Water Resour Manage (2016) 30:2831–2850
DOI 10.1007/s11269-016-1326-6

Integrated Decision Support System for Prognostic and Diagnostic Analyses of Water Distribution System Failures

M. Shafiqul Islam[1] · **Rehan Sadiq**[2] · **Manuel J. Rodriguez**[3] · **Homayoun Najjaran**[2] · **Mina Hoorfar**[2]

Received: 1 October 2015 / Accepted: 7 April 2016 /
Published online: 12 April 2016

Abstract This paper presents an innovative decision support system (DSS) for prognostic and diagnostic analyses of water distribution system (WDS) failures. The framework of the DSS is based on four novel models developed and published by the authors of this paper. The four models include reliability assessment model, leakage potential model, leakage detection model, and water quality failure potential model. Information obtained from these models together with external information such as customer complaints, lab test results (if any), and historical information are integrated using Dempster-Shafer (D-S) theory to evaluate prognostic and diagnostic capabilities of the DSS. The prognostic capabilities of the DSS provide hydraulic and water quality states of a WDS whereas the diagnostic capabilities of the DSS help to identify the failure location with minimal time after the occurrence and will help to reduce false positive and false negative predictions. The framework has 'unique' capacity to bring the modeling information (hydraulic and Quality), consumer complaints, historical failure data, and laboratory test information under a single platform to perform a prognostic and diagnostic investigation of WDS failures (hydraulic and Quality). The proof of concept of the DSS has been demonstrated using data used in published four articles. The outcomes of this research widely addressed the uncertainties associated with WDS which improves the efficiency and effectiveness of diagnosis and prognosis analyses of WDS. It is expected that the developed integrated framework will help municipalities to make informed decisions to increase the safety, reliability and the security of public health.

✉ M. Shafiqul Islam
aitmsi@gmail.com

[1] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

[2] Okanagan School of Engineering, University of British Columbia, Kelowna, BC, Canada

[3] École Supérieure d'aménagement Du Territoire, Université Laval, Québec City, QC, Canada

Keywords Prognostic · Diagnostic · Reliability · Uncertainty · Decision support system

1 Background

Water distribution system (WDS) is the lifeline of the human civilization. A well-maintained WDS is an asset to a community. From the beginning of water supply history to current day, the main purpose of a WDS is to deliver safe water with desirable quantity, quality, and continuity (pressure) to the consumers in a cost effective manner. However, in many cases, WDS fails to meet its objective either due to structural and associated hydraulic failure (SHF) or water quality failure (WQF). The SHF causes water losses and interrupts water supply to the consumers with desirable quantity and continuity with desired pressure whereas the WQF may pose a serious threat to consumers' health. A breach in the structural integrity makes the WDS vulnerable for contaminant intrusion, and may compromise the water quality in the distribution system. Commonly, WDS bears certain structural integrity issues which manifest by leakage or percentage of non-revenue water (NRW) and the resulting degraded water quality.

Whatever might be the reason for the failure, both likelihood and the consequence of a WDS failure can be reduced to an acceptable limit if the prognostic analysis and necessary preventive measures are taken on time. In case of a failure, consequence of the failure can be reduced significantly if necessary interventions are taken in a minimal time after its occurrence. The interventions could be for day-to-day operation and maintenance (O&M) activities such as failure detection, location and repair or for long-term improvement, rehabilitation and replacement. Therefore, utility managers require tools for informed interventions and better decision making.

Numerous studies on a WDS failure investigation and asset prioritization for the long-term improvement have been reported in the literature (Alegre 1999; Allen et al. 2004; Besner et al. 2001; Engelhardt et al. 2000; Farley 2001; Francisque et al. 2009; Li 2007; Poulakis et al. 2003; Sadiq et al. 2010a, b; Storey et al. 2010). Although the related literature acknowledged the high levels of uncertainties, most of them did not address the uncertainties in the analyses, and even when incorporated, uncertainties were poorly addressed. In many cases, the asset prioritization models prioritize wrong assets which do not require intervention at a given point in time. In a similar fashion, during the failure diagnosis, these models either produce many cost-incurring false positive alarms of failure or fail to detect the actual failure events by generating false negative alarms. In this research, an efficient decision support system has been developed which addresses the uncertainties in an integrated manner that can provide better confidence in the prognostic and diagnostic investigations. The developed framework will guide long-term improvement in WDS management and will help in day-to-day-operation and maintenance (O&M) such as failure detection, location and repair/replacement.

To support prognostic and diagnostic investigation of water distribution system, a group authors from the University of British Columbia recently published a series of four independent research articles (Islam et al. 2011, 2012, 2013, 2014) for estimating various aspects of water distribution system failure prognostic and diagnostic analyses. In those articles, four models have been presented- (i) reliability assessment model (ii) leakage potential model (iii) leakage detection model, and (iv) WQF potential model. Although these research articles were independent study, they were integral parts of the integrated forensic analysis of WDS failure. In this article, a proof-of-concept of the integrated prognostic and diagnostic decision support system has been demonstrated and described how different information obtained from the developed models and other independent sources interact coherently to provide final decisions about hydraulic & water quality state and detect failure events (if any) in the water distribution system. Following sections provide basic descriptions of the published models.

1.1 Reliability Assessment Model

The current practice of reliability assessment is based on the simulation results of water distribution system which is modeled using uncertain parameters such as pipe roughness, nodal demands, pipe diameters, reaction kinetics etc. As the model-independent parameters are uncertain, the estimated reliability will subject to uncertainty. This model provides an estimate of the level of uncertainties in traditional reliability by expressing reliability as R (U, μ), where U is the utility gained from a WDS and μ is the *belief* in the calculated utility. The utility component of the proposed reliability of WDS has been defined as a function of hydraulic utility and water quality utility. The hydraulic utility has been defined as function pressure utility and demand utility. Water quality utility has been defined in terms of residual chlorine concentration. Details procedure for estimating utility and belief have been discussed in Islam et al. (2014). For the integrated DSS, the complement of hydraulic utility and water quality utility termed as *hydraulic disutility* and *water quality disutility* have been used as input parameters.

