Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access
Drinking Water
Engineering and Science
Discussions

Real-Time Hydraulic Interval State Estimation for Water Transport Networks: a Case Study

Stelios G. Vrachimis, Demetrios G. Eliades, and Marios M. Polycarpou

KIOS Research Center for Intelligent Systems and Networks
Department of Electrical and Computer Engineering
University of Cyprus, Nicosia, Cyprus

Correspondence to: Stelios G. Vrachimis (vrachimis.stelios@ucy.ac.cy)

Abstract. Hydraulic state estimation in water distribution networks is the task of estimating water flows and pressures in the pipes and nodes of the network based on some sensor measurements. This requires a model of the network, as well as knowledge of demand outflow and tank water levels. Due to modeling and measurement uncertainty, standard state-estimation may result in inaccurate hydraulic estimates without any measure of the estimation error. This paper describes a methodology for generating hydraulic state bounding estimates based on interval bounds on the parametric and measurement uncertainties. The estimation error bounds provided by this method can be applied to estimate the unaccounted-for water in water distribution networks. As a case study, the method is applied to a transport network in Cyprus, using actual data in real-time.

1 Introduction

Hydraulic state estimation in Water Distribution Networks (WDN) is a challenging task due to the presence of modeling uncertainties, such as structural uncertainty introduced by skeletonization of the network, parameter uncertainty of pipe roughness coefficients and uncertainty in water demands. While this last uncertainty can be reduced by the use of real-time flow measurements, these measurements come with their own instrument uncertainties and noise (Hutton et al., 2014).

In standard state estimation techniques, statistical characterization of sensor measurement error is needed to give more weight to measurements originating from more accurate sensors. Using the weighted least squares method the nodal demands are adjusted to fit the constraints imposed by the measurements and produce the most probable state estimate (Davidson and Bouchart, 2006). Another approach is the Kalman Filter (KF) method which provides a solution for the network state based on the available measurements. The standard KF performs poorly in nonlinear looped WDN due to the use of a linearized system model (Kang et al., 2009). Overall, the above methods generate a point in state-space and are referred to as *point state estimation* (Andersen et al., 2001).

Most point state estimation methods assume a known statistical characterization of the measurement error. This could lead to significant estimation errors, especially in the case when pseudo-measurements are used, which are estimates determined from population densities and historical data. The use of pseudo-measurements may be necessary when there are not enough sensors to guarantee the observability of the network. In this case, no measure of the estimation error is available. Additionally,

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access Drinking Water Engineering and Science Discussions

in order for point state estimation methods to produce feasible solutions, model calibration is required *a priori* or during state estimation (Savic et al., 2009; Kang and Lansey, 2011).

An alternative approach for the representation of measurement and model parameter uncertainty is the use of bounds. In contrast to traditional point state estimation methods, the use of bounding uncertainty can provide upper and lower bounds on the state variables. This method is referred to as *interval state estimation*. In this work a hydraulic interval state estimation methodology is described and its use is demonstrated with a case study of a real transport network of a large water utility in Cyprus. A possible application of this method for estimating the unaccounted-for water in the network is also presented.

The use of measurement bounds for the representation of measurement uncertainty and their incorporation into the state estimation cost function was introduced by Bargiela and Hainsworth (1989). Interval state estimation was developed by Gabrys and Bargiela (1997) as the so-called set-bounded state estimation problem. An implicit state estimation technique for leakage detection for an idealized grid network under steady conditions was presented by Andersen and Powell (2000). A straightforward method for interval state estimation is the use of Monte-Carlo simulations, which under some assumptions converge to the true uncertainty bounds by randomly generating and evaluating a large number of parameter sets or realizations (Eliades et al., 2015).

In many applications, such as leakage detection and contamination detection, the derivation of a range of possible values for the state of the WDN provides useful information for event and fault detection methodologies. Hydraulic state bounds can be used to generate bounds on chlorine concentration in the water network or other chemicals in the water, by taking into consideration the uncertainty on decay rate (Vrachimis et al., 2015). When additionally this bounded estimate is generated in real-time, it helps to reduce the time of detecting water leakages and prevent catastrophic scenarios such as water contamination.

