511.

Misiunas, D., Vítkovský, J. P., Olsson, G., Simpson, A. R., and Lambert, M. F. (2003). "Pipeline burst detection and location using a continuous monitoring technique." *Proceedings of the 7th International Conference on Computing and Control in the Water Industry*, London, UK.

Misiunas, D., Vítkovský, J. P., Olsson, G., Simpson, A. R., and Lambert, M. F. (2005b). "Pipeline break detection using the transient monitoring." *Journal of Water Resources Planning and Management*, vol. 131, no. 4, pp. 316–325.

Misiunas, D., Vítkovský, J. P., Olsson, G., Simpson, A. R., and Lambert, M. F. (2004). "Burst detection and location in pipe networks using a continuous monitoring technique." *Proceedings of the 9th International Conference on Pressure Surges*, Chester, UK.

Mitchell, M. (1996). "*An introduction to genetic algorithms*." Massachusetts Institute of Technology Press, Cambridge, USA.

Montana, D., and Davis, L. (1989). "Training feedforward neural networks using genetic algorithms." *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Detroit, USA.

Moody, J. E. (1992). "The effective number of parameters: an analysis of generalization and regularization in nonlinear learning systems." in "*Advances in neural information processing systems: vol. 4*." Morgan Kaufmann, San Francisco, USA.

Moré, J. J. (1977). "The Levenberg-Marquardt algorithm: implementation and theory." in "*Numerical analysis, lecture notes in mathematics*." Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, Berlin, Germany.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Transactions of the American Society of Agricultural and Biological Engineers*, vol. 50, no. 3, pp. 885–900.

Mounce, S. R., and Boxall, J. B. (2010). "Implementation of an on-line artificial intelligence district meter area flow meter data analysis system for abnormality detection: a case study." *Water Science & Technology: Water Supply*, vol. 10, no. 3, pp. 437–444.

Mounce, S. R., and Machell, J. (2006). "Burst detection using hydraulic data from water distribution systems with artificial neural networks." *Urban Water Journal*, vol. 3, no. 1, pp. 21–31.

Mounce, S. R., Boxall, J. B., and Machell, J. (2007). "An artificial neural network/fuzzy logic system for DMA flow meter data analysis providing burst identification and size estimation." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Mounce, S. R., Boxall, J. B., and Machell, J. (2008). "Online application of ANN and fuzzy logic system for burst detection." *Proceedings of the 10th Annual International Symposium on Water Distribution Systems Analysis*, Kruger National Park, South Africa.

Mounce, S. R., Boxall, J. B., and Machell, J. (2010a). "Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows." *Journal of Water Resources Planning and Management*, vol. 136, no. 3, pp. 309–318.

Mounce, S. R., Day, A. J., Wood, A. S., Khan, A., Widdop, P. D., and Machell, J. (2002). "A neural network approach to burst detection." *Water Science and Technology: A Journal of the International Association on Water Pollution Research*, vol. 45, no. 4-5, pp. 237–246.

Mounce, S. R., Farley, B., Mounce, R. B., and Boxall, J. B. (2010c). "Field testing of optimal sensor placement and data analysis methodologies for burst detection and location in an urban water network." *Proceedings of the 9th International Conference on Hydroinformatics*, Tianjin, China.

Mounce, S. R., Machell, J., and Boxall, J. B. (2006). "Development of artificial intelligence systems for analysis of water supply system data." *Proceedings of the 8th Annual International Symposium on Water Distribution Systems Analysis*, Cincinnati, USA.

Mounce, S. R., Mounce, R. B., and Boxall, J. B. (2011a). "Novelty detection for time series data analysis in water distribution systems using support vector machines." *Journal of Hydroinformatics*, vol. 13, no. 4, pp. 672–686.

Mounce, S. R., Mounce, R. B., and Boxall, J. B. (2012). "Identifying sampling interval for event detection in water distribution networks." *Journal of Water Resources Planning and Management*, vol. 138, no. 2, pp. 187–191.

Mpesha, W., Chaudhry, M. H., and Gassman, S. L. (2002). "Leak detection in pipes by frequency response method using a step excitation." *Journal of Hydraulic Research*, vol. 40, no. 1, pp. 55–62.

Mpesha, W., Gassman, S. L., and Chaudhry, M. H. (2001). "Leak detection in pipes by frequency response method." *Journal of Hydraulic Engineering*, vol. 127, no. 2, pp. 134–147.

