Pipe burst diagnostics using evidence theory

Josef Bicik, Zoran Kapelan, Christos Makropoulos and Dragan A. Savić

ABSTRACT

This paper presents a decision support methodology aimed at assisting Water Distribution System (WDS) operators in the timely location of pipe bursts. This will enable them to react more systematically and promptly. The information gathered from various data sources to help locate where a pipe burst might have occurred is frequently conflicting and imperfect. The methodology developed in this paper deals effectively with such information sources. The raw data collected in the field is first processed by means of several models, namely the pipe burst prediction model, the hydraulic model and the customer contacts model. The Dempster–Shafer Theory of Evidence is then used to combine the outputs of these models with the aim of increasing the certainty of determining the location of a pipe burst within a WDS. This new methodology has been applied to several semi-real case studies. The results obtained demonstrate that the method shows potential for locating the area of a pipe burst by capturing the varying credibility of the individual models based on their historical performance.

Key words | decision support, diagnostics, evidence theory, pipe burst, water distribution system

Josef Bicik (corresponding author)
Zoran Kapelan
Dragan A. Savić
Centre for Water Systems,
College of Engineering,
Mathematics and Physical Sciences,
University of Exeter,
North Park Road,
Exeter EX4 4QF,
UK
E-mail: *j.bicik@exeter.ac.uk*

Christos Makropoulos
School of Civil Engineering,
National Technical University of Athens,
Heroon Polytechneiou 5,
Athens GR-157 80,
Greece

NOTATION

Θ	frame of discernment
$m(\ )$	basic probability assignment
Bel	Belief function
Pl	Plausibility function
BetP	Pignistic probability function
K	conflicting probability mass

INTRODUCTION

The operation of Water Distribution Systems (WDS) is a complex process, relying on the experience of operators who often have to base their decisions on scarce and incomplete information. Under normal operating conditions the behaviour of WDS is understood relatively well and can be simulated using hydraulic models. However, when pipe bursts occur, the lack of information makes the diagnostics task difficult. Pipe bursts cause water and energy losses (Colombo & Karney 2002), and can also lead to flooding of properties (Cooper *et al.* 2000) and intrusion of contaminants into the WDS (Sadiq *et al.* 2006). Timely detection and location of pipe bursts is therefore of primary interest to water utilities worldwide in order to improve their customer service, minimise leakage, preserve resources and thus minimise impact on the environment.

Pipe burst prediction models have been developed in order to model the deterioration of underground assets (Kleiner & Rajani 2001; Berardi *et al.* 2008; Wang *et al.* 2009). However, such models are more suitable for strategic planning and cannot be utilised on their own to support operational decisions, e.g. to locate a pipe burst in the system in real-time. With recent advances in sensor technologies, wireless pressure and flow sensors have been widely deployed to monitor the state of the WDS in real-time (Mounce *et al.* 2010). Their data have been used in combination with model-based methodologies in attempting to detect and locate leakage or pipe bursts within a WDS. Andersen & Powell (2000)

doi: 10.2166/hydro.2010.201

presented an implicit state estimation technique to locate a burst and demonstrated the methodology on a simple looped network without explicitly taking into account uncertainty and measurement errors. Poulakis *et al.* (2003) developed a Bayesian probabilistic framework for pipe burst detection and showed the capability of the methodology to identify the most likely burst location in a synthetic case study. Wu *et al.* (2010) used genetic algorithms to optimise the pressure-dependent emitter locations and coefficients as possible leakage points and illustrated the methodology in a real-life network. Misiunas *et al.* (2006) used the EPANET (Rossman 2000) hydraulic solver to find a burst location by comparing the fit between the modelled and measured pressures in a WDS. Also a number of transient-based leak/burst detection and location techniques have emerged recently (e.g. Brunone 1999; Kapelan *et al.* 2003*a*; Misiunas *et al.* 2006). A comprehensive review of these methods was recently provided by Colombo *et al.* (2009) and Puust *et al.* (2010). The latter group of authors concluded that transient-based techniques, relying on more expensive pressure transducers, are not yet ready for widescale use by water utilities. Despite the progress achieved there is little evidence that any of the above methods, when used on their own, is ready to be applied in real-life conditions for near real-time decision support of WDS operations.

In this paper, a methodology for combining the outputs of several models (including a Pipe Burst Prediction Model (PBPM), an extended period simulation Hydraulic Model (HM) and a Customer Contacts Model (CCM)) is proposed, to improve the potential for reliable and rapid identification of the possible locations of a pipe burst. This is essential to water companies, reflecting a proactive approach that attempts to detect and resolve failures in the WDS before they start affecting customers. Proactive response is not always possible (e.g. due to the time required to receive and process data from the field or dispatch a leakage team) and in some situations the water company can only react after a problem is first reported by its customers. In the proposed methodology, information provided by individual models is fused together, using the Dempster–Shafer (DS) Theory of Evidence (Shafer 1976). The combined output, which encapsulates the varying credibilities of the individual models, provides the spatial distribution of *Belief* and *Plausibility* of failure of any pipe in the WDS to support the decision-making process by an operator. This evidential reasoning approach further reduces the information load faced by operators and increases confidence in the results that are supported by several models.

DEMPSTER–SHAFER THEORY

The DS theory, also known as Evidence Theory, was first formulated in the late 1970s by Dempster (1967) and later on extended and formalised by Shafer (1976). DS theory can be used for inference in the presence of incomplete and uncertain information, provided by different, independent, sources. A significant advantage of DS theory is its ability to deal with missing information and to estimate the imprecision and conflict between different information sources.

Sentz & Ferson (2002) discussed the foundations of DS theory and provided a review of its applications in various disciplines including classification and recognition, decision-making, engineering and optimisation, fault detection and failure diagnostics, etc. Evidence theory has also been used in water-related applications. Démotier *et al.* (2003) applied DS theory to risk analysis of water treatment processes. Sadiq & Rodriguez (2005) and Sadiq *et al.* (2006) used DS theory to interpret water quality data. Li (2007) used DS theory to aggregate risk levels in a hierarchical risk assessment of components, subsystems and the overall water supply system. Bai *et al.* (2008) used Dempster's combination rule in a hierarchical aggregation of evidence for condition assessment of buried pipes.

The DS theory operates on a "frame of discernment" Θ, which is a finite set of mutually exclusive and exhaustive propositions. Unlike in traditional Bayesian models (Bayes 1763), probability mass can be assigned to subsets of the frame of discernment Θ using a Basic Probability Assignment (BPA), typically denoted $m(A)$, where A is a non-empty subset of Θ. DS theory defines two fundamental functions: *Belief* (*Bel*) and *Plausibility* (*Pl*):

$$Bel : 2^{\Theta} \rightarrow [0,1] \quad and \quad Bel(A) = \sum_{B \subseteq A} m(B) \tag{1}$$

$$Pl : 2^{\Theta} \rightarrow [0,1] \quad and \quad Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{2}$$

where B is a non-empty subset of Θ.

