Environmental Modelling & Software 40 (2013) 78–87

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Linking distribution system water quality issues to possible causes via hydraulic pathways

W.R. Furnass*, S.R. Mounce, J.B. Boxall

Pennine Water Group, Department of Civil and Structural Engineering, The University of Sheffield, Sir Frederick Mappin Building, Mappin Street, Sheffield S1 3JD, UK

ARTICLE INFO

Article history:
Received 29 November 2011
Received in revised form
6 July 2012
Accepted 11 July 2012
Available online 13 September 2012

Keywords:
Distribution networks
Water quality
Data-driven modelling
GIS
PostGIS

ABSTRACT

Our limited understanding and quantification of the variety and complexity of chemical, physical and biological reactions and interactions occurring within drinking water distribution systems currently prohibit the development of a deterministic model of water quality. The causes of known water quality anomalies can however be investigated through mining the large volumes of water quality, hydraulic and asset data currently being collected by utility companies.

The data-driven methodology described here permits historical cause–effect linkages to be identified in a scalable, largely automatable fashion. Under Distribution System Integrated Modelling (DSIM), spatio-temporal searches within the set of pipes that typically lie upstream of a known water quality anomaly are used to identify possible causes. Understanding of the flow paths that connect causes and effects are derived from the results of hydraulic network simulations.

DSIM was used to investigate contacts regarding discolouration and smell/taste issues from customers within a Water Supply Zone in England, UK, over a six-year period. 17.6% of discolouration issues and 17.4% of smell/taste issues were linked to maintenance jobs using the methodology, much smaller proportions than were identified using radial cause searches. The DSIM search results contained a greater proportion of one-to-one linkages and so are less ambiguous than the results of the radial spatio-temporal searches. DSIM was found to be a useful and informative tool for data mining multiple water quality related datasets.

1. Introduction and previous research

The quality of drinking water in the developed world is typically very high with regards to international guidelines and national regulations. In England and Wales the periodic sampling of potable water at treatment works, treated water reservoirs and customers' taps is mandated. In 2010 99.96% of the water samples taken to satisfy the water industry regulators' monitoring requirements were in compliance with regulations (DWI, 2011). There is however a need to ensure that the health impact and the aesthetics of water supplied by water providers continue to be satisfactory.

In many developed countries drinking water distribution system (DWDS) infrastructures are ageing and processes such as corrosion continue to impact on water quality in networks that contain older metal pipework (Kirmeyer, 2002, p. 89). As a result, regulators and water providers are now developing and implementing capital and operational measures for managing water quality in DWDS (DWI, 2002), in addition to the sound catchment management and treatment practices already in place.

The designs of these measures are in part informed by evidence of events and activities within DWDS that have previously had a detrimental effect on the quality of water at customers' taps. In addition, theoretical threats to water quality from intra-network causes are being explored through research. For example:

- 15% of the cases of diarrhoea in a UK study by Hunter et al. (2005) were thought to be associated with bursts and pressure losses within DWDS.
- Payment et al. (1997) considered the possible causes of pathogen ingress to include cross-connections, pipe replacements and the manipulation of valves and hydrants. Besner et al. (2007) then related variations in water quality to such maintenance activities.
- Research is being conducted into how transient pressure waves (resulting from valve closures, for example) can cause contaminant ingress via back-siphonage (LeChevallier et al., 2003; Boyd et al., 2004; Collins et al., 2011).

* Corresponding author.
E-mail address: wrfurnass1@sheffield.ac.uk (W.R. Furnass).

1364-8152/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2012.07.012

- From one study it was calculated that on average 0.23 low pressure events, each of which had the potential to cause contaminant ingress, occurred per year per 1000 population served (WRc, 2008).

It should be noted that a) there is typically much uncertainty associated with the results of epidemiological studies such as Hunter et al. (2005), b) that intra-network water quality issues can have a variety of anthropogenic causes (e.g. power failures causing pump trips; maintenance activities conducted by water companies; structural failures, possibly due to a lack of maintenance/renewal; deliberate contamination) and c) several sets of industry guidelines have been developed in which methods are presented for greatly reducing the risk of operations and maintenance activities negatively impacting on water quality (e.g. Ainsworth and Holt, 2004).

Water providers are not only concerned with the chemical and microbiological quality of supplied water but its aesthetics too. In one study 41% of the complaints made by customers to English and Welsh water providers were regarding drinking water aesthetics, 37% of which pertained to discolouration (Vreeburg and Boxall, 2007).

A considerable proportion of discolouration incidents can be associated with 'known activities' within DWDS. In the period 2006–2008 the Drinking Water Inspectorate for England and Wales attributed 47% of issues to planned works, 6% to pump failures, 16% to valve failures/replacements, 9% to mains damage, 10% to connections, 6% to reservoir issues and 6% to treatment works issues (Husband and Boxall, 2011).

