1918
TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY

RAIDO PUUST Probabilistic Leak Detection in Pipe Networks Using the SCEM-UA Algorithm

Probabilistic Leak Detection in Pipe Networks Using the SCEM-UA Algorithm

RAIDO PUUST

TALLINN 2007

THESIS ON CIVIL ENGINEERING F12

Probabilistic Leak Detection in Pipe Networks Using the SCEM-UA Algorithm

RAIDO PUUST

TUT
PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Civil Engineering
Department of Mechanics

Dissertation was accepted for the defence of the degree of Doctor of Philosophy on September 24, 2007.

Supervisors: Prof Tiit Koppel, Faculty of Civil Engineering.

Prof Dragan Savic, Centre for Water Systems, School of Engineering, Computer Science and Mathematics University of Exeter, UK.

Dr Zoran Kapelan, Centre for Water Systems, School of Engineering, Computer Science and Mathematics University of Exeter, UK.

Opponents: Prof Orazio Giustolisi, Engineering Faculty of Taranto, Technical University of Bari, Italy.

Dr Olivier Piller, Networks, Water Treatment and Water Quality Research Unit, Cemagref Bordeaux, France.

Defense of the thesis: November 01, 2007.

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for any degree or examination.

/Raido Puust /

ISSN 1406-4766
ISBN 978-9985-59-728-6

EHITUS F12

Lekete tõenäosuse määramine torustikes algoritmi SCEM-UA abil

RAIDO PUUST

TTÜ
KIRJASTUS

Dissertation was accepted for the commencement of the degree of Doctor of Philosophy on Septmeber …, 2007.

Juhendajad: Prof Tiit Koppel, Ehitusteaduskond.

Prof Dragan Savic, Centre for Water Systems, School of Engineering, Computer Science and Mathematics University of Exeter, UK.

Dr Zoran Kapelan, Centre for Water Systems, School of Engineering, Computer Science and Mathematics University of Exeter, UK.

Oponendid: Prof Orazio Giustolisi, Engineering Faculty of Taranto, Technical University of Bari, Italy.

Dr Olivier Piller, Networks, Water Treatment and Water Quality Research Unit, Cemagref Bordeaux, France.

Kaitsmise aeg: November 01, 2007.

Deklareerin, et käesolev doktoritöö, mis on minu iseseisva töö tulemus, on esitatud Tallinna Tehnikaülikooli doktorikraadi taotlemiseks ja selle alusel ei ole varem taotletud akadeemilist kraadi.
/Raido Puust /

ISSN 1406-4766
ISBN 978-9985-59-728-6

“There is no water-supply in which some unnecessary waste does not exist and there are few supplies, if any, in which the saving of a substantial proportion of that waste would not bring pecuniary advantage to the Water Authority…”
(Hope, 1892).

Table of Contents

List of Tables

List of Figures

Introduction

General

The real life water distribution system can be described by several mathematical models. The mathematical model should reflect the real life system as closely as possible. To do that it needs to be calibrated. Calibration can be done in two main aspects: quantity and quality. In this thesis water distribution systems (*WDS*) are analysed only by water quantity. Hydraulic models that are used here belong to two well known groups: (a) the steady state flow model and (b) the multi steady state model.

Hydraulic models are mostly calibrated against pipe roughness and/or unknown demands. By definition the calibration is the process where certain numbers of water distribution parameters are adjusted until the model presents the real life system as closely as possible. Although calibration process has been a manual task and in some cases it still is, a breakthrough came when calibration of *WDS* was formulated and solved as an optimisation problem. In practice manpower is still needed even in highly intelligent optimisation procedures and the development of fully automatic calibration for water network systems has been so far complicated or even impossible. Engineering knowledge is still needed in a sophisticated software system – to direct and/or analyse the results.

Overall, the calibration or optimisation problem in practice is very much simplified. The simplifications are mostly made because of lack of data. Insufficient amount of valuable data will lead us to the final solution that is not unique or in other words it is not the only solution to our problem. The calibration process is formulated through an objective function, a set of calibration parameters that are needed to calibrate, and analyses of one or more optimisation runs. Optimisation methods, in general, can be divided into two groups: global search space (searching a solution from a chain of hills) and local search space (searching a solution from a particular hill only). Some effort has recently been made to put both ideologies together, giving us a much faster and more accurate method than taking one alone. Among relatively large number of optimisation methods to choose from, very little attention is paid to the aspect which type of analyses should be made. Are these deterministic ones (where a fixed value is given as a final solution) or probabilistic ones (where in addition to a value also its probability is given). This thesis shows that in a highly non-linear environment and in case of lack of data probabilistic optimisation methods give more information about a system. Including also the errors that exist in all real systems, giving a result with one deterministic value should let us ask: “What makes you believe that this answer is correct and the only one in a multi-dimensional parameter space?”

Objective of the Thesis

The main objective of this thesis is to introduce a probabilistic (*Bayesian* base) approach in water distribution system for leakage studies that has not been used in that way before. Practically applicable to any kind of liquid distribution system, the objectives of the current thesis are as follows:

- To introduce a probabilistic calibration approach for *WDS* hydraulic models because the lack of data in well-known deterministic models does not always give us a unique, stable solution or a solution at all. The probabilistic calibration approach can give us a solution even in case of data deficiency. It is still shown that as more data become available, the results can be recalculated to decrease the uncertainties. This way, additional information is provided to a decision maker, when compared to the existing modelling approaches.
- Effective decisions about calibration parameters must be carefully chosen before the analysis. This thesis considers that a priori the model is calibrated against pipe roughness and only leakages are unknown. Choosing the reasonable boundaries to the possible calibration parameters will make the calculations more reliable.
- Practical benefits regarding the proposed calibration approach are tested with several case studies (artificial and real life system) to illustrate the current approach applicability.

Layout of the Thesis

This thesis has eight chapters. After the introductory chapter an extensive overview of the leakage studies is given (*Chapter 1*). The main algorithms that have been commonly used so far are described in *Chapter 2*. In the next chapter (*Chapter 3* an introduction into probabilistic modelling approach is given and the basics of the algorithm that is applied in this thesis are described. The leak modelling that is used in current thesis is described in *Chapter 4* followed by the description of software environment (*Chapter 5*) where the theory is put together into one easily adapted user interface. Case studies are analysed in *Chapter 6* and relevant conclusions are drawn and suggestions for future research are made in the final chapter (*Chapter 7*). A more detailed outline of the thesis is given below.

In *Chapter 1* a review of leakage studies since late 1960's is presented. Leakage studies are divided into three groups that are reviewed. Although the current research is dealing with leak detection only, it makes sense to give an overview also of other leakage studies (*leak assessment*, *leak control/management*). The idea is not to mislead the user but to inform that in practice a combination of methods is

used to detect leakages. The reviews are grouped by methodologies and their advantages and disadvantges mentioned.

In *Chapter 2* some main algorithms are described that have been used in water distribution system calibration/optimisation procedure. This chapter gives a background of current methods available and their advantages/disadvantages. Algorithm basic steps are given and a list of references given where more information can be found.

In *Chapter 3* the main part of probabilistic approach description is given. Starting from an introductory part to stochastic modelling, it will continue with *Bayesian* networks overview and its application in a *Markov Chain Monte Carlo* type sampler named as the *Shuffled Complex Evolution Metropolis* algorithm.

In *Chapter 4* the leakage detection methodology that has been used in the current thesis is presented.

In *Chapter 5* the software development will be looked at more closely. The base of the software system is *MATLAB* programming environment that is used to combine the hydraulic calculation software *EPANET2* (described first) and the *Shuffled Complex Evolution Metropolis* algorithm (layout and user identifiable parameters described) into one easily manageable user interface (discussed at the end of the chapter).

In *Chapter 6* multiple calibration case studies are presented with the following main objectives: (1) to compare the probabilistic calibration approach that has been put together in *Chapters 4 – 6* with some well-known deterministic calibration algorithms bringing out some main statistics with the results; (2) to test the probabilistic calibration approach in a more realistic (but artificially created) environment to find out its capabilities with additional possibilities to present the results; (3) to apply the methodology in a real life system based on *Rakvere Water (Estonia)* network data.

In *Chapter 7* a summary is given. Thereafter relevant conclusions are drawn followed by suggestions for future research work.

Acknowledgements

This work has been carried out at the Department of Mechanics, Tallinn University of Technology and partly at the Centre for Water Systems, University of Exeter, UK.

I would like to express my appreciation to all of my colleagues in Tallinn, especially to my supervisor Tiit Koppel and Leo Ainola for fruitful discussions.

I would like to express my sincere thanks to the people at the Centre for Water Systems, especially to Dragan Savic and Zoran Kapelan for their supervison.

Rakvere Vesi AS is thanked for providing the data for analysis.

1. Literature Review

1.1 Introduction

1.1.1 Leakage studies

Today's hydraulic pipelines might work at severe conditions – several network elements must work together to serve the water for consumers with its best economical value. William Hope (Hope, 1892) declared in his paper: "*There is no water-supply in which some unnecessary waste does not exist and there are few supplies, if any, in which the saving of a substantial proportion of that waste would not bring pecuniary advantage to the Water Authority…"* One field of study in fluid networks that always can bring some substantial savings is leakage minimization. The leakages have interested water authorities for decades. Although we might find information about countries where the overall leakage levels are really small (for example in Netherlands it is 3 – 7% of distribution input, Beuken et al., 2006) in most networks there is always a chance to decrease overall leakage levels. Controversially to a couple of low leakage networks we can even today find networks where there are more leaks than billed water (more than 50% of water is unaccounted, including leaks).

The causes of leaks are different and so are the methods that can be used to detect them. One of the breakthroughs in leakage studies was the knowledge when pipe leaks were linked to fluid pressure (Germanopoulos, 1985). The higher the pressure at leak position, the higher the leak flow and vice versa. The main causes of leak are bad connection, pipe corrosion or mechanical damaging. Those occasions affect leak behaviour and separate analyses/methods are used to detect them. We are interested in reducing leakages mainly because of two aspects: (1) leakages cause inefficient energy distribution through the network; (2) leakages may affect water quality in respect of infection that could come up with transients in water networks in low pressure conditions. A good review that looks at both aspects is provided by Colombo and Karney (2002). *Figure 1.1* shows how complicated task the leaks are playing in water distribution system.

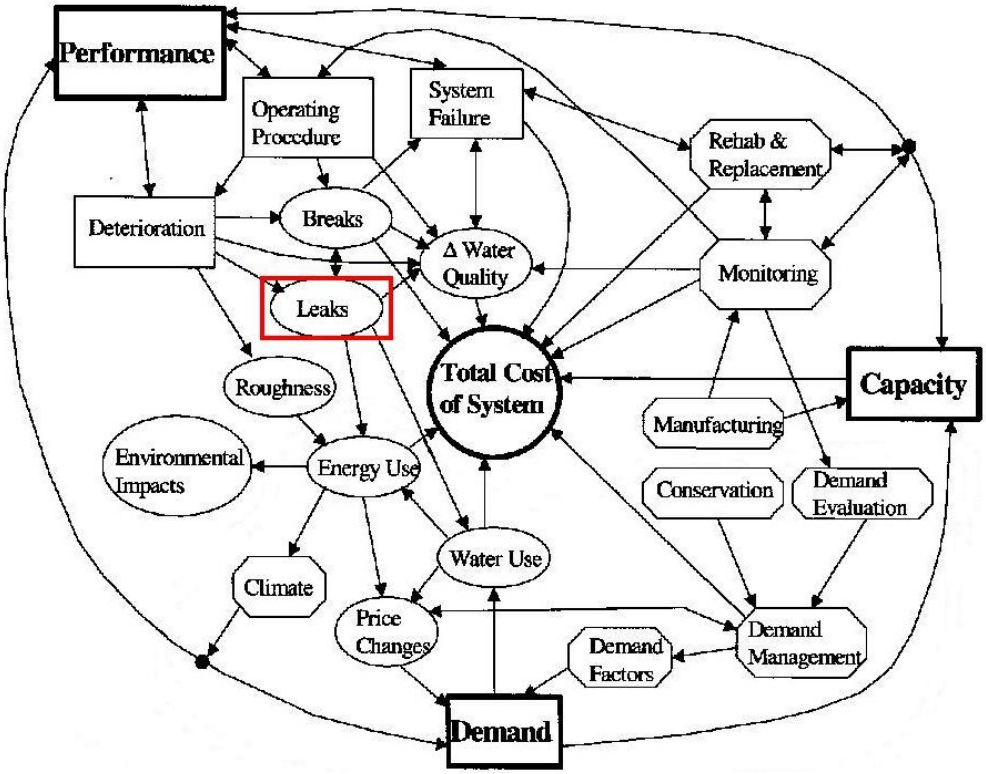

Figure 1.1 Leaks are a crucial part of the water distribution system labyrinth [Adapted from Colombo and Karney (2003, Figure 1)]

1.1.2 Previous review papers

There are few papers that concentrate only on leakage historical reviews. Most of them are part of a project or report. Morris Jr. (1967) describes the principal causes and remedies of water main breaks. It is a good starting point for those who are just starting leakage studies as it gives an overview of the causes that lead to water pipe breaks in systems. Depending on pipe material the causes are different or in some cases some of those might not happen at all. The comparison of key attributes of different leak detection methods can be found from (Cist and Schutz, 2001). A classification of leakage detection methods is also reported by Liou et al. (2003). A review of calibration methods in water pipelines (including leaks) is provided by Kapelan (2002).

1.1.3 The outline of the paper

Leakages in pipe networks have been studied by several authors. In this review we investigate these study areas dividing them into three main groups – leakage

assessment methods, leakage detection methods and leakage control/management models. Several other groupings can also be made to separate different methods/models from each other. The main idea is to present a historical review, how different methods have been developed in last decades and what are their advantages/disadvantages in today's perspective. Also this paper gives a broad list of references that concludes the studies.

The review has the following sections: *Chapter 1.2* gives the definition of leak and how it is modeled on a pipe. *Chapter 1.3* focuses more closely on leak historical research based on their classifications. *Chapter 1.4* concludes the paper. *Appendix A* is reserved for the equations used in leak modelling history.

1.2 Pipe leak – definition

1.2.1 Introduction

When leakage studies are carried out in a certain network, then the result normally will be an estimate how many losses there are (all unaccounted water that is not paid for). The easiest way to express it is in percentage value (how much water is lost from overall). Although this number is the easiest way to understand it is not recommended by several authorities since the size of denominator (i.e. the distribution input) varies according to water use habits, making the leakage levels seem lower in countries which use more water.

1.2.2 The basics of leak detection and modelling

Many of the today's softwares in the fluid modelling field include a module that can be used for leak detection (Rossman, 2000; AWDM, 2003). A leak is normally modelled as an independent outflow from a pipe/node. The flow in pipes, in essence, can be divided into the steady and unsteady state. Both flow states are well researched and more information can be found from any hydraulic book. In leakage studies a clear division is made, if the developed methodology is applicable in both regimes. Overall, leaks are investigated against discrete or continuous pressure/flow/sound measurements. The flow pattern in the unsteady state is more informative than in the steady state, but also more difficult to analyse as the unsteady flow behaviour is affected by pipe/node elements in a way that is only understandable when tests are made in laboratory conditions (where all pipe features are known). Wiggert (1968) investigates how leaks affect unsteady flow in pipes. When transients are introduced in pipe systems (pressure fluctuations in a system due to the regulation of flow area, pump/valve tripping) then every outflow (pipe connection) in a system affects its pattern – in general its amplitude decreases in non-reflected environment. Wiggert also introduces the leak modelling equation derived from lateral orifice outflow using *Bernoulli*'s simplified relation. The

equation that is used by many research studies (see *Appendix A*) can be written down:

$$q = C_d b\sqrt{2gH} \tag{1.1}$$

where q = volume rate per unit length; C_d = discharge coefficient; b = width of the slot; g = gravitational acceleration; H = pressure head. Equation (1.1) says that when pressure increases at leak position, then the outflow also increases. Although equation (1.1) is widely used in many research studies the user must be aware that emitter equation can lead to misleading results when the pressure is negative (Todini, 2003). The implication for leakage modelling is that one can determine a value of $C_d b$ to represent a leak and place that value at locations in the distribution system that correspond to leaks. The model of the system would then accurately determine the behaviour of the leaks and the effect of pressure modification. The key, of course, is that a modeller is never certain of the combination of leaks and whether the locations are correct.

1.2.3 More advanced studies in leakage determination

Having the mathematical model is not enough for leak detection. We need to know about the system as much as possible, even what type of leakages we are seeking. For example a research by O'Day (1982) introduces two terms for leaks registration: main leak and main break. Main breaks represent structural failures and main leaks represent inadequate joint sealing. Depending on the leak type we can rewrite equation (1.1) taking into account the type's special characteristics. This will be more closely described in *Appendix A*.

Recently special attention has also been paid to leak shape on the pipeline. Pressure affection to different leak openings in PVC pipes can be found in Cassa et al. (2005). It was experimentally tested that the leakage exponent value rises with size broadening. Further studies on different leakage exponents that could depend on leak hydraulics, pipe material behaviour, soil hydraulics and water demand can be found from Greyvenstein and Van Zyl (2005); Walski et al. (2006); Noack and Ulanicki (2006). Van Zyl and Clayton (2005) make additional remarks that when the leak is analyzed as a pressure dependent, the demand should follow the same procedure. There is a clear resemblance between the equations for leakage and demand elasticity. Pipe features including leakage as a hole in a pipe is investigated also by Beck et al., (2005a, b). The signal processing technique called cross-correlation was used in the laboratory set-up. The method was tested only for leak location but not for leak size. In Coetzer et al. (2006) the hydraulics of leakage through small circular holes in plastic pipes is investigated further.

1.3 Leak classification

1.3.1 Leakage assessment methods

In general leakage assessment estimates the quantity of water lost (water audit). It can be done for a single leak, leaks in some area or for the whole network. This group of study can be broken down into two simple sub-classes: (a) top-down leak assessment and (b) bottom-up leak assessment. As the top-down type is simply paper audits that utilize existing data without further fieldworks it can also be called a *desk top* exercise. In other words, the available data are transferred from one report to another. There are different table formats available. One could refer to (TWDB, 2005) for an example. The next step in the water audit might be the bottom-up type leakage assessment. In the following subsections both methods are more closely analyzed.

1.3.1.1 Top-down type leakage assessment

The general water audit deals with the value of lost water. That is calculated using the equation (1.2):

Water loss = water produced – water billed (or consumed) (1.2)

When talking about leakage, a clear distinction must be made between two terms: leakage and water loss. The *International Water Association* (Farley and Trow, 2003) has defined water loss as:

Water loss = real losses + apparent losses (1.3)

Table 1.1 IWA standard for international water balance and terminology

System input volume	Authorized consumption	Billed	Billed metered consumption (including water exported)	Revenue water
			Billed unmetered consumption	
		Unbilled	Unbilled metered consumption	Non-revenue water
			Unbilled unmetered consumption	
	Water losses	Apparent losses	Unauthorized consumption	
			Customer metering inaccuracies	
		Real losses	Leakage on transmission and/or distribution mains	
			Leakage and overflows at utility's storage tanks	
			Leakage on service connections up to point of customer metering	

In the top-down leakage survey the equation (1.2) and (1.3) are the most important ones. Real losses and apparent losses are described in *Table 1.1*.

If we step into a top-down type leakage assessment the following terms might be taken under consideration in addition to *Table 1.1* (TWDB, 2007):

- Technical performance indicators (total real loss/length of mains, total real loss/number of service connections/per day)
- Financial performance indicators

Despite its generality and simplicity (in sense of the required data amount), top-down leakage assessment should be carried out regularly. Gathering such information helps to decide what should be the next step in leakage studies for the particular network. More information about leak general assessment can be found from Stenberg (1982), Thornton (2002), Farley and Trow (2003), and Scott and Barrufet (2003). The latest reports about average losses in the UK (based on operated areas by different water service companies) can be found from (AHL, 2006). The latest guidance notes for leak location and repair were published by (Pilcher, 2003; Pilcher et al., 2007).

1.3.1.2 Bottom-up type leakage assessment

Bottom-up type leakage assessment can be called the second part of the audit process. This procedure is implemented when the utility has confirmed the data used in the top-down portion. It includes each aspect of the utility: billing records, the distribution system, accounting principles, etc. The audit's main purpose is to find out the efficiency of the water distribution system and the goals to achieve them. The bottom-up audits requirement is that the data used here should be as accurate and current as possible.

Leakage rates vary over the 24-hour period in response to changes in pressure. They are at their maximum at night when the demand is at its minimum. Therefore the flow measurements at night present the best method of determining leakage levels. For most districts, it has been found that multiplying night leakage flow rates by 20 hours gives a satisfactory estimate of the daily leakage (Stenberg, 1982). In this section attention is paid to general leakage studies and performance indicators found from literature.

In section *1.3.1.1* we already mentioned some measures that characterize the selection of leak location activity. Several countries have their own traditions. For example the sample measures could be: percentage of average daily flow (USA, France); m^3/km of mains/hours (German, Japan); litres/property/hour (UK) and litres/service connection/hour. These indications do not take into account component analysis techniques (Farley and Trow, 2003) that was introduced in the early 1990s and shown in *Table 1.1*. Therefore, additional measures have to be made. The result would be a performance measure. Several performance measures exist. The paper by Tuhovčák et al. (2005) compares four of them. Those indexes are presented in *Table 1.2*.

Table 1.2 Various performance indicators

Description	Calculation formula
Non Revenue Water (NRW)	$NRW = \frac{NRW(in\ pressure\ zone)}{Water\ delivered\ to\ pressure\ zone} \cdot 100\%$
Specific Leakage per km (JUVNF)	$JUVNF = \frac{NRW(in\ pressure\ zone) \cdot 10^3}{Length\ of\ water\ mains\ in\ pressure\ zone} \left[m^3 / km / year\right]$
Infrastructure Leakage Index (ILI)	$ILI = \frac{NRW\ (m^3 / year)}{Unavoidable\ annual\ real\ losses\ (m^3 / year)}$
Economical Leakage Index (ELI)	$ELI = (Economical\ Index) \cdot (Losses\ Index), \quad (Losses\ Index) = \frac{JUVNF}{3600}$

The most commonly used leak index nowadays is *ILI* (Lambert, 2003; Farley and Trow, 2003). The drawbacks of the other methods are: Although *NRW* is widely used because of its simplicity, it is not right to use such an index when consumption decreases (seasonally or annually, or due to demand management measures), the percentage of real losses increases even if the volume of real losses remains unchanged; *JUVNF* (*ELI*) is a more objective indicator of water losses due to the technical condition of the network but its disadvantage is that it does not include the effect of pressure relations in the pressure zone.

The *ILI* measures how effectively the infrastructure activities (see *Figure 1.2*), such as repairs, active leakage control and pipeline/assets management – are being managed at the current operating pressure.

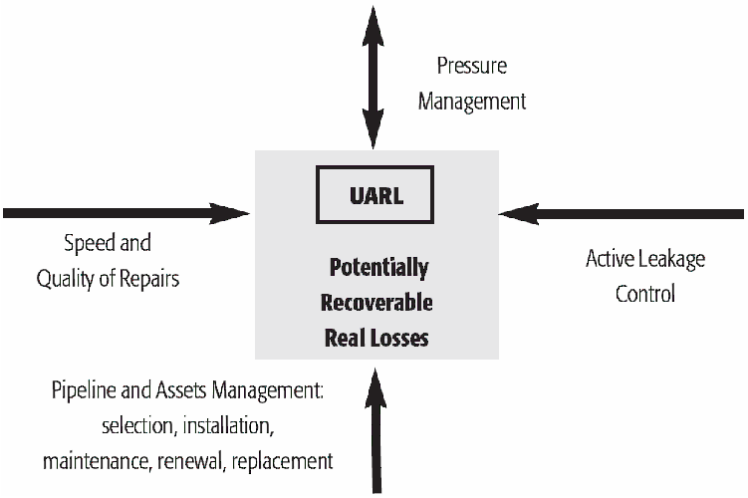

Figure 1.2 Basic methods of managing real losses

The advantages of using *ILI* are that it can be consistently applied across a range of utilities and that it is a measure of what can be achieved after giving the condition of infrastructure. Its key disadvantage is that it is not easily understood by non technical readers. Additionally it does not take into account the relative costs of

leakage management (and other marginal costs, like environmental costs) and it is not able to define what level of reduction is economically feasible.

Let's come back to the leakage equation first presented in *Chapter 1.2*. This equation can be written down in a more general form as:

$$q = k \cdot H^{\beta} \tag{1.4}$$

where $k = C_d b \cdot (2g)^{\beta}$. Leakage flow depends on the exponent β that previously has been taken simply 0.5. In literature much attention is paid to that exponent (more recently in Garzon-Contreras and Thornton, 2006). The reason to bring that equation in is that this exponent can be calculated using the *ILI* index. In Thornton and Lambert (2005) the equation is given as:

$$\beta = 1.5 - (1 - 0.65 / ILI) \cdot \frac{p}{100} \tag{1.5}$$

where p = the percentage of rigid pipes in network.

Bottom-up leakage assessment also includes the methodologies where equation (1.4) is used for general studies (Almandoz, 2003, 2005**)** in leak assessments. In these studies additional components are added – demand patterns. In those instances the measurements must be made at minimum night flow. Minimum night flows are also used when the starting value of the leak exponent is needed (in iterative processes). As this is clearly a system parameter, new calculations must be made regularly (before every new assessment). Iterative leak assessment has been used in Burrows et al. (2003) where EPANET software was used for hydraulic simulations. In essence the equality between water balance calculation (here it is called unaccounted for water) and the sum of leaks at the node level is sought. Nodal leakage interpretation is one of the strongest drawbacks in cases when leakage might not appear at the node at all. On the other hand, the methodology herein does not try to localize the leak but only to carry out the general assessment including also system unique parameters (pressure dependent leakage).

Although the classical minimum night flow analysis (Araujo et al., 2003a; Covas et al., 2006; Garcia et al., 2006) can reduce real losses (leakage) considerably there are many other methods for leak assessment. Some of those are simply used to redistribute the leakage in the network using various characteristics (Ainola et al., 2000; 2003) to get better understanding about the leakage levels (divisions) in the water distribution system. Additionally, fuzzy sets and systems are used in (Mamlook and Al Javyousi, 2003) and in the transient environment (Da Silva et al., 2005), the generalized likelihood ratio test (Mukherjee and Narasimhan, 1996), stratified random sampling with leak flow gaging (Arreguin-Cortes and Ochia-Alejo, 1997), standard weighted least squares state-estimation (Andersen and Powell, 2000), redistribution of leaks in respect of the screening method or sequential statistical analysis (Carpentier and Cohen, 1991; Buchberger and Nadimpalli, 2004), the virtual distortion method (Holnicki-Szulc et al., 2005) and

head driven simulation method (Tabesh et al., 2005; Cheung et al., 2005), the static and time delay neural networks approach (Mounce and Machell, 2006).

Fuzzy sets and systems are a part of the decision support systems that were developed in 1960s. At first mainly space industry used such technology, but it has also been applied to other engineering systems including water networks. Although this kind of methodology can be applied very well to WDS-s, it has some disadvantages – namely there is a clear need to use engineering knowledge for weights estimation before the analysis (like the pipe type, pipe age, pumping system, pipe diameter and demand). The statistical analysis for the detection of the magnitude or leaks is expected to be more accurate than conventional water audits and computationally less intensive than dynamic hydraulic models (Buchberger and Nadimpalli, 2004). On the other hand, the continuous high resolution measurements (for example every 1s, 10s, etc) of discharge taken at one (or more) location (in DMA) along a pipe are needed. This can be problematic in some cases. High resolution data measurements are not used quite often, especially when talking about bounded DMA. The location of data acquisition systems must be carefully planned in such case studies.

1.3.2 Leakage detection methods

Previously the leak assessment studies were carried out. The quantification and possible identification of loss groups (real and apparent) were described. The next step would be the leakage detection (does the leak exist and its localization). Various forms of leak detection have been around for many years and it is a process which has recently become very high-tech. Is it an equipment or non-equipment based technology; the common practice is that a particular method/solution is used in conjunction with the others. Therefore, the following chapter involves three sub-sections. Firstly, more general leakage detection methods available in literature are researched. Thereafter, moving through the different technologies the chapter concludes the list of precise technologies where the leakage can be pinpointed at its exact underground/submerged location.

1.3.2.1 Leakage awareness methods

The term leak awareness is used to get the understanding about the network if the leak exists in a particular area. It does not give any information about its true location. In general the hydraulic model is needed for the leakage awareness test. Various hydraulic models have been proposed to detect leaks in water distribution systems. Those methods usually involve a calibration/optimisation technique to analyse the different areas of the network. Most commonly the calibration/optimisation involves the minimization of difference between measured and calculated network parameters. Those can be the network's nodes pressures/heads, flows in pipes or both at the same time. This minimization gives

the solution to an inverse problem (Pudar and Liggett, 1992; Stathis and Loganathan, 1999). It was done in a steady state regime that brought up an important need – the calibration of the system must be well known in order to have confidence in the leak areas that are the result of the method. In fact, most systems that are used for leak detection are not well calibrated. Liggett (Liggett, 1993; Liggett and Chen, 1994, 1995) extended this method to transient flow. Already Nicholas (Nicholas, 1990) used the real-time transient analysis. The leak detection method was formulated by using a mass balance approach to compensate for the rate of change of inventory in the pipeline. It was noted that uncertainties in fluid properties limit the leak detection sensitivity and that the measurement noise causes time delays in the detection of leaks. Many other developments published an on-line computational technique to analyze hydraulic transients caused by leaks on a single pipeline (Silva et al., 1996).

Fluid transients are used to probe the pipeline in much the same way as the radar and sonar are applied to locate and identify objects. The use of fluid transients for leak detection has gained popularity over the last decade as they travel at high speeds and are susceptible to changes in the internal conditions of the pipe. Transient response of a pipeline with a distributed leak was investigated by Wiggert (Wiggert, 1968). It was found that the transients in a pipeline were greatly affected by the magnitude of the distributed lateral flow. The effects of the demands on the transients in a field pipe network test was researched by McInnis and Karney (1995) where a similar distributed leak model based on the method of characteristics was used.

Using transients for leak detections starts to develop this point forward mainly because of the capability of simultaneous calibration and leak detection (Chen, 1995). At this stage the leak is sought only at nodal points and a large amount of data is a requirement (at each node in the network).

Working with transients, the pressure transducers are used because they can give instant readings whereas most flow meters do not react instantaneously to a change in flow (Chen, 1995). The merit function therefore would be written as:

$$E = \sum_{i=1}^{M} \left(h_i^m - h_i \right)^2 \tag{1.6}$$

where M is the number of measurement, h_i^m is the measured head at node i, and h_i is the computed head at node i. That needs to be minimized over the entire network for the period of the calculation by choosing calibration parameters so that it makes the computed head as close as possible to the measured heads.

Various minimization algorithms have been used to do the job: *Levenberg-Marquardt (LM)* (Pudar and Liggett, 1992); *adjoint LM* (Chen, 1995), *Genetic Algorithms (GA)* (Simpson and Vitkovský, 1997; Vitkovský and Simpson, 1997; Vitkovský et al., 1999, 2000, 2003a; Tang et al., 2001; Araujo et al., 2003a; Saldarriaga et al., 2006; Wu and Sage, 2006), *Genetic Algorithms with Levenberg-Marguardt (GALM)* (Kapelan et al., 2000), *Shuffled Complex Evolution (SCE)*

(Vitkovský et al., 2001), *Artificial Neural Networks* (Caputo and Pelagagge, 2002), *Shuffled Complex Evolution Metropolis (SCEM)* (Puust et al., 2006). Additionally following progressives were made: Vitkovsky and Simpson (1997) analysed two encoding schemes (continuous and discrete) and concluded that a discrete coded scheme produces better results; prior estimates (pseudo measurements) were taken into account in the minimization problem by Kapelan (Kapelan, 2002; Kapelan et al., 2000, 2003a, 2003b, 2004) to avoid the ill-posed problem (that is there is no solution, no unique solution or the solution is unstable), to improve the accuracy of the estimated calibration parameters and to increase the seed of the convergence process. Attention must be paid on that prior estimates work better with pipe friction factors as these are less sensitive than leak effective areas. Several attempts have been made to use hybrid optimisation models (*GALM*) for calibration. Using both algorithm's capabilities at their best stage the model run-time savings and increased accuracy can be established.

