Failure Experience Improvement System (FEIS) for water supply systems

Aditya Lukas, Ernest Mayr, Max Ruhri, Harald Katzmair and Reinhard Perfler

ABSTRACT

The Failure Experience Improvement System (FEIS) is a software tool that was developed in order to contribute to a minimization of hazardous events and failures within water supply systems and thus to achieve increased water safety. Based on the analysis of failure systems by applying Social Network Analysis (SNA) to the water supply infrastructure, the FEIS enables water utilities to identify causes and effects of failure events and to locate vulnerable points in their infrastructure. Failure events and the relations between them are the basis for the FEIS database. This database draws upon information on failure events which have occurred in practice at water utilities in Austria and on a literature review and survey of guidelines. The FEIS, which is accessed online, is currently used by six Austrian water utilities for development and test purposes. It provides both graphical visualization of the failure network and analytical indicators to evaluate failure events. In this way, it supports the utilities in identifying corrective actions in order to minimize the probability of failure occurrence and to limit the damage to the system once a failure has occurred.

Key words | failure database, failure management, network analysis, risk management, water supply

Aditya Lukas (corresponding author)
Ernest Mayr
Reinhard Perfler
Institute of Sanitary Engineering and Water Pollution Control,
University of Natural Resources and Life Sciences Vienna,
Muthgasse 18,
1190 Wien,
Austria
E-mail: *aditya.lukas@boku.ac.at*

Max Ruhri
Harald Katzmair
FAS Research,
Network Analysis for Science and Business,
Müllnergasse 3/1,
1090 Wien,
Austria

INTRODUCTION

Water supply systems are essential for the functioning of a society and economy and are thus regarded as critical infrastructure (USA Patriot Act 2001; Commission for the European Communities 2005). Concepts of total quality management and risk assessment are therefore increasingly being used to assure safe drinking water. The World Health Organization (WHO) stresses the importance of the Water Safety Plan (WSP) concept, which aims at ensuring the safety of drinking water through the use of comprehensive risk identification, assessment and management (WHO 2006; Bartram *et al*. 2009). Various approaches have been developed to apply vulnerability and risk assessment to water supply systems and to implement the WSP approach: Jayaratne (2008) and Mälzer *et al*. (2010), for example, describe the implementation of the WSP using a semi-quantitative matrix for risk assessment as suggested by the WHO. Wienand *et al*. (2009) used Geographical Information Systems (GIS) to support semi-quantitative risk assessment in the catchment area. Miller *et al*. (2005) provide a review of different catchment risk assessment approaches. Sadiq *et al*. (2004) present a fuzzy-based methodology to analyse risks associated with water quality. Quantitative Microbial Risk Assessment (QMRA) is described by Smeets *et al*. (2010) and Medema & Smeets (2009). The development and application of methods to predict pipe failure rates and to assess the reliability of pipes has been described in a number of publications (for example, Watson *et al*. 2004; Almoussawi & Christian 2005; Berardi *et al*. 2008; Tabesh *et al*. 2009). Mays (2004) gives an overview of methods for vulnerability assessment for drinking water systems used in the USA.

Failures in water supply systems can have dramatic consequences for the supplied population, farming and industrial processes. Generally, the term failure refers to the condition of not meeting an intended objective. In the case of water supply, the general objective is to provide continuously an

doi: 10.2166/hydro.2011.125

adequate quantity of high-quality water in a sustainable manner. The term failure in the context of this work therefore includes all events that can lead to the condition of not meeting this objective. Such failures include events that immediately affect water supply, such as pipe or pump failures and the introduction of contaminants into the network, but also actions with indirect or long-term effects on the water supply system, such as insufficient expenditure for mains rehabilitation. In this case, for example, a failure can be defined as the failure to comply with the expenditure level determined by the rehabilitation strategy. Failures in water supply systems often have similar root causes and show similar failure propagation, which can be generalized and corrected. Therefore, there is a high potential to prevent failures in networked infrastructures by failure analysis and by sharing failure experiences with others. The Failure Experience Improvement System (FEIS), which was funded by the Austrian programme for security research, aims to make use of this potential for the water supply sector. This use is achieved by systematic collection of failure causes and their effects in a database, analysis of the resulting failure network and visualization of failure propagation using Social Network Analysis (SNA). Based on the analysis of the failure network in the water supply infrastructure, causal interrelationships can be identified and vulnerable points in the system can be localized. Thus, the FEIS supports risk management by enabling water utilities to identify and assess potential failure events in their water supply system. Consequently, corrective actions can be introduced in order to minimize both the probability of failure occurrence and the damage to the system once a failure has occurred. A permanent minimization of failures and hazards within the water supply system results in increased water safety and contributes to uninterrupted service and crisis prevention.

Failure reporting systems and SNA

The FEIS draws upon two existing concepts: failure reporting systems and SNA. Failure reporting systems have for the first time been used in the aviation industry and were transferred to other industries, such as petrochemical processing, steel production, military operation and healthcare. These so-called Incident Reporting Systems (IRS), which are voluntary and nonpunitive, have been shown to produce large amounts of essential process information unobtainable by other means (Barach & Small 2000). Nowadays, more and more safety strategies from industry and especially from aviation are being implemented into healthcare, with anaesthesia being the first profession in healthcare to introduce IRS (Thomeczek & Ollenschläger 2006; Rooksby *et al.* 2007). Consequently, the WHO (2005) published draft guidelines for Adverse Event Reporting and Learning Systems, which aim at gaining information by generalizing and analysing similar cases from other institutions. In this way, failure reporting systems support the identification of hazards and risks and provide important information that helps to prevent future failures. Rahman *et al.* (2009), for example, used public failure reports to identify failure sources and to investigate failure patterns in information and communication technology infrastructure.

SNA, on the other hand, is a set of methods that have been developed to investigate the relational aspects of social structures. The pioneers of SNA came from sociology and social psychology and anthropology. Basically, social network data consist of measurements on a variety of relations for one or more sets of actors. In the graph theory literature, actors are frequently referred to as nodes which are joined by lines (Scott 1994; Wasserman 1994). Accordingly, a graph comprises a set of nodes $N = \{n_1, n_2, \ldots, n_g\}$ and a set of lines $L = \{l_1, l_2, \ldots, l_L\}$. In addition, values can be added to lines in order to represent the strength or intensity of a relation between two nodes. A valued graph therefore consists of three sets of information: a set of Nodes N, a set of Lines L and a set of Values $V = \{v_1, v_2, \ldots, v_L\}$ attached to the lines. Figure 1 shows an example of an unvalued directed graph. These data can also be represented in a two-way matrix of the size $g \times g$. The entry x_{ij} of the matrix equals 1 if node n_i is incident to node n_j, and 0 otherwise (see Table 1). For a valued graph, the elements of the matrix represent the values of the lines between the

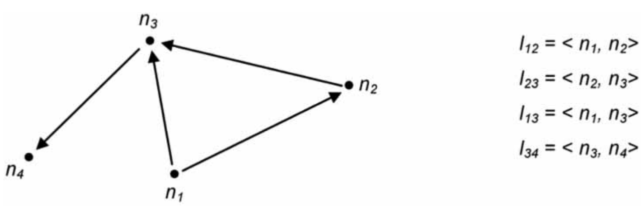

Figure 1 | Example of an unvalued directed graph.

nodes n_i and n_j. A detailed introduction to SNA, including notation, graph theory and matrix operations, is provided by Wasserman (1994).

