Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 89 (2014) 750 – 757

Procedia Engineering

www.elsevier.com/locate/procedia

16th Conference on Water Distribution System Analysis, WDSA 2014

SWMM5 Toolkit Development for Pollution Source Identification in Sewer Systems

B.K. Banik[a,*], C. Di Cristo[b], A. Leopardi[b]

[a]*ETeCoS[3], Dipartimento di Ingegneria Civile e Meccanica, Università degli Studi di Cassino e del Lazio Meridionale, Cassino 03043, Italy*
[b]*Dipartimento di Ingegneria Civile e Meccanica, Università degli Studi di Cassino e del Lazio Meridionale, Cassino 03043, Italy*

Abstract

The Storm Water Management Model (SWMM) is considered one of the standard tools for modeling sewer systems, but it does not have the programmer's toolkit. In this work a toolkit library is presented, which has been built for running wastewater quality simulations from an outside environment. An example of the developed Toolkit use is furnished applying it to a pollution source identification (PSI) problem, expressed as an optimization problem. The presented example shows the capability of the Toolkit in providing a useful tool for performing network analysis from other applications and demonstrates the good performance of the proposed PSI methodology.

Peer-review under responsibility of the Organizing Committee of WDSA 2014

Keywords: Toolkit; Sewer system; SWMM5; Optimization problems; Pollution source.

1. Introduction

Wastewater systems in urban areas can be different in typology and size, but they all include a collection system, which can be combined or sanitary sewers. The good functioning of the wastewater treatment plant (WWTP) strongly depends on the quality of the wastewater. Moreover, watercourses are more susceptible to contamination when they are linked with sewer systems, because stormwater outfalls can discharge pollution from various catchment sources, such as, vehicle emission, building and road corrosion and erosion, animal faeces, street litter deposition, fallen leaves and grass residues and spills. Serious short-term pollution can arise from Combined Sewer Overflows (CSOs) too.

* Corresponding author. Tel.: +3907762994381; fax: +39 (0)776299 3939.
E-mail address: bk.banik@unicas.it

1877-7058
Peer-review under responsibility of the Organizing Committee of WDSA 2014
doi:10.1016/j.proeng.2014.11.503

Nomenclature

C_0	concentration of the contaminant at the source
C_{it}^{obs}	measured concentration at sensor *i* at time step *t*
C_{it}	simulated concentration value at sensor *i* at time step *t*
D	duration of the contamination event
F	objective function
i	sensor index
L	node index of the contaminant source
N_s	total number of sensors
t	time step
T_0	start time of the contamination event
t_0	time of the first detection of the contamination at sensors
t_c	current time

In order to control those situations, modeling wastewater hydraulic and quality in sewer systems is an important basic requirement for adequate management strategies [1]. Simplified quality models have been linked to hydraulic flow models of sewer systems for some time. By in the mid-1980s, deterministic flow models have become so popular and widespread in their use, such natural tools for the drainage engineer, that they seemed an appropriate step to develop deterministic quality models of sewer-flow. Physically-based deterministic models have become available for general use, also for quality modeling. Actually there is a greater variety of modeling packages for simulating both flow and quality in sewers, such as SWMM (Storm Water Management Model), MOUSE TRAP, Wallingford packages etc [1]. The USEPA's SWMM [2] is a dynamic rainfall-runoff simulation model that computes runoff quantity and quality from, primarily, urban areas. Runoff and pollutant loads generated from precipitation in different subcatchments are functioned by the runoff component of SWMM. This runoff transported through a system of pipes, channels, storage/treatment devices, pumps and regulators are operated by the routing portion. SWMM tracks the quantity of runoff generated within each subcatchment, evaluating the flow rate and depth in each pipe and channel during a simulation period comprised of multiple time steps. SWMM has also the ability of analyzing the buildup, washoff, transport and treatment of a number of water quality constituents during either wet or dry weather flow conditions. In water distribution systems (WDSs), starting from water quality modeling, many efforts have been spent for studying the detection of contamination events [3, 4, 5].