Bel corresponds to the total mass of evidence, which supports a proposition and all of its subsets, whereas *Pl* corresponds to the total mass of evidence, which is not in contradiction to a proposition (Shafer 1976).

In this study, a binary frame of discernment Θ (Safranek *et al.* 1990), is used, comprising two propositions ("*Burst*" and "*NoBurst*") representing the likelihood of occurrence/non-occurrence of a burst. The power set 2^{Θ} is thus formed by the following subsets: (Ø,{*Burst*},{*No Burst*},{*Burst*, *No Burst*}), where the subset {*Burst*, *NoBurst*} represents the whole frame of discernment Θ and any probability mass assigned to this subset corresponds to a lack of knowledge (i.e. ignorance). The chosen definition of the binary frame of discernment implies that the process of identifying the location of a burst pipe is similar to a classification problem where a value of belief is calculated for every pipe in the WDS indicating the likelihood of that pipe being the true (i.e. {*Burst*}) or false (i.e., {*NoBurst*}) burst location.

Dempster's rule of combination (Shafer 1976) is an inherent part of DS theory which allows information from different, independent sources of evidence to be combined. It is defined as follows:

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - K} \quad when\ A \neq \emptyset \tag{3}$$

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{4}$$

$$m_{1,2}(\emptyset) = \emptyset \tag{5}$$

where $m_{1,2}$ is the combined BPA, m_1, m_2 are the BPAs of independent sources of evidence, K represents the level of conflict amongst the evidence, and A, B and C are non-empty subsets of Θ.

Since the introduction of Dempster's rule, various other combination rules have been developed (Sentz & Ferson 2002). In this work, Yager's combination rule (Yager 1987) and the PCR5 combination rule (Smarandache & Dezert 2006) were used, in addition to Dempster's rule, to observe their different behaviour and performance in the process of information fusion. These rules differ in the way they distribute conflicting probability mass K amongst the propositions of Θ. Dempster's rule distributes the conflicting mass equally amongst all propositions of Θ, Yager's rule attributes all conflicting mass to Θ and the PCR5 rule proportionally redistributes partial conflicting masses amongst propositions involved in the partial conflict.

To make decisions based on belief functions, Smets & Kennes (1994) proposed a model of transformation, based on the assumption that "beliefs manifest themselves at two mental levels: the 'credal' level where beliefs are entertained and the 'pignistic' level where beliefs are used to make decisions". Based on the principle of insufficient reason, Smets & Kennes (1994) defined the pignistic probability function *BetP* as follows:

$$BetP(B) = \sum_{A \in 2^{\Theta}} m(A) \frac{|B \cap A|}{|A|}. \tag{6}$$

The pignistic probability function (*BetP*) is a measure that can be used to present the outputs of the information fusion process to the decision-maker and will be later utilised in performance evaluation of the information fusion methodology.

INFORMATION SOURCES

This section provides an example of information sources and models that could be utilised to support the process of locating a pipe burst within a District Metered Area (DMA). It is acknowledged here that the data required to use or develop some of the models might not be readily available to all water utilities. Thanks to the flexibility of DS theory, any kind of information providing an indication of the likelihood of a particular pipe bursting in the WDS can be combined to reduce the lack of knowledge about the location of the failed pipe and increase the confidence in its correct identification. This research utilises three information sources that are considered to be independent: (a) a PBPM output, (b) a CCM output and (c) an HM output. As discussed by Marashi *et al.* (2008) and Bi *et al.* (2008) the assumption of their independence is realistic. This particular set of information sources was chosen because of their general availability to many water utilities worldwide and they do not prevent other information sources from being used (see the conclusions for examples). The first source of information (i.e. based on the pipe burst prediction model output) is treated as a static indicator of pipe burst occurrence whereas the other two remaining sources

can be dynamic and provide new information as it becomes available (e.g. when another customer complaint is received or when the hydraulic model is updated with new real-time measurements obtained from field sensors).

Pipe burst prediction model

A PBPM is used to obtain expected burst frequencies for every pipe in the studied WDS during the current month. The particular choice of PBPM depends on the availability of data and is not important for the methodology shown here as long as the independence of the model outputs used in the information fusion holds (Bai *et al.* 2008; Marashi *et al.* 2008).

Customer contacts model

The current methods of detection and location of pipe bursts aim to notify the control room personnel of any abnormal conditions before a failure starts affecting customers. However, frequently, large pipe bursts are first reported by customers (i.e. when leaked water emerges on the surface). In situations where no explicit pipe burst detection mechanisms are in place, customers reporting the locations of bursts are the only means of (reactive) response to control leakage. Despite being a very strong indicator of a burst location, customer contacts are imperfect and cannot be entirely trusted. A CCM was developed under the assumption that a burst pipe is located in the proximity of the location reported by a customer. The coordinates of the geocoded location of a burst (i.e. easting and northing) reported by a customer were used in this work. Furthermore, the CCM used a weighted distance to reduce the influence of outliers (i.e. misleading customer contacts) in situations when multiple customer contacts were received. The mathematical formulation of the model is as follows:

$$\textit{Criterion Measurement}_{CCM}(i) = \min_{j}(\text{dist}(i, CC_j) \times w_j) \tag{7}$$

$$w_j = \frac{\text{dist}(CC_j,\ C)}{\sum_{k=1}^{N_{CC}} \text{dist}(CC_k,\ C)} \tag{8}$$

where i is the index of a pipe, dist is the shortest Euclidean distance between the burst location reported by a customer and the pipe, CC_j is a customer contact j, w_j is a weight reflecting the significance of a particular customer contact (i.e. the lower the value of w_j the more significant a given customer contact is), N_{CC} is the total number of customer contacts associated with a particular pipe burst and C is the centroid of all customer contacts related to the pipe burst.