1.2 Leakage Potential Model

The leakage potential model evaluates potential for leakage of any component of WDS or a system as a whole. After extensive literature review (e.g., Fares and Zayed 2009; Farley 2001; Farley and Trow 2003; Lambert 2001; Lambert et al. 1999; May 1994) and state-of-the-practice information, a list of 22 basic factors has been identified that directly (e.g., pressure) and/or indirectly (e.g., traffic movement) influence the leakage. Based on those basic 22 leakage influencing factors and their interrelationship, a leakage potential evaluation model has been developed (Islam et al. 2012). The model shows a unique way to evaluate the impacts of different influencing factors in leakage. The model brings the pressure dependent background leakage under a risk-based methodology. In the current study, *leakage potential* has been used as an input parameter for the integrated DSS.

1.3 Leakage Detection and Location Model

The leakage detection model is based on a parameter termed as *indices of leakage propensity*[1] (ILPs) which are estimated for all pipes and nodes in the system. An ILP provides relative influence of pressure drop by an occurrence of leakage in the system. To estimate this parameter, possible minimum, most likely, and maximum pressure at each node need to be estimated. The parameter ILPs have been used as a surrogate to indicate leakage in the distribution system. The presence of leakage has been identified by comparing ILPs of different nodes and pipes, and topology of the pipe network. Details of the methodology and proof of concepts have been discussed in Islam et al. (2011).

[1] The term index of leakage propensity (ILP) is the ratio of deviation of monitored flow from the most likely value to deviation of extreme value from most likely value

1.4 WQF Potential Model

To develop this model water quality parameters have been divided into two groups - the *symptoms* of WQF and the *causes* of WQF. A list 9 symptoms (i.e., taste, odor, etc.) and 17 *causes* (free residual chlorine, ammonia, etc.) have been identified to develop a causal relationship of those parameters. The *causes* of failure have been considered as alternatives of failure modes and *the symptoms* of the failure as criteria. The TOPSIS (a technique for order preference by similarity to ideal solution) has been used to estimate relative influences of *causes of WQF* on *symptoms of WQF.* Finally, an ordered weighted averaging (OWA) operator has been used to aggregate the impacts of different parameters which have been termed as WQF potential. The parameter WQF potential has been used as an input for the integrated DSS. The extended version of this model has been reported in Islam et al. (2013).

2 Methodology

Figure 1 shows the conceptual integration of the proposed framework. Considering *modes of failure* and *modes of investigation*, the integration process has divided into four parts- hydraulic prognostic investigation (H, P), hydraulic diagnostic investigation (H, D), water quality prognostic investigation (Q, P), and water quality diagnostic investigation (Q, D).

As a part of the prognostic investigation, integrated hydraulic state of the WDS has been evaluated based on the outputs of leakage potential model (Islam et al. 2012), calculated hydraulic utility (disutility) model (Islam et al. 2014) and historical pipe burst data. Integrated water quality state of the WDS has been evaluated based on the outputs of WQF potential model (Islam et al. 2013), calculated water quality utility

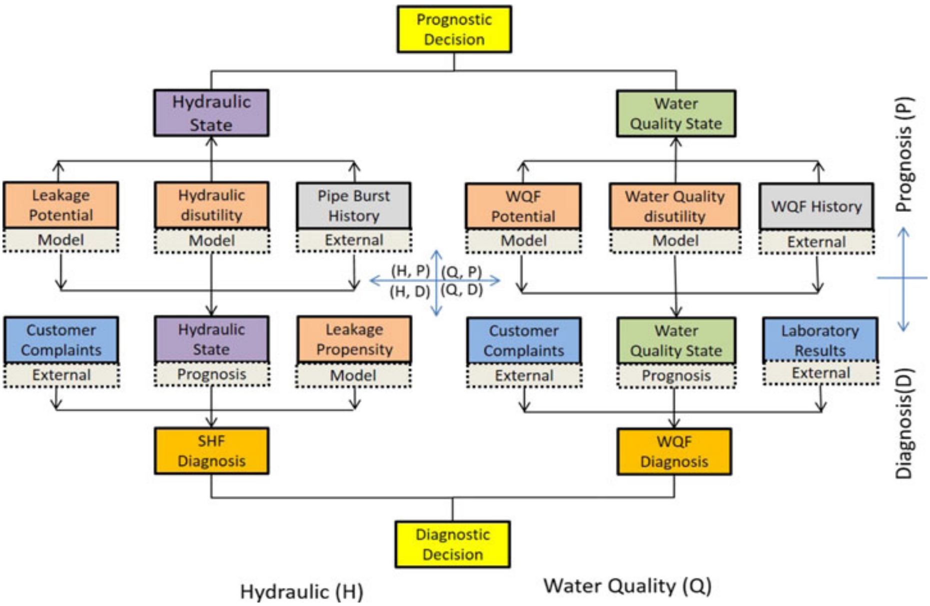

Fig. 1 Conceptual integration of proposed framework

(disutility) model (Islam et al. 2014), and historical WQF data. Customer complaints, hydraulic state and leakage propensity have been used to diagnosis SHF, and customer complaints, water quality state and laboratory test results have been used to diagnosis WQF in WDS. In addition to the information obtained from four models, historical pipe bursts and water quality complaints data can be collected from utilities' regular data repository. Laboratory test results data can be obtained, if any test results available, from utilities' environmental laboratories to integrate into the decision-making process.

Information coming from the different models have been used directly in the integration process. However, historical pipe burst data, historical WQF data, hydraulic and water quality consumer complaints and water quality laboratory test results have been harmonized. Due to the variation of the importance of data in space and data type, the two adaptation factors called, E*quivalent on-Spot Complaints* (*ESC*) and *Equivalent Illness Complaints* (*EIC*) have been introduced. While ESC has been used to harmonize the spatial issues, EIC has been used to harmonize different types of water quality complaints and laboratory test results. Finally, an integrated decision support system has been developed based on the model results and harmonized data under the framework of Dempster–Shafer (D-S) theory of evidence (Shafer 1976).

2.1 Equivalent on-Spot Complaint (ESC)

The complaints received by water companies are spatially distributed. Bicik et al. (2011) mentioned that the consumer complaints are strong indicators of water distribution failure. However, all of them are not trustworthy with the same confidence level. A complaint received from the spot of the failure shall have a higher level of confidence than a complaint received from other than the failure location. Therefore, different complaints such as pipe burst complaints and water quality complaints have been converted to an equivalent on spot complaint (ESC). An ESC has been defined as a complaint which has a confidence level equivalent to a complaint reported from the spot of the failure. Equation 1 has been used to estimate an ESC of a complaint:

$$ESC = e^{-ax} \tag{1}$$

where, a is a constant based on the expert opinion, x is the Euclidian distance from the complaining consumer to the candidate node/pipe in km.