Mukherjee, J., and Narasimhan, S. (1996). "Leak detection in networks of pipelines by generalized likelihood ratio method." *Industrial and Engineering Chemistry Research*, vol. 35, no. 6, pp. 1886–1893.

Mukhopadahyay, S., and Srivastava, G. P. (2000). "Characterisation of metal loss defects from magnetic flux leakage signals with discrete wavelet transform." *Non Destructive Testing and Evaluation International*, vol. 33, no. 1, pp. 57–65.

Murphy, A. H. (1995). "The coefficients of correlation and determination as measures of performance in forecast verification." *Weather and Forecasting*, vol. 10, pp. 681–688.

Murphy, K. P. (2001). "The Bayes net toolbox for MATLAB." *Computing Science and Statistics*, vol. 33, pp. 331–350.

Murray, S., Ghazali, M., and McBean, E. A. (2011). "Real-time water quality monitoring: assessment of multisensor data using Bayesian belief networks." *Journal of Water Resources Planning and Management*, vol. 138, no. 1, pp. 63–70.

Murthy, Z. V. P., and Vengal, J. C. (2006). "Optimization of a reverse osmosis system using genetic algorithm." *Separation Science and Technology*, vol. 41, no. 4, pp. 647–663.

Myers, D. E. (1982). "Matrix formulation of cokriging." *Mathematical Geology*, vol. 14, no. 3, pp. 249–257.

Nadi, F., Agogino, A., and Hodges, D. (1991). "Use of influence diagrams and neural networks in modeling semiconductor manufacturing processes." *Institute of Electrical and Electronics Engineers Transactions on Semiconductor Manufacturing*, vol. 4, no. 1, pp. 52–58.

Nakhkash, M., and Mahmood-Zadeh, M. (2004). "Water leak detection using ground penetrating radar." *Proceedings of the 10th International Conference on Ground Penetrating Radar*, Delft, Netherlands.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models, part I - a discussion of principles." *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290.

Nason, G. P., and Sachs, R. V. (1999). "Wavelets in time series analysis." *Philosophical Transaction of the Royal Society*, vol. 357, pp. 2511–2526.

Neapolitan, R. (1990). "*Probabilistic reasoning in expert systems: theory and algorithms*." John Wiley and Sons, New York, USA.

Nelson, M. C., and Illingworth, W. T. (1990). "*A practical guide to neural nets*." Addison-Wesley, Harlow, UK.

Neuroshell2 manual (1996). "*Ward systems group inc., artificial intelligence software for science and business*." http://www.wardsystems.com/manuals/neuroshell2/ (Accessed 31 May 2012).

Nicklow, J., Reed, P., Savić, D. A., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., and Zechman, E. (2010). "State of the art for genetic algorithms and beyond in water resources planning and management." *Journal of Water Resources Planning and Management*, vol. 136, no. 4, pp. 412–432.

O'Brien, E., Murray, T., and McDonald, A. (2003). "Detecting leaks from water pipes at a test facility using ground penetrating radar." *Proceedings of Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, Valencia, Spain.

Page, E. S. (1955). "Control charts with warning lines." *Biometrics*, vol. 42, no. 1-2, pp. 243–257.

Palau, C. V., Arregui, F. J., and Carlos, M. (2012). "Burst detection in water networks using principal component analysis." *Journal of Water Resources Planning and Management*, vol. 138, no. 1, pp. 47–54.

Palau, C. V., Arregui, F. J., and Ferrer, A. (2003). "Using multivariate principal component analysis of injected water flows to detect anomalous behaviors in a water supply system. A case study." *Proceedings of the 2nd International Conference on Efficient Use and Management of Water in Urban Areas*, Tenerife, Spain.

Parzen, E. (1962). "On estimation of a probability density function and mode." *Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076.

Pearl, J. (1987). "Evidential reasoning using stochastic simulation of causal models." *Artificial Intelligence*, vol. 32, no. 2, pp. 245–258.

Pearl, J. (1988). "*Probabilistic reasoning in intelligent systems*." Morgan Kaufmann, San Francisco, USA.

Pearl, J. (2000). "*Causality: models, reasoning, and inference*." Cambridge University Press, New York, USA.

Percival, D. B. (2000). "*Wavelet methods for time series analysis*." Cambridge University Press, Cambridge, UK.