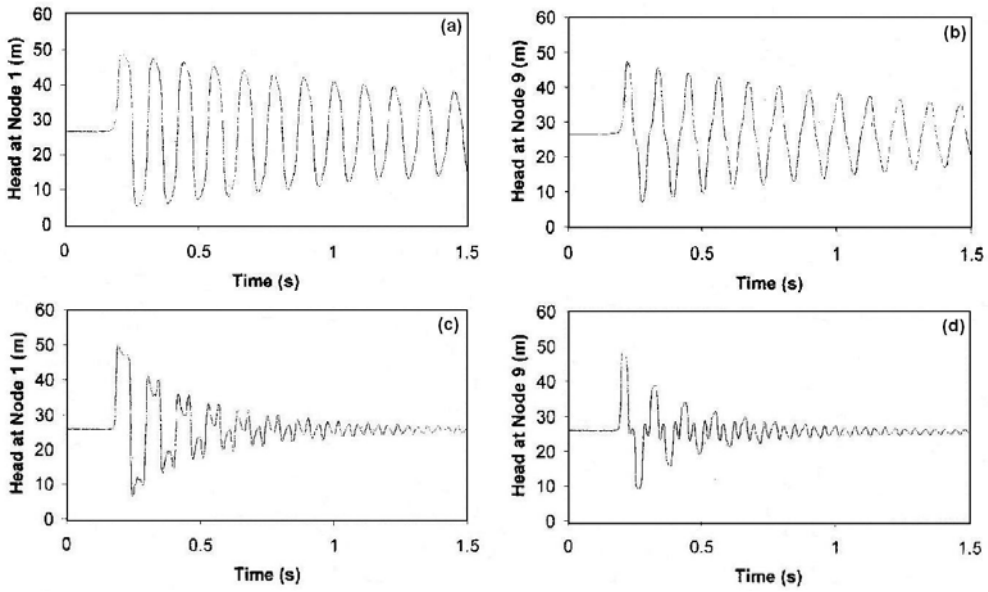

Figure 1.3 Transient analyses for leak detection
Head variations measured at the same nodes for no-leak case studies (a, b) and for leak case studies (c, d). Considerable head damping can be seen when leaks exists in the system. [Adapted from Vitkovský et al. (2001, Figures 3, 4)].

In general it is found that global optimization algorithms (*GA*, *SCE*) have the following advantages: they do not need to take numerical derivatives (i.e. sensitivities) or to calculate *Jacobians* (main sources of errors) and although the calculations need more computer power, they tend to give the global minimum rather than converge to a local minimum.
Inverse transients with *GA* have been tested for leak detection also on laboratory set-ups (Tang et al., 2001) and in real-life systems. Laboratory tests have shown

that it is vital that unsteady behaviour be able to be modelled without leaks initially (Vitkovský et al., 2001).

Time or *frequency analyses* of the pressure wave signal for leak detection are used in (Jönsson, 1995; Brunone, 1999, Brunone and Ferrante, 1999, 2001, 2004; Covas and Ramos, 1999; Covas et al., 2001, Wang et al., 2006). *Standing wave difference method* is studied by Covas (Covas et al., 2005a). In these studies the leak reflected wave is analysed through impulse response method. Wang (Wang et al., 2006) propose a simple methodology how to incorporate the time-domain analytical solution with the transients to a pipe section where two monitoring locations across a leak are used (similar to the acoustical leak detection techniques).

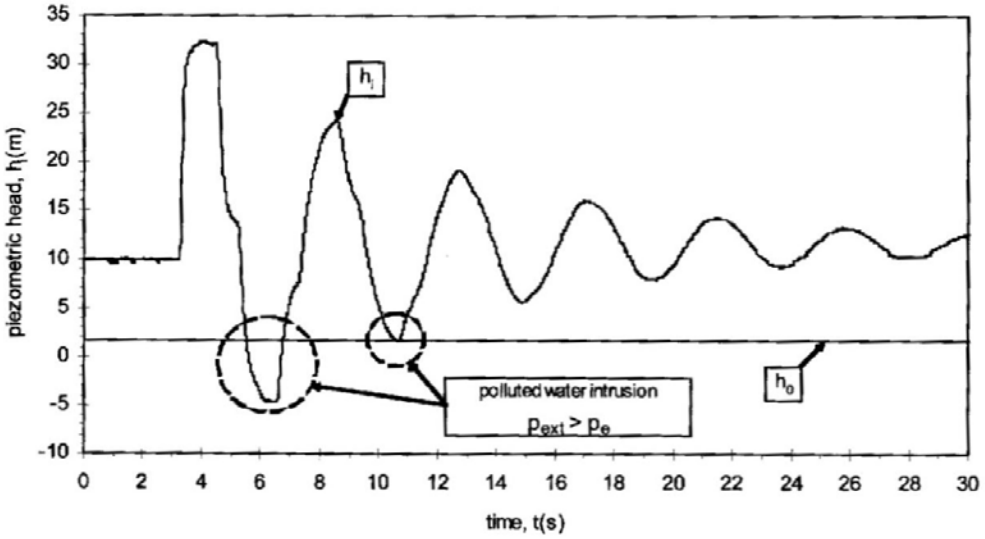

Figure 1.4 Pressure time-history at the leak during the transient due to the closure of the end valve

Even though the transient attenuates quickly, the risk of compromising water quality exists (where pressure falls below the tank level, *h0* and the ambient pressure external to the leak, pext is higher than the internal pipe pressure at the leak, *pe*). [Adapted from Brunone et al. (2004, Figure 1)].

In such cases the leak's discharge must change a pressure wave signal. In other words the leak's discharge must be sensitive to the pressure/head variation during the transient regime, otherwise such methods do not work. In addition to that these methods work only on well-known pipelines where all pipe features (junctions, bends, etc) are known. Frequency analyses deal with the dominant frequencies of the damped free oscillation by a *Fourier* analysis. This method is very sensitive to the topology of the system because the frequency associated with the leak is dependent on the leak's position, the leak's magnitude, and the presence of other singularities and the free damping of the system (Ferrante et al., 2001a, Ferrante and Brunone, 2003a, Wang et al., 2002). Such interpretation is also known as the modal damping method (MDM) of modelling the pipeline. Due to these artefacts leaks in some parts of the pipe can remain hidden (falling onto shadow zones).

Frequency analyses or *wavelet analyses* (harmonic analyses) are sometimes called the next step in pressure wave decryption (Stoianov et al., 2001). Ferrante (Ferrante and Brunone, 2001b, 2003b) describes two additional parameters that frequency analyses take the advantage of: scale and translation parameters that able to zoom in very short-lived in time and high-frequency phenomena, such as transient signals or singularities in functions. Moreover, wavelets are often compared to a microscope for their ability to reveal particular aspects of the signal at different scales just by adjusting the focus. Different wavelets exist. Continuous wavelets are explained by Ferrante and Brunone (2003b) and discrete ones are introduced in the later paper (Ferrante et al., 2005). In general *continuous wavelet analyses* give a better estimate of the leak location and *discrete ones* give a better reconstruction of the signal together with an estimate of the size of the leak. Computational power can be saved when the analyses are used together. Still practical tests are needed to show the usefulness of the methodology. The problem is that wavelet analyses are not good at systems where signals strongly decay in time (as those typical of transients in pressurized pipe systems).

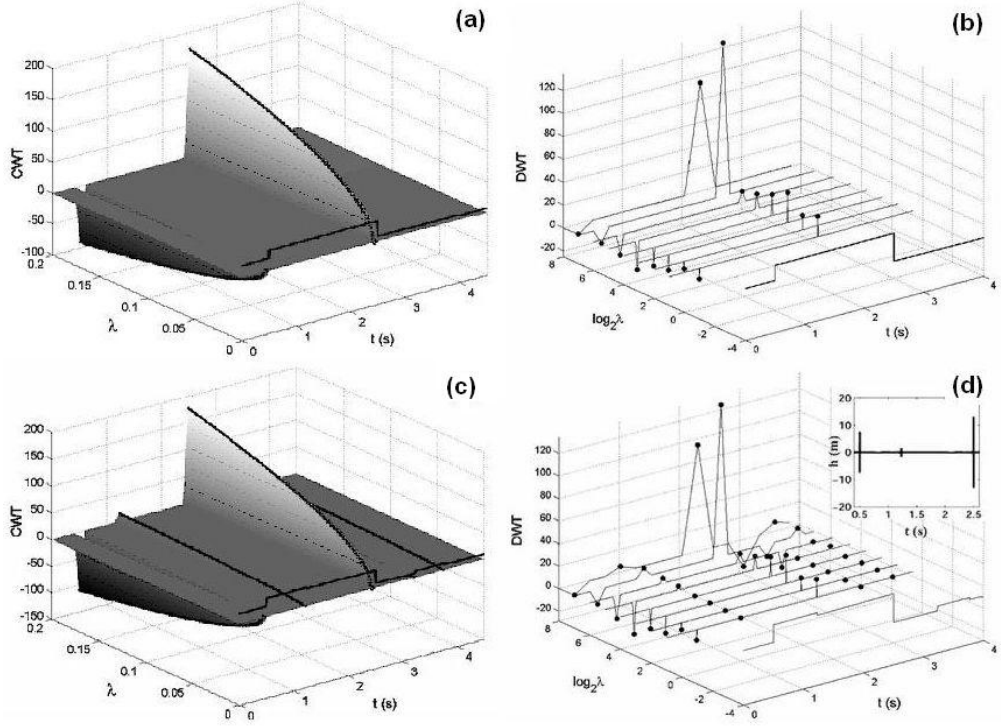

Figure 1.5 Continuous wavelet transform (CWT) and discrete wavelet transform (DWT) for no-leak (a, b) and for leak (c, d) case studies

It is possible to show the presence of the leak of a diameter equal to 1.49 mm in both cases. It is observed that chains of maxima appear in figures (c) and (d) in correspondence with the instants t = 1.23 s and t = 3.23 s, which are not present in figures (a) and (b). [Adapted from Ferrante et al. (2005, Figures 2 - 5)].

In case studies where time analyses have been applied to investigate leaks in looped networks (Covas and Ramos, 2001) it was found that the leak tends to spread all over the neighbourhood nodes. Several suggestions are made, including checking the synchronization of measurements all over the field.

In inverse transient analyses much attention is also paid to test different transient solvers. Covas (Covas et al., 2003) compares 'linear elastic' and 'linear viscoelastic'. It is shown that the linear elastic transient solver is very imprecise in the description of transient events in plastic pipes. Therefore for correct leak locations the 'linear viscoelastic" is recommended to use. Additional transient models for leak detection are tested by Wang et al. (2003). Extensive analyses on real pipelines are available in Covas et al. (2005b).

The *Time marching* and *Cauchy algorithm* (*initial value problem*) for leak detection was introduced by Liou (Liou and Tian, 1994, 1995; Liou, 1994). Both algorithms rely on recognising a leak discrepancy pattern that is specific only to leaks and continuous monitoring of pressures and flows on both ends (boundaries) of the analysed pipeline section is required. The *Cauchy* approach amplifies signals and noise while the time marching algorithm attenuates signals. Therefore, the amplification makes the *Cauchy* approach ineffective for highly frictional systems (and also when the data are noisy – because of the amplified noise its amplitude hides the change in the measured data).

Liou (Liou, 1998) presented the impulse response method that involves cross correlations between the low amplitude pseudorandom binary disturbance input and the systems output. This approach, like the statistical methods, does not use a mathematical model for the transient pipeline hydraulics. Dealing with data noise and extracting information from noisy data is the main focus of such approaches. As this is a reflection-based method, the impulse response function is unaffected if the opposite end of the pipeline from the generation source (e.g. reservoir, tank) is connected to a network (Vitkovský et al., 2003b).

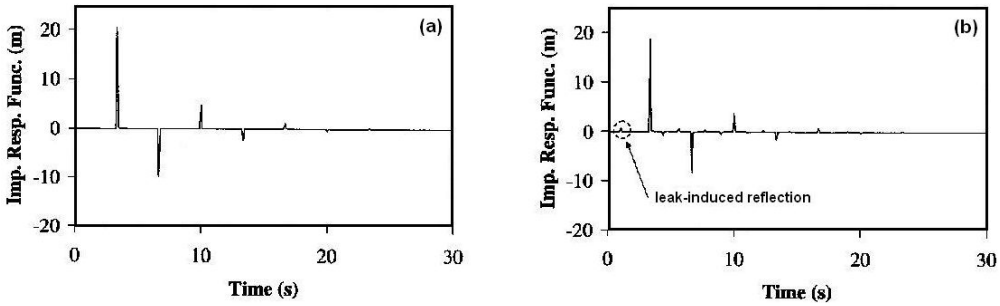

Figure 1.6 Impulse response functions for the non-leaking (a) and leaking (b) cases.
The first leak-induced reflection for the leaking case determines the correct location and size of the leak. The secondary reflection is negligible compared to the main leak reflection. [Adapted from Vitkovský et al. (2003b, Figure 4)].

The impulse response method with *Genetic Algorithms* is extended by Kim (Kim, 2005) considering unsteady friction for turbulent flow. On the other hand, the uncertainty of real pipeline systems (the noise of hydraulic data) has not been

considered here. Still if a comparison with other transient based methodologies had to be made, the impulse response method takes an advantage in platform independency, no discretization of the pipeline is needed and the shape of the generated transient is not important (in Mpesha et al., 2002 and Lee et al., 2003 it must be sinusoidal).

The frequency response method using a step excitation for leak detection was applied by Mpesha (Mpesha et al., 2001; 2002) and Lee (Lee et al., 2003; 2005a; 2005b) which is essentially based on the transfer matrix method (TMM). The leak is indicated by the presence of lower peaks in between the primary peaks in the experimental frequency response diagram, a leak can be detected as there is a difference between those peaks. If more than one leak is sought, the *Fourier* decomposition shows spikes at all locations (Lee et al., 2003). Lee (Lee et al., 2005a) presents two different approaches in the frequency domain analysis (the inverse resonance method that deals with fitting the numerical frequency response function model to the measured one; and the resonance peak-sequencing method that compares the actual shape with the already known response shapes).

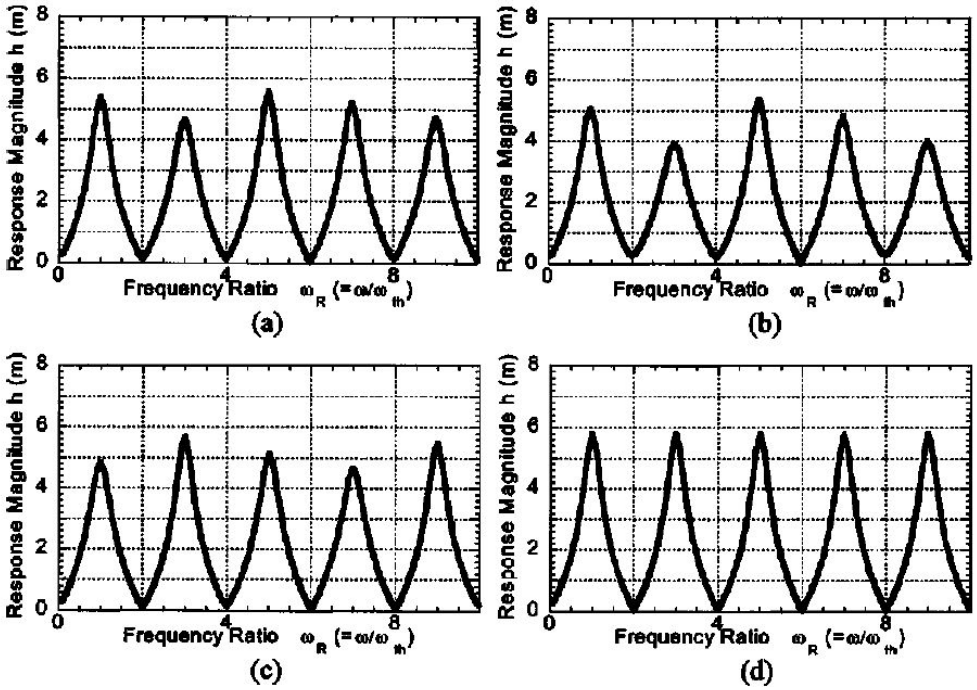

Figure 1.7 Impact of changing leak size and position on the frequency response diagram
It is extracted at the inline valve at downstream boundary: (a) leak at 700 m, $C_d b = 0.00014$ m^2; (b) $C_d b = 0.00028$ m^2; (c) leak at 1400 m, $C_d b = 0.00014$ m^2; (d) no leak. [Adapted from Lee et al. (2005a, Figure 6)].

One works with the other making the leak detection more efficient and limiting the size of the search space. Some limitations that needs to be known before using frequency response functions are: perfectly symmetric leak locations on the pipeline

will manifest themselves as a single leak and leaks located exactly at the mid-point of the pipe will be undetected; in addition to that because of linearized equations the transients involved must be kept with small magnitudes to avoid the effect on nonlinearity (Lee et al., 2005b). Lee (Lee et al., 2006) presents a solution where both symmetric and antisymmetric boundary conditions can be used. The oscillating inline valve is replaced by the side-discharge solenoid valve for sharper transient signals. When comparing with other leak detection methods the advantages of such technique are: (a) no-leak benchmark is needed; (b) no forward modelling is needed (like pipe roughness, pipe size determination); (c) system noise does not affect the performance. The method has to be tested on real pipelines to prove its validity/effectiveness. For comparative studies where the TMM and MDM methods are compared, see the paper by Zecchin et al. (2006).

In an uncertain case it is necessary to carry out to see how it affects the results. Although it is mentioned that a leak produces a higher pressure amplitude peak than other discontinuities, the whole system must be known in a no leak situation. Therefore, the methodology is available mainly as real-time leak detection on single pipelines. The advantage is that this methodology needs only one valve operation to take pressure measurements at a certain frequency (no need to vary the period of valve oscillations as many other methodologies based on transients).

Zecchin et al (2005) incorporate stochastic framework to model noise within a frequency-domain model that has not previously been used for leak detection. Maximum likelihood estimation (MLE) was used to estimate the unknown model parameters (like leak location and leak size). It was noted that if a priori parameter knowledge does not exist, the objective surface could be extremely flat and no information about leaks in a system can be extracted. It is one of the rare occasions where the importance of noise is incorporated into the leak analysis and is connected with the final results.

The determination of the power spectrum of computed and measured hydraulic transients was performed by Jönsson (2003) in order to try to distinguish the effect of a leak on the spectrum. It was concluded that no extra peaks are found in the spectra which could be attributed to the leak. Therefore, no evidence has been found that the spectral analysis should have the potential of adding any information on leaks in a pipeline.

Other leak detection methods that have been used for leak awareness include: based on the Liapunov criteria (Abhulimen and Susu, 2004), based on the Kullback information and Kalman filter (Wang et al., 1991; Verde, 2001, Lesyshen, 2005), based on the Shannon's entropy function (Robert-Nicoud et al., 2005), based on the cepstrum analysis (Taghvaei et al., 2006), based on the Bayesian system identification methodology (Poulakis et al., 2003; Rougier, 2005). The key point why to use Bayesian interface in the leakage studies is that normally we deal with different kind of errors that are not always possible to include in calculations. Therefore, to make more sense, the final discrete value is bounded with a certain probability that gives us more information about the result reliability. The drawback is that usually such procedures need more computer power for calculations. In general, probabilistic models are also used to conduct critical analyses, where small

cracks in pipes might cause the overall failure of a much bigger system (like cooling systems in nuclear power plants, Rahman et al., 1997).

1.3.2.2 Leakage localization methods

The purpose of leakage localization is to limit the leak search area. Still it is not possible to pin-point a leak. Methods that suits into this group include: acoustic loggers, step-testing, ground motion sensors, ground penetrating radars.

Normal flow in water pipelines occurs at low enough velocities not to generate any sound or vibrations. The whole picture changes when water flows from a leak, - energy in the system is released and converted suddenly into other forms of energy. This includes vibrations that are also referred to as leak sounds, which are conducted along the pipe and through the soil to the ground level. The object of acoustic leak tracing equipment is to convert vibrations at the listing point into audible sound. Logged data are analysed statistically, e.g. by means of frequency analysis of leak noise levels, to detect the presence of leaks.

Acoustical leak detection methodologies are used by Moyer (1983), Hough (1988), Hessel et al., (1999), Hunaidi and Chu (1999), Miller et al. (1999), Lockwood et al. (2003), Shimanskiy et al. (2003), Bracken and Hunaidi (2005) and Muggleton et al. (2006). Acoustical leak detection with assessment studies can be found in Hunaidi (2005).

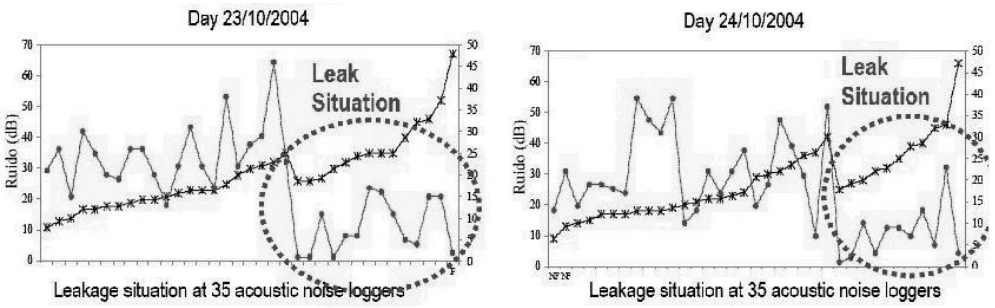

Figure 1.8 Results of acoustic noise loggers in two consecutive days (after repairs)
"Line with crosses" – noise amplitude and "Line with dots" – noise dispersion. Leakage situation corresponds to "line with crosses" above "line with dots". [Adapted from Covas et al. (2006, Figures 8)].

Acoustical methods have also been used in conjunction with *neural networks* that are able to analyse different leakage types (Hessel et al., 1999). Hough (1988) tests a special method based on *pressure/temperature change relationship*. It has a clear disadvantage that measurements should be made over a reasonable time-period (usually 24h) to avoid differences within the pipe wall and ambient temperatures. Much modern technology investigating leakages through temperature change is based on *infrared radiation image* (Hunaidi et al., 2000). It is not well documented yet but the questions that need to be answered are mostly dealing with survey time –

when and where (ambient conditions) it is most useful to use such methodology. *Electromagnetic surveys* (ground-penetrating radar, ground motion sensor) have also been used for leak detection (Lockwood et al., 2003; O'Brien et al., 2003). Methods like this could be used to locate leaks in water pipes by detecting either an underground void created by leaking water as it circulates near the pipe or by detecting anomalies in the pipe depth as measured by radar. Ground-penetrating radar is similar in principle to seismic and ultrasound techniques. The main disadvantages of such methods are that these can mislead to improper conclusions if a metal piece (or other anomalies) exists in the ground; and because of soil scanning, small leaks that do not cause considerable change in soil saturation, cannot be detected.

The acoustical methods have the following advantages: no need to excavate to install the sensors, only one measurement point is needed. The main disadvantages are as follows: the presence of air in the system (during the pipe fill-up stage) reduces sound propagating properties (because the propagation of any significant sounds from a leak in a water-filled pipe is through the water in the pipe and not through the pipe wall); pipe material strongly affects the sound attenuation (attenuation is greater in lead, PVC, PE and rubber pipes and smaller in steel, cast iron, copper, concrete and asbestos-cement pipes – Stenberg, 1982); the so called masking effect of leaks should be considered (two different leaks hide each other) and after repair the survey should be remade; any background noise (continuous traffic, cooling compressors, etc.), might disturb the survey.

1.3.2.3 Leakage pinpointing

Leakage pinpointing includes methodologies that are the most accurate in today's leak detection surveys. Three main groups that can be described here are based on (a) gas injection (Field and Ratcliffe, 1978, Hunaidi et al., 2000; Farley and Trow, 2003); (b) leak noise correlator (Grunwell and Ratcliffe, 1981; Cascetta and Vigo, 1992; Gao et al., 2004, 2005, 2006; Hunaidi et al., 2004; Muggleton et al., 2004, Muggleton and Brennan, 2004, 2005) and (c) pig-mounted acoustic sensing (McNulty, 2001). Their historical appearance is shown at *Figure 1.9*.

Leak noise correlators were first introduced commercially into the marketplace in the late 1970s (Thornton, 2002) but technology has been improved even over the last few years. *Water Research Centre* (*WRC*) in *England* was one of the leading research institutions who were able to apply the methodology to real pipelines. To correlate the sound from a leak, two microphones are located in contact with the pipe or valve stems at the same time, with one microphone on each side of the leak (Grunwell and Ratcliffe, 1981; Stenberg, 1982). The sound is compared in the correlator, which is capable of determining the difference in time for sound to reach the correlator. Knowing the speed of sound in the pipe, it is then easy to calculate the distance to the leak, which will be independent of the geophone, traffic noise, etc. The distance between sensors can be read from distribution system maps when the correlator is used in survey mode but it should be measured onsite accurately

when it is used in pinpointing mode. Leak noise correlators can be used over quite long lengths of pipe, depending on the pipe diameter, material and the lack of ambient noise, which could interfere with leak sounds.

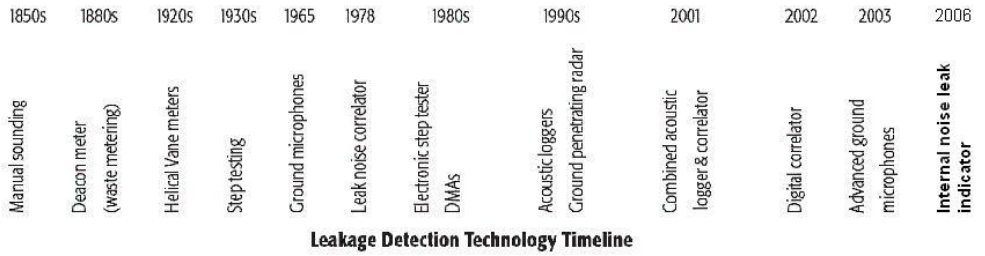

Figure 1.9 Several leak detection methods by historical appearance [Adapted from Pilcher et al. (2007, Figure 17)].

Most correlators can manage in excess of 500 m reaching even up to 3'000 m in ideal conditions. The lower end, particularly on plastic pipe, may be as low as 10 or 15 m. Therefore if sounding is undertaken every 200 m (quite common that this distance is dictated by fire hydrants interval) but the pipes are plastic, there is a very good chance that unless the leak just happens to be next to a hydrant of fitting which is being sounded, the leak could very easily be missed. Plastic pipes in these situations are more closely researched by Gao (Gao et al., 2004, 2005, 2006). Most significant improvement can be established in multi-leak and coherent noise situations when acceleration signals are used (among pressure and velocity signals). Attention must be drawn to the fact that when using noise correlators, the pipe segmentations (for example tees) need to be known. Imagine a pipe system with a branch from which a leak exists at some distance. When noise correlators are placed on the main pipe (at both sides of the tee), the leak is shown at the tee and not on the actual position (distance). Therefore, misleading information can be derived for excavation. These situations can be ruled out when the system itself is well known (at branch levels).

Gas injection and tracing techniques are not so often used for leak pinpointing, mainly because the other techniques have had more success. Anyway, this technology has advantages as a collateral method. The most common tracer gases used are sulphur hexafluoride (SF_6) and industrial hydrogen. Field and Ratcliffe (1978) used the SF_6 technique. The main disadvantage of the SF_6 method is that bar-holes have to be made in the ground at 1 m intervals along the pipeline, to allow the gas to collect and be traced. Nowadays industrial hydrogen (Hunaidi et al., 2000) has swapped the SF_6 method, mainly because of the following advantages: (a) speed of tracing (no need for bar-holing), (b) capability to find multiple leaks in a single section of the pipe (for example during a step test), (c) finding leaks on unexpected loops and bends that makes accurate correlation difficult (Farley and Trow, 2003). The main disadvantages are: (a) the time needed for the gas to surface is relatively long and (b) it might be a concern that gas could be trapped near the ceiling of water-filled pipes and thus could not escape if leaks are not near the top of the pipe.

Therefore, gas injection method may be in some situations impractical for routine leak surveys or pinpointing.
The pig-mounted (in-pipe) acoustic technique has also been used for leak detection (EPSRC, 2002; McNulty, 2001). This technique requires under-pressure insertion of a microphone (or pair of microphones) into the main. The velocity of water carries the microphone to the leak position whereas the noise and its position are recorded all the time. The so called intelligent pigs (*pipeline inspection gauges*) are highly sophisticated, computerised, self-contained devices inserted into the pipeline and propelled forward by the liquid flowing through the line, recording information as they go. They were originally developed to remove deposits which could obstruct or retard flow through a pipeline (EPSRC, 2002). Inline pigs are used to carry different kinds of sophisticated measuring devices such as magnetic flux leakage (Mukhopadahyay and Srivastava, 2000), hydroscope (Makar and Chagon, 1999) or ultrasonic tools (Willems and Barbian, 1998) along the pipeline. In general these tools need that the pipe is cleaned and therefore it makes it difficult to apply the methodology to old pipes (where heavy corrosion can exist). Access to the inside of the pipeline is also needed. Also as pigs are in contact with the pipe wall, its effect on water quality must be considered before a survey is performed.

1.3.3 Leakage control/management models

Leakage control/management can be classified into two groups: (a) passive (reactive) leakage control and (b) active leakage control. Passive leakage control is a policy of responding only to leaks and bursts reported by the public (in some cases also by the company's own staff). Active leakage control deals with management policies and processes used to locate and repair unreported leaks from the water company supply system and customer supply pipes.

1.3.3.1 Passive policy

Many water utilities still take a passive attitude of waiting for the trouble to come and repairing it only when leaks become self-evident with several inevitable problems associated for customers (Ramos et al., 2001). Passive control of leaks (waiting for consumer complaints) can lead to a percentage of losses around 40%. According to experts, a percentage of 10 – 15 % of losses is considered fairly satisfactory and from this point forward it seems not to be economically feasible to invest in leakage control and repair (see *Figure 1.10.*).

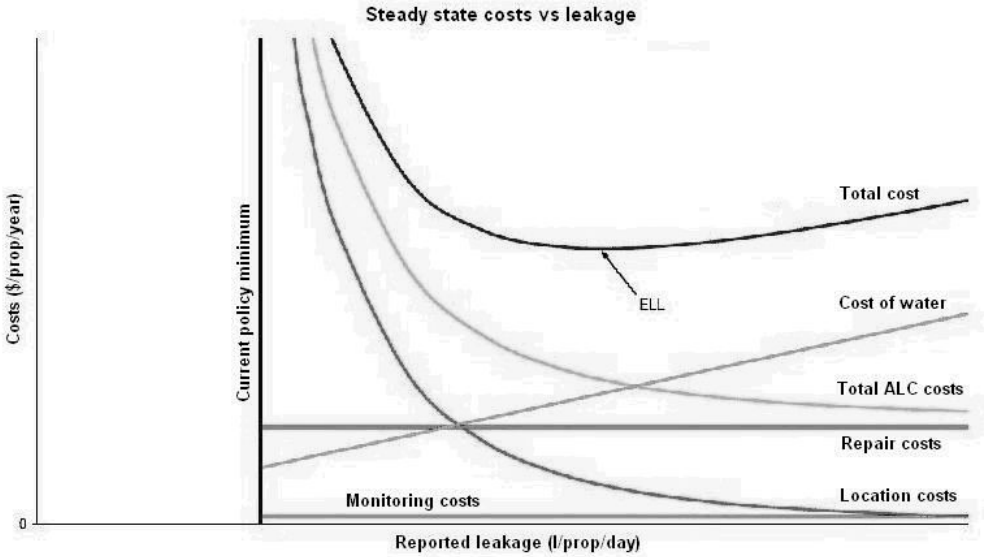

Figure 1.10 Typical economic level of leakage (ELL) analysis
Curve 'ELL' is the total cost of operation i.e. cost of leakage control plus cost of water production. As it can be seen, the curve will be high initially due to the high cost of leakage detection required to achieve very low levels of leakage. The point at which the total cost is lowest will be the short run economic level of leakage. ALC – Active Leakage Control. [Adapted from ATKINS (2005, Figure 2)].

The main problems with passive control policies are that a rupture cannot be detected if it occurs in a small branch, particularly if it is out of use, and also considerable leak losses (leak locations) cannot become visible because of soil structure.

1.3.3.2 Active policy

Active leakage policy includes the following general techniques: (a) active pressure management, (b) active leakage control (looking for unreported bursts), (c) sectorisation and (d) economic intervention. The most appropriate leakage control policy will mainly be dictated by the characteristics of the network and local conditions, which may include financial constraints on equipment and other resources (Farley and Trow, 2003). The main factor governing choice, however, is the value of the water – if a particular methodology is economic for the savings achieved.

Few active leakage control methodologies have been published: the point-analysis technology (Farmer et al., 1988), a technology based on Newton iteration matrix (Dunlop, 1999), the neural recognition system (Gabrys and Bargiela, 1999), discretizing the spacial variable for multi-leak detection (Verde, 2005), belief rule based expert systems (Xu et al., 2007), pressure transients with one measurement

point (Misiunas et al., 2003, 2005a, 2005b, 2006), the GIS based on-line evaluation model (Tabesh and Delavar, 2003).

All such methods are assume that pressure/flow data are available continuously so that when the leak appears, the systems ability to compare data in respect of rapid (or uncommon) change (or pre specified faulty patterns) in pressure/flow values detects it. Experts are able to provide a set of rules to distinguish patterns between operations under normal and leak situations. It has been shown that the learning stage can start with a random rule base and therefore pre-knowledge does not have to be provided which can be assumed to be a clear advantage in leak detection technologies. The drawback of such systems is that these methods are available mainly for pipelines and not for pipe networks. Those methods are most suitable to apply for environmentally critical fluid pipelines (where oil or gas is transported).

