Identifiability analysis for pressure sensors positioning

Valeria Puleo[1], Silvia Tarantino[2], Gabriele Freni[3], Goffredo La Loggia[4]
[1,2,4] Department of Civil, Environmental, Aeronautic, and Chemical Engineering, University of Palermo, Viale delle Scienze, Ed. 8, 90100 Palermo, Italy.
[3] School of Engineering and Architecture, University of Enna "Kore", Cittadella Universitaria, 94100 Enna, Italy
[1]*valeria.puleo@unipa.it*

ABSTRACT

The identifiability analysis is investigated as sampling design method aimed to the leakage detection in looped water distribution networks. The preliminary ranking of the candidate nodes for the pressure sensors positioning is performed by running several hydraulic simulations and calculating sensitivity functions. The reduced subset of nodes and their sensitivities are then used to perform the identifiability analysis by calculating the collinearity index which provides the maximum number of sensors and their location into the network. The index selects the nodes according to their sensitivities to several leakages scenarios, simulated in EPANET by changing the emitter coefficient of the leakages function both with a One-At-Time and Monte Carlo approach. The collinearity index also identifies the subset of the pressure monitoring nodes with the lowest correlation (redundancy) between the measurements. The method is applied to the benchmark network Apulian.

Keywords: Leakages, Identifiability, Sensor positioning

1 INTRODUCTION

Water preservation is one of the great challenges of modern cities in which the efficient use of the resource is still difficult to achieve. Water utilities can contribute to water savings by means of an efficient water losses management. Several methodologies and technologies are available to improve the efficiency in term of leakages reduction. Numerous studies have been carried out to improve the knowledge about leakages, by means of experimental [1] and numerical modelling [2], as well as to detect and reduce them [3]. The detection methods usually require the definition of a District Metered Area (DMA) in which flow meters and pressure sensors allow to verify the presence of leakages. Afterwards, different techniques are adopted to locate the leaks and then it is possible to replace or repair the damaged pipes.

Due to the extension of the water distribution networks (WDNs), the detection step can take long time during which some parts of the system are put out of service. As consequence, the water utilities deal with the complaining customers while they lose a great amount of water and spend money. Therefore, model-based techniques, including WDNs hydraulic modelling coupled with several algorithms for leakages detection, can allow more effective and less costly water losses control strategies.

In this context, the paper investigates the application of the identifiability analysis for the pressure sensor location in looped water distribution networks aiming to leakages detection. The method determines the nodes which can be candidates as monitoring points and it selects the more reliable sensors combination by means of the collinearity index. The study is applied to the benchmark network Apulian.

1.1 Background

The optimal positioning of monitoring sensors, or simply sampling design, has been previously addressed with respect to several purposes [4] such as monitoring for baseline system characteristics, detection of contamination events or leaks, compliance for maintenance of system performance and calibration of network models. Several sampling design techniques have been developed for model calibration [4, 5], for contamination events [6] and leakages detections [7]. These studies demonstrated that the sensors location is crucial to provide good and reliable information for the specific purpose.

