

Investigating a spatially constrained geostatistical technique for approximate leak/burst event location

Shaun Boatwright^{1, 2, *}, Michele Romano², Stephen Mounce¹, Kevin Woodward² and Joby Boxall¹

¹ Pennine Water Group, Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, UK

² United Utilities, Lingley Mere Business Park, Warrington, WA5 3LP, UK

*shaun.boatwright@uuplc.co.uk

ABSTRACT

Responding quickly to leaks and bursts in water distribution systems can provide a number of benefits to water companies. Therefore, methods which allow a water company to determine the approximate location of leaks and bursts in a timely manner are desirable. This paper presents 10 examples of a spatially constrained inverse-distance weighted interpolation technique for determining the approximate location of a new leak or burst which has been tested using pressure data collected during a number of bursts simulated by opening fire hydrants in a real water distribution system. The results of the examples indicate that the SCIDW technique can reduce the proportion of the DMA which needs to be searched but that there are a number of factors, which have been identified as a result of these examples, which affect the leak/burst localisation performance which are not currently accounted for by the methodology.

Keywords: Leak/burst localisation, inverse-distance weighted interpolation, spatial constraint

1 BACKGROUND

Leak and pipe burst events in Water Distribution Systems (WDS) are a compelling issue for the water industry worldwide as the loss of large volumes of treated and frequently pumped water is environmentally and economically damaging. Furthermore, these events have a negative impact on the water companies' operational performance and can impact upon large numbers of customers by, for example, causing water supply problems and disruption to businesses. The impacts of leaks and bursts are closely linked to the time between their occurrence and repair (often referred to as the lifecycle of a leak/burst [1]). As water companies are still typically reliant on their customers reporting problems with the water supply (such as low pressure / no water or a visible event) and on the use of resource intensive acoustic techniques [1], alternative and more efficient ways to minimise the leak/burst lifecycle are needed.

Continuous, near real-time monitoring of WDS's hydraulic parameters (i.e. pressure and flow) has been the focus of much research in recent times, due to its' potential to help mitigate the leak/burst problem. A number of innovative data-driven leak/burst detection techniques which can analyse the large amounts of data now being collected from WDSs have been developed and some have even been implemented by water companies [2, 3], with wider proliferation ongoing. These techniques have demonstrated the ability to minimise the leak/burst awareness time by effectively detecting, at the district metered area (DMA) level, leaks and bursts as soon as they occur. Taking this a step further, methods which utilise hydraulic data to determine the approximate location of leaks and bursts within a DMA have recently emerged [4-8]. These methods can allow a further reduction in the response time to leak/burst events by providing an indication of their likely location and, hence,

reducing the amount of time taken to find their precise location (i.e., pinpointing on a pipe) using methods such as leak noise correlators [9]. Key to the success of the approximate leak/burst location methods is the additional instrumentation required. On the one hand, the selection of suitable sensor locations to maximise sensitivity to leaks/bursts whilst minimising the required number of instruments (and hence cost) is very important and a number of methodologies have been developed for this purpose [10-12]. On the other hand, geostatistical interpolation techniques, which enable the inference of the value of a variable at an unmeasured location using measurements from nearby sensors, have also been proposed [5] to improve the approximate leak/burst location accuracy achievable with a given number of instruments. To ensure improved accuracy, however, geostatistical interpolation techniques for approximate leak/burst location must be able to account for the topology of the particular network in which they are being used by using an appropriate distance function. Therefore, previous work by the authors [7] presented a spatially constrained inverse distance weighted spatial interpolation technique that utilises the layout of pipes within a DMA to direct the interpolation to occur along the pipes only (instead of taking the Euclidean distance between sensors), demonstrating the potential to significantly enhance the leak/burst approximate location performance. This paper aims to build upon previous work by examining a number of applications of a constrained geostatistical technique which can determine the approximate location of a leak/burst in a DMA, to determine which factors influence its' performance.

The remainder of the paper shall be broken down as follows; section 2 will outline the methodology, section 3 will describe a case study containing the preliminary results of a number of applications of a geostatistical leak location technique, section 4 contains discussion of the results and planned future work and section 5 will examine the conclusions which have arisen as a result of the work included in this paper.