The paper is organized as follows: Section 2 formulates the problem of hydraulic state estimation and describes a methodology to solve this problem based on the Interval Hydraulic Interval State Estimator (IHISE) algorithm. In Section 3, a case study is presented in which this method is applied to a real transport network as one module of *AquaRisk*, a real-time cloud-based water distribution system monitoring platform. Finally, we discuss the application of this method for estimation of unaccounted-for water in the network.

2 Hydraulic Interval State Estimation

A water transport network is modeled using a directed graph, for which nodes represent water sources, junctions of pipes and water demand locations and the links represent pipes. Each pipe is indicated by the index j, where $j \in \{1,..,n_p\}$ and n_p is the number of pipes. These are characterized by pipe length, diameter and roughness coefficient, parameters which are generally assumed known. Pipe parameters are used to compute the Hazen-Williams (H-W) resistance coefficients r_j, which are in turn used to formulate the energy conservation equations of a water network (Boulos et al., 2006).

Modeling uncertainty in a WDN is considered in this work to arise from insufficient knowledge of pipe parameters. The uncertain parameters are represented using intervals, with the actual value of the parameter being within a corresponding interval. For notational convenience, the parameters representing intervals will be written in bold font. Any uncertainty parameters in

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access Drinking Water Engineering and Science Discussions

pipe j will be included in $\mathbf{r_j} \in \left[r_j^l,\, r_j^u\right]$. The interval parameter $\mathbf{r_j}$ is the uncertain H-W coefficient for pipe j, with r_j^l and r_j^u being the lower and upper bound of each coefficient respectively.

Nodes are indicated by the index i, where $i \in \{1,..,n_u\}$ and n_u is the number of nodes with unknown head, thus excluding the nodes that represent water sources. In this work we consider water transport networks in which sensors measure all the water demands at nodes, which typically, are the inflows of District Metered Areas (DMAs). Measurements arrive at a fixed time interval from sensors that may not be accurate, and each measurement is associated with a certain measurement error. The uncertainty of each measurement is given as a percent error of the measurement and it is modeled as an interval with the measurement being the median value of the interval. Measured water demand at node i is then given by the interval $\mathbf{q_{ext,i}} = \left[q_{ext,i}^l,\, q_{ext,i}^u\right]$, where $q_{ext,i}^l$ is the lower bound on water demand and $q_{ext,i}^u$ is the upper bound.

The unknown state vector of the WDN is denoted by $x = [q^\top\ h^\top]^\top \in \mathbb{R}^n$, where $h \in \mathbb{R}^{n_u}$ are the unknown heads at nodes, $q \in \mathbb{R}^{n_p}$ are the water flows in pipes and $n = n_p + n_u$. These are computed by formulating the conservation of energy and mass equations, as formulated by Todini and Pilati (1987). The matrix formulation for a general looped water distribution system, which also includes the uncertain parameters and variables as intervals (written in bold), is given by:

$$\begin{bmatrix} \mathbf{A_{11}}(\mathbf{q}) & A_{12} \\ A_{21} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{h} \end{bmatrix} = \begin{bmatrix} -h_{ext} \\ \mathbf{q_{ext}} \end{bmatrix}, \tag{1}$$

where $\mathbf{A_{11}}(\mathbf{q}) \in \mathbb{R}^{n_p \times n_p}$ is a diagonal matrix containing the nonlinear terms $\mathbf{r_j}|\mathbf{q_j}|^{\nu-1}$, $\nu = 1.852$ is a constant associated with the H-W coefficient and $h_{ext} \in \mathbb{R}^{n_p}$ is a vector that contains the known heads in each equation. For simplicity, we assume that measurements of the tank levels are available, thus h_{ext} is known.