To the knowledge of the authors, SNA has not previously been applied to investigate failure propagation in water supply systems. A similar approach, the Fault Tree Analysis (FTA), has been used to model risks in drinking water supply (Ezell *et al.* 2000a, b; Risebro *et al.* 2007; Rosén *et al.* 2010). A fault tree represents interactions between different events and shows how the events may lead to system failure. In contrast to SNA, however, FTA is a quantitative approach and requires probabilities for each event causing a failure.

DATABASE DESIGN AND DATA COLLECTION

The FEIS is based on network analysis of failure propagation. The underlying database of the FEIS basically consists of relations between causes of failures and their effects in a water supply system. Since the effect of a failure can in turn be the cause of another failure, effect and failure are both referred to as an event. Failure events in the database are represented as nodes; the relations between the events as directed lines (arcs) pointing from one node to another. Moreover, lines between nodes are assigned the values: (i) time lag; and (ii) impact. The property time lag characterizes the interval of time between a cause and its effect whereas the property impact describes the severity of the effect.

Collection of failure events and relations

To populate the database with cross-linked failure events two methods were applied. First, a literature review and a survey of guidelines were carried out to collect data on possible failure events. Guidelines and standards usually describe an ideal state of technical infrastructure and operating procedures. They were therefore used to identify discrepancies from these ideal conditions as potential failure events. Second, data on failures that have occurred in practice were collected from six water utilities using electronic questionnaires. In addition to failures that have actually occurred, so-called 'near miss' events were also included in the data collection. A 'near miss' event is defined as a serious error that has the potential to cause a failure event but fails to do so because of chance or because it is intercepted (WHO 2005). The project was carried out in cooperation with six water utilities (which supply a total of about 500,000 people) with different supply structures in order to receive a great variety of possible failure events. The collected events were described in short and categorized according to a hierarchical three-part schema: (i) category ('where in the system did the failure event occur'); (ii) system element ('what was affected'); and (iii) failure event ('what happened'). Figure 2 provides an overview of the categories and shows an example of the categorization of a failure event.

After categorization of the collected data, the failure events were cross-linked and visualized using Pajek, a software for performing network analysis (Nooy *et al.* 2005). During the review of this network, relations were added and adjusted and the properties time lag (see Table 2) and impact (see Table 3) were determined for each relation based on the nature of the failure events. For the data collection in the FEIS, the time lag has been defined in rough classes (hours, days, weeks, months). Although it can sometimes be difficult to determine the time lag in cases where the awareness time is quite long (for example, minor pipe breaks, which lead to progressive deterioration of backfill and bedding of pipes), this classification of the time lag should be adequate for the collective database of the FEIS in order to allow some sort of distinction between fast and slow propagating cause–effect chains.

In this manner, a total of about 1,200 event relations were collected and a failure network was established. The resulting network consists of directed graphs; their direction equals the direction of failure propagation, which is in the majority of cases the flow direction of the water through the water supply system. Nodes in the periphery of the network represent events that occur in the beginning of the water supply process (for example, in the groundwater protection

Table 1 | Matrix representing the graph of Figure 1

	n_1	n_2	n_3	n_4
n_1	–	1	1	0
n_2	0	–	1	0
n_3	0	0	–	1
n_4	0	0	0	–

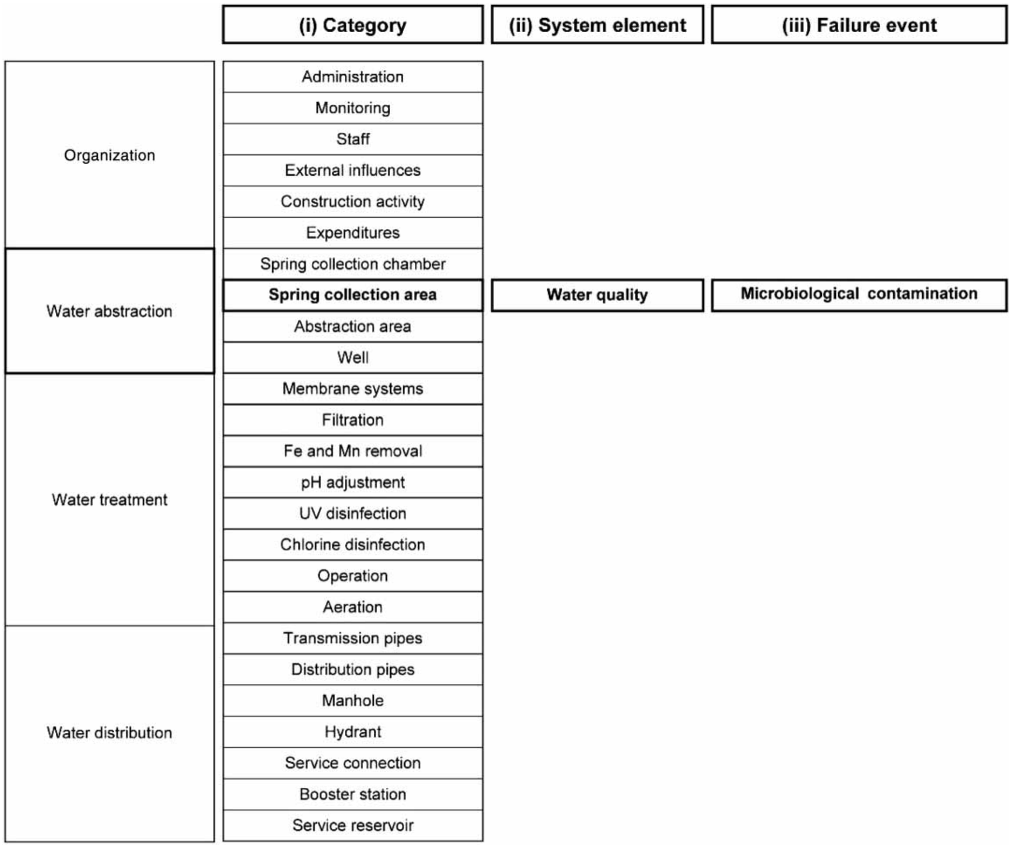

Figure 2 | Overview of the categories used in the FEIS database. Each of the categories belongs to one of the system sections: organization, water abstraction, water treatment or water distribution. The given failure event occurs in the spring collection area ('where in the system did the failure event occur'). The affected system element is the water quality ('what was affected'), which is impaired by microbiological contamination ('what happened').