Hydraulic model

An extended-period simulation hydraulic model was used to locate a burst in a WDS by simulating its effects (i.e. an increase in flow and drop in pressure) and comparing them with values obtained from pressure and flow sensors deployed in the field. An estimated magnitude of the burst flow is first provided by a detection system able to discover abnormally high inflows into a DMA (Misiunas *et al.* 2006; Mounce & Machell 2006; Romano *et al.* 2009). It was assumed here that the detected pressure and flow anomalies were caused by the burst of a single pipe within a DMA (i.e. no simultaneous pipe bursts were considered due to their very low probability of occurrence). Under this assumption, an extra demand equal to the estimated burst flow was then added to the centre of every pipe to simulate the effects of a burst on that location. The pressure boundary conditions of the hydraulic model were set according to the data obtained from inlet pressure sensors at the time when the burst was first detected. The flow balance of the hydraulic model was established by proportionally scaling the average customer demands at time t to match the measured inflow into the DMA at the same time, excluding the estimated burst flow (i.e. customer demands = DMA inflow − burst flow). The likelihood of any pipe bursting in the system was then indicated by a sum of squared errors between the observed and modelled pressures calculated as follows:

$$\textit{Criterion Measurement}_{HM}(i) = \sum_{s=1}^{N_S}\sum_{t=1}^{T}(P_{1,s}(t) - P_{2,s}(t))^2 \tag{9}$$

where i is an index of the burst pipe in the hydraulic model, s is an index of a node where a pressure sensor is located, N_S is the total number of pressure sensors in the DMA, T is the number of pressure measurements available (i.e. different times), $P_{1,s}(t)$ is the modelled pressure at time t at node s and $P_{2,s}(t)$ is the measured pressure at time t at node s.

Flow measurements inside a DMA were not utilised since these are not typically available in real-life systems (at least not in the UK) due to the higher cost of flowmeters in comparison to pressure sensors.

INFORMATION FUSION

Each of the information sources described above provides a single output (i.e. criterion measurement) for each pipe in the WDS reflecting the likelihood (i.e. a normalised value of the criterion measurement) of occurrence of a burst on that pipe. The individual information sources used are not considered to be fully reliable and each may be associated with a different level of credibility. In order to improve the combined confidence in the location of a pipe burst, the information from all available sources is fused using the DS theory by applying a suitable combination rule.

Before the outputs of individual models can be combined, the criterion measurements need to be transformed into BPAs, each representing the exact belief in the given proposition (i.e. {*Burst*}, {*NoBurst*}) as well as the degree of ignorance (i.e. {*Burst*, *NoBurst*}). For this purpose a two-step procedure was adapted from Beynon (2005). The criterion measurement values were first converted to confidence factors using a suitable normalisation function and then transformed into BPAs as shown in Figure 1.

Beynon (2005) used a sigmoid normalisation function to transform criterion measurements into confidence factors that were mapped to corresponding BPAs. Similar to Safranek *et al.* (1990), Beynon (2005) applied simple symmetric functions defined by two parameters A and B to map confidence factors to BPAs. On the other hand, Sadiq *et al.* (2006) used trapezoids, typical for fuzzy sets, to obtain BPAs directly from criterion measurements. In this work, however, the type of normalisation functions (i.e. linear, sigmoid, one-sided Gaussian and logit function) as well as the shape of the mapping functions (defined by eight parameters, i.e. four points A_1, B_1, A_2 and B_2 as shown in Figure 1) were determined for each of the input models based on its performance during calibration in a number of historical cases. The mapping function describing $m(\{Burst\})$ is a non-decreasing function whereas the function describing $m(\{NoBurst\})$ is a non-increasing function. Once the evidence for every pipe in the network is transformed to BPAs the individual pieces can be combined using a combination rule (Equation (3)). The actual rule used is determined as part of a calibration procedure so that the ensemble of the combination rule, the normalisation and mapping functions gained the maximum benefit according to the criteria outlined in the results and discussion section. A numerical example illustrating the information fusion process can be found in the appendix.

CASE STUDY

The proposed methodology was applied to a case study based on data from a real system in North Yorkshire, UK.

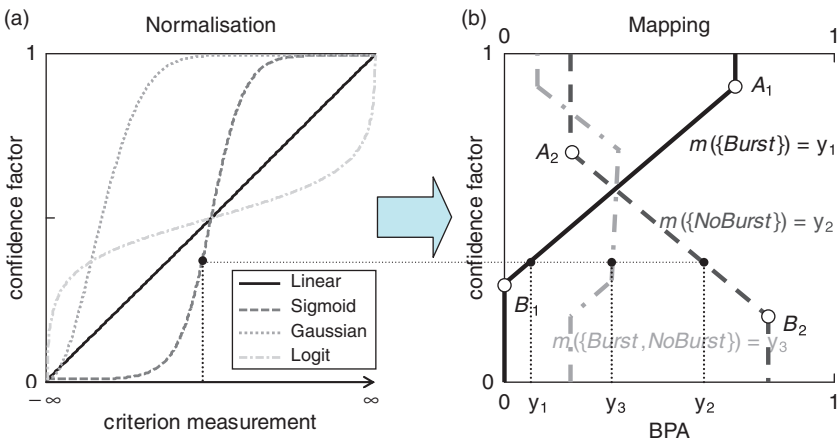

Figure 1 | Transformation of measurement criteria into BPAs based on Beynon (2005).

The studied DMA (see Figure 2) was an urban, highly looped network with two inlets and no exports, supplying water to over 4500 customers.

The available dataset contained information about water main repairs from a work management system, customer contact data and asset data providing required inputs into the PBPM. A regression-based PBPM providing the expected burst frequency of a pipe as a function of its material, diameter, age, soil type, land use and weather conditions was used here. The specific expression and the related coefficients used in this work will not be reported here as they fall outside the scope of this paper.

In order to calibrate the DS model it was necessary to obtain details about a number of historical pipe bursts. During the period from April 2002 to April 2008 54 pipe bursts were recorded in this DMA at locations shown in Figure 2. Customers reported 65% of the pipe bursts either 24 h before the burst was repaired or during the same day that the repair took place. Based on this, it was assumed here that a burst pipe was repaired the same day that an anomaly was detected. The time window over which customer contacts were considered to be related to a particular burst event was established by performing spatial analysis of customer contacts and work management system data of a large number of DMAs. The size of the window was chosen as the best trade-off maximising the number of customer contacts associated with pipe bursts and minimising the distance of those contacts from the location of the burst pipe.

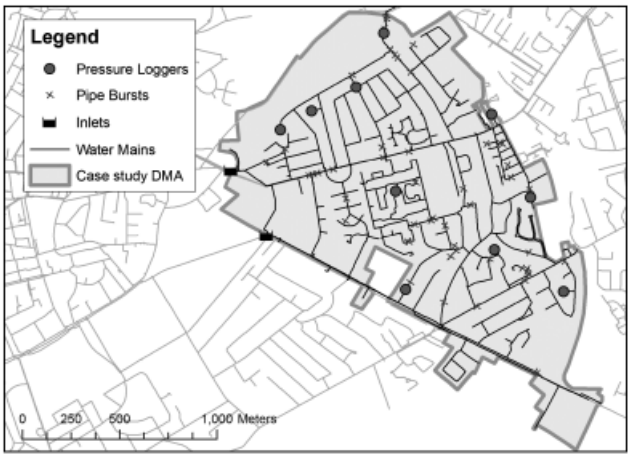

Figure 2 | An overview of the case study area.