Perelman, L., and Ostfeld, A. (2010). "Bayesian networks for estimating contaminant source and propagation in a water distribution system using cluster structure." *Proceedings of the 12th Annual International Symposium on Water Distribution Systems Analysis*, Tucson, USA.

Pilcher, R. (2002). "The benefits of effective maintenance and improvement of district meter areas." *Proceedings of the International Water Association Conference: Leakage Management-A Practical Approach*, Lemesos, Cyprus.

Pilcher, R., Hamilton, S., Chapman, H., Ristovski, B., and Strapely, S. (2007). "Leak location and repair guidance notes." *Proceedings of the International Water Association Conference: Water Loss 2007*, Bucharest, Romania.

Pourret, O., Naim, P., and Marcot, B. (2008). "*Bayesian networks: a practical guide to applications*." John Wiley and Sons, Chichester, UK.

Powell, R. S. (1992). "*On-line monitoring for operational control of water distribution networks*." Thesis (PhD), University of Durham, Durham, UK.

Pudar, R. S., and Liggett, J. A. (1992). "Leaks in pipe networks." *Journal of Hydraulic Engineering*, vol. 118, no. 7, pp. 1031–1046.

Puust, R., Kapelan, Z., Savić, D. A., and Koppel, T. (2006). "Probabilistic leak detection in pipe networks using the SCEM-UA algorithm." *Proceedings of the 8th Annual International Symposium on Water Distribution Systems Analysis*, Cincinnati, USA.

Puust, R., Kapelan, Z., Savić, D. A., and Koppel, T. (2010). "A review of methods for leakage management in pipe networks." *Urban Water Journal*, vol. 7, no. 25, pp. 25–45.

Pyzdek, T. (1989). "*What every engineer should know about quality control*." Marcel Dekker, New York, USA.

Ramos, H., Reis, C., Ferreira, C., Falcao, C., and Covas, D. (2001). "Leakage control policy within operating management tools." *Proceedings of the 6th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Rechenberg, I. (1973). "*Evolutionsstrategie: optimierung technischer systeme nach prinzipien der biologischen evolution*." Frommann-Holzboog, Stuttgart, Germany.

Rechenberg, I. (1978). "Evolutionsstrategien." in "*Simulationsmethoden in der medizin und biologie*." Springer-Verlag, Berlin, Germany.

Redner, R., and Walker, H. (1984). "Mixture densities, maximum likelihood and the EM algorithm." *Society for Industrial and Applied Mathematics Review*, vol. 26, no. 2, pp. 195–239.

Reeves, C. R., and Taylor, S. J. (1998). "Selection of training data for neural networks by a genetic algorithm." in "*Parallel problem solving from nature*." Springer-Verlag Berlin and Heidelberg GmbH & Co. K, Berlin, Germany.

Reis, F. R., and Chaudhry, F. H. (1999). "Hydraulic characteristics of pressure reducing valves for maximum reduction of leakage in water supply networks." *Proceedings of the 5th International Conference on Computing and Control in the Water Industry*, London, UK.

Reis, F. R., Porto, R. M., and Chaudhry, F. H. (1997). "Optimal location of control valves in pipe networks by genetic algorithm." *Journal of Water Resources Planning and Management*, vol. 123, no. 6, pp. 317–326.

Roberts, L. (2005). "*SPC for right-brain thinkers: process control for non-statisticians*." American Society for Quality: Quality Press, Milwaukee, USA.

Rosenblatt, F. (1958). "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological Review*, vol. 65, no. 6, pp. 386–408.

Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function." *Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837.

Rossman, L. A. (2000). "*EPANET2 User's Manual*." United States Environmental Protection Agency, Cincinnati, USA.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning internal representations by error propagation." in "*Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*." Massachusetts Institute of Technology Press, Cambridge, USA.

Rumelhart, D. E., McClelland, J. L., and the Parallel Distributed Processing Research Group (1986). "*Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations.*" Massachusetts Institute of Technology Press, Cambridge, USA.

Ruskai, M. B. (1992). "*Wavelets and their applications*." Jones and Bartlett Publishers, Sudbury, USA.

Russell, S., and Norvig, P. (2010). "*Artificial intelligence: International Version: a modern approach*." 3rd edition, Pearson, USA.

Sage, P. (2005). "Developments in use of network models for leakage management at United Utilities North West." *Proceedings of the Chartered Institution of Water and Environmental Management North West and North Wales Branch Water Treatment and Distribution Conference*, London, UK.

