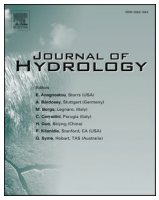
Journal of Hydrology 628 (2024) 130515

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Forecasting and optimization for minimizing combined sewer overflows using Machine learning frameworks and its inversion techniques

Zeda Yin [a,*], Yasaman Saadati [b,c], Arturo S. Leon [a], M. Hadi Amini [b,c], Linlong Bian [a], Beichao Hu [d]

[a] Department of Civil and Environmental Engineering, College of Engineering and Computing, Florida International University, Miami, FL, USA
[b] Sustainability, Optimization, and Learning for InterDependent Networks Laboratory (solid lab), FIU, Miami, FL, USA
[c] Knight Foundation School of Computing and Information Sciences, College of Engineering and Computing, Florida International University, Miami, FL, USA
[d] Department of Mechanical and Materials Engineering, College of Engineering and Computing, Florida International University, Miami, FL, USA

ARTICLE INFO

Keywords:
Combined Sewer Overflow
Prediction
Optimization
Machine Learning

ABSTRACT

Combined sewer overflows (CSOs), which typically occur during heavy rainfall events, pose significant threats to both public health and the environment. These threats encompass various concerns, including contamination of drinking water. Numerous studies have developed strategies aimed at mitigating the adverse effects of CSOs. These strategies include Green Infrastructure, Integrated Planning, and Smart Control Strategies. Among these, Smart Control Strategies have gained the most traction due to their exceptional cost-effectiveness. Nevertheless, the existing control methods face a challenge in striking the right balance between precision and computational efficiency. While employing full numerical methods as predictive models can provide high accuracy, they often prove inefficient in terms of runtime, especially when applied to real-world complex combined sewer systems. Conversely, reduced-order models tend to offer quicker results but may sacrifice accuracy. To address this issue, we propose an exploration of various mainstream machine learning models for CSO predictions. Additionally, we introduce a novel approach known as "inversion of neural networks" to bridge the gap between prediction and optimization. This innovative method enables us to use a single neural network for both CSO prediction and optimization tasks, resulting in a significant enhancement in terms of computational efficiency. The accuracy of our predictive approach has been validated through simulation results. In terms of optimization performance, it provides similar outcomes to the genetic algorithm, while significantly improving computational speed.