The distance x is constant for a particular complaint and a decision maker has no control over it. However, a decision maker can model his attitude to the confidence level by controlling the parameter, a.

2.2 Equivalent Illness Complaints (EIC)

A water company may receive consumer complaints about various kinds of water quality issues. However, the severities of the different type complaints are not the same. A complaint about bad taste does not have the same level severity for a complaint about reported waterborne disease. To address the variation of different types of water quality complaint, a common scale, termed as Equivalent Illness Complaint (EIC) has been introduced. An EIC has been defined as a water quality complaint which has a severity level equivalent to a complaint that causes illness for the consumers. Equation 2 has been used as measured EIC of a water quality complaint.

$$EIC = \sum_{i=1}^{N} C_i \times W_i, \quad where, \quad W_i \in (0, \ 1) \tag{2}$$

where i is the different types of complaints, C_i is the number of different types of complaints, and W_i is the weight of the each complaint with respect to a complaint of illness to the consumer. Although a decision maker has no control on the number of different types of complaints, a decision maker can model his attitude by controlling W_i based on their expert judgments.

2.3 Dempster–Shafer Theory

Dempster-Shafer (D-S) theory is a mathematical evidence theory and considered as a generalization of the Bayesian approach (Sadiq et al. 2006). The groundbreaking work was introduced by Dempster (1967) and extended by Shafer (1976). Three important concepts - *basic probability assignment* (*BPA*), *belief function*, and *plausibility* are the souls of D-S theory and operates under on a "frame of discernment", Θ. The concept of D-S theory has been described in the various literature. Readers are suggested to consult with different D-S theory-based literature (e.g., Bicik et al. 2011; Islam 2012; Li 2007; Sadiq and Rodriguez 2005; Sadiq et al. 2006) for the basic discussion and detail interpretation on the D-S theory.

2.3.1 Estimation of BPAs

D-S theory has been implemented to evaluate hydraulic & the water quality state of WDS (prognosis) and diagnostic investigation of different kind hydraulic & water quality failures. In all cases, the frame of discernment consists of two propositions (binary frame of discernment) of different kinds of failures, and described by a universal set, Θ = {L, NL} where *L* represents *likelihood* and *NL* represents *not likelihood* of different kind of incidences. The power set of incidences consists of the following subsets: (, {L}, {NL}, {L, NL}). In case of a hydraulic failure, {L} represents the *likelihood* of hydraulic failure; {NL} represents the *not likelihood* of hydraulic failure and {L, NL} represents complete ignorance about hydraulic failure. In a similar way, the water quality failure frame of discernment has been implemented where {L} represents the *likelihood* of water quality failure; {NL} represents *not likelihood* of water quality failure and {L, NL} represents complete ignorance about water quality failure. The proposed integration framework (Fig. 1) has four different phases. In each phase, three different sources of evidence have been used. Table 1 shows different sources of evidence used in different phases of integration.

Some of the evidences are available along the pipe lengths, some are at nodes, and some can be available both along pipes and at the nodes. For example, the leakage potential and the historical pipe bursts evidence are available along the pipes whereas the hydraulic utilities are evaluated at different nodes. However, the prognostic and diagnostic investigation can be carried out using either nodal or pipe evidence. Therefore, either the nodal evidences are required to convert to the pipe evidences

Table 1 Different phases of integration and sources of evidence

Phases	Phase names	Evidences	Location
1	Hydraulic prognosis (H, P)	• Leakage potential model • Hydraulic utility model • Historical pipe burst data	• Pipe • Node • Pipe
2	Hydraulic diagnosis (H, D)	• Consumer complaints • Hydraulic state • Leakage propensity	• Pipe/Node • Pipe/Node • Pipe/Node
3	Quality prognosis (Q, P)	• WQF potential model • Water quality utility model • Historical WQF data	• Node • Node • Pipe
4	Quality diagnosis (Q, D)	• Consumer complaints • Water quality state • Laboratory results	• Pipe/Node • Pipe/Node • Pipe/Node

or the pipe evidences are required to convert to the nodal evidences. In this case, the nodal evidences have been converted to pipe evidences by averaging nodal evidences of two connecting nodes and the pipe evidences have been converted using zone of influences estimated by the half of the connecting pipe length to a node.

In both cases, evidences are harmonized based on the appropriate adjustment factors (ESC & EIC). In case of a evidence available along a pipe, Euclidian shortest distance from the complaint location to the pipe location and in case of a evidence available at a node, Euclidian shortest distance from the complaint location to the node have been used in ESC and EIC calculations. From the homogenized evidences, BPAs have been evaluated using two steps procedure adapted from Safranek et al. (1990).

At the first step of BPAs estimation, the historical pipe bursts, consumer complaints, historical WQFs, laboratory results and leakage propensities have been converted to normalized confidence factors using a suitable normalization function. However, leakage potential, WQF potential, hydraulic and water quality utilities are already on a scale of (0, 1). Therefore, further normalizations are not necessary for them. Numerous (e.g., Sigmoid, Gaussian, Logit, and even a Linear) normalization functions can be used to estimate equivalence confidence factors. A sigmoid function has been adapted in this study. Equation 3 shows the sigmoid function used for normalization (Beynon 2005; Gerig et al. 2000):

$$cf_i\left(v_{j,i}\right) = \frac{1}{1 + e^{-k_i\left(v_{j,i}-\theta_i\right)}} \tag{3}$$

where, k_i describes the steepness of the function and θ_i determines the offset of v axis. The solutions of this function are controlled by two controlling variables, k_i and θ_i. According to Safranek et al. (1990), above function satisfy following criteria:

(i) $cf_i(v_{j,i})$ is a monotonic increasing (or decreasing) function.
(ii) $cf_i(v_{j,i}) = 1$; if the measurement v implies certainty in the hypothesis
(iii) $cf_i(v_{j,i}) = 0$; if the measurement v implies certainty in not the hypothesis
(iv) $cf_i(v_{j,i}) = 0.5$; if the measurement v favors neither the hypothesis nor not the hypothesis.