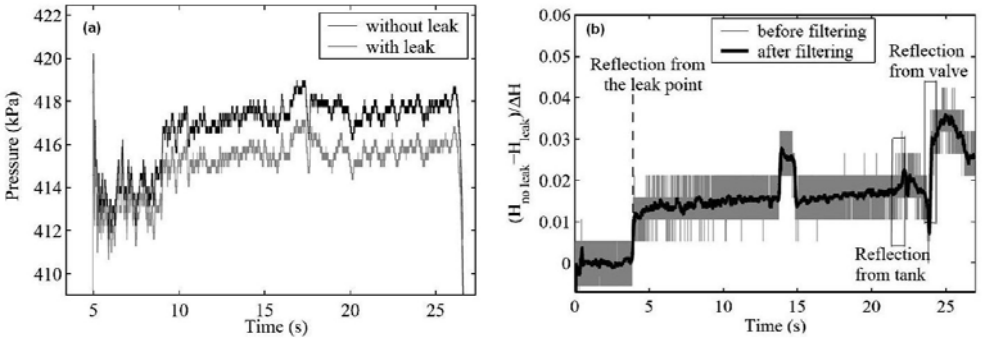

Figure 1.11 On-line leak analysis. Comparison of pressure traces measured with and without leakage

The change in difference between the two traces (a) indicates the presence of a leak. The actual difference between measured pressures can be analyzed to get better resolution as shown in (b). [Adapted from Misiunas et al. (2006, Figures 9, 10)].

Pressure transients used for on-line leak detection need to be further developed in the areas when the leak appears in normal pressure transients conditions. So far this situation is left out of the analysis because otherwise it would be difficult to compare the pressure signal from the real leak with the pressure signal of normal operation (including transients caused by pump start-up, valve closures etc.).

Active pressure management has been used for leak reduction in several water distribution systems. Pressure management has been around for many years in various forms but only in the last few years the use of advanced pressure control has been used on a wide basis in system optimization and in loss reduction and management programs. Consider a simple example when pressure is changed from P_0 to P_1, the leakage rate changes from L_0 to L_1. Therefore, $L_1 = L_0\left(P_1 / P_0\right)^{\beta}$, where β is a leakage exponent and normally taken in between 0.5 – 2.95 depending on leak types. *Table 1.3* shows different values of β derived from field tests.

When a zone with fixed-area leakage has a leak rate of 100 m^3/s at 60 $m\ H2O$ and the pressure is reduced to 35 $m\ H2O$, what will be the savings on leakage rate? Of course the answer depends on the leakage exponent β.

Table 1.3 Summaries of exponents b derived from field tests.
[Adapted from Garzon-Contreras and Thornton (2006, Table 1)].

Country	Number of zones tested	Range of exponents β	Average exponent β
United Kingdom (1970's)	17	0.70 to 1.68	1.13
Japan (1979)	20	0.63 to 2.12	1.15
Brazil (1998)	13	0.52 to 2.79	1.15
United Kingdom (2003)	75	0.36 to 2.95	1.01
Cyprus (2005)	15	0.64 to 2.83	1.47
Brazil (2006)	17	0.73 to 2.42	1.4
Totals	**157**	**0.36 to 2.95**	**1.14**

For example when the $\beta = 0.5$, the answer would be 24 m^3/s and if the $\beta =$ 1.15, then the savings would be 46 m^3/s. The conclusion can be drawn that when interpreting any kind of pressure reduction we have to know what type of leakages dominate (Thorton, 2002; Thorton and Lambert, 2005).

A note by Germanopoulos (1985) states that the purpose is not to evaluate the consumer flows and leakage losses with great accuracy but to give a network simulation model which can be used to approach more realistically certain problems in water distribution system operation.

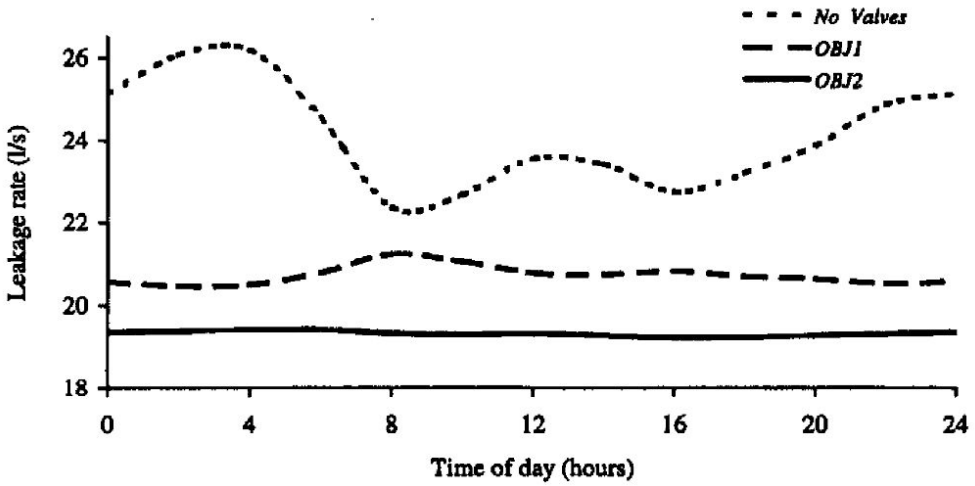

Figure 1.12 Rate of leakage (with and without optimized valve control, i.e. pressure regulation)
Lines 'OBJ1' and 'OBJ2' indicate different optimization model types. [Adapted from Vairavamoorthy and Lumbers (1998, Figure 6)].

May (1994) gives an interesting formula to get the boundaries when we can talk about major leaks in a particular system using the pressure-leakage relationship formula. Different approaches for pressure dependent leakage interpretation have been published (please see the *Appendix A* for further information) and there is no doubt that a particular equation might work well in one system but not in the others. Pressure-leakage relationship as a basis active pressure management is usually achieved with valve control.

Pressure reduction in water distribution systems is normally achieved through pressure reducing valves. The objective of pressure reduction is to ensure the target pressure at any given zone/area/node that satisfies the customers and not more. Germanopoulos (Germanopoulos, 1995) proposed an indirect formulation of the leakage minimisation problem, under the same assumption of independence between intervals (keeping certain minimum pressure in the network). At this stage the problem solving was linear that did not strictly guarantee leakage minimisation. Different automatic approaches are available where pressure reduction is solved by using a computer algorithm/program: sequence of quadratic programming (SQP) (Vairavamoorthy and Lumbers, 1998), SQP with genetic algorithms (Reis and Chaudhry, 1999; Araujo et al, 2003b, 2006), SQP with parallel computing (Hernandez et al., 1999; Alonso et al., 2000), the generalised gradient method (Ulanicki et al., 2000; Ulanicka et al., 2001), maximization of a likelihood function (Tucciarelli et al., 1999). The performance of different kinds of pressure valves for leakage reduction is described in (Dias et al., 2005) and specifically concentrated for case studies on real networks (Martinez et al., 1999). A leak redistribution algorithm for getting more precise leak distribution in a real network is also described in (Ainola et al., 2001).

In general pressure reduction will raise the system performance in respect of water quality and service targets. Active policy can be used before the general leak detection or afterwards. It is recommended to do it after leak assessment studies because normally active policies need considerable investments. Pressure management needs the knowledge for finding best suitable points in the network to place (or relocate) pressure reducing valves. The most sensitive factor is a terrain shape. For a flat terrain it gives the best results. Other fields where pressure reducing gives an immediate effect: (a) minimising the total cost of water supply; (b) reducing excess static pressure (minimizing the bursts risks); (c) smoothing out pressure variations; (d) increasing unacceptably low water velocities in pipes. In some cases there are certain times of day when the pressure reduction is feasible.

1.4 Conclusions

From the mid 1950s the leak assessment/detection studies have been paid a considerable amount of attention in pipelines/pipe networks. The purpose of this review is to span the most influenced methodologies (and their developments) in leakage studies.

Leakage studies start with making use of proper terminologies. IWA has done a lot of work in that respect, proposing an international standard for water balance and terminology. Leak assessment studies (bottom-up, top-down) separate the water balance components from each other and general conclusions have to be made, deciding where to go further (which leakage methodology to use). Decisions about leakage methodology involve the determination of the leakage type and its interpretation (equation) in the model. Research papers at a very high level are available where only the leak shape on the pipe is under consideration and how it affects the modelling. The modelling side of leakage equation has evolved from the 1980s. The break-through was when the models incorporated a pressure-dependent part in the equation.

As far as the loss of revenue is concerned, the systems with high leakage will almost always see a positive benefit from pressure management. This is especially true for systems with high water production or purchase costs. In some cases where the revenue cannot be tolerated, the pressure management can be limited to night hours – legitimate consumption is at its lowest and system pressures are at their highest.

Pressure transients have been given considerable attention to leakage studies. From the late 1960s it has been used as a basis for different approaches. As leak location accuracy depends on the leak size and location, flow regime, location where the transient event is generated, pipe material, system configuration (e.g. branches, pipe diameter) and presence of intermediate obstacles, it is very difficult to apply it as one and only. Quite an important factor for such technologies was how many measurement stations were needed. Some of those were able to accomplish trustful results with a single measurement point (reduced costs of methodology). The *Fourier analysis* is quite commonly used to distinguish the pressure waves. One of its drawbacks is that the *Fourier analysis* loses time information and is therefore not well suited for analysing transient signals. To retain the advantages of the analysis both in the frequency and time domain, the wavelet analysis is applied to pressure signals. In real situations the attenuation of the pressure wave is a very important parameter to be taken into account for the precision of the leak location computational method. Any complex waveform created by branches and loops and demands reduces the possibilities that leak will be distinguished from other pipe features (*Leak Reflection Method*).

The most accurate technologies are worked out for single pipelines. Noise correlators started to appear into the field from the late 1970s. Such methodologies have been improved very rapidly and nowadays it is quite a common practice that the authorities carry out regular surveys with leak correlator systems. It should be noted that all acoustic leakage testing methods suffer the calibration issue (this has to be determined by empirically based laboratory tests). The parameters include: wall thickness, size and type of defects, pressure ranges, medium with different viscosity, ground conditions.

The computer power available today has enhanced the possibilities to make sophisticated analyses using a specific mathematical algorithm for leak awareness/detection.

In the near future it is likely that stochastic modelling approaches make their move. Leakages are very little researched in terms of their probabilities. In a distribution network, the flow itself is noisy (so that even a perfect sensor measuring flow or pressure would output a non-steady signal). Furthermore, the noise amounts to dynamic noise due to the system's interaction with the environment. Therefore, it is not possible to build an accurate non-linear model completely describing the system that incorporates the uncertain data.

To solve the leakage in the water distribution systems, sophisticated leak detection techniques are required. The best method of leak analyses and control is very tightly bound with the type of system and customer base in the utility. Despite the wide research topic and numerous leak detection techniques, the one and only leakage methodology has not been worked out. Moreover, some of the methodologies are labour intensive, imprecise or only appropriate to a limited area of the network. This ensures that the development continues and new methodologies are put into practise. Whilst new systems may appear, the most likely developments will be in the combining of current systems.

2. Calibration algorithms

2.1 Introduction

Given a set of observations, we usually want to summarize the data by fitting it to a model that depends on adjustable parameters. Sometimes the model is simply a convenient class of functions, such as polynomials or *Gaussians*, and the fit supplies the appropriate coefficients. At other times, the model's parameters come from an underlying theory that the data are supposed to satisfy. Modelling can also be used as a kind of constrained interpolation, where you want to extend a few data points into a continuous function, but with an underlying idea of what that function should look like. When dealing with pipe equations in network studies (mass and energy balance) we end up with series of nonlinear equations. In principle these equations can be solved in a variety of ways. If they were linear, the simple least-squares method would suffice (Pudar and Liggett, 1992). They could be linearized about a trial solution, the least-squares solution performed to find a new solution and then iterated or some so called global optimisation techniques can be used instead. There are several methods commonly used in pipe network studies and hereby we give short descriptions to some of these with relevant references.

2.2 LM (Levenberg-Marquardt)

The primary solution criterion is the minimization of the difference – in the least-squares sense – of the measured and calculated heads. With nonlinear dependences, the minimization must proceed iteratively (Press et al., 1988). Consider the minimization problem given by equation (2.1) (in matrix representation).

$$\sum_{l=0}^{M-1} \alpha_{kl} \delta a_l = \beta_k \qquad (2.1)$$

where δa_l represents the correction of kth parameter of current minimization step; expressed as:

$$\delta a_l = \text{constant} \times \beta_l \qquad (2.2)$$

Equation (2.1) converges in a parabolic manner near the minimum. However, at a point far from the minimum, especially where the curvature of the merit function (measures the agreement between the data and the model with a particular choice of parameters) is negative, convergence is not guaranteed. An efficient method that deals with such minimizations is the *Levenberg-Marquardt* method. Equations (2.1)

& (2.2) can be combined if we introduce a new matrix α' by the following prescription:

$$a'_{jj} \equiv a_{jj}(1+\lambda)$$

$$a'_{jk} \equiv a_{jk} \quad (j \neq k)$$

replacing equations (2.1) & (2.2) by

$$\sum_{l=0}^{M-1} a'_{kl}\delta a_l = \beta_k \tag{2.3}$$

Near the solution the *Levenberq-Marquardt* method uses *Newton*-type approximation by setting very small values of λ and equation (2.3) approaches equation (2.1). Equation (2.3) is then solved until there is no change in **a** (final set of fitted parameters). For very large λ, the diagonal terms of the curvature matrix dominate and the matrix equations degenerate into separate equations that yield a parameter correction vector in the same direction as the gradient of β vector in equation (2.2), or the opposite direction of the gradient of the merit function. Starting with a moderate λ value, if the solution is improved by the use of equation (2.3), λ is decreased – perhaps by a factor of 10. When this does not bring in the improvement of the merit function, that iteration is discarded, λ is increased, and equation (2.3) is applied again taking a shorter step in the downhill direction. The process is repeated until the convergence criteria are satisfied. The *Levenberg-Marquardt* method is computationally intensive because the first and second derivatives of the merit functions are involved, or to be exact, the evaluations of the *Jacobian* matrix are required. Several articles can be found where the *Levenberq-Marquardt* method has been enhanced to improve its ability to find the proper solution (Hill, 1998; Cooley and Naff, 1990). In general, two additional parameters are introduced: (1) *Marquardt*'s parameter (improves the performance for ill-conditioned problems); (2) the damping parameter (varies in the interval 0…1, ensures that relative parameter changes are smaller than some pre-specified values and damps oscillations that may occur when parameter vectors within two sequenced iterations define opposite directions).

In summary, the *LM* method described herein consists of the following steps:

1. Initialize parameters and *LM* method variables.

2. Solve the forward problem for model dependent variables.

3. Calculate the *Jacobian*, scale matrix and residual vector.

4. Calculate *Marquardt*'s parameter and upgrade the vector.

5. Update the damping factor and determine new parameter values.

6. If a termination criterion is met, terminate the calculation; otherwise go back to *Step 2*.

2.3 GA (Genetic Algorithm)

GA is a member of a class of probabilistic search algorithms based on Darwin's principle of the survival of the fittest (Goldberg, 1989; Cui and Kuczera, 2003). There are many research papers describing the application of *GA* to water resources and it is growing. Goldberg and Kuo (1987) were the first to apply *GA* to a pipe network optimization problem. From that timeframe the number of applications of GA to pipe network optimization have been reported (Murphy et al., 1993; Davidson and Goulter, 1995; Dandy et al., 1996; Savic and Walters, 1997; Montesinos et al., 1999). They found that *GA* were capable of finding acceptable solutions, and in certain cases, *GA* yielded better results than other optimization techniques. *GA* have been applied to several other kinds of problems, such as calibration of rainfall-runoff models (Wang, 1991; Franchini, 1996) and multi-objective groundwater modeling (McKinney and Lin, 1994; Cieniawski et al., 1995). *GA* has so far had little application in reservoir systems optimization. Esat and Hall (1994) applied a *GA* to a problem with four reservoirs. Fahmy et al. (1994) solved a reservoir system with a *GA* and concluded that *GA* had potential in application to large river basin systems. Oliveira and Loucks (1997) showed that *GA* can be used to identify effective operating solutions when they are applied to multi-reservoir systems. Sharif and Wardlaw (2000) showed that the *GA* is robust and can be easily applied to complex reservoir problems. The idea behind a *GA* is that a randomly generated set of decision variables are firstly created (Cui and Kuczera, 2003). These decision variables are usually encoded as strings of binary digits or real numbers. Each string, formed by concatenating the decision variable values, represents a single potential solution that can be used to evaluate the objective function. The entire population of such strings forms a generation. Genetic operators are applied to the current population to form the individuals for the next generation. This study employs the most commonly used operators, such as selection, crossover, and mutation, and the lesser-known operators, such as inversion and population selection. The population of designs is allowed to evolve through successive generations until stopping criteria are met.

The *Genetic Algorithm* has been developed further and some common implementations are parameters like *selection, crossover, inversion, population selection strategy*. As there is considerable scope for customizing the *GA*, hereby some details about the adaptations to the standard *GA* are given.

Selection
A tournament selection is used to generate the new members of the next generation. Three individuals are selected at random from the current population. Those individuals are then pitted against each other in a tournament, and the one with the best objective function value wins. The tournament produces parents who will then produce offspring using crossover and inversion operators.

Crossover
A crossover is the partial exchange of bits between two parent strings to form two offspring or child strings. Goldberg (1989) described many different methods of performing the *GA* crossover operation. The most straightforward is the one point crossover. The crossover occurs if $r < P_{\text{cross}}$, where r is a random number uniformly distributed between 0 and 1, and P_{cross} is a predefined probability for the crossover. It begins by selecting two strings at random from the population using tournament selection. A crossover point L is randomly selected, $2 \le L \le k-1$, where k is the length of the binary string. Two child strings are then created by swapping the parent substrings.

Inversion
An inversion operator is applied independently after the *GA* operators of the tournament, crossover, and mutation have produced the child population. The idea behind inversion is to produce orderings in which beneficial genetic material is more likely to survive. It provides a type of nondestructive noise that helps the crossover to escape local maxima (Whitley, 1987). For each string, if a random number is greater than the user defined probability P_{invert} , the string remains unchanged; otherwise, it is modified by the inversion operator. Under the inversion, two different bits along the decision string are chosen at random and swapped.

Population Selection Strategy
The crossover operator performs two complementary search functions — exploitation, which provides new decisions for further testing within the neighborhood of the best solution in the current population; and exploration, which introduces decisions from the whole search space into the population to maintain global diversity and thus avoid being trapped at local optima (Goldberg, 1989). The traditional *GA* theory almost completely places the burden of exploitation and exploration on the crossover, forcing a tradeoff between population diversity and selective pressure. Increasing the selective pressure tends to reduce diversity but it increases the search speed; decreasing the selective pressure helps to maintain diversity but it results in a slower, though more robust, search. According to Eshelman and Schaffer (1991), some of the burden of preserving good genetic material can be transferred from the crossover to the selection procedure. A suitable selection method, called the population selection strategy (*PSS*), merges the child and parent populations and selects the best individuals to make up the parent population for the next generation. This ensures the preservation of good genetic material from the parent population, thereby allowing the crossover to concentrate

on maintaining diversity. With this strategy in place, a very high crossover probability P_{cross} can be used to vigorously explore the decision space. This maximizes the chance of avoiding premature convergence on a flat region.

In summary, the customized *GA* used in this study consists of the following steps:

1. An initial population of *s* designs, each consisting of *n* decision variables ranging from 0 to 1, is randomly and uniformly sampled. The simulation model evaluates the objective function for each initial design. The decisions are encoded into binary strings.

2. A tournament selection is used to generate the new designs for the next generation. The *GA* terminates if all of the new designs are the same as the designs from the previous population.

3. A crossover is performed, decision by decision, for each pair of new designs with probability P_{cross}. Mutations are carried out with low probability on a bit-by-bit basis in the strings. An inversion is then performed with probability *P* invert for each decision.

4. The strings are then decoded into their real value equivalents and the objective function of the new design is evaluated.

5. The PSS method is used to form the new generation from the child and parent populations.

6. *Steps 2–5* are repeated until stopping criteria are satisfied. The search stops if the best design does not improve over five successive generations, or if the number of evaluations exceeds a predefined maximum (10,000).

2.4 HGA (Hybrid Genetic Algorithm)

The *HGA* stands for the *Hybrid Genetic Algorithm* and has been used by several authors in different fields of studies (water networks – Kapelan, 2002; aerospace industry – Vicini and Quagliarella, 1998). The *HGA* consists of two different search algorithms – the *GA* (*Genetic Algorithm*, see *Chapter 2.3*) and the *LM* (*Levenberg-Marquardt*, see *Chapter 2.2*). The idea is to utilise the *GA* and *LM* methods in the areas where their performance is best. In other words: the *GA* can be used for effective global search and the *LM* for effective local search. Such ways the time can be reduced that the *GA* spends on global search. Also, the error can be minimized that *LM* causes when its subjective choice of the starting point is badly chosen. The main problem of using the two-staged optimisation method that Kapelan (2002) points out is how to decide when to stop the *GA* and start using the *LM* method. When the *GA* does not get into the proximity of the global optimum

within a critical number of iterations (depending on the computer resources that are available), the *LM* method will fail to identify and it will climb to a local optimum instead of a global one.

To improve the *HGA* local search (comparing with the *GA*), a new operator is introduced. Every certain number of generations, an intermediate population is created from the existing one by using standard *GA* operators (crossover, mutation, etc.). After that, in order to create a new population, few chromosomes are selected from that intermediate population and transformed by application of the local search operator (*LSO*) with pre-specified probability rate. The *LSO* operator then transforms analysed chromosomes by applying the *LM* method for the number of iterations (typically one or two). Chromosomes for the *LSO* transformation may be selected by following strategies: (a) with best fitnesses; (b) using the existing *GA* selection operator (roulette-wheel, etc.); (c) select randomly.

In general the *HGA* has the following steps:

1. Create initial population and evaluate fitnesses.

2. Number of generations and count equals 0 (starting value).

3. Create new population using standard *GA* operators: selection, crossover, mutation.

4. Update the value of parameters: number of generations and count.

5. If the count equals the number of generations where intermediate population is generated, go to *Step 6*, otherwise continue from *Step 11*.

6. Select a number of chromosomes to be transformed using one of three *HGA* strategies.

7. Decode the selected chromosomes.

8. Apply the *LSO* operator with pre-specified probability.

9. Encode the selected chromosomes with new fitnesses.

10. Count parameter equals 0.

11. Termination of the *GA*? If the answer is no, go back to *Step 3*, otherwise to *Step 12*.

12. Stop calculation.

2.5 SCE (Shuffled Complex Evolution Method)

The *SCE* algorithm for function optimization was developed by Duan et al. (1992; 1994). This method was originally designed to deal with the difficult problems encountered in the calibration of conceptual watershed models. It incorporates the best features from several existing methods, including competitive evolution, the combination of random and deterministic strategies, the concepts of controlled random search, and complex shuffling. In essence, the *SCE* begins with an initial population of points sampled randomly from the feasible space. The population is partitioned into one or more complexes, each containing a fixed number of points. Each complex is allowed to evolve based on a competitive evolution technique that uses the simplex method to direct the search in the correct direction. Periodically, the entire population is shuffled and points are reassigned to new complexes to enable information sharing. This shuffling strategy reduces the chance of complexes being trapped on flat regions and thus converging prematurely. As the search progresses, the entire population tends to converge toward the neighborhood of the global optimum, provided the initial population size is sufficiently large.

A more formal description of the algorithm follows.

1. Generate initial samples. Sample s = p*m designs x1 , x2 ,...,xS, randomly from the search space using a uniform distribution, where p is the number of complexes, each having m = 2n + 1 designs, with n being the number of decision variables.

2. Calculate the objective function value f i of each design xi .

3. Rank designs. Sort all *s* designs in ascending order of their corresponding objective function values and store them in an array $\mathbf{D} = \{x_{(i)}, f_{(i)}, i = 1,...,s\}$, so that $i = 1$ represents the design with the smallest function value.

4. Partition into complexes. Sequentially assign D into p complexes A1, A2,...,Ap such that D1 is assigned to Complex 1, D2 is assigned to Complex 2, and so on. Each complex contains m designs.

5. Evolve each complex Ak (k = 1,...,p) independently using the competitive complex evolution algorithm, which is based on the *Nelder-Mead* simplex algorithm and is described in detail by Duan et al. (1992).

6. Shuffle complexes. Insert the new A1, A2,...,Ap into D and sort D in order of increasing objective function value.

7. If the stopping criteria are not met, go to Step 4. The SCE search will stop if the number of evaluations exceeds a predefined maximum (10,000), or if the

smallest objective function value does not improve within a specified tolerance, here, 10-6, over two consecutive generations.

The *SCE* method contains many probabilistic and deterministic components that are controlled by algorithmic parameters. Similar to the *GA* method, these parameters have to be chosen carefully. Duan et al. (1994) presented the results of several experimental studies using different algorithmic parameter setups. The most sensitive algorithmic parameter is the number of complexes *p*. The experiments made by Duan et al. (1992; 1994) showed that the greater the degree of difficulty of the problem, the larger the number of complexes required to find the global optimum.

2.6 Summary and conclusions

Obviously different algorithms are developed for various problem solving tasks. All have their advantages and disadvantages as did the ones that were described earlier. *Table 2.1* lists some of those.

Table 2.1 Main advantages and disadvantages of commonly used algorithms

	LM	GA	HGA	SCE
Main advantages	Fast and efficient local search in terms of CPU time	Multi-modal search space, quick in global search	Reduced CPU search time for finding an optimal solution	Can reliably find the global minimum in the parameter space
		Reduced chance to get into the local optimum	The variance of best fitnesses from multiple runs is smaller than with GA	
		No need to calculate derivatives, or perform any numerically demanding operations like matrix inversions	More accurate calibration parameter values because of using combination of algorithms	
		Both discrete and continuous decision variables can be used		
		Can be used for generating a population of good solutions		
		Parallel calculations (multi processors/PC) routines available		
Main disadvantages	Can be easily trapped into a local optimum	CPU time expensive	Higher number of parameters than in GA	Difficult to find a unique "best" parameter set
	Can improve only a single solution at any one time	Multiple runs should be done to ensure the global optimum	Sensitivities need to be calculated due to involvement of LM method	The tendency to collapse to a single region of attraction (global minimum)
	Requires calculations of derivatives that may cause numerical error and is also time consuming	High number of model parameters that needs to be tuned to get the best out of GA		CPU time expensive; 25% less than GA (when compared single CPU-s)
	Requires a good starting point for finding global not local optimum	Penalty function should be carefully chosen as it affefcts the GA search performance		Limited capabilities to develop parallel computing (for example to exploit parallelism within each complex)
		Slow in local search		
		The risk of premature convergence		

3. Probabilistic (stochastic) modelling approach

3.1 Introduction

The development of methods for the detection of leaks in pipelines and pipe networks continues to be a very active research area. Basically all of them have the same underlying structure – to solve an inverse problem subject to parametric uncertainty and measurement error. In general the problem is formulated as an optimisation. From the literature review it is not hard to see the deficiency in an optimization approach even in the cases where sufficient resources are available to ensure that a solution is found and that solution is indeed a global optimum in the parameter space. In this thesis a pipe network model is parameterized by leak coefficients at each node (see *Chapter 4*). Pressure observations are supplied for a subset of the nodes (flow measurements can be used as well) and the sum of the squared differences between the observations and the model output is minimized over the possible values of the parameters. The result shows us a point estimate for the leak coefficients. This point estimate conveys no measure of uncertainty. Although there are generalizations that use the curvature of the optimand to infer a variance matrix, this type of approximation is only valid in the presence of large amounts of effectively independent data – something that is seldom in case of pipe networks (Rougier, 2005).

The key point of the thesis is that we use the probabilistic approach to include the uncertainty factor in leak detection studies. This is mainly done because a leak is only the first stage in a complex decision how to proceed. The *Decision theoretic framework* in which actions are selected based on both probabilities and consequences can be used to make potentially costly decisions under uncertainty.

The key result is to compute a joint probability distribution over leak location and size taking account of parametric uncertainty and measurement errors. An expert knowledge about the ways how the particular pipe network might leak is also taken account. In general every optimization approach is actually a probabilistic approach but with very strong (or in other words, unrealistic) assumptions about the distributions of the uncertain quantities. In this thesis the probabilistic approach is based on the *Bayesian inference schema*.

3.2 Bayesian networks

Contrary to deterministic models, which try to build associative relationships between the input variable, the main underlying concept of Bayesian networks is the causal relationship. In probability calculus, causal relationships are represented by the notion of conditional independence. Bayes' theorem states that (Box and Tiao, 1973)

$$P(\theta \mid y) = \frac{P(y \mid \theta)P(\theta)}{P(y)} \tag{3.1}$$

where $P(\theta)$ = prior probability of θ with no knowledge of observation *y*, also called prior distribution of θ, or the distribution of θ *a priori*; $P(y)$ = prior probability of *y* with no knowledge of θ; $P(\theta \mid y)$= posterior probability of θ after observation *y*, also called the posterior distribution of θ given *y*, or the distribution of θ *a posteriori*; $P(y \mid \theta)$ = probability of observation *y* given that θ is true. The ratio $\frac{P(y \mid \theta)}{P(y)}$ is called the normalized likelihood. The *Bayesian* prior probability $P(\theta)$ is updated to a posterior probability $P(\theta \mid y)$ that reflects an observation *y*. Given the data y, $P(y \mid \theta)$ may be regarded as a function not of y but of θ. It is called the *likelihood function* of θ for given y and can be written $l(\theta \mid y)$. We can thus write the *Bayes'* formula as

$$P(\theta \mid y) = l(\theta \mid y)P(\theta) \tag{3.2}$$

Bayes' theorem states that the probability distribution for θ posterior to the data y is proportional to the product of the distribution for θ prior to the data and the likelihood for θ given *y*. In other words we can write:

posterior distribution $\propto$ *likelihood* $\times$ *prior distribution*

The likelihood function $l(\theta \mid y)$ plays a very important role in *Bayes'* formula (Box and Tiao, 1973). Through this function the data *y* modifies prior knowledge of θ, therefore it can be regarded as representing the information about θ coming from the data. The likelihood function is defined up to a multiplicative constant, meaning that multiplication by a constant leaves the likelihood unchanged. This is also in accord with the role it plays in *Bayes'* formula, since multiplying the likelihood function by an arbitrary constant will have no effect on the posterior distribution of θ. The constant will cancel upon normalizing the product on the

right hand side of equation (3.2). It is only the relative value of the likelihood which is of importance. Equation (3.2) is appealing because it provides a mathematical formulation of how previous knowledge may be combined with the new knowledge. It allows us to continually update information about a set of parameters θ as more observations are taken. Let's have an example. Suppose that there is an initial sample of observations y_1, the *Bayes'* formula gives:

$$P(\theta \mid y_1) \propto P(\theta) l(\theta \mid y_1) \tag{3.3}$$

Supposing that the second sample of observations y_2 is distributed independently of the first sample, then

$$P(\theta \mid y_2, y_1) \propto P(\theta) l(\theta \mid y_1) l(\theta \mid y_2) \propto P(\theta \mid y_1) l(\theta \mid y_2) \tag{3.4}$$

The expression (3.4) is precisely of the same form as equation (3.3) except that $P(\theta \mid y_1)$, the posterior distribution for θ given y_1, plays the role of the prior distribution for the second sample. Without doubt this process can be repeated any number of times. In other words having *n* independent observations, the posterior distribution, if desired, can be recalculated after each new observation. Thereafter at the *m*th stage the likelihood associated with the *m*th observation is combined with the posterior distribution of θ after *m-1* observations giving the new posterior distribution:

$$P(\theta \mid y_1, \ldots, y_m) \propto P(\theta \mid y_1, \ldots, y_{m-1}) l(\theta \mid y_m), \quad m = 2, \ldots, n \tag{3.5}$$

where

$$P(\theta \mid y_1) \propto P(\theta) l(\theta \mid y_1).$$

In general *Bayes'* theorem describes, in a fundamental way, the process of learning from experience, and shows how knowledge about the state of nature represented by θ is continually modified as new data becomes available (Box and Tiao, 1973).

Conceptually this updating process mimics the reasoning of people presented with new information about uncertain phenomena (Dawsey et al., 2006). For a classic illustration of this idea, originally presented in The Economist (2000), consider a newborn infant on her first day. After observing her first sunset, she wishes to determine the probability that the sun will rise again using *Bayes'* theorem. The infant puts a black marble and a white marble into a bag to represent her initial estimate that there is an equal probability that the sun will rise again or not. Each day that the sun rises, she puts another white marble into the bag. After one morning, the probability of sunrise increases from 0.5 (1 white marble / 2 total marbles) to 0.75 (2 white marbles / 3 total marbles). The probability is 0.8 after the second day, 0.833 after the third, and so on until the child has near certainty that the

sun will continue to rise every day. In much the same way, the prior probability of a hypothesis, $P(\theta)$ is revised to a posterior probability after an observation has been made $P(\theta \mid y)$.

Bayesian networks are a good solution when taking into account the pipe history and its characteristics (Babovic and Drecourt, 2000). They are based on *Bayes'* theorem and model causal relationships between variables. The relationship is represented by an arrow from the cause to the consequence. If y causes θ, variable y is called a *parent* and variable θ is called a *child*. When variable θ has many parents $y_1,...,y_n$, these can interact with each other. Therefore, it is necessary to get a specification of $P(\theta \mid y_1,...,y_n)$.

A Bayesian network is defined as follows:

- A set of variables and a set of directed edges between variables
- Each variable has a finite set of mutually exclusive states
- The variables together with the edges from a directed acyclic graph (i.e. there is no directed path that leads from y_1 to y_n such that $y_1 = y_n$).
- To each variable θ with parents $y_1,...,y_n$, there is attached a conditional probability table $P(\theta \mid y_1,...,y_n)$.