Equation (1) represents a system of nonlinear equations, which include interval parameters and it is referred to in the literature as a Nonlinear Interval Parametric (NIP) problem (Kolev, 2014). The objective is to find the smallest interval state vector $\boldsymbol{x} = [\boldsymbol{q}^\top\ \boldsymbol{h}^\top]^\top$ that contains all the solutions of this system of equations for every value contained in the interval parameters. To solve the NIP problem given in (1), an algorithmic technique named Iterative Hydraulic Interval State Estimation (IHISE) was developed by the authors. The IHISE method comprises of five steps: 1) Find initial bounds on the state vector x; 2) Use interval linearization to remove nonlinear terms from (1) and transform them into a system of Linear Interval Parametric (LIP) equations; 3) Formulate a Linear Program (LP) using the system of LIP equations; 4) Solve the LIP problem; 5) Iteratively tighten the bounds on x and approximate the solution of the NIP problem.

Figure 1 illustrates how this technique is implemented in a real-time framework. At discrete time instant k, the measurements from the sensors in the network are received, which include the water outflow $q_{ext}(k)$ and the water level in tanks. The measured tank level at each time instant is used to calculate the known head vector $h_{ext}(k)$ of the network. Since these equations only depend on the current time instant k, the discrete time notation is omitted. The uncertainty of these measurements is inserted by converting them into intervals with the measurement as the mean value. The hydraulic equations of (1) are then formulated using the new measurements. Modeling uncertainty is represented by including the interval parameters $\boldsymbol{r_j}$ into the equations.

The first step of the IHISE algorithm is to impose initial bounds on the state vector x. The initial bounds should be an outer interval solution of (1) (Kolev, 2014). An outer interval solution includes all the point solutions of (1), but it is not the smallest

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access
Drinking Water Engineering and Science Discussions

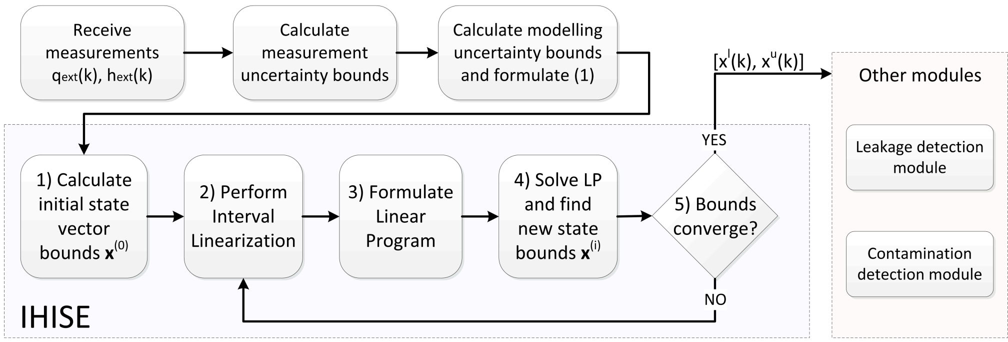

Figure 1. A diagram illustrating how the IHISE algorithm works in a real-time framework

possible interval. Bounds on the unknown head vector h can be chosen using physical properties of the network such as the minimum head of each node and the maximum head that pumps and water sources can add to the network. After finding an initial interval for the unknown heads $\boldsymbol{h}^{(0)}$, the special structure of (1) can be exploited and, using interval arithmetic, the initial bounds for the flows $\boldsymbol{q}^{(0)}$ can be calculated.

In the second step, the nonlinear terms present in (1) need to be linearized in order for the system of equations to be transformed into a LIP problem and solved (Kolev, 2004a). This is achieved using interval linearization (Kolev, 2004b). Given a range of values for the state $\boldsymbol{x}$ in which interval linearization will be performed, each of the nonlinear functions is enclosed between two lines and an interval term represents the linearization uncertainty. In the third step, the LIP equations are formulated into a LP with constraints. The interval terms in these equations are transformed into constraints of the LP and a suitable cost function ensures that the solution of this problem will give either the minimum or maximum of a certain state.