The normalized evidences in different phases have been mapped to BPAs for {L}, {NL} and {L, NL} using the following equations (Eqs. 4, 5, and 6):

$$m_{j,i}(\{L\}) = \begin{bmatrix} 0 & , \quad cf_i(v_{j,i}) \le A1 \\ \frac{B_1}{1-A_1-A_2} cf_i(v_{j,i}) - \frac{A_1 B_1}{1-A_1-A_2} & , \quad 1-A_2 > cf_i(v_{j,i}) > A1 \\ B_1 & , \quad cf_i(v_{j,i}) \ge 1-A_2 \end{bmatrix} \tag{4}$$

$$m_{j,i}(\{NL\}) = \begin{bmatrix} B_3 & , \quad cf_i(v_{j,i}) \le A_4 \\ B_3 + \frac{[cf_i(v_{j,i}) - A_4][B_3 - B_2]}{1-A_3-A_4} & , \quad 1-A_3 > cf_i(v_{j,i}) > A_4 \\ B_2 & , \quad cf_i(v_{j,i}) \ge 1-A_3 \end{bmatrix} \tag{5}$$

$$m_{j,i}(\{L, \ NL\}) = 1 - m_{j,i}(\{L\}) - m_{j,i}(\{NL\}) \tag{6}$$

The Eqs. 4, 5, and 6 are controlled by one variable and seven controlling parameters (A_1, A_2, A_3, A_4, B_1, B_2, and B_3; $\subseteq 1$). A decision maker can define these controlling parameters based on their expert judgment. Figure 2 shows the definition of the controlling variables and the general steps for the transformation of input data into BPAs.

After conversion of evidences into BPAs, combination rules have been applied to evaluate the joint evidences of different kinds of failures.

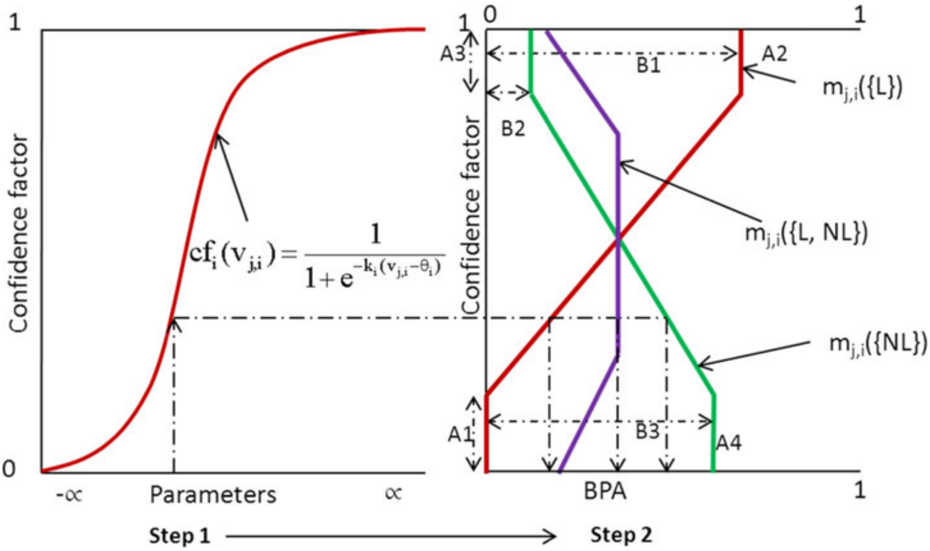

Fig. 2 Transformation of data into BPAs (adapted from Beynon 2005)

2.3.2 Rules of Combination

Dempster's rule of combination has been implemented which combines two BPAs, m_1 and m_2 and estimates the joint BPA, m_{1-2} emphasizing on the agreement of the sources of evidences and ignoring the conflicting evidence (Sentz and Ferson 2002). Equation 7 expresses the mathematical formula of Dempster's combination rule (Klir and Folger 1988):

$$m_{1-2}(A) = \frac{\sum_{B \bigcap C=A} m_1(B)m_2(C)}{1-K} \quad when \quad A \neq \phi; \\ and \quad m_{1-2}(\phi) = 0 \\ where \quad K = \sum_{B \bigcap C=\phi} m_1(B)m_2(C) \tag{7}$$

The term K represents the degree of conflict between two different sources B and C and 1-K is the normalization factors.

3 Implementation of Decision Support System

The integration process has been demonstrated using a simplified version of a part of a WDS. Figure 3 shows the layout of the example WDS. Details of the network information have been published in Islam et al. 2011, 2014.

3.1 Data Preparation

The implementation of the DSS depends on the model dependent secondary data and primary data as additional evidences. The model dependent secondary data include leakage potential, WQF potential, hydraulic utility, water quality utility and indices of leakage propensity in different nodes. The additional primary data include historical pipe bursts & water quality failures, recent consumer complaints, and water quality laboratory test reports. The historical water quality failures and recent water quality complaint parameters include the number of reported illness, color and turbidity, a drop of free residual chlorine and odor and so forth. Recent consumer complaints include pipe bursts and water quality complaints received from consumers. For the all recent complaints, distances from the source of complaints have been considered. In case of a laboratory test report, the sample locations (i.e., distances) with respect to the nearest node/pipe also have been considered as recorded evidences. Since the leakage potential model and WQF potential model have been demonstrated using different WDSs than the network used in this article, typical average values of leakage potential and the WQF potential have been used in all cases. The other parameters such as hydraulic utility & water quality utility and ILPs have been used as same as described earlier in Islam et al. 2011, 2014. Table 2 shows the synthetic number of pipe bursts and leakage potentials for different pipes and historical water quality failure events and WQF potentials in different nodes.