The main force of the *Bayesian* networks is that they use a kind of relationship that is intuitive for the human brain: causality. This makes it possible to ask many people from different horizons to contribute to the model. The second interesting aspect of *Bayesian* networks, which makes them especially adapted to different kinds of problems, is that any cause to the parameter change can be taken in account. Even if no data is available, a specialist can provide information, which can be used for building the network. Both knowledge from data and experience can be gathered in the same network to build one model to predict unknown parameters/occurrences (for example pipe bursts).

The *Bayesian* network is found to be useful also in power system analysis. The case study with the pipe type cable at high pressure is performed in Tylman and Anders (2006).

There are a number of exact and approximate algorithms that have been developed to infer posterior probabilities for *Bayesian networks*. In this thesis the *Shuffled Complex Evolution Metropolis* (*SCEM-UA*) algorithm is used and accommodated for the inverse problem of leak detection in pipe networks.

3.3 SCEM-UA algorithm

The *Shuffled Complex Evolution Metropolis* (*SCEM-UA*) algorithm (Vrugt et al., 2003a) was developed in collaboration between the *University of Amsterdam* (*The*

Netherlands) and the *University of Arizona* (*USA*). *SCEM-UA* is a modification of the global optimisation method *SCE-UA* (The *Shuffled Complex Evolution* algorithm, see *Chapter 3*) developed by Duan et al. (1992). *SCEM-UA* is, fundamentally, a *Markov Chain Monte Carlo* (*MCMC*) sampler. The candidate point is generated from a previous iteration point using some proposal distribution (multivariate normal distribution is used here). The *SCEM-UA* algorithm operates by merging the strengths of the *Metropolis* algorithm, controlled random search, competitive evolution, and complex shuffling in order to continuously update the proposal distribution and evolve the sampler to the posterior target distribution (Vrugt et al., 2003a).

The goal in hydrological models is to predict outputs from inputs. These models are indexed by parameters, which may (or may not) be physically interpretable. Assuming that the mathematical structure of the model is essentially predetermined and fixed, the classical approach to model calibration is to find the best attainable values of the parameters such that the vector of error terms is as close to zero as possible. The general form of that model can be written:

$$y^* = f(x \mid \theta) + e \tag{3.6}$$

where $f(x \mid \theta)$ = a vector of model outputs; x = a vector of model inputs; θ = a vector of model parameters that need to be calibrated; y^* = a vector as model measured (i.e. observed) output values; e = a model predictions errors. The goal of searching for a single optimal representation of equation (3.6) is, however, questionable (Vrugt et al., 2003a). Most likely when a search is carried out in the feasible parameter space that is close to the global optimum it will reveal many behavioural parameter sets with quite similar performance in reproducing the observed data. While classical statistics consider the model parameters θ in equation (3.6) to be fixed but unknown, the *Bayesian* statistics treat these values as probabilistic variables having a joint posterior *probability density function* (*PDF*). *PDF* captures the probabilistic beliefs about the parameters θ in the light of the observed data y^*. The posterior PDF $P(\theta \mid y)$ is proportional to the product of the likelihood function and the prior PDF (see *Chapter 3.2*). This prior PDF with probability density (or mass) function $P(\theta)$ summarizes information about θ before any data are collected. Normally the prior information is expressed with some realistic lower and upper bounds on each of the parameters, thereby defining the feasible parameter space giving a uniform (noninformative) prior distribution on this rectangle.

Supposing that the residuals are mutually independent, *Gaussian* distributed, with constant variance, the likelihood of a parameter set $\theta^{(t)}$ for describing the observed data y^* can be computed using the following equation:

$$l\left(\theta^{(t)} \mid y^*\right) = \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{e_i\left(\theta^{(t)}\right)}{\sigma}\right)^2\right] \tag{3.7}$$

Box and Tiao (1973) showed that the influence of σ can be integrated out when a noninformative prior is assumed in the form of $P(\theta) \propto \sigma^{-1}$. And the following form of the posterior density of $\theta^{(t)}$ can be reached:

$$P\left(\theta^{(t)} \mid y^*\right) \propto \left[\sum_{i=1}^{N} e_i\left(\theta^{(t)}\right)^2\right]^{-\frac{1}{2}N} \tag{3.8}$$

More information about the *Bayesian inference scheme* can be found from Box and Tiao (1973).

Working with non-linear models the surface of $P(\theta \mid y)$ might have multiple local optima and discontinuous derivatives (Duan et al., 1992) and an explicit expression of the joint and marginal probability density functions is often not possible. Therefore the *MCMC* samplers are used. These are very well suited to deal with the peculiarities encountered in the posterior *PDF* (Vrugt et al., 2003a). *Markov Chain* schemes give us a general approach for sampling from the posterior probability distribution $P(\theta \mid y)$. The most general and earliest *MCMC* algorithm is known as the *Metropolis-Hastings* (*MH*) algorithm. *Hastings* took the sampling method introduced by *Metropolis* in 1953 (Metropolis et al., 1953) and modified it to get a more generalized one (Hastings, 1970). The basic *MH* algorithm set-up is given as follows (Vrugt et al., 2003a):

1. Randomly start at a location in the feasible parameter space, $\theta^{(t)}$, compute the posterior density, $P\left(\theta^{(t)} \mid y\right)$, relevant to this point according to equation (3.7) or (3.8).
2. Generate a new configuration $\theta^{(t+1)}$ from $z\left(\theta \mid \theta^{(t)}\right)$, where $\theta^{(t+1)}$ is called a candidate point and z is called the proposal distribution.
3. Evaluate $P\left(\theta^{(t+1)} \mid y\right)$ using equation (3.7) or (3.8) and compute $\Omega = P\left(\theta^{(t+1)} \mid y\right) / P\left(\theta^{(t)} \mid y\right)$.
4. Randomly sample a uniform label Z over the interval 0 to 1.
5. If $Z \leq \Omega$, accept the new configuration. Reject the candidate point, when $Z > \Omega$ and remain at the current position, that is, $\theta^{(t+1)} = \theta^{(t)}$.
6. Increment t. If t is less than a prespecified number of draws, return to *Step 2*.

The *MH* algorithm always allows candidate points (jumps) into a region of higher posterior probability but it will also explore regions with lower posterior probability

with probability Z. The proposal distribution z is known to be very crucial for the convergence properties of the *Markov Chain*. When the proposal distribution is too large, too many candidate points are rejected, and therefore the chain slowly covers the target distribution. Controversially when the proposal distribution is too small, too many candidate points are accepted, and the chain traverses slowly through the parameter space, thereby resulting in slow convergence. Therefore an important challenge is to design samplers that rapidly converge to the global minimum in the parameters space, while maintaining sufficient occupation of the lower posterior probability regions of the parameters space (Vrugt et al., 2003a). The *SCEM-UA* algorithm is a modification from the *SCE-UA* algorithm (see *Chapter 2.5*). The main goal for the new algorithm was to prevent the collapse of the algorithm into the relatively small region of a single best parameter set (it is typical for many evolutionary search algorithms). The basic algorithm set-up is as follows:

1. Generate sample. Sample s points $\{\theta_1, \theta_2, \ldots, \theta_s\}$ randomly from the prior distribution and compute the posterior density $\{P(\theta^{(1)} \mid y), P(\theta^{(2)} \mid y), \ldots, P(\theta^{(s)} \mid y)\}$ of each point using equation (3.7) or (3.8).
2. Rank points. Sort the s points in order of decreasing posterior density and store them in array $D[1:s, 1:n+1]$, where n is the number of parameters, so that the first row of D represents the point with the highest posterior density.
3. Initialize parallel sequences. Initialize the starting points of the parallel sequences, $S^1, S^2, \ldots, S^q$, such that S^k is $D[k, 1:n+1]$, where $k = 1, 2, \ldots, q$.
4. Partition into complexes. Partition the s points of D into q complexes $C^1, C^2, \ldots, C^q$, each containing m points, such that the first complex contains every $q(j-1)+1$ ranked point, the second complex contains every $q(j-1)+2$ ranked point of D, and so on, where $j = 1, 2, \ldots, m$.
5. Evolve each sequence. Evolve each of the parallel sequences according to the *Sequence Evolution Metropolis* (*SEM*) algorithm outlined later in this chapter.
6. Shuffle complexes. Unpack all complexes C back into D, rank the points in order of decreasing posterior density, and reshuffle the s points into q complexes according to the procedure specified in *Step 4*.
7. Check convergence. Check the *Gelman and Rubin* (*GR*) convergence statistic. If the convergence criteria are satisfied, stop; otherwise return to *Step 5* (see *Chapter 4* for more information about the *GR* convergence statistic).

The increased initial random sample provides an extensive exploration of the parameter space, increasing the chance of finding the global optimum of the prescribed density function (Vrugt et al., 2003a). On the other hand, the use of a

number of parallel sequences with different starting points enables an independent exploration of the search space, thereby allowing that the optimization problem has more than one region of attraction and enables the use of heuristic tests to judge whether convergence of the sequences to a limiting distribution has been achieved. During the evolution process the use of complexes enables the collection of information gained about the search space by each individual sequence. And if these sequences are shuffling, it enhances the survivability of the sequences by a global sharing of the information gained independently by each parallel sequence. In overall, this series of operations results in a robust *MCMC* sampler that conducts a robust and efficient search of the parameter space.

The *SEM* algorithm (see *Step 5* in *SCEM-UA*) produces new candidate points in each of the parallel sequences S^k by generating draws from an adaptive proposal distribution by using the information induced in the m samples of C^k. An outline of the *SEM* algorithm is given as follows (Vrugt et al., 2003a):

1. Compute the mean μ^k and covariance structure $\sum^k$ of the parameters of C^k. Sort the m point in complex C^k in order of decreasing posterior density and compute the ratio of the posterior density of the first (*"best"*) to the posterior density of the last (*"worst"*) member of C^k.
2. Compute α^k, the ratio of the mean posterior density of the m points in C^k to the mean posterior density of the last m generated points in S^k.
3. If α^k is smaller than a predefined likelihood ratio, T, generate a candidate point, $\theta^{(t+1)}$, by using a multinormal distribution centred on the last draw, $\theta^{(t)}$, of the sequence S^k, and covariance structure $c_n{}^2 \sum^k$, where c_n is a predefined jumprate. Go to *Step 5*, otherwise continue with *Step 4*.
4. Generate offspring, $\theta^{(t+1)}$, by using a multinormal distribution with mean μ^k and covariance structure $c_n{}^2 \sum^k$, and go to *Step 5*.
5. Compute the posterior density, $P\left(\theta^{(t+1)} \mid y\right)$, of $\theta^{(t+1)}$ using equation (3.7) or (3.8). If the generated candidate point is outside the feasible parameter space, set $P\left(\theta^{(t+1)} \mid y\right)$ to zero.
6. Compute the ratio $\Omega = P\left(\theta^{(t+1)} \mid y\right) / P\left(\theta^{(t)} \mid y\right)$ and randomly sample a uniform label Z over the interval 0 to 1.
7. If $Z \le \Omega$, accept the new candidate point. Reject the candidate point, when $Z > \Omega$ and remain at the current position in the sequence, that is, $\theta^{(t+1)} = \theta^{(t)}$.
8. Add the point $\theta^{(t+1)}$ to the sequence S^k.
9. If the candidate point is accepted, replace the best member of C^k with $\theta^{(t+1)}$, and go to *Step 10*; otherwise replace the worst member (m) of C^k with

$\theta^{(t+1)}$, provided that Γ^k is larger than the predefined likelihood ratio, T, and $P\left(\theta^{(t+1)} \mid y\right)$ is higher than the posterior density of the worst member of C^k.

10. Repeat the *Steps* 1 – 8 L times, where L is the number of evolution steps taken by each sequence before complexes are shuffled.

Candidate points in the *SEM* algorithm are generated using an adaptive multinormal proposal distribution with the mean identical to the current draw in the sequence and the covariance matrix corresponding to the structure induced in the m points of complex k (Vrugt et al., 2003a). It should be noted that when the mean posterior density of the last m generated points in sequence k is significantly smaller than the mean posterior density of the m points in the corresponding complex k, the centre of the proposal distribution is temporarily switched to the mean of the points in the complex. This behaviour in the *SEM* algorithm (*Step 4*) significantly reduces the chance that individual sequences get stuck in a local non-productive region of attraction causing further improvement in the mixing of the sequences. After generating a new candidate point, the posterior density relevant to this point is computed and the *Metropolis-annealing* (Metropolis et al., 1953) criterion is used to judge whether the candidate point should be added to the current sequence or not. In *Step 9* it is decided which member of the current complex k should be replaced with the point $\theta^{(t+1)}$. When the candidate point is accepted, $\theta^{(t+1)}$ automatically replaces the best member of the complex. Controversially, when the candidate point is rejected, $\theta^{(t+1)}$ replaces the worst point in complex k provided that Γ^k is larger than the predefined likelihood ratio, T, and the posterior density relevant to $\theta^{(t+1)}$ is higher than the posterior density corresponding to the worst point of complex k. Therefore when Γ^k is larger than a prior defined large number (i.e., $T > 10^5$), there is sufficient reason to believe that the covariance of the proposal distribution is specified too big as members with a too low probability are still present in C^k. The replacement of the worst member of C^k in this particular situation will facilitate convergence to a limiting distribution.

The *SCEM-UA* algorithm differs from the traditional *MCMC* samplers: it is an adaptive sampler, where the covariance of the proposal or sampling distribution is periodically updated in each complex during the evolution to the posterior target distribution using the information from the sampling history induced in the transitions of the generated sequences (Vrugt et al., 2003a). The *SCEM-UA* algorithm is different from the original *SCE-UA* algorithm presented by Duan et al. (1992) (see *Chapter 2.5*) in two important ways:

1. The downhill *Simplex* method in the competitive complex evolution algorithm outlined by Duan et al. (1992) is replaced by a *Metropolis-annealing* covariance based offspring approach, thereby avoiding a deterministic drift toward a single mode.
2. The *SCEM-UA* algorithm does not further subdivide the complex into subcomplexes during the generation of the offspring (candidate points) and

uses a different replacement procedure, to counter any tendency of the search to terminate occupations in the lower posterior density region of the parameter space.

The *SCEM-UA* algorithm has been mainly used to solve hydrological problems (Vrugt et al., 2003a; 2003b; 2006). In this thesis *SCEM-UA* is applied to pipe networks to solve the problem of leak detection. In *Chapter 6* the algorithm core elements (including those that are specifically designed for the leakage problem) are laid out more specifically. The main advantages and disadvantages of using *SCEM-UA* are:

Advantages	Disadvantages
- Generates explicit estimates of parameter uncertainty in a single optimization run (estimates the entire posterior parameter distribution); - Generates prediction uncertainty bounds on the model simulations; - Generates improved estimates of parameter sensitivity and correlation in the full parameter space. This information is very useful to estimate what complexity is warranted by the calibration data; - Generates useful estimates of the most informative measurements and time varying parameters in the recursive mode.	- The method is computationally demanding, especially for high-dimensional state/parameter estimation problems; - The method relies on correct model structure and *Bayesian* statistics. - The method ignores input, output and model structural uncertainty. In principle this could be incorporated in the identification procedure.

4. Leakage detection methodology

A water distribution system (*WDS*) hydraulic model is defined by a mass conservation equation written for each node in the network and an energy conservation equation written for each link (i.e. pipe) in the network:

$$\sum_j Q_{ij} = Q_D^i + Q_L^i \tag{4.1}$$

$$H_i - H_j = K_{ij} Q_{ij} \left|Q\right|_{ij}^{n-1} \tag{4.2}$$

where Q_{ij} = the flow in the link (i.e. pipe) connecting nodes *i* and *j*; Q_D^i = demand (known outflow) at node *i*; Q_L^i = leak at node *i* (leaks are assumed to occur at network nodes only); H_i = total head at node *i*; K_{ij} = pipe resistance coefficient; n is the energy loss equation exponent (1.85 for the *Hazen-Williams* and 2 for the *Darcy-Weisbach* headloss equation). In the literature, several different approaches exist to define leak flow Q_L^i in equation (4.1). Leaks are usually modelled as pressure dependant flows. A number of papers exist in the literature which try to define the effect of pressure on leakage more precisely by analysing the behaviour of pipes made of different materials and leaks with different shape openings (Almandoz et al., 2005; Van Zyl and Clayton, 2005). Please see *Appendix A* for more information about different interpretations of the leakage models. Here, the following leakage model is used (Germanopoulos, 1985):

$$Q_L^i = K_i \cdot \left(H_i - z_i\right)^\beta \tag{4.3}$$

where $K_i = C_{d,i} A_i \cdot (2g)^\beta$; $C_{d,i}$ = leak discharge coefficient; A_i = leak area (the product of $C_{d,i}$ and A_i is also known as the effective leak area) ; g = gravitational constant; H_i = total head at leak node *i;* z_i = node *i* elevation (assuming atmospheric water pressure in the surrounding soil); and β = leak/emitter exponent. The value of β typically varies between 0.5 and 2.8 with a median of 1.15 (WRc, 1994). Germanopoulos (1985) identified a mean value of β to be 1.18. In the case studies shown here, different leakage exponents are used (ranging from 0.5 to 1.1).

The leakage-based *WDS* hydraulic model defined in equations (4.1)-(4.3) can be conceptualised as follows:

$$y = f(x \mid \theta) \tag{4.4}$$

where y = a vector of model outputs (nodal heads H_i and pipe flows Q_{ij}); x = a vector of all *WDS* hydraulic model inputs (e.g. network configuration data, various pipe data, nodal demands, etc.) ; θ = a vector of model parameters that need calibrating (e.g. leak areas A_i). Introducing a vector y^* as model measured (i.e. observed) output values, *WDS* model prediction errors e can be defined as follows:

$$e = y^* - y \tag{4.5}$$

Equations (4.4) – (4.5) can be rewritten as (recalling equation 4.6):

$$y^* = f(x \mid \theta) + e \tag{4.6}$$

The above equation represents the conceptual formulation of the non-linear regression approach. In the approach used here, parameter values a (i.e. leak areas) are assumed to be random variables following a joint posterior probability density function (*PDF*). The prior *PDF*s of parameters capture the probabilistic beliefs about parameter θ values prior to solving the inverse problem. Here, non-informative prior parameter *PDF*s are assumed to be uniformly distributed on the interval defined by the assumed low and high search bounds. Informative prior parameter *PDF*s are based on the *Gaussian* distribution with assumed mean and standard deviation values. The posterior parameter *PDF*s capture the probabilistic beliefs about parameter values in the light of the observed data y^*.

Assuming that the residuals e are mutually independent, each having the exponential power density $E(\sigma, \gamma)$, the likelihood of a parameter set θ for describing the observed data y^* can be computed using the following equation (Box and Tiao, 1973):

$$P(\theta \mid y^*, \gamma) \sim \left[\sum_{i=1}^{N} \left(\frac{e_i(\theta)}{\sigma} \right)^{2/(1+\gamma)} \right]^{-N(1+\gamma)/2} \tag{4.7}$$

where parameter γ takes the value of 0 when residuals are assumed to be normally distributed, γ = 1 when residuals are double exponentially distributed and, finally, $\gamma \to -1$ when residuals tend to a uniform distribution. In the case of a non-informative prior *PDF* of θ in the form $p(\theta, \sigma \mid \gamma) \sim \sigma^{-1}$, Box and Tiao (1973) showed that the influence of σ can be integrated out, leading to (in case of *Gaussian* residuals $\gamma = 0$) a simple least squares problem:

$$\textit{Minimise} \quad E = \sum_{i=1}^{N} e_i^2 , \tag{4.8}$$

where E is total calibration error.

The leakage detection methodology presented here is defined as an optimisation-based inverse problem driven by the objective defined in equation (4.8) subject to the following set of constraints: (a) *WDS* hydraulic model equations (4.1) – (4.3) and (b) assumed prior *PDF*s for all calibration parameters (leak areas). In *Section 5.2.2* it is described how the leakage model is incorporated into optimization process.

5. Software development

In this section the software environment is described. The software application consists of three main parts. Firstly, the freely available *EPANET DLL* library is described, followed by Matlab programming environment. The third part, the *SCEM-UA* algorithm, is finally overdrawn. The *SCEM-UA* algorithm is modified inside the *Matlab* environment where leakage specific subsections are also added for current studies. All this is bound into a user-interface system (developed for the current work) that is meant to help to tune the different parameters for the future analysis.

5.1 Matlab environment

EPANET is a program that analyzes the hydraulic and water quality behaviour of water distribution systems (Rossman, 2000). A distribution system or simply a network consists of pipes, nodes (pipe junctions), pumps, valves and storage tanks or reservoirs. *EPANET* tracks the flow of water in each pipe, the pressure at each node, the height of water in each tank, and the concentration of a chemical species throughout the network during a simulation period comprised of multiple time steps. In this thesis the water quality aspect is left out, and only the hydraulics aspect is used in the leak detection analysis. *EPANET* is a free software, developed by the *Water Supply and Water Resources Division of the U.S. Environmental Protection Agency's National Risk Management Research Laboratory* (*US EPA*) and can be downloaded from *World Wide Web* (US EPA). *EPANET* uses the so called *"Gradient Method"* to solve the flow continuity and headloss equations that characterize the hydraulic state of the pipe network at a given point in time (Rossman, 2000). Besides free availability of the EPANET software it also allows us to customize its features (or connect to another interface/software) through the *EPANET Programmer's Toolkit* (US EPA). It is a dynamic link library (*DLL*) of functions that allows developers to customize *EPANET's* computational engine for their own specific needs. The functions can be incorporated into 32-bit *Windows* applications written in *C/C++*, *Delphi Pascal*, *Visual Basic*, or any other language that can call functions within a *Windows DLL*. The *Programmer's Toolkit* file is named *EPANET2.DLL* and is distributed with *EPANET*. Additionally the *Toolkit* comes with several different header files, functions definition files, and *.lib* files that simplify the task of interfacing it with *C/C++*, *Delphi* and *Visual Basic* code. In this thesis we use *MATLAB* to connect the *EPANET* dynamic link library and the *SCEM-UA* algorithm.

5.2 Matlab environment

To solve the leakage detection problem, the *SCEM-UA* source code (SCEM-UA) was modified and linked to the *EPANET2* software in the *MATLAB 7.3* programming environment. The basic connection schema is shown in *Figure 5.1*. Hereby the subsections of this environment are described.

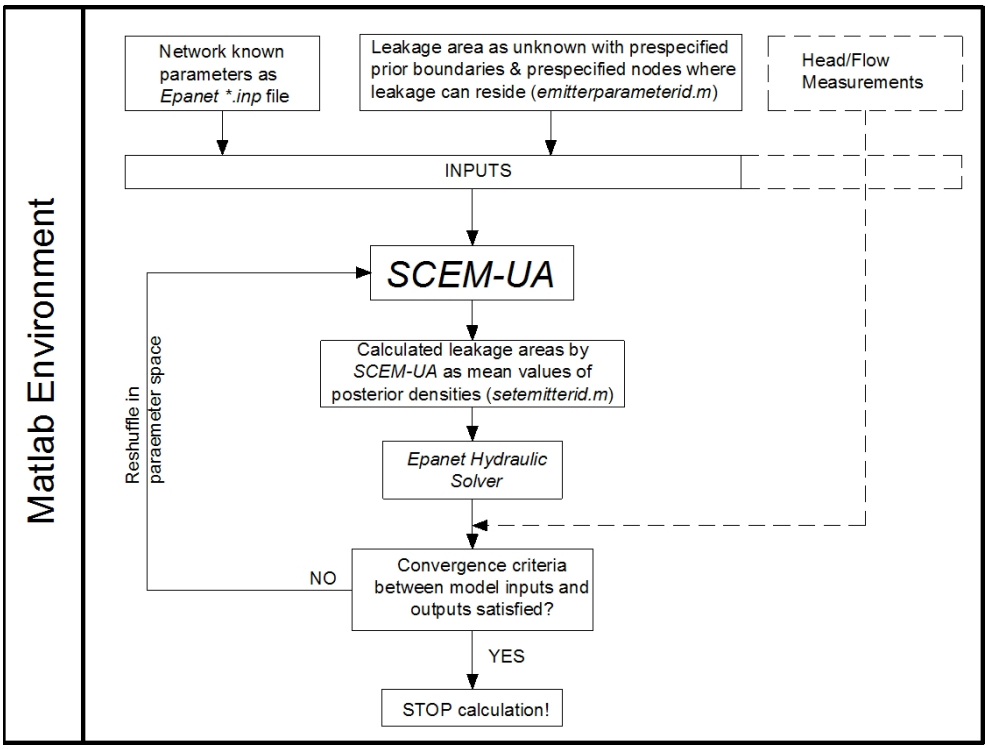

Figure 5.1 EPANET and SCEM-UA connection through MATLAB programming environment

5.2.1 SCEM-UA layout

In this section we describe the implementation of the *SCEM-UA* algorithm in a systematic workstation-based calibration strategy. This version of the *SCEM-UA* source code is a modified version from the original one, created by Jasper A. Vrugt (*University of Amsterdam, Netherlands*). The base version of this can be downloaded from the *World Wide Web* (SCEM-UA). The modifications were carried out by the *Centre for Water Systems* (*University of Exeter, UK*) and by the author. These modifications were needed to link *SCEM-UA* to the *EPANET2* software (Rossman, 2000) in the *MATLAB* programming environment. The basic block scheme about software links is presented in *Figure 5.1*. Now the attention is

paid to file levels (*Matlab .m* files, *Epanet library* files and parameters files). The simplest description of the three different file levels is presented in *Table 5.1*.

Table 5.1 File types used in programming environment

File level (type)	**Description**
MATLAB .m files	Programming environment. All procedures are described in these files: File readings, executing the model (*SCEM-UA*, *SEM* algorithms). Leak modelling definitions are also described in these files. Results are written into *.mat* files.
EPANET library files	Defines the data structures that need to be followed in *MATLAB*. Used for data analysis through hydraulic solver.
Parameters files (network *.inp*, measurements *.dat*)	Reading network parameters, measurements (pressures, flows). The final output of the model is given with the same filetype as its input (*.inp*)

The descriptions of the *SCEM-UA* and *SEM* algorithms can be found in *Chapter 3.3*. For the *MATLAB* programming environment the steps in *Chapter 3.3* are put side by side with accompanying *.m* files in *MATLAB* (Vrugt et al., 2003c). *Figure 5.2* gives an idea of these function names for different steps in *SCEM-UA* and *SEM*. In the next section re-reference is made to the same function names with additional *.m* files that are used for connection to *EPANET* dynamic link library files.

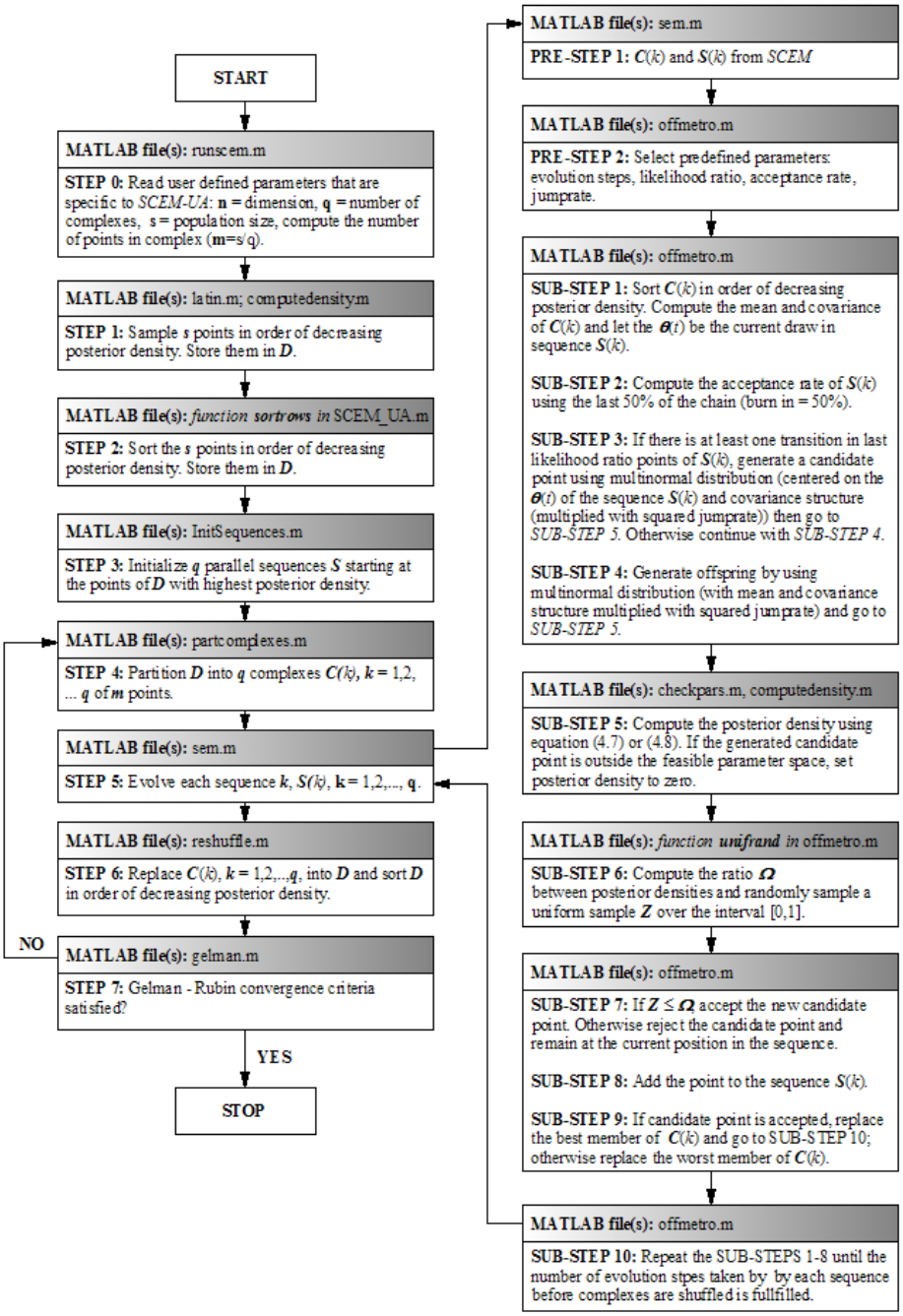

Figure 5.2 The flowchart of algorithms SCEM-UA and SEM in conjunction with MATLAB functions

5.2.2 EPANET and SCEM-UA connection at file level

In *Chapter 5.2.1* the *SCEM-UA* layout in *MATLAB* was presented. In this chapter the flowchart for the whole environment is given where more attention is paid to connections within *EPANET* and *SCEM-UA*. The parameters that need to be defined before the analysis can start are more closely investigated in *Chapter 5.2.3*. *Figure 5.3* describes the environment that makes up the analysis system at the file level. For more information about *SCEM-UA* please refer to *Chapter 3.3*.

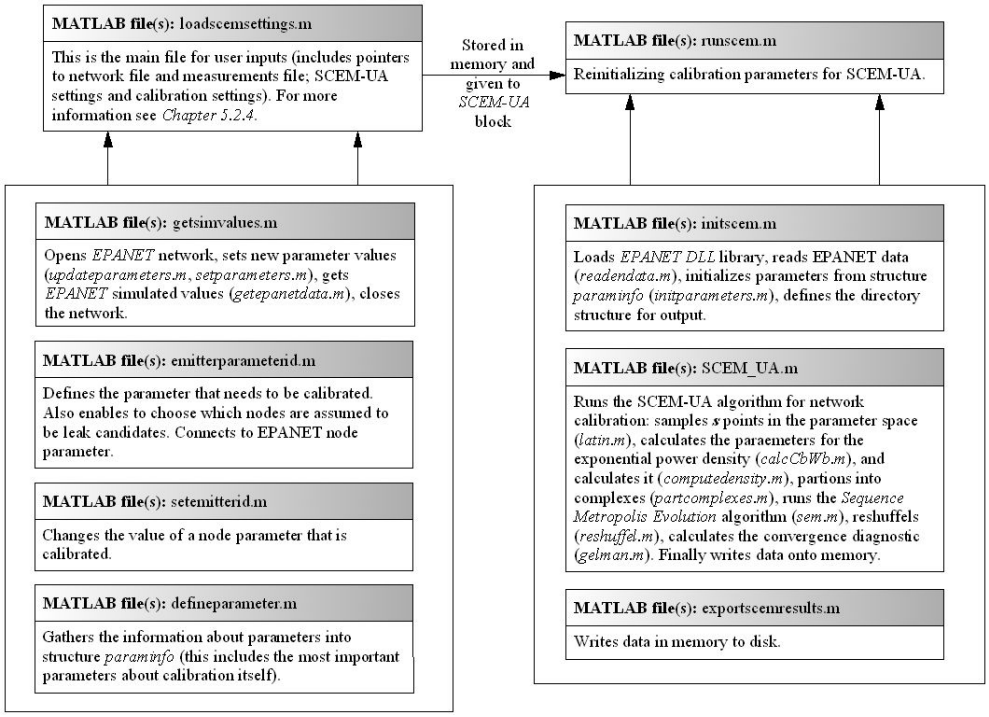

Figure 5.3 Flowchart for analysis system in MATLAB environment at the file level

5.2.3 User interface system for analysis

In this section the developed user interface for the analysis is discussed. Also, the main parameters of *SCEM-UA* are given and described. *Figure 5.4* presents the graphical user interface for calculations.