A widely-accepted theory is that discolouration is due to cohesive layers of particulate matter being stripped from pipe walls due to a change in pipe wall shear stress (Husband et al., 2008). The manipulation of values during maintenance operations therefore has the potential to cause discolouration.

While combined compliance rates in the England and Wales are very high, unacceptable chemical, aesthetic and biological water quality failures do occur, a proportion of which are thought to occur within DWDSs. Given this, and the desire to operate DWDSs in a more proactive manner, there is a need to better understand water quality within DWDSs, particularly with respect to the number of water quality anomalies that may be due to known or planned activities.

Analytical modelling of the activities and events that can affect water quality in DWDSs is impeded by the latter's structural heterogeneity and by the complexity of the physical, chemical and microbiological processes involved. Data-driven investigations into the causes of water quality fluctuations can potentially elucidate causal relationships without the need for computationally-intensive analysis.

Jaeger et al. (2002) suggested that data regarding water quality incidents and possible causal events/activities could be mined for cause–effect relationships. They stated that such data mining needs to take into account the hydraulics of DWDS as flow paths determine if and how cause and effect are linked. However, the methodology proposed by Jaeger et al. (2002) for determining the causes of water quality issues quantifies the separation of cause and effect in Euclidean (planar) space. This is thought to be inappropriate for determining if and how an event at one location in a labyrinthine DWDS has influenced water quality at another as hydraulic connectivity and network path distances are not considered. Okabe et al. (2006) argued that planar spatial analysis techniques are inappropriate for studying network-constrained data (such as water quality incidents or mains repair data) as they can produce many false-positive results. Even so, planar/Euclidean searches continue to be the common method for determining the causes of water quality issues because of their simplicity and ease of implementation. Others have used GIS systems and either visual assessment of data (e.g. Eng et al. (1999)) or planar spatial searches (e.g. Kistemann et al. (2001)) to explore the causes of DWDS water quality issues.

The methodology of Jaeger et al. (2002) was developed further in its incorporation into the *IMADSIG* software (Trepanier et al., 2006; Besner et al., 2007). With this tool, flows can be traced back from the location of a water quality incident to upstream boundaries such as treatment works. This software allows the set of pipes identified through such 'back-tracing' to be visualised along with the results of spatio-temporal queries of the various data model datasets. Human judgement is then applied to associate 'qualitative probabilities' with cause–effect linkages.

The results of analysing historical data from DWDSs in seven cities were mixed (Besner et al., 2007): no causes could be identified for 0–38% of coliform-positive samples, for 36–53% of the customer complaints and for 8–83% of the heterotrophic plate counts. Besner et al. considered the efficacy of the methodology to be limited by the number of unrecorded events and activities, to inaccurate spatio-temporal references, to the low sampling frequency and to hydraulic models not being sufficiently representative.

2. Aims and objectives

A scalable, largely automatable methodology called Distribution System Integrated Modelling (DSIM) has been developed for linking historical water quality issues to known events by network topologies and hydraulics. The objectives of the research were to:

1. Develop a method for combining water quality, asset, topological and hydraulic data to form an integrated network data model, giving consideration to the quantity, quality and value of the datasets used.
2. Devise a method for determining the *region of influence* hydraulically upstream of a given water quality issue. This is the set of pipes that lies upstream of a particular network location, given typical directions of flow. Possible causes of that issue can be found by spatio-temporally searching the issue's region of influence.
3. Apply the developed methodology to datasets that describe a particular distribution system to assess its efficacy.

The application of the methodology for identifying a probable cause for a reported discolouration complaint is illustrated in Fig. 1.

3. Developing and analysing an integrated data model

The developed methodology is presented in this section and section 4 features a case study that demonstrates its application.

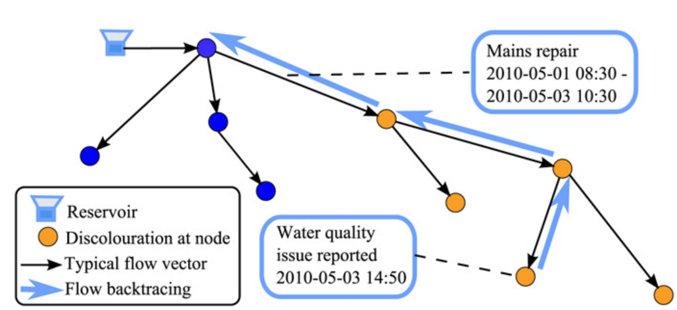

Fig. 1. Searching within an integrated network model for upstream causes of a historical water quality issue.