This topic has not been addressed in sewer systems, even if the individuation and the elimination of illicit intrusions is a very important aspect of their management policy. Recently, the wastewater and stormwater management is evolving from a simple flood and sanitary control to a whole environmental protection function. In many countries, such as EU and US states, operators are required to obtain a permit for discharges in sewer systems from the regulatory authority. So, a very important aspect of the sewer system management policy is represented by the detection and the elimination of illicit intrusions. Those events may generate problems to waste water treatment plant and/or to the final recipient water body. Moreover, for combined sewer systems intrusions are more probable as they are realized through open channel flow networks and because the collection networks are geographically dispersed and have multiple access points, which are generally not monitored. For all these reasons, in urban drainage systems (UDSs) the identification of the source of the illicit intrusion, along with the inflow characteristics is of particular interest. A first difference in treating the source contamination identification in UDSs respect to WDSs is in the definition of a contaminant event. In the present research work it is assumed that a contamination is represented by the intrusion of a pollutant substance, different from the usual composition of the wastewater of the network. In these terms, the detection of a contamination in a UDS is becoming a real possibility, since the implementation in the recent years of new real-time monitoring sensor systems instead of the standard analytical procedure measurements [6]. Assumed that, based on monitoring system measurements, an event detection procedure indicates a contamination is in act, in this paper a methodology to identify the contaminant source characteristics in a sewer network is proposed. This pollution source identification (PSI) methodology is formulated as an optimization problem, which minimizes an objective function. In particular, it has been developed in order to individuate the contamination source location, along with its main characteristics, represented by the input concentration, starting time and intrusion duration. The storm water management model SWMM is used as hydraulic and water quality simulator, while a Genetic Algorithm (GA) code is adopted for solving the optimization problem. One of the major limitations in using the SWMM software in the PSI methodology is the unavailability of a complete programmer's toolkit as the hydraulic network simulator EPANET [7] does. Several researchers [8, 9, 10] have made their effort to improve the existing EPANET toolkit, but no such effort have been seen so far in the case of SWMM. The current SWMM dynamic link library (DLL) consists of just nine functions, which are insufficient to perform any simulation from another platform (e.g. C++). In order to integrate the SWMM simulator with the proposed automated PSI methodology an ad-hoc SWMM-TOOLKIT is necessary to establish a communication from an outside environment.

In doing that some 22 additional functions have been created for retrieving information about network nodes and time patterns, as well as for setting new values during the extended period simulation from a C++ platform. An application of the SWMM-Toolkit for running the proposed PSI methodology is presented performing an example test considering the literature network Example 8 of SWMM manual [11].

2. The pollution source identification methodology

To date, most of the researches on contamination event detection are primarily focused on WDSs due to high concern about the public health in case a pollutant is introduced into the system either deliberately or accidentally. The current literature reflects how WDSs security related research has evolved in two seemingly separate directions [12]: (1) the contamination detection with optimal sensors displacement for an early warning [13, 14, 15, 16, 17] and (2) the pollution source identification problem. The studies that deal with the PSI problem [3, 12, 18, 19, 20, 21, 22, 23] are devoted to essentially individuate the intrusion point location, along with the main characteristics of the inflow. In particular, [24] proposed a genetic algorithm-based contamination source detection model which was further embedded in a statistical framework for quantifying the uncertainty of a contamination source detection outcome. Very rare efforts have been spent for studying the effect of an illicit injection in an UDS [25, 26], and in particular for the pollution source identification problem. The primary goal of a PSI problem is to estimate quickly the source characteristics that best explain the observed contamination data. In the following an original PSI methodology for UDS is presented, which involves the recognition of the injection location, the input start time, duration and magnitude. However, the use of such kind of methodology requires sensor measurements and a strategy in order to individuate that an event is occurring, as the one proposed for WDS in [27]. In this preliminary presentation, it is assumed that input pollutant is a conservative substance, whose concentration is measured through ideal sensors, without errors. The PSI problem is formulated searching the minimum of a time-dependent objective function, which minimize the difference between the simulated and the measured concentration values, furnished by the sensors. The mathematical formulation of the used dimensionless objective function F is:

$$F = \sum_{t=t_0}^{t_c} \sum_{i=1}^{N_s} \left(\frac{C_{it}^{obs} - C_{it}(L, C_0, T_0, D)}{(C_{it}^{obs} + C_{it}(L, C_0, T_0, D))/2} \right)^2 \quad (1)$$

where L = node index of the contaminant source; T_0 = start time of the contamination event; t_0 = time of the first detection of the contamination at the sensors; t_c = current time step; C_0= concentration of the contaminant at the source; D = duration of the contamination event; C_{it}^{obs} = measured concentration at sensor i at time step t; C_{it} = simulated concentration value at sensor i at time step t; i = sensor index; t = time step; and N_s = total number of sensors. For solving the optimization problem the simple Genetic Algorithm (GA) of the GAlib [28] is used. A schematic description of the interaction of PSI methodology with the SWMM and the GA codes is shown in Fig. 1.

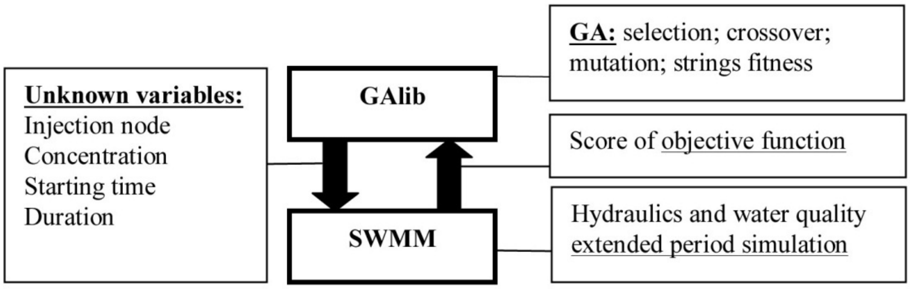

Fig. 1. Schematic representation of the PSI methodology

In the first step the GA defines the initial population. In more detail, the phenotypes/strings/solutions, randomly generated by GA, consist of four unknown variables: (1) Contamination injection node, represented by an integer number; (2) Contaminant concentration, represented by a real number; (3) injection starting time, represented by an integer number; and (4) injection duration, represented by an integer number. In the second step, using the values of the four variables furnished from GA, SWMM performs the hydraulic and water quality simulation and calculates the objective function (Eq.1). In the third step the fitness of each individual solution, based on the objective function score obtained from SWMM, is evaluated by GA. Depending on the fitness value new population is generated through mutation and crossover. Steps 2 and 3 are repeated until a fixed number of generations are achieved.

3. The SWMM toolkit description

The idea of making a SWMM toolkit derived from the necessity of implementing the methodology, described in the previous paragraph, for solving a PSI problem in sewer systems. So it is definitely not a complete SWMM programmer's toolkit, but it responds to the main requirements for performing a wastewater quality simulation. The programmer's toolkit is a DLL of functions that allows developers to customize SWMM's computational engine for their own specific needs. It provides a series of functions that allows programmers to perform the SWMM's (the engine consists of 46 C-code file and 19 header files) water quality solution engine to their own application. The functions were written in ANSI standard C, in the file swmm.c, which provides supervisory control over the program. They can be incorporated into 32-bit Windows applications written in C/C++ or in any other language that can call functions within a Windows DLL. The Toolkit DLL file is named SWMM5.DLL. The Toolkit comes with one header file and one SWMM5.lib file that simplify the task of interfacing it with C/C++ code. The data flow diagram for performing the PSI problem with the proposed methodology is shown in Fig. 2. In detail, the input processor module receives a description of the simulated network from a SWMM input file (.INP). Then, a second input file is represented by the SENSOR text file (.TXT), containing the information about the time series of the pollutant concentration measured in different sensor nodes. The file's contents are parsed, interpreted, and stored in a shared memory area. Both SWMM hydraulics and water quality solver modules carry out an extended period simulation. During this process both the computed hydraulic and water quality results for each preset reporting interval are written to an unformatted (binary) output file (.OUT). Finally, if requested, a report writer module reads back the computed simulation results from the binary output file (.OUT) for each reporting period and writes out selected values to a formatted report file (.RPT). Any error or warning messages generated during the run are also written to this file. Different Toolkit functions have been generated to carry out all of these steps under the programmer's control, including the ability to read or modify some of the most important system's global data, related to the PSI problem. In particular, the newly built Toolkit functions can be classified as five types according to the tasks they perform. The first type, which consists of two functions, is basically for opening and closing the SWMM toolkit system. The second one is for retrieving information about network nodes such as: node index, id, type and value of a particular parameter of that node. The third type retrieves information about the used time pattern and it has five functions. The fourth group is for getting other relevant information of the network like: number of network components of a specified object type, flow unit, starting/ending/reporting time and step, duration etc. The penultimate group of functions is able to set new values for node parameters and time pattern. Finally, the last group, which consists of only one function, is for customizing a simulation. The function *swmm_step,* already present in the original DLL provided by the US EPA, was modified to *OWNswmm_step,* because the previous one didn't meet all the requirements of the PSI problem having only one argument "elapsed time". In the new function another argument, "objective function score", has been added as in each time step the PSI methodology demands the score of an objective function.