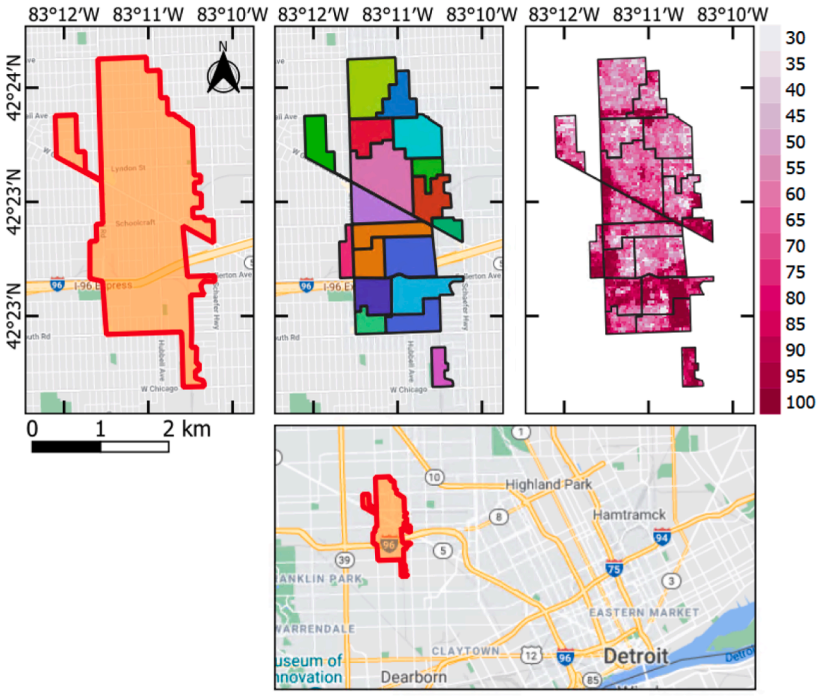
1. Introduction

A substantial portion of urban regions in Europe and North America continues to rely on combined sewer systems (Lund et al., 2020). The combined sewer system collects rainwater runoff, domestic sewage, and industrial wastewater in the same pipe (Jean et al., 2018). During heavy stormwater events, combined sewer overflows (CSOs) occur when the incoming flow exceeds the capacity of a sewer system (Gu et al., 2017; Zhao et al., 2017). When this occurs, untreated sewage and industrial wastewater discharge directly to nearby streams, rivers, and other water bodies. Consequently, this would result in significant environmental concerns (Brokamp et al., 2017; Botturi et al., 2021), and threaten public health (Ten Veldhuis et al., 2010; Gasperi et al., 2012; García et al., 2017). Recognized as one of major contributor to water pollution (Gooré Bi et al., 2015), the United States Environmental Protection Agency (EPA) has developed a set of policies and initiated long-term control plans aimed at minimizing the adverse environmental consequences associated with this issue (Moffa, 1997; Tao et al., 2017). Multiple approaches to minimize the impact of CSOs have been investigated by previous researchers, including investigations into Green/Gray Infrastructure (Cohen et al., 2012; Mancipe Muñoz, 2015; Tavakol-Davani et al., 2016; McGarity et al., 2017; Jean et al., 2021), Integrated Planning (Zukovs & Marsalek, 2004; Autixier et al., 2014; Fu et al., 2019; Matthews et al., 2000, and Smart Control Strategies (Rathnayake & Anwar, 2019; Bachmann-Machnik et al., 2021; Lund et al., 2020; Van Der Werf et al., 2023). In recent years, Smart Control Strategies have gained popularity owing to their cost-effectiveness (Lund et al., 2018; Zhang et al., 2018).

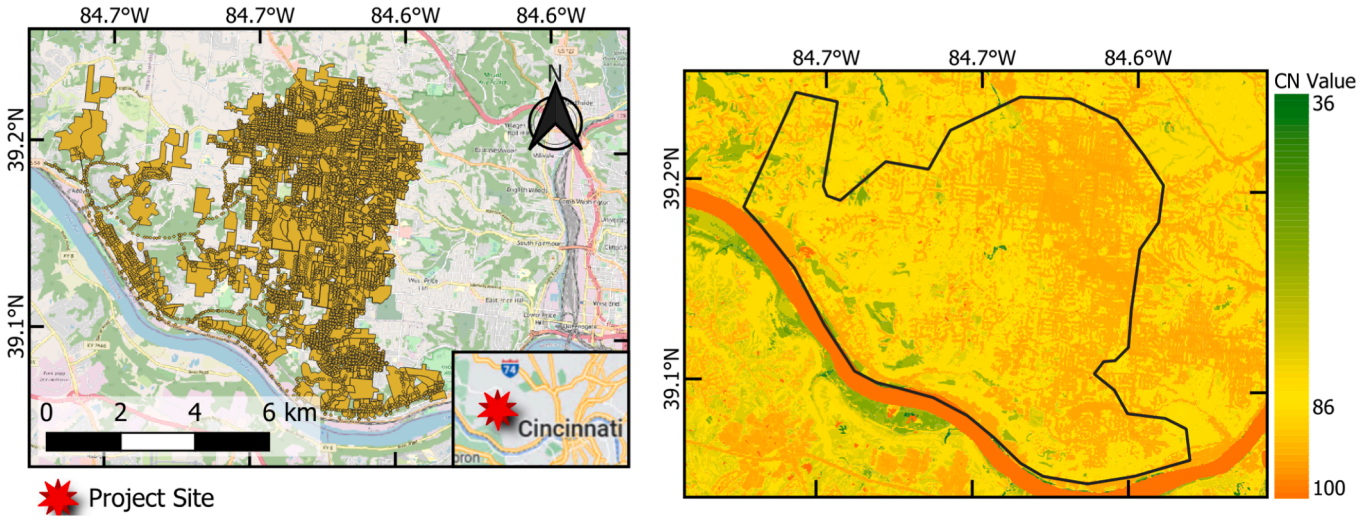
Throughout the past decades, the control algorithm has undergone substantial evolutionary progressions with regards to control strategy.

* Corresponding author.
E-mail address: zyin005@fiu.edu (Z. Yin).

https://doi.org/10.1016/j.jhydrol.2023.130515
Received 7 August 2023; Received in revised form 30 October 2023; Accepted 2 November 2023
Available online 21 November 2023



(a)



(b)

Fig. 1. Schematic of the study area, sub-catchment division, and impervious percentages. (a) Puritan Fenkell/Seven-Mile Collection System; (b) Metropolitan Sewer District of Greater Cincinnati.