Table 2 | Time lag

Time lag	Description
Instantaneous	Within hours or shorter
Short term	Within days
Medium term	Within weeks
Long term	Within months or longer

Table 3 | Impact

Impact	Probability that the event n_i causes the consequence n_j
Weak	[0–0.25]
Moderate	[0.25–0.50]
Strong	[0.50–0.75]
Very strong	[0.75–1]

zones). Nodes located in the centre of the network are events that occur in the end of the water supply process (for example, on service connections). Both peripheral and central nodes can be connected to networks outside of the water supply process. Thus, peripheral nodes can have external causes (for example, vandalism or natural hazards) whereas central nodes can have external consequences (for example, illness of water users or damage to roads caused by water leaks).

ANALYSIS OF FAILURE PROPAGATION AND VISUALIZATION WITH THE FEIS

The features and the user interface of the FEIS tool were discussed and developed in workshops together with the cooperating water utilities. The software tool is accessed online by the water utilities and offers both graphical visualization of the failure network (cause–effect chains) and analytical indicators to evaluate failure events.

Analytical indicators are given for each event. At the time of writing, the indicators are calculated based on the failure network described by the matrix X without the values impact and time lag. As shown in the introduction, the entries x_{ij} of X

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1g} \\ x_{21} & x_{22} & \dots & x_{2g} \\ \dots & \dots & \dots & \dots \\ x_{g1} & x_{g2} & \dots & x_{gg} \end{bmatrix} \quad (1)$$

therefore equal either 1 or 0. On the one hand, the indicator 'dependence' (indegree of a node) shows how dependent a specific event is on upstream events. On the other hand, the indicator 'influence' (outdegree of a node) shows how strongly a specific event influences downstream events (Wasserman 1994). Moreover, these indicators are calculated for both the one-step neighbourhood (Direct Influence DI or Direct Dependence DD) and the 10-step neighbourhood (System Influence SI or System Dependence SD). In this way, it is possible to distinguish between events that are interlinked in long chains (and thus important for the whole system) and events that have only short cause–effect chains. The longest path length in the current FEIS database is 18 steps; the average path length is 4.7 steps. The chosen 10-step neighbourhood therefore equals approximately twice the average path length and provides robust results. As the occurrence of longer paths decreases exponentially with the path length, the calculation of the indicators for larger neighbourhoods would result only in marginal differences.

An event with a high DI has an effect on many subsequent events. Thus, a node from which many arcs originate has a higher DI than a node with few originating arcs. The calculation only takes the one-step neighbourhood into account. Therefore, events with a high DI pose great risks for their direct neighbourhood, but not necessarily for the overall system. For calculation of the DI, the number of arcs originating from a node is determined:

$$\mathrm{DI}(n_i) = \sum_{j=1}^{g} x_{ij} \quad (2)$$

where x_{ij} is an element of the matrix X and g is the size of the square matrix X.

The DD is the reverse of the DI. An event with a high DD is caused by many events. Thus, a node to which many arcs point has a higher DD than a node which is a receiver of only a few arcs. Failure events with a high DD easily occur if the direct neighbourhood of the node does not function correctly. The DD of a node is given by:

$$\mathrm{DD}(n_j) = \sum_{i=1}^{g} x_{ij} \quad (3)$$

For example, the DI of the node n_1 in Figure 1 is calculated as the sum over the first row of the 4-by-4 matrix X given in Table 1:

$$X = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

$$\mathrm{DI}(n_1) = x_{11} + x_{12} + x_{13} + x_{14} = 0 + 1 + 1 + 0 = 2 \quad (5)$$

and the DD of the node n_4 is calculated as the sum over the fourth column of the matrix X:

$$\mathrm{DD}(n_4) = x_{14} + x_{24} + x_{34} + x_{44} = 0 + 0 + 1 + 0 = 1 \quad (6)$$

In contrast to the DI, which only takes the one-step neighbourhood into account, the calculation of the SI includes the 10-step neighbourhood of the failure network. Failure events with a high SI provoke events which in turn have many consequences. If an event with a high SI occurs, it is likely that the consequences affect large sections of the water supply system. For example, the contamination of a well has a higher SI than the contamination that occurs at a single service connection. In order to calculate the SI, paths of length between one and 10 steps have to be considered. The entries of the matrix X^k give the number of paths of length k between each pair of nodes. Using the rules for matrix multiplication, the element $x_{nm}^{(2)}$ of X^2 is calculated as:

$$x_{nm}^{(2)} = \sum_{i=1}^{g} x_{ni} \cdot x_{im} \quad (7)$$

where x_{ni} and x_{im} are elements of the matrix X and g is the size of the square matrix X.

According to the Einstein sum convention, an index variable appearing twice implies that one sums over all of its possible values (Ehlotzky 2007). Using this convention, the above equation can be written without the summation operator:

$$x_{nm}^{(2)} = x_{ni} \cdot x_{im} \tag{8}$$

Similarly, the element $x_{nl}^{(3)}$ of X^3 is:

$$x_{nl}^{(3)} = x_{nm}^{(2)} \cdot x_{ml} = x_{ni} \cdot x_{im} \cdot x_{ml} \tag{9}$$

where x_{ni}, x_{im} and x_{ml} are elements of the matrix X.

Therefore, the element $x_{ij}^{(k)}$ of X^k can be calculated as:

$$x_{ij}^{(k)} = \overbrace{x_{il} \cdot x_{lm} \cdot x_{mn} \cdot \ldots x_{qj}}^{k} \tag{10}$$

where $x_{il}, x_{lm}, x_{mn}, \ldots x_{qj}$ are elements of the matrix X and k is the path length.

In order to obtain the SI of a node, the sum over the rows of each matrix X^k has to be calculated:

$$\mathrm{SI}(n_i) = \sum_{k=1}^{10} \sum_{j=1}^{g} x_{ij}^{(k)} \tag{11}$$

where x_{ij} is an element of the matrix X, g is the size of the square matrix X and k is the path length.

The SD is the reverse of the SI. Failure events with a high SD are caused by events which have in turn many causes. These events are likely to occur more often. The SD of a node is given by:

$$\mathrm{SD}(n_j) = \sum_{k=1}^{10} \sum_{i=1}^{g} x_{ij}^{(k)} \tag{12}$$

For example, the SI of the node n_1 in Figure 1 is calculated as the sum over the first row of each matrix X^k. The elements $x_{ij}^{(k)}$ of the matrices X^k can be calculated according to Equation (10). The element $x_{13}^{(2)}$ of the matrix X^2, for instance, is:

$$\begin{aligned} x_{13}^{(2)} &= x_{11} \cdot x_{13} + x_{12} \cdot x_{23} + x_{13} \cdot x_{33} + x_{14} \cdot x_{43} \\ &= 0 \cdot 1 + 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 = 1 \end{aligned} \tag{13}$$

where x_{ij} are the elements of the matrix X given in Equation (4).

