Model predictive control based on artificial intelligence and EPA-SWMM model to reduce CSOs impacts in sewer systems

Khalid El Ghazouli a,b,*, Jamal El Khatabi a, Aziz Soulhi c and Isam Shahrour a

a Laboratoire de Génie Civil et géo-Environnement, Univ. Lille, IMT Lille Douai, Univ. Artois, Yncrea Hauts-de-France, ULR, 4515 - LGCgE, Lille F-59000, France
b Laboratoire d’Analyse des Systèmes, Traitement de l’Information et Management Industriel, Université Mohammed V, Rabat, Morocco
c National Higher School of Mines, Agdal Rabat, Morocco
*Corresponding author. E-mail: elghazouli.khalid@gmail.com

ABSTRACT

Urbanization and an increase in precipitation intensities due to climate change, in addition to limited urban drainage systems (UDS) capacity, are the main causes of combined sewer overflows (CSOs) that cause serious water pollution problems in many cities around the world. Model predictive control (MPC) systems offer a new approach to mitigate the impact of CSOs by generating optimal temporally and spatially varied dynamic control strategies of sewer system actuators. This paper presents a novel MPC based on neural networks for predicting flows, a stormwater management model (SWMM) for flow conveyance, and a genetic algorithm for optimizing the operation of sewer systems and defining the best control strategies. The proposed model was tested on the sewer system of the city of Casablanca in Morocco. The results have shown the efficiency of the developed MPC to reduce CSOs while considering short optimization time thanks to parallel computing.

Key words: artificial intelligence, combined sewer overflows, genetic algorithm, model predictive control, real-time control, sewer network

HIGHLIGHTS

- Model predictive control of smart sewer networks.
- Artificial Neural Networks and parallel computing enhance the proactivity of the MPC.
- Real-Time Control.
- Combined sewer overflows reduction.

INTRODUCTION

As results of urbanization and climate change, world agglomerations are facing major environmental issues, particularly those related to pollution that impacts waterbodies. In many cities, the existing combined sewer systems cannot convey all the polluted water to wastewater treatment plants during rain events (Zhao et al. 2019), leading to frequent combined sewer overflows (CSOs). The pollution released by sewer networks can significantly impact the ecosystem by unbalancing its kinetics through the increase of the concentrations of microbiological, mineral, and organic pollutants, thereby leading to oxygen depletion and a rise in eutrophication (Chocat 1997; McLellan et al. 2007; Weyrauch et al. 2010; Passerat et al. 2011; Phillips et al. 2012; Brokamp et al. 2017). Climatic factors, such as the quantity and intensity of precipitation, are key factors that determine the severity of CSO discharges (Botturi et al. 2020). According to future meteorological projection scenarios, a substantial increase in storm intensities and frequencies will be recorded, thereby causing more frequent CSOs (Yazdanfar & Sharma 2015; Alves et al. 2016; Jean et al. 2018).