2 METHODOLOGY

Two techniques for determining the approximate location of a leak/burst in a DMA were presented previously by the authors [6, 7]. The first of these utilised statistical process control (SPC – see [13] for details) in order to determine the change in pressure at each sensor location which can be attributed to a leak or burst, once suspected or detected using one of many leak detection techniques available [14,15]. The sensor locations were subsequently prioritised for searching based upon the magnitude of the change calculated by the leak location system (LLS) by comparing the behaviour of each pressure signal, during a user-specified time window, to its previous behaviour prior to the existence of the leak or burst. The LLS was tested on a number of engineered events (bursts simulated by opening fire hydrants) in 17 real DMAs to validate its performance and the results showed encouraging performance (see [6] for more details of the testing procedure). Building upon this, improvements were made to the LLS methodology and it was tested on another set of engineered event data collected from 13 different DMAs within the same WDS. The results showed that the improvements to the methodology enabled a modest improvement in the results, allowing reduction of the search area to less than 1/3 of a DMA in 58% of the test events (see the case study in [7] for details). A significant limitation of the method was that once the sensor closest to the leak/burst was correctly identified no further information about the likely location of the leak or burst in relation to the sensor was available, limiting the accuracy of the method. With this in mind, a novel spatially constrained inverse-distance weighted interpolation (SCIDW) technique for

approximate leak/burst location was developed which utilised the pressure drop calculated for each sensor location to calculate the value of pressure drop for all locations within the DMA to overcome the aforementioned limitation. This technique was based upon inverse-distance weighted (IDW) [16] interpolation which is commonly used for inferring the value of an unmeasured variable based upon a number of measured points distributed throughout the space being considered.

The SCIDW technique previously demonstrated by the authors [7] relies on data coming from larger numbers of pressure sensors than are typically installed in UK DMAs. Once the presence of a leak or burst has been confirmed, and prior to performing the SCIDW interpolation technique, the pressure drop at each sensor location caused by the leak/burst was calculated. This is achieved by comparing the mean pressure at each sensor location during the period coinciding with the minimum night flow (between 03:00 and 04:00) when the leak or burst was running, with the mean pressure for the same time window for a number of days prior to the initiation of the leak/burst (referred to the “normal dataset” from this point forward). Based upon previous studies, 14 days has been used in this case study to prevent a small number of anomalous days (for example, days where network operations, such as temporary re-zoning, have been carried out) being selected as the normal dataset which could then lead to large variations in the mean and thus incorrect mean values for the normal dataset. Once the pressure drop was calculated for each sensor location the interpolation procedure as set out in [7] was followed to produce a “heatmap” showing the interpolated pressure drop for the entire DMA. In order to determine which parts of the DMA should be searched a threshold was used whereby any values of pressure drop which fall above a pre-determined threshold (as a % between the minimum and maximum values of calculated pressure drop in the DMA) are included in the search area. Based upon a limited investigation a value of 60% has been used in the case study. Further explanation with regard to the selection of the threshold and a sensitivity study can be found in the section 3.2.

3 CASE STUDY

3.1 Description

In order to determine which factors impact upon the effectiveness of the previously described SCIDW interpolation technique, it has been applied to 10 engineered events conducted in a single DMA over a several month period. The DMA used to test the SCIDW interpolation technique contains 16.3km of mains, 1664 residential properties and is fed by gravity. The engineered events in this DMA were conducted in two phases; during phase 1 there were an additional (in addition to the single pressure and flow instruments United Utilities currently uses to monitor all of its DMAs) 5 pressure sensors installed in the DMA and during phase 2 there were an additional 7 pressure sensors installed in the DMA (i.e. 2 more than during phase 1). For some of the engineered events (T2, T4 and T10) not all of the additional pressure sensors were used to perform the SCIDW interpolation. This was because a number of days of data was missing from the normal dataset due to temporary failures of these instruments, leading to erroneous values for the pressure drop being calculated in those cases. The engineered events were carried out between 02:00 and 04:00, to coincide with the period of minimum night flow, in order to reduce the risk of supply problems and disruption to customers. Details of the installation location of each additional pressure sensor, as well as the location of the hydrants opened for each engineered event can be found in figure 1(a) below. Table 1 summarises the distances between the opened hydrants and each of the additional pressure sensors installed in the DMA during each engineered event. The distances included in table