In this example, the maximum path length is 3. Therefore, the elements of matrices X^k with $k > 3$ will be 0. The matrices X^2 and X^3 are:

$$X^2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{14}$$

$$X^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{15}$$

and the SI of the node n_1 is given by:

$$\begin{aligned} \mathrm{SI}(n_1) &= (x_{11} + x_{12} + x_{13} + x_{14}) + (x_{11}^{(2)} + x_{12}^{(2)} + x_{13}^{(2)} + x_{14}^{(2)}) \\ &\quad + (x_{11}^{(3)} + x_{12}^{(3)} + x_{13}^{(3)} + x_{14}^{(3)}) \end{aligned} \tag{16}$$

$$\begin{aligned} \mathrm{SI}(n_1) &= (0 + 1 + 1 + 0) + (0 + 0 + 1 + 1) + (0 + 0 + 0 + 1) \\ &= 5 \end{aligned} \tag{17}$$

Similarly, the SD of the node n_4 in Figure 1 is calculated as the sum over the fourth column of each matrix X^k:

$$\begin{aligned} \mathrm{SD}(n_4) &= (x_{14} + x_{24} + x_{34} + x_{44}) + (x_{14}^{(2)} + x_{24}^{(2)} + x_{34}^{(2)} + x_{44}^{(2)}) \\ &\quad + (x_{14}^{(3)} + x_{24}^{(3)} + x_{34}^{(3)} + x_{44}^{(3)}) \end{aligned} \tag{18}$$

$$\begin{aligned} \mathrm{SD}(n_4) &= (0 + 0 + 1 + 0) + (1 + 1 + 0 + 0) + (1 + 0 + 0 + 0) \\ &= 4 \end{aligned} \tag{19}$$

The indicators are normalized and mapped as values between 0 and 1. The risk catalogue allows events to be ranked in order of importance according to these indicators and shows high-consequence or high-dependence events. By these means, one can determine how critical an event is in comparison to others and centres of failure-clusters in the system can be localized. Measures at these nodes can eliminate a high failure potential and therefore minimize the probability of failure occurrence. After the next software update, the indicators will be calculated using the impact value to weight the strength of the relations. Thus, the elements x_{ij} of the matrix X will be weighted using data on the impact as given in Table 3. As a result, the matrix X, which is currently a binary matrix, will contain different entries x_{ij} depending on the impact. This way, the quantification of the indicators will better reflect different probabilities. However, at the time of writing, it is still unclear how the time lag can be included in the calculation in a meaningful way. On the one hand, a failure event that comes into effect after a long delay after being triggered gives the water utility more time to respond to the hazard. On the other hand, it can be more difficult to identify the root cause of such a failure.

The graphical features of the FEIS include (i) visualization relating to a single event and (ii) visualization for a whole category. In the first case, a single event is selected through the hierarchical three-part schema or a keyword search. The cause–effect chains for the selected event can be visualized for the one-, two- or three-step neighbourhood. Relations between failure events that have occurred in practice are represented by arrows with solid lines, whereas arcs connecting potential failure events from literature or guidelines are represented by arrows with dashed lines. The size of the circles representing the nodes relates to the selected indicator. Figures 3 and 4 show the three-step neighbourhood of the node 'Spring collection area – Water quality – Microbiological contamination' for the indicators SI and SD, respectively. The labelled nodes in Figure 3 represent one of the causing failure chains originating from the category 'Abstraction area'. The origin of this chain, the node 'Abstraction area – Wellhead protection area – Unregulated land use', is represented by a circle of greater diameter

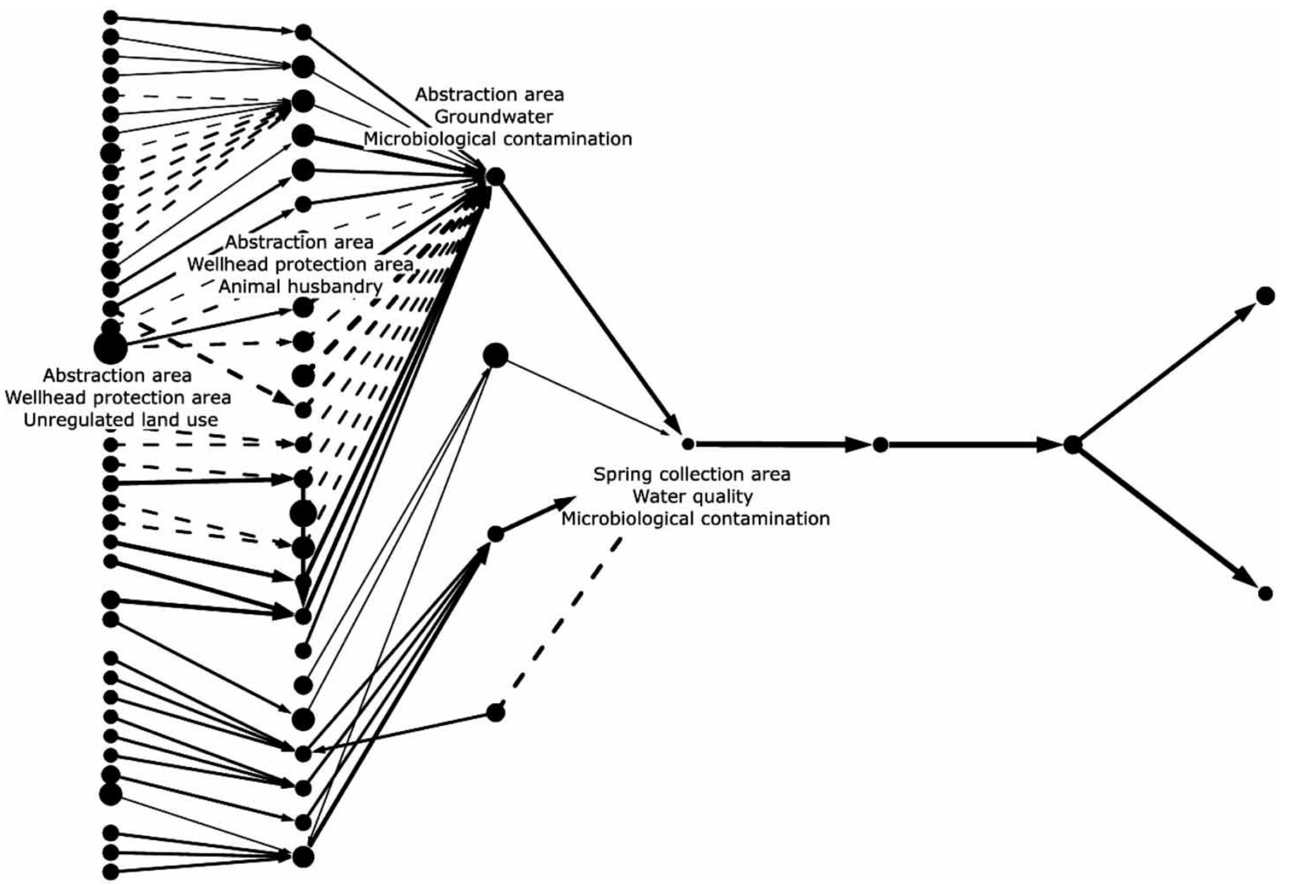

Figure 3 | Visualization of the three-step neighbourhood of the node 'Spring collection area – Water quality – Microbiological contamination'. The size of the circles relates to the indicator SI.

compared to the other nodes, indicating its greater SI. Figure 4, on the other hand, relates to the SD. The nodes 'Service reservoir – Water quality – Microbiological contamination' and 'Distribution pipes – Water quality – Microbiological contamination' have a particularly high SD since they are triggered by numerous failure chains not displayed in the three-step neighbourhood of the node 'Spring collection area – Water quality – Microbiological contamination'.