Multiple research studies have discussed the impact of CSOs on ecosystems. Viviano et al. (2017) demonstrated through a complete monitoring scheme based on caffeine and turbidity that more than 50% of the total phosphorus of the Lambro River is due to sewer network overflows from rainy weather. Studies conducted by Phillips et al. (2012) and Launay et al. (2016) affirmed that even if the volume of CSOs represents a small part of the annual volume of wastewater, their impact is not negligible and can contribute up to 95% of the annual pollution caused by various pollutants.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).
Several actions can minimize CSOs and improve the receiving environment quality. Green infrastructures play a significant role in limiting peak flows and pollution, and they also relieve the downstream sewer network by regulating the flow or infiltrating water. Storage basins allow the storage of a large volume of polluted water during rain events and release it back to sewer networks once the rainfall events are over. However, the construction of basins remains complicated because of the lack of space, construction, and maintenance costs (Garofalo et al. 2017). One of the emerging ways to reduce CSOs is by performing the advanced real-time control of urban drainage systems (UDS) based on model predictive control (MPC), which computes optimal control strategies on the basis of deterministic rain forecasts. MPC has exhibited the efficient and cost-effective management of sewer systems to reduce pollution and energy consumption through several research case studies. Lund et al. (2020) used an integrated stormwater inflow control to mitigate CSOs in Copenhagen by dynamically controlling stormwater inflow to the combined sewer system in real time. This control was performed with an MPC on the basis of convex optimization including a linear internal surrogate. The MPC was tested on 18 rainfalls that have caused CSOs. Four of the 18 events were avoided, and the total CSO volume was reduced by 98.4% of the potential reducible volume. In addition, Bonamente et al. (2020) demonstrated through a study conducted on a sewer system that an MPC based on the NSGA II optimization algorithm can save energy consumption up to 32% and an overflood of approximately 10%. Rathnayake & Faisal Anwar (2019) also successfully applied an MPC to the combined sewer network of Liverpool in the United Kingdom using the NSGA II and SWMM models. The proposed model minimized the pollution load in the receiving water body and wastewater treatment and pumping costs in the sewer system. Although the proposed algorithm produced satisfactory results, the solution algorithm could not be applied to real-time control due to the simulation time needing to be improved. Considering that the reaction time in urban sewer systems is usually short, for UDS with many decision variables, such as a high number of gate valves or flow-regulating structures, the algorithm’s running time may be very long and unsuitable for real-time control purposes.

This paper aims to fill the gap in the MPC research field by presenting the development and investigation of the performance of an MPC, based on robust genetic algorithms (GAs) and neural networks that have the advantage of performing fast calculations that fit the needs of such systems. Further, it aims to demonstrate the benefits of MPC associated with parallel computing that offers a sufficient lead time to define the global optimal control strategies of weir gate valves to reduce CSOs in the smart and durable city.

MATERIALS AND METHODS

The control of sewer networks is based on regulated structures (e.g., gates and pumps), most of the time controlled locally, and does not depend on communication with other facilities or the other parts of the sewer system. Local control strategies may represent a good solution in the case of one actuator, but in the case of complex sewer systems where many actuators must operate jointly, a dynamic global control system becomes necessary. In dynamic systems, control actions are based on the time-varying requirements of interests in a sewer system, the water system load, and the watershed dynamic processes.

The current work aims to develop a robust dynamic MPC system for implementing global control strategies aiming at reducing CSOs through the dynamic management of gate valves. The MPC system combines a supervisory control and data acquisition (SCADA) that receives information and measures from monitoring sensors and implements control actions, hydraulic modeling software for flow conveyance, and artificial intelligence algorithms for forecasting and control optimization purposes. MPC is based on the following three main parts (Figure 1):

1. The first part concerns the forecast of wastewater and rainwater flows at watershed outlets representing strategic control points.
2. The second part comprises real-time modeling to better represent the network state at any time.
3. The third part is about predictive modeling and optimization.

Flow forecasting in sewer networks

Without a robust model, forecasting flows in sewer networks constitute significant uncertainties for operators. Short-term forecasts are an essential component for any MPC system and significantly improve the reaction time. Nevertheless, given that the main sewer system is combined and considering the spatial variation of rainfalls, having a flow forecast of dry and rainy weather flows is necessary. The forecasting of stormwater discharges is performed with a stormwater forecasting model (SWFM) based on the NARX neural networks, which have the advantage of performing fast calculations and providing quick and accurate stormwater discharges for anticipatory models (El Ghazouli et al. 2019). This model takes rainfall...
forecasts as inputs once available and returns as output stormwater discharges. Given that a nowcasting method based on extrapolation gives reasonable values for short-term (0–180 min) rainfall forecasts (Bowler et al. 2006; Berenguer et al. 2011), the SWFM will run with the latest updated rainfall forecasts every hour.

Instantaneous dry weather flows at watershed outlets are predicted with a wastewater flow forecasting model (WWFFM). The WWFFM is an artificial neural network (ANN) black-box model that handles nonlinear problems taking real-time water consumption and previous wastewater flow records as inputs (El Ghazouli et al. 2021). The output of the model is a 5-h forecasted wastewater flow time series. The combination of the SWFM and the WWFFM gives combined sewer discharge inputs for the SWMM model for flow conveyance and optimization purposes.