1 represent the shortest distance from the opened hydrant to each sensor by travelling via the DMA pipes only. The distances were calculated by converting a GIS layer containing all of the DMA pipes to a weighted undirected graph [16]. This means that the DMA pipes are converted to a network of nodes which are connected by links. The total distance when travelling from one node to another, via several intermediate links/nodes, can be calculated as the sum of the individual links travelled. For this purpose Dijkstra’s algorithm [17] has been used. In order to achieve this processing of the GIS layer was required in order to remove duplicate pipes (i.e. 2 pipes which have the same start node and end node and differ in length). In the case where duplicate pipes were removed, the longest pipes were removed to ensure that the minimum distance was being travelled in each case. Figure 1 also show the pipe properties for each pipe in the test DMA which are used in section 4 to qualitatively analyse the results of the SCIDW interpolation.

Engineered Event	Distance from each additional pressure sensor to the opened hydrant (m)						
	1	2	3	4	5	6	7
T1	1327	316*	361	983	698	-	-
T2	1327 ⁺	316*	361	983	698	-	-
T3	1020	720	255*	1312	544	-	-
T4	1086 ⁺	744	639	1382	133*	-	-
T5	140*	1317	874	1931	813	-	-
T6	1035	654	549	1293	174*	-	-
T7	1082	397	94*	1064	629	744	552
T8	781	1137	694	1751	633	1432	249*
T9	1767	537	779	680	1162	150*	1237
T10	2232	1002	1244	227*	1627	789 ⁺	1702

Table 1. Summary of distances between the opened hydrant and additional pressure sensors for each engineered event in the test DMA (* denotes the sensor closest to the opened hydrant for each engineered event whilst + denotes a sensor which was faulty during the engineered event)

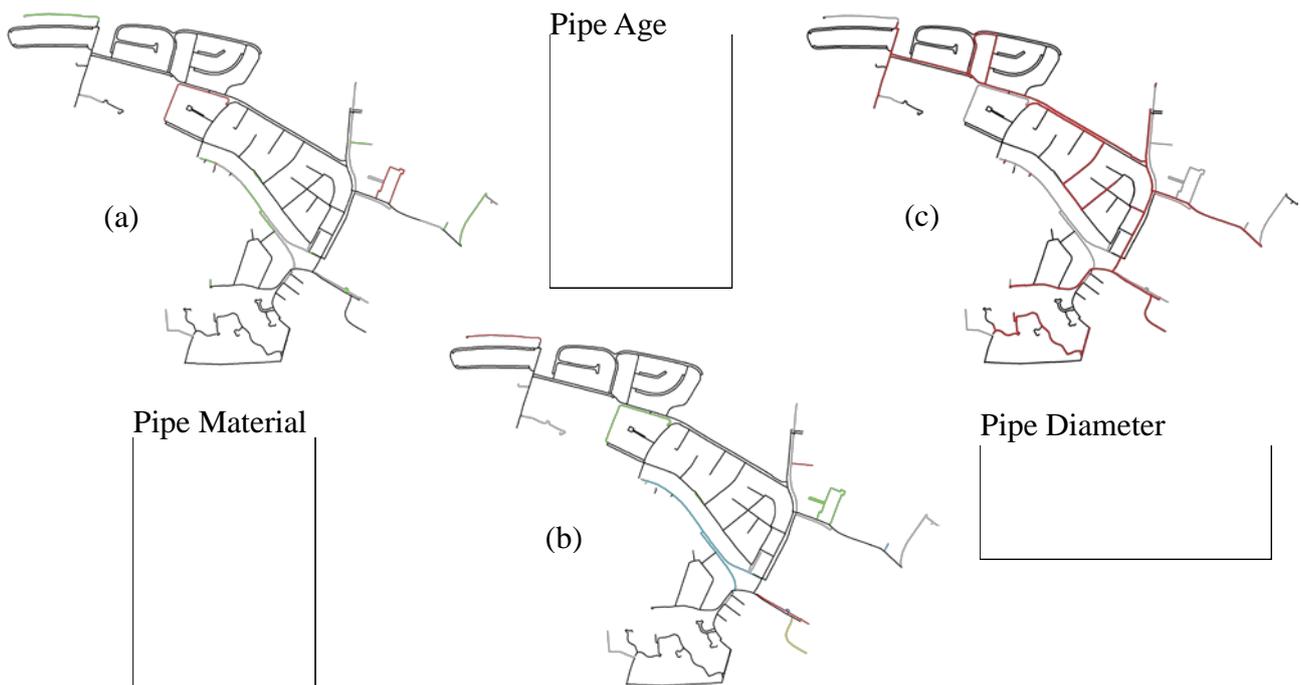


Figure 1. GIS layer of the test DMA showing for each pipe (a) pipe material, (b) pipe age and (c) pipe diameter