Farley et al. [7] proposed an approach aimed to place pressure loggers in the most leak-sensitive locations by means of an evaluation matrix for detecting new pipe bursts. The methodology was validated by using a real-size network in which leak flows have been generated by the hydrant openings. The same authors [8] later combined the Jacobian sensitivity matrix together with a genetic search approach for the leakages localization. Zheng and Yuan [9] developed ad two-stage method for which the sensor locations have been optimized to cover the maximum number of leakage events. Namely, the leak-events are randomly generated by means of the Monte Carlo method and simulated into a hydraulic solver. Simulation results are elaborated in an evaluation matrix populated with pressure changes at each node for each leak-event. With respect to Farley et al. [7] multiple leakage nodes are simulated and they also introduced the pressure logger accuracy to convert the evaluation matrix to a binary database. Such database is input to a competent genetic algorithm to find the optimal sensor locations. Genetic algorithms have been also applied in Perez et al. [10]. Christodoulou et al. [11] proposed a greedy-search heuristic based on entropy theory for the sensors placement such as acoustic, pressure, or flow sensors acting on pipe segments. In Steffelbauer et al. [12] the sensor placement was solved through a non-binarized leak sensitivity matrix with a projection-based leak isolation approach. Moreover, the hydraulic model parameters uncertainty effects on the measurements have been considered in the analysis. A differential evolution algorithm performed the leakage localization. The results showed that the most sensitive nodes to the leakages can also be to the demand fluctuations. Blesa et al. [13] defined a robustness percentage index to evaluate the sensor placement strategies based on the fault sensitivity matrix for different leak magnitudes and operating scenarios (e.g. changes in demands). A clustering analysis reduced the number of the candidate sensors. A significative variation on the leak localization was observed when the operating points changed. The sensitivity analysis features were also applied in Bort et al. [14] which compared three alternatives for sensors positioning: empirical considerations, sensitivity analysis results and correlation analysis of the pressure measurements nodes. Casillas et al. [15] solved the optimal positioning of pressure sensors and leakages detection by means of the Leak Signature Space method coupled with genetic algorithms or particle swarm optimization techniques. Leakages have been modelled as an extra demand in the hydraulic solver.

Hence, the sampling design studies applied to the leakages detection have been recently addressed in literature and they reveal to be open to further developments. This paper can contribute to the model-based methodologies for sensor positioning aimed to the water losses control strategies.

2 METHODS

Leakages position can be deduced by measuring pressure in some points (nodes) of the water distribution network through the development of model-based techniques based on hydraulic simulation models. The effectiveness of such procedure relies on the good knowledge of the network topology, of the water demand and the pipes roughness. Several methodologies have been developed to provide leakages detection and hydraulic model calibration at the same time.

Therefore, the pressure sensors positioning has a great importance to have reliable and robust model responses, even if it depends on the specific analysis purpose.

Herein, the proposed methodology aims to select the pressure monitoring nodes for leakages detection by coupling the water distribution network hydraulic simulation model with the identifiability analysis. The nodes selection is done among those which are more sensitive with respect to different leakages positions and uncorrelated from each other to avoid redundant information. To sake of simplicity, in this study, both the water demands and the pipes roughness are assumed known. Then, only the parameters describing the leakages in the water distribution network model are modified during the simulation runs.

2.1 The identifiability analysis

The identifiability analysis is a preliminary phase in the model analysis for parametric estimation. For a given model structure and a defined configuration of the input and output data, it is possible to find out the unknown model parameters in the ideal case for which model and output variables are not affected by errors. The identifiability analysis generally deals with two problems: the model structure selection and the parameters identification [16]. Physically based considerations often impose the model structure. In several applications, such as in large environmental simulation models, the parameters identification is very important. In these cases, the parameter evaluation cannot be univocally determinate starting from the available measured data [17], rather the objective becomes finding physically reasonable values of the parameters able to describe adequately the model. In non-linear complex model, there are several reasonable values, then the identifiability analysis finally aims to find out more information about these parameters. In Freni et al. [18] this approach was applied on an integrated urban drainage system model to evaluate the number of parameters which can be reliably calibrated with few measurement data. Clearly, in system characterized by reduced complexity, the analysis of sensitivity functions should be sufficient as identifiability criterion.

In this paper, the identifiability analysis is applied to evaluate the nodes where the pressures (output variables) are more sensible to leakages (model input parameters) variation, in order to arrange pressure sensors for the leak detection. A water distribution network model can be described by a general set of equation $y = f(\vartheta)$ where y is the vector of the *n* output variables and ϑ the vector of the *m* independent model parameters. Sensitivity functions can be defined to have information about the raw dependency of the output with model parameters (eq. 1) and to find the relative impact of different parameters on the output variables of the model (eq. 2):

$$s_{ij} = \frac{\Delta\vartheta_j}{ys_i}\frac{\partial y_i}{\partial \vartheta_j} \qquad (1)$$

$$\tilde{s}_{ij} = \frac{s_{ij}}{\sqrt{\sum_{k=1}^{n} s_{kj}^2}} \qquad (2)$$

where $\Delta\vartheta_j$ is the variation range of the *j*th parameter; ys_i is the reference value of the output variable y_i; s_{kj} are the sensitivities of the *n* modelling outputs to the variation of the same parameter ϑ_j . The sensitivity functions can be allocated in the *n*x*m* matrices, respectively, *S* and $\tilde{S}$. Each entry in the sensitivity matrix considers the variations of pressures between the no-leak scenario and the leak-scenarios.