The use of the hydraulic model as a source of evidence required a relatively high number of pressure sensors in the network, depending on its size and topology in order to achieve an acceptable performance. Water companies in the UK typically do not monitor pressure at a sufficient number of locations in the WDS. Ten pressure sensors were deployed in the case study area in 2009 at locations indicated in Figure 2. In principle, an optimal sampling methodology (Bush & Uber 1998; Kapelan *et al.* 2003*b*) should be used to determine the exact number and location of pressure sensors. In practice this is not always the case (indeed it was not the case in the examined network). Such an optimal sampling methodology also needs to examine the cost–benefit trade-offs associated with the deployment of multiple sensors. This cost–benefit might rule out the use of a hydraulic model as a source of evidence for certain WDS.

Throughout the period from 2002 until 2008 pressure and flow data were not collected in sufficient quantity, nor was an online pipe burst detection system (Mounce *et al.* 2010), capable of providing estimates of the abnormal burst flows, in place. Therefore the inputs into the hydraulic model (i.e. pressure and flow measurements, and estimated burst flow magnitude) had to be synthetically generated. A medium burst (between 4.5 and 5.5 l/s, i.e. around 15% of the peak demand) was first simulated as a fixed demand added to the centre of a pipe nearest to the location obtained from the work management system. Smaller burst flows, which do not cause sufficient head losses, might even be impossible to locate at all given the typical accuracy of pressure sensors. Pressures in the system obtained at demand nodes closest to the real location of sensors were recorded and used as reference pressures representing a pipe burst situation. Uniformly distributed noise of 2% and 7.5% was added to the reference pressures and nodal demands, respectively, to reflect real-life conditions more closely. These figures are representative of the pressure sensors used and real-life demand conditions in the DMA. Without adding any noise the HM would always find the right location of the burst and would significantly outperform the remaining information sources. It was assumed that the magnitude of the burst flow was known (i.e. provided by a burst detection system) and no noise was added to this input parameter at this stage (it was, however, considered during the sensitivity analysis). The detectable burst flow in this DMA was significantly

higher than regular consumer demand at any network node, hence the noise in pressure measurements played a more important role.

The complete dataset comprising 54 historical pipe bursts was split into a calibration set comprising 41 cases and a validation set comprising 13 cases (approx. ratio 75% calibration/25% validation). The split between calibration and validation data was done in such a way that both datasets had similar properties (e.g. in terms of the number of customer contacts received and the performance of individual models). The calibration procedure aimed to determine the most suitable normalisation and mapping functions as well as the combination rule that would produce the best combined results. The resulting mapping function of the CCM tailored specifically for the case study DMA is shown in Figure 3 as an example. The most suitable normalisation function for the PBPM was the sigmoid function and for the HM and the CCM, it was the logit function. Dempster's rule yielded better results in view of the calibration objectives than Yager's and the PCR5 combination rules. Note that the above findings should be considered case-specific and should not be generalised to other situations. The same methodology can, however, be used in other cases to identify appropriate normalisation functions and combination rules.

As can be seen from Figure 3, the mapping function captures the different behaviour of the analysed model. In the case of the CCM, it can be observed that, in a large number of cases, customers reporting a burst were located in close proximity to the pipe burst. However, a portion of customer contacts was misleading, which explains the shape of the mapping function in Figure 3.

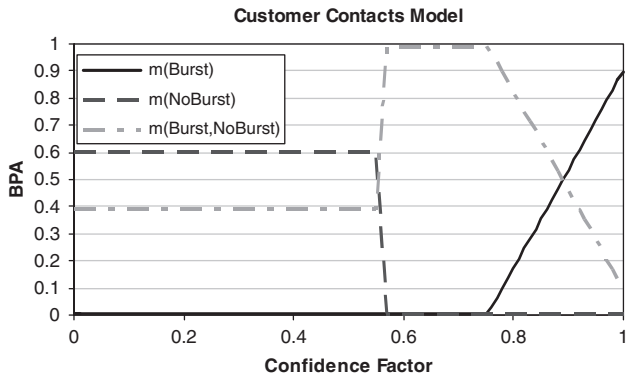

Figure 3 | Optimised mapping function of the CCM.

RESULTS AND DISCUSSION

The main aim of information fusion applied in the context of pipe burst diagnostics is to identify hotspots, comprising a small number of pipes, where the burst is most likely to be located. Figure 4 illustrates the performance of the DS model for a historical pipe burst. In this case, the burst was reported by two customers and therefore all three sources of evidence were available.

The accuracy of the PBPM was limited and a large number of pipes received the same value of confidence factor (see Figure 4(a)). The HM performed poorly in this particular case and identified two possible pipe burst hotspots, with the most likely location being far from the burst pipe (see Figure 4(b)). One of the customer contacts was received from a location in close proximity to the burst pipe whereas the other one was more than 250 m away from the burst location (see Figure 4(c)). Based on the input of the CCM, the DS model attributed higher levels of *BetP*(*Burst*) to the pipes in the second pipe burst hotspot previously identified by the HM, supporting the proposition that this was the true location (i.e. according to a record in the work management system that a burst was repaired there) of the burst (see Figure 4(d)). The pipes close to the second customer contact, which was further away from the true location of the burst, received a lower level of *BetP*(*Burst*). Therefore a field investigation, based on the results of the DS model, would focus on the first customer contact and thus reduce the time for repair, reducing the amount of water lost from the system and the possible follow-on (socio-economic) impact on customers. Figures 4(e) and (f) show the spatial distribution of *Belief* (i.e. probability mass supporting the hypothesis that a burst is located in a particular area) and *Plausibility* (i.e. probability mass not contradicting the hypothesis that the burst is located there), respectively.