3.2 Results

The results of the analysis using SCIDW interpolation for each of the 10 engineered events in the test DMA are presented below. The percentage defines the proportion of the DMA to be searched.

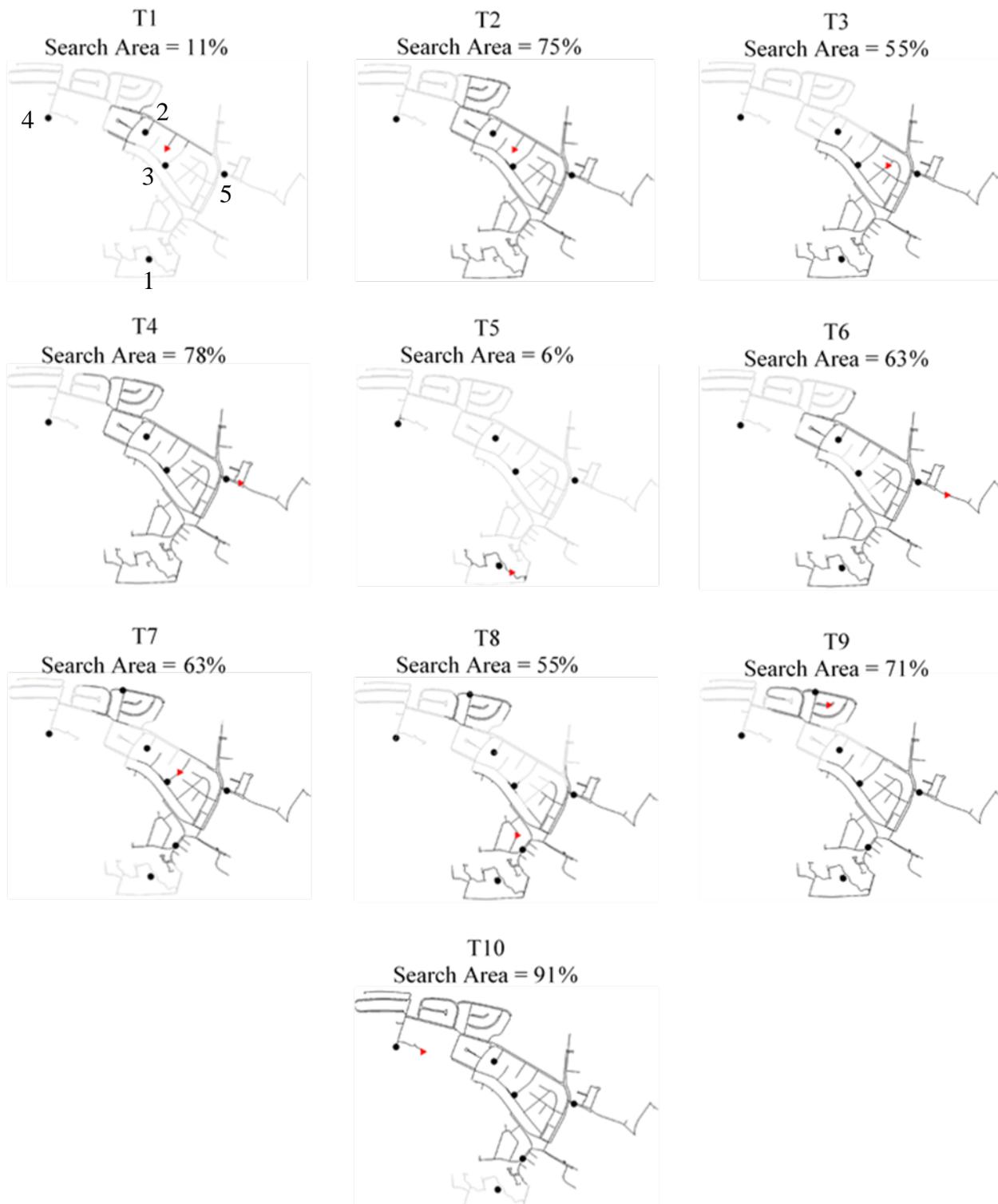


Figure 2. Search area (black pipes) produced by using SCIDW interpolation with a 60% threshold. Additional pressure sensors (black dots) and opened fire hydrants (red triangles) are shown for each of the 10 engineered events.

It is evident that the size of the search area varies considerably, ranging from as small as 6% of the DMA (T5) to 91% (T10) and having an average of 57% across all 10 engineered events which have been analysed. This represents a reduction of almost forty percent (or approximately 700 properties) of the DMA, which can be considered significant. For the majority of the engineered events the SCIDW produced a search area in the range between 55-75% indicating that moderate reductions in the search area can be achieved in most cases. There are two more interesting cases are those where the SCIDW technique enables significant reductions in the search area and effectively guides the field team directly to a small area around the leak/burst (such as for T1 and T5). In both of these cases the sensor closest to the engineered event registered a much larger pressure drop than any of the other sensors enabling the interpolation to correctly exclude the majority of the DMA.

As described previously, a 60% threshold has been used to define the pipes which, following the application of the SCIDW, have a calculated pressure drop high enough to be considered for searching. This threshold is a critical parameter to the performance of the SCIDW technique and, as such, was selected following a study to investigate how it affects both the average size of the resulting search areas and the percentage of correct burst location (i.e. how often the resulting search area contains the location of the engineered event). A range of thresholds between 50 and 85% (in 5% increments) were trialled and the resulting average search area size and percentage of correct burst location were calculated for each, which are shown in figure 3. In order to select the threshold it is desirable to maximise the percentage of correct leak/burst locations whilst minimising the search area. With this in mind the ratio between the two was calculated and plotted as grey triangles in figure 3. 