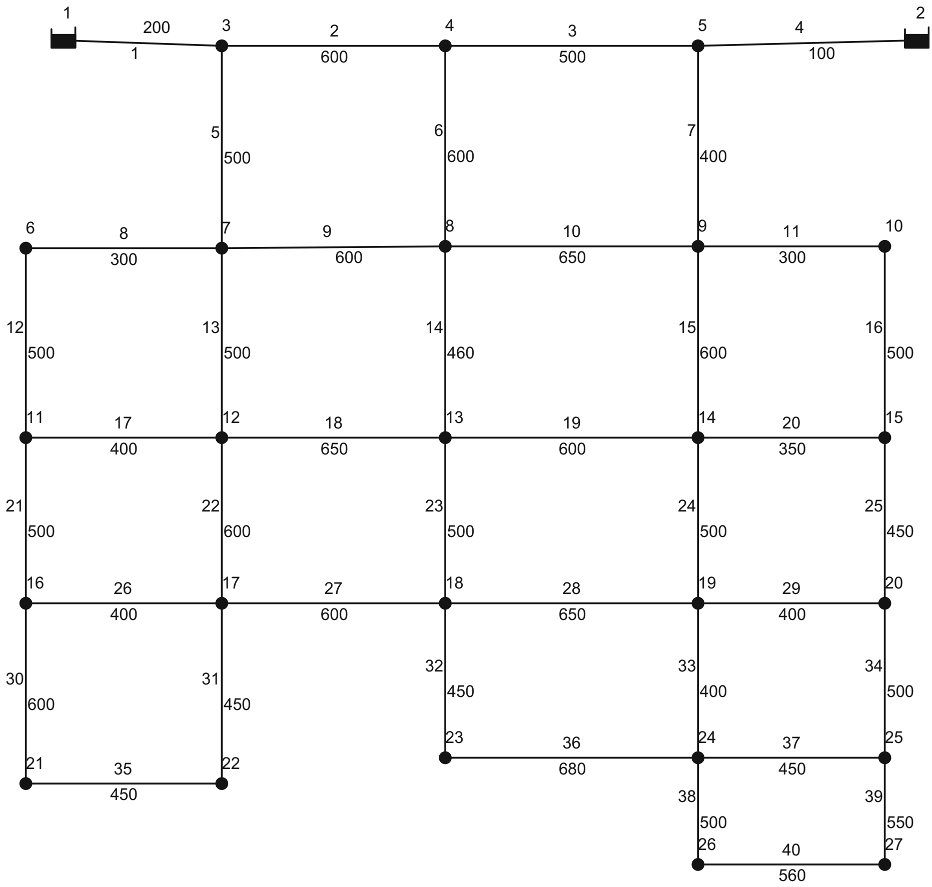

Fig. 3 Layout of the study example WDS (pipe length not to the scale, Islam et al. 2011, 2014)

The network failure condition has been simulated by disconnecting the failed pipe from the remaining network which has been modeled in EPANET by making pipe's initial status to 'close'. In this case, pipe 29 was modeled as a failed pipe. From Islam et al. (2014), it was found that consumers can get full hydraulic utility from the WDS under the normal condition. However, in case of a failure (pipe 29), there is an issue of hydraulic disutility in some nodes. Due to the pipe 29 failure, the most critical duration of the day is 6–8 am and during this period, the WDS has highest demand multiplier of 1.49. The estimated hydraulic disutility at nodes 18, 22, 23, 24 and 25 during the period of 6–8 am are 0.976, 0.192, 1, 1 and 1, respectively. However, other nodes have full utilities during the day. It is noted that in spite of WDS failure, no violation of desired water quality utility has been considered. Under the normal operating condition, it is expected that the WDS will continue to operate without any hydraulic and quality consumer complaints or any positive laboratory test results. However, in a failure condition, it is expected that utility company will receive different kinds of consumer complaints based on the type of failures. In case of water quality events, the water company itself may find some quality issues in their regular laboratory test programs. To mimic the failure condition, consumer complaints are assumed in different locations of WDS over the time.

Table 2 Primary data used in proof of concept of DSS

Pipe/ Node number	No. of pipe burst	Leakage potential	Historical quality failure (nodal)	Nodal WQF potentials
1	0	0.34	0	0.27
2	0	0.41	0	0.38
3	1	0.37	0	0.31
4	0	0.33	0	0.30
5	0	0.45	1	0.44
6	0	0.35	0	0.28
7	0	0.42	0	0.37
8	1	0.33	2	0.26
9	0	0.48	0	0.42
10	0	0.42	0	0.33
11	0	0.34	1	0.32
12	1	0.36	2	0.35
13	0	0.49	5	0.49
14	0	0.49	2	0.39
15	0	0.41	1	0.39
16	0	0.38	0	0.34
17	0	0.32	0	0.26
18	0	0.38	0	0.36
19	1	0.31	1	0.24
20	0	0.31	0	0.30
21	1	0.36	0	0.32
22	0	0.31	0	0.21
23	0	0.38	0	0.32
24	0	0.45	2	0.37
25	3	0.49	0	0.44
26	0	0.48		
27	0	0.47		
28	0	0.48		
29	2	0.48		
30	0	0.49		
31	0	0.31		
32	1	0.40		
33	0	0.37		
34	3	0.48		
35	0	0.43		
36	0	0.36		
37	4	0.40		
38	0	0.50		
39	3	0.50		
40	1	0.40		

Table 3 shows non-zero primary data under hydraulic and water quality failure conditions as additional evidences. It is also noted that the distances from the complaint sources to the candidate nodes have been considered as a single value of 0.3 km. Parametric values for ESC and EIC have been assumed based on expert judgement. The values of these parameters can vary one expert to another expert; however, the impact of the variations should not be significant to dilute the model results. Details have been described in Islam (2012).

3.2 Estimation of BPAs

After gathering relevant data, the BPAs are evaluated following the procedure described earlier. Primary data are normalized using a sigmoid function (Eq. 3). The normalization of different evidences is controlled by two parameters (k_i & θ_i) whereas the BPAs are controlled by seven parameters (A_1, A_2, A_3, A_4, B_1, B_2, and B_3). Table 4 shows the controlling parameters used for the normalization and the BPAs estimation. Using the normalized values of evidences, the BPAs are evaluated using Eqs. 4, 5 and 6. After the estimation of individual BPAs, the Dempster rule combination has been applied.