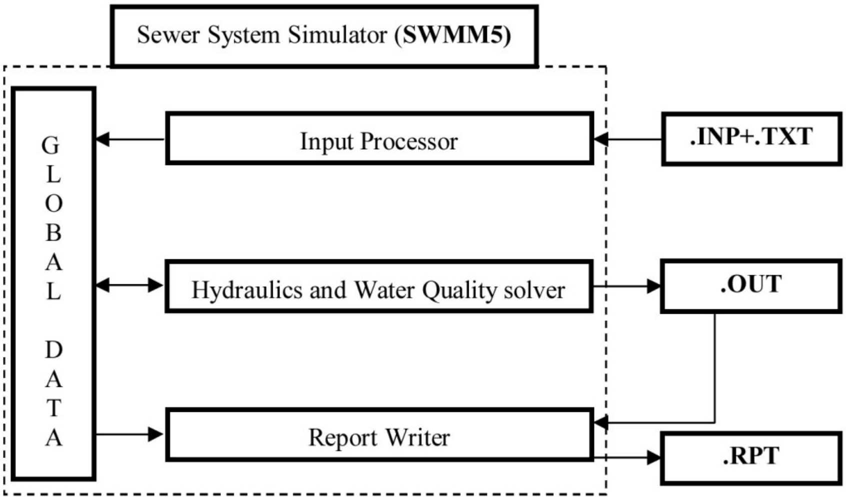

Fig. 2. Data flow diagram of SWMM-PSI Toolkit

The main toolkit functions used in performing the PSI methodology are:

- Calling *OWNswmm_open* function to open the toolkit system along with the SWMM input file and the SENSOR text file, which contain the measurements of the time series concentration of the pollutants observed through the sensors.
- Using *SWMMsetXXX* series of functions to change system characteristics.
- Running a simulation time step using the *OWNswmm_step* function. If needed, *SWMMsetXXX* series of functions can be used to re-set system characteristics.
- Finally calling *OWNswmm_close* function to close all files and release system memory.

4. Application and results

In order to show an application of the proposed PSI methodology with the SWMM-Toolkit the literature network reported in the Example 8 of the SWMM application manual [11] has been considered, even if some characteristics have been modified. The example system, schematically represented in Fig. 3, is a combined sewer and it consists of 31 nodes (28 junctions, 2 outfalls and 1 storage unit), 35 links (29 conduits, 1 pump, 1 orifice and 4 weirs). The served area is 11.74 hectares and the system is close to a natural stream, represented in Fig. 3 with an ash colored line with C label. Two outlets are also reported, where one is the WWTP, while the other one discharges combined sewer overflow during wet weather.

In Fig. 3 the sewer pipes are reported with thick black line with P label and they drain the rain from the subcatchments and/or dry weather flow. The interceptor, indicated with thin black line with I label, are pipes designed to capture the sanitary flows during dry weather periods and convey them to the WWTP. The flow regulators (weirs and orifices) and the pump station are also reported in Fig. 3. In the presented test, only dry weather flow (DWF) is considered herein with the inflow values reported in Table 1 in the input nodes indicated in Fig. 3. The input discharges follow a typical 24 hour variable pattern shown in Table 2.