The architectures of the SWFM and the WWFFM neural networks include two layers; namely, a hidden layer and an output layer. The tan-sigmoid nonlinear transfer function was used in the hidden layer, and the unbounded linear transfer function that transforms the weighted sum inputs of the neurons to an output was employed in the output layer. During the learning phase, the weights and biases of the SWFM and the WWFFM ANN were adjusted according to the Levenberg–Marquardt back-propagation (LMBP) algorithm to minimize the error between the neural network output and measured data. In addition, the suitable lag parameter of the WWFFM was assessed through a cross-correlation analysis between the main causal variable and the output data. The lag parameter was defined as the concentration time of the watersheds for the SWFM, and the dataset was split into three subsets using the divide block method for the forecasting models. The first subset, representing 70% of the data, is the training set used to find the model parameters by computing the gradient and updating the network weights and biases, and the second subset, depicting 15% of the dataset, is the validation set. Additionally, during the training, the validation set error was monitored to prevent the increase of errors that leads to overfitting based on the early stopping method. The remaining 15% of the dataset was utilized as a test set to evaluate the performances and the generalization error in the final models. The SWFM and WWFFM NARX-NN were trained in their open-loop form and turned to their closed-loop form to perform multistep-ahead time series forecasting. Various neurons were tested, and the best training, testing, and validation results were obtained with a hidden layer with 10 neurons after several trials.

Real-time modeling

The MPC system uses the EPA SWMM engine to compute flow conveyance on the basis of a simplified model that comprises the main branches of the interception system of the sewer networks of Casablanca, as can be seen in Supplementary Material, Fig. S1. The model is connected to a SCADA and continuously updated with real-time data. Flow rates at the outlet of the watersheds and orifice opening status are automatically set to the same values as those observed in the field. Moreover,
the real-time model is automatically updated every hour with the last few hours of data. At the end of each model run, a Hot-start file containing the system boundary conditions is generated and will be employed for the next simulation with stormwater and dry weather flow forecasts as inputs.

**Optimization of the operational system**

The optimization section consists of algorithms able to set optimal operating control strategies while considering various constraints, such as the maximum flow of the wastewater treatment plant (WWTP). GAs are widely utilized and present their efficiency to solve many optimization problems in many fields, specifically in water. We can cite the works of Tayfur et al. (2009) for predicting peak flows, Li et al. (2020) for water resource management, Bostan et al. (2019) for the optimal design of shock dampers, Montes et al. (2020) for predicting bedload sediment transport in sewer networks, and Hassan et al. (2020) for the optimal design of sewer works. Therefore, GAs are chosen to optimize the sewer system operating as part of this work. GAs are inspired by the evolution theory of Darwin (1859) and, more specifically, natural selection, reproduction, and the survival of the fittest. In addition, they belong to the larger class of evolutionary algorithms. They were also developed by Holland (1962) in the 1970s and were popularized in the late 1980s and the early 1990s (Davis 1987; Goldberg 1989; Alliot & Schiex 1993; Forrest 1993).

The MPC is designed to be robust enough to define optimal strategies for managing the sewer system. The optimization of the objective function that we seek to minimize is reducing CSOs by maximizing the treated volume of polluted water at the WWTP. The optimization work involves adjusting the system decision variables that correspond to gate valve opening status at each time of the simulation horizon during different generations considering the operation of constraints, which are mainly the maximum flow capacities at the entrance of the WWTP, by applying a succession of operators to population individuals. The flow chart below describes the distinct steps of GAs.