Performance comparison

Table 1 shows the performance of the DS model and of the individual models both in the calibration and validation cases. These were further split, depending on the presence of customer contacts (CC). The comparison was based on the ranking (pipe burst candidates were sorted in descending order of their likelihood) of the real burst pipe according to

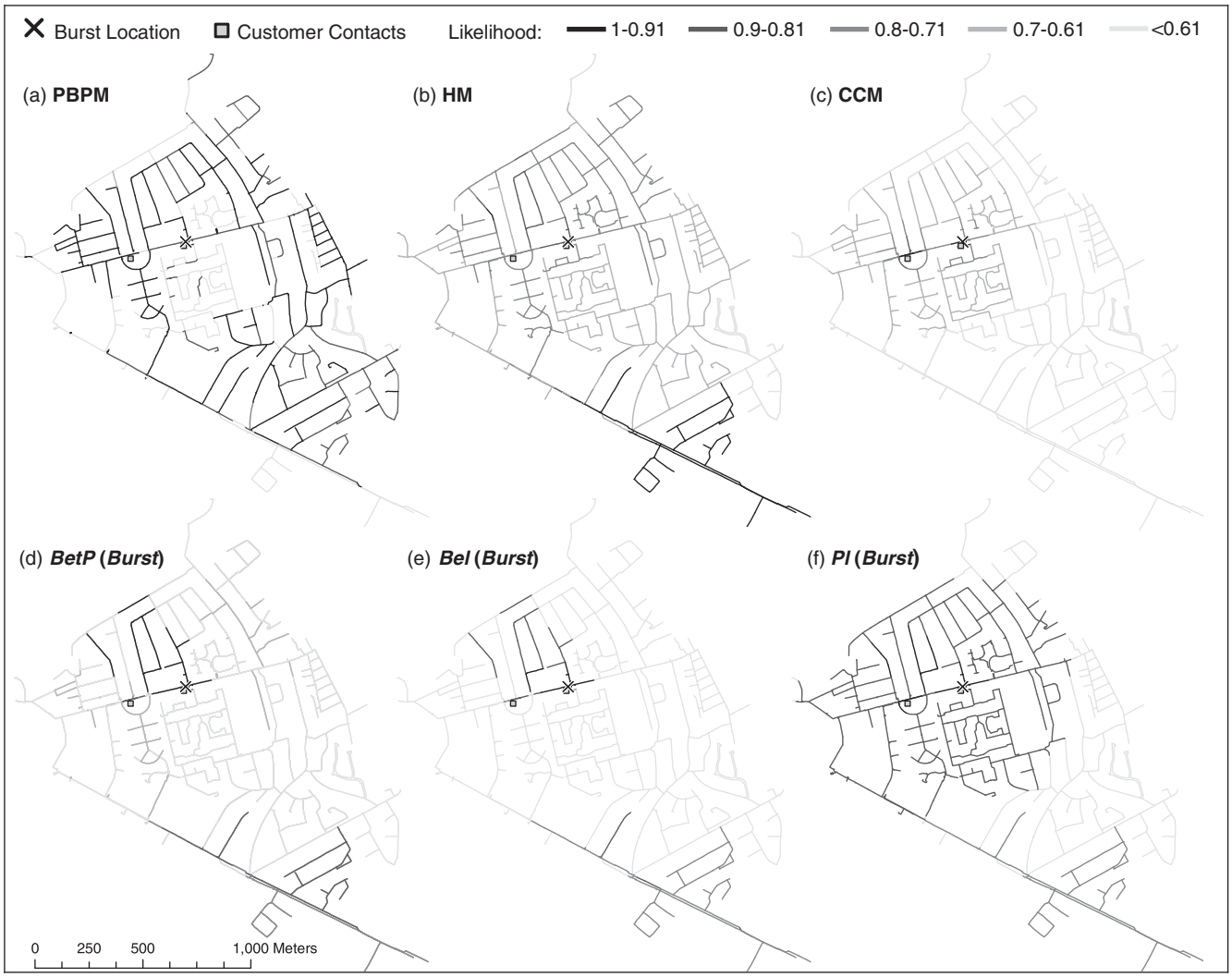

Figure 4 | Example output from the (a) PBPM, (b) HM and (c) CCM and the DS model: (d) *BetP*(*Burst*), (e) *Bel*(*Burst*) and (f) *Pl*(*Burst*).

the output of the DS model (i.e. the *BetP*(*Burst*)) and the ranking assigned by individual models (i.e. criterion measurements). The performance of any model was considered good if the real burst location was among the top 10 burst candidates identified by the respective model. As can be seen from Table 1 none of the individual input models, i.e. the PBPM,

Table 1 | An overview of the performance of the DS model

	Rank of real burst location < 10				Rank of real burst location < 50			
Scenario	**DS model**	**PBPM**	**HM**	**CCM**	**DS model**	**PBPM**	**HM**	**CCM**
Calibration (no CC)	28.6%	0.0%	14.3%	0.0%	71.4%	7.1%	42.9%	0.0%
Calibration (CC)	74.1%	0.0%	29.6%	66.7%	85.2%	11.1%	66.7%	66.7%
Validation (no CC)	0.0%	0.0%	0.0%	0.0%	40.0%	0.0%	40.0%	0.0%
Validation (CC)	87.5%	0.0%	62.5%	62.5%	87.5%	0.0%	75.0%	75.0%

HM and CCM, was able to achieve the above goal in all of the situations (i.e. 54 historical pipe bursts) considered in the case study. The degree of success in identifying the location of a burst pipe varied significantly amongst the models. According to this assessment criterion the overall performance of the DS model was, on average, in every scenario either equally good or better than any of the individual models. A similar performance can be observed in Table 1 where the number of potential burst candidates was increased from 10 to 50.

Evaluating the benefits of information fusion algorithms is not simple and using only the measure above would not reflect the additional advantages of this approach. A particular model might fail to identify the correct burst location according to the criteria used above but can, on the other hand, still identify a number of locations where the burst pipe is unlikely to be located. To take this fact into account and to compare the quality of the output of the DS model and the individual models, the following set of performance indicators was established:

1. *Likelihood concentration.* For the method to be useful operationally, it is important that the likelihood of burst occurrence assigned to the pipes near the real burst location is higher than the likelihood assigned to pipes further away. This can be expressed using the ratio of the average likelihood of occurrence of the burst assigned to pipes close to the true burst location over the average likelihood of burst occurrence assigned to all remaining pipes. The higher this ratio is, the better the overall performance of a particular model. The set of pipes in the proximity of the true burst location was assumed here as the 10 topologically nearest pipes. Given that the average length of the pipes in the case study area was 30 m and that the network was highly looped, such resolution should be considered acceptable.
2. *Certainty*. According to Yager (2004), Shannon entropy (Shannon 1948) was used to characterise the certainty of the outputs of the individual models and the DS model. The entropy of an information source (i.e. output of a particular model) was calculated using Equation (10) and its certainty can be expressed using Equation (11). The higher the certainty of a particular model the better was its performance:

$$H = -\sum_{k=1}^{N_P} p_k(Burst)\ln(p_k(Burst)) \tag{10}$$

$$Certainty = 1 - \frac{H}{\ln(N_P)} \tag{11}$$

where H is Shannon entropy, p_k is either the normalised $BetP_k(Burst)$ or the normalised value of the confidence factor of a potential incident (pipe) k in the case of the DS model and the individual models, respectively, and N_P is the number of potential incidents (i.e. pipes) in the system

The results of the comparison based on the two additional criteria suggested above are shown in Table 2, which indicates in how many calibration and validation cases was the DS model better than the individual models (values above 50% indicate that the DS model on average improved over the prediction of an individual model and 100% means that the DS model was better in all considered cases than a particular individual model). Again, cases are further split into scenarios where customer contacts were and were not available.