The identifiability analysis was carried out through the collinearity index by adapting the formulation proposed in Freni et al. [18]. The analysis starts with a sensitivity ranking of the pressure monitoring nodes by averaging the sensitivities of the modelling output y_i to the variations of the m parameters (eq. 3).

$$\bar{s}_i = \frac{1}{m}\sum_{j=1}^{m}\left|s_{ij}\right| \tag{3}$$

In the present study, sensitivity indices have been calculated by means of both the One-At-Time (OAT) and the Monte Carlo Simulation (MCS) runs. The OAT approach envisages the variation of a parameter by maintaining constant the others, then the local impact of one input parameter variations on the model outputs can be obtained. The Monte Carlo simulations are obtained by varying randomly all model parameters, the global impact of one parameter variations on the model outputs can be evaluated while other parameters change contemporarily.

The analysis of the sensitivity index $\bar{s}_j$ allows to reduce the number of nodes to be investigated for the pressure sensors positioning. The nodes with sensitivity lower than a defined threshold is considered non-identifiable. The choice of such threshold should be carefully done to avoid deleting too many or too few nodes [19].

The identifiability analysis is performed on the remaining nodes by involving gradually a larger number of these. First, only one node is chosen as candidate for sensor positioning according to the sensitivity ranking. Then, this one is combined with any other in the network model and the collinearity index is calculated as follows:

$$\gamma = \frac{1}{\sqrt{\min\left(EV[\tilde{S}\,\tilde{S}^T]\right)}} \tag{4}$$

The couple of nodes providing the lower value of collinearity index is considered identifiable. The analysis continues by adding other nodes and selecting the one providing the lower increment in the collinearity index until the increment is higher than a user-defined threshold or all nodes are added to the subset.

Finally, the analysis provides the number and the list of the pressure monitoring nodes for the specific water distribution network model. The method can be applied by running single or extended period simulation (EPS). In this last case, the hourly sensitivity indices can be averaged prior to evaluate the collinearity index.

2.2 The case study

The method presented above was tested on the benchmark Apulian network [20]. It is a small-sized network constituted of 23 nodes and 34 pipes. The network is fed by a reservoir with fixed hydraulic head, there are no tanks, valves or pumping stations. The leakages are modeled as *new* nodes, placed halfway in the pipes except for the pipe connected to the reservoir. The leak-nodes elevations are computed as average of the elevations of the pipe connecting nodes. Therefore, the simulated leakage-nodes are 33. The network scheme is depicted in Figure 1. The 24-hours EPSs have been carried out by means of the EPANET 2.0 [21] Dynamic-Link Library in MATLAB [22] and leakages have been modelled with the emitter function (eq. 5) which describes the relationship between the leakages flow rate q_{leak} and the pressure P. Only the emitter coefficient α is varied while the exponent β is set to 0.5.

$$q_{leak} = \alpha P^{\beta} \tag{5}$$

Different values of α are calculated through the eq. 5 for the OAT approach by fixing q_{leak} as a rate of the network inlet flow and assuming the hydrostatic pressure over each node. The resulting emitter coefficients are reported in Table 1. The simulations, as mentioned above, were performed by running a leakage, for each emitter node, one at time. The resulting OAT simulations have been 33 for each test. Whereas the Monte Carlo analysis was carried out by randomly varying α in a range between zero (no-leak) and a maximum value (equal to Test 3), 1000 simulations have been performed. The eligible sensor locations are limited to the original network nodes. The maximum number of sensors are identified according to the collinearity index values.