First, an initial population (P (0)) of gate valve status is generated. This population contains chromosomes distributed in the solution space with varied genetic materials defined in a given interval. The fitness of each chromosome is evaluated, and the best individuals are selected for reproduction. Once the selection operation is conducted for the current population (P (t)), the crossover operator is then applied. The fundamental role of the crossover operation is to enrich the population diversity by manipulating the structure of chromosomes through the recombination of information present in the genetic heritage of individuals. The newly generated individuals are then submitted to the mutation operator. This operator allows GAs to better search the space of solutions by changing the allelic value of the gene with a very low mutation probability, which is generally between 0.01 and 0.001 (Fogel et al. 2000). The last step in the iterative process is to incorporate new solutions into the current population. The new solutions are added to the current population, thereby replacing the old solutions. This process is repeated until a stop criterion is met, and this criterion can be a determined number of generations or a particular value of the objective function (Figure 2).

**Application for the sewer network of Casablanca (Morocco)**

The sewer network of the eastern part of Casablanca is used to implement and evaluate the proposed MPC. The total area of the studied sector is approximately 41.6 km² (Figure 3). It comprises parallel watersheds with combined sewer networks.

The main sewer network outlets of these watersheds are historically discharged into the sea. An interceptor has been designed to convey dry weather flows to the treatment plant and discharges the excess water into the sea by rainy weather via frontal or side weirs. Additionally, it consists of a pipe of 9.5 km with a maximum capacity constrained by the capacity of a lift station (6.5 m³/s) downstream at the entrance to the WWTP. At each branch level, the interception system (Figure 4) is equipped with gate valves that control the flow diverted to the WWTP. These valves are modeled as orifices in the SWMM model. Today, by rainy weather and in the absence of control strategies, the gate valves are closed, and the polluted water is discharged directly into the sea.

The developed algorithm for the sewer network of Casablanca considers one objective function (Obj *Fun*) that is formulated to minimize CSOs by maximizing the total treated volume. The mathematical expression of the objective function is given in Equation (1):

\[
\text{Obj } \text{*Fun*} = - \sum_{i=1}^{n} (V_{x,i})
\]

where \(V\) is the intercepted volume at branch \(x\) at a given timestep \(i\).
The above-mentioned objective function is under a set of nonlinear constraints. A constraint is set such that the instantaneous flow rate at the WWTP entrance should not exceed its maximum capacity. The mathematical expression of the constraint is given in Equation (2):

\[ Qt \leq Q_{\text{max}} \]  

(2)

where \( Qt \) and \( Q_{\text{max}} \) are the flow rates at the entrance of the WWTP at timestep \( t \).

The number of decision variables (\( n \)) corresponds to the states (\( S_t \)) of the orifices for the simulation duration. The decision variable values are bounded between 0, which corresponds to a closed valve, and 1 represents a totally opened valve. The MPC result is an optimal gate valve control schedule with a 30-min timestep.

The MPC was initially run on the basis of a one-year return period (YRP) double-triangle rainfall hyetograph with a total duration of 1 h, an intense duration of 10 min, and a simulation time of 3 h. The optimization calculations were first performed in serial computing for different population sizes composed of 10, 20, 40, 60, 80, and 100 individuals over 100 generations to choose the best population size. This optimization problem comprises 36 decision variables corresponding to the states of six orifices for a simulation horizon of 3 h with six timesteps of 30 min.

Once the suitable numbers of population and generation are chosen, MPC is then evaluated on a real rainfall event recorded on 11 December 2017. The event has been marked by two successive rainfalls with a return period of 5 years, a cumulative rain of 38 mm, and a total duration of 6 h. For this event, MPC runs a simulation every 2 h with a 5-h simulation horizon.

For problems with a high number of decision variables, the running time of the algorithm may be very long. Parallel computation is a technique that reduces computational time by distributing the population and evaluating their fitness over several workers. Parallelization tasks are conducted for this event, and a comparison of the performance times according
to a different number of threads (workers) is performed. All the simulations are performed on the same desktop PC (CPU: AMD Ryzen Threadripper 3970X 4.5 GHz, RAM: 128 GB).

RESULTS AND DISCUSSION

Serial computing
Simulation results for distinct population sizes (Table 1 and Figure 5) applied to one YRP rainfall event validate that the MPC demonstrates its ability to reduce CSOs into the sea by maximizing the volume treated without exceeding the maximum flow at the WWTP. We can notice that population size significantly impacts the final solution determined by GAs.

