

XVIII International Conference on Water Distribution Systems Analysis, WDSA2016

State Estimation for Water Distribution Networks in the Presence of Control Devices with Switching Behavior

Francesco Fusco^a and Ernesto Arandia^{a*}

^a*IBM Research, Damastown Industrial Estate, Mulhuddart, Dublin 15, Dublin, Ireland*

Abstract

In process monitoring and control, state estimation is the fundamental tool for processing redundant and noise-corrupted measurements in order to provide reliable estimates of the state of a system. In the context of water distribution networks (WDNs), state estimation has been proposed as the core technology which can enable various applications ranging from real-time monitoring and control to anomaly diagnosis, such as leak detection and localization. Measurements are typically available from sparse and often scarce telemetry sensors, such as flow at the inlet of a district metered area (DMA) and pressure at some nodes, or from utility estimates, for example prior estimates of the nodal demands. The problem consists of using the available measurements to reconstruct an estimate of the state variables and is solved iteratively by minimizing the weighted least squares (WLS) of the differences between the measurements and model predictions, typically with gradient methods. WDN state estimation in the presence of control devices, such as pressure reducing valves, remains an open problem due to the complexity in modeling efficiently their switching behavior. Control elements prevent from obtaining an explicit function of the measurements with respect to the state variables for all possible switching statuses. In this paper, an extension to traditional state estimation methods is proposed, which only requires a minor modification of existing WLS solvers based on gradient methods. Based on residual analysis, conditions are given in order to verify correct convergence at the end of a state estimation or to identify changes in connectivity due to opening/closing of control elements before proceeding to a new run. The method does not require including explicit binary variables to model the state of control elements, which would require complex heuristic-based solvers and would present scalability challenges for large networks with many such elements. Results on a real-world test case with two PRVs are reported to demonstrate the effectiveness and of the proposed method.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the organizing committee of the XVIII International Conference on Water Distribution Systems

* Corresponding author. Tel.: +353 83 163 5403
E-mail address: ernestoa@ie.ibm.com

Keywords: State estimation; pressure control valve; water distribution systems.

1. Introduction

State estimation is concerned with providing a complete representation of the internal condition or status of a system at a given instant of time and is applicable to virtually all areas of engineering and science [1]. In process monitoring and control, state estimation is a powerful tool to process collected measurements and to filter redundant and noise-corrupted ones in order to provide reliable estimates of state variables [2].

In the context of water distribution networks (WDNs), state estimation has been proposed as the core technology which can enable various applications ranging from real-time monitoring and control to anomaly diagnosis, such as leak detection and localization. The state estimation methods that have been reported for online monitoring of WDNs emphasize applications such as water demand estimation and forecasting [3, 4, 5]. These methods predict hydraulic behavior of the system based on prior estimation of demands and further correct the predictions using supervisory control and data acquisition (SCADA) measurements of variables such as flows and pressures. Use of state estimation for detection and localization of leaks was also widely studied, see for example [6, 7, 8].

The state estimation problem consists of using available measurements to reconstruct an estimate of the system hydraulics, i.e., nodal demands and flows throughout the WDN. The measurements are typically available from sparse telemetry sensors, such as flow at the inlet of a district metered area (DMA) or at some pipes and pressure at some nodes, or from utility estimates, for example prior estimates of the nodal demands. As detailed in [6], by appropriately choosing the state of the system as the set of nodal demands and flows at some links, an explicit relation between the measurements and the state can be found. The problem can then be efficiently solved, for example by minimizing the weighted least squares (WLS) difference between measurements and model predictions using gradient methods.

As pointed out by [8], WDS state estimation in the presence of control devices, such as pressure reducing valves, remains an open problem due to the complexity in modeling the discontinuities produced when the devices change status. In particular, measurements cannot typically be expressed as an explicit function of the system hydraulics, and accurately modeling the switching behavior would require the introduction of binary variables. As a consequence, the state estimation problem would increase in complexity often requiring complex heuristic methods with challenges in scalability for large networks with many control devices.