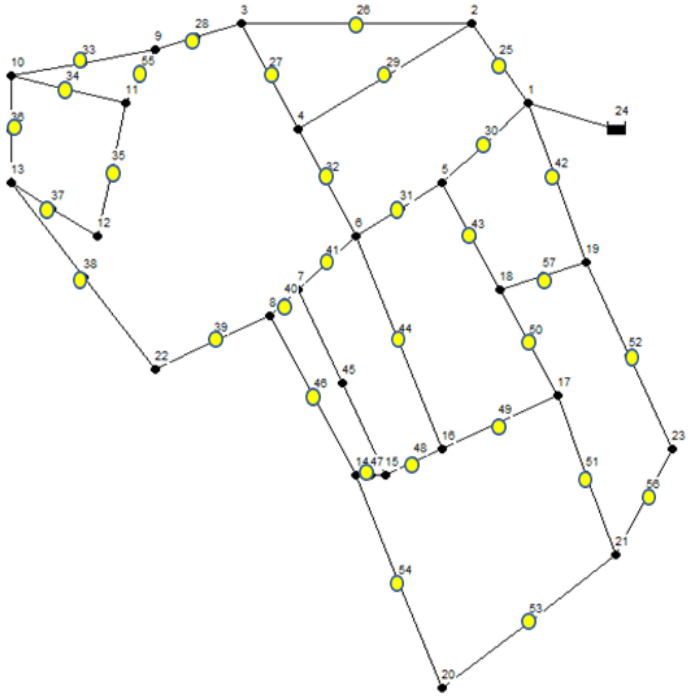

Figure 1. The Apulian network with emitter nodes (in yellow)

Table 1. The tested emitter coefficients for the OAT and MCS analysis

	Test 1 (OAT)	Test 2 (OAT)	Test 3 (OAT)	Test 4 (MCS)
Emitter coefficient α (l/(s m^{0,5})	0.3295	0.4943	0.6590	0 ÷ 0.6590
Leakages maximum flow (l/s)	1.71	2.56	3.42	3.42

3 RESULTS AND DISCUSSION

The number of the eligible pressure monitoring nodes is reduced by selecting the mean of the average sensitivity index $\bar{s}$ as threshold. The resulting sensor locations (Nodes ID) are reported in Table 2. Such subset is used to calculate the collinearity index by changing the number of sensors installed both for the OAT and MCS analysis. In Figure 2 the collinearity index grows as the number of sensors increases. The index value has a greater increment from four to five sensors. Test 2 and 3 show similar results (for sake of clarity, in Figure 2, they are reported with the same marker). Test 1 has some differences especially when more than five sensors are considered. Test 4 (MCS) provides values of the collinearity index greater than the previous ones. The obtained results allow some considerations:

- as the collinearity index increases, also the correlation between measurements grows;
- the selection of four sensors for Apulian network is sufficient to guarantee the most sensitive and less correlated points of measurements;
- the OAT analysis may be influenced by the selection of the parameter values, especially if the derived leakages flow is very low respect to the user demand (Test 1);
- the MCS analysis (Test 4) demonstrates that uncorrelated measurements are not easy to achieve in system characterized by high redundancy such as in the looped water networks.

The corresponding sensor monitoring points are reported in Table 3. The OAT tests (Test 1, 2 and 3) show the same selection except for the 4-sensors scenario in Test 1 where the node 16 is replaced by the nearby node 15. The MCS analysis provides a different selection for the 3-sensors scenario by preferring the node 16 to node 4, the former has a higher average sensitivity respect to the latter. Test 2, 3 and 4 show the same solution in the case of 4-sensors scenario. These results, compared with those presented in Bort et al. [14], which applied a correlation analysis, differ only for the node 13 in place of the nearby node 12.