The analysis of Figure 6, which presents the evolution of fitness according to generation numbers, verifies that the larger the population size, the more the algorithm converges toward an optimal solution in the first iterations of the calculation. We can also observe that the number of generations becomes negligible on the convergence as soon as we exceed the generation equal to 60.

Computation time is a critical component in MPC systems, and it must be reduced. Supplementary Material, Fig. S2 depicts the linear relationship between the computation time and the population size. The analysis of the various results confirms that population size and a generation number equal to 60 appear to be a good compromise between the algorithm performance in finding an optimal solution and the computation time. Hence, population size and a number of generations equal to 60 are applied for the parallel calculation for the 11 December 2017 rainfall event.

Parallel computing
For the 11 December 2017 rainfall event, the MPC performed four runs with a 5-h simulation horizon. These simulations are updated every 2 h to consider recent and accurate sewer network states, boundary conditions, and rainfall forecasts, allowing definition of an operating schedule for the various control valves over 5 h with a 30-min timestep for each simulation (Figure 7).
For each run, the SWFM and the WWFFM generate accurate forecasted flows (Supplementary Material, Fig. S3) used as inputs for the SWMM model. The NARX-NN allows the MPC to benefit more proactively from the relatively fast computing time of neural networks within 2 sec compared with a conventional hydraulic model that takes approximately 6 min for each run.

For the rainfalls of 11 December 2017, a comparison of the performance times according to the divergent numbers of threads (workers) is performed. Supplementary Material, Fig. S4 illustrates that parallel computing can significantly reduce computing time from 4,092 s for four threads to 890 s for 64 threads on the same processor. However, parallel

**Figure 4** | Interception system.

**Table 1** | Optimized flow for different population sizes

| Population size | 10  | 20  | 40  | 50  | 60  | 80  | 100 |
|-----------------|-----|-----|-----|-----|-----|-----|-----|
| Treated water volume (m³) | 53,978 | 53,674 | 55,110 | 55,153 | 54,707 | 54,950 | 54,950 |
computing performance decreases exponentially as a function of the number of threads employed in the simulation. Above 32 threads, the number of threads has no significant impact on the solution time. Moreover, for more proactivity, the computation time could be reduced further by parallelizing computation on several processors.

Figure 8 exhibits that the MPC developed as part of this work gives good results and enhances the efficiency of the sewer system of Casablanca, thereby allowing conveyance to the WWTP, in addition to strict wastewater, a volume of 146,830 m$^3$ of polluted rainwater without exceeding the maximum flow of the WWTP. The MPC managed the rainy event of 11 December 2017 by generating a schedule (Supplementary Material, Fig. S5) for controlling the gate valves at the interception structures. Furthermore, we can notice that a lack of a control strategy would cause disturbances and flooding at the lifting station at the entrance of the treatment plant, with a peak flow rate exceeding 8.5 m$^3$/s.

**CONCLUSIONS**

This paper presented a novel MPC based on neural networks for predicting flows and a GA for optimizing the operation of the sewer system of Casablanca. The MPC demonstrates high efficiency in reducing CSOs in the receiving environment by generating optimal temporally and spatially varied dynamic control strategies of the gate valves of the sewer system. This MPC could benefit municipalities around the world facing environmental issues and their consequences, such as CSOs and floods. As part of this work, we determined the best population size and the generation number that present a compromise between the algorithm performance in finding an optimal solution and the computation time. The parallelization of calculations allows the reduction of the computation time considerably and makes the MPC more proactive. However, since the MPC relies on weather forecasts, handling the uncertainty is required to successfully introduce the use of rainfall forecasts in operational management systems. Long-term verification analysis is needed to assess the quality of the forecasts for a particular sewer system. If this is satisfactory, a verification analysis is needed to test the decision rules and control strategies, given the forecasts, so that operators can see the effect of potential management strategies and avoid disasters.
Figure 6 | Fitness with populations equal to (a) 10; (b) 20; (c) 40; (d) 60; (e) 80; and (f) 100.
In the continuity of this work, a pollution measurement campaign will be conducted at the outlet of each watershed to determine the impact of CSOs on the natural environment and will be completed by two-dimensional hydraulic modeling of the dispersion of pollution into the sea. Thus, on the basis of these future results, the objective function will be adapted by weighting the volumes intercepted at the level of each branch according to the identified impact risk.