60% was chosen on this basis. Further discussion with regard to the selection of this threshold and the effects that it has on the performance of the SCIDW technique are included in section 4.

4 DISCUSSION AND FUTURE WORK

Some discussion points have arisen as a result of the 10 applications of the SCIDW technique included in section 3.2 of this paper, which will be explored further here. These are primarily concerned with identifying and explaining those factors which influence the performance of the SCIDW technique, in terms of the localisation accuracy achieved in each of the examples. Firstly, one such factor which could be the cause of a number of the cases where a large search area was produced is the highly looped/connected nature of this DMA. Good design of WDSs often calls for multiple pathways to any given point in a DMA so as to provide resilience. This aims to ensure that during catastrophic failure of a single pipe, flow to all properties within the DMA can still be maintained. This, however, also allows the network to compensate for the additional flow caused by a leak or burst and masks the reduction in pressure that may otherwise have resulted. Therefore the performance of the SCIDW for leaks/bursts in the areas of the DMA which are highly interconnected (such as between sensors 3 and 5) tends to reduce. By examining the search area produced in the cases of T3, T7 and T8 it can be seen that the SCIDW technique cannot accurately determine the location of leaks and bursts in this area. In addition to this there is no pressure sensor placed in this area which means that the calculation of the pressure drop in this area can only be known by utilising information from other nearby measured locations. On the contrary, in the case where the method performed best (T5), the opened hydrant and the nearest additional pressure sensor to it were located in an area supplied by a single pipe.