As can be seen from Table 2, the DS model yields better results (e.g. DS > PBPM) in terms of the likelihood concentration in a higher number of cases compared to the individual models. The DS model was significantly better than the

Table 2 | Performance of the DS model compared with PBPM, HM and CCM based on spatial distribution of the likelihood of potential pipe bursts

	Likelihood concentration (%)			Certainty (%)		
Scenario	**DS > PBPM**	**DS > HM**	**DS > CCM**	**DS > PBPM**	**DS > HM**	**DS > CCM**
Calibration (No CC)	100.0	100.0	–	85.7	28.6	–
Calibration (CC)	96.3	100.0	100.0	96.3	44.4	100.0
Validation (No CC)	80.0	80.0	–	80.0	0.0	–
Validation (CC)	100.0	100.0	100.0	100.0	75.0	100.0

PBPM and CCM in view of the certainty criterion: however, in some situations, it performed worse than the HM. This fact is most apparent in scenarios where no customer contacts were received and only the outputs of the HM and PBPM were combined. In such situations the most likely locations of the burst pipe typically form a number of scattered hotspots rather than a relatively well-confined area as shown in Figure 4(d). Despite this fact, the use of the PBPM as an information source still yields certain benefits as illustrated in Table 1.

Sensitivity analysis

To investigate the sensitivity of individual model outputs as well as the DS model output to the noisy inputs, global sensitivity analysis using Monte Carlo simulation (1000 samples) was performed on the example presented in Figure 4. The selected case represented a suitable scenario since at least two of the individual models (i.e. the HM and the CCM) performed acceptably and therefore the effect of the added noise could be observed. Various levels of uniformly distributed noise, as indicated in Table 3, were added to the inputs of the individual models, namely the HM (observed pressures, demands and estimated burst flow) and the CCM (easting and northing). Adding noise to the PBPM would be problematic and given its relatively low credibility it would not make a significant difference in this case. The "AVG rank" of the PBPM was therefore the same for all scenarios and had the value of 742.0 out of 1052 potential incidents (i.e. poor performance in this case).

The "AVG rank" shown in Table 3 is the rank as described in the performance comparison section, averaged over 1000 samples. The lower the value of the "AVG rank" the better the performance of a particular model is. As can be seen from Table 3 the combined results were, in all scenarios (for this particular case), slightly worse than those of the best model (such information is, however, unknown until the burst is located by a field technician). On the other hand, the DS model outputs are, to some extent, less sensitive to the noise added to the inputs of individual models. If the performance of only one of the models degrades significantly, the two remaining models (the CCM or HM in particular) would still influence the combined results so that they did not degrade as fast as the worst model. However, in cases where the quality of evidence of the most influential input models (i.e. the HM and the CCM) deteriorated at the same

Table 3 | Results of a global sensitivity analysis

Scenario	Burst flow	Pressure noise	Demands noise	Burst flow noise	CC noise	AVG DS rank	AVG HM rank	AVG CCM rank
A	5	1.0%	5.0%	0.5%	0.01%	6.4	4.8	10.2
B	5	2.0%	10.0%	1.0%	0.01%	7.2	5.4	10.2
C	3	2.0%	7.5%	1.0%	0.01%	62.6	84.0	10.2
D	5	3.0%	10.0%	2.0%	0.02%	24.8	16.5	14.4
E	5	4.0%	10.0%	2.0%	0.02%	42.9	52.5	14.4
F	7	3.0%	10.0%	5.0%	0.03%	8.1	5.2	20.1
G	5	2.0%	7.5%	2.0%	0.03%	8.4	5.5	20.1
H	5	2.0%	7.5%	4.0%	0.03%	9.9	5.6	20.1
I	5	3.0%	7.5%	1.0%	0.03%	26.4	16.3	20.1
J	3	2.0%	7.5%	1.0%	0.03%	66.9	84.0	20.1
K	5	1.0%	7.5%	0.5%	0.05%	8.1	4.8	37.5
L	5	2.0%	10.0%	1.0%	0.05%	8.9	5.4	37.5
M	3	2.0%	10.0%	1.0%	0.05%	78.0	84.0	37.5
N	7	2.0%	10.0%	2.0%	0.08%	8.2	4.9	70.1
O	7	5.0%	10.0%	2.0%	0.08%	26.5	9.8	70.1

time (e.g. because of the amount of noise present in the data or due to low burst flow; illustrated in scenarios D and I in Table 3), then the combined results were worse than those of any of the two key input models.

CONCLUSIONS

Locating a pipe burst within a DMA using data-driven or conventional model-based methods is a challenging problem. The main constraint of such methods is typically the lack of data or insufficient calibration of the models used. Under such conditions of uncertainty, when no single model is able to provide a satisfactory answer, it is beneficial to combine the outputs from several models, based on different inputs, in order to improve the confidence in the overall result. This paper presents a methodology based on the Dempster–Shafer theory which combines evidence from several independent sources/models (i.e. a pipe burst prediction model, a hydraulic model and a customer contacts model) to locate a pipe burst within a DMA. It is argued that this methodology is able to fully exploit all information sources available in a WDS control room and reduce the information load that needs to be processed by a human operator and facilitate targeted field investigations.

A limiting factor to a wider application of hydraulic models in near-real-time burst diagnostics is the unavailability of pressure and flow data in sufficient quantity and quality. In certain WDS, deployment of a sufficient number of sensors might be uneconomical since the potential benefits from timely burst identification would not justify the cost of the sensors. However, strengthening requirements on delivered levels of service and customer satisfaction might support more investment in monitoring technology in the not-too-distant future which, coupled with the availability of cheaper sensors due to technological advances, may tip the balance of the cost–benefit analysis. Water utilities in the UK have only recently started to collect such data and even now it is still difficult to find a sufficient number of pressure monitoring points. The lack of field data prevented the application of the methodology to a real-life system. A more detailed analysis to verify the performance of the methodology on a range of real cases is required. The results obtained in a number of semi-real historical pipe bursts suggest that the method (depending on the quality of the input evidence) is capable of identifying the most likely area of the pipe burst. Initial calibration and maintenance of the mapping curves (e.g. Figure 3), which reflect the credibility of the models used, is not straightforward and poses a challenge that needs to be addressed.