**DATA AVAILABILITY STATEMENT**

All relevant data are included in the paper or its Supplementary Information.

**REFERENCES**

Alliot, J. M. & Schiex, T. 1993 *Intelligence artificielle et informatique théorique*. Cepadues, Toulouse, France.

Alves, A., Sanchez, A., Vojinovic, Z., Seyoum, S., Babel, M. & Brdjanovic, D. 2016 Evolutionary and holistic assessment of Green-grey infrastructure for CSO reduction. *Water* 8 (9), 402.

Berenguer, M., Sempere-Torres, D. & Pegram, G. G. S. 2011 SBMcast – an ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *Journal of Hydrology* 404 (3–4), 226–240.

Bonamente, E., Termite, L. F., Garinei, A., Menculini, L., Marconi, M., Piccioni, E., Biondi, L. & Rossi, G. 2020 Run-time optimisation of sewer remote control systems using genetic algorithms and multi-criteria decision analysis: CSO and energy consumption reduction. *Civil Engineering and Environmental Systems* 37 (1–2), 62–79.

Bostan, M., Akhtari, A. A., Bonakdari, H. & Jalili, F. 2019 Optimal design for shock damper with genetic algorithm to control water hammer effects in complex water distribution systems. *Water Resources Management* 33 (5), 1665–1681.
Botturi, A., Ozbayram, E. G., Tondera, K., Gilbert, N. I., Rouault, P., Caradot, N., Gutierrez, O., Daneshgar, S., Frison, N., Akyol, Ç., Foglia, A., Eusebi, A. L. & Fatone, F. 2020 Combined sewer overflows: a critical review on best practice and innovative solutions to mitigate impacts on environment and human health. Critical Reviews in Environmental Science and Technology 51 (15), 1585–1618.

Bowler, N. E., Pierce, C. E. & Seed, A. W. 2006 STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. Quarterly Journal of the Royal Meteorological Society 132 (620), 2127–2155.

Brokamp, C., Beck, A. F., Muglia, L. & Ryan, P. 2017 Combined sewer overflow events and childhood emergency department visits: a case-crossover study. Science of the Total Environment 607–608, 1180–1187.

Chocat, B. 1997 Encyclopédie de l’hydrologie urbaine et de l’assainissement. Tec & doc-Lavoisier, Paris, pp. 529–543.

Darwin, C. 1859 In: On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life (Murray, J., ed.). John Murray, London.

Davis, L. 1987 Genetic Algorithms and Simulated Annealing. Morgan Kaufmann Publishers, Pitman, London.

El Ghazouli, K., El Khattabi, J., Shahrouj, I. & Soulihi, A. 2019 Comparison of M5 model tree and nonlinear autoregressive with exogenous inputs (NARX) neural network for urban stormwater discharge modelling. MATEC Web of Conferences 295, 02002. https://doi.org/10.1051/matecconf/201929502002.

El Ghazouli, K., El Khattabi, J., Shahrouj, I. & Soulihi, A. 2021 Wastewater flow forecasting model based on the nonlinear autoregressive with exogenous inputs (NARX) neural network. H2Open Journal. https://doi.org/10.2166/h2oj.2021.107.

Fogel, D. B., Bäck, T. & Michalewicz, Z. 2000 Evolutionary Computation: Basic Algorithms and Operators. Institute of Physics Publishing Bristol, PA.

Forrest, S. 1993 Genetic algorithms: principles of natural selection applied to computation. Science 261, 872–878.

Garofalo, G., Giordano, A., Piro, P., Spezzano, G. & Vinci, A. 2017 A distributed real-time approach for mitigating CSO and flooding in urban drainage systems. Journal of Network and Computer Applications 78, 30–42.

