Calibration of parameters of water supply network model using genetic algorithm

Tomasz Boczar [1,*], *Norbert* Adamikiewicz[2], and *Włodzimierz* Stanisławski[3]

[1] Opole University of Technology, ul. Prószkowska 76, 45-758 Opole, Poland
[2] Miejskie Wodociągi i Kanalizacja w Kędzierzynie-Koźlu Sp. z o.o., ul. Filtrowa 14, 47-200 Kędzierzyn-Koźle, Poland
[3] State Higher Vocational School in Nysa, ul. Armii Krajowej 7, 48-300 Nysa, Poland

Abstract. Computer simulation models of water supply networks are commonly applied in the water industry. As part of the research works, results of which are presented in the paper, OFF-LINE and ON-LINE calibration of water supply network model parameters using two methods was carried out and compared. The network skeleton was developed in the Epanet software. For optimization two types of dependent variables were subjected: the pressure on the node and volume flow in the network section. The first calibration method regards to application of the genetic algorithm, which is a build in plugin - "*Epanet Calibrator*". The second method was related to the use of function *ga*, which is implemented in the MATLAB toolbox Genetic Algorithm and Direct Search. The possibilities of application of these algorithms to solve the issue of optimizing the parameters of the created model of water supply network in both cases: OFF-LINE and ON-LINE calibration was examined. An analysis of the effectiveness of the considered algorithms for different values of configuration parameters was performed. Based on the achieved results it was stated that application of the *ga* algorithm gives higher correlation of the calibrated values to the empirical data.

1 Introduction

Water supply networks are one of the main components of the technical infrastructure of cities and settlements. Due to water losses there is a necessity for operational cost reduction, which may be performed through continuing development of new supporting methods for monitoring and diagnostics of these facilities. Apart from the extensions and narrowing of the pipe cross section, sediment causing changes in surface roughness, etc., cracks, causing water leaks, are the most common damage to the water supply infrastructure. But only some of them are manifested by the outflow of water to the surface, the rest are invisible and difficult to locate [1,2]. The problem is growing in the areas of cities, where as a result of the exploitation of deposits, e.g. coal or calcium, located below the city, the occurrence of the failures is more common. This issue will have an impact also after the cessation of mining. Current methods of locating failures in water supply systems use monitoring devices. These methods are, however, less effective in case of increasingly common PVC and PE pipes [3,4,5]. For that reason, Authors performed research studies aimed at implementation of a device for signaling of occurrence and localization of a defect in the water supply network. The main component of the system for detecting and locating leaks in a water supply network is a computer model that simulates network operation. This paper presents the results of research works dealing with the calibration procedure of the proposed mathematical water supply network.

2 Methodology

The object under the study was the water supply network of Kedzierzyn-Kozle, Poland. It is build from the following elements: two pumping stations (water treatment plant and water treatment station), two zonal pumping stations located in Slawiecice and Cisowa, transmission pipelines with valves, junction valves and the water recipients. The system of pipelines consists of buses of a total length of ~153 km and of junction valves of total length of ~74 km. The network covers an area of 6392 x 12322 m (excluding wholesale customers in the neighboring communities). The city has an area of 123.4 km^2 where ~64000 residents live. Fig. 1 shows the general structure of the network that was created in the EPANET environment.

The water supply network simulation model before application was calibrated using empirical data, obtained during experimental measurements. Within carried out research works, the OFF-LINE parameters were adjusted using two methods. The first method was to use a genetic algorithm implemented in EPANET using the "EPANET Calibrator" plugin, while the second method involved the use of the MATLAB toolbox for Genetic Algorithm and Direct Search, using the *ga* function.

[*] Corresponding author: t.boczar@po.opole.pl

Both OFF-LINE and ON-LINE adjustment use the ready-to-use features available in the environment and the applicability of these algorithms to solve the optimization problem of the water supply model was investigated.

Fig. 1. Screenshot of the network created in the EPANET with the map in the background.

The "EPANET Calibrator" plug-in used during the OFF-LINE adjustment process was created by the "Resan Labs" team from Brazil. This solution makes it possible to correctly and repeatedly estimate the optimal parameters of model parameters. The calculation result is saved to a file that can then be further analyzed in the EPANET environment.