It is concluded that a major strength of the proposed methodology is that it has the potential to learn from the performance of individual models during the calibration stage and successfully apply this knowledge to unseen cases. As information about new pipe bursts becomes progressively available, the DS model can be recalibrated in order to better reflect the evolving performance of the input models. Moreover, additional models suggesting the location of a burst pipe (e.g. based on the information of third parties working in the system, weather information, etc.) can be readily incorporated, acting as additional information sources, to further improve the performance of and benefits from information fusion.

ACKNOWLEDGMENTS

The work on the NEPTUNE project was supported by the UK Science and Engineering Research Council, grant EP/E003192/1 and industrial collaborators. In particular, the authors would like to express their gratitude to Mr. Ridwan Patel from Yorkshire Water Services and Dr. Steve Mounce from the Pennine Water Group for their kind assistance. The authors also thank the anonymous reviewers for their constructive comments.

REFERENCES

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

Bai, H., Sadiq, R., Najjaran, H. & Rajani, B. 2008 Condition assessment of buried pipes using hierarchical evidential reasoning model. *J. Comput. Civil Engng.* **22**(2), 114–122.

Bayes, T. 1763 An essay towards solving a problem in the doctrine of chances. *Phil. Trans. R. Soc.* **53**, 370–418.

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

Beynon, M. J. 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.

Bi, Y., Guan, J. & Bell, D. 2008 The combination of multiple classifiers using an evidential reasoning approach. *Artif. Intell.* **172**(15), 1731–1751.

Brunone, B. 1999 Transient test based technique for leak detection in outfall pipes. *J. Wat. Res. Plann. Mngmnt.* **125**(5), 302–306.

Bush, C. A. & Uber, J. G. 1998 Sampling design methods for water distribution model calibration. *J. Wat. Res. Plann. Mngmnt.* **124**(6), 334–344.

Colombo, A. F. & Karney, B. W. 2002 Energy and costs of leaky pipes: toward a comprehensive picture. *J. Wat. Res. Plann. Mngmnt.* **128**(6), 441–450.

Colombo, A. F., Lee, P. & Karney, B. W. 2009 A selective literature review of transient-based leak detection methods. *J. Hydroenviron. Res.* **2**(4), 212–227.

Cooper, N. R., Blakey, G., Sherwin, C., Ta, T., Whiter, J. T. & Woodward, C. A. 2000 The use of GIS to develop a probability-based trunk mains burst risk model. *Urban Wat.* **2**(2), 97–103.

Démotier, S., Denœux, T. & Schön, W. 2003 Risk assessment in drinking water production using belief functions. In: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (T. D., Nielsen, N. L., Zhang (Eds.)) pp.319–331. Springer, Berlin.

Dempster, A. P. 1967 Upper and lower probabilities induced by a multivalued mapping. *Annal. Math. Statist.* **38**(2), 325–339.

Kapelan, Z., Savic, D. A. & Walters, G. A. 2003a A hybrid inverse transient model for leakage detection and roughness calibration in pipe networks. *J. Hydraul. Res.* **41**(5), 481–492.

Kapelan, Z., Savic, D. A. & Walters, G. A. 2003b Multiobjective sampling design for water distribution model calibration. *J. Wat. Res. Plann. Mngmnt.* **129**(6), 466–479.

Kleiner, Y. & Rajani, B. 2001 Comprehensive review of structural deterioration of water mains: statistical models. *Urban Wat.* **3**(3), 131–150.

Li, H. 2007 *Hierarchical Risk Assessment of Water Supply Systems*. PhD thesis, Loughborough University, Loughborough, Leicestershire.

Marashi, S. E., Davis, J. P. & Hall, J. W. 2008 Combination methods and conflict handling in evidential theories. *Int. J. Uncertainty Fuzziness Knowledge-Based Syst.* **16**(3), 337–369.

Misiunas, D., Vítkovský, J., Olsson, G., Lambert, M. & Simpson, A. 2006 Failure monitoring in water distribution networks. *Wat. Sci. Technol.* **53**(4–5), 503–511

Mounce, S. R., Boxall, J. B. & Machell, J. 2010 Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows. *J. Wat. Res. Plann. Mngmnt.* **136**(3), 309–318.

Mounce, S. R. & Machell, J. 2006 Burst detection using hydraulic data from water distribution systems with artificial neural networks. *Urban Wat. J.* **3**(1), 21–31.

Poulakis, Z., Valougeorgis, D. & Papadimitriou, C. 2003 Leakage detection in water pipe networks using a Bayesian probabilistic framework. *Probab. Engng. Mech.* **18**(4), 315–327.

Puust, R., Kapelan, Z., Savic, D. A. & Koppel, T. 2010 A review of methods for leakage management in pipe networks. *Urban Wat. J.* **7**(1), 25–45.

Romano, M., Kapelan, Z. & Savic, D. A. 2009 Bayesian-based online burst detection in water distribution systems. In: *Proc. 10th International Conference on Computing and Control for the Water Industry, CCWI 2009 "Integrating Water Systems", Sheffield, UK* (eds J. Boxall and C. Maksimović) pp. 331–337. CRC Press, Florida.

Rossman, L. A. 2000 *EPANET 2 Users Manual*. US Environmental Protection Agency, Cincinnati, OH

Sadiq, R., Kleiner, Y. & Rajani, B. 2006 Estimating risk of contaminant intrusion in water distribution networks using Dempster–Shafer theory of evidence. *Civil Engng. Environ. Syst.* **23**(3), 129–141.

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

Safranek, R. J., Gottschlich, S. & Kak, A. C. 1990 Evidence accumulation using binary frames of discernment for verification vision. *IEEE Trans. Robotics Automation* **6**(4), 405–417.

Sentz, K. & Ferson, S. 2002 *Combination of Evidence in Dempster–Shafer Theory*. SAND 2002–0835, Sandia National Laboratories, Albuquerque, NM.

Shafer, G. A. 1976 *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shannon, C. E. 1948 A mathematical theory of communications, I and II. *Bell Syst. Tech. J.* **27**, 379–423.

Smarandache, F. & Dezert, J. 2006 *Advances and Applications of DSmT for Information Fusion II (Collected Works)*. American Research Press Rehoboth, NM.

Smets, P. & Kennes, R. 1994 The transferable belief model. *Artif. Intell.* **66**(2), 191–234.

Wang, Y., Zayed, T. & Moselhi, O. 2009 Prediction models for annual break rates of water mains. *J. Perf. Constr. Facil.* **23**(1), 47–54.

Wu, Z. Y., Sage, P. & Turtle, D. 2010 Pressure-dependent leak detection model and its application to a district water system. *J. Wat. Res. Plann. Mngmnt.* **136**(1), 116–128.