Saldarriaga, J. G., Araque Fuentes, D. A., and Castaneda Galvis, L. F. (2006). "Implementation of the hydraulic transient and steady oscillatory flow with genetic algorithms for leakage detection in real water distribution networks."

Proceedings of the 8th Annual International Symposium on Water Distribution Systems Analysis, Cincinnati, USA.

Sánchez, E. H., Ibáñez, J. C., and Cubillo, F. (2005). "Testing applicability and cost effectiveness of permanent acoustic leakage monitoring for loss management in Madrid distribution network." *Proceedings of the International Water Association Conference: Leakage 2005*, Halifax, Canada.

Sang, Y. F., Wang, D., and Wu, J. C. (2010a). "Entropy-based method of choosing the decomposition level in wavelet threshold de-noising." *Entropy*, vol. 12, no. 6, pp. 1499–1513.

Sang, Y. F., Wang, D., and Wu, J. C. (2010b). "Uncertainty analysis of decomposition level choice in wavelet threshold de-noising." *Entropy*, vol. 12, no. 12, pp. 2386–2396.

Sang, Y. F., Wang, D., Wu, J. C., Zhu, Q. P., and Wang, L. (2009). "Entropy-based wavelet de-noising method for time series analysis." *Entropy*, vol. 11, no. 4, pp. 1123–1147.

Savić, D. A., Boxall, J. B., Ulanicki, B., Kapelan, Z., Makropoulos, C., Fenner, R., Soga, K., Marshall, I. W., Maksimovic, C., Postlethwaite, I., Ashley, R., and Graham, N. (2008). "Project Neptune: improved operation of water distribution networks." *Proceedings of the 10th Annual International Symposium on Water Distribution Systems Analysis*, Kruger National Park, South Africa.

Savić, D. A., and Walters, G. (1997). "Genetic algorithms for least cost design of water distribution networks." *Journal of Water Resources Planning and Management*, vol. 123, no. 2, pp. 67–77.

Schaefli, B., Maraun, D., and Holschneider, M. (2007). "What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology." *Advances in Water Resources*, vol. 30, no.12, pp. 2511–2525.

Schwefel, H. -P. (1981). "*Numerical optimization for computer models*." John Wiley and Sons, Chichester, UK.

Schwefel, H. -P. (1995). "*Evolution and optimum seeking*." John Wiley and Sons, New York, USA.

Seaford, H. (1994). "Acoustic leak detection through advanced signal-processing technology." *Noise and Vibration Worldwide*, vol. 25, no. 5, pp. 17–18.

Shaw Cole, E. (1979). "Methods of leak detection: an overview." *Journal of American Water Works Association*, vol. 71, no. 2, pp. 73–75.

Sheng, Y. (1996). "Wavelet transform." in "*The transforms and applications handbook*." 2nd edition, CRC Press, Bosa Roca, USA.

Shepard, D. (1968). "A two-dimensional interpolation function for irregularly-spaced data." *Proceedings of the 23rd Association for Computing Machinery National Conference*, New York, USA.

Shewhart, W. (1931). "*Economic control of quality of manufactured product*." Van Nostrand Reinhold, New York, USA.

Shihab, K. (2005). "Modeling groundwater quality with Bayesian techniques." *Proceedings of the 5th International Conference on Intelligent Systems Design and Applications*, Wrocław, Poland.

Shinozuka, M., Liang, J., and Feng, M. Q. (2005). "Use of supervisory control and data acquisition for damage location of water delivery systems." *Journal of Engineering Mechanics*, vol. 131, no. 3, pp. 225–230.

Silva, R. A., Buiatti, C. M., Cruz, S. L., and Pereira, J. A. F. R. (1996). "Pressure wave behaviour and leak detection in pipelines." *Proceedings of the 6th European Symposium on Computers and Chemical Engineering*, Rhodes, Greece.

Sipper, M. (2002). "*Machine nature: the coming age of bio-inspired computing*." McGraw-Hill, New York, USA.

Skipworth, P. J., Saul, A., and Engelhardt, M. (2000). "Distribution network behaviour - extracting knowledge from data." *Proceedings of Water Network Modelling for Optimal Design and Management*, Exeter, UK.