These visualizations support water utilities in identifying possible causes and effects of failures. Figures 5–7 show an example of the visualization process for some particular failure chains, starting at the failure event 'Service connection – Water quality – Not safe for drinking'. As shown in Figure 5, there are two main upstream chains: 'Microbiological contamination' and 'Chemical contamination'. In order to follow a particular failure chain, nodes of interest can be set as new points of origin, for example the node 'Service reservoir – Water quality – Microbiological contamination'. Figure 6 shows the three-step neighbourhood of the new initial node, which has a variety of upstream nodes, including the node 'Service reservoir – Reservoir compartment – Water condensation'. This node can in turn be set as new point of origin, as shown in Figure 7. Accordingly, causes of water condensation in the reservoir include insufficient thermal insulation caused by erosion or poor design of the reservoir.

The visualization for a whole category provides an overview of the cause–effect chains of the selected category. Figure 8 shows the visualization of the category 'Spring collection area' relating to the indicator SD. Peripheral nodes with a high SD represent upstream connections to other categories (for example, 'Abstraction area – Groundwater – Microbiological contamination'), whereas peripheral nodes with low SD represent source events (for example, 'Spring collection area – Cover of aquifer – Soil erosion'). The

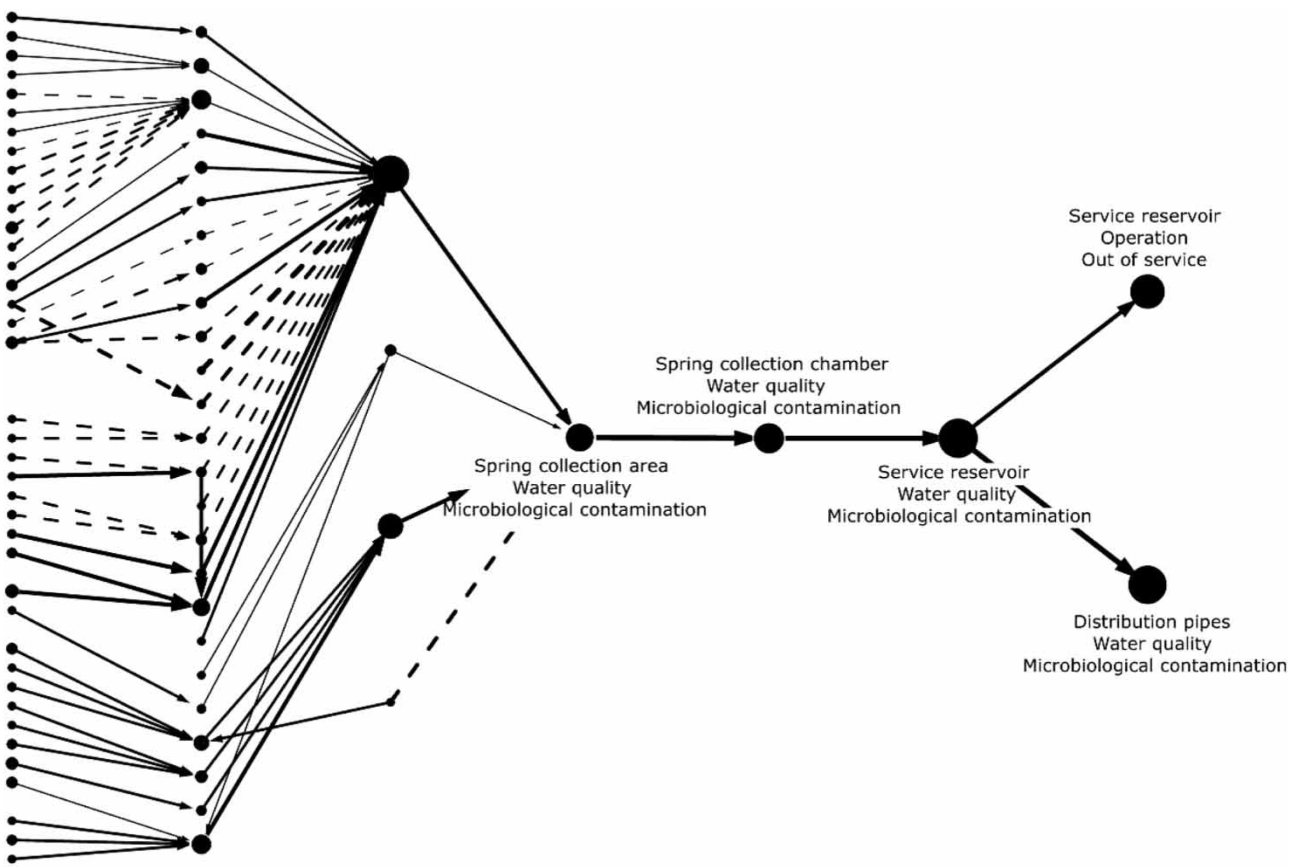

Figure 4 | Visualization of the three-step neighbourhood of the node 'Spring collection area – Water quality – Microbiological contamination'. The size of the circles relates to the indicator SD.