Table 2. The sensor nodes ranking with respect to the average sensitivity $\bar{s}$ for the OAT simulations

Nodes ID	**23**	**12**	**13**	**16**	**21**	**15**	**10**
Average sensitivity $\bar{s}$	*1.415*	*1.401*	*1.307*	*1.288*	*1.272*	*1.241*	*1.232*
Nodes ID	**20**	**17**	**4**	**22**	**11**	**14**	**9**
Average sensitivity $\bar{s}$	*1.132*	*1.051*	*1.038*	*1.014*	*1.010*	*0.998*	*0.970*

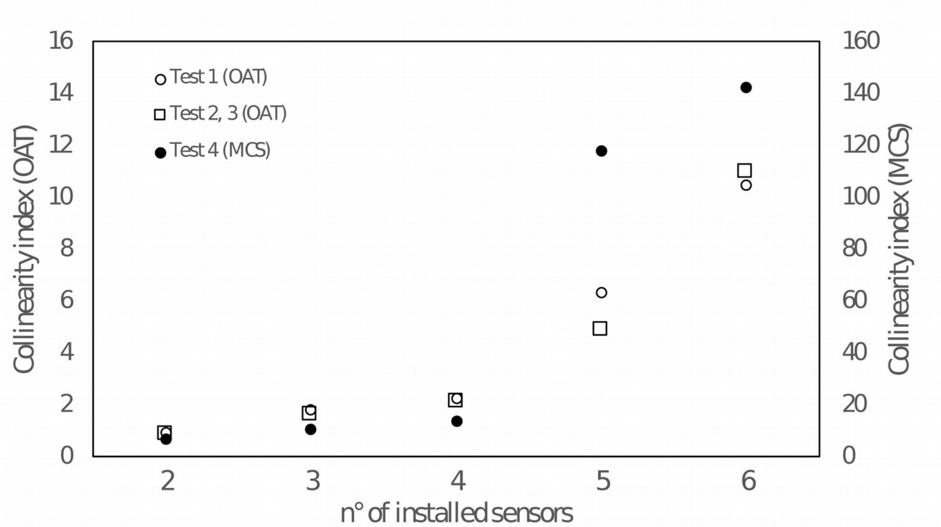

Figure 2. The collinearity index values vs number of installed sensors

Table 3. Sensor placement using the identifiability analysis

n° of installed sensors	Nodes ID			
	Test 1	**Test 2**	**Test 3**	**Test 4**
2	12, 23	12, 23	12, 23	12, 23
3	4, 12, 23	4, 12, 23	4, 12, 23	12, 16, 23
4	4, 12, 15, 23	4, 12, 16, 23	4, 12, 16, 23	4, 12, 16, 23

4 CONCLUSIONS

The application of the identifiability analysis was investigated as sampling design method for the pressure monitoring sensor positioning in looped water distribution networks. The main objective of the analysis was the selection of the sensor location for leakages detection. The collinearity index was applied to identify the subset of eligible points which are more sensitive to the leakages and provide less correlated measurements. Test 2, 3 and 4 provided the same sensor combination in the case of four installed sensors which can be considered the maximum number for the Apulian network. The comparison with the literature references [14] has been encouraging. The MCS analysis has been most reliable respect to OAT, it considers the uncertainty linked to both magnitude and positioning of the leakages in the water distribution network. However, to apply the identifiability analysis to a real system the uncertainty linked to the nodes demand as well as the pipe roughness has to be considered. Future developments of this research will address these issues.

References

[1] M. De Marchis, C. M. Fontanazza, G. Freni, V. Notaro, and V. Puleo, "Experimental Evidence of Leaks in Elastic Pipes," Water Resources Management, vol. 30, no. 6, pp. 2005-2019, 2016.

[2] A. Cassa, J. van Zyl, and R. Laubscher, "A numerical investigation into the effect of pressure on holes and cracks in water supply pipes," Urban Water Journal, vol. 7, no. 2, pp. 109-120, 2010.

[3] A. O. Lambert, "International report: Water losses management and techniques," Water Science and Technology: Water Supply, vol. 2, pp. 1-20, 2002.

[4] C. A. Bush and J. G. Uber, "Sampling design methods for water distribution model calibration," Journal of Water Resources Planning and Management-Asce, vol. 124, no. 6, pp. 334-344, 1998.