Smeti, E. M., Thanasoulias, N. C., Kousouris, L. P., and Tzoumerkas, P. C. (2007). "An approach for the application of statistical process control techniques for quality improvement of treated water." *Desalination*, vol. 213, no. 1-3, pp. 273–281.

SOPAC (2011). "*Water, sanitation and hygiene: pipeline and asset management*." Pacific Islands Applied Geosciences Commission, http://www.pacificwater.org/pages.cfm/water-services/water-demand-management/what-water-demand-management/four-methods-of-real-loss-management/pipeline-asset-management.html (Accessed 17 Jan 2011).

South East Water (2005). "Leak detection gets smart." *Institution of Water Officers Journal*, no. 26.

Spiegelhalter, D., Franklin, R., and Bull, K. (1989). "Assessment criticism and improvement of imprecise subjective probabilities for a medical expert system." *Proceedings of the 5th Workshop on Uncertainty in Artificial Intelligence*, Mountain View, USA.

Srirangarajan, S., Allen, M., Preis, A., Iqbal, M., Lim, H. B., and Whittle, A. J. (2010a). "Water main burst event detection and localization." *Proceedings of the 12th Annual International Symposium on Water Distribution Systems Analysis*, Tucson, USA.

Srirangarajan, S., Allen, M., Preis, A., Iqbal, M., Lim, H. B., and Whittle, A. J. (2012). "Wavelet-based burst event detection and localization in water distribution systems." *Journal of Signal Processing Systems*, pp. 1–16.

Srirangarajan, S., Iqbal, M., Allen, M., Preis, A., Fu, C., Girod, L., Wong, K. -J., Lim, H. B., and Whittle, A. J. (2010b). "Real-time burst event detection in water distribution systems." *Proceedings of the 9th Association for Computing Machinery/Institute of Electrical and Electronics Engineers International Conference on Information Processing in Sensor Networks*, New York, USA.

Stampolidis, A., Soupios, P., Vallianatos, F., and Tsokas, G. N. (2003). "Detection of leaks in buried plastic water distribution pipes in urban places - a case study." *Proceedings of the 2nd International Workshop on Advanced Ground Penetrating Radar*, Delft, Netherlands.

Stein, E. M., and Shakarchi, R. (2003). "*Fourier analysis: an introduction*." Princeton University Press, Princeton, USA.

Stenberg, R. (1982). "*Leak detection in water supply systems*." The Swedish Water and Waste Water Works Association, Stockholm, Sweden.

Stephens, M. L., Lambert, M. F., Simpson, A. R., Vítkovský, J. P., and Nixon, J. B. (2004). "Field tests for leakage, air pocket, and discrete blockage detection using inverse transient analysis in water distribution pipes." *Proceedings of the World Water and Environmental Resources Congress*, Salt Lake City, USA.

Stephens, M. L., Misiunas, D., Lambert, M. F., Simpson, A. R., Vítkovský, J. P., and Nixon, J. B. (2005b). "Field verification of a continuous transient monitoring system for burst detection in water distribution systems." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Stephens, M. L., Simpson, A. R., Lambert, M. F., and Vítkovský, J. P. (2005a). "Field measurements of unsteady friction effects in a trunk transmission pipeline." *Proceedings of the World Water and Environmental Resources Congress*, Anchorage, USA.

Sterling, M., and Bargiela, A. (1984). "Minimum norm state estimation for computer control of water distribution systems." *Institution of Electrical Engineers Proceedings F*, vol. 131, no. 2, pp. 57–63.

Stiber, N. A., Pantazidou, M., and Small, M. J. (1999). "Expert system methodology for evaluating reductive dechlorination at TCE sites." *Environmental Science and Technology*, vol. 33, no. 17, pp. 3012–3020.

Stiber, N. A., Small M. J., and Pantazidou, M. (2004). "Site-specific updating and aggregation of Bayesian belief network models for multiple experts." *Risk Analysis*, vol. 24, no. 6, pp. 1529–1538.

Stow, C. A., Roessler, C., Borsuk, M. E., Bowen, J. D., and Reckhow, K. H. (2003). "Comparison of estuarine water quality models for total maximum daily load

development in Neuse river estuary." *Journal of Water Resources Planning and Management*, vol. 129, no. 4, pp. 307–314.

Sturm, R., and Thornton, J. (2005). "Proactive leakage management using district metered areas (DMA) and pressure management – is it applicable in North America?" *Proceedings of the International Water Association Conference: Leakage 2005*, Halifax, Canada.