In this paper, a simple extension of existing state estimation solvers based on gradient methods, as introduced in [6], is proposed. The method allows to successfully deal with the presence of multiple control elements with minor impact on the computational complexity. Based on classical residual analysis, criteria for successful convergence of the state estimator are defined, such that correct estimation of the status of control elements can be verified. The same conditions allow to identify, eventually, a change in connectivity due to opening/closing of a control elements, so that a new state estimation run can be found. Depending on the number and the status of control elements, one to a number of simple state estimation runs is required to converge to the correct solution, thus keeping the computational cost reasonable even in the most complex cases.

Nomenclature

q	vector of nodal demands	$\Phi(x)$	state-measurement mapping function
C	connectivity matrix	$\Psi(x)$	hydraulic constraints
Q	vector of link flows	ε	measurement noise
ΔH	vector of head losses	W	weighting matrix
h	vector of nodal heads	Σ	measurement error covariance matrix
h_F	vector of fixed heads	x	state vector
Λ	connectivity matrix of spanning tree	y	vector of observed hydraulic quantities
V	connectivity matrix of co-tree	z	measurement vector
(\cdot)	head loss equations in spanning tree	F	Jacobian matrix
$\Gamma_V(\cdot)$	head loss equations in co-tree		

After reviewing the details of the state estimation model in section 2, the extension of the solution algorithm is described in section 3. The effectiveness of the methodology is tested in a real-world WDN with two pressure reducing valves, and results are reported in section 4. Conclusions are given in section 5.

2. State Estimation Model

In order to specify the state estimation problem, the hydraulic model of the WDN needs to be expressed in the following explicit form:

$$y = f(x) + \varepsilon, \quad (1)$$

where $y \in \mathbb{R}^p$ is the vector of observed hydraulic quantities (pressures, flows, demands), $x \in \mathbb{R}^q$ is the hidden state vector, which is the minimum set of variables that is sufficient to specify the complete hydraulic solution for the WDN, and ε represents model uncertainty (sensor noise, imprecise equations), typically as Gaussian white noise with zero mean and given covariance matrix Σ . A convenient expression of the hydraulic model of a WDN in the form (1) was derived in [6, 9] and it is detailed as follows.

The equilibrium between demand and flows can be written as:

$$q = CQ, \quad (2)$$

where $q \in \mathbb{R}^n$ are the nodal demands and $Q \in \mathbb{R}^m$ are the link flows, for a WDN with n junctions and m links, specified by the connectivity matrix $C \in \mathbb{R}^{n \times m}$. Similarly, the head-loss across each link is related to the nodal heads through the following:

$$\Delta H = -C^T h + h_F, \quad (3)$$

where $h \in \mathbb{R}^n$ are the nodal heads, $\Delta H \in \mathbb{R}^m$ are the head-losses and $h_F \in \mathbb{R}^m$ is a vector representing the elevation of the fixed-head nodes, essentially representing the boundary conditions of the WDN. The link between (2) and (3) is given by the hydraulic head-loss relations that exist for each link. In the case of pipes, the head-loss equation at link j can be written in the following form:

$$\Delta H_j = r_j |Q_j|^{k_j} \text{sign}(Q_j), \quad (4)$$

where a number of empirical laws exist expressing the parameters r_j and k_j with respect to physical properties of the pipe (length, diameter, roughness) and, in some cases, to the dynamic flow conditions. Typical expressions are the Hazen-Williams or Darcy-Weisbach equations [10]. In the case of control elements such as valves, the head-loss equation can be more complicated and typically exhibits a switching behavior depending on the hydraulic conditions of the network. For a pressure-reducing valve (PRV), for example:

$$\Delta H_j = \begin{cases} h_i - K_i & \text{if } K_i < h_i < h_{i'} \quad (\text{active}) \\ r_j Q_j^{k_1} & \text{if } h_i \leq K_i < h_{i'} \quad (\text{open}), \\ \infty & \text{if } h_{i'} \leq h_i \quad (\text{closed}) \end{cases} \quad (5)$$

where K_i is a constant valve setting parameter, $h_{i'}$ is the pressure of the node downstream of node i (positive flow according to the valve orientation), $r_j = K_{hw} l_j / (\mu_j^{k_1} / d_j^{k_2})$ is the Hazen-Williams resistance factor assuming a fully open valve behaves as a short smooth pipe, $l_j = 2d_j$ is the length of the short pipe (twice the diameter), K_{hw} is the Hazen-Williams constant, μ_j is the roughness coefficient and d_j is the internal diameter of the valve. For notational convenience, in the following, we express the head-loss equation at any link as a generic nonlinear function $\Delta H = \Gamma(Q_j, h_i, h_{i'})$.