Goldberg, D. E. 1989 Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Publishing Company, Boston, Massachusetts, USA.

Hassan, W. H., Attea, Z. H. & Mohammed, S. S. 2020 Optimum layout design of sewer networks by hybrid genetic algorithm. Journal of Applied Water Engineering and Research 8 (2), 108–124.

Holland, J. H. 1962 Outline for a logical theory of adaptive systems. Journal of the ACM 9 (3), 297–314.

Jean, M. É., Duchesne, S., Pelletier, G. & Pleau, M. 2018 Selection of rainfall information as input data for the design of combined sewer overflow solutions. Journal of Hydrology 565, 559–569.

Launay, M. A., Dittmer, U. & Steinmetz, H. 2016 Organic micropollutants discharged by combined sewer overflows – characterisation of pollutant sources and stormwater-related processes. Water Research 104, 82–92.

Li, M., Zhang, J., Cheng, X. & Bao, Y. 2020 Application of the genetic algorithm. In: Water Resource Management (Atiquzzaman, M., Yen, N. & Xu, Z., eds). Springer Singapore, Big Data Analytics for Cyber-Physical System in Smart City. Singapore, pp. 1681–1686.

Lund, N. S. V., Borup, M., Madsen, H., Mark, O. & Mikkelsen, P. S. 2020 CSO reduction by integrated model predictive control of stormwater inflows: a simulated proof of concept using linear surrogate models. Water Resources Research 56 (8), 1–15.

McLellan, S. L., Hollis, E. J., Depas, M. M., Van Dyke, M., Harris, J. & Scopec, C. O. 2007 Distribution and fate of Escherichia coli in Lake Michigan following contamination with urban stormwater and combined sewer overflows. Journal of Great Lakes Research 33 (3), 566.

Montes, C., Berardi, L., Kapelan, Z. & Saldañariaga, J. 2020 Predicting bedload sediment transport of non-cohesive material in sewer pipes using evolutionary polynomial regression – multi-objective genetic algorithm strategy. Urban Water Journal 17 (2), 154–162.

Passerat, J., Ouattara, N. K., Mouchel, J. M., Rocher, V. & Servais, P. 2011 Impact of an intense combined sewer overflow event on the microbiological water quality of the Seine River. Water Research 45 (2), 893–903.

Phillips, P. J., Chalmers, A. T., Gray, J. L., Kolpin, D. W., Foreman, W. T. & Wall, G. R. 2012 Combined sewer overflows: an environmental source of hormones and wastewater micropollutants. Environmental Science & Technology 46 (10), 5336–5343.

Rathnayake, U. & Faisal Anwar, A. H. M. 2019 Dynamic control of urban sewer systems to reduce combined sewer overflows and their adverse impacts. Journal of Hydrology 579, 124150.

Tayfur, G., Barbetta, S. & Moramarco, T. 2009 Genetic algorithm-based discharge estimation at sites receiving lateral inflows. Journal of Hydrologic Engineering 14 (5), 463–474.

Viviano, G., Valsecchi, S., Polesello, S., Capodaglio, A., Tartari, G. & Salerno, F. 2017 Combined use of evapotranspiration and water fluxes to evaluate the impact of CSOs on water river quality. Water, Air, & Soil Pollution 228 (9), 330.

Weyrauch, P., Matzinger, A., Pawlowsky-Reusing, E., Plume, S., von Seggern, D., Heinzmann, B., Schroeder, K. & Rouault, P. 2010 Contribution of combined sewer overflow to trace contaminant loads in urban streams. Water Research 44 (15), 4451–4462.

Yazdanfar, Z. & Sharma, A. 2015 Urban drainage system planning and design – challenges with climate change and urbanization: a review. Water Science and Technology 72 (2), 165–179.

Zhao, W., Beach, T. H. & Rezgui, Y. 2019 Automated model construction for combined sewer overflow prediction based on efficient LASSO algorithm. IEEE Transactions on Systems, Man, and Cybernetics: Systems 49 (6), 1254–1269.

First received 30 May 2021; accepted in revised form 15 November 2021. Available online 29 November 2021