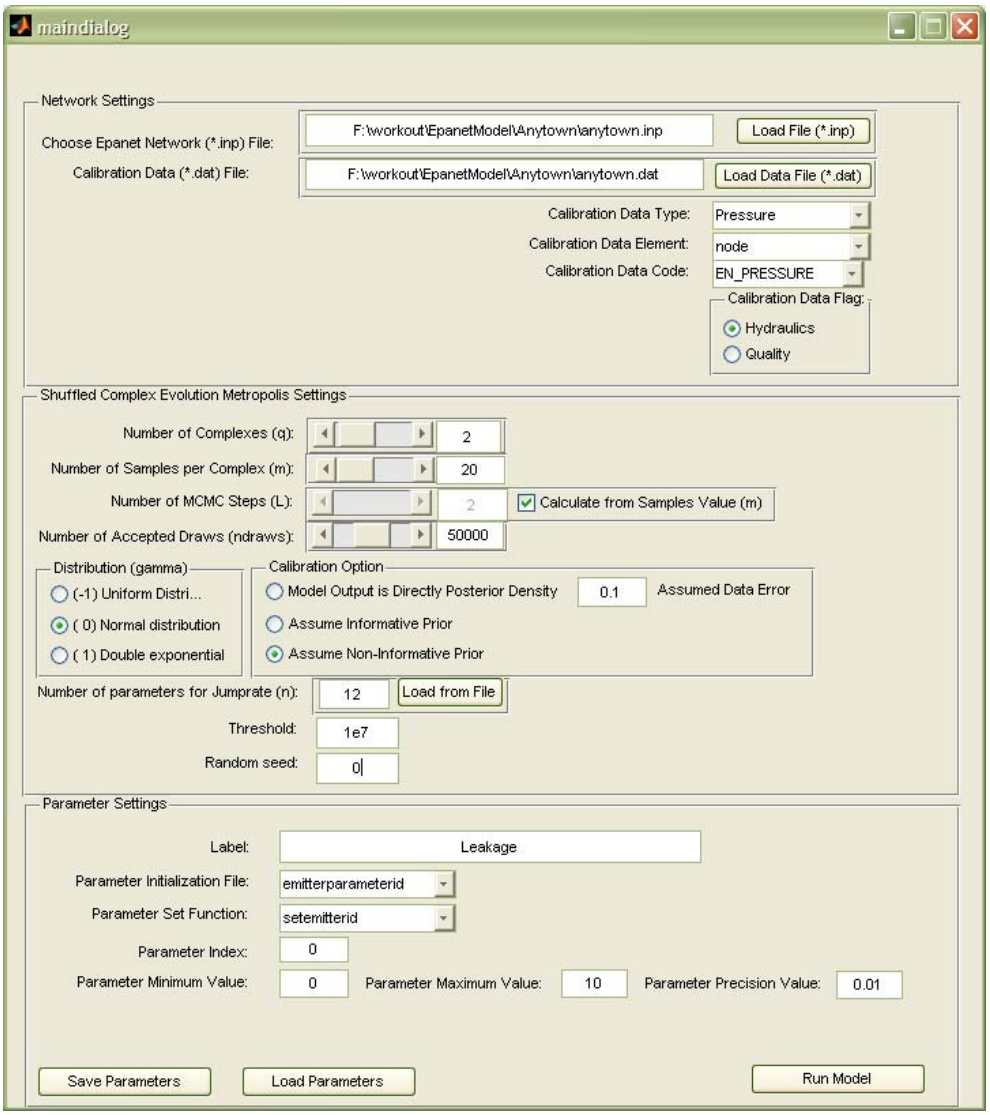

Figure 5.4 The user interface system

The user interface system that the user needs to work with before the start of calculation is divided into three main sections: *Network Settings*, *SCEM-UA Settings* and *Parameter Settings*. After the input of values the network model can be run. Hereby, a short description of main user input parameters is given.

Network Settings
Here model architecture, measurements and calibration type are chosen.

Choose EPANET Network (.inp) File* – Model architecture that needs to be calibrated against leakages is chosen here. It is a standard *EPANET* model input file that can be chosen from the hard drive pushing a button *Load File (*.inp)*.

Calibration Data (.dat) File* – For calibration the measurements are needed. This file is written in the standard *EPANET* data file format that describes all nodes at all times when there is a measurement data available. The user can again choose it from the hard-drive pushing a button *Load Data File (*.dat)*.

Calibration Data Type – At this time two main types of data are available for input: *Pressure* and *Flow*. Also, their respective derivatives can be used (*Head* and/or *Velocity*).

Calibration Data Element – Measurements are taken either on the pipe or at the node. Therefore in this section there are two possible values: *node* or *link*.

Calibration Data Code – A code in *EPANET DLL Library* that is specific for a certain type of data. When pressure is used, the code will be *EN_PRESSURE*.

Calibration Data Flag – The analysis can be run in two separate modes. At this time only *Hydraulics* is taken into account. The second option would be *Quality*.

***Shuffled Complex Evolution Metropolis* Settings**
These parameters are specific for the *SCEM-UA* algorithm. Depending on network architecture these values should be carefully chosen. Some information about the meaning of each value is given at each section. Please refer also to Vrugt et al. (2003a) for more information.

Number of Complexes (q) – For simple problems with an uncorrelated or correlated Gaussian target distribution a small number of generated sequences/complexes ($q \leq 5$) will usually be sufficient.

Number of Samples per Complex (m) – Usually this value is 10 or 20. Higher values might be chosen for complex models. Normally the values of *q* and *m* are found (or adjusted) during a simple preliminary sensitivity analysis. Every network might have its own pair of parameters. These parameters also affect the time of analysis. Therefore, it is suggested that too large values are not taken. The recommended values are shown in this work in the case studies section.

Number of MCMC Steps (L) – In general, this value can be calculated from previous value using the equation $L = m/10$. The user can choose if that equation is used

(making a check mark into the box *Calculate from Samples Values (m)*) or this value can be entered manually. This parameter is important for the *Sequence Evolution Metropolis* (*SEM*) algorithm.

Number of Accepted Draws (ndraws) – This value determines how long the calculations should be carried on. If this value is met, the calculation ends even if the *Gelman Rubin Convergence* is not met. Calculations also end when *Gelman Rubin Convergence* criterion is met before this value. The convergence criterion for *Gelman Rubin* is given as:

$$\sqrt{SR} = \sqrt{\frac{i-1}{i} + \frac{q+1}{q \cdot i}\frac{B}{W}} \tag{5.1}$$

Where $\sqrt{SR}$ = quantitative convergence diagnostic also called the scale reduction score, based on the within and between chain (sequences) variances (Gelman and Rubin, 1992); i = number of iterations within each sequence; q = number of complexes; B = the variance between the q sequence means; W = the average of the q within-sequence variances for the parameters under consideration. The product of i and q is identical to the total number of draws, with the *Metropolis-Hastings* or *SCEM-UA* sampler. When the potential scale reduction is near 1, the conclusion is made that each set of the q complexes of i iterations is close to the target distribution. Convergence, equal to 1, is sought for each parameter that is calibrated. However, because a score of unity is difficult to achieve, Gelman and Rubin (1992) recommend using values less than 1.2 to declare convergence to a stationary distribution.

Distribution (gamma) – Three options are available: *Uniform Distribution*, *Normal distribution*, *Double exponential*. By default *Normal distribution* is used. Please refer back to *Chapter 4* (equations (4.7) & (4.8)).

Calibration Option – Three options are available: *Model Output is Directly Posterio Density*, *Assume Informative Prior*, *Assume Non-Informative Prior*. By default *Assume Non-Informative Prior* is used. Other selections have also an additional parameter *Assumed Data Error*. This value is ignored at default option.

Number of parameters for Jumprate (n) – This value is used for jumprate evaluation. It is recommended that jumprate is calculated using the following equation: $2.4/\sqrt{n}$ (Vrugt et al., 2003a). Value n indicates the number of calibration parameters and therefore it can be read automatically (or inserted manually) from the file using a button *Load from File*. This value belongs to the *Sequence Evolution Metropolis* (*SEM*) algorithm.

Threshold – Predefined likelihood ratio in the *Sequence Evolution Metropolis* (*SEM*) algorithm. Default value is $T = 10^6$.

Random seed – Indicates the start point in the parameter space where calculations are started. Normally the calculations are averaged over a range of different random seed values to get more certain values.

Parameter Settings
This section defines problem set-up parameters.

Label – Simple text that is used to describe the problem category.

Parameter Initilization File – Describes the calibration parameter type. The *MATLAB m-file* is chosen to indicate the type of problem. At this stage two different selections can be made: (a) *emitterparameterid*. Here nodes that are assumed to be leak candidates can be predefined; (b) *emitterparameter*. All nodes that are found from the *EPANET* input file are automatically assumed to be leak candidates and are calibrated against the leak coefficient.

Parameter Set Function – Companion to the previous selection. Set the model run values. Two selections can be made: (a) *setemitterid*. Companion to *emitterparameterid*; (b) *setemitter*. Companion to *emitterparameter*.

Parameter Index – Used as a mid calculation value. At this stage this value is equal to 0.

Parameter Minimum/Maximum Value – Defines the limits of each calibration parameter. In this case it is the emitter coefficient in *EPANET*. These limits can be determined by simple leak assessment to find out the maximum possible leak value per node.

Parameter Precision Value – Determines the precision for the calibration parameter. Default value is 0.01.

The chosen parameters can be saved into the file for easier selection in the later analysis (using buttons *Save Parameters* and *Load Parameters*). The values defined in this user interface are loaded onto file *loadscemsettings.m* and based on that file the model is run. The analysis of calculation results is more closely looked in *Chapter 6*.

6. Leak detection case studies

6.1 Introduction

In this chapter three leak calibration case studies are presented with the main objectives as follows: (1) verification of the probabilistic leak calibration approach presented in *Chapter 4* with perfect measurement data; (2) verification of the probabilistic leak calibration approach with imperfect measurement data; (3) verification of the probabilistic leak calibration approach applied to a real network.

In the *Case study 01*, the calibration approach is verified based on the literature case study that was presented by Pudar and Liggett (1992). The calibration process is fully directed meaning that both the network model and observed data are artificially created. Normally this is the first stage of checking a new calibration algorithm to see how it performs when results are fully predictable and checkable.

In *Case study 02*, the calibration approach is verified on a much bigger network involving the larger, more complex pipe network with more realistically modelled leaks. The model used here (named *Anytown Network*) is used by several authors, more recently Kapelan et al. (2005) used it for pipe roughness calibration. The model is calibrated as a multiple loading condition steady-state model. Additionally imperfect measurement data is used to generate a more realistic example.

In *Case study 03*, the probabilistic leak calibration approach is applied to the real network data of the town of Rakvere (Estonia).

6.2 Case study 01 - Perfect measurement data

6.2.1 Problem description

The main purpose of this case study is to test the performance of the *SCEM-UA* algorithm for detecting leaks in a simple pipe water network with perfect pressure measurement data (see *Table 6.1*). The pipe network used in this case study is the simple network used previously by Pudar and Liggett (1992). The network layout and data are shown in *Figure 6.1*. Three different leak cases are considered: (1) one leak occurring at node 2; (2) two leaks occurring simultaneously at nodes 2 and 4 and (3) three leaks occurring simultaneously at nodes 2, 4 and 7. The assumed values of leak areas can be found in *Table 6.3*. The perfect measurement data is the pressure head data obtained by running the *EPANET2* software with the ('unknown') leak areas set equal to the corresponding assumed (i.e. true) values.

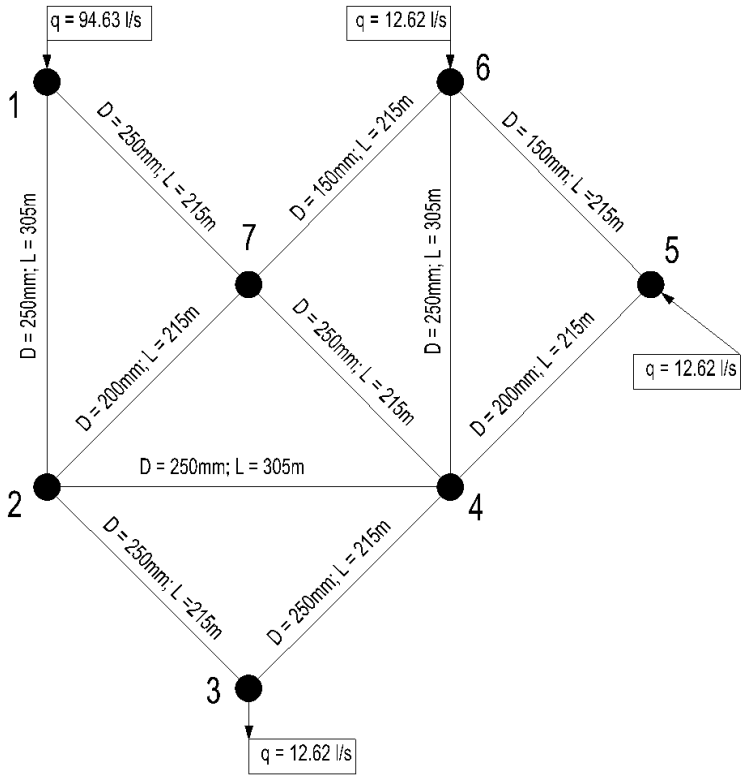

Figure 6.1 Case study network as seen in Pudar and Liggett (1992). All pipes have Hazen-Williams coefficients of 120.

The perfect pressure measurement data is assumed to be available at nodes 2-7. Three nodes are considered as possible leak candidates in all above leak cases: nodes 2, 4 and 7. Nodal leak flow is defined using the orifice equation, i.e. equation (4.3) with $\beta = 0.5$. Note that all the inverse problems solved here are over-determined, but not by much (6 measurements vs. 3 calibration parameters).

Table 6.1 Pressure head measurements (m)

Case study	Node					
	2	3	4	5	6	7
Leak at node 2	27.59	26.06	27.84	28.24	28.18	28.49
Leaks at nodes 2 and 4	25.98	24.24	25.81	26.27	26.25	27.06
Leaks at nodes 2, 4 and 7	25.93	24.18	25.75	26.2	26.18	26.99

6.2.2 Analysis

Three leakage detection problems (presented in *Table 6.1*) were solved for the 'unknown' K leak factors (see equation (4.3)) using the *SCEM-UA* algorithm described in *Chapter 3*. The *SCEM-UA* algorithm parameters used are shown in *Table 6.2*. These values are guided by the values suggested by Vrugt et al. (2003a)

but are based on a limited sensitivity analysis. Each of the three leak detection problems is solved by running the *SCEM-UA* algorithm 51 times, each time with a different random seed (i.e. different random initial population).

Table 6.2 SCEM-UA parameters used in Case study 01

Option for posterior PDF	Number of Complexes	Samples per Complex	Number of MCMC steps	Jump rate	Threshold	Parameter min value*	Parameter max value*
Non-Informative Prior	2	60	6	1.39	1.00E+07	0	15

** Note: The range of possible values of K was chosen to represent a leak from [no leak] to [two times the assumed leak size].*

Table 6.3 Case study 01 results

Leak Area (node index as subscript)	True values	SCEM-UA results		Results by [Pudar et al, 1992] using Levenberg-Marquardt method
		Mean	Standard deviation	Mean
Case study 1.1. One leak (Node 2), three leak candidates (Nodes 2, 4, 7)				
A_2	0.00232	0.00234	0.000102	0.00232
A_4	0	3.75E-05	6.97E-05	0
A_7	0	3.75E-05	5.36E-05	0
Case study 1.2. Two leaks (Nodes 2 and 7), three leak candidates (Nodes 2, 4, 7)				
A_2	0.00232	0.00232	5.63E-05	0.00232
A_4	0.00232	0.00228	5.63E-05	0.00232
A_7	0	4.56E-05	3.49E-05	0
Case study 1.3. Three leaks (Nodes 2, 4 and 7), three leak candidates (Nodes 2, 4, 7)				
A_2	0.00232	0.00231	5.36E-05	0.00232
A_4	0.00232	0.00232	5.36E-05	0.002317
A_7	0.00009	9.39E-05	3.49E-05	0.0001

The best leak area values obtained in the above process (average means and standard deviations from 51 *SCEM-UA* runs) are shown in *Table 6.3* together with the corresponding optimal deterministic values obtained by Pudar and Liggett (1992) who used the *Levenberg-Marquardt* method. From this *Table 6.3* it can be concluded that the *SCEM-UA* algorithm is capable of determining the network leaks under the conditions analysed in this case study.

6.2.3 Summary

In this case study the small leak calibration problem was carried out to compare the proposed probabilistic leak detection algorithm with previously published results

where the *Levenberg-Marquardt* method was used (Pudar and Liggett, 1992). This case study was presented to apply and verify the probabilistic leak calibration approach presented in *Chapter 4* of this thesis.

6.3 Case study 02 – Imperfect measurement data

6.3.1 Problem description

The previous case study showed that the *SCEM-UA* can successfully detect leaks when perfect measurement data is used. The aim of this case study is to further test the performance of the leak detection methodology outlined in *Chapter 5* on a more realistic example involving a larger, more complex pipe network with more realistically modelled leaks. The leak detection problem formulated and solved here is based on the *Anytown* network shown in *Figure 6.2*. The same network was used before to test the *SCEM-UA* performance when calibrating the *WDS* hydraulic model for the grouped pipe roughness coefficients (Kapelan et al., 2005). The *Anytown* network has a single source from which the water is pumped into the network. Two tanks (both with the same initial water level of 77 m in all loading conditions) are used for providing the water to satisfy the network demands at peak times.

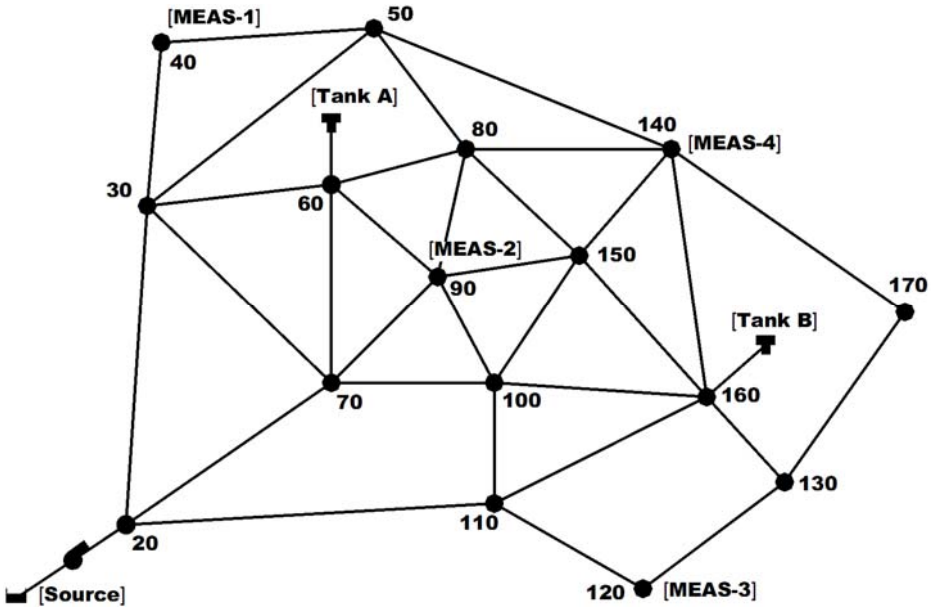

Figure 6.2 Anytown Network Layout

The measurement data in this case study consists of the 'observed' pressure heads at four nodes (see *Figure 6.2*) for five loading conditions, as specified in Kapelan et al. (2005). Imperfect pressure measurement data is created artificially by adding the normally distributed noise (with zero mean and standard deviation of 0.1 m) to perfect measurement data. Perfect measurement data was created by running the

EPANET2 software for the assumed leak(s) areas (see below for details). Nodal leak flow is defined by the model used in equation (4.3) with $\beta = 1.1$.

6.3.2 Analysis

Compared with the previous case study the number of iterations has been increased after first appearance of the convergence (see equation 5.1). The calculations run until there are 1000 iterations after the convergence. It should be noted that during the 1000 iterations the convergence criteria might rise and 1000 test values are needed again. Setting the limiting factor (maximum number of iterations) reasonably low the calculation process can be halted. It means that there is no convergence between all the unknowns and it is possible that some parameter tuning is needed before recalculation. *Figures 6.3* and *6.4* show the convergence statistics at two different parameters set up. *Figure 6.3* uses only 20 iterations after convergence, but *Figure 6.4* uses 1000 iterations. The overall result when using more iterations is much more stable and also more certain.

Using more iterations after the convergence ensures that the results are stable. It is also noted that when using more iterations, the parameters like "samples per complex" and "number of MCMC steps" can be decreased to *win back* the computer processing time.

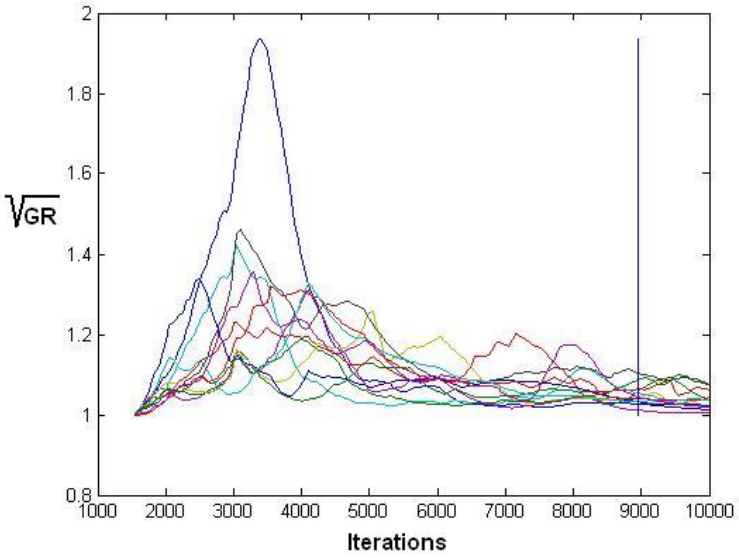

Figure 6.3 Gelman Rubin statistics with additional 20 iterations after convergence

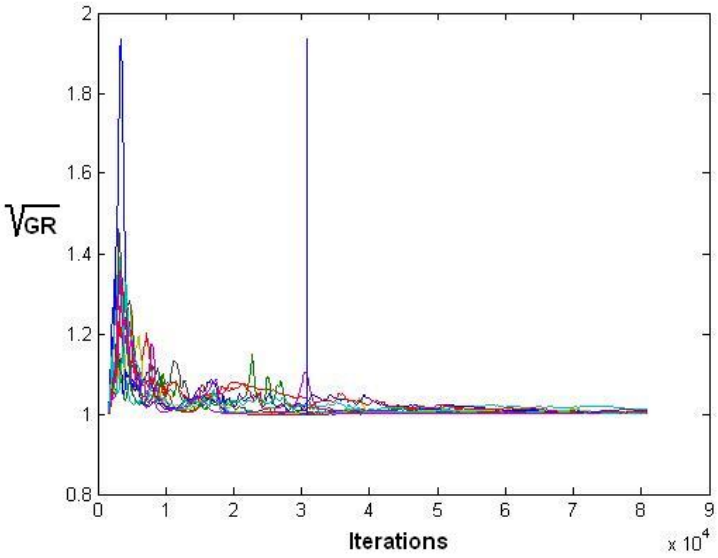

Figure 6.4 Gelman Rubin statistics with additional 1000 iterations after last convergence (Value 1.1 is needed for 1000 iterations to get final results)

Analyses are carried out for two different leak case studies: (a) Single leak and (b) Two leaks. In both cases the *SCEM-UA* algorithm parameters used are shown in *Table 6.4*.

Table 6.4 SCEM-UA parameters used in Case study 02

Option for posterior PDF	Number of Complexes	Samples per Complex	Number of MCMC steps	Jump rate	Threshold	Assumed Data Error	Parameter min value	Parameter max value
		6 nodes / 12 nodes						
Informative Prior	5	300	10	0.98/0.69	1.00E+07	0.1	0	10

In case of a single leak, a leak is assumed to exist at node 150. Two leakage detection problems are solved here: (1) assuming all 12 non-measurement nodes are possible leak candidates and (2) assuming selected 6 nodes are possible leak candidates. The results obtained are in *Table 6.5*. The values shown in this table represent the average values from multiple *SCEM-UA* runs with different random seeds. As it can be seen from *Table 6.5*, the *SCEM-UA* algorithm can detect the most probable leak location (with high mean and low standard deviation) in both cases. As expected, the model run with fewer leak candidates gives better solutions. The non-leak nodes have relatively high uncertainties (in terms of standard deviations) associated with obtained leak areas. This is caused by the errors in measurement data and also by the high sensitivity of particular nodes (like inflow nodes: 20, 60, and 160).

Table 6.5 Results for a single leak case in Case study 02

Node ID	True Value	12 leak candidates		6 leak candidates	
		Mean	Stdev	Mean	Stdev
20	0	0.07	0.11	x	x
30	0	0.02	0.03	0.01	0.01
50	0	0.01	0.02	0.01	0.01
60	0	0.37	0.32	0.36	0.29
70	0	0.05	0.08	x	x
80	0	0.06	0.06	x	x
100	0	0.09	0.09	x	x
110	0	0.02	0.03	0.01	0.01
130	0	0.01	0.02	x	x
150	2.00	1.59	0.25	1.88	0.07
160	0	0.19	0.22	x	x
170	0	0.02	0.03	0.01	0.01

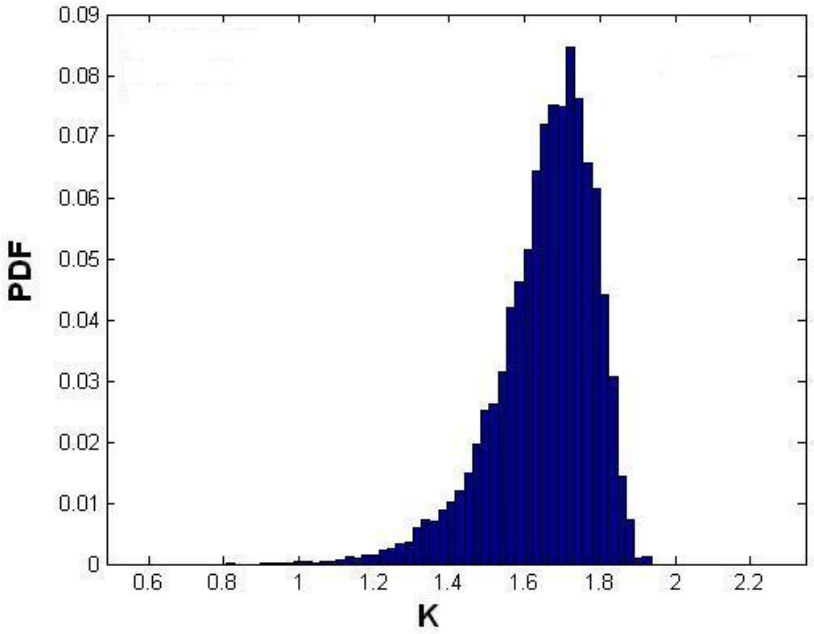

Figure 6.5 Normalized PDF curve for node 150 (the 17th run out of 51: the most probable value is 1.86)

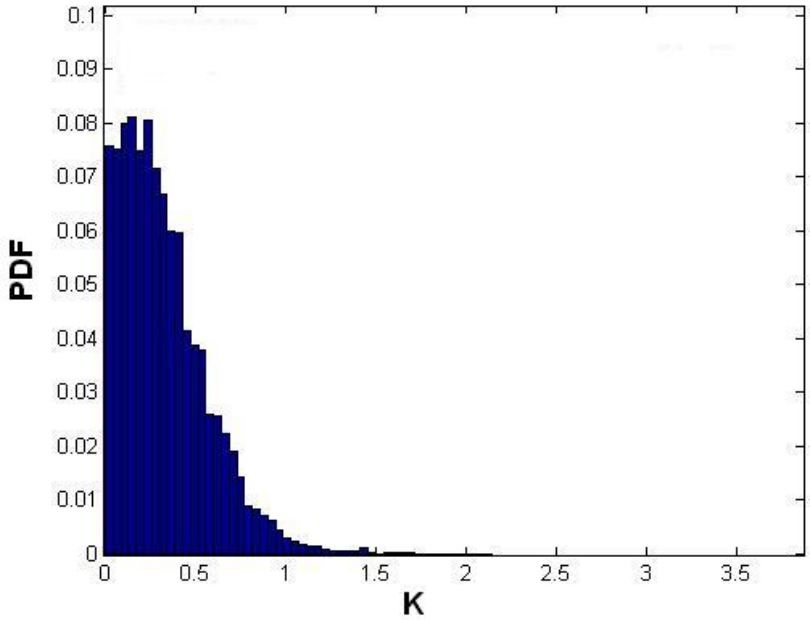

Figure 6.6 Normalized PDF curve for node 60 (the 17th run out of 51: the most probable value is 0.05)

The obtained posterior *PDF*s for leak areas at nodes 150 and 60 in a 12 leak candidate inverse problem are shown in Figures 6.5 and 6.6, respectively. Based on these two figures the following can be noted: (1) the most probable leak node is node 150 (note also from *Table 6.5* that the leak area obtained for this node is much closer to the 'true' value than for node 60), (2) the most probable value (obtained from the corresponding *PDF* curve) is always closer to the true value than the overall mean itself.

In case of two leaks, leaks are assumed to exist at nodes 30 and 170. Again, two inverse problems are solved: (1) assuming all 12 non-measurement nodes are leak candidates and (2) assuming a reduced set of 6 leak candidates (see *Table 6.6* for details). The results obtained are shown in *Table 6.6* and represent averages from multiple *SCEM-UA* runs with different random seeds.

Table 6.6 Results for a two leak case in Case study 02

Node ID	True Value	12 leak candidates		6 leak candidates	
		Mean	Stdev	Mean	Stdev
20	0	0.32	0.29	x	x
30	0.5	0.28	0.11	0.45	0.05
50	0	0.06	0.04	0.02	0.02
60	0	0.27	0.24	0.17	0.15
70	0	0.06	0.07	x	x
80	0	0.02	0.02	x	x
100	0	0.04	0.05	x	x
110	0	0.10	0.12	0.05	0.07
130	0	0.13	0.07	x	x
150	0	0.02	0.04	0.01	0.01
160	0	0.34	0.34	x	x
170	3.00	1.81	0.39	2.80	0.16

From *Table 6.6* it can be seen that in the leak detection case with 12 leak candidates, the 'true' leak nodes 30 and 170 have considerably less uncertainty associated with the leak areas obtained when compared to the other nodes. At the same time, the mean leak area values obtained are different from their true values. The two likely reasons for this are as follows: (1) presence of noise in the pressure measurements and (2) relative lack of the measurement information available (only 20 pressure heads used to determine a joint *PDF* of 12 leak areas). These observations are further justified by analysing the results obtained in the case with 6 leak candidates. As it can be seen from *Table 6.6*, in this case, the mean leak area values obtained are getting closer to their true values and the corresponding standard deviations are reduced when compared to the case with 12 leak candidates.

The marginal *PDF*s obtained for leak *K* values at nodes 30 and 170 in the case with 6 leak candidates are shown in *Figures 6.7* and *6.8*, respectively. From the results it can be seen that some values have almost the same standard deviation as the mean itself. This is mainly because they are situated in highly sensitive zones (close to system inflows, like nodes 20, 60, 160).