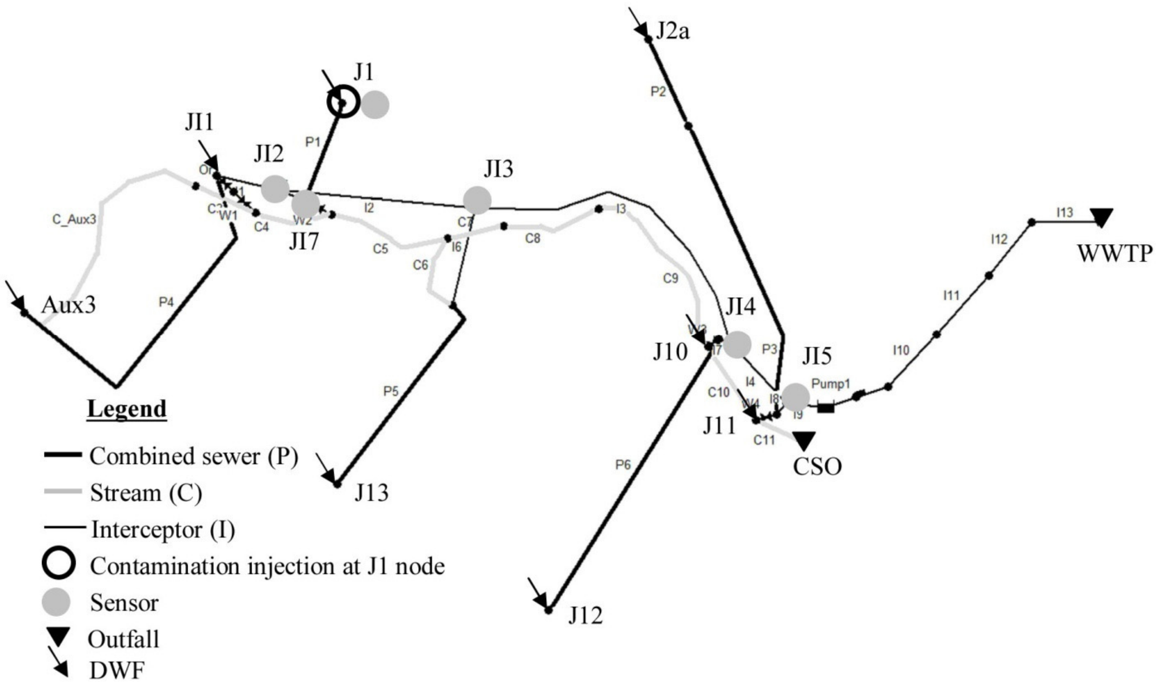

Fig. 3. Scheme of the example network

Table 1. DWF values in different nodes

Node	J1	J2a	J10	J11	J12	J13	Aux3	JI1
Flow (l/s)	2.27	2.83	0	0	3.54	3.48	2.27	3.48

Table 2. Hourly DWF pattern

Hour	1	2	3	4	5	6	7	8	9	10	11	12
Multiplier	0.4	0.3	0.2	0.2	0.5	0.8	1	1.8	1.6	1.4	1.2	1
Hour	13	14	15	16	17	18	19	20	21	22	23	24
Multiplier	1.1	1.1	1.2	1.3	1.3	1.4	1.4	1.5	1.3	1	0.6	0.5

Two input files are mandatory in order to accomplish the PSI methodology. One is the SWMM input file (.inp), with all data describing the network; the second is a text file SENSOR having the time series concentration measurements from the monitoring stations. In the presented test 6 sensors are assumed in the nodes J1, JI2, JI3, JI4, JI5 and JI7 (Fig. 3). The measured synthetic data have been generated through a SWMM quality simulation considering as "true" contamination scenario an injection of a conservative contaminant at node J1 with a concentration of 2 mg/l starting at 8 AM with duration of 3 hours. The simulation was run from 6:00 to 14:00, with a routing time step of 30 s. This duration was fixed considering the wastewater travelling time in the network. The first detection of the contamination was revealed by sensor J1 at 10:00. The generated synthetic sensor text file, along with the SWMM input file was used for running the PSI methodology. The following GA parameters were chosen: number of population =100; number of generation= 100; mutation probability= 0.01 and crossover probability= 0.9. Applying the PSI methodology presented in paragraph 2, the obtained values of the four PSI variables (Source Node, Input Concentration, Starting Time and Duration) are summarized in Table 3 as Case 1, showing that the exact solution is obtained. This result was expected, since it represents just a preliminary test for verifying the capability of the Toolkit in providing a useful instrument for performing the proposed PSI methodology or in general for running network analyses from other applications. Further investigations for testing the PSI methodology in more complex situations are necessary, such in wet weather conditions or considering measurement errors. Moreover, additional tests on a real case study are also planned in future studies.