Sumer, D., Gonzalez, J., and Lansey, K. (2007). "Real-time detection of sanitary sewer overflows using neural networks and time series analysis." *Journal of Environmental Engineering*, vol. 133, no. 4, pp. 353–363.

Suo, L., and Wylie, E. B. (1989). "Impulse response method for frequency-dependent pipeline transients." *Journal of Fluids Engineering*, vol., 111, no. 4, pp. 478–483.

Suzuki, J. (1978). "Learning Bayesian belief networks based on the MDL principle: an efficient algorithm using the branch and bound technique." *Annals of Statistics*, vol. 6, pp. 461–464.

Swets, J. (1988). "Measuring the accuracy of diagnostic systems." *Science*, vol. 240, no. 4857, pp. 1285–1293.

Swift, J. A. (1995). "*Introduction to modern statistical quality control and management*." St. Lucie Press, Delray Beach, USA.

Tabesh, M., Asadiani Yekta, A. H., and Burrows, R. (2005). "Evaluation of unaccounted for water and real losses in water distribution networks by hydraulic analysis of the system considering pressure dependency of leakage." *Proceedings of the 8th International Conference on Computing and Control in the Water Industry*, Exeter, UK.

Taghvaei, M., Beck, S. B. M., and Staszewski, W. J. (2006). "Leak detection in pipelines using cepstrum analysis." *Measurement Science and Technology*, vol. 17, no. 2, pp. 367–372.

Takagi, H. (1997). "Introduction to fuzzy systems, neural networks, and genetic algorithms." in "*Intelligent systems: fuzzy logic, neural networks, and genetic algorithms*." Kluwer Academic Publishers, Norwell, USA.

Tang, K. W., Brunone, B., Karney, B. W., and Rossetti, A. (2000). "Role and characterization of leaks under transient conditions." *Proceedings of the Joint Conference on Water Resources Engineering and Water Resources Planning and Management*, Minneapolis, USA.

Tang, K. W., Karney, B. W., and Brunone, B. (2001). "Leak detection using inverse transient calibration and GA: some early successes and future challenges." *Proceedings of the 6th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Technolog (2011). "*Technolog limited, engineering solutions for the utilities*."

http://www.technolog.com/default.aspx (Accessed 31 May 2012).

Thiesson, B. (1995). "Accelerated quantification of Bayesian networks with incomplete data." *Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining*, Montreal, Canada.

Thomas, A. M., Metje, N., Rogers, C. D. F., and Chapman, D. N. (2006). "GPR interpretation as a function of soil response complexity in utility mapping." *Proceedings of the 11th International Conference on Ground Penetrating Radar*, Columbus, USA.

Thornton, J. (2002). "*Water loss control manual*." McGraw-Hill, New York, USA.

Thornton, J., Sturm, R., and Kunkel, G. (2008). "*Water loss control*." 2nd edition, McGraw-Hill, New York, USA.

Tobler, W. (1970). "A computer movie simulating urban growth in the Detroit region." *Economic Geography*, vol. 46, no. 2, pp. 234–240.

Torrence, C., and Compo, G. P. A. (1998). "Practical guide to wavelet analysis." *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61–78.

Trenchless Technology Network (2002). "*Underground mapping, pipeline location technology and condition assessment*." Infrastructure Engineering and Management Research Centre, University of Birmingham, Birmingham, UK.

Tucciarelli, T., Criminisi, A., and Termini, D. (1999). "Leak analysis systems by means of optimal valve regulation." *Journal of Hydraulic Engineering*, vol. 125, no. 3, pp. 277–285.

Tukey, J. W. (1977). "*Exploratory data analysis*." Addison-Wesley, Reading, USA.

TWGWW (1980). "*Leakage control policy and practice*." Technical Working Group on Waste of Water, Report no. 26, National Water Council, Department of the Environment, UK.

UKWIR (1994). "*Managing leakage series of reports: A. Summary report - B. Reporting comparative leakage performance - C. Setting economic leakage targets - D. Estimating unmeasured water delivered - E. Interpreting measured night flows - F: Using night flow data - G. Managing water pressure - H. Dealing with customers' leakage - J. Leakage management techniques, technology and training*." United Kingdom Water Industry Research, London, UK.

UKWIR (1999a). "*A manual of DMA practice*." United Kingdom Water Industry Research, London, UK.