4 Results and Discussion

4.1 Hydraulic Prognostic Investigation (H, P)

The hydraulic prognosis analysis is based on the performance measure of WDS in past, present and future conditions which are reflected in historical pipe bursts, hydraulic disutility, and leakage potential of different nodes, respectively. Based on these evidences, the state of WDS for a *likely failure* (*L*), *not likely failure* (*NL*) and ignorance (*L* or *NL*) of failure under normal

Table 3 Non-zero primary data under failure condition

Time/Node No.	Reported cumulative consumer complaints												Non-compliance lab. report						
	Pipe bursts		Illness				Colors						FRC			Turbidity			
	20	21	12	13	19	24	12	13	15	17	19	25	13	26	27	11	13	17	21
0–2	0	1	2	2	1	3	1	2	1	2	1	2	1	1	1	1	1	1	1
2–4	0	1	3	3	1	3	1	2	1	2	1	3	1	1	1	1	1	1	1
4–6	2	1	3	4	1	3	1	2	1	2	1	4	1	1	1	1	1	1	1
6–8	3	1	3	5	1	3	1	2	1	2	1	5	1	1	1	1	1	1	1
8–10	4	1	3	6	1	3	1	2	1	2	1	5	1	1	1	1	1	1	1
10–12	4	1	3	6	1	3	1	2	1	2	1	5	1	1	1	1	1	1	1
12–14	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1
14–16	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1
16–18	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1
18–20	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1
20–22	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1
22–24	4	1	3	6	1	3	1	3	1	2	1	5	1	1	1	1	1	1	1

Table 4 Controlling parameters for normalization and BPAs estimation

Controlling Parameters	Normalization		BPAs						
	k	θ	A_1	A_2	A_3	A_4	B_1	B_2	B_3
Leakage potential	–	–	0.1	0.1	0.1	0.1	0.8	0.1	0.5
Hydraulic utility	–	–	0.1	0.1	0.1	0.1	0.8	0.1	0.5
Historical pipe burst	0.5	1	0.1	0.1	0.1	0.1	0.8	0.1	0.5
WQF potential	–	–	0.2	0.1	0.1	0.1	0.8	0.1	0.5
Water quality utility	–	–	0.2	0.1	0.1	0.1	0.8	0.1	0.5
Historical WQF	0.1	2	0.2	0.1	0.1	0.1	0.8	0.1	0.5
Hydraulic consumer complaints	0.3	1.5	0.1	0.1	0.2	0.1	0.9	0.1	0.5
Leakage propensity	0.3	1.5	0.1	0.1	0.2	0.1	0.9	0.1	0.5
Water quality consumer complaints	0.9	2	0.0	0.1	0.1	0.1	1.0	0.1	0.4
Laboratory results	0.9	2	0.0	0.1	0.1	0.1	0.7	0.1	0.7

operating condition as well as under failure condition have been estimated. Table 5 shows the hydraulic prognosis results of top four likely pipes (in order) under normal condition.

The pipes with the most likely, the 2nd, the 3rd, and the 4th most likely chances of failure are no. 34, 39, 37, and 25, respectively. The values of $\{L\}$ and $\{NL\}$ are complementary. If a node has highest $\{L\}$ value, it has lowest $\{NL\}$ value among all the nodes. It is observed that the $\{L\}$ values have proportional relation to the available pressure profile and inverse relation with the available water demand. The time period of 6-8 am has the highest demand and lowest available pressure. During this period, the $\{L\}$ is lowest and $\{NL\}$ highest which indicates lower chances of failure. It is understandable that at low pressure, the chance of a pipe failure is minimal. However, in this model, this relationship is valid until the nodal hydraulic utilities are to a certain limit. In any case, hydraulic utility falls down from the desired level, the prognosis results shows the increase of chances for failures, in spite of pressure drop. This has been reflected from a similar solution under a failure condition. Prognostic analysis has been performed under a failure condition and found a certain level of changes in prognostic analysis which is not enough for failure detection. Details of this analysis has been reported in Islam (2012). If the network continues to run without any failure events, a daily prognosis results should not change significantly. To identify the failure location, a diagnostic investigation is necessary.

4.2 Hydraulic Diagnostic Investigation (H, D)

The hydraulic diagnostic analysis is based on the hydraulic prognosis analysis results/hydraulic state of WDS, consumers' complaints and estimated leakage propensities at different nodes. Based on these evidences, *likely failure* (L), *not likely failure* (NL) and ignorance (L *or* NL) of failure of particular condition have been estimated. Following the methodology described previously, hydraulic diagnosis analysis has been performed under the normal condition and found that both diagnostic and prognostic investigation identifies the same node as the most probable location of the failure. As no significant changes in the BPAs ($\{L\}$, $\{NL\}$, $\{L, NL\}$) from the normal operating condition are not observed, it does not provide any indication of failure in the system. Details of this analysis has been reported in Islam (2012). As depicted in Table 6,

Table 5 Hydraulic prognosis results of top four likely pipes (in order) under normal condition

Time (hour)	First				Second				Third				Fourth			
	Pipe No.	{L}	{NL}	{L, NL}	Pipe No.	{L}	{NL}	{L, NL}	Pipe No.	{L}	{NL}	{L, NL}	Pipe No.	{L}	{NL}	{L, NL}
0–2	34	0.50	0.43	0.06	39	0.46	0.48	0.07	37	0.44	0.49	0.07	25	0.39	0.53	0.08
2–4	34	0.50	0.44	0.06	39	0.45	0.48	0.07	37	0.44	0.49	0.07	25	0.39	0.53	0.08
4–6	34	0.45	0.48	0.07	39	0.39	0.53	0.08	37	0.39	0.54	0.08	25	0.35	0.57	0.08
6–8	34	0.34	0.57	0.08	37	0.29	0.63	0.09	25	0.28	0.63	0.09	39	0.28	0.63	0.09
8–10	34	0.39	0.54	0.08	37	0.33	0.59	0.08	39	0.33	0.59	0.08	25	0.31	0.60	0.09
10–12	34	0.40	0.53	0.08	37	0.34	0.58	0.08	39	0.34	0.58	0.08	25	0.32	0.60	0.09
12–14	34	0.43	0.50	0.07	39	0.38	0.55	0.08	37	0.37	0.55	0.08	25	0.34	0.57	0.08
14–16	34	0.45	0.48	0.07	39	0.40	0.53	0.08	37	0.39	0.53	0.08	25	0.36	0.56	0.08
16–18	34	0.41	0.52	0.07	37	0.35	0.57	0.08	39	0.35	0.57	0.08	25	0.33	0.59	0.08
18–20	34	0.40	0.52	0.07	37	0.34	0.57	0.08	39	0.34	0.57	0.08	25	0.32	0.59	0.08
20–22	34	0.42	0.51	0.07	39	0.36	0.56	0.08	37	0.36	0.56	0.08	25	0.33	0.58	0.08
22–24	34	0.45	0.49	0.07	39	0.39	0.53	0.08	37	0.39	0.54	0.08	25	0.35	0.57	0.08