In order to obtain a model in the form in (1), suitable for state estimation, a state vector needs to be defined such that all possible hydraulic measurements can be expressed in an explicit form. As proposed in [6], a convenient choice for the state is based on the separation of the connectivity matrix in spanning tree and co-tree (i.e. the set of closure links):

$$C \triangleq [\Lambda : V]. \quad (6)$$

In (6), $\Lambda \in \mathbb{R}^{n \times n}$ is the connectivity matrix of a spanning tree of the network graph, with the convenient property of being invertible or $\det(\Lambda) \neq 0$, and $V \in \mathbb{R}^{n \times (m-n)}$ is the connectivity matrix of the closure links. Based on the separation in (6), the hydraulic equations given in (2) and (3) can be rewritten such that heads, flows and demands can be explicitly written as follows:

$$q = \Lambda Q_\Lambda + V Q_V \quad (7)$$

$$Q_\Lambda = \Gamma_\Lambda^{-1}(-\Lambda^T h + h_\Lambda) \quad (8)$$

$$Q_V = \Gamma_V^{-1}(-V^T h + h_V) \quad (9)$$

$$h = (-\Lambda^T)^{-1} \Gamma_\Lambda (\Lambda^{-1} q - \Lambda^{-1} V Q_V) + (\Lambda^T)^{-1} h_\Lambda, \quad (10)$$

where the vector of head-loss equations is separated in $\Gamma_\Lambda(\cdot)$ and $\Gamma_V(\cdot)$ to denote links in the spanning tree and closure links, respectively, as well as the fixed-head elevations in $h_F \triangleq [h_\Lambda^T, h_V^T]^T$.

Based on (7) through to (10), by choosing the state vector to comprise the nodal demands and the flow in the closure links, namely $x \triangleq [q^T, Q_V^T]^T$, of dimensionality $p = m$, one can immediately obtain an explicit state estimation model of the desired form as in (1). In particular, any combination of hydraulic sensors measuring demand, flows or pressures, defining a measurement vector z , can always be related to x with an explicit non-linear equation, i.e.,

$$z = \Phi(x) + \varepsilon_z, \quad (11)$$

with an additional noise term being added to model sensor noise or model uncertainty. The measurement equation in (11), however, is not sufficient to specify the full hydraulics of the WDN and needs to be complemented by a set of $m - n$ additional constraints making the connection between the tree and co-tree of the WDN, which are mathematical abstractions necessary to obtain an explicit model. In particular, as detailed in [6], the following hydraulic consistency constraints can be derived from (10) and (9):

$$\Gamma_V(x) = V(\Lambda^T)^{-1}[\Gamma(\Lambda^{-1}q - \Lambda^{-1}VQ_V) - h_\Lambda] + h_V \triangleq \Psi(x) = 0, \quad (12)$$

which represents the fact that the sum of head-losses across each loop created by a closure link is zero.

The measurement equation in (11) and the hydraulic constraints in (12) can be ultimately combined to obtain the desired state estimation model in (1), as follows:

$$y \triangleq \begin{bmatrix} z \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \Phi(x) \\ \Psi(x) \end{bmatrix} + \begin{bmatrix} \varepsilon_z \\ \varepsilon_0 \end{bmatrix} \triangleq f(x) + \varepsilon. \quad (13)$$

In (13) a noise term for the constraints, ε_0 , was also introduced for modeling convenience. In practice one can assign this noise a very low uncertainty, or once can use to relax the constraint to accommodate for some modeling uncertainties.