[5] W. B. F. de Schaetzen, G. A. Walters, and D. A. Savic, "Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms," Urban Water, vol. 2, no. 2, pp. 141-152, 2000.

[6] A. Ostfeld and E. Salomons, "Optimal layout of early warning detection stations for water distribution systems security," Journal of Water Resources Planning and Management-Asce, vol. 130, no. 5, pp. 377-385, 2004.

[7] B. Farley, S. R. Mounce, and J. B. Boxall, "Field testing of an optimal sensor placement methodology for event detection in an urban water distribution network," Urban Water Journal, vol. 7, no. 6, pp. 345-356, 2010.

[8] B. Farley, S. R. Mounce, and J. B. Boxall, "Development and Field Validation of a Burst Localization Methodology," Journal of Water Resources Planning and Management, vol. 139, no. 6, pp. 604-613, 2013.

[9] Y. W. Zheng and S. Yuan, "Optimizing pressure logger placement for leakage detection and model calibration," presented at the WDSA 2012: 14th Water Distribution Systems Analysis Conference, 24-27 September 2012, Adelaide, South Australia, 2012.
[10] R. Perez, V. Puig, J. Pascual, J. Quevedo, E. Landeros, and A. Peralta, "Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks," Control Engineering Practice, vol. 19, no. 10, pp. 1157-1167, 2011.
[11] S. E. Christodoulou, A. Gagatsis, S. Xanthos, S. Kranioti, A. Agathokleous, and M. Fragiadakis, "Entropy-Based Sensor Placement Optimization for Waterloss Detection in Water Distribution Networks," Water Resources Management, vol. 27, no. 13, pp. 4443-4468, 2013.
[12] D. Steffelbauer, M. Neumayer, M. Günther, and D. Fuchs-Hanusch, "Sensor Placement and Leakage Localization Considering Demand Uncertainties," Procedia Engineering, vol. 89, pp. 1160-1167, 2014.
[13] J. Blesa, F. Nejjari, and R. Sarrate, "Robustness Analysis of Sensor Placement for Leak Detection and Location under Uncertain Operating Conditions," Procedia Engineering, vol. 89, pp. 1553-1560, 2014.
[14] C. M. G. Bort, M. Righetti, and P. Bertola, "Methodology for Leakage Isolation Using Pressure Sensitivity and Correlation Analysis in Water Distribution Systems," Procedia Engineering, vol. 89, pp. 1561-1568, 2014.
[15] M. V. Casillas, L. E. Garza-Castanon, and V. Puig, "Optimal Sensor Placement for Leak Location in Water Distribution Networks using Evolutionary Algorithms," Water, vol. 7, no. 11, pp. 6496-6515, 2015.
[16] R. Brun, P. Reichert, and H. R. Künsch, "Practical identifiability analysis of large environmental simulation models," Water Resources Research, vol. 37, no. 4, pp. 1015-1030, 2001.
[17] E. Walter and L. Pronzato, "On the identifiability and distinguishability of nonlinear parametric models," Mathematics and Computers in Simulation, vol. 42, no. 2, pp. 125-134, 1996.
[18] G. Freni, G. Mannina, and G. Viviani, "Identifiability analysis for receiving water body quality modelling," Environmental Modelling & Software, vol. 24, no. 1, pp. 54-62, 2009.
[19] G. Freni, G. Mannina, and G. Viviani, "Assessment of the integrated urban water quality model complexity through identifiability analysis," Water Research, vol. 45, no. 1, pp. 37-50, 2011.
[20] O. Giustolisi, D. Savic, and Z. Kapelan, "Pressure-driven demand and leakage simulation for water distribution networks,", Journal of Hydraulic Engineering-Asce, Article vol. 134, no. 5, pp. 626-635, May 2008.
[21] L. Rossman, "EPANET 2 Users Manual." Cincinnati: National Risk Management Research Laboratory, 2000.
[22] Mathworks, "Using Matlab," ed. Natik, MA: Mathworks Inc., 2011.