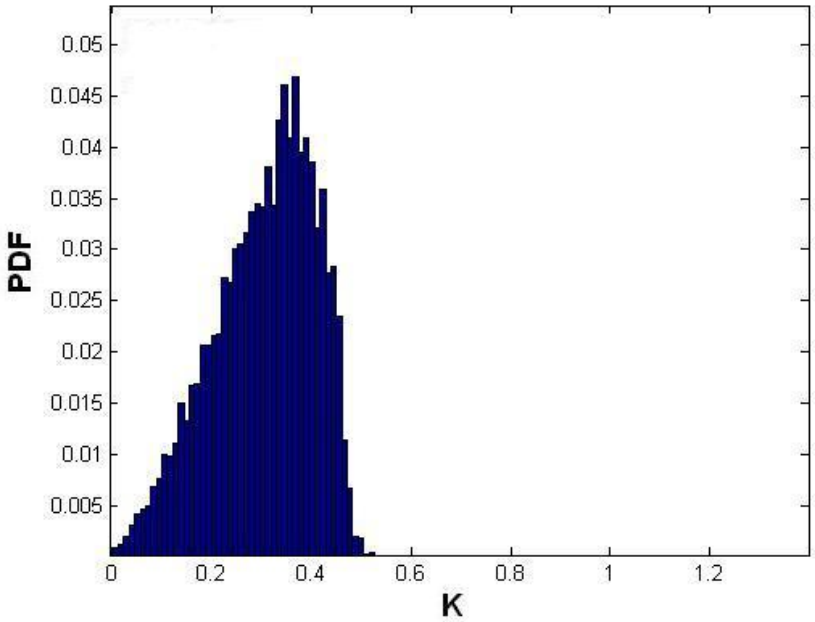

Figure 6.7 Normalized PDF curve for node 30 (the 25th run out of 51: the most probable value is 0.45)

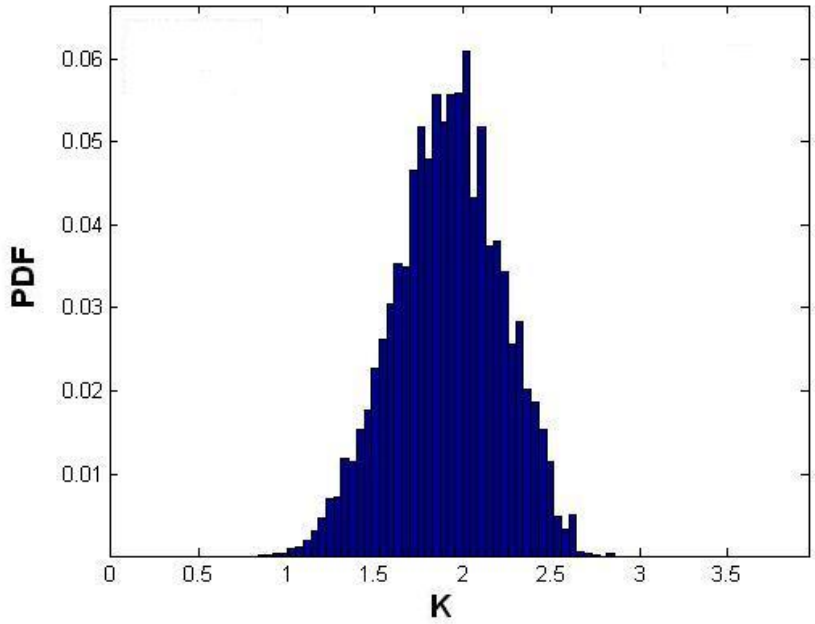

Figure 6.8 Normalized PDF curve for node 170 (the 25th run out of 51: the most probable value is 2.51)

This study demonstrates that nodes with leaks can be separated from the nodes without leaks even in the presence of noise measurement. The magnitude of value K differs in high and low pressure zones within wide ranges. Therefore this should be taken under consideration when setting up prior *PDF*-s (parameter minimum and maximum values, see for example *Table 6.4*). These estimates are made for every node (these are not global values for the system overall leakage) and it is not reasonable to enlarge these boundaries unrealistically. The existing information encapsulated in the form of the system designer's/operator's knowledge and/or experience could and should be utilised when identifying the nodal leak candidates and specifying the corresponding prior *PDF* parameters.

6.3.3 Summary

This case study was carried out with a much bigger (and realistic) network. This network had five loading conditions and multiple leak locations with their respective mean values were detected. This case study was presented to apply and verify the probabilistic leak calibration approach with imperfect measurement data.

6.4 Case study 03 – Real measurement data

6.4.1 Problem description

The previous case studies had shown that the methodology is able to detect leak areas in artificial situations. The next step therefore, would be the real network case study. The town of Rakvere (Estonia) was chosen as an example. The town's network is given in *Figure 6.9*. It has one input (indicated as *INPUT* in the picture) and 8 separately measured zones (see *Figure 6.10*).

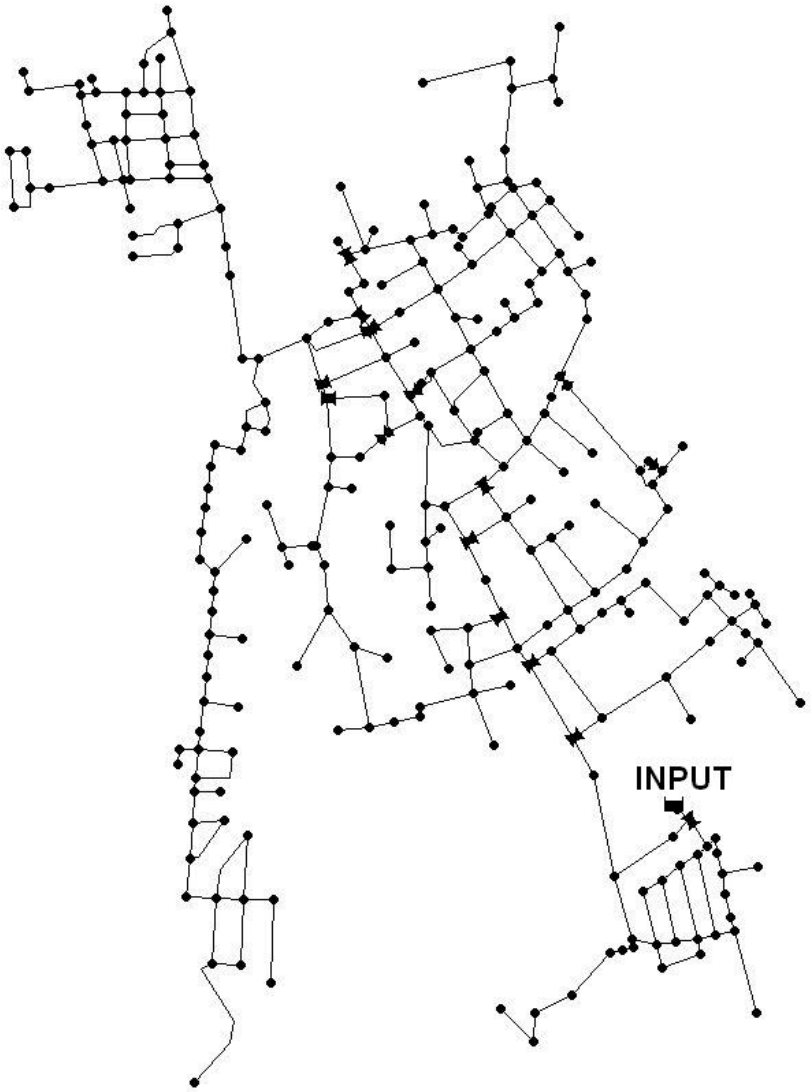

Figure 6.9 Water network of a real water distribution system (town of Rakvere)

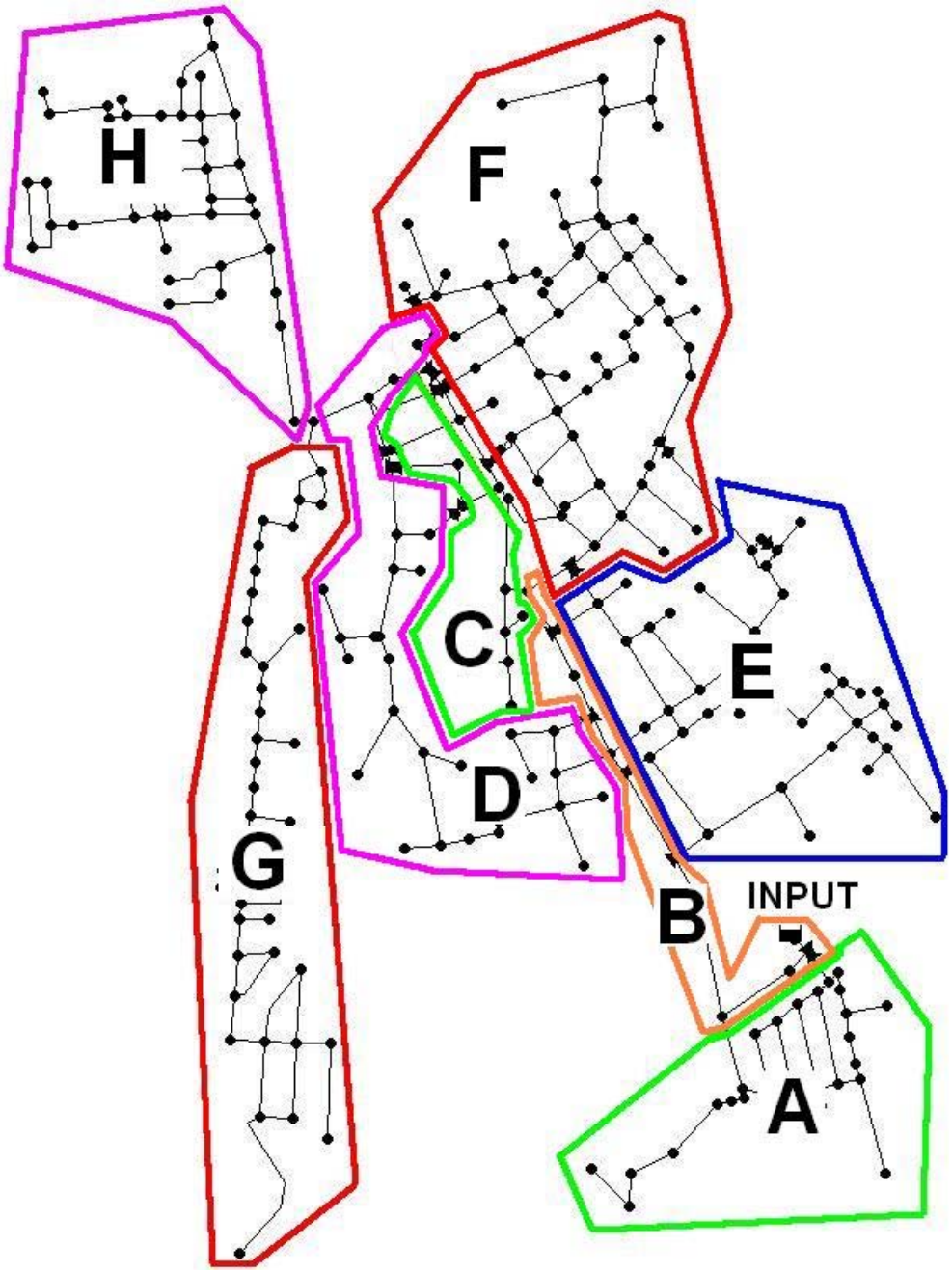

Figure 6.10 Rakvere water network's district metered areas

The water network was calibrated before the leakage studies. As the network is with separated district metered areas, *Zone F* was taken as a pilot study and used for the analysis in the current work. *Zone F* was separated from the whole model using the measured flow and pressure data at the zone's entrance point. Demands at the nodes are calculated based on the client's database and derived from the month average.

Respective time patterns are taken from the zone's pressure/flow data. Measurement point locations (pressure measurements) are indicated in *Figure 6.11*.

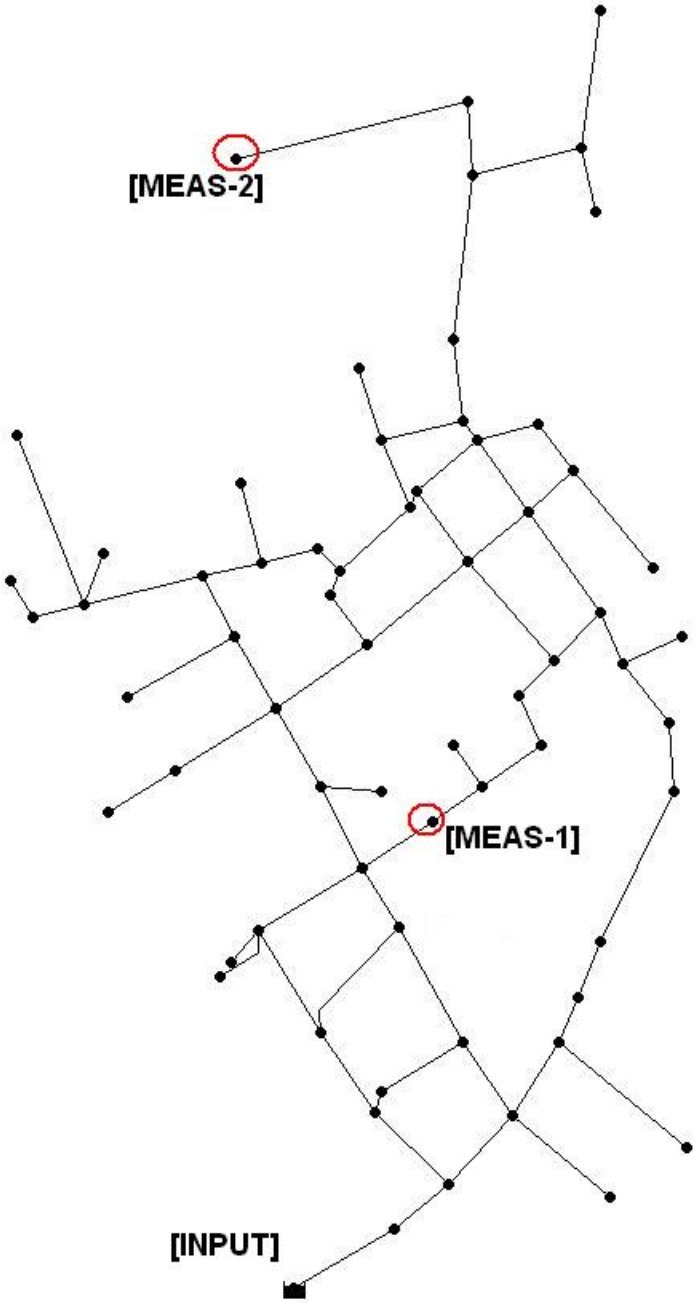

Figure 6.11 District metered area – Zone F

Measurements were available at two locations. Pressure loggers were recording with a 30 second interval. The main scope was to use fire hydrants as leak sources and let the water out to imitate the leak situation. One case study involved approximately 15 minutes of free flow, thereafter the fire hydrant was closed completely. Different flow rates were used at night and in the daytime. Hereby the analyses are carried out for night time. The calibrated model (in respect of pipe roughnesses) was used in conjunction with the *SCEM-UA* algorithm to detect the leak location.

6.4.2 Analysis

The case study analyzed here is based on a one leak case study that is simulated with a fire hydrant at around two o'clock at night. The test of the fire hydrant lasted from *01:55* to *02:20* including the hydrant's opening, stabilization and closing. The approximate fire hydrant flow was *2 L/s* (this was derived from the approximate flow level rise at the pump station). The average demand at this time was about *1.7 L/s*. Considering the approximate demand and hydrant flow value, the simulated leak was proposed to imitate *50%* additional flow. The *EPANET* network model is used with 26 time steps (from *01:55 – 02:20* every minute) for the leak localization. Network nodes that are considered as possible leak locations are shown in *Figure 6.12*.

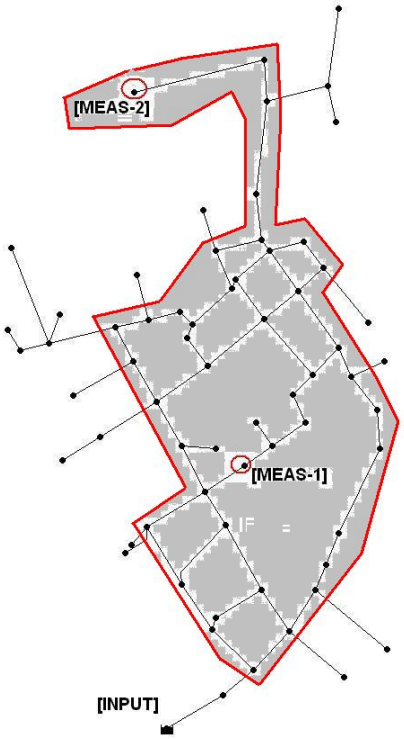

Figure 6.12 Nodes in the shadowed area are used as leak candidates (43 nodes)

At this stage 43 nodes were considered as possible leak candidates. Nodal leak flow is defined by the model used in equation (4.3) with β = 1.0 (taken as an approximate value). SCEM-UA parameters used are in *Table 6.7*.

Table 6.7 SCEM-UA parameters used in Case study 03

Option for posterior PDF	Number of Complexes	Samples per Complex	Number of MCMC steps	Jump rate	Threshold	Assumed Data Error	Parameter min value	Parameter max value
		43 nodes						
Non-Informative Prior	50	30	3	0.37	1.00E+07	0.00	0	0.2

The results obtained are in *Table 6.8*. An emitter coefficient of 0.05 is expected at one node. This is derived from the district metered area's average flow at that time.

Table 6.8 Results for a real network - Case study 03

Leak coefficients		
min P	0.0002	
avg P	0.0167	
max P	0.0824	
Stdev P	0.0203	
		Stdev
min (Mean)	0.0166	0.0167
avg (Mean)	0.0342	0.0215
max (Mean)	0.0952	0.0433

In *Table 6.8* the minimum, maximum and average value of leak coefficients from 43 nodes are shown. It is an easier interpretation to show the results. Value *P* means the most probable value. It can be concluded that there is no clear distinction between no-leak and leak nodes at this time. This can be seen from the average values *avg P* and *avg (Mean)*. It was stated beforehand that only one node with an approximate leak coefficient of 0.05 is sought. From *Table 6.8* it can be seen that there are too many nodes with too big leak coefficients (because the average is high). *Figure 6.13* maps the average (mean) and most probable (P) values for the respective nodes from *Table 6.8*. Why is the value at the *max (Mean)* position so high? Almost twice as high as it was expected. *SCEM-UA* randomly generates samples for all nodes, and thereafter these values are recalculated in respect of pressure measurements. When the next iteration is evaluated, it checks if there is an improvement and stops if there is not. So it can be concluded that for some points the pressure measurement locations are too far away.

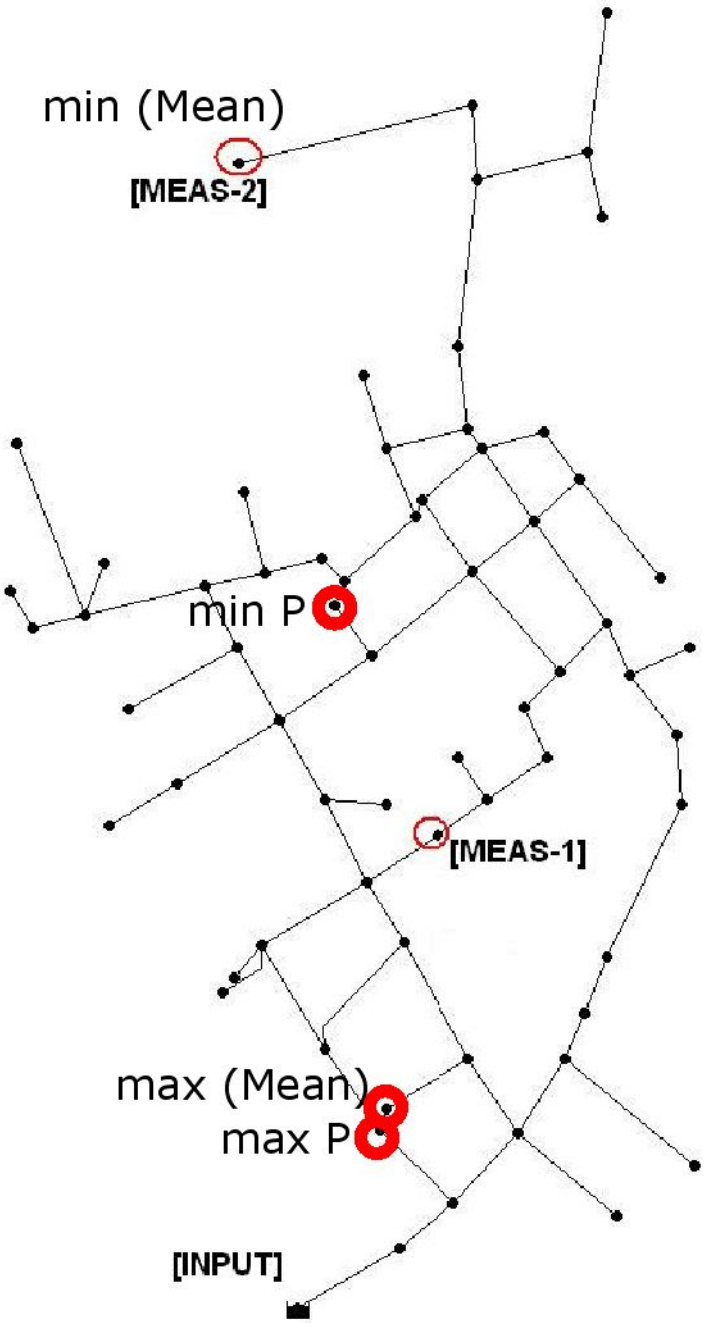

Figure 6.13 Test values from the first run showing that pressure measurements do not support points away from the measurement location (max (Mean) is too high to be true).

The next calculation is made with fewer nodes. It was also demonstrated in *Case Study 02* that after preliminary calculations concentration on fewer nodes might improve the results. *Figure 6.14* shows the points that were taken into calculation.

SCEM-UA parameters are the same as in *Table 6.7* except the parameter *Jump rate* = 0.58.

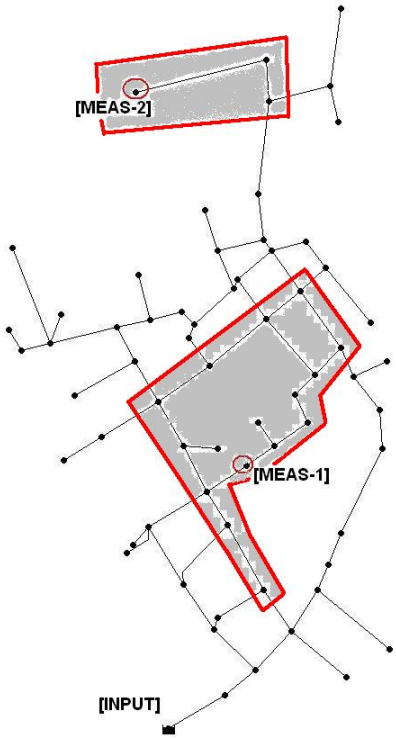

Figure 6.14 Reassumed possible leak locations to check the sensitivity of pressure measurement locations (17 nodes)

The results obtained are in *Table 6.9*. Again an emitter coefficient of 0.05 at one node is searched and much lower values at other nodes. The minimum and maximum values are plotted in *Figure 6.15*.

Table 6.9 Results for a real network - Case study 03 – calculation 2

Leak coefficients		
min P	0.0002	
avg P	0.0258	
max P	0.0888	
Stdev P	0.0255	
		Stdev
min (Mean)	0.0142	0.0112
avg (Mean)	0.0255	0.0140
max (Mean)	0.0679	0.0426

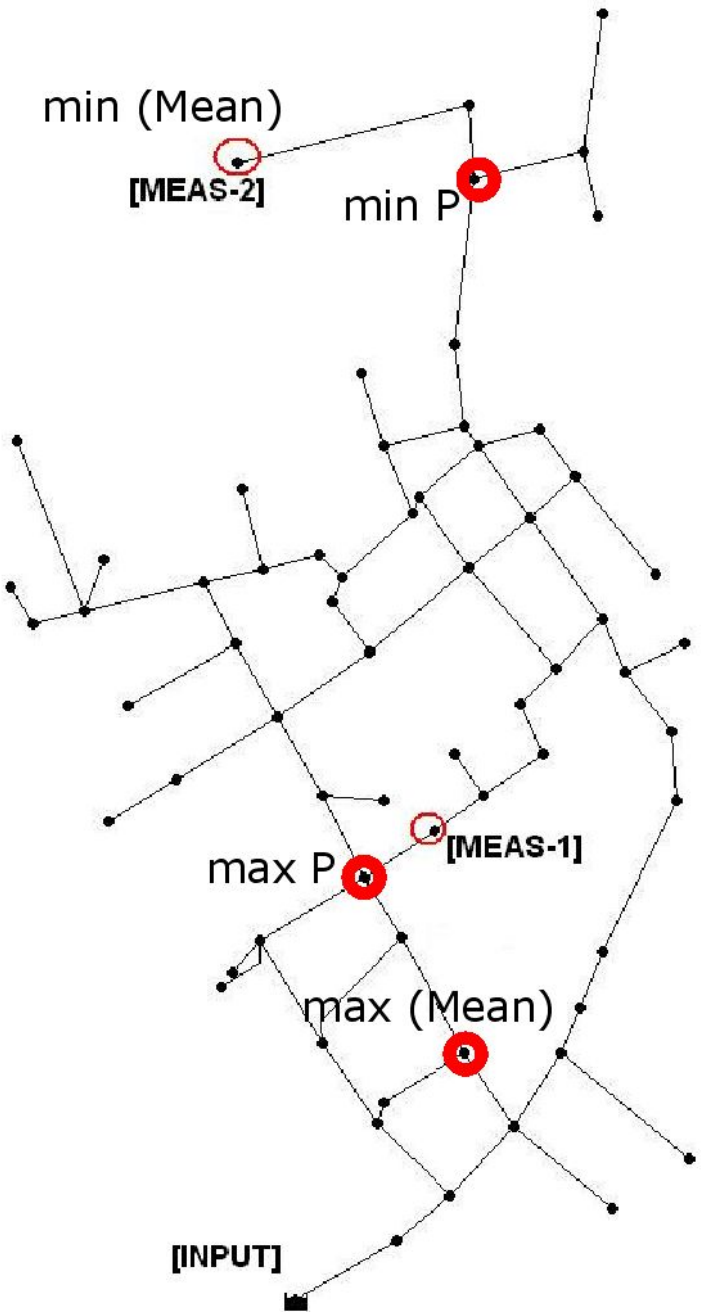

Figure 6.15 Test values from the second run showing that pressure measurements do not support points away from the measurement location (max (Mean) is too high to be true).

It can be concluded from *Figure 6.15* that the most probable minimum and maximum values are very close to the measurement points. On the other hand, the calculations show that there are several other nodes that have too large leakage

coefficients. When the leakage coefficients are summed up, the overall value of 0.5 is ten times higher than it is expected. Therefore, the results are not realistic because the input flow to the system rises from an expected value of 3 L/s to 17 L/s. Another test is carried out where a data error is expected in the measurements (See also *Case Study 02*). The *SCEM-UA* algorithm allows to assume the data error. The measurements at two locations are made with a 0.5 m step. This value is used as an error. The results obtained are in *Table 6.10*. An emitter coefficient of 0.05 is searched at one node and much lower values at other nodes.

Table 6.10 Results for a real network - Case study 03 – calculation 3

Leak coefficients		
min P	0.0014	
avg P	0.0280	
max P	0.1365	
Stdev P	0.0337	
		Stdev
min (Mean)	0.0119	0.0082
avg (Mean)	0.0290	0.0164
max (Mean)	0.0773	0.042

There is not a big difference. The main problem that occurs is that leak coefficient values are too high at too many nodes. The reason for this is not something extraordinary. Simply, because of low velocities in pipes the pressure at measurement locations is not affected by recalculations. The average velocity at the zone's input is 0.05 m/s at night. A test was carried out in the daytime. Still, the flows are very low, rising up to 0.09 m/s at the main pipe. A simple test showed that when an occasional leak is added to the node then the pressure measurements at fixed locations are affected only a few nodes away. The response would be in the range of 0.05 m (two nodes away) to 0.10 m (one node away). Some improvement can be reached when the pressure measurement locations are at main pipes (where the velocities are bigger). Still, the overall change is an additional 5 cm of pressure change. It means that in these conditions too many pressure measurement points are needed which is not sensible. Another option would be to use flow meters at main diversions or minimizing the pipe diameters to increase the velocities.

6.4.3 Summary

This case study was carried out with a realistic network. The main focus was to investigate the leakages that were generated at one point with an open hydrant. Hydrant flow was approximately 50% of the average demand in the region. The results showed that it is impossible to get good results in low velocity networks. It was recommended that flow meters in such situations can improve the overall results and make sense of probabilistic leak calibration in such cases.

6.5 Summary and conclusions

In this chapter, several case studies were carried out with the following main objectives: (a) to verify the leak calibration of the water distribution network model (suggested in *Chapter 4*) where probabilistic methodology is used (b) to demonstrate and analyse that the proposed model can be used to detect and locate leakages starting from a bigger network area (more leak candidates) and thereafter moving closer to the predicted suspicious leakage areas (more concentration on the particular network area) to adjust the modelled values.

The main conclusions regarding the calibration of the WDS hydraulic models against leakages are as follows:

1. The probabilistic leak calibration approach is successfully applied to steady and multiple steady-state loading condition of WDS hydraulic models. Firstly a small artificial network model was used. Secondly more realistic artificial network model was used. Various analyses were conducted to detect the leak values and locations in probabilistic terms.

2. For more reliable results, water distribution models should be calibrated for a limited number of calibration parameters. An analysis in probabilistic terms shows that when leak location areas are found, an additional analysis in that area, improves the final results (i.e. parameter certainties increase). In overall, the available observation data should be used properly to decrease the parameter uncertainties. Keeping the analysis at an over-determined level (more measurements than calibration parameters) helps to get more reliable results.

3. The prior information can help to increase the certainties of calibration parameters. Before the analysis the prior intervals for calibration parameters should be realistically chosen (i.e. conducting the basic leak assessment studies to find the average quantities for losses and from there to find the maximum limit per element in terms of calibration parameter type).

4. In case of real network data (including imperfect measurement data) the sensitivities of measurement point locations might change the calibrated values considerably. Since normally the water authorities have some fixed measurement stations and it is not possible to change their locations by time basis, it is recommended to recompense the available measurement data with seasonal mobile measurement points. These measurement points can be relocated when the network's architecture or loadings change to get more sensitive measurement data.

5. Despite the fact that the method has shown promising results on artificially generated realistic network data, a problem appears when the leakages are searched in low velocity networks. Relocating the pressure measurement

stations does not help much when low flows exist. Adding more pressure measurement stations makes the overall performance better. Still, preliminary calculations can show the sensitivities of particular network nodes against pressures and it might happen that using pressure measurements in conjunction with flow meters gives much more reliable results and is more economic than using only pressures. In general, the lack of pressure/flow data in low-flow systems gives unpredictable results with other commonly known algorithms.

7. Summary, conclusions and recommendations for further work

7.1 Summary

7.1.1 Thesis summary

The main objective of this thesis is to point out that as model calibration always involves uncertainty, it is natural to use probability to express the results. The *Bayesian* type calibration approach called the *Shuffled Complex Evolution Metropolis* has been applied to a water distribution system. Its main characteristics are as follows:

- The objective function in the calibration problem is defined as a weighted least square type.
- Previously tested only in hydrological models, the *Shuffled Complex Evolution Metropolis* algorithm is used in the optimisation process with its excellent statistical capabilities that help to draw/analyse the final results and conclusions about the calibration as a whole.
- Two main types of hydraulic models are used: (a) the steady-state flow model, and (b) the multi-steady state model.
- Special attention in calibration process is paid on the model's capabilities to work in an uncertain environment (including measurement errors).
- Although it is emphasised that more data will normally lead us to better results, sensitive data is out of scope in the current thesis, mainly because of the fact that most of the WDS systems have fixed measurement stations. Still, there is a possibility to use seasonal, additional measurement points to include new valuable data about the system's behaviour as time passes.
- The calibration procedure is tested in several case studies where each of them fulfils a particular task to show the method’s positive applicability.

7.1.2 Summary of the present work contribution

The main contributions of the work presented in this thesis regarding the WDS modelling are as follows:

- A novel probabilistic leak detection method with user definable prior information about the assumed leak areas which is applicable to any pipe network configuration.
- The possibilities to choose the model error predictions before calculations start and use it in conjunction with real network data.

The main contributions of the work presented in this thesis regarding the WDS hydraulic model calibration are as follows:

- The probability density function can be continuously updated when additional information about the system becomes available.
- The possibility to continue with post-calibration analyses or concentrating on a new leak area that has previously been found to be the most probable. Using available data in a more concentrated area might help to find the most critical leak areas.

7.2 Conclusions

7.2.1 Probabilistic leak calibration in pipeline systems

The main conclusions regarding the probabilistic leak calibration in pipeline systems are as follows:

- Before any WDS calibration procedure is involved the particular aims which the calibration should achieve should be clearly identified. The main question that should be asked before the calibration would be: "What the calibrated model is later used for?" Stating the goals helps to identify the types of data that needs to be collected. Even if some data seems to be unimportant, it might be useful in the post-calibration process or for analysing some anomalies in the results.
- Every real WDS has some uncertainties included and therefore defining the calibration procedure as a probabilistic one gives additional information to the decision maker. In essence, the calibration procedure is a complicated task, and engineering knowledge is still needed to choose the right data to be used in the calibration/optimisation.
- The methodology outlined here has the further advantage of being able to capture the knowledge and experience of various *WDS* staff in the form of (marginal) prior posterior density functions of assumed leak areas. The *Bayesian* recursive procedure can then be used to continuously update these posterior density functions when and if the additional measurement data becomes available.
- One of the clear advantages of probabilistic leak calibration in pipeline systems is its capability to give answers even with an insufficient amount of measurement data. As new measurement data gets available the previously received results can be improved through the update of posterior density functions. This way, additional information is provided for a decision maker, when compared to the existing modelling (Burrows et al., 2003; Ainola et al., 2000; Kapelan et al., 2000; Dunlop, 1999) approaches (which normally result in a single number, i.e. leak area size, determined for each nodal leak candidate).
- Using reasonable predictions (i.e. engineering knowledge) about the calibration parameters' boundaries helps to get meaningful results. These pre-definitions can be made before the calibration procedure using some basic leak assessment studies.