Further details for the derivation of the model in (13) can be found in [6, 9]. Note that this is a steady-state model and all relationships are to be thought as applying at each point in time independently, which is a valid assumption for typical sampling rates of analysis of WDN, in the range of a few minutes to one or two hours.

3. State Estimation Solution Algorithm

Given the model in (13), a state estimate can then be calculated from the following WLS problem:

$$\hat{x} = \underset{x}{\operatorname{argmin}} J(x) = [y - f(x)]^T W [z - f(x)], \quad (14)$$

via the following iterative scheme based on the Gauss-Newton method:

$$\hat{x}_{k+1} = \hat{x}_k + (F_k^T W F_k)^{-1} F_k^T W [y - f(\hat{x}_k)], \quad (15)$$

where F_k is the Jacobian of $f(x)$ calculated around \hat{x}_k . The weighting matrix is typically set to the inverse of the measurement error covariance matrix $W = \Sigma^{-1} = \text{blockdiag}(\Sigma_z^{-1}, \Sigma_0^{-1})$. Line-search methods can be applied to the iterative scheme in (15) in order to improve convergence rate in strongly non-linear regions of the objective function, by avoiding erroneous over-estimation of the iteration step [11].

3.1. Extended Solution Algorithm with Control Elements

In the presence of control elements such as the ones described in section 2, the algorithm update equation in (15) cannot be immediately calculated, since there is no explicit form to evaluate $f(\hat{x}_k)$ from the state estimate at step k . Given that the state vector is defined as set of nodal demands, q , and closure links Q_V , as described in section 2, the steps to evaluate the measurement equation $f(x)$, that is any set of hydraulic quantities (demands, flows, heads) are the following:

1. Demands are known.
2. Calculate flows in spanning tree, Q_A , using (8). Flows in co-tree are known.
3. Calculate heads using (10), which requires evaluating head-loss equations across the spanning tree, $\Gamma_A(\cdot)$.

In the presence of flow or pressure control elements, steps 2 and 3 need to be followed by a check to validate that the assumed element status is compatible with the head-loss equation being used. In the case of PRVs, for example, the step 3 is evaluated by assuming open status so that a head-loss can be calculated from the flows. After step 3, the head at all nodes is available and the conditions in (5) can then be evaluated in order to determine the appropriate status of the PRVs. In particular:

4. For each PRV i , if the head at the upstream node, h_i , is found to be higher than the setting K_i , the status is active and the corresponding head-loss is updated to $\Delta H_i = h_i - K_i$.

Note that the case of closed status is not checked. This case needs extra care, because a closed valve would require the whole model to be updated to reflect the change in connectivity. A refactoring of the spanning tree and co-tree separation is also required and, therefore, also a refactoring of the choice for the state variable. Similarly, a PRV assumed in closed status is considered as an open link for the whole state estimation run.

Before changing the status of the element in such a way to affect the graph, we propose to keep the elements as closed or active/open throughout the whole state estimation run, until termination. At this point, residual analysis can be applied in order to determine whether the state estimator was able to converge successfully. In particular, the state estimation residuals are defined as:

$$r \triangleq y - f(\hat{x}), \quad (16)$$

and it can be shown how they should belong to a normal distribution with covariance matrix given by [7]:

$$\Omega = \Sigma - F(F^T \Sigma^{-1} F)^{-1} F^T. \quad (17)$$

A statistical test can then be designed to determine whether the residuals are within the expected distribution with a certain confidence. For example, a threshold of $3\sqrt{\omega_i}$ applied to the magnitude of the residual r_i , where ω_i is the i -th diagonal element of Ω , would flag an erroneous value with 99.73% confidence. If some of the residuals denote erroneous convergence, it is likely that some of the model assumptions are incorrect. By validating the following flow condition at the suspicious valves:

$$Q_i \geq 0, \quad (18)$$

one can get indications of whether a valve is being kept open or active erroneously. A negative flow would in fact indicate that the condition $h'_i \leq h_i$ applies in (5) and that the valve should therefore be closed. A new state estimation run can then be performed until no anomalous residuals are found. For the opposite case of PRV assumed closed, the estimated flow will obviously be zero and condition (18) cannot be used. However, the residuals (16) can still give indication of erroneous convergence, and where pressure sensors are available at sensitive locations the corresponding residuals will give clear indication that the assumptions about a close-by PRV may be wrong.