To get an interval solution of the whole state vector x, in step four, the LP formulated is solved for all the states by changing the cost function. At the end of this step, an interval solution for the linearized system of equations is derived. The new bounds on state vector x are then checked for convergence in step five. The criterion for convergence is the change of bounds to be smaller than a specified small number ϵ. The algorithm then gives the final state vector bounds calculated as the result. Otherwise, the new bounds calculated are used as initial bounds and the algorithm re-iterates from step two.

3 Case study: Limassol, Cyprus

This study uses data from a real water transport sub-network in Limassol, which is the second largest city in Cyprus. An illustrative diagram of the network is shown in Figure 2. The network has a tree structure and it comprises of 16 demand nodes, 16 links which represent pipes and one water tank. Flow sensors (F) are installed at demand nodes which are entrances to DMAs, and a water level sensor (L) is installed in the tank. Sensor measurements arrive at fixed five-minute intervals. The tank's water input originates from four water sources, of which three are water dams and one is a desalination unit. The water inflow q_0 coming from these sources is measured with a flow sensor. The water outflow q_1 of the tank is not directly measured.

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access Drinking Water Engineering and Science Discussions

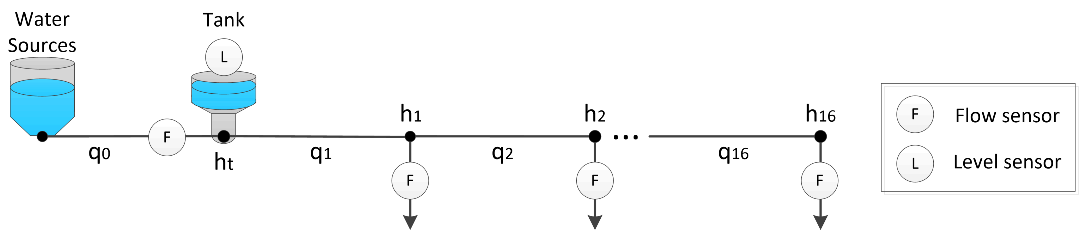

Figure 2. Illustrative diagram of the Limassol water transport sub-network of this case study.

3.1 Real-time Hydraulic Interval State Estimation

The implementation of this case study in real-time is based on the AquaRisk platform, which was first designed at the KIOS Research Center for Intelligent Systems and Networks at the University of Cyprus, and later further developed by a start-up enterprise. AquaRisk is a cloud-based platform for real-time monitoring of WDN against hydraulic and quality events. A model of the transport network was created as an EPANET input file. Using the AquaRisk platform, one can select the dates with available sensor data and request a state estimation. The available measurements from demand nodes and the level of the tank are then retrieved and a data validation process takes place in order to replace certain data errors (e.g. missing data, outliers).

Sensor measurements have an uncertainty which is defined by the installed sensor's specifications. The measurements given by the flow sensors are within $\pm 2\%$ of the actual flow at those locations. Modeling uncertainty is also present in the form of pipe parameter uncertainty. For this case study we assumed a total uncertainty of $\pm 2\%$ on the Hazen-Williams coefficient which is calculated using pipe parameters. Using the IHISE algorithm, bounds on water flows and pressures in the network are generated using the flow measurements at demand nodes and the tank level measurements, by taking into account measurement and modeling uncertainty. For illustration purposes, flow and pressure estimates using a real-time EPANET-based state estimator were also generated. The state estimates for a selected pipe and node, accompanied by its corresponding uncertainty bounds generated by the IHISE algorithm, are shown in Figure 3.

3.2 Calculating unaccounted-for water using bounds on state estimates

A challenge that the Water Board of Limassol was facing, was an unexplained variation in mass balance calculations, i.e. the difference between the volume of water entering and exiting the transport network. The problem was not trivial to solve because there is no sensor measuring the tank outflow q_1 and all sensors have measurement errors. As a result, it was not clear whether the differences were because of uncertainty in measurements or some other reason. Since the IHISE algorithm generates bounds on the flows in the network based on measurement and modeling uncertainties, it was used to investigate if there was mass imbalance in the network.