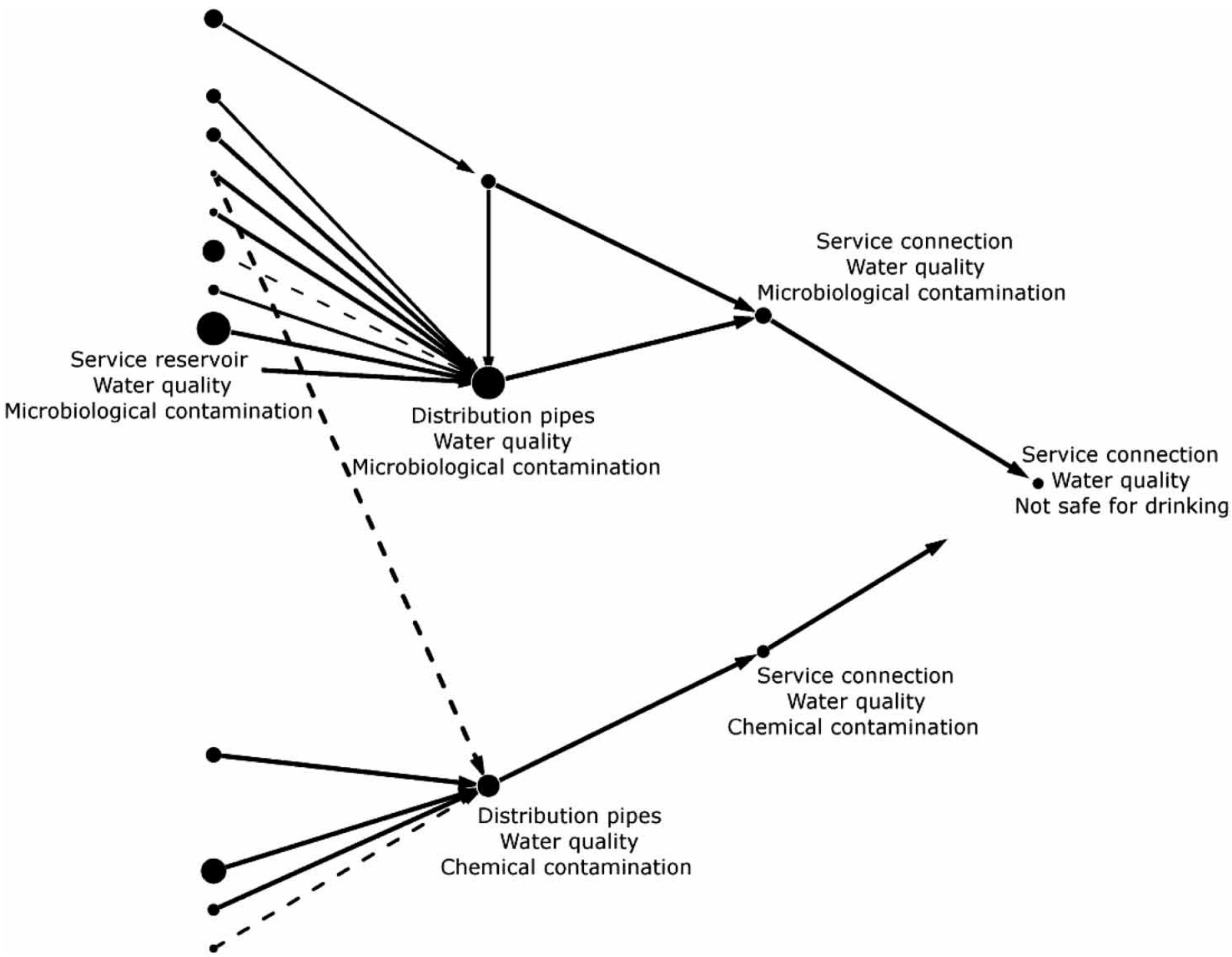

Figure 5 | Visualization of the three-step neighbourhood of the node 'Service connection – Water quality – Not safe for drinking'.

displayed network can be expanded upstream or downstream from each node in order to follow failure chains connected to other categories. These visualizations give a logical and quick overview of the interrelationships in the failure network.

In addition to the visualization in the previous figures, the Flash-based online visualization of the FEIS makes use of colours for intuitive visualization. The colours of the nodes represent either different categories or indicator values. A traffic light colour code system is used to visualize numerical values. Time lag and impact factor are represented by arrow colour and thickness, respectively. Moreover, functional chains are assigned to each node. These functional chains describe the nature of the failure event (microbiological water quality, chemical water quality, physical water quality, water quantity, externalities/environment, structural element). The selection of one functional chain will hide the others, which allows the user to focus on one at a time, for example only water quantity related failure chains. Using the individualization feature, water utilities can hide nodes in order to customize the visualization to match their individual system. More detailed information on the failure events and follow-up actions by the utilities can also be entered in public text fields to enhance exchange of experiences. In this way, the FEIS supports documentation and knowledge sharing within and between utilities. The collective database of the FEIS helps water utilities to consider cause–effect chains that may otherwise have remained unnoticed. The potential of the FEIS therefore lies in the

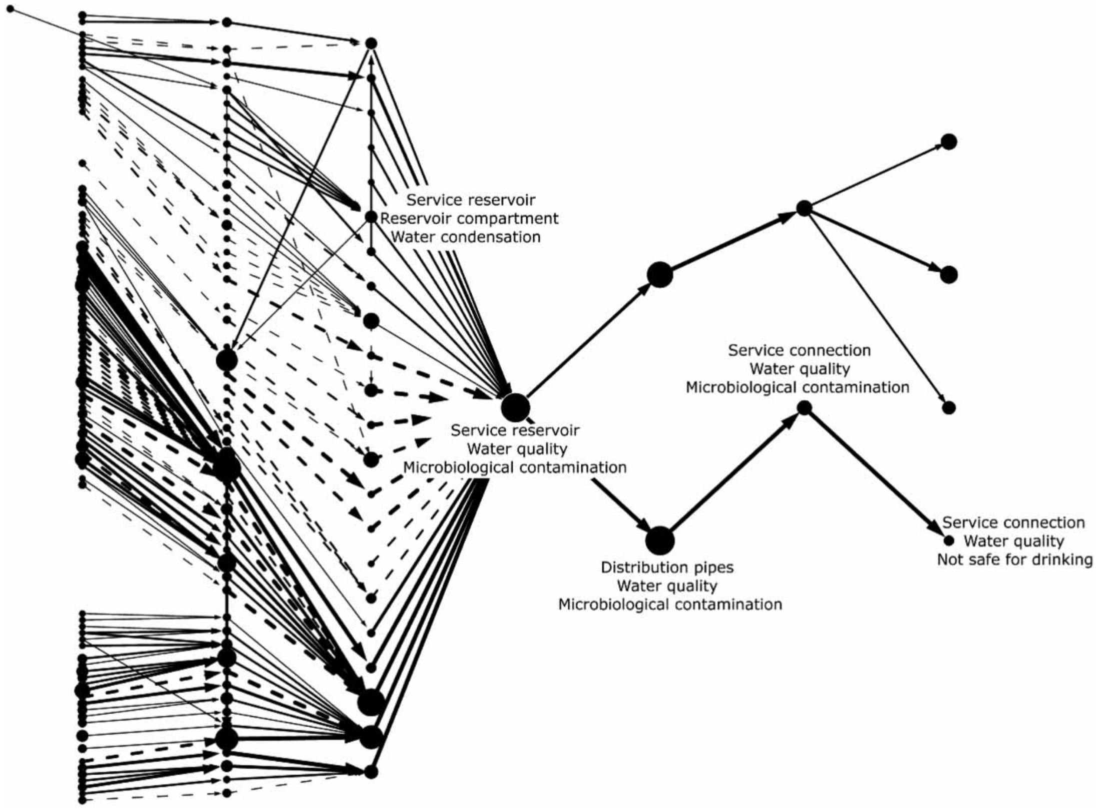

Figure 6 | Visualization of the three-step neighbourhood of the node 'Service reservoir – Water quality – Microbiological contamination'.