4. Results

A real-world example, shown in Fig. 1, is used in order to show how the algorithm behaves in the presence of two PRVs. The network is a reduced version of the WDS of a medium size Italian city. The network is first mentioned by [12] and consists of 68 junctions, 3 reservoirs, and 99 pipes. Two PRVs are positioned as indicated in the Figure.

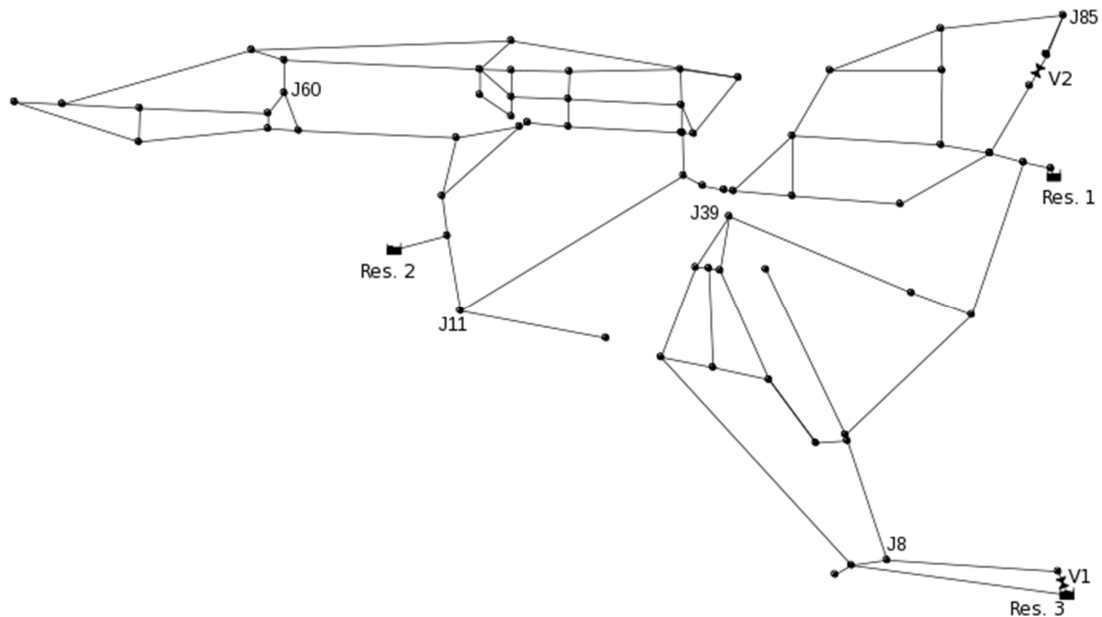


Fig. 1. Pescara WDN. Sensor measurements are available at junctions J8, J11, J39, J60, J85. The PRVs are indicated as V1 and V2.

It is assumed that the available sensor measurements are the flow at the inlet pipes, P19, P54, P90, P93, with 0.1 L/s standard error and that pressure loggers are available at junctions J8, J11, J39, J60, J85, with 1m standard error. The demands are assumed to be known at all junctions with large 20% standard error. The state estimation is initialized by assuming the two PRVs to be active or fully open.

Table 1. State estimation results with active valves. V1: setting 45 m, status *active*; V2: setting 40 m, status *active*. Successfully solved after first state estimation step.

Sensor	Unit	Actual	Measured	Estimated #1
Head J8	[m]	38.54	38.63	37.74
Head J11	[m]	48.74	48.32	48.95
Head J60	[m]	32.28	30.63	30.90
Head J39	[m]	33.43	34.05	33.56
Head J85	[m]	41.19	41.78	41.16
Flow V2	[L/s]	8.34	-	8.48

A first experiment is setup where the settings of the two PRVs are such they are both in active status. Table 1 summarizes the results of the state estimation by comparing the true value, the measured value (true value plus noise) and the estimate value at the five pressure loggers. Clearly, the first state estimation run successfully converges and gives pressure estimates that are closer to the true values with respect to the noisy measurements. The table also shows the value of the estimated flow at the one of the PRVs, V10, which is in line with the true value.