Considering the uncertainties defined in the previous section, the IHISE algorithm generated interval estimates for the tank outflow, indicated here by $\boldsymbol{q}_1^a(k)$. The tank inflow measurement, indicated by q_0, is not used by the IHISE algorithm and it can be utilized to check the mass balance in the network. To do this, an additional estimate of the tank outflow, indicated by $\boldsymbol{q}_1^b(k)$,

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access Drinking Water Engineering and Science Discussions

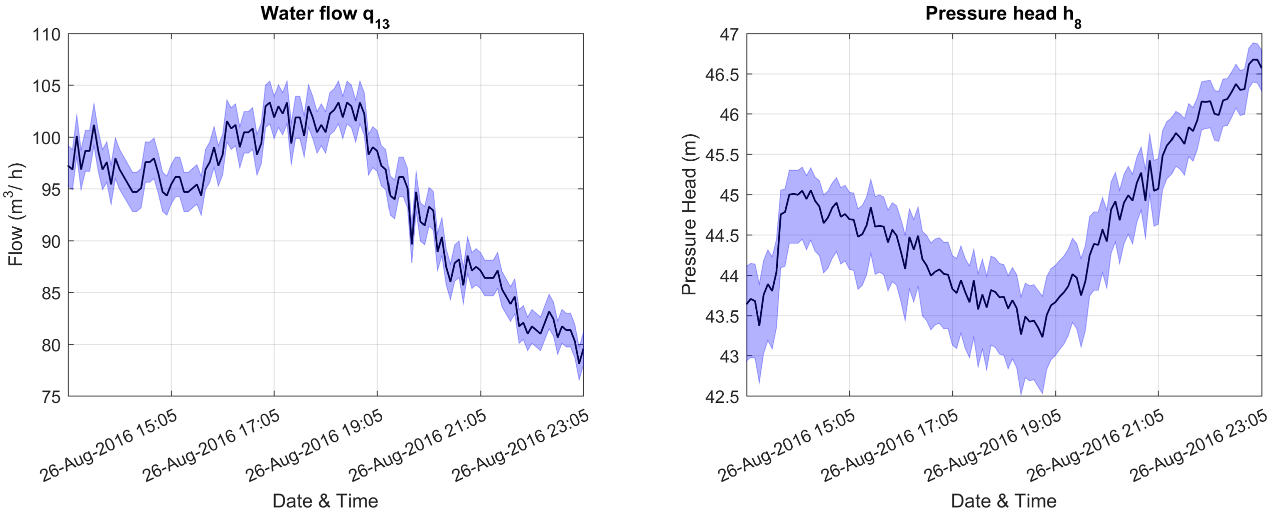

Figure 3. State estimate (black line) and bounds on this estimate using the IHISE algorithm (blue area) for the water flow in a selected pipe (left) and the head at a selected node (right).

was calculated using the inflow measurement $\boldsymbol{q}_0$ and the tank water level measurement $\boldsymbol{h}_t$. Because these are also uncertain, they are represented by intervals. As defined by the sensor specifications, the tank inflow sensor uncertainty is set at $\pm 2\%$ and the level sensor uncertainty is set at $\pm 0.5\%$ of the measurement value. The additional tank outflow estimate $\boldsymbol{q}_1^b(k)$ is calculated using the following equations which model the water tank and are solved using interval arithmetic:

$$\begin{aligned} \boldsymbol{q}_1^b(k) &= \boldsymbol{q}_0(k) - (\alpha_t/\Delta t)\ \Delta\boldsymbol{h}_t(k) \\ \Delta\boldsymbol{h}_t(k) &= \boldsymbol{h}_t(k) - \boldsymbol{h}_t(k-1), \end{aligned} \tag{2}$$

where α_t is the base area of the tank and Δt is the measurement time step.

The comparison of the two sets of bounds, $\boldsymbol{q}_1^a$ and $\boldsymbol{q}_1^b$, for a certain period of time is shown in Figure 4 (left). What we observe from this comparison is that the two sets of bounds do not overlap at all time steps, even though uncertainties are considered in the calculations. This indicates that there is unaccounted-for water due to background leakages or unmeasured demand locations in the water transport network, or due to some other metering error in the tank inflow.