Table 6 Hydraulic diagnosis results of top four likely nodes (in order) under failure condition

Time (hour)	First				Second				Third				Fourth			
	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}
0–2	20	0.69	0.30	0.01	25	0.54	0.45	0.01	27	0.39	0.59	0.02	17	0.39	0.59	0.02
2–4	20	0.70	0.29	0.01	25	0.55	0.44	0.01	27	0.40	0.58	0.02	17	0.38	0.60	0.02
4–6	20	0.66	0.33	0.01	25	0.49	0.49	0.02	17	0.36	0.62	0.02	27	0.35	0.63	0.02
6–8	20	0.93	0.06	0.00	25	0.85	0.15	0.00	27	0.75	0.24	0.01	26	0.71	0.28	0.01
8–10	20	0.66	0.33	0.01	25	0.41	0.57	0.02	17	0.33	0.65	0.02	15	0.31	0.67	0.02
10–12	20	0.67	0.32	0.01	25	0.42	0.57	0.02	17	0.33	0.65	0.02	15	0.31	0.67	0.02
12–14	20	0.71	0.28	0.01	25	0.46	0.52	0.02	17	0.35	0.63	0.02	15	0.33	0.65	0.02
14–16	20	0.72	0.27	0.01	25	0.48	0.50	0.02	17	0.36	0.62	0.02	15	0.34	0.64	0.02
16–18	20	0.68	0.31	0.01	25	0.43	0.55	0.02	17	0.34	0.64	0.02	15	0.32	0.66	0.02
18–20	20	0.67	0.32	0.01	25	0.42	0.56	0.02	17	0.33	0.64	0.02	15	0.31	0.67	0.02
20–22	20	0.69	0.30	0.01	25	0.44	0.54	0.02	17	0.34	0.64	0.02	15	0.32	0.66	0.02
22–24	20	0.71	0.28	0.01	25	0.47	0.51	0.02	17	0.36	0.62	0.02	22	0.34	0.64	0.02

diagnostic investigation results under the failure condition have changed in a considerable extent indicating a probable failure in the system. Most importantly, the diagnostic investigation results (Table 6) have identified the Node 20 as the most probable leaky node in the system which was the leaky node synthetically created by assigning an extra leakage demand.

4.3 Water Quality Prognostic Investigation (Q, P)

Similar to the hydraulic prognosis and diagnosis analyses, water quality prognosis and diagnosis analyses have been carried out. Following the methodology described earlier, Table 7 shows the water quality prognosis results under normal condition. First couples of hour simulation results in Table 7 need to be ignored as these low level of concentrations are due to simulation delay. Actual results will show up after these periods. From Table 7, it is found that the pipe 15 has highest chances of water quality failure. It is noted that, due to the hydraulic failure at a node, no impact on water quality failures were observed because the hydraulic changes did not impact the water quality failure in the distribution system in this case. However, to check the sensitivity of the water quality prognosis model, a minor change was incorporated (5 % increase) of water quality failure potential under normal condition. From that case, changes are also observed in the chances of water quality failure. However, the changes were not enough to change the ranking of nodes for the chances of failure. Details of this analysis has been reported in Islam (2012).

4.4 Water Quality Diagnostic Investigation (Q, D)

The water quality diagnostic analysis is based on the quality prognosis analysis results/water quality state of WDS, consumers' complaints and routine/special laboratory test results. Based on these evidences, *likely failure* (L), *not likely failure* (NL) and ignorance (L or NL) of failure of the particular condition have been estimated. Water quality diagnosis analysis has been performed under normal condition and found that no significant changes in the BPAs ($\{L\}$, $\{NL\}$, $\{L, NL\}$) from the normal condition are observed, it does not provide an indication of failure in the system. Details of this analysis has been reported in Islam (2012).

However, under the failure condition, from Table 8, it is evident that the diagnostic investigation results have changed in considerable extents indicating a probable failure in the system. Most importantly, it is seen (Table 8) that the diagnostic investigation results identify the different node (Node 13) as the most probable location of water quality failure in the system which is consistent with the received consumer complaints and laboratory test results.

5 Summary and Conclusions

This paper presents an innovative diagnostic and prognostic analysis framework for WDS failures. The prognostic capability of the framework will increase the confidence in predictive state analysis of WDS which will help to reduce the likelihood of failures. The diagnostic capabilities of the framework will reduce false positive and false negative predictions, which will help to identify the failure location in WDS with minimal time after the occurrence which will also minimize the consequences of a failure. This paper developed an innovative way to bring modeling information (hydraulic and Quality), consumer complaints, historical failure

Table 7 Water quality prognosis results of top four likely nodes (in order) under normal condition

Time (hour)	First				Second				Third				Fourth			
	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}
0–2	27	0.81	0.16	0.03	26	0.79	0.18	0.03	25	0.76	0.20	0.04	20	0.67	0.28	0.05
2–4	27	0.32	0.57	0.11	15	0.27	0.62	0.11	11	0.26	0.63	0.11	16	0.25	0.64	0.11
4–6	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
6–8	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
8–10	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
10–12	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
12–14	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
14–16	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
16–18	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
18–20	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
20–22	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11
22–24	15	0.27	0.62	0.11	11	0.26	0.63	0.11	27	0.26	0.63	0.11	16	0.25	0.64	0.11

Table 8 Water quality diagnosis results of top four likely nodes (in order) under failure condition

Time (hour)	First				Second				Third				Fourth			
	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}	Node No.	{L}	{NL}	{L, NL}
0–2	27	0.87	0.13	0.00	26	0.86	0.14	0.00	25	0.83	0.16	0.00	24	0.80	0.19	0.00
2–4	27	0.49	0.50	0.01	13	0.46	0.53	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01
4–6	13	0.48	0.51	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
6–8	13	0.50	0.49	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
8–10	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
10–12	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
12–14	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
14–16	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
16–18	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
18–20	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
20–22	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01
22–24	13	0.52	0.47	0.01	12	0.45	0.54	0.01	15	0.43	0.55	0.01	27	0.42	0.56	0.01

data and laboratory test information under a single platform for prognostic and diagnostic investigation of WDS. The outcomes of this research broadly addressed the uncertainties associated with WDS which improves the efficiency and effectiveness of diagnosis and prognosis analyses.