In the past, heuristic control is the major control strategy that was widely used in the engineering field (Box, 1978; Fuchs et al., 1997; Fuchs & Beeneken, 2005; Gavrilas, 2010; Kroll et al., 2018; Van Der Werf et al., 2022). However, the implementation of hierarchical control heavily relies either on the expertise and practical knowledge of the specific sewer system or on optimization techniques in order to determine the appropriate set-points for actuators (Schütze et al., 2002; Lund et al., 2018). Thus, a few drawbacks are unavoidable when using heuristic control. Firstly, the determination of actuators set-points are typically based on the synthetic data derived from regression analysis, therefore, their fixed configuration settings restrict their adaptability to diverse rainfall-runoff events (Tian et al., 2022a). Furthermore, the experience and practical knowledge for heuristic control are significantly varied in each individual case. Consequently, the generalization of heuristic control to diverse catchments becomes challenging, particularly in cases where substantial variations exist among the catchments (Van Der Werf et al., 2022).

Model predictive control, as another control strategy, demonstrates superior adaptability and generalization capabilities (Pleau et al., 2005; Mollerup et al., 2016; Lund et al., 2018; Tian et al., 2022a). All model

predictive control methods need three essential elements: forecasting precipitation, predictive (nowcasting) model, and optimization algorithms (Sadler et al., 2019; Tian et al., 2022a). In terms of the border scope of optimal flood control, multiple previous researches have investigated the various combinations of these three elements. To name a few noteworthy examples, Genetic Algorithms with HEC-RAS (Leon et al., 2020; Albo-Salih et al., 2022), Pattern Search with HEC-RAS (Leon et al., 2021), Genetic Algorithms with HEC-HMS (Tang et al., 2020), Genetic Algorithms with SWMM (Sadeghi et al., 2022), Evolutionary Algorithm with SWMM (Yazdi, 2019), among others. In the context of minimizing combined sewer overflows, a few notable efforts have been made employing this technique (Zhao et al., 2017; Rathnayake & Anwar, 2019; Peng et al., 2021; El Ghazouli et al., 2022; Li et al., 2022). Over the past few years, the field of machine learning has experienced a significant surge in popularity within both research and practical applications. Some previous studies have employed diverse machine learning models on CSOs prediction (Maltbie et al., 2021; Rosin et al., 2021; Yin et al., 2022; Bakhshipour et al., 2023), and applied reinforcement learning approach on CSOs optimization problems (Lund et al., 2020; Mullapudi et al., 2020; Balla et al., 2022; Tian et al., 2022a; Tian et al., 2022b; Yin et al., 2023; Zhang et al., 2023).

However, there are several research gaps that need to be addressed. Firstly, predicting CSOs poses a unique challenge as it differs from other time series problems characterized by long histories of continuous patterns (e.g. water stage prediction is riverine system). In the CSO problem, the CSO rate remains zero for approximately 99 % of the time and experiences sudden spikes during a few hours before returning to zero. This atypical pattern makes it difficult for standard time series prediction techniques, involving data collection and model construction, to accurately capture CSO behavior. Previous research has tried to overcome this issue by shortening the prediction time (Rosin et al., 2021; Yin et al., 2022; Bakhshipour et al., 2023), but this approach significantly reduces the applicability of the model in real-world scenarios. Secondly, CSOs only occur during heavy rainfall events, making it challenging to collect a sufficient amount of training data from historical events. In this study, we only focus on the rainfall induced CSO. The infiltration induced CSO is considered as another topic (Liu et al., 2018; Su et al., 2020). Many previous control strategy studies have relied on hypothetical or regression rainfall data with return periods to build the training dataset, but relying solely on hypothetical rainfall data can compromise the accuracy of predictions during real rainfall events. Integrating both historical event data and hypothetical rainfall data into the training dataset presents difficulties, as they often do not align in the same dimension. Thus, a novel approach is required to effectively combine both types of data for training the CSO prediction model. Thirdly, it's important to note that all the previous research has focused on testing their method in very small and highly simplified sewer systems. In the real world, sewer systems are much more complex. Therefore, the prediction and optimization performance of these methods used in past research cannot be guaranteed for real-world applications. Last but most importantly, the current control methods struggle to keep balance between accuracy and the total computational time required for optimization. The utilization of full numerical methods as predictive models provides high accuracy but proves inefficient in terms of time, especially when applied to real-world complex combined sewer systems. Conversely, the use of reduced-order models typically delivers faster results but often falls short in terms of accuracy. To the best of the author's knowledge, all current methods still rely on conventional model predictive control frameworks, which are built upon non-gradient-based optimization algorithms and numerical models. While this framework can indeed provide remarkably accurate outcomes, it is extremely computational inefficient. Currently, there is no existing method that can effectively bridge the information gap between simulation and optimization because gradient-based optimizers cannot be directly integrated with either full or reduced-order numerical simulations. Therefore, it becomes imperative to establish a bridge between predictive model and optimization frameworks to minimize CSOs and address all other optimization challenges.