In order to get an estimate of the unaccounted-for water, it is assumed that this is constant for a certain period of time and is denoted by the constant θ. To calculate this constant, the following function $\theta = F(\boldsymbol{q}_1^a, \boldsymbol{q}_1^b)$ is used, where $F(.)$ is a function which takes as input the two sets of bounds $\boldsymbol{q}_1^a$ and $\boldsymbol{q}_1^b$ for a given period of time and maximizes the overlapping area between them. It does this by calculating the optimal constant θ which must be added to the IHISE bounds $\boldsymbol{q}_1^a$, assuming that the difference is due to a leakage in the transport network. Figure 4 (right) illustrates the correction when θ is added to the IHISE bounds $\boldsymbol{q}_1^a$, resulting in a maximum overlapping area between the two sets of bounds.

The advantage of this approach in comparison to just using point state estimates, is that the use of bounds takes into account the effect of measurement and modeling uncertainties, thus enabling the network operator to distinguish between possible

Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2017-18
Manuscript under review for journal Drink. Water Eng. Sci.
Discussion started: 6 June 2017

Open Access
Drinking Water
Engineering and Science
Discussions

problems in the network and uncertainty. The results of this approach could not confirm whether the difference in unaccounted-for water was due to background leakages in the water transport network, or due to some metering error in the tank inflow. When these findings were presented to the Water Board of Limassol for further investigation, it was eventually validated that there was a metering error at the tank inflow.

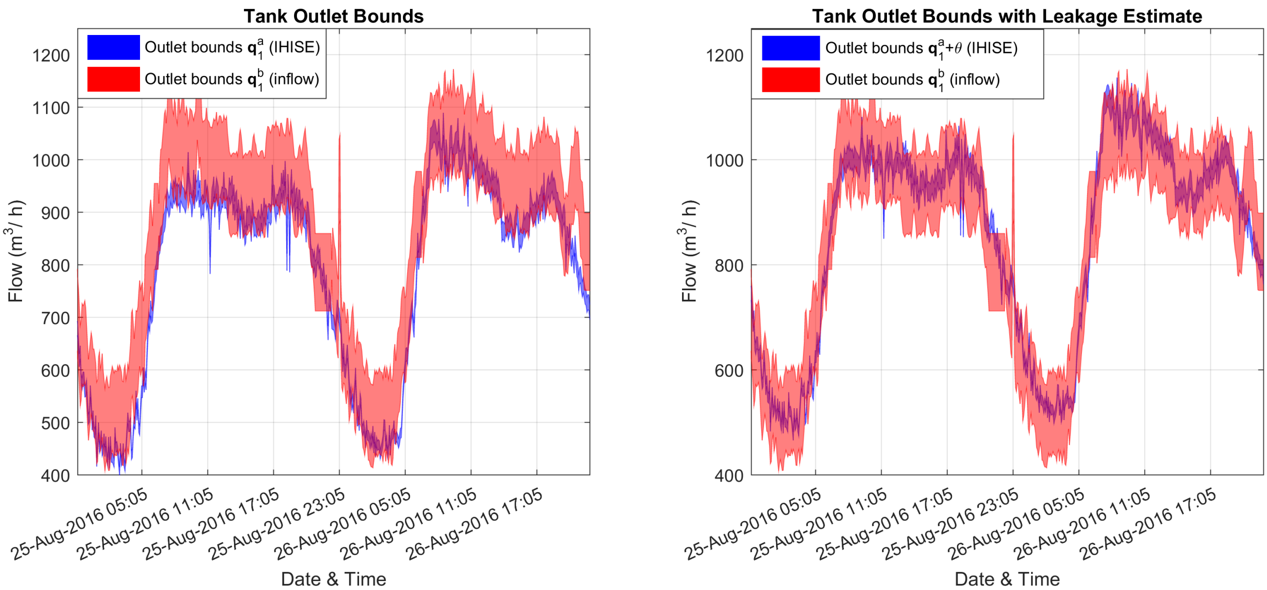

Figure 4. Left: Tank outflow bounds generated using: 1) the IHISE algorithm and 2) the tank inflow and level. Right:The same bounds with the leakage estimate added to the IHISE generated bounds

4 Conclusions

In this work we described a methodology for real-time hydraulic interval state estimation, to monitor water transport networks. Using real-time uncertain measurements from a real transport network, the proposed Iterative Hydraulic Interval State Estimation algorithm generates bounds on hydraulic states of the network, by taking into account the measurement uncertainty and modeling uncertainty in the form of uncertain pipe parameters. The applicability of this methodology was demonstrated by using it to estimate the unaccounted-for water in the network.