Table 3. PSI results.

Solution	Source Node	Concentration (mg/l)	Starting Time (hr)	Duration (hr)
Exact Solution	J1	2.0	8:00	3
Case 1	J1	2.03529	8:00	3

5. Conclusions

The paper is focused on the presentation of newly built toolkit functionality for SWMM, realized in order to perform more complex applications related to wastewater hydraulic and quality simulations, such as optimization methodologies that require running many system analyses with modified input parameters. Different Toolkit functions have been generated, including the ability to read or modify some of the most important system's global data. In particular, the new Toolkit functions can be classified as five types according to the tasks they perform. The first type is basically for opening and closing the SWMM toolkit system, while the second one is for retrieving information about network nodes. The third type retrieves information about the used time pattern and the fourth group is for getting other relevant information of the network like number of network components of a specified object type, flow unit, starting/ending/reporting time and step, duration etc. The penultimate group of functions is able to set new values for node parameters and time pattern, and then the last group is for customizing a simulation.

The idea of making a SWMM toolkit derived from the necessity of implementing a methodology for solving a pollution source identification (PSI) problem in sewer systems. The proposed methodology expresses the PSI as an optimization problem which was solved using a Genetic Algorithm code combined with SWMM. Moreover, a preliminary test in applying the proposed PSI methodology to a literature sewer system is presented, showing the capability of the Toolkit in providing a useful tool for performing network analysis from other applications. The results demonstrate also a good performance of the PSI methodology, even if it has to be tested in more complex situations and on a real case study, which is the plan for future studies.

Acknowledgements

The first author would like to thank the EU for the financial support through the Erasmus Mundus Joint Doctorate Programme ETeCoS[3] (Environmental Technologies for Contaminated Solids, Soils and Sediments), grant agreement FPA n° 2010-0009.

References

[1] D. Butler, J.W. Davies. Urban drainage. 1st ed. New York: Taylor & Francis, 2004.
[2] LA. Rossman Storm water management model user's manual. Version 5.0. National Risk Management Research Laboratory, Office of Research and Development: US Environmental Protection Agency, 2010.
[3] C. Di Cristo, A. Leopardi, Pollution source identification of accidental contamination in water distribution networks. J. Water Resour. Plann. Manage., 134(2) (2008)197-202.
[4] A. Preis, A. Ostfeld Hydraulic uncertainty inclusion in water distribution systems contamination source identification. Urban Water J., 8(5) (2011) 267-277.
[5] L. Liu, E.M. Zechman, G. Mahinthakumar, S.R. Ranjithan. Identifying contaminant sources for water distribution systems using a hybrid method. Civ Eng Environ Syst 29(2) (2012) 123-136.
[6] N. Al-Dasoqi, A. Mason, R. Alkhaddar, A. Al-Shamma'a. Use of Sensors in Wastewater Quality Monitoring-a Review of Available Technologies. In proc. World Environmental and Water Resources Congress, 2011, pp. 354.
[7] L.A. Rossman, The epanet programmer's toolkit for analysis of water distribution systems. In Proc. Annual Water Resources Planning and