Yager, R. R. 1987 On the Dempster–Shafer framework and new combination rules. *Inf. Sci.* **41**(2), 93–137

Yager, R. R. 2004 On the determination of strength of belief for decision support under uncertainty – Part II: Fusing strengths of belief. *Fuzzy Sets Syst.* **142**(1), 129–142.

First received 17 November 2009; accepted in revised form 28 June 2010. Available online 26 November 2010

APPENDIX

To illustrate the actual process of information fusion used in this work a simplified example of one potential incident (i.e. pipe segment "P1") and two sources of evidence (i.e. the PBPM and HM only) is presented here.

1. The PBMP and the HM are run for every considered potential incident in a DMA and for the selected potential incident ("P1") return the following result: $CriterionMeasurement_{\mathrm{PBPM}}$ ("P1") = 481 bursts/1000 km/yr (burst rate) $CriterionMeasurement_{\mathrm{HM}}$ ("P1") = 5.42 m^2 (SSE).
2. Confidence factors are then obtained after normalising criterion measurements of every considered model using a suitable normalisation function (i.e. sigmoid function for the PBPM and logit function for the HM):

$$ConfidenceFactor_{\mathrm{PBPM}}(\text{"P1"}) = 0.998$$

$$ConfidenceFactor_{\mathrm{HM}}(\text{"P1"}) = 0.635.$$

3. From the value of the confidence factor the BPAs are obtained for each of the models using their mapping functions (see, e.g., Figure 1(b)). The actual mapping functions of the PBPM and the HM are not presented in this paper:

$$m_{\mathrm{PBPM}}(\{Burst\}) = 0.357,\ m_{\mathrm{PBPM}}(\{No\ Burst\}) = 0.014,\ m_{\mathrm{PBPM}}(\{Burst\},\ \{No\ Burst\}) = 0.629$$

$$m_{\mathrm{HM}}(\{Burst\}) = 0.000,\ m_{\mathrm{HM}}(\{No\ Burst\}) = 0.130,\ m_{\mathrm{HM}}(\{Burst,\ \{No\ Burst\}) = 0.870.$$

4. According to Equations (1), (2) and (6) the *Bel*, *Pl* and *BetP* structures can be calculated as follows:

$$Bel_{\mathrm{PBPM}}(\{Burst\}) = m_{\mathrm{PBPM}}(\{Burst\}) = 0.357$$

$$\begin{aligned} Pl_{\mathrm{PBPM}}(\{Burst\}) &= m_{\mathrm{PBPM}}(\{Burst\}) + m_{\mathrm{PBPM}}(\{Burst\}, \{No\ Burst\}) = 0.357 + 0.629 \\ &= 1 - 0.014 = 0.986 \end{aligned}$$

$$\begin{aligned} BetP_{\mathrm{PBPM}}(\{Burst\}) &= [Pl_{\mathrm{PBPM}}(\{Burst\}) + Bel_{\mathrm{PBPM}}(\{Burst\})]/2 \\ &= [0.986 + 0.357]/2 = 0.672 \end{aligned}$$

$$Bel_{\mathrm{HM}}(\{Burst\}) = m_{\mathrm{HM}}(\{Burst\}) = 0.000$$

$$\begin{aligned} Pl_{\mathrm{HM}}(\{Burst\}) &= m_{\mathrm{HM}}(\{Burst\}) + m_{\mathrm{HM}}(\{Burst\}, \{NoBurst\}) = 0.000 + 0.870 \\ &= 1 - 0.130 = 0.870 \end{aligned}$$

$$\begin{aligned} BetP_{\mathrm{HM}}(\{Burst\}) &= [Pl_{\mathrm{HM}}(\{Burst\}) \\ &\quad + Bel_{\mathrm{HM}}(\{Burst\})]/2 = 0.435. \end{aligned}$$

5. Once the BPAs are obtained, Dempster's combination rule defined in Equations (3)–(5) can be applied:

$$\begin{aligned} K &= m_{\mathrm{PBPM}}(\{Burst\}) \times m_{\mathrm{HM}}(\{No\ Burst\}) \\ &\quad + m_{\mathrm{HM}}(\{Burst\}) \times m_{\mathrm{PBPM}}(\{No\ Burst\}) \\ &= 0.357 \times 0.130 + 0.000 \times 0.014 = 0.046 \end{aligned}$$

$$\begin{aligned} m_{\mathrm{PBPM,HM}}(\{Burst\}) &= [m_{\mathrm{HM}}(\{Burst\}) \times m_{\mathrm{PBPM}}(\{Burst\}) \\ &\quad + m_{\mathrm{HM}}(\{Burst\}) \times m_{\mathrm{PBPM}}(\{Burst\}, \{No\ Burst\}) + m_{\mathrm{PBPM}}(\{Burst\}) \\ &\quad \times m_{\mathrm{HM}}(\{Burst\}, \{No\ Burst\})]/(1 - K) \\ &= [0.000 \times 0.014 + 0.000 \\ &\quad \times 0.629 + 0.357 \\ &\quad \times 0.870]/(1 - 0.046) = 0.326 \end{aligned}$$

$$\begin{aligned} m_{\mathrm{PBPM,HM}}(\{No\ Burst\}) &= [m_{\mathrm{HM}}(\{No\ Burst\}) \\ &\quad \times m_{\mathrm{PBPM}}(\{No\ Burst\}) \\ &\quad + m_{\mathrm{HM}}(\{No\ Burst\}) \\ &\quad \times m_{\mathrm{PBPM}}(\{Burst\}, \{No\ Burst\}) \\ &\quad + m_{\mathrm{PBPM}}(\{No\ Burst\}) \\ &\quad \times m_{\mathrm{HM}}(\{Burst\}, \{No\ Burst\})]/(1 - K) \\ &= 0.130 \times 0.014 + 0.130 \\ &\quad \times 0.629 + 0.014 \\ &\quad \times 0.870]/(1 - 0.046) = 0.1 \end{aligned}$$

$$\begin{aligned} m_{\mathrm{PBPM,HM}}(\{Burst),\ \{NoBurst\}) &= 1 - m_{\mathrm{PBPM,HM}}(\{Burst\}) \\ &\quad - m_{\mathrm{PBPM,HM}}(\{NoBurst\}) \\ &= 1 - 0.326 - 0.1 \\ &= 0.574. \end{aligned}$$

The corresponding belief structures *Bel*, *Pl* and *BetP* can then be easily calculated using Equations (1), (2) and (6), respectively. Given the associativity of Dempster's rule the combined results obtained above could again be combined with evidence from the CCM. If other combination rules (e.g. Yager's or PCR5) were applied, their quasi-associative versions would have to be used since the fusion results should be independent of the order in which evidence is combined.