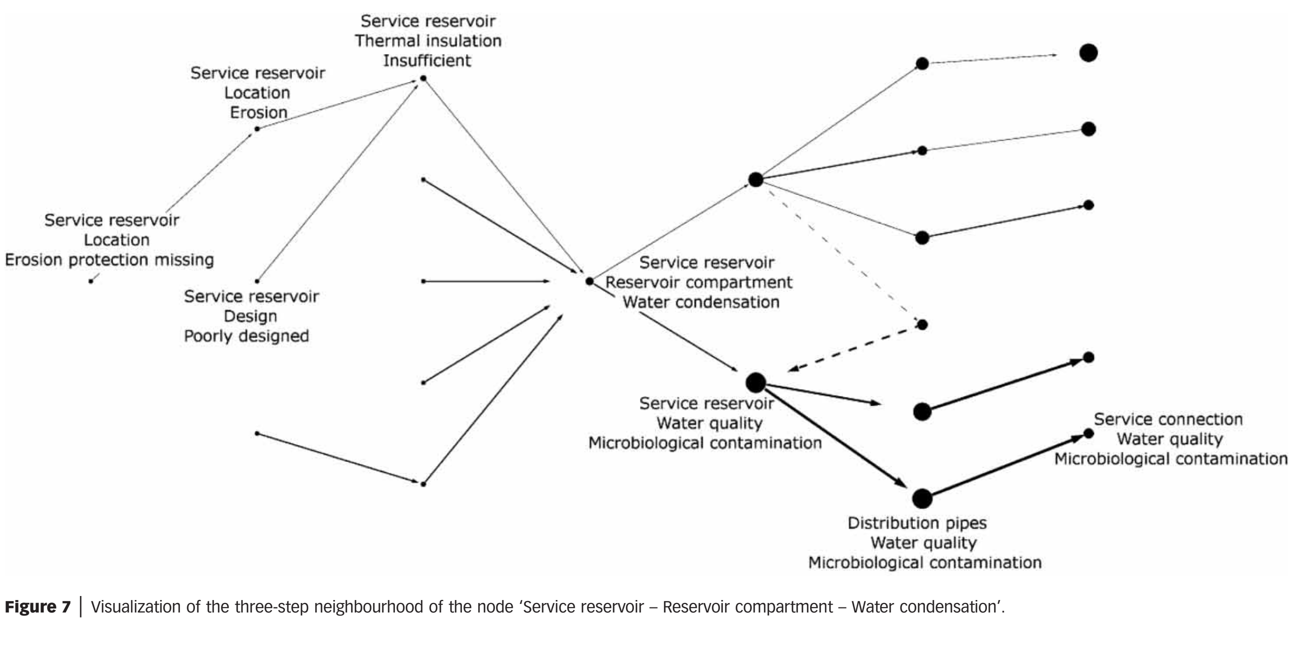

Figure 7 | Visualization of the three-step neighbourhood of the node 'Service reservoir – Reservoir compartment – Water condensation'.

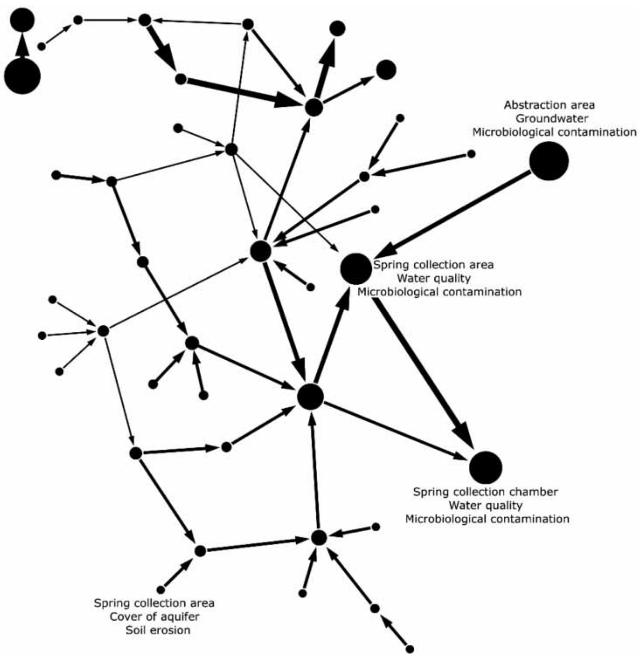

Figure 8 | Visualization of the category 'Spring collection area' relating to the indicator SD.

extendable database, which enhances exchange of experiences with failures to help other utilities to avoid similar failures in their operation. In order to update the FEIS database with new failure events, water utilities can submit failure events to the developers of the FEIS using an electronic questionnaire. After reviewing, categorizing and linking the new failure event to the existing failure network, it is added to the database. At this stage, the analytical indicators are recalculated by the FEIS.

Discussion of further development: individual datasets

As mentioned above, water utilities can hide nodes in order to customize the visualization to match their individual system. Currently, this individualization is limited to the visualization itself and does not influence the calculation of the indicators DI, DD, SI and SD. However, it is intended to further develop the software to make individually adapted FEIS calculations possible. This development requires a database with separate individual datasets besides the collective dataset that comprises all possible nodes and relations, and a dynamic calculation of the indicators. The individual dataset can be adapted to the individual water supply system by adding or cancelling nodes, and by modifying time lag and impact values. The indicators are immediately updated by dynamic calculation if the individual dataset is changed. It is important to note that the step length k for the calculation of SI and SD should be adapted depending on the average path length in the individual dataset. Moreover, if a hydraulic model of the network is available, it can be used to determine time lag and impact values and thus to calibrate the individual FEIS dataset. The simulation of hydraulic and water quality behaviour provides travel times and the concentration of a chemical contaminant throughout the network. Travel times and concentrations from the hydraulic model can be used to refine values for time lag and impact, respectively.

Currently, the severity of different external effects, such as water disruption for a hospital and water disruption for public parks, is not considered in the collective database of the FEIS because it does not reflect the structure of the individual water supply system. In order to distinguish between different severities of external effects, the individual water supply system should be considered. The individual dataset would allow the user to add different nodes for water disruption for a hospital and water disruption for public parks and to define different severities of the effects. Subsequently, the defined severities can be used to weight all upstream relations of these nodes. This way, the calculation of the indicators DI, DD, SI and SD would account for different external effects.

CONCLUSIONS

The present paper describes the design process and the functionality of the FEIS, which enables the user to analyse and visualize failure propagation in water supply systems. It is based on the systematic documentation of failure events and analysis of the relationships between them using SNA. The FEIS supports: (i) identification of potential failures, (ii) decision making, and (iii) knowledge management. Identification of potential failures is facilitated by visualization of the failure networks in the water supply infrastructure and importance-ranking of failure events, which highlights critical nodes in the system. This information can be used to support strategic planning of corrective actions in both normal operation and emergencies, of maintenance and rehabilitation, and of design improvements in the long term. In this way, the FEIS contributes to a minimization of failures, which leads to increased efficiency and cost reduction. Moreover, it provides a platform for systematic documentation and exchange of experiences with failures, helping other utilities to avoid similar failures in their operation. Therefore, the FEIS database is intended as an expandable database; additional failure events should be added and included in the network analysis. By reporting and adding a new failure event into the database, the information on this failure and its causes and consequences in the failure network are shared with other users of the FEIS. In this way, the FEIS supports a safety culture in which preventable failures and their causes can be identified. The more utilities report failures to the FEIS, the more comprehensive the database will be and the more benefits the user will receive. At the moment, additional failure events cannot be reported online but have to be submitted to the developers of the FEIS using an electronic questionnaire. One of the objectives of the further development of the FEIS is to improve this process of continuous documentation of failure events by introducing an online reporting system. Other objectives are to refine the calculation of the analytical indicators by using impact and time lag to weight the strength of the relations between nodes and to make individually adapted FEIS calculations possible.