Management Conference, 1999, 8. 2.2:10-22.
[8] M. Guidolin, P. Burovskiy, Z. Kapelan, D. Savić, CWSNET: An Object-Oriented Toolkit for Water Distribution System Simulations. In Proc. Water Distribution Systems Analysis, 2010. pp. 1-13.
[9] M.S. Morley, R.M. Atkinson, D.A. Savic, G.A. Walters, OpenNet: An application independent framework for hydraulic network representation, manipulation & dissemination. Hydroinformatics Conference 2000 (CD-ROM publication).
[10] J.E. Van Zyl, J. Borthwick J, A. Hardy, OOTEN: An objected-oriented programmers toolkit for EPANET. In Proc. Advances in Water Supply Management CCWI 2003.
[11]J. Gironás, L.A. Roesner, J. Davis, L.A. Rossman, Supply W. Storm water management model applications manual. National Risk Management Research Laboratory, Office of Research and Development: US Environmental Protection Agency, 2009.
[12] M.E. Tryby, M. Propato, S.R. Ranjithan. Monitoring design for source identification in water distribution systems. J. Water Resour. Plann. Manage. 136(6) (2010) 637-646.
[13]L. Cozzolino, C. Mucherino, D. Pianese, F. Pirozzi. Positioning, within water distribution networks, of monitoring stations aiming at an early detection of intentional contamination. Civ. Eng Environ Syst 23(3) (2006) 161-174.
[14]L. Cozzolino, R. Della Morte, A. Palumbo, D. Pianese, Stochastic approaches for sensors placement against intentional contaminations in water distribution systems. Civ Eng Environ Syst 28(1) (2011) 75-98.
[15]F. Edthofer, J. Van den Broeke, J. Ettl, W. Lettl, A. Weingarthner, Reliable online water quality monitoring as basis for fault tolerant control. In Proc. Control and Fault-Tolerant Systems (SysTol), IEEE, 2010, pp. 57-62.
[16]A. Ostfeld, E. Salomon, Optimal layout of early warning detection stations for water distribution systems security. J. Water Resour. Plann. Manage. 130(5) (2004) 377-385.
[17]A. Ostfeld, A. Kessler, I. Goldberg, A contaminant detection system for early warning in water distribution networks. Eng Optimiz 36(5) (2004) 525-538.
[18] J.J. Huang, E.A. McBean, Data mining to identify contaminant event locations in water distribution systems. J. Water Resour. Plann. Manage 135(6) (2009) 466-474.
[19]C. Di Cristo, A. Leopardi, Closure to "Pollution source identification of accidental contamination in water distribution networks", J. Water Resour. Plann. Manage. 136(2) (2010) 292-294.
[20] C.D. Laird, L.T. Biegler, B.G. van Bloemen Waanders. Mixed-integer approach for obtaining unique solutions in source inversion of water networks. J. Water. Resour. Plann. Manage 132(4) (2006) 242-251.
[21] A. Preis, A. Ostfeld, Genetic algorithm for contaminant source characterization using imperfect sensors. Civ Eng Environ Syst 25(1) (2008) 29-39.
[22] A.E. De Sanctis, F. Shang, J.G. Uber. Real-time identification of possible contamination sources using network backtracking methods. J Water Resour. Plann. Manage. 136(4) (2010), 444-453.
[23] E.M. Zechman, S.R. Ranjithan. Evolutionary computation-based methods for characterizing contaminant sources in a water distribution system. J. Water Resour. Plann. Manage. 135(5) (2009) 334-343.
[24] A. Preis, A. Ostfeld. Hydraulic uncertainty inclusion in water distribution systems contamination source identification. Urban Water J. 8(5), (2011) 267-277.
[25]R. Field, R. Pitt, M. Lalor, M. Brown, W. Vilkelis, E. Phackston. Investigation of dry-weather pollutant entries into storm-drainage systems. J Environ. Eng. ASCE 120(5) (1994) 1044-1066.
[26]K. Irvine, M.C. Rossi, S. Vermette, J. Bakert, K. Kleinfelder. Illicit discharge detection and elimination: Low cost options for source identification and track down in stormwater systems. Urban Water J., 8(6) (2011) 379-395.
[27]N. Oliker, A. Ostfeld. Minimum volume ellipsoid classification model for contamination event detection in water distribution systems. Environ. Modell. Softw. 57 (2014) 1-12.
[28] M. Wall, GAlib: A C++ library of genetic algorithm components. Mechanical Engineering Department: Massachusetts Institute of Technology, 1996.