The use of the "Epanet Calibrator" plug-in requires the network to be split into separate regions (sectors). In each node or pipeline of a given sector, the value of the sought parameter is the same. In case of the roughness parameter, the areas in which the type and age of the infrastructure are similar should be selected. Fig. 2 shows the network divided into sectors.
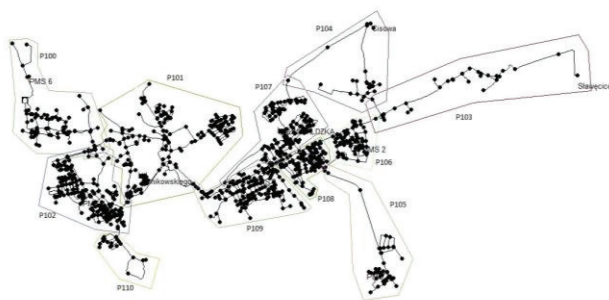
Fig. 2. Screenshot of the network created in the EPANET.

After establishing the sectors, reference values should be entered. The "EPANET Calibrator" optimizes model parameters for two types of dependent variables: water pressure in a given node and volume flow in a given network segment.

With a dedicated interface, one can enter the value of each parameter in successive time units manually, or one can import a time sequence directly from a CSV file. After the optimization process is started, the algorithm will modify the model parameters so that the resulting dependent variables correspond to the reference values. As part of the carried out works values measured on the object that were entered using CSV files were used.

In the next step, one determines which model parameters will be optimized. The EPANET Calibrator allows one to choose from three options:

- *roughness* of the pipe walls,
- the amount of water *leakage*, which is determined by the emitter coefficient,
- *roughness* and *leakage* at the same time. In the considered case, this option was applied.

The algorithm used in the plug-in seeks to minimize the error between measured (reference) values and the simulation obtained by both of the following values:

- network pressure values in a strictly defined network nodes, which are: p_{ZUW}, p_{SUW}, p_{PMS1}, p_{PMS2}, p_{PMS3}, p_{PMS4}, p_{PM5}, p_{PMS6} and suction pressure at two pumping stations: p_{CIS}, p_{SLA},
- volumetric flow in strictly defined pipes, which are: v_{ZUW}, v_{SUW}.

The EPANET Calibrator allows one to set values for four selected genetic algorithm parameters. The drawback of this plug-in is that there is no information on the progress of the calculation except for the successive iteration error.

Parameters of the genetic algorithm applied in the EPANET Calibrator were defined as follows:

- population size = 200,
- number of generations = 500,
- probability of mutation = 0.01,
- probability of crossing = 0.9,
- the optimized parameter = *roughness*,
- dependent variable = *pressure* (network pressure).

3 Calibration results

For the mentioned settings, the mean square error (RMSE) values in all ten sectors were calculated and the obtained results are presented in Table 1 as compared to data before optimization. Further, exemplary graphics representations of the results of the algorithm are presented in Figs. 3 and 4.
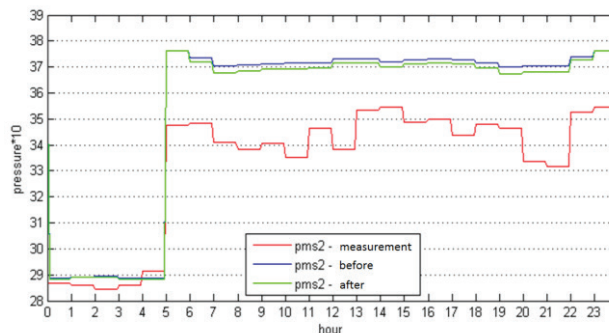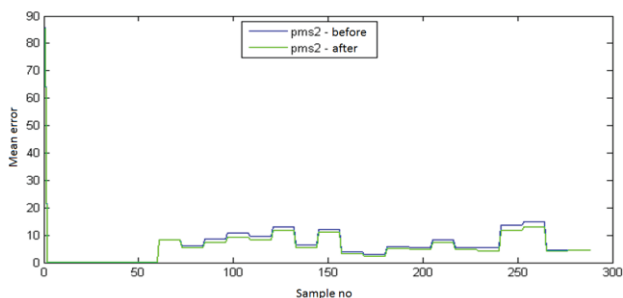

Fig. 3. Summary of genetic algorithm for PMS 2- pressure in measurement point no 2: the measured values (red), simulation result before optimization (blue), result of optimization (green).

