Author's response to:

Interactive comment on "Online data processing for proactive water distribution network operation" by J. Machell et al.

Anonymous Referee #1

Received and published: 9 September 2013

General comments:

The paper does not describe the development of new technologies, and no detailed description of the applied technologies. The scientific level of the paper is therefore limited.

The special interest in this paper is combining of existing techniques in one monitoring system, and implementing this to monitor a real water system, using real data. This makes it sufficiently interesting for publication.

The discussion section previews possible future applications / developments. As a result, this section is a bit hypothetical (all kind of possible applications, íf certain developments have taken place).

Specific comments:

1. Page 261, Title, word "operation". A water distribution network is a quite static system, and there is little to be operated. Why not use the word "monitoring" rather than "operation".

> The objective of this work was to demonstrate that distribution network flow and pressure data collected *via* monitoring can be automatically analysed and used to plan and manage certain network operations in a proactive, rather than a reactive, manner.
>
> "Operation" is the recognised word in use within the UK water industry for the day to day control and functioning of a distribution network. The use of "monitoring" would imply watching and/or recording information/data but not necessarily acting on it.

2. Page 263, line 24. Other papers also describe methods to identify optimal sensor locations (like Ami Preis, et al, 2011; James-A. Goulet, et al (2013)

> The references quoted will be added prior to publication. With hindsight, we would also like to suggest these additional references be added:
>
> Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., Peralta, A., (2011). Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. Control Engineering Practice, 19, 1157–1167

R. Perez, S. de las Heras, J. Aguilar, J. Pascual, A. Peralta. District management areas characterisation in water network based on clustering. Water Science Technology, 9 (5) (2009), pp. 591–600

3. Page 263, line 25 / 267, line 26: I think the typical situation in the UK is described.

Utilities in the Netherlands started earlier making robust sensor locations in the networks (some with permanent power and communication facilities), and typically the data quality is better. However, the number of sensors in the network is typically much lower

The description reflects the position within the collaborating UK water company at the time the work was carried out for the described case study. It was not intended to compare against other work in the field / other countries.

4. Page 269, line 8: Not all detection algorithms return binary classification. The method described by Romano / Kapelan / Savic (2012) generates a probability value between 0 and 1.

It is correct to say that some detection algorithms produce outputs that are not binary, such as the Romano et al. 2012 reference.

The system described in this paper delivers a fuzzy output score and a % confidence for the classification window – neither of which is binary. The outputs however must ultimately produce the binary classification of 'alarm' or 'non-alarm' depending on the cut-off criteria. We therefore recommend the following rewording to resolve this reviewers comment:

"Event detection algorithms work by obtaining data, performing some analysis and then returning outputs such as probabilities or fuzzy values. These are then processed into a binary classification i.e. generate an alarm or not."

5. Page 269, line 22: The trade off between false alarms and non-detection of smaller burst event occurs with any method and not only with flat-line thresholds, even with the method described in this paper (by setting the a% confidence value)

This is an issue for all event detection systems, but more significant for flat lines where the level has been set arbitrarily / without any of the positive aspects of data driven methodologies.

6. Page 270, line 25: The number of alerts from the flat line system is ridiculous (equals to 2.5 alerts per sensor per day). Obviously, the threshold values are chosen too low. This can't be considered as a good configured monitoring system.

It was not suggested it was a well configured system. The figures reflect a specific situation at the time the field work was completed, and is one good reason for promoting more proactive use of existing data and improved techniques to make alarm management less onerous. Flat line alarm thresholds are problematic in dynamic networks where flows, pressures and thresholds change continually.

The initial set up of the alarm thresholds was probably valid at the time. However, they were not re-assessed often enough to *keep* them valid through changes in network

operational characteristics and/or changes in demand; hence the large number of alarms generated.

7. Page 270, line 26: Classifying events as “abnormal” is somewhat questionable. I would almost say that any alert can be classified as “abnormal” (for there has to be some deviant value in the measurements, in order to generate an alert). Classifying such events is a bit subjective. And the method is evaluated to working OK, though an alert is raised in a situation without any collapse of the network.

When noise and spurious data is removed, and the data is then re-checked and smoothed, water distribution network flow and pressure values can fall within a number of repeatable patterns dependent upon the regime applied to dynamic element operation at any given time i.e. which source(s) to use, what pressures to maintain within various different parts of the interconnected networks, turn over time of service reservoir volumes etc. In this case “abnormal” refers to being different from planned events such as re-zoning by changing boundary valves, or changing a pumping schedule, and is simply one word that could have been used to denote this.

Abnormal changes that we might wish to know about, but that are not going to result in anything as catastrophic as a collapse of the network, also occur in the data. Indeed, these are the ones most likely to provide early warning, and time, for operators to proactively manage the network to minimise any potential reduction of service level effects of an event.

8. Page 275, line 17: Add reference to Mounce, Mounce and Boxall (2012)

This reference will be added prior to publication.