The main objectives of this paper are to address the previously mentioned research gaps comprehensively. Firstly, we aim to develop a machine learning framework capable of utilizing both historical and hypothetical data. In this paper, we have demonstrated that our framework have the ability to perform long-term predictions of the total CSO volume over the entire duration of a rainfall event. Additionally, the model exhibited high prediction accuracy in the real world existing complex combined sewer systems. Secondly, we proposed a novel approach called "inversion of neural networks" to bridge the machine learning prediction and optimization. The details of "inversion of neural networks" are shown in section 2.6. Through this method, we demonstrate the high accuracy and extreme computation efficiency of the proposed framework, allowing for real-time updating of gate positions based on the latest weather forecast. This bridging of machine learning prediction and optimization promises significant feasibility and effectiveness in practical applications.

2. Methodology

2.1. Study area

To prove our framework has capability to predict and optimize the real-world complex combined sewer systems, we selected two existing combined sewer systems within the United States as the study area in this paper. The first combined sewer system is Puritan Fenkell/Seven-Mile Collection System, which exhibits a more standardized dimension in terms of isolated sewer system. The second selected study area for our research is the Metropolitan Sewer District of Greater Cincinnati, which represents an intricate and complex sewer system within a large urban setting. The schematic of the study area, sub-catchment division, and impervious percentages for both Puritan Fenkell/Seven-Mile Collection System and Metropolitan Sewer District of Greater Cincinnati are shown in Fig. 1 (a) and 1 (b), respectively.

Detroit (42.3314° N, 83.0458° W) stands as Michigan's largest and most populous city, and it also holds the distinction of being the largest city in the United States on the border with Canada. On the northwest side of Detroit lies the Puritan Fenkell/Seven-Mile Collection System (PFSMC), which is owned by the Great Lakes Water Authority. The Puritan Fenkell/Seven-Mile Collection System is a relatively isolated system comprising two CSO treatment facilities. These two facilities are specifically designed to handle minor storms, offering screening, settling, skimming, and disinfection processes to ensure water quality before discharging excess water into the local river during more severe storms.

Cincinnati (39.1031° N, 84.5120° W) is a city in Ohio who lies on the northern banks where the Licking and Ohio rivers meet. On the west side of the city, you can find the Metropolitan Sewer District of Greater Cincinnati. This system is responsible for collecting and transporting flows from different combined/separate sewer areas to the Muddy Creek wastewater treatment plant (WWTP) during minor storms. However, during heavy storms, there are 16 Combined Sewer Overflow (CSO) outfalls that can directly release untreated water into the Ohio River.

2.2. Data generation

The training data for our framework can be obtained from either historical monitoring or numerical simulation. Ideally, using historical monitoring data holds greater potential for enhancing prediction accuracy because all numerical models rarely achieve 100 % accuracy in simulating all scenarios. However, monitoring the flow and water depth at a single specific location for several months can incur costs amounting to thousands of dollars so that makes long-time monitoring the entire sewer system economically unfeasible. Thus, we employ Environmental Protection Agency Storm Water Management Model (EPA SWMM) to