UKWIR (1999b). "*Leakage estimation from night flow analysis*." United Kingdom Water Industry Research, London, UK.

Ulanicka, K., Bounds, P., Ulanicki, B., and Rance, J. (2001). "Pressure control of a large scale water distribution network with interacting water sources: a case study." *Proceedings of the 6th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

USGS (2011). "*Earth's water distribution*." United States Geological Survey, http://ga.water.usgs.gov/edu/waterdistribution.html (Accessed 11 Dec 2011).

Vairavamoorthy, K., and Lumbers, J. (1998). "Leakage reduction in water distribution systems: optimal valve control." *Journal of Hydraulic Engineering*, vol. 124, no. 11, pp. 1146–1154.

Valdés, J. L., Castelló, J. (2003). "The management of distribution networks by sectors and the Barcelona experience." *Proceedings of the 2nd International Conference on Efficient Use and Management of Urban Water Supply*, Tenerife, Spain.

Van Der Kleij, F. C., and Stephenson, M. J. (2002). "Acoustic logging - the Bristol water experience." *Proceedings of the International Water Association Conference: Leakage Management-A Practical Approach*, Lemesos, Cyprus.

Vapnik, V. N. (1998). "*Statistical learning theory*." John Wiley and Sons, New York USA.

Vidakovic, B. (1999). "*Statistical modeling by wavelets*." John Wiley and Sons, New York, USA.

Vision (2005). "*Technology news from Primayer*." Issue no. 7, http://www.primayer.co.uk/downloads/PDF_English/newsletters/VISION_7.pdf (Accessed 4 Aug 2012).

Vítkovský, J. P., and Simpson, A. R. (1997). "*Calibration and leak detection in pipe networks using inverse transient analysis and genetic algorithms*." Research Report no. R157, University of Adelaide, Adelaide, Australia.

Vítkovský, J. P., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2007). "Experimental observation and analysis of inverse transients for pipeline leak detection." *Journal of Water Resources Planning and Management*, vol. 133, no. 6, pp. 519–530.

Vítkovský, J. P., Lee, P. J., Stephens, M. L., Lambert, M. F., Simpson, A. R., and Liggett, J. A. (2003). "Leak and blockage detection in pipelines via an impulse response method." *Proceedings of the International Conference on Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, Valencia, Spain.

Vítkovský, J. P., Simpson, A. R., and Lambert, M. F. (1999). "Leak detection and calibration of water distribution system using transient and genetic algorithms." *Proceedings of the 26th Annual Water Resources Planning and Management Conference*, Tempe, USA.

Vítkovský, J. P., Simpson, A. R., and Lambert, M. F. (2000). "Leak detection and calibration using transients and genetic algorithms." *Journal of Water Resources Planning and Management*, vol. 126, no. 4, pp. 262–265.

Vítkovský, J. P., Simpson, A. R., and Lambert, M. F. (2002). "Minimization algorithms and experimental inverse transient leak detection." *Proceedings of the Water Resources Planning and Management Conference*, Roanoke, USA.

Vítkovský, J. P., Simpson, A. R., Lambert, M. F., and Wang, X. J. (2001). "An experimental verification of the inverse transient technique." *Proceedings of the 6th Conference on Hydraulics in Civil Engineering*, Hobart, Australia.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S. (2003). "A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters." *Water Resources Research*, vol. 39, no. 8, article no. 1201.

Walters, G. A., and Smith, D. K. (1995). "Evolutionary design algorithm for optimal layout of tree networks." *Engineering Optimization*, vol. 24, no. 4, pp. 261–281.

Wang, X. J. (2002). "*Leakage and blockage detection in pipelines and pipe network systems using fluid transients*." Thesis (PhD), University of Adelaide, Adelaide, Australia.

Wang, X. J., Lambert, M. F., Simpson, A. R., and Vítkovský, J. P. (2001). "Leak detection in pipelines and pipe networks: a review." *Proceedings of the 6th Conference on Hydraulics in Civil Engineering*, Hobart, Australia.

Webster, R., and Oliver, M. (2001). "*Geostatistics for environmental scientists*." John Wiley and Sons, Chichester, UK.

Weigend, A. (1994). "On overfitting and the effective number of hidden units." in "*Proceedings of the 1993 Connectionist Models Summer School."* Lawrence Erlbaum Associates Inc., Mahwah, USA.