In the next step, performance analysis of the algorithm for different values of the configuration parameters was performed. The obtained profit was compared by increasing the size of initial population or

number of generations. Fig. 5 shows the evolution of the genetic algorithm for an initial population of 200 individuals and 500 generations.

Fig. 6 shows the opposite situation: 500 individuals and 200 generations. By analyzing the curves in Figs. 5 and 6, it can be stated that a greater number of population does not imply better or faster calculations. For 200 individuals, a satisfactory result was reached already after forty minutes of calculations, while for 500 individuals the algorithm achieved optimum results already in the sixteenth minute of operation.

Fig. 4. Summary of genetic algorithm for PMS 2- pressure in measurement point no 2: error before optimization (blue), error after optimization (green).
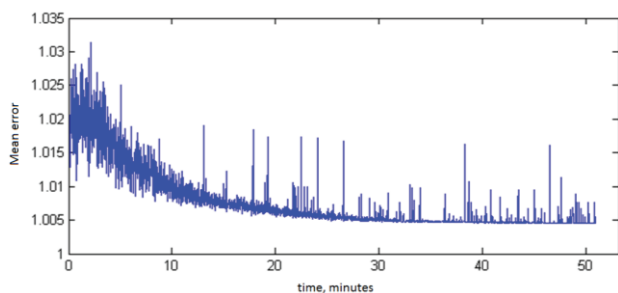
Fig. 5. The process of optimization in time. Population of 200 individuals, number of generations: 500.
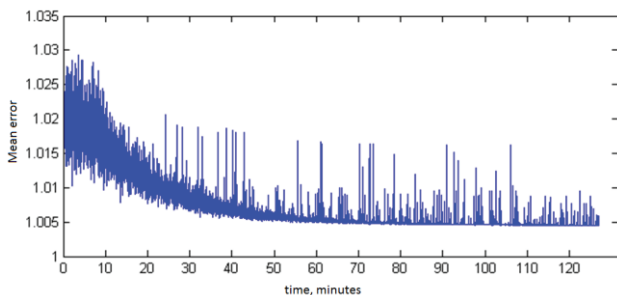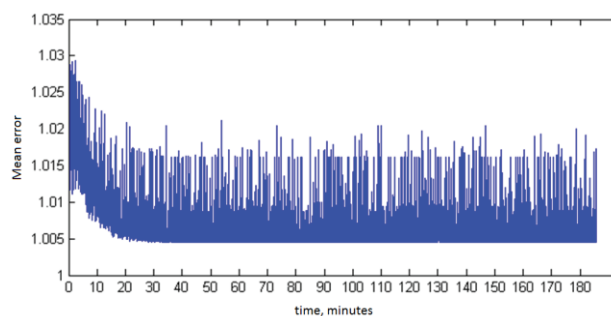

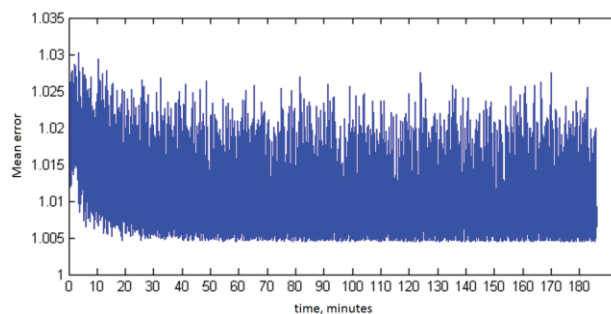**Fig. 6**. The process of optimization in time. Population of 500 individuals, number of generations: 200.

The undoubted advantage of a large number of populations is the widening of the scope of the search while maintaining the convergence of the algorithm to the final result. Modification of mutation parameters would give a similar or better extension effect, but it could cause problems with maintaining convergence to the desired optimum value.