7.2.2 Bayesian type, SCEM-UA algorithm

The main conclusions regarding the usefulness of the *Bayesian* type, *Shuffled Complex Evolution Metropolis* algorithm in leak calibration are as follows:

- Both the leak size and the associated error (i.e. uncertainty) can be determined in a single, optimisation type model run (even more, the whole joint parameter of probability density function is determined).
- The *SCEM-UA* methodology does not require any post-processing type procedure (e.g. the *First Order Second Moment*) together with the associated assumptions (e.g. linearity and normality) to be made to determine the above uncertainties.
- The *SCEM-UA* methodology does not require any derivatives to be calculated (possible discontinuity issues and numerical errors).
- The *Bayesian* recursive procedure used in the *SCEM-UA* methodology enables the specification of prior information on calibration parameters in a more flexible, probabilistic rather than deterministic way.
- The main disadvantage when using *SCEM-UA* seems to be the computational effort required to converge to a posterior target distribution. This effort, however, is not expected to be significantly different from the alternative approaches where e.g. *Genetic Algorithms* are used to determine the (mean) leak areas followed by the *First Order Second Moment* model application to calculate the associated uncertainties. Also, computational time could always be reduced by using parallel and/or faster computers.

7.3 Recommendations for further work

Regarding the calibration of WDS hydraulic models the recommendations for further research are as follows:

- Incorporation of another leak model type into the calibration procedure. Several leak models exist in literature (e.g. Tucciarelli et al., 1999) introducing some additional information about the network (number of pipes, pipe length and its age or break probability rate that are attached to a particular node). As this procedure normally increases the number of leak candidates (or other type of calibration parameters) the statistical spread-sheet type analysis is recommended at first.
- Alternative *WDS* models (e.g. extended period simulation) might come into effect when analysing real life systems.
- Use of additional type of measurements (e.g. flows). Quite often the real network is measured at the same time against pressures and flows. Using this additional data in a proper way can bring in new (more accurate) information about the calibration procedure.
- The use of engineering knowledge and experience in the leak detection process should be no way disregarded. Every effort should be made to improve its effectiveness and usability in the calibration procedure.

Kokkuvõte

Doktoritöös kasutatakse veevõrgumudeli kalibreerimiseks tõenäosusteooriat, kuna arvutusmudeli algandmed sisaldavad endas alati ebamäärasusi. Vedelikuga täidetud torusüsteemi lekete määramiseks kasutatud tõenäosusel põhinevat mudelit pole varem sellisel kujul esitatud. Peamised töö eesmärgid on: (a) Rakendada tõenäosusel põhinevat kalibreerimise algoritmi veevõrgu hüdraulilises mudelis, kuna kättesaadavate andmete puudulikkus ei anna alati deterministlikul lähenemisviisil ühtset, stabiilset lahendit. Tõenäosusel põhinev lähenemisviis kalibreerimisel võib anda lahendi ka puudulike andmete korral, samas edasine andmete täiendamine võimaldab tulemuste ebamäärasusi vähendada, mis on oluline lisaväärtus võrreldes traditsiooniliste meetoditega; (b) Tõestamaks meetodi praktilisi eeliseid, esitatakse erinevate näidete varal (tehis- ja reaalne veevõrk) selle rakendamise võimalusi.

Veevõrgu süsteemile on rakendatud *Bayes'i* tüüpi kalibreerimise algoritmi *Shuffled Complex Evolution Metropolis* (*SCEM-UA*). Algoritmi peamised karakteristikud on: (a) sihifunktsioon põhineb vähimruutude meetodil; (b) *SCEM-UA* algoritmi on varasemalt testitud hüdroloogiliste mudelite juures, kus rõhutatakse selle suurepäraseid statistilise omadusi, mida saab edukalt ära kasutada tulemuste esitamiseks ning kalibreerimisülesande kui terviku lahendamiseks; (c) doktoritöös kasutatakse nii statsionaarset kui ajas muutuvat hüdraulilist mudelit; (d) oluline rõhk on pööratud mudeli kalibreerimisele ebamäärastes tingimustes; (e) doktoritöös ei pöörata tähelepanu andmete tundlikkusele, kuna enamikes veevõrgusüsteemides on kasutusel fikseeritud mõõtepunktid, samas võimaldatakse tulemuste parandamist peale lisamõõtmiste teostamist; (f) kalibreerimise algoritmi kasutatakse erinevatel testvõrkudel, kus iga eraldi võetud osa täidab kindlat eesmärki töös.

Doktoritöö peamine panus veevõrgu modelleerimises väljendub järgmistes põhipunktides: (a) uudse, tõenäosusel põhineva, lekete määramise meetodi kasutusele võtmine ühes kasutajapõhiste määrangutega, millega defineeritakse otsitavate lekete parameetrid ning iseloom mistahes torustiku süsteemis; (b) võimalus valida enne arvutuse läbiviimist mudeli vea piirid, mis ennekõike leiab suuremat kasutusvõimalust reaalsetel veevõrgumudelitel; (c) tõenäosusfunktsiooni saab jooksvalt uuendada, kui lisainformatsioon mudeli kohta tehakse kättesaadavaks (sh lisamõõtmised); (d) võimalus teostada järelanalüüsi, mis keskendub kindlale (kitsamale) veevõrgu piirkonnale, kus varasemalt leiti kõige tõenäosemad lekked (kasutades olemasolevat andmepanka kitsamas piirkonnas võimaldab efektiivsemalt leida kriitilisemad lekked).

Doktoritöös esitatud meetod võimaldab kaasata ka kogemustel põhinevaid teadmisi, mida saab lekete puhul väljendada eeldefineeritava maksimaalse lekke suurusena mingis piirkonnas. Kasutades *Bayes*' rekursiivset protseduuri võimaldatakse uuendada tõenäosusfunktsioone lisaandmete saamisel. Üks suuremaid puuduseid deterministlike mudelite juures on asjaolu, et tulem esitatakse ühe ainsa numbrina, arvestamata seejuures mudeli ebamäärasusi.

Peamised eelised kasutamaks *Bayes'* tüüpi *SCEM-UA* algoritmi lekete kalibreerimisel on järgmised: (a) nii lekke suurus kui selle viga (ehk ebamäärasus) määratakse ühe ja sama arvutusprotseduuri käigus (veelgi enam, arvutuse käigus määratakse tõenäosusfunktsioon otsitavale suurusele); (b) *SCEM-UA* algoritm ei vaja järeltöötlusel põhinevat (nt *First Order Second Moment*) protseduuri ega ka nendega seotud eelduste (lineaarsus, reeglipärasus) kasutamist ebamäärasuste kindlaks tegemiseks; (c) *SCEM-UA* algoritm ei vaja tuletiste arvutamist, mis võib põhjustada arvutusvigu; (d) *Bayes'* rekursiivse protseduuri kasutamine *SCEM-UA* algoritmis võimaldab määratleda kalibreerimise lähteparameetrite piire oluliselt paindlikumal viisil. Üks suuremaid puuduseid *SCEM-UA* algoritmi juures on selle mahukas ressursivajadus, et saavutada lähteandmetest tulenev sihtjaotus. Samas ei ole antud juhul tegemist erandiga, kuna teised alternatiivsed algoritmid (nt geneetiline algoritm) kasutavad nn järeltöötlust arvutamaks seotud ebamäärasused. Doktoritöös kasutatavat algoritmi saab tööle rakendada korraga mitmes arvutis, kus iga arvutuse eesmärgiks on seatud üksteisest erinev lähtepunkt. Viimaste kombineerimisel (kogudes kokku tulemid erinevatest arvutitest) saab leida ühtse (summaarse) lahendi.

Lähtuvalt veevõrgu hüdraulilise mudeli kalibreerimisest võib edasise töö eesmärkideks võtta järgmised täiendused: (a) mõne teise üldtuntud lekke mudeli lisamine kalibreerimise algoritmi, mis ühtlasi seob lekke otsimise veel mõne veevõrgumudeli parameetriga (nt torude arv, toru pikkus, toru vanus ning toru purunemise tõenäosus) – kuna antud protseduuri lisamine üldjuhul kasvatab lekke kandidaatide arvu (või üldiselt kalibreeritavate parameetrite arvu), siis esmalt võiks soovitada tulemuste järelanalüüsi tabelarvutuste põhiselt; (b) alternatiivse veevõrgumudeli kasutamine (dünaamiline mudel); (c) andmete lisatüüpide kaasamine (nt vooluhulk) – üsna tihti mõõdetakse reaalses võrgus nii rõhku kui ka vooluhulkasid, mida saaks edukalt ära kasutada kalibreerimise täpsustamisel.

Bibliography

Papers presented by the candidate

Vassiljev, A., Minguell Font, L., Puust, R. (2007). Use of pressure differentials for calibration of the operational water distribution system. In: Water Management Challenges in Global Change: International conference of computing and control for the water industry (CCWI2007) and suistainable urban water management (SUWM2007), B. Ulanicki, K. Vairavamoorthy, D. Butler, P. L. M. Bounds, F. Ali Memon, eds., Taylor & Francis, UK, 205 - 209.

Koppel, T., Ainola, L., Puust, R., (2007). A mathematical model for the determination of leakage in mains and water distribution networks. Proceedings of the Estonian Academy of Sciences. Engineering, 13(1), 3 - 16.

Puust, R., Kapelan, Z., Savic, D., Koppel, T. (2006). Probabilistic leak detection in pipe networks using the SCEM-UA algorithm. In: Proceedings of WDSA 2006: Water Distribution System Analysis Symposium, Cincinnati, Ohio, USA, CDROM Edition.

Vassiljev, A., Koppel, T., Puust, R. (2005). Calibration of the model of an operational water distribution system. In: Proceeding of the 8th International Conference on Computing and Control for the Water Industry (CCWI): 8th International Conference on Computing and Control for the Water Industry (CCWI), D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume I, 155-159.

Koppel, T., Puust, R., Ainola, L. (2004). An inverse approach for leakage identification in water distribution systems. In: XXIII Nordic Hydrological Conference, A. Järvet, ed., Tallinn, Estonia, Tartu University Press, NHP report 48, Volume I, 342-349.

List of references

Abhulimen, K. E. and Susu, A. A. (2004). Liquid pipeline leak detection system: model development and numerical simulation. *Chemical Engineering Journal*, 97, 47-67.

Ainola, L., Koppel, T., Tiiter, K., Vassiljev, A. (2000). Water network model calibration based on grouping pipes with similar leakage and roughness estimates. *Proceedings of Joint Conference on Water Resources Engineering and Water Resources Planning & Management*, Minneapolis, Minn., USA.

Ainola, L., Koppel, T., Kändler, N. and Vassiljev, A. (2001). Principle of locality and its application to water network model calibration. *Proceedings of Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume I, 111-120.

Ainola, L., Koppel, T., Kändler, N. (2003). Diurnal dynamic analysis of water distribution networks. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 165-171.

Almandoz, J., Cabrera, E., Gil, J. and Pellejero, I. (2003). Evaluation of leakage by means of night flow measurements and analytical discrimination. A comparative study. *Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management)*, E. Cabrera and E. Cabrear Jr., eds., Volume I, 327-341.

Almandoz, J., Cabrera, E., Arregui F., Cabrere Jr., E. and Cobacho, R. (2005). Leakage assessment through water distribution network simulation. *Journal of Water Resources Planning and Management*, Vol. 131, No. 6, November 1, 458-466.

Alonso, J. M., Alvarruiz F., Guerrero, D., Hernandez, V., Ruiz, P. A., Vidal, A. M., Martinez, F., Vercher, J. and Ulanicki, B. (2000). Parallel computing in water network analysis and leakage minimization. *Journal of Water Resources Planning and Management*, Vol. 126, No. 4, July/August, 51-260.

American Water Works Association (AWWA). (1992). Water audit and leak detection guidebook. Department of Water Resources, State of California, USA, p. 164.

American Water Works Association Water Loss Control Committee (AWWA). (2003). "Applying worldwide BMPs in water loss control." *J. Am. Water Works Assoc.*, 95 (8), 65-79.

Andersen, J. H. and Powell, R. S. (2000). Implicit state-estimation technique for water network monitoring. *Urban Water*, 2, 123-130.

Araujo, L. S., Coelho, S. T. and Ramos, H. M. (2003a). Estimation of distributed pressure-dependent leakage and consumer demand in water supply networks. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon (Eds.), *Swets & Zeitlinger*, Lisses, 119-128.

Araujo, L. S., Ramos, H. M. and Coelho, S. T. (2003b). Optimisation of the use of valves in a network water distribution system for leakage minimisation. Proceedings of the International Conference on Advances in Water Supply Management. Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 97-107.

Araujo, L. S., Ramos, H., Coelho, S. T. (2006). Pressure control for leakage minimisation in water distribution systems management. Water Resources Management, 20, pp. 133-149.

Arreguin-Cortes, F. I., Ochia-Alejo, L. H. (1997). Evaluation of water losses in distribution networks. Journal of Water Resources Planning and Management, Vol. 123, No. 5, September/October, pp. 284 – 291.

ATKINS (2005). IPART – Potential leakage requirements for Sydney Water. Final Report, 35 p.

Authority of the House of Lords (AHL). (2006). Water Management. Volume I, 8th *Report of Session 2005-06*, Science and Technology Committee, London, p. 161.

(AWDM) Advanced Water Distribution Modeling and Management. (2003). Haestad Methods, Haestad Press, 751 pp. 2003.

Babovic, V. and Drecourt, J. (2000). Data mining approach to pipe burst risk assessment. *Proceedings of International Symposium: Water Network Modelling for Optimal Design and Management*, G. Walters & D. Savic, eds., Centre for Water Systems, University of Exeter, pp. 93-107.

Beck, S. B. M., Curren, M. D., Sims, N. D. and Stanway, R. (2005a). Pipeline network features and leak detection by cross-correlation analysis of reflected waves. *Journal of Hydraulic Engineering*, Vol. 131, No. 8, August 1, 715-723.

Beck, S. B. M., Bagshaw, N. M., and Yates, J. R. (2005b). Explicit equations for leak rates through narrow cracks. *International Journal of Pressure Vessels and Piping*, 82, 565-570.

Beuken, R. H. S., Lavooij, C. S. W., Bosch, A., Schaap, P. G. (2006). Low leakage in the Netherlands confirmed. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Box, G. E. P., and Tiao, G. C. (1973). Bayesian Inference in Statistical Analysis, *Addison-Wesley-Longman*, Reading, Mass., p. 601.

Bracken, M. and Hunaidi, O. (2005). Practical aspects of acoustical leak location on plastic and large diameter pipe. *Leakage 2005 Conference Proceedings*, Halifax ,Canada.

Brunone, B. and Ferrante, M. (1999). On leak detection in single pipes using unsteady-state tests. *Proceedings of the IASTED International Conference: Modeling and Simulation (MS '99)*, Philadelphia, Pennsylvania, USA, 268-272.

Brunone, B. (1999). Transient Test-Based Technique for Leak Detection in Outfall Pipes. *Journal of Water Resources Planning and Management*, *ASCE*, 125(5), 302-306.

Brunone, B. and Ferrante, M. (2001). Detecting leaks in pressurized pipes by means of transients. *Journal of Hydraulic Research*, 39(5), 539 – 547.

Brunone, B. and Ferrante, M. (2004). Pressure wave as a tool for leak detection in closed conduits. *Urban Water Journal*, 1(2), 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, July 1, 321-329.

Burrows, R., Mulreid, G. and Hayuti, M. (2003). Introduction of a fully dynamic representation of leakage into network modelling studies using Epanet. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 109-118.

Caputo, A.C. and Pelagagge, P.M. (2002). An inverse approach for piping networks monitoring. *Journal of Loss Prevention in the Process Industries*, 15, 497-505.

Carpentier, P., Cohen, G. (1991). State estimation and leak detection in water distribution networks. Civil Engineering Syst., 1991, Vol. 8, pp. 247 - 257.

Cascetta, F. and Vigo, P. (1992). Location and assessment of water leakage. *Measurement + Control*, Volume 25, 269 – 275.

Cassa, A. M., Van Zyl, J. E. and Laubscher, R. F. (2005). A numerical investigation into the behaviour of leak openings in uPVC pipes under pressure. *Proceedings of*

CCWI2005, Water Management for the 21st Century, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, pp. 155 – 160.

Chen, L.-C. (1995). Pipe network transient analysis – the forward and inverse problems, *PhD Thesis*, Faculty of the Graduate School, Cornell University, p. 163.

Cheung, P. B., Van Zyl, J. E., Reis, L. F. R. (2005). Extension of EPANET for pressure driven demand modeling in water distribution system. Proceedings of CCWI2005, Water Management for the 21st Century, University of Exeter, UK, 2005, Volume 2, pp. 143 – 148.

Cieniawski, S. E., Wayland, E., and Ranjithan, S. (1995). Using genetic algorithms to solve a multi-objective groundwater monitoring problem. *Water Resour. Res.,* 31(2), 399-409.

Cist, D. B., Schutz, A. E. (2001). State of the art for pipe & leak detection. *Report by Geophysical Survey Systems, Inc.*, 8p.

Coetzer, A. J., van Zyl, J. E., Clayton, C. R. I. (2006). An experimental investigation into the turbulent-flow hydraulics of small circular holes in plastic pipes. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA.

Colombo, A. F. and Karney B. W. (2002). Energy and costs of leaky pipes: Toward comprehensive picture. *Journal of Water Resources Planning and Management*, Vol. 128, No. 6, November 1, 441-450.

Colombo, A. F. and Karney B. W. (2003). The labyrinth of water distribution systems: demand, energy and climate change. Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management). Edited by Enrique Cabrera and Enrique Cabrera Jr. Volume I, 2003, pp. 239 – 246.

Cooley, R.L., and Naff, R.L. (1990). Regression Modelling of Ground-Water Flow, Techniques of Water Resources Investigations. *USGS*, Book 3, Chapter B4, p. 232.

Covas, D., Ramos, H. (1999). Leakage Detection in Single Pipelines Using Pressure Wave Behaviour. *Proceedings of Water Industry Systems: modelling and optimization applications*, D. Savic and G. Walters, eds., *Research Studies Press Ltd.*, Baldock, Hertfordshire, England, Volume I, 287-299.

Covas, D., Stoianov, I., Butler, D., Maksimovic, Ć., Graham, N. and Ramos, H. (2001). Leakage detection in pipeline systems by inverse transient analysis – from theory to practice. *Proceedings of Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume II, 3-16.

Covas, D. and Ramos, H. (2001). Hydraulic Transients used for Leakage Detection in Water Distribution Systems. *4th Conference on Water Pipeline Systems: Managing Pipeline Assets in an Evolving Market,* York, UK, 227-241.

Covas, D., Graham, N., Maksimović, Ć., Ramos, H., Kapelan, Z., Savic, D. and Walters, G. (2003). An assessment of the application of inverse transient analysis for leak detection: Part II – Collection and application of experimental data. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 79-87.

Covas, D., Ramos, H. and de Almeida, A. B. (2005a). Standing wave difference method for leak detection in pipeline systems. *Journal of Hydraulic Engineering*, Vol. 131, No. 12, December 1, 1106-1116.

Covas, D., Ramos, H., Young, A., Graham, N. and Maksimovic, C. (2005b). Uncertainties of leak detection by means of hydraulic transients: from the lab to the field. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, 143 – 148.

Covas, D., Ramos, H., Lopes, N., Almeida, A. B. (2006). Water pipe system diagnosis by transient pressure signals. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Cui, L.-J., and Kuczera, G. (2003). Optimizing urban water supply headworks using probabilistic search methods. *Journal of Water Resources Planning and Management*, Vol. 129, No. 5, September 1, 380 – 387.

Dandy, G. C., Simpson, A. R., and Murphy, L. J. (1996). 'An improved genetic algorithm for pipe network optimization. *Water Resour. Res.,* 32(2), 449-458.

Da Silva, H. V., Morooka, C. K., Guilherme, I. R., da Fonseca, T. C. and Mendes, J. R. P. (2005). Leak detection in petroleum pipelines using a fuzzy system. *Journal of Petroleum Science and Engineering*, 49, 223-238.

Davidson, J.W., and Goulter, I. C. (1995). 'Evolution program for design of rectilinear branched networks. *J. Comput. Civ. Eng.,* 9(2), 112-121.

Dawsey, W. J., Minsker, B. S., VanBlaricum, V. L. (2006). Bayesian belief networks to integrate monitoring evidence of water distribution system contamination. *Journal of Water Resources Planning and Management*, Vol. 132, No. 4, July 1, 234-241.

Dias, N., Covas, D. and Ramos, H. (2005). Leakage reduction and customer level of service improvement by pressure management. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, 137-142.

Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.,* 28(4), 1015-1031.

Duan, Q., Sorooshian, S., and Gupta, V. K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J. Hydrol.,* 158, 265-284.

Dunlop, E. J. (1999). A dynamic hydraulic model for water distribution network simulation. *Proceedings of Water Industry Systems: modelling and optimization applications*, D. Savic and G. Walters, eds., *Research Studies Press Ltd.*, Baldock, Hertfordshire, England, Volume I, 103-123.

The Economist. (2000). In praise of Bayes. Sept. 30, 356(8190), p. 83.

EPSRC Engineering Programme Network in Trenchless Technology (2002). Underground mapping, pipeline location technology and condition assessment. The University of Birmingham – Infrastructure Engineering and Management Research Centre, p. 77.

Esat, V., and Hall, M. J. (1994). Water resources system optimization using genetic algorithms. *Hydroinformatics '94 Proc., 1st Int. Conf. on Hydroinformatics*, A. Verwey, et al., eds., Balkema, Rotterdam, The Netherlands, 225-231.

Eshelman, L. J., and Schaffer, J. D. (1991). Preventing premature convergence in genetic algorithm by preventing incest. *Proc., 4th Int. Conf. on Genetic Algorithms*, R. Belew and L. Booker, eds., *Morgan Kaufman* (San Mateo), 115-122.

Fahmy, H. S., King, J. P., Wentzel, M. W., and Seton, J. A. (1994). Economic optimization of river management using genetic algorithm. *Proc., ASAE Int. Summer Meeting*, Paper No. 943034, American Society of Agricultural Engineers, St. Joseph, Mich.

Farley, M., Trow, S. (2003). Losses in water distribution networks. *IWA Publishing*, London, p. 282.

Farmer, E., Kohlrust, R., Myers, G. and Verduzco, G. (1988). Leak detection tool undergoes field tests. *Oil & Gas Journal*, Dec 19, 48-53.

Ferrante, M., Brunone, B., Rossetti, A. G. (2001a). Harmonic analysis of pressure signal during transients for leak detection in pressurized pipes. *In: Proceedings ofthe 4th International Conference on Water Pipeline Systems*, York, UK. *BHR Group Limited*, 259-276.

Ferrante, M., Brunone, B., Rossetti, A. G. (2001b). Leak detection in pressurized pipes by means of wavelet analysis. *In: Proceedings of the 4th International Conference on Water Pipeline Systems*, York, UK. *BHR Group Limited*, 243-255.

Ferrante, M., Brunone, B. (2003a). Pipe system diagnosis and leak detection by unsteady-state tests. 1. Harmonic analysis. *Advances in Water Resources*, 26, 95-105.

Ferrante, M., Brunone, B. (2003b). Pipe system diagnosis and leak detection by unsteady-state tests. 2. Wavelet analysis. *Advances in Water Resources*, 26, 107-116.

Ferrante, M., Brunone, B., Meniconi, S. and Almadori, C. (2005). Wavelet analysis of numerical pressure signals for leak monitoring. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 1, 329 – 334.

Field, D. B., Ratcliffe, B. (1978). Location of leaks in pressurised pipelines using sulphur hexafluoride as a tracer. *Water Research Centre*, *Technical Report 80*, p. 30.

Franchini, M. (1996). Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall runoff models. *Journal of Hydrological Science,* 41(1), 21-40.

Gabrys, B. and Bargiela A. (1999). Neural networks based decision support in presence of uncertainties. *Journal of Water Resources Planning and Management*, Vol. 125, No. 5, September/October, 272-280.

Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., and Hunaidi, O. (2004). A model of the correlation function of leak noise in buried plastic pipes. *Journal of Sound and Vibration*, 277, 133-148.

Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., and Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. *Journal of Sound and Vibration*, 283, 927 – 941.

Gao, Y., Brennan, M. J., Joseph, P. F. (2006). A comparison of time delay estimators for the detection of leak noise signals in plastic water distribution pipes. *Journal of Sound and Vibration*, 292, 552–570.

Garcia, V. J., Cabrera, E., Cabrera (Jr.), E. (2006). The minimum night flow method revisited. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Garzon-Contreras, F., Thornton, J. (2006). Pressure-leakage relationships in urban water distribution systems. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Gelman, A., Rubin, D. B. (1992). Inference from iterative simualtion using multiple sequences. *Statist. Sci.*, 7, 457-472.

Germanopoulos, G. (1985). A technical note on the inclusion of pressure dependent demand and leakage terms in water supply network models. *Civ. Engng Syst.*, Vol. 2. 171-179.

Germanopoulos, G. (1995). Valve control regulation for reducing leakage. *Improving efficiency and reliability in water distribution systems*, E. Cabrera, A. F. Vela, eds., *Kluwer Academic Publishers*, 165-188.

Goldberg, D. E., and Kuo, C. H. (1987). Genetic algorithms in pipeline optimization. *J. Comput. Civ. Eng.*, 1(2), 128-141.

Goldberg, D. E. (1989). Genetic algorithms in search, optimization and machine learning, *Addison-Wesley*, Reading, Mass.

Goodwin, S. J. (1980). The results of the experimental programme on leakage and leakage control. *Water Research Centre, Technical Report 154*, p. 52.

Greyvenstein, B. and van Zyl, J. E. (2005). An experimental investigation into the pressure leakage relationship of some failed water pipes. *Leakage 2005 Conference Proceedings*, Halifax ,Canada.

Grunwell, D., Ratcliffe, B. (1981). Location of underground leaks using the leak noise correlator. *Water Research Centre, Technical Report 157*, p. 40.

Hernandez, V., Martinez, F., Vidal, A. M., Alonso, J. M., Alvarruiz, F., Guerrero, D., Ruiz, P. A. and Vercher, J. (1999). HIPERWATER: A high performance computing EPANET-based demonstrator for water network simulation and leakage minimisation. *Proceedings of Water Industry Systems: modelling and optimization applications*, D.Savic and G. Walters, eds., *Research Studies Press Ltd.*, Baldock, Hertfordshire, England, Volume I, 141-153.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97-109.

Hesssel, G., Schmitt, W., Van Der Vorst, K. and Weiss, F.-P. (1999). A neural network approach for acoustic leak monitoring in the VVER-440 pressure vessel head. *Progress in Nuclear Energy*, Vol. 34, No. 3, 173-189.

Hill, M.C. (1998). Methods and Guidelines for Effective Model Calibration. *Report No. 98-4005*, USGS, p. 90.

Holnicki-Szulc, J., Kolakowski, P., and Nasher, N. (2005). Leakage Detection in Water Networks, *Journal of Intelligent Material Systems and Structures*, Vol. 16, issue 3, pp. 207-219.

Hope, W. (1892). The waste of water in public supplies, and its prevention. P*roc. Instn. Civ. Engrs. Mun. Engr.*, (republished 1996), 115, 68-75.

Hough, J. E. (1988). Leak testing of pipelines uses pressures and acoustic velocity. *Oil & Gas Journal*, Nov 21, 35-41.

Hunaidi, O., and Chu, W. T. (1999). Acoustical characteristics of leak signals in plastic water distribution pipes. *Applied Acoustics*, 58, 235-254.

Hunaidi, O., Chu, W., Wang, A., Guan, W. (2000). Detecting leaks in plastic pipes. Journal AWWA, Volume 92 (2), pp. 82-94.

Hunaidi, O., Bracken, M., Gambino, T., Fricke, C. (2004). Acoustic methods for locating leaks in municipal water pipe networks. *International Conference on Water Demand Management*, Dead Sea, Jordan, 1-14.

Hunaidi, O. (2005). Economic comparison of periodic acoustic surveys and DMA-based leakage management strategies. Leakage 2005 Conference Proceedings, Halifax ,Canada, 2005, 12 – 15 September.

Jönsson, L. (1995). Leak detection in pipelines using hydraulic transients – laboratory measurements. *Department of Water Resources Engineering*, University of Lund, Sweden, p. 40.

Jönsson, L. (2003). Hydraulic transients in a pipeline with a leak. *VA-FORSK Rapport Nr. 28*, p. 322.

Kapelan, Z., Savic, D. A. and Walters, G. A. (2000). Inverse transient analysis in pipe networks for leakage detection and roughness calibration. *Proceedings of International Symposium: Water Network Modelling for Optimal Design and Management*, G. Walters and D. Savic, eds., Centre for Water Systems, University of Exeter, 143-159.

Kapelan, Z. (2002). Calibration of WDS hydraulic models, *PhD Thesis*. University of Exeter, Exeter, p. 334.

Kapelan, Z. S., Savic, D. A., Walters, G. A. (2003a). A hybrid inverse transient model for leakage detection and roughness calibration in pipe networks. *Journal of Hydraulic Research*, Vol. 41, No. 5, 481-492.

Kapelan, Z., Savic, D., Walters, G. A., Covas, D., Graham, I. N. and Maksimović, Ć. (2003b). An assessment of the application of inverse transient analysis for leak detection: Part I – Theoretical considerations. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 71-78.

Kapelan, Z., Savic, 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*, 1(2) 129-143.

Kapelan, Z., Savic, D. A., and Walters, G. A. (2005). Advanced calibration of water distribution models using the Bayesian type procedure. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 1, 167-172.

Kim, S. H. (2005). Extensive development of leak detection algorithm by impulse response method. *Journal of Hydraulic Engineering*, Vol. 131, No. 3, March 1, 201-208.

Lambert, A, Brown T., Takizawa, M. and Weimer, D. (1999). To review of indicators for real losses from to water supply systems. *IESA, J Water SRT – Aqua*, 48. December, 227-237.

Lambert, A. (2003). Assessing non-revenue water and its components: a practical approach. Water 21, August 2003, pp. 50 – 51.

Lee, P. J., Vitkovský, J. P., Lambert, M. F., Simpson, A. R., Liggett, J. A. (2003). Frequency response coding for the location of leaks in single pipeline systems. Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management). Edited by Enrique Cabrera and Enrique Cabrera Jr. Volume I, 2003, pp. 371 – 378.

Lee, P. J., Vitkovsky, 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, July 1, 596-604.

Lee, P.J., Vitkovsky, J. P., Lambert, M. F., Simpson, A. R. and Liggett J. (2005b). Leak location using the pattern of the frequency response diagram in pipelines: a numerical study. *Journal of Sound and Vibration*, 284, 1051-1073.

Lesyshen, R. M. (2005). Water transmission line leak detection using extended Kalman filtering, *PhD Thesis*, University of Saskatchewan, Saskatoon, Canada, p. 118.

Lee, P. J., Lambert, M. F., Simpson, A. R., Vitkovský, J. P., Liggett, J. (2006). Experimental verification of the frequency response method for pipeline leak detection. Journal of Hydraulic Research, Vol. 44, No. 5, pp. 693-707.

Liggett, A. (1993). Network monitoring and the algorithmic location of leaks under steady and unsteady conditions. *Water supply systems: State of the art and future trends*, E. Cabrera and F. Martinez, eds., *Computational Mechanics Publication*, Souhtampton, 253-270.

Liggett, J.A., and Chen, L.-C. (1994). Inverse transient analysis in pipe networks. *Journal of Hydraulic Engineering*, *ASCE*, 120(8), 934-955.

Liggett, J. A., and Chen, L.-C. (1995). Monitoring water distribution systems. The inverse method as a tool for calibration and leak detection. *Improving efficiency and reliability in water distribution systems*, E. Cabrera, A. F. Vela, eds., *Kluwer Academic Publishers*, 107-132.

Liou, J. C. P. and Tian, J. (1994). A Comparision of Two Transient-Based Leak Detection Algorithms. *The 1994 International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, Chicago, Illinois, FED-Vol. 198 and PVP-Vol. 291, 35-37.

Liou, J. C. P. (1994). Mass Imbalance Error of Waterhammer Equations and Leak Detection. *Journal of Fluids Engineering*, March, Vol. 116, 103-109.

Liou, J. C. P. and Tian, J. (1995). Leak Detection – Transient Flow Simulation Approaches. *Journal of Energy Resources Technology*, September, Vol. 117, 243-248.

Liou, C. P. (1998). Pipeline leak detection by impulse response extraction. Journal of Fluids Engineering, Vol. 120, December, pp. 833-838.

Liou, J. C. P., Hall, R. J., and McMahon, M. C. (2003). Hazardous liquid leak detection techniques and processes. *General Physics Corporation, Report No. DTRS56-02-D-70037-01*. April, p. 47.

Liu, K. T. H. (1969). The numerical analysis of water supply networks by digital computer. *Proc. 13th Congress, International Association for Hydraulic Research*, Vol. 1, Sep., 36-42.

Lockwood, A., Murray, T., Stuart, G. and Scudder, L. (2003). A study of geophysical methods for water leak location. *Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management)*, E. Cabrera and E. Cabrear Jr., eds., Volume I, 379-386.

Makar, J.M. and Chagnon, N. (1999). Inspecting systems for leaks, pits, and corrosion. Journal of American Water Works Association, Vol. 91, No. 7, pp. 36-46.

Mamlook, R., Al-Jayyousi, O. (2003). Fuzzy sets analysis for leak detection in infrastructure systems: a proposed methodology. Clean Techn Environ Policy, 6, pp. 26 – 31.