Table 2. State estimation results with one valve closed. V1: setting 45 m, status *active*; V2: setting 15 m, status *closed*. Status of V2 erroneously initialized to *open*. Successfully solved after second state estimation step.

Sensor	Unit	Actual	Measured	Estimated #1	Estimated #2
Head J8	[m]	38.55	39.15	37.03	37.03
Head J11	[m]	48.70	49.27	48.80	48.80
Head J60	[m]	32.10	32.78	32.71	32.71
Head J39	[m]	33.45	34.45	33.67	33.66
Head J85	[m]	18.55	19.54	17.14	18.74
Flow V2	[L/s]	0.00	-	-1.61	0.00

In a second experiment, the setting of V2 is changed to 15 m, so that its status switches from active to closed. After the first state estimation run, as shown in Table 2, one of the pressure estimates significantly deviates from the true value and the corresponding residual, calculated from (16), is larger than 3 times its standard deviation indicating a convergence failure with 99.73% confidence. Figure 2 shows the value of the residuals, with the $3\sqrt{\omega}$ interval band, for at the five pressure loggers. Interestingly the node at which we observe non-convergence is J85, which is downstream of V2 as shown in Fig. 1. Also, the estimated value of the flow at V2 is significantly below zero which violates the condition in (18). Based on these indications, the status of V2 is changed to closed and the second run of the state estimation successfully converges, as indicated in Table 2, with the pressure residuals all being within the desired confidence, as shown in Fig. 2.

In a final experiment, the setting of V2 is restored to 40 m, so that its status switches back from closed to active. Table 3 summarizes the results obtained when the first state estimation run erroneously assumes the status of V2 to be closed. Two of the pressure estimates significantly deviates from the true value and the corresponding residual is larger than 3 times its standard deviation, as also detailed in Fig. 3. In particular, a quite significant deviation (greater than 7 m) is observed at the node J85, which is downstream of V2. This is strong indication that a wrong assumption about V2 status is being considered and after correction to open/active the second state estimation run correctly converges close to the true hydraulic conditions, as detailed in Table 3, with residuals within the confidence bounds, as indicated in Fig. 3.

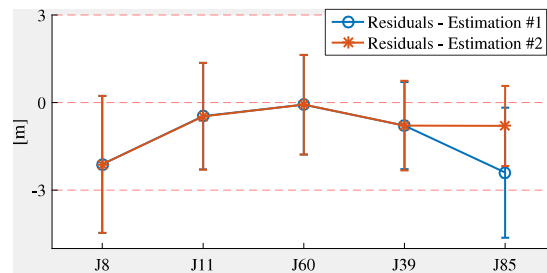


Fig. 2. Residuals at pressure sensors after the two state estimation steps. V1: setting 45 m, status *active*; V2: setting 15 m, status *closed*. Status of V2 erroneously initialized to *open*.

5. Conclusion

An extension of existing state estimation solvers based on the gradient methods introduced in [6], is proposed above. The method successfully handles the existence of multiple control devices with a minor increase in the computational complexity of the original approach. The correct estimation of the status of control devices is verified using criteria for successful convergence based on classical residual analysis. The criteria also allows to detect a change in connectivity caused by opening or closing of control elements, so that a new state estimation run can be performed to refine the solution.

The computational performance depends on the number and the status of the control elements. Depending on the initial assumptions about valves status, one to a number of runs is required in order to achieve convergence. Each state estimation run can be solved in very efficient time and multiple runs could be easily parallelized to consider multiple statues assumptions where there is uncertainty. For the example considered in section 4, of a mid-sized real-world WDN (about 100 state variables) a single state estimation run was solved in a fraction of a second with a medium-range dual core machine, using a Matlab scripting environment and with limited code optimization.