This phenomenon shows the results of optimization attempts shown in Figs. 7 and 8. The algorithm searching a larger solution space significantly minimizes the error. However, single individuals cannot keep their genetic information long enough for the analysis of a narrowed neighborhood to enable finding the best result.

The results of calculations performed using the EPANET calibration tool were not satisfactory and the and the mean squared error was not significantly reduced.

Fig. 7. The process of optimization in time. Population of 200 individuals, number of generations equal to 400, modified probability of mutation = 0.1.

Fig. 8. The process of optimization in time. Population of 200 individuals, number of generations equal to 400, modified probability of mutation = 0.5.

Thus, in order to enhance the calibration results, the genetic algorithm implemented in MATLAB within the function *ga* was applied in combination with the epanet2.dll dynamic link library. For optimization the emitter parameter value was subjected in ten defined network sectors. On the prepared network, using the defined sectors, optimization was started with the algorithm configured according to the following settings:
• population size = 200,
• number of generations = 400,
• probability of mutation = 0.01,
• probability of crossing = 0.9,
• optimized parameter - node leakage = *emitter*,
• dependent variable: network *pressure* in 10 sectors,
• initial values: *roughness* = 0.1, *emitter* = 0.04.

As result of optimization the mean squared error value was significantly reduced, as shown in Table 1. It should be noted that successful optimization of the parameters using the "EPANET Calibrator" plug-in, is not possible when using automate diagnostic processes

e.g. in an expert system, since there exist no correct interfaces to the system.

Table 1. Summary of optimization results – RMSE error values before optimization and values obtained using EPANET build in genetic algorithm and MATLAB tollbox.

Location – sector name	RMSE **before optimization**	RMSE **after optimization with EPANET**	RMSE **after optimization with MATLAB**
p_{ZUW}	0.3104	0.3083	0.2767
p_{SUW}	0.7968	0.7777	5.4206
p_{PMS1}	1.5688	1.4049	0.4892
p_{PMS2}	6.5923	5.8241	1.0097
p_{PMS3}	1.1115	0.9184	0.5122
p_{PMS4}	10.7789	9.7080	1.9766
p_{PMS5}	11.2838	10.609	0.9033
p_{PMS6}	8.4838	8.0646	0.9485
p_{CIS}	10.1014	9.3956	2.7899
p_{SLA}	15.5286	14.248	1.7254
Average	**6.6556**	**6.1250**	**1.6052**

4 Conclusions

The paper depicts the procedure of calibration of water supply network model using two methods: a build-in EPANET calibrator and an algorithm using MATLAB toolbox. The process considers optimization of the *roughness* and *leakage* parameters, which are estimated based on measurement data by using genetic algorithm. The motivation behind the performed comparative analysis was to show the difference between the accuracy of the model parameters calculated using the two considered methods.

Based on the research task performed it was shown that it is crucial to optimize the model parameters in order to gather a better fitted simulation model to the empirical data. The calibration and optimization tasks are necessary for achieve a precise simulation model, which may be applied as an module in an expert system for detection and localization of leakages in water supply networks. Example of such an expert system an interested reader may find in [5].

Acknowledgment

References

1. H. Jia, W. Wei, K. Xin, Front. Environ. Sci. Eng. **2**, 57-62 (2008)
2. R. Schmid, *Review of modelling software for piped distribution networks. 2nd Ed.,* Swiss Centre for Development Cooperation in Technology and Management (2002)
3. B. Ulanicki, A. Zehnpfund, F.M. Alzamora, *Proc. 2nd Intl. Conf. on Hydroinformatics,* Zurich, (1996)
4. A. Preis, A.J. Whittle, A. Ostfeld, L. Perelman, J. Water Resour. Plan. Manage. **137,** 343-351 (2010)
5. D.A. Savic, Z.S. Kapelan, P.M.R. Jonkergow, Urban Water J. **6**, 3-22 (2009)
6. T. Boczar, D. Zmarzły, P. Frącz, *E-proceedings of the 37th IAHR World Congress*, Kuala Lumpur, Malaysia (2017)