Extension of this work, will use the generated bounds to apply fault-detection methods that detect and localize leakages in the network. Additionally, the bounds on hydraulic states of the network will be used to generate bounds on water quality states, since the dynamics of hydraulic and quality states of a water network are interconnected.

Acknowledgements. This research is funded by the European Research Council under the ERC Advanced Grant ERC-2011- AdG-291508 (FAULT-ADAPTIVE).

Open Access Drinking Water Engineering and Science Discussions

References

Andersen, J. H. and Powell, R. S.: Implicit state-estimation technique for water network monitoring, Urban Water Journal, 2, 123–130, 2000.

Andersen, J. H., Powell, R. S., and Marsh, J. F.: Constrained state estimation with applications in water distribution network monitoring, International Journal of Systems Science, 32, 807–816, 2001.

Bargiela, A. and Hainsworth, G. D.: Pressure and flow uncertainty in water systems, Water Resources Planning and Management, 115, 212–229, 1989.

Boulos, P. F., Lansey, K. E., and Karney, B. W.: Comprehensive water distribution systems analysis handbook for engineers and planners, American Water Works Association, 2006.

Davidson, J. W. and Bouchart, F. J.: Adjusting nodal demands in SCADA constrained real-time water distribution network models, Journal of Hydraulic Engineering, 132, 102–110, 2006.

Eliades, D. G., Stavrou, D., Vrachimis, S. G., Panayiotou, C. G., and Polycarpou, M. M.: Contamination event detection using multi-level thresholds, in: Proceedings of 13th Computing and Control for the Water Industry Conference, CCWI 2015, vol. 119, pp. 1429–1438, 2015.

Gabrys, B. and Bargiela, A.: Integrated neural based system for state estimation and confidence limit analysis in water networks, in: Proceedings of European Simulation Symposium ESS'96, pp. 398–402, 1997.

Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L., and Savic, D. A.: Dealing with uncertainty in water distribution system models : A framework for real-time modeling and data assimilation, Water Resources Planning and Management, 140, 169–183, 2014.

Kang, D. and Lansey, K.: Demand and roughness estimation in water distribution systems, Water Resources Planning and Management, 137, 20–30, 2011.

Kang, D. S., Pasha, M. F., and Lansey, K.: Approximate methods for uncertainty analysis of water distribution systems, Urban Water Journal, 6, 233–249, 2009.

Kolev, L. V.: A method for outer interval solution of parametrized systems of linear interval equations, Reliable Computing, 10, 227–239, 2004a.

Kolev, L. V.: An improved interval linearization for solving nonlinear problems, Numerical Algorithms, 37, 213–224, 2004b.

Kolev, L. V.: Parameterized solution of linear interval parametric systems, Applied Mathematics and Computation, 246, 229–246, 2014.

Savic, D. A., Kapelan, Z. S., and Jonkergouw, P. M.: Quo vadis water distribution model calibration?, Urban Water Journal, 6, 3–22, 2009.

Todini, E. and Pilati, S.: A gradient algorithm for the analysis of pipe networks, in: Proceedings of International Conference on Computer Applications for Water Supply and Distribution, pp. 1–20, 1987.

Vrachimis, S. G., Eliades, D. G., and Polycarpou, M. M.: The backtracking uncertainty bounding algorithm for chlorine sensor fault detection, in: Proceedings of 13th Computing and Control for the Water Industry Conference, CCWI 2015, vol. 119, pp. 613–622, 2015.