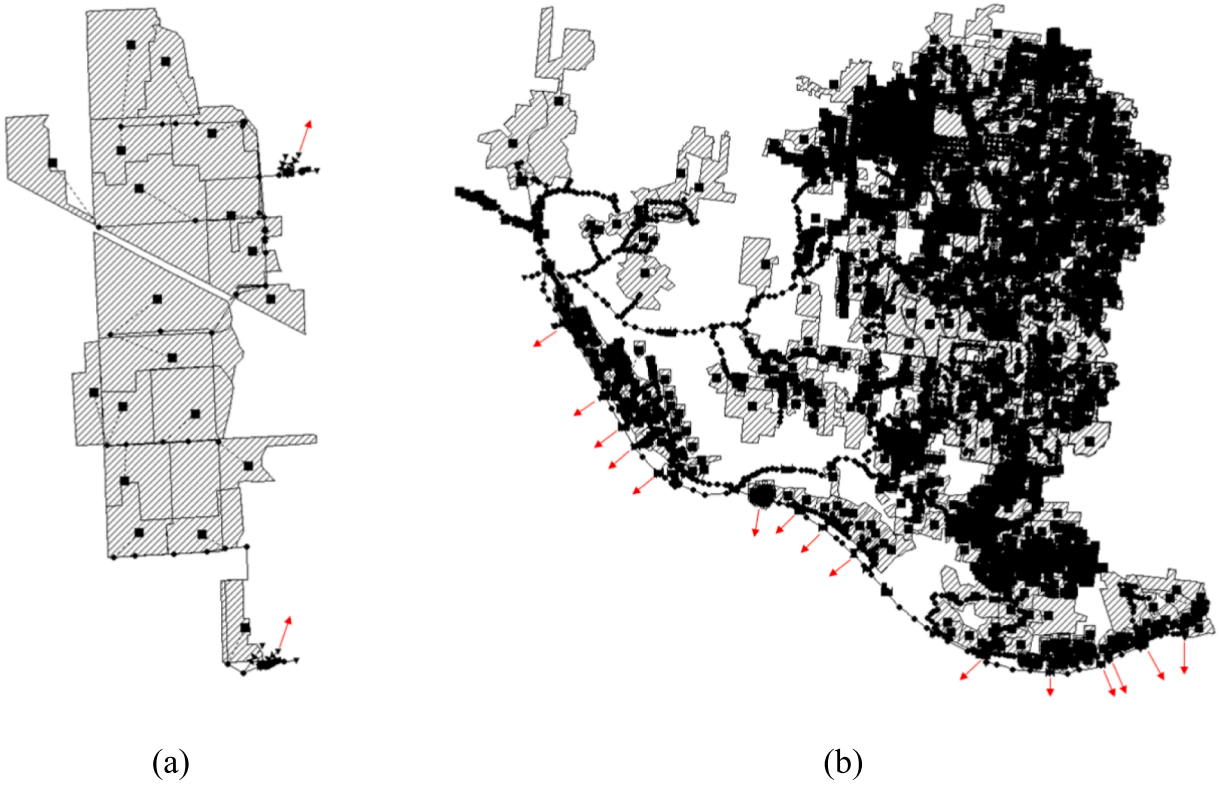


Fig. 2. Schematic of study area in SWMM model. (a) Puritan Fenkell/Seven-Mile Collection System; (b) Metropolitan Sewer District of Greater Cincinnati.

generate the required data in this study. The schematic of the study area in the SWMM model is shown in Fig. 2. For the convenience purpose, the Puritan Fenkell/Seven-Mile Collection System will be referred to as "C-town" and the Metropolitan Sewer District of Greater Cincinnati will be referred to as "D-town" for the rest sections in this paper. The CSO location and direction can be found at red arrow.

The C-town SWMM model is relatively simpler. It consists of 19 sub-catchments, 37 junctions, 6 outfalls, 4 storage units, 38 conduits, 3 pumps, 7 orifices, and 4 weirs. Among the 6 outfalls, two are directly connected to the CSO treatment facilities, while the remaining 4 outfalls possess the potential to release untreated wastewater into the local river. There are currently four existing weirs in place to prevent CSO occurrences during regular stormwater events. These weirs not only serve the purpose of preventing CSO, but they also provide us with a convenient means of tracking the CSO rate because the CSO rate at each location is actually equal to the corresponding weir flow rate. Conversely, the D-town SWMM model is significantly more complex. It consists of 2592 sub-catchments, 5147 junctions, 31 outfalls, 64 storage units, 6099 conduits, 11 pumps, 268 orifices, and 34 weirs. In the D-town system, out of the total 31 outfalls, 16 are connected to the CSO treatment facilities, leaving the remaining 15 with the potential to discharge untreated wastewater into the local river. Similar to C-town, all 15 of these locations are equipped with existing weirs that effectively prevent CSO events during regular stormwater events. Thus, we can employ the same method to track the CSO rate in D-town as well. Due to the complexity of our domain, it is not feasible to include all the necessary details in this text. Therefore, we have uploaded both SWMM model files on GitHub for reference. You can find them at the following address: [https://github.com/ZedaYin/Predicting-and-Minimizing-Combined-Sewer-Overflows].