In addition to this other factors influence the hydraulics of the DMA. Therefore, It is clear that use of a purely distance based function for calculating the pressure drop at unmeasured locations is not entirely satisfactory. The variation of pressure in a DMA is dependent upon the physical characteristics of the pipes (such as roughness or diameter) in addition to the flow determined by upstream network conditions and customer water demand. It is important that the interpolation can account for the physical characteristics of the pipes in the DMA and future development should be aimed at incorporating this. For example, by considering the distribution of pipe diameter, material and age (which each affects the head loss along a length of pipe), as per figure 1, it can be seen that the test DMA contains larger pipes (above 100mm) near the inlet and through the centre of the DMA with smaller pipes branching from these. Leaks and burst taking place on or near to these large, plastic pipes (as in the case of T4 and T6) may not produce a noticeable pressure drop as the additional flow of 50 litres per minute can be small in comparison with the “normal” flow. It is worth noting here that the variation in pipe diameter throughout the DMA is much greater than for either material or age. The pipes in the DMA are predominantly polyethylene which are less than 30 years old but the sizes range from 180mm to 50mm. It is proposed, to overcome this shortcoming in the methodology, that the pressure loss along individual pipes (from an offline hydraulic model) in the DMA should be used in place of length. It is expected that this would more accurately represent the operation of the DMA during by accounting for the flow magnitude and direction as well as the physical properties of the pipes in the DMA. However, this leads to the question of how well calibrated the model needs to be to allow accurate localisation of leaks and bursts and this will form a part of the work to incorporate hydraulic parameters into the SCIDW.

In two of the examples (T4 and T6) the location of the engineered event is further away from the inlet than the nearest additional pressure sensors, making the prediction of that point using the SCIDW an extrapolation and not interpolation. The suitability of inverse-distance weighting for extrapolation outside of the convex hull (the smallest polygon containing all measured data points) is not known, particularly for the constrained case which is presented here. This is because traditional inverse-distance weighted interpolation is based upon the proximity, and hence assumed similarity, of the measured and unmeasured points being considered in the DMA. Extrapolation leads to an increase in this distance and as such the relationship between them becomes more uncertain. This is particularly pertinent for pressure measurements which tend to be more affected by upstream changes in flow rather than those downstream as in T4 and T6. This is not the only reason why inverse-distance weighted interpolation may not be the most suitable for this application, however. IDW interpolation has the limitation that the calculated values of the variable being inferred must be between the largest and smallest measured values of the variable. For this application it is expected that the largest value of pressure drop will be at or near the point of the engineered event. However when using IDW interpolation the location of maximum pressure drop will always be calculated at one of the pressure sensor locations. To overcome this, use of a geostatistical technique such as Kriging [18], will allow the value of maximum pressure drop to be at any point in the DMA.

It must also be noted that the average size of the search area has been affected by faulty sensors in 3 of the engineered events. In the case study, faulty sensors were excluded, which affects the size of the search area as the area around the faulty sensor is now calculated based upon the values of the other sensors which are further away. This tends to have the effect of making the pipes around the faulty sensor take a value similar to the nearest working sensor. If we consider the average search area for only the engineered events which had all of additional pressure sensors working then the

average search area drops from 57% to 46%. This can be seen most easily by examining the search areas for both T2 and T4. In both of these examples sensor 1 is faulty and hence the SCIDW bases the value of pressure drop in the area around sensor 1 predominantly on the values of pressure drop at sensors 3 and 5. As sensor 3 and 5 are both closer to the opened hydrant for these engineered events the value of pressure drop around sensor 1 is artificially elevated and therefore incorrectly included in the search area. This therefore raises a question about the selection of the number and location of sensors which has not so far been addressed in this paper. A number of sensor placement algorithms [10-12] are available which can automatically select the “optimal” location for a specified number of sensors in a DMA. Given that these were not used, and that the additional pressure sensors in the test DMA were placed according to engineering judgement, it is expected that a significant increase in the performance of the SCIDW interpolation could be achieved if a sensor placement algorithm was incorporated into the methodology. Having said this, even when using a sensor placement algorithm, faulty sensors will still impact upon the localisation accuracy of any interpolation technique. Therefore a more suitable strategy for dealing with faulty sensor data, as and when it arises, is required.

5 CONCLUSIONS

Due to a foreseen increase in the amount of data being collected from WDS, methods which can analyse this data to allow decisions to be made by water companies are required. This paper has presented 10 applications of the SCIDW interpolation technique which can allow the approximate location of a leak/burst to be calculated by analysing the data collected from a number of pressure sensors distributed in a real DMA. This technique, by considering the topology of the DMA in question, can narrow down the location of a leak/burst to less than 60% of the DMA on average. Indeed, in 2 of the engineered events which have been given as examples the search area was reduced to a small fraction of the DMA (11% for T1 and 6% for T5) further demonstrating the potential of this method to be used operationally to improve the response time to leak/burst events by reducing the time taken to find them prior to repair. Having said this, several issues/limitations of the technique have been identified in this paper which further the understanding of the performance of this technique and will allow further development of this method, as well as the development of other similar interpolation techniques in the future, to help mitigate the leak/burst localisation problem.