To avoid lumping of the failure likelihood with consequences, the consequences of failures have been intentionally avoided. The developed framework can be extended by incorporating consequences component of risk which will increase the capacity of the proposed framework. Due to data limitations, the different models of the developed framework have been implemented in different WDSs. To demonstrate the integration of the proposed framework, pipe burst, water quality failure, laboratory test results are synthetically generated. Implementation of the developed models with real world problems is also recommended. To identify the optimum number of sensors for leakage and WQF detection and location, further studies on sensors optimization and placement are recommended.

Acknowledgments This research has been carried out as a part of NSERC-SPG (Strategic Project Grants) project funded by Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would like to express sincere gratitude to the anonymous reviewers for their suggestions to improve the quality of the article.

References

Alegre H (1999) Performance indicators for water supply systems – current trends and on-going projects. In: Drought management planning in water supply systems. Kluwer Publishers, Netherlands, pp 148–178

Allen MJ, Edberg SC, Reasoner DJ (2004) Heterotrophic plate count bacteria–what is their significance in drinking water? Int J Food Microbiol, Elsevier 92(3):265–274

Besner MC, Gauthier V, Barbeau B, Millette R, Chapleau R, Prevost M (2001) Understanding distribution system water quality. J Am Water Works Assoc, AWWA 93(7):101–114

Beynon MJ (2005) A novel technique of object ranking and classification under ignorance: an application to the corporate failure risk problem. Eur J Oper Res 167(2):493–517

Bicik J, Kapelan Z, Makropoulos C, Savić D a (2011) Pipe burst diagnostics using evidence theory. J Hydroinf 13(4):596

Dempster A (1967) Upper and lower probabilities induced by a multivalued mapping. Ann Stat 28:325–339

Engelhardt M, Skipworth P, Savic D, Saul A, Walters G (2000) Rehabilitation strategies for water distribution networks: a literature review with a UK perspective. Urban Water, Elsevier 2(2):153–170

Fares H, Zayed T (2009) Risk assessment for water mains using fuzzy approach. Building a Sustainable Future - Proceedings of the 2009 Construction Research Congress, ASCE:1125–1134

Farley M (2001) Leakage management and control: a best practice training manual. WHO, Geneva

Farley M, Trow S (2003) Losses in water distribution networks: a practitioner's guide to assessment, monitoring and control. (First, ed.), IWA Publishing

Francisque A, Rodriguez MJ, Sadiq R, Miranda LF, Proux F (2009) Prioritizing monitoring locations in a water distribution network: a fuzzy risk approach. J Water Supply: Res Technol AQUA, IWA 58(7):488–509

Gerig G, Welti D, Guttmann C, Colchester A (2000) Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial MR data. Med Image Anal 4(1):31–42

Islam M (2012) Water distribution system failures: an integrated framework for prognostic and diagnostic analyses. Doctoral thesis, School of Engineering, The University of British Columbia, Kelowna

Islam MS, Sadiq R, Rodriguez MJ, Francisque A, Najjaran H, Hoorfar M (2011) Leakage detection and location in water distribution system using a fuzzy-based methodology. Urban Water J 8(6):351–365

Islam MS, Sadiq R, Rodriguez MJ, Francisque A, Najjaran H, Naser B, Hoorfar M (2012) Evaluating leakage potential in water distribution systems: a fuzzy-based methodology. J Water Supply: Res Technol - AQUA 61(4):240–252

Islam MS, Sadiq R, Rodriguez MJ, Najjaran H, Hoorfar M (2013) Evaluating water quality failure potential in water distribution systems: a Fuzzy-TOPSIS-OWA-based methodology. Water Resour Manag 27(7):2195–2216

Islam MS, Sadiq R, Rodriguez MJ, Najjaran H, Hoorfar M (2014) Reliability assessment for water supply systems under uncertainties. Water Resour Plann Manag 140(4):468–479

Klir GJ, Folger T (1988) Fuzzy Sets, Uncertainty, and Information. Prentice-Hall, Englewood Cliffs

Lambert A (2001) What do we know about pressure/leakage relationships in distribution systems. Proc. IWA Specialised Conference: System Approach to Leakage Control and Water Distribution Systems Management, Brno, Czech Republic, pp 89–96

Lambert A, Brown T, Takizawa M, Weimer D (1999) A review of performance indicators for real losses from water supply systems. J Water Supply Res Technol: AQUA, Int Water Serv Assoc 48(6):227–237

Li H (2007) Hierarchical risk assessment of water supply systems. Loughborough University, Loughborough

May J (1994) Pressure dependent leakage. World Water Environ Eng 17:10

Poulakis Z, Valougeorgis D, Papadimitriou C (2003) Leakage detection in water pipe networks using a Bayesian probabilistic framework. Probabilistic Eng Mech, Elsevier 18(4):315–327

Sadiq R, Rodriguez MJ (2005) Interpreting drinking water quality in the distribution system using Dempster-Shafer theory of evidence. Chemosphere 59(2):177–188

Sadiq R, Kleiner Y, Rajani B (2006) Estimating risk of contaminant intrusion in distribution networks using fuzzy rule-based modeling. Proceedings of the NATO Advanced Research Workshop on Computational Models of Risks to Infrastructure, Primosten, pp 318–327

Sadiq R, Kleiner Y, Rajani B (2010a) Modelling the potential for water quality failures in distribution networks: framework (I). J Water Supply Res Technol AQUA 59(4):255–276

Sadiq R, Rodriguez MJ, Tesfamariam S (2010b) Integrating indicators for performance assessment of small water utilities using ordered weighted averaging (OWA) operators. Expert Syst Appl, Elsevier 37(7):4881–4891

Safranek RJ, Gottschlich S, Kak AC (1990) Evidence accumulation using binary frames of discernment for verification vision. Robot Autom, IEEE Trans, IEEE 6(4):405–417

Sentz K, Ferson S (2002) Combination of evidence in Dempster-Shafer theory, Technical Report, Report No. SAND2002-0835, Sandia National Laboratories, 2002

Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton

Storey MV, van der Gaag B, Burns BP (2010) Advances in on-line drinking water quality monitoring and early warning systems. Water Res, Elsevier 6(4):741–747