Table 3. State estimation results with active valves. V1: setting 45 m, status *active*; V2: setting 40 m, status *active*. Status of V2 erroneously initialized to *closed*. Successfully solved after second state estimation step.

Sensor	Unit	Actual	Measured	Estimated #1	Estimated #2
Head J8	[m]	38.54	37.94	35.69	36.98
Head J11	[m]	48.74	49.23	48.69	48.67
Head J60	[m]	32.28	33.02	32.77	32.87
Head J39	[m]	33.43	35.15	31.65	34.27
Head J85	[m]	41.19	40.99	33.12	41.24
Flow V2	[L/s]	8.34	-	0.00	6.76

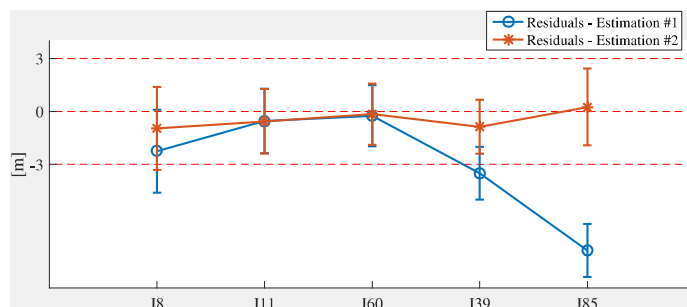


Fig. 3. Residuals at pressure sensors after two state estimation steps. V1: setting 45 m, status *active*; V2: setting 40 m, status *active*. Status of V2 erroneously initialized to *closed*.

As objective of future work, it will be interesting to understand how the approach would scale to an arbitrary number and type of control elements. From the methodological point of view, further research is required in order to understand the impact of wrong assumptions about control elements on the pressure and flow residuals, such to more formally identify the root-case element in case non-convergence is detected. Finally, robustness of the method to other types of anomalies needs to be investigated.

References

- [1] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, New Jersey: John Wiley & Sons, 2006.
- [2] S. Pajić, "Power system state estimation and contingency constrained optimal powerflow - a numerically robust implementation," Ph.D. dissertation, Worcester Polytechnic Institute, 2007.
- [3] J. W. Davidson and F. J. C. Bouchart, "Adjusting nodal demands in SCADA constrained real-time water distribution network models," *J. Hydraul. Eng.*, vol. 132, pp. 102–110, 2006.
- [4] F. Shang, J. G. Uber, B. G. van Bloemen Waanders, D. Boccelli, and R. Janke, "Real time water demand estimation in water distribution system," in *8th Annual water Distribution Systems Analysis Symposium*, 2006.
- [5] A. Preis, A. Whittle, A. Ostfeld, and L. Perelman, "Efficient hydraulic state estimation technique using reduced models of urban water networks," *J. Water Res. Plann. and Manag.*, vol. 137, no. 4, pp. 343–351, 2011.
- [6] J. H. Andersen and R. S. Powell, "Implicit state-estimation technique for water network monitoring," *Urban Water*, vol. 2, no. 2, pp. 123–130, 2000.
- [7] F. Fusco, B. Eck, and S. McKenna, "Bad data analysis with sparse sensors for leak localisation in water distribution networks," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, Aug 2014, pp. 3642–3647.
- [8] D. Jung and K. Lansey, "Water distribution system burst detection using a nonlinear kalman filter," *J. Water Res. Plann. and Manage.*, vol. 141, no. 5, p. 04014070, 2015.
- [9] F. Fusco and A. Ba, "Fault diagnosis of water distribution networks based on state-estimation and hypothesis testing," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, 2012, pp. 886–892.
- [10] Q. B. Travis and L. W. Mays, "Relationship between Hazen-William and Colebrook-White roughness values," *Journal of Hydraulic Engineering*, vol. 133, no. 11, pp. 1270–1273, 2007.
- [11] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 2006.
- [12] C. Bragalli and C. D'Ambrosio, "On the optimal design of water distribution networks: a practical minlp approach," *Optimization and Engineering*, vol. 13, no. 2, pp. 219–246, 2012.