Weigend, A., and Rumelhart, D. E. (1992). "Predicting sunspots and exchange rates with connectionist network." *Proceedings of the Workshop on Non-Linear Modelling and Forecasting*, Santa Fe, USA.

Weil, G. J. (1993). "Non contact, remote sensing of buried water pipeline leaks using infrared thermography." *Proceedings of the 20th Anniversary Conference of the American Society of Civil Engineers' Water Resources Planning and Management Division,* Seattle, USA.

Weil, G. J., Graf, R. J., and Forister, L. M. (1994). "Remote sensing pipeline rehabilitation methodologies based upon the utilization of infrared thermography." *Urban Drainage Rehabilitation Programs and Techniques, American Society of Civil Engineers,* New York, USA.

Western Electric Company (1958). "*Statistical quality control handbook*." 2nd edition, AT&T Technologies, Indianapolis, USA.

Whaley, R. S., Nicolas, R. E., and Reet, J. V. (1992). "*A tutorial on software based leak detection methods*." Technical Report, Pipeline Simulation Interest Group, http://env1.kangwon.ac.kr/leakage/2010/knowledge/papers/techniques/9204.pdf (Accessed 17 May 2011).

Wheeler, D. J. (1983). "Detecting a shift in process average: tables of the power function for X charts." *Journal of Quality Technology*, vol. 15, no. 4, pp. 155–170.

Willems, H., and Barbian, O. A. (1998). "Operational experience with inline ultrasonic crack inspection of German crude oil pipelines." *Proceedings of the 7th European Conference on Non Destructive Testing*, Copenhagen, Denmark.

Wirahadikusumah, R., Abraham, D. M., Iseley, T., and Prasanth, R. K. (1998). "Assessment technologies for sewer system rehabilitation." *Automation in Construction*, vol. 7, no.4, pp. 259–270.

Wu, Z. Y., Farley, M., Turtle, D., Dahasahasra, S., Mulay, M., Boxall, J. B., Mounce, S. R., Kleiner, Y., and Kapelan, Z. (2011). "*Water loss reduction*." Bentley Institute Press, Exton, USA.

Wu, Z. Y., and Sage, P. (2006). "Water loss detection via genetic algorithm optimization-based model calibration." *Proceedings of the 8th Annual International Symposium on Water Distribution Systems Analysis*, Cincinnati, USA.

Wu, Z. Y., and Sage, P. (2007). "Pressure dependent demand optimization for leakage detection in water distribution systems." *Proceedings of the 9th International Conference on Computing and Control in the Water Industry*, Leicester, UK.

Wu, Z. Y., and Walski, T. (2005). "Self-adaptive penalty approach compared with other constraint-handling techniques for pipeline optimization." *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 181–192.

Wu, Z. Y., Sage, P., and Turtle, D. (2010). "Pressure-dependent leak detection model and its application to a district water system." *Journal of Water Resources Planning and Management*, vol. 136, no. 1, pp. 116–128.

Wu, Z. Y., Sage, P., Turtle, D., Wheeler, M., Hayuti, M., Velickov, S., Gomez, C., and Hartshorn, J. (2008). "Leak detection case study by means of optimizing emitter locations and flows." *Proceedings of the 10th Annual International Symposium on Water Distribution Systems Analysis*, Kruger National Park, South Africa.

Yao, X. (1999). "Evolving artificial neural networks." *Proceedings of the Institute of Electrical and Electronics Engineers*, vol. 87, no. 9, pp. 1423–1447.

Zadeh, L. A., Fu, K. S., Tanaka, K., and Shimura, M. (1975). "*Fuzzy sets and their applications to cognitive and decision processes*." Academic Press, New York, USA.

Zeng, X., and McMechan, G. A. (1997). “GPR characterization of buried tanks and pipes.” *Geophysics*, vol. 62, no. 3, pp. 797–806.

Zhang, N. L. (1996). “Irrelevance and parameter learning in Bayesian networks.” *Artificial Intelligence*, vol. 88, no. 1-2, pp. 359–373.

Zhang, N. L., and Poole, D. (1996). “Exploiting causal independence in Bayesian network inference.” *Journal of Artificial Intelligence Research*, vol. 5, pp. 301–328.

Zhou, F., Huai-Cheng, G., Yun-Shan, H., and Chao-Zhong, W. (2007). “Scientometric analysis of geostatistics using multivariate methods.” *Scientometrics*, vol. 73, no. 3, pp. 265–279.