6 ACKNOWLEDGEMENTS

The authors wish to thank the innovation team at United Utilities and the engineering and physical sciences research council (EPSRC) for funding this work under a STREAM IDC research project (grant number EP/L015412/1).

References

- [1] R. Puust, Z. Kapelan, D. A. Savic and T. Koppel, “A review of methods for leakage management in pipe networks,” *Urban Water J.*, vol. 7 (1), pp. 25-45, 2010.
- [2] S. R. Mounce, J. B. Boxall and J. Machell, “Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows,” *J. Water Resour. Plan. Manage.*, vol. 136 (3), pp. 309–318, 2010.
- [3] M. Romano, Z. Kapelan and D. A. Savic, “Automated detection of pipe bursts and other events in water distribution systems,” *J. Water Resour. Plan. Manage.*, vol. 140 (4), pp. 457-467, 2014.

- [4] B. Farley, S. R. Mounce and J. B. Boxall, “Development and field validation of a burst localization methodology,” *J. Water Resour. Plan. Manage.*, vol. 139 (6), pp. 604-613, 2013.
- [5] M. Romano, Z. Kapelan and D. A. Savic, “Geostatistical techniques for approximate location of pipe burst events in water distribution systems,” *J. Hydroinform.*, vol. 15 (3), pp. 634–651, 2013.
- [6] M. Romano, Z. Kapelan and K. Woodward, “Testing of a statistical process control based system for approximate location of pipe bursts and leaks in a UK water distribution system,” *Proc. Water Distribution Systems Analysis Conference*, Cartagena de Indias, Colombia, 2016.
- [7] S. Boatwright, M. Romano, S. R. Mounce, K. Woodward and J. B. Boxall, “Approximate location of leaks and bursts in a district metered area using statistical process control and geostatistical techniques,” *Proc. Computing and Control for the Water Industry Conference*, Amsterdam, Netherlands, 2016.
- [8] D. De Silva, J. Mashford and S. Burn, “Computer Aided Leak Location and Sizing In Pipe Network,” *Urban Water Security Research Alliance Technical Publication No. 17* (2nd Edition), 2011.
- [9] J. Thornton, “Water loss control manual,” *McGraw-Hill*, New York, 2002.
- [10] D. P. O. Imschoot, W. R. Furnass, S. R. Mounce and J. B. Boxall, “Flow-pressure sensor placement optimisation for pipe burst localisation in a water distribution network,” *Proc. Computing and Control for the Water Industry Conference*, Amsterdam, Netherlands, 2016.
- [11] D. Steffelbauer and D. Fuchs-Hanusch, “Efficient sensor placement or leak localization considering uncertainties,” *Water Resources Management*. vol. 30 (14), pp. 5517-5533, 2016.
- [12] J. Blesa, F. Nejjari and R. Sarrate, “Robust sensor placement for leak location: Analysis and design,” *J. Hydroinform.*, vol. 18 (1), pp. 136-148, 2016.
- [13] W. Shewhart, “Economic control of quality of manufactured product,” *Van Nostrand Reinhold*, New York, 1931.
- [14] M. Bakker, J. H. G. Vreeburg, M. Van De Roer and L. C. Rietveld, “Heuristic Burst Detection Method Using Flow and Pressure Measurements,” *J. Hydroinform.*, vol. 16 (5), pp. 1194-1209, 2014.
- [15] K. Aksela, M. Aksela and R. Vahala, “Leakage Detection in a Real Distribution Network Using a SOM,” *Urban Water J.*, vol. 6 (4), pp. 279–89, 2009.
- [16] J. Bang-Jensen and G. Gutin, “Digraphs: Theory, Algorithms and Applications,” *Springer-Verlag*, London, 2000.
- [17] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269 – 271, 1959.
- [18] D. Zimmerman, C. Pelvic, A. Ruggles and M. P. Armstrong, “An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting,” *Math. Geol.*, vol. 31 (4), pp. 375–390, 1999.