Martinez, F., Signes, M., Savall, R., Andrés, M. and Ponz, R. (1999). Construction and use of a dynamic simulation model for the Valencia metropolitan water supply and distribution network. *Proceedings of Water Industry Systems: modelling and optimization applications*, D. Savic and G. Walters, eds., *Research Studies Press Ltd.*, Baldock, Hertfordshire, England, Volume I, 155-174.

May, J. (1994). Pressure Dependent Leakage. *World Water and Environmental Engineering*, October 1994, p. 10.

McInnis, D., Karney, B. W. (1995). Transients in distribution networks: field tests and demand models. Journal of Hydraulic Engineering. 1995, Vol. 121 (3): p. 218 – 231.

McKinney, D. C., and Lin, M. D. (1994). Genetic algorithm solutions of groundwater management models. *Water Resour. Res.,* 30(6), 1897-1906.

McNulty, J. G. (2001). An acoustic-based system for detecting, locating and sizing leaks in water pipelines. *4th International Conference on WATER PIPELINE SYSTEMS Managing Pipeline Assets in an Evolving Market*, York, UK, 217-225.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21, 1087-1091.

Miller, R. K., Pollock, A. A., Watts, D. J., Carlyle, J. M., Tafuri, A. N. and Yezzi Jr, J. J. (1999). A reference standard for the development of acoustic emission pipeline leak detection techniques. *NDT&E International*, 32, 1-8.

Misiunas, D., Vitkovsky, J., Olsson, G., Simpson, A. R. and Lambert, M. F. (2003). Pipeline burst detection and location using a continuous monitoring technique. *Proceedings of the International Conference on Advances in Water Supply Management,* Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, pp. 89-96.

Misiunas, D., Vitkovsky, J., Olsson, G., Simpson, A. and Lambert, M. (2005a). Pipeline break detection using pressure transient monitoring. *Journal of Water Resources Planning and Management*, Vol. 131, No. 4, July 1, 316-325.

Misiunas, D., Vitkovsky, J., Olsson, G., Lambert, M. and Simpson, A. (2005b). Failure monitoring in water distribution networks. *Proc. 2nd IWA Conference on Instrumentation, Control and Automation*, Busan, Korea.

Misiunas, D., Lambert, M., Simpson, A. (2006). Transient-based periodical pipeline leak diagnosis. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Montesinos, P., Garcia-Guzman, A., and Ayuso, J. L. (1999). Water distribution network optimization using a modified genetic algorithm. *Water Resour. Res.,* 35(11), 3467-3473.

Morris Jr., R. E. (1967). Principal Causes and Remedies of Water Main Breaks. *Southwest Section Meeting*, Austin, Texas, 782-798.

Mounce, S. R., Machell, J. (2006). Burst detection using hydraulic data from water distribution systems with artificial neural networks. Urban Water Journal, Vol. 3, No. 1, March, pp. 21-31.

Moyer, E., Male, J. W., Moore, C., and Hock, G. (1983). The economics of leak detection and repair – a case study. *American Water Works Association*, January, 29-35.

Mpesha, W., Gassman, S.L. and Chaudhry, M. H. (2001). Leak detection in pipes by frequency response method. *Journal of Hydraulic Engineering*, Vol. 127 (2), 134-147.

Mpesha, W., Chaudhry, M. H., 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.

Muggleton, J. M., Brennan, M. J., Linford, P. W. (2004). Axisymmetric wave propagation in fluid-filled pipes: wavenumber measurements in vacuo and buried pipes. *Journal of Sound and Vibration*, 270, 171-190.

Muggleton, J. M., Brennan, M. J. (2004). Leak noise propagation and attenuation in submerged plastic water pipes. *Journal of Sound and Vibration*, 278, 527-537.

Muggleton, J. M., and Brennan, M. J. (2005). Axisymmetric wave propagation in buried, fluid-filled pipes: effects of wall discontinuities. *Journal of Sound and Vibration*, 281, 849-867.

Muggleton, J. M., Brennan, M. J., Pinnington, R. J., Gao, Y. (2006). A novel sensor for measuring the acoustic pressure in buried plastic water pipes. *Journal of Sound and Vibration,* 295, 1085-1098.

Mukherjee, J., Narasimhan, S. (1996). Leak detection in networks of pipelines by the generalized likelihood ratio method. Ind. Eng. Chem. Res., 35, pp. 1886 – 1893.

Mukhopadhyay, S. and Srivastava, G.P. (2000). Characterisation of metal loss defects from magnetic flux leakage signals with discrete wavelet transform. NDT&E International, 33, pp. 57-65.

Murphy, L. J., Simpson, A. R., and Dandy, G. C. (1993). Design of a network using genetic algorithms. *Water, J. Aust. Water Wastewater Assoc.*, 20(4), 40-42.

Nicholas, R. E. (1990). Leak detection on pipelines in unsteady flow. *In: Forum on unsteady flow*, *Winter annual meeting of ASME*, New York, p. 23–5.

Noack, C., Ulanicki, B. (2006). Modelling of soil diffusibility on leakage characteristics of buried pipes. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

O'Brien, E., Murray, T. and McDonald, A. (2003). Detecting leaks from water pipes at a test facility using ground-penetrating radar. *Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management)*, E. Cabrera and E. Cabrear Jr., eds., Volume I, 395-403.

O'Day D.K. (1982). Organizing and analyzing leak and break data for making main replacement decisions. *Journal of the American Water Works Association (AWWA)*, November, 589-594.

Oliveira, R., and Loucks, D. P. (1997). Operating rules for multireservoir systems. *Water Resour. Res.,* 33(4), 839-852.

Pilcher, R. (2003). Leak detection practices and techniques: a practical approach. *Water 21*, December, 44-45.

Pilcher, R., Hamilton, S., Chapman, H., Ristovski, B., Strapely, S. (2007). Leak location and repair guidance notes. International Water Association. Water Loss Task Forces: Specialist Group Efficient Operation and Management, p. 67.

Poulakis Z., Valougeorgis, D. and Papadimitriou, C. (2003). Leakage detection in water pipe networks using a Bayesian probabilistic framework. *Probabilistic Engineering Mechanics*, 18, 315-327.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1988). Numerical Recipes: The Art of Scientific Computing. *Cambridge University Press*, p. 702.

Pudar, R. S., Liggett, J. A. (1992). Leaks in Pipe Networks. *Journal of Hydraulic Engineering, ASCE*, 118(7), 1031-1046.

Puust, R., Kapelan, Z., Savic, D., Koppel, T. (2006). Probabilistic leak detection in pipe networks using the SCEM-UA algorithm. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Rahman, S., Ghadiali, N., Wilkowski, M., Paul, D. (1997). A computer model for probabilistic leak-rate analysis of nuclear piping and piping welds. *Int. J. Pres. Ves. & Piping*, 70, 209-221.

Ramos, H., Reis, C., Ferreira, C., Falcao, C. and Covas D. (2001). Leakage control policy within operating management tools. *Proceedings of Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume I, 61-72.

Reis, L. F. R. and Chaudhry, F. H. (1999). Hydraulic characteristics of pressure reducing valves for maximum reduction of leakage in water supply networks. *Proceedings of Water Industry Systems: modelling and optimization applications*, D. Savic and G. Walters, eds., *Research Studies Press Ltd.*, Baldock, Hertfordshire, England, Volume I, 259-267.

Robert-Nicoud, Y., Raphael, B., Smith, I. F. C. (2005). Configuration of measurement systems using Shannon's entropy function. *Computer and Structures*, 83, 599-612.

Rossman, L.A. (2000). Epanet2 User Manual. Washington, USA: US EPA.

Rougier, J. (2005). Probabilistic leak detection in pipelines using the mass imbalance approach. *Journal of Hydraulic Research*, Vol. 43, No. 5, 556-566.

Saldarriaga, J. G., Fuentes, D. A. A., Galvis, L. F. C. (2006). Implementation of the hydraulic transient and steady oscillatory flow with genetic algorithms for leakage

detection in real water distribution networks. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Savic, D. A., and Walters, G. A. (1997). Genetic algorithms for least cost design of water distribution networks. *J. Water Resour. Plan. Manage.,* 123(2), 67-77.

SCEM-UA. SCEM-UA source code. http://www.science.uva.nl/ibed/ index_en.html, 28.06.2006.

Scott, S. L., Barrufet, M. A. (2003). Worldwide assessment of industry leak detection capabilities for single & multiphase pipelines. *Project report prepared for the minerals management service under the MMS/OTRC cooperative research agreement*, *1435-01-99-CA-31003*, *Task Order 18133*, August 6, p. 125.

Sharif, M., and Wardlaw, R. (2000). Multireservoir systems optimization using genetic algorithms: Case study. *J. Comput. Civ. Eng.,* 14(4), 255-263.

Shimanskiy, S., Iljima, T., and Naoi, Y. (2003). Development of microphone leak detection technology on Fugen NPP. *Progress in Nuclear Energy*, Vol. 43, No. 1-4, 357-364.

Silva, R. A., Buiatti, C. M., Cruz, S. L. and Pereira, J. A. F. R. (1996). Pressure wave behaviour and leak detection in pipelines. *Computers chem.. Engng.*, Vol. 20, 491-496.

Simpson, A. R. and Vitkovsky, J. P. (1997). A Review of Pipe Calibration and Leak Detection Methodologies for Water Distribution Networks. *Water in the Balance: Proceedings of the Australian Water & Wastewater Association 17th Federal Convention,* Melbourne, Australia, 680-687.

Stathis, J. A. and Loganathan, G.V. (1999). Analysis of pressure-dependent leakage in water distribution systems. *In: Preparing for the 21st Century*, *29th Annual Water Resources Planning and Management Conference*, E. M. Wilson, ed., Tempe, Arizona, USA, p. 43.

Stenberg, R. (1982). Leak detection in water supply systems. *The Swedish Water and Waste Water Works Association*, Sweden, p. 138.

Stoianov, I., Karney, B., Covas, D., Maksimović, Ć. and Graham, N. (2001). Wavelet processing of transient signals for pipeline leak location and quantification. *Proceedings of Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume II, 65-76.

Tabesh, M. and Delavar, M. R. (2003). Application of integrated GIS and hydraulic models for unaccounted for water studies in water distribution systems.

Proceedings of the International Conference on Advances in Water Supply Management, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 129 -135.

Tabesh, M., Asadiani Yekta, A. 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 CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, 125-130.

Taghvaei, M., Beck, S. B. M., Staszewski, W. J. (2006). Leak detection in pipelines using cepstrum analysis. Meas. Sci. Technol., 17, pp. 367-372.

Tallinna Vesi: *Annual Report* 2003, p. 90.

Tang, K. W., Karney, B. and Brunone, B. (2001). Leak detection using inverse transient calibration and GA – Some early successes and future challenges. *Proceedings of Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume II, 55-64.

Thornton, J. (2002). Water loss control manual. New York: McGraw-Hill. p. 646.

Thornton, J., Lambert, A. (2005). Progress in practical prediction of pressure: leakage, pressure: burst frequency and pressure: consumption relationships. Leakage 2005 Conference Proceedings, Halifax ,Canada, 2005, 12 – 15 September.

Todini, E. (2003). A more realistic approach to the 'extended period simulation' of water distribution networks. *Proceedings of the International Conference on Advances in Water Supply Management*, Maksimovic, Butler & Memon, eds., *Swets & Zeitlinger*, Lisses, 173-183.

Tucciarelli, T., Criminisi, A. and Termini, D. (1999). Leak Analysis in Pipeline Systems by Means of Optimal Valve Regulation. *Journal of Hydraulic Engineering*, *ASCE*, 125(3), 277-285.

Tuhovčák, L., Svoboda, M., Sviták, Z., Tothova, K. (2005). The technical audit of water distribution network using the different leakage indicators. Leakage 2005 Conference Proceedings, Halifax ,Canada, 2005, 12 – 15 September.

TWDB (2005). Water loss manual. Texas Water Development Board, Austin, Texas. 38 p.

TWDB (2007). Texas Water Development Board, Water audit reporting form, http://www.twdb.state.tx.us/assistance/conservation/Municipal/Water_Audit/HB3338ReportingForm.pdf, 03.04.2007.

Tylman, W., and Anders, G. J. (2006). Application of probabilistic networks for decision support in power system analysis. *Energy*, Volume 31, Issue 14, November, 2874-2889.

UK Water Industry (1994). Managing Leakage. Interpreting Measured Night Flow (Report E). *Engineering and Operations Committee*, October, p. 63.

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 Water Software Systems: theory and applications*, B. Ulanicki, B. Coulbeck and J. P. Rance, eds., Volume II, 41-53.

Ulanicki, B., Bounds, P. L. M., Rance, J. P. and Reynolds, L. (2000). Open and closed loop pressure control for leakage reduction. *Urban Water*, 2, 105-114.

US EPA. EPANET Programmer's Toolkit, http://www.epa.gov/nrmrl/wswrd/ /epanet.html, 28.12.2006.

Vairavamoorthy K. and Lumbers, J. (1998). Leakage reduction in water distribution systems: optimal valve control. *Journal of Hydraulic Engineering*, Vol. 124, No. 1, November, 1146-1154.

Van Zyl, J.E. and Clayton, C. R. I. (2005). The effect of pressure on leakage in water distribution systems. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, 131-136.

Verde, C. (2001). Multi-leak detection and isolation in fluid pipelines. *Control Engineering Practice*, 9, 673-682.

Verde, C. (2005). Accommodation of multi-leak location in a pipeline. *Control Engineering Practice*, 13, 1071-1078.

Vicini, A., and Quagliarella, D. (1998). Airfoil and Wing Design Through Hybrid Optimization Strategies. *Proc. 16th Applied Aerodynamics Conference,* Albuquerque, New Mexico.

Vitkovský, J. P. and Simpson, A. R. (1997). Calibration and Leak Detection in Pipe Networks Using Inverse Transient Analysis and Genetic Algorithms. *Research Report No. R 157*, August. Department of Civil and Environmental Engineering, The University of Adelaide, p. 97.

Vitkovsky, J. P., Simpson, A. R., Lambert, M. F. (1999). Leak detection and calibration of water distribution systems using transients and genetic algorithms.

Preparing for the 21st Century: Proceedings of the 26th Annual Water Resources Planning and Management Conference, *ASCE*, Tempe, Arizona.

Vítkovský, J.P., Simpson, A.R., Lambert, M. (2000). Leak Detection and Calibration Using Transients and Genetic Algorithms. *Journal of Water Resources Planning and Management*, 126(4), 262-265.

Vitkovský, J. P., Lambert, M. F., Simpson, A. R. and Wang, X. J. (2001). An Experimental Verification of the Inverse Transient Technique. *Conference on Civil Engineering Hydraulics*, *Institution of Engineers*, Hobart, Australia, 373-380.

Vitkovský, J. P., Liggett, J. A., Simpson A. R., Lambert, M. F. (2003a). Optimal measurement site locations for inverse transient analysis in pipe networks. *Journal of Water Resources Planning and Management*, Vol. 129, No. 6, November 1, 480-492.

Vitkovský, J. P., Lee, P. J., Stephens, M. L., Lambert, M. F. and Simpson, A. R. (2003b). Leak and blockage detection in pipelines via an impulse response method. *Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management)*, E. Cabrera and E. Cabrear Jr., eds., Volume I, 423-430.

Walski, T., Bezts, W., and Posluszny, E. T. (2006). Modeling leakage reduction through pressure control. *Journal AWWA*, April, 147-155.

Wang, G. Z., Fang, C. Z. and Wang, K. F. (1991). State Estimation and Leak detection and location in pipeline. *Industrial Electronics, Control and Instrumentation. Proceedings. IECON '91*, Vol. 1, 155-160.

Wang, Q. J. (1991). The genetic algorithm and its application to calibrating conceptual rainfall runoff models. *Water Resour. Res.*, 27(9), 2467-2472.

Wang, X.-J., Lambert, M. F., Simpson, A. R., Liggett, J. A. and Vitkovský, J. P. (2002). Leak detection in pipelines using the damping of fluid transients. *Journal of Hydraulic Engineering*, Vol 128 (7), 697-711.

Wang, X.-J., Simpson, A. R., Lambert, M. F., Vitkovský, J. P. and Liggett, J. A. (2003). Simulation of transients in a looped laboratory pipe network. *Proceedings of PEDS 2003 (Pumps, Electromechanical Devices and Systems Applied to Urban Water Management)*, E. Cabrera and E. Cabrear Jr., eds., Volume II, 791-798.

Wang, X.-J., Simpson, A. R., Lambert, M. F. (2006). An analytical solution for the transients in a pipeline with a variable boundary condition: leak detection in pipe networks using coded transients. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Vela, A. F., Espert, V., Fuertes, V.S. (1995). General overview of unaccounted for water in water distribution systems. *Improving efficiency and reliability in water distribution systems*, E. Cabrera, A. F. Vela, eds., *Kluwer Academic Publishers*, 61-95.

Whitley, D. (1987). Using reproductive evaluation to improve genetic search and heuristic discovery. *Proc., 2nd Int. Conf. on Genetic Algorithms*, *Lawrence Erlbaum*, Cambridge, Mass., 108-115.

Willems, H. and Barbian, O.A. (1998). Operational Experience with Inline Ultrasonic Crack Inspection of German Crude Oil Pipelines. NDT.net, Vol. 3, No. 11.

Wiggert, D. C. (1968). Unsteady flows in lines with distributed leakage. *ASCE Journal of the Hydraulics Division*, Vol. 94, 143-161.

WRc (1994). UK Water Industry: Managing Leakage. Interpreting measured night flows (Report E). *Engineering and Operations Committee*, October, p. 63.

Vrugt, J. A., Gupta, H. V., Bouten, W., Sorooshian, S. (2003a). A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*, Vol. 39, No. 8, 1201.

Vrugt, J. A., Dekker, S. C., Bouten, W. (2003b). Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage. *Water Resources Research*, Vol. 39, No. 9, 1251.

Vrugt, J. A., Gupta, H. V., Bouten, W., Sorooshian, S. (2003c). Shuffled Complex Evolution Metropolis (SCEM-UA) Algorithm. *Manual*, p. 24. Available online at http://www.science.uva.nl/ibed/index_en.html, 28.06.2006.

Vrugt, J. A., Gupta, H. V., Dekker, S. C., Sorooshian, S., Wagener, T., Bouten, W. (2006). Application of stochastic parameter optimization to the Sacramento soil moisture accounting model. *Journal of Hydrology*, 325, 288-307.

Wu, Z. Y., Sage, P. (2006). Water loss detection via genetic algorithm optimization-based model calibration. *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA, CDROM Edition.

Xu, D.-L., Liu, J., Yang, J.-B., Liu, G.-P., Wang, J., Jenkinson, I., Ren, J. (2007). Inference and learning methodology of belief-rule-based expert system for pipeline leak detection. *Expert Systems with Applications*, Volume 32, Issue 1, January, 103-113.

Zecchin, A.C., White, L. B., Lambert, M. F. and Simpson, A. R. (2005). Frequency-domain hypothesis testing approach to leak detection in a single fluid line. *Proceedings of CCWI2005, Water Management for the 21st Century*, D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume 2, 149-154.

Zecchin, A. C., Lambert, M. F., Simpson, A. R., White, L. B. (2006). Laplace-domain comparison of linear models of a reservoir-pipe-valve system with a leak. 8th Annual Water Distribution Systems Analysis Symposium, Cincinnati, Ohio, USA, 27 - 30 August 2006.

Appendix A: Leakage models in literature

The most widely used equation can be written in general form:

$$Q = A \cdot B^{\alpha} \qquad \text{(A.1)}$$

Where Q = leak flow and the other constants varies by the author.

Table A.1 Notations for equation (A.1) by the author

Notations	**References (by first author)**
$A = C_d A_l \sqrt{2g}$ C_d = discharge coefficient; A_l = leak area; g = gravitational acceleration B = pressure head at leak; α = leak exponent = 0.5	Colombo (2002), Wiggert (1968), Pudar (1992), May (1994), Liggett (1994), Simpson (1997), Vitkovsky (1997, 2003, 2003b), Covas (1999), Sathis (1999), Brunone (1999, 2001, 2004), Wang (2002), Lee (2005a), Ramos (2001), Verde (2005), Vela (1995), Mpesha (2001). Rossman (2000).
α = leak exponent = 1.1	Hernandez (1999), Ulanicka (2001), Puust (2006).
α = leak exponent = 0.5 – 1.5	Dunlop (1999), Ulanicki (2000).
α = leak exponent = 1.18	Alonso (2000).
$B^{\alpha} = \sqrt{B} + 2EB^{3/2} + E^2 B^{5/2}$ E = takes into account pipe physical and structural characteristics	Van Zyl (2005).

Another equation that takes account also pipe characteristics can be written in general form:

$$Q = c \cdot P_{ch} \cdot B^{\alpha} \qquad \text{(A.2)}$$

Table A.2 Notations for equation (A.2) by the author

Notations	**References (by first author)**
c = network's constant $P_{ch} = L^{1.18}$ L = pipe length α = leak exponent = 1.18	Germanopoulos (1985, 1995), Tabesh (2003), Vairavamoorthy (1998).
$P_{ch} = 0.5 \cdot L^{1.18}$ (summed over nodes)	Araujo (2003), Dias (2005).

$P_{ch} = L \cdot D^{d} \cdot e^{a\tau}$ D = pipe diameter d = field test's exponent e = pipe material's coefficient τ = the pipe age α = leak exponent = 1.1 – 1.2	Martinez (1999).
$P_{ch} = L$	Reis (1999).
c = 1 $P_{ch} = \sum P_{surf} L_{surf}$ $\sum$ - product for pipes coming and leaving from each node P_{surf} = half of the pipe surface L_{surf} = leak surface per unit pipe surface α = leak exponent = 1.18	Tucciarelli (1999).
c = leakage coefficient of proportionality $P_{ch} = (0.5)^{b} \cdot L \cdot a \cdot p_j$ a = function of pipe diameter, age and its material $B = \left(p_i + p_j\right)$ p_i, p_j = pressures at pipe ends $\alpha = b - 1$ $b = 1.18$	Ainola (2000)

Appendix B: Curriculum Vitae

1. Personal information

First and Last Name:	Raido Puust
Date of birth, City:	18.11.1978, Pärnu
Nationality:	Estonian

2. Contacts

Address:	Suur-Veski 13-2, 80021, Pärnu
Phone:	(+372) 514 7924
E-mail address:	raido.puust@mail.ee

3. Education

Name of Educational Institution	**Graduate time**	**Education (speciality/degree)**
Tallinn University of Technology	2003	Engineering Physics / MSc
Tallinn University of Technology	2001	Engineering Physics / BA
Tartu University, School of Exact Sciences	1997	Mathematics
Pärnu Hansa Gymnasium	1997	Secondary

4. Language skills (basic-, intermediate- or high level)

Language	**Level**
Estonian	Native language
English	High level
German	Basic level

5. Further training

Time of Attendance	**Name of Educational Institution or Name of Organization**
Sept., 2005 – August, 2006	Centre for Water Systems, University of Exeter, UK.
June, 2001	Haestad software's training (WaterCAD, StormCAD, SewerCAD)

November, 1999	Grenoble University (Grande Platforme Tournate de Grenoble ”Coriolis”, France)

6. Career

Period	Name of University, Scientific Institution, Other Organization	Position
2004 – to date	Tallinn University of Technology	Researcher
2003 – 2004	Tallinn University of Technology	Engineer
June 2002 – Sept., 2002	Water Service (Belfast, UK)	Designer
2001 – 2003	Talone Design Office	Designer
2001 – 2002	Tondi Üks	Designer
Oct. 2000 – Dec., 2000	Satakunta Polytechnic (Pori, Finland)	Assistant
June 1999 – Dec., 1999	Estonian Marine Institute	Assistant
1998 – 2001	Tallinn University of Technology	Contract job

7. Research activity

2003 – to date: Member of *American Society of Civil Engineers*

8. Degree information

Modelling of pulsating flow with thermal load. Master’s Degree.
Origin of turbulence in an accelerated flow. Bachelor’s Degree.

9. Honours & Awards

Prize of the Tallinn City Enterprises "Applied Research Project 2006", 2006
Best Applied Research Project in Tallinn University of Technology, 2006
Archimedes Foundation, Kristjan Jaak stipend, 2005
TUT Development Foundation (Tallinna Vesi) PhD student stipend 2005
TUT Development Foundation (Teede REV-2) master’s stipend, 2002

10. Research area

Modelling of hydraulic networks. Leakages in pipe networks.

11. Other projects

Methods for analysis of the efficiency and hydraulic reliability of water supply systems.
Investigation of the influence of different diffuse pollution sources on the nitrogen and phosphorus runoff from watersheds in Estonia.
Fluid-boundary dynamic interaction.

12. Publications

Vassiljev, A., Minguell Font, L., Puust, R. (2007). Use of pressure differentials for calibration of the operational water distribution system. In: Water Management Challenges in Global Change: International conference of computing and control for the water industry (CCWI2007) and suistainable urban water management (SUWM2007), B. Ulanicki, K. Vairavamoorthy, D. Butler, P. L. M. Bounds, F. Ali Memon, eds., Taylor & Francis, UK, 205 - 209.

Koppel, T., Ainola, L., Puust, R., (2007). A mathematical model for the determination of leakage in mains and water distribution networks. Proceedings of the Estonian Academy of Sciences. Engineering, 13(1), 3 - 16.

Puust, R., Kapelan, Z., Savic, D., Koppel, T. (2006). Probabilistic leak detection in pipe networks using the SCEM-UA algorithm. In: Proceedings of WDSA 2006: Water Distribution System Analysis Symposium, Cincinnati, Ohio, USA, CDROM Edition.

Vassiljev, A., Koppel, T., Puust, R. (2005). Calibration of the model of an operational water distribution system. In: Proceeding of the 8th International Conference on Computing and Control for the Water Industry (CCWI): 8th International Conference on Computing and Control for the Water Industry (CCWI), D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume I, 155-159.

Koppel, T., Puust, R., Ainola, L. (2004). An inverse approach for leakage identification in water distribution systems. In: XXIII Nordic Hydrological Conference, A. Järvet, eds., Tallinn, Estonia, Tartu University Press, NHP report 48, Volume I, 342-349.

Ainola, L., Ekholm, A., Koppel, T., Lähdeniemi, M., Puust, R. (2002). Boundary layer with heat load in transitional accelerating pipe flow. In: Hydrodynamics V, Theory and Application: 5th international conference on hydrodynamics, H. H. Hwung, J. F. Lee, K. S. Hwang, eds. Taiwan: ICHD2002 Local Organizing Committee, 523 - 528.

Signature **Date:**

Appendix C: Elulookirjeldus

1. Isikandmed

Ees- ja perekonnanimi:	Raido Puust
Sünniaeg ja –koht:	18.11.1978, Pärnu
Kodakondsus:	Eesti

2. Kontaktandmed

Aadress:	Suur-Veski 13-2, 80021, Pärnu
Telefon:	(+372) 514 7924
E-posti aadress:	raido.puust@mail.ee

3. Hariduskäik

Õppeasutus (nimetus lõpetamise ajal)	**Lõpetamise aeg**	**Haridus (eriala/kraad)**
Tallinna Tehnikaülikool	2003	loodusteaduste magistrikraad
Tallinna Tehnikaülikool	2001	loodusteaduste bakalaureuse kraad
Tartu Ülikool, Täppisteaduste kool	1997	matemaatika valdkond
Pärnu Hansagümnaasium	1997	keskharidus

4. Keelteoskus (alg-, kesk- või kõrgtase)

Keel	**Tase**
Eesti	emakeel
Inglise	kõrgtase
Saksa	algtase

5. Täiendõpe

Õppimise aeg	**Õppeasutuse või muu organisatsiooni nimetus**
Sept., 2005 – August, 2006	Centre for Water Systems, University of Exeter, UK.
Juuni, 2001	Haestad tarkvara koolitus (WaterCAD, StormCAD, SewerCAD)
November, 1999	Grenoble Ülikool (Grande Platforme Tournate de Grenoble ”Coriolis”, Prantsusmaa)

6. Teenistuskäik

Töötamise aeg	Ülikooli, teadusasutuse või muu organisatsiooni nimetus	Ametikoht
2004 - tänaseni	Tallinna Tehnikaülikool	Teadur
2003 – 2004	Tallinna Tehnikaülikool	Insener
Juuni, 2002 – Sept., 2002	Water Service (Belfast, UK)	Projekteerija
2001 – 2003	Projektibüroo Talone AS	Projekteerija
2001 – 2002	AS Tondi Üks	Projekteerija
Okt. 2000 – Dets. 2000	Satakunta Polütehnik (Pori, Soome)	Assistent
1999 (7 kuud)	Eesti Mereinstituut	Assistent
1998 – 2001	Tallinna Tehnikaülikool	Lepinguline töötaja

7. Teadustegevus

2003 – tänaseni: American Society of Civil Engineers liige

8. Kaitstud lõputööd

Soojuskoormusega pulseeriva voolamise modelleerimine. Magistrikraad.
Turbulentsi teke kiirenevas voolus. Bakalaureuse kraad.

9. Teaduspreemiad ja –tunnustused

Tallinna Linna Ettevõtlusameti auhind "Rakenduslik teadustöö 2006", 2006
TTÜ parim rakenduslik teadustöö, 2006
SA Archimedes, Kristjan Jaak stipendium, 2005
SA TTÜ Arengufondi (Tallinna Vesi) doktoriõppe stipendium, 2005
SA TTÜ Arengufondi (Teede REV-2) magistriõppe stipendium, 2002

10. Teadustöö põhisuunad

Hüdrauliliste võrkude modelleerimine. Lekked veevõrgus.

11. Teised uurimisprojektid

Veevarustussüsteemide efektiivsuse ja hüdraulilise töökindluse analüüsi meetodid.
Eesti valgaladelt lämmastiku- ja fosforiühendite hajuallikate väljakannet mõjutavate tegurite uurimine.
Vedeliku dünaamiline koostoime piiretega.

12. Publikatsioonid

Vassiljev, A., Minguell Font, L., Puust, R. (2007). Use of pressure differentials for calibration of the operational water distribution system. In: Water Management Challenges in Global Change: International conference of computing and control for the water industry (CCWI2007) and suistainable urban water management (SUWM2007), B. Ulanicki, K. Vairavamoorthy, D. Butler, P. L. M. Bounds, F. Ali Memon, eds., Taylor & Francis, UK, 205 - 209.

Koppel, T., Ainola, L., Puust, R., (2007). A mathematical model for the determination of leakage in mains and water distribution networks. Proceedings of the Estonian Academy of Sciences. Engineering, 13(1), 3 - 16.

Puust, R., Kapelan, Z., Savic, D., Koppel, T. (2006). Probabilistic leak detection in pipe networks using the SCEM-UA algorithm. In: Proceedings of WDSA 2006: Water Distribution System Analysis Symposium, Cincinnati, Ohio, USA, CDROM Edition.

Vassiljev, A., Koppel, T., Puust, R. (2005). Calibration of the model of an operational water distribution system. In: Proceeding of the 8th International Conference on Computing and Control for the Water Industry (CCWI): 8th International Conference on Computing and Control for the Water Industry (CCWI), D. Savic, G. Walters, R. King, S.-T. Khu, eds., University of Exeter, UK, Volume I, 155-159.

Koppel, T., Puust, R., Ainola, L. (2004). An inverse approach for leakage identification in water distribution systems. In: XXIII Nordic Hydrological Conference, A. Järvet, eds., Tallinn, Estonia, Tartu University Press, NHP report 48, Volume I, 342-349.

Ainola, L., Ekholm, A., Koppel, T., Lähdeniemi, M., Puust, R. (2002). Boundary layer with heat load in transitional accelerating pipe flow. In: Hydrodynamics V, Theory and Application: 5th international conference on hydrodynamics, H. H. Hwung, J. F. Lee, K. S. Hwang, eds. Taiwan: ICHD2002 Local Organizing Committee, 523 - 528.

Allkiri **Kuupäev:**

DISSERTATIONS DEFENDED AT TALLINN UNIVERSITY OF TECHNOLOGY ON *CIVIL ENGINEERING*

1. **Heino Mölder**. Cycle of investigations to improve the efficiency and reliability of activated sludge process in sewage treatment plants. 1992.

2. **Stellian Grabko**. Structure and properties of oil-shale Portland cement concrete. 1993.

3. **Kent Arvidsson**. Analysis of interacting systems of shear walls, coupled shear walls and frames in multi-storey buildings. 1996.

4. **Andrus Aavik**. Methodical basis for the evaluation of pavement structural strength in Estonian Pavement Management System (EPMS). 2003.

5. **Priit Vilba**. Unstiffened welded thin-walled metal girder under uniform loading. 2003.

6. **Irene Lill.** Evaluation of Labour Management Strategies in Construction.

7. **Juhan Idnurm**. Discrete analysis of cable-supported bridges. 2004.

8. **Arvo Iital**. Monitoring of Surface Water Quality in Small Agricultural Watersheds. Methodology and optimization of monitoring network. 2005.

9. **Liis Sipelgas**. Application of satellite data for monitoring the marine environment. 2006.

10. **Ott Koppel**. Infrastruktuuri arvestus vertikaalselt integreeritud raudtee-ettevõtja korral: hinnakujunduse aspekt (Eesti peamise raudtee-ettevõtja näitel). 2006.

11. **Targo Kalamees**. Hygrothermal criteria for design and simulation of buildings. 2006.