Despite the willingness of the water utilities to use and to contribute to the FEIS due to the potential benefits of failure documentation and analysis, it has been experienced that they were initially reluctant to use the term failure. Communication and open discussion of failure events are a delicate issue, especially in water supply because of the fear of creating insecurity among consumers. Moreover, the reporter of a failure event might be at risk of blame. Therefore, it is important to ensure that reports are anonymous and to encourage water utilities and their staff to report a wide range of failure events. An open-minded handling of failure events and a strong organizational culture of continuous learning is necessary for successful risk minimization. The FEIS can also be used to train staff by visualizing the impact of failure events and thus to develop awareness of the importance of failure management. Failure analysis and management also have great potential in other infrastructures. Consequently, the possibility of transferring the FEIS approach to other areas of infrastructure (for example, wastewater or gas infrastructure) should be investigated.

ACKNOWLEDGEMENTS

The authors wish to thank Nettah Yoeli-Rimmer for linguistic revision of the manuscript and the anonymous reviewers for critical review of the paper and helpful comments. The development of the FEIS was funded by the Austrian programme for security research of the Austrian Federal Ministry for Transport, Innovation and Technology.

REFERENCES

Almoussawi, R. & Christian, C. 2005 Fundamentals of quantitative risk analysis. *J. Hydroinformatics* **7**, 61–77.

Barach, P. & Small, S. D. 2000 Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *Br. Med. J.* **320** (7237), 759–763.

Bartram, J., Corrales, L., Davison, A., Deere, D., Drury, D., Gordon, B., Howard, G., Rinehold, A. & Stevens, M. 2009 *Water Safety Plan Manual: Step-by-Step Risk Management for Drinking-water Suppliers*. WHO, Geneva.

Berardi, L., Kapelan, Z., Giustolisi, O. & Savic, D. A. 2008 Development of pipe deterioration models for water distribution systems using EPR. *J. Hydroinformatics* **10** (2), 113–126.

Commission for the European Communities 2005 Green paper on a European programme for critical infrastructure protection, COM (2005) 576 final.

Ehlotzky, F. 2007 *Angewandte Mathematik für Physiker (Applied Mathematics for Physicists)*. Springer, Berlin, Heidelberg.

Ezell, B. C., Farr, J. V. & Wiese, I. 2000a Infrastructure risk analysis model. *J. Infrastruct. Syst.* **6** (3), 114–117.

Ezell, B. C., Farr, J. V. & Wiese, I. 2000b Infrastructure risk analysis of municipal water distribution system. *J. Infrastruct. Syst.* **6** (3), 118–122.

Jayaratne, A. 2008 Application of a risk management system to improve drinking water safety. *J. Water Health* **6** (4), 547–557.

Mälzer, H. J., Staben, N., Hein, A. & Merkel, W. 2010 Identification, assessment, and control of hazards in water supply: experiences from water safety plan implementations in Germany. *Water Sci. Tech.* **61** (5), 1307–1315.

Mays, L. W. 2004 Vulnerability assessment, emergency response planning: summary of what's available. In: *Water Supply Systems Security* (L. W. Mays, ed.). McGraw-Hill, New York.
Medema, G. & Smeets, P. 2009 Quantitative risk assessment in the water safety plan: case studies from drinking water practice. *Water Sci. Tech. Water Supply* **9**, 127–132.
Miller, R., Whitehill, B. & Deere, D. 2005 A national approach to risk assessment for drinking water catchments in Australia. *Water Sci. Tech. Water Supply* **5**, 123–134.
Nooy, W., Mrvar, A. & Batagelj, V. 2005 *Exploratory Social Network Analysis with Pajek*. Cambridge University Press, Cambridge.
Rahman, H. A., Beznosov, K. & Martí, J. R. 2009 Identification of sources of failures and their propagation in critical infrastructures from 12 years of public failure reports. *Int. J. Critical Infrastructures* **5** (3), 220–224.
Risebro, H. L., Doria, M. F., Andersson, Y., Medema, G., Osborn, K., Schlosser, O. & Hunter, P. R. 2007 Fault tree analysis of the causes of waterborne outbreaks. *J. Water Health* **5** (1), 1–18.
Rooksby, J., Gerry, R. M. & Smith, A. F. 2007 Incident reporting schemes and the need for a good story. *Int. J. Med. Informat.* **76** (1), 205–211.
Rosén, L., Lindhe, A., Bergstedt, O., Norberg, T. & Pettersson, T. J. R. 2010 Comparing risk-reduction measures to reach water safety targets using an integrated fault tree model. *Water Sci. Tech. Water Supply* **10**, 428–436.
Sadiq, R., Kleiner, Y. & Rajani, B. 2004 Aggregative risk analysis for water quality failure in distribution networks. *JWSRT – AQUA* **53** (4), 241–261.
Scott, J. 1994 *Social Network Analysis*. Sage Publ., London.
Smeets, P. W. M. H., Rietveld, L. C., Van Dijk, J. C. & Medema, G. J. 2010 Practical applications of quantitative microbial risk assessment (QMRA) for water safety plans. *Water Sci. Tech.* **61** (6), 1561–1568.
Tabesh, M., Soltani, J., Farmani, R. & Savic, D. 2009 Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. *J. Hydroinformatics* **11** (1), 1–17.
Thomeczek, C. & Ollenschläger, G. 2006 Fehlermeldesysteme aus jedem Fehler auch ein Nutzen? Bedeutung von Fehler- und Incident-Reporting-Systems in Industrie und Medizin (Error reporting systems – Benefits even from errors? Importance of error and incident reporting systems in industry and medicine). *Rechtsmedizin* **16**, 355–360.
USA Patriot Act, HR 3162 2001 SEC. 1016 'Critical Infrastructures Protection Act of 2001'.
Wasserman, S. 1994 *Social Network Analysis*. Cambridge University Press, Cambridge.
Watson, T. G., Christian, C. D., Mason, A. J., Smith, M. H. & Meyer, R. 2004 Bayesian-based pipe failure model. *J. Hydroinformatics* **6**, 259–264.
WHO 2005 *WHO Draft Guidelines for Adverse Event Reporting and Learning Systems*. WHO, Geneva.
WHO 2006 *Guidelines for Drinking-water Quality: Incorporating First Addendum*. Vol. 1, Recommendations, 3rd edition. WHO, Geneva.
Wienand, I., Nolting, U. & Kistemann, T. 2009 Using Geographical Information Systems (GIS) as an instrument of water resource management: a case study from a GIS-based Water Safety Plan in Germany. *Water Sci. Tech.* **60** (7), 1691–1699.

First received 3 October 2010; accepted in revised form 3 August 2011. Available online 7 November 2011