In this study, we used both historical rainfall data and synthetic rainfall data in this study. Regarding the historical rainfall data, we have acquired records from rain gauges spanning the past 20 years. However, due to the presence of the existing weirs in both systems, CSOs have only occurred during instances of heavy rainfall events. As a result, this greatly reduces the effective training dataset, with our outputs having a value of zero approximately 99 % of the time. To encompass more useful simulation scenarios, we incorporated synthetic rainfall data obtained from the local Intensity Duration Frequency (IDF) analysis. In accordance with the guidelines provided by the City of Detroit Water and Sewerage Department Manual, the local rainfall characteristics in C-town align with the NOAA Atlas 14 Precipitation-Frequency Atlas and can be represented by equation (1). Similarly, as indicated by the City of Cincinnati Division of Stormwater Management, the IDF equation for the D-town area can be expressed as equation (2). Both agencies recommend utilizing the Natural Resources Conservation Service (NRCS) Midwest-Southeast (MSE) type 3 to obtain the 24-hour continuous rainfall distribution.

$$i = \frac{38.416T^{0.208}}{(12.3258 + D)^{0.841}} \tag{1}$$

$$i = e^{C_1 + C_2 \ln(D) + C3(\ln(D))^2} \tag{2}$$

where i denotes rainfall intensity in inch per hour, T represents return period in years, and D stands for the rainfall duration in minute, the coefficient C_1, C_2, and C_3 can be obtained from the table provided in the manual.

One of the main objectives of this framework involves the optimization stage. Consequently, our predictive machine learning models must be capable of providing responses to changes in the control units. Manually inputting various sets of time-series data into the SWMM graphical user interface (GUI) can be a time-consuming process. To address this issue, we employed PySWMM, a Python software package, which allowed us to efficiently create and manipulate the time-series control unit states (such as orifice gate states) in our model. In order to expand the training dataset and incorporate variations in the controlled gates, we utilized a uniform distribution, ranged from zero

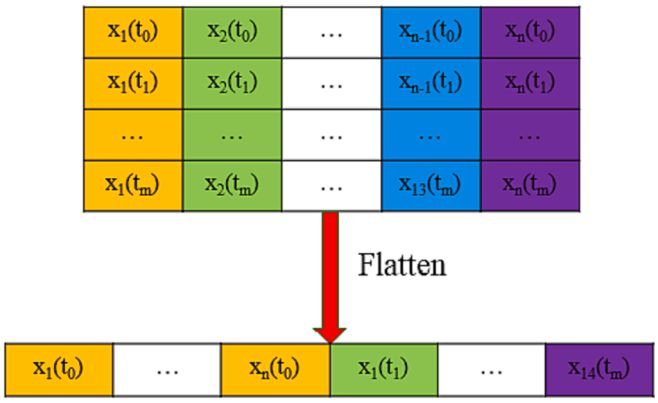


Fig. 3. The preprocessing and reconstruction of input variables.

(fully closed) to 1 (fully open), to generate time-series data for each individual controlled orifice gate. This approach allowed us to incorporate a wider range of scenarios and increase the diversity of our training data.

2.3. Data preprocessing

To ensure a precise representation of the physics and practical applicability of the framework in real-world scenarios, we integrated the exact same input variables used by the SWMM model into our machine learning framework. These input variables only consist of rainfall intensities, outfall water depth, and gate states. Also, we assumed that initial conditions for both CSO system are empty. In this study, we also assume all the existing structures are undamaged. The damage of structure induced CSO could be found on the other related work (Liu et al., 2021).

Typically, in machine learning time series forecasting problems, incorporating a historical window of past data is common for achieving more accurate results. However, we discovered that including such a history window can disrupt predictions during the initial timesteps when the control unit undergoes changes, as observed by Shi et al. (2023). Moreover, unlike other time-series forecasting projects, such as predicting tide elevation, our dataset is more like event-based due to the occurrence of Combined Sewer Overflows (CSOs) only during heavy stormwater events. Therefore, the information contained within the past window at each location appears to lack significance or relevance, and we did not use this past window in this study.

There are two mainstream preprocessing techniques that are commonly used in the machine learning field. The first method is to organize all covariates and targets at each timestep into a 2D structure, and subsequently flattening the entire table into a 1D representation that preserves the temporal features. The second technique involves incorporating time indexes (e.g., date, time, day of week, date of moth, etc.) into the 2D time-feature table. Time indexes can offer significant advantages when predicting obvious periodic features such as tides or power consumption. However, the second method requires a long continuous time window in both training and testing, which is hard to achieve in our event-based dataset. Furthermore, our CSO prediction is not periodic at all. Thus, the first technique is the best solution in this paper. The details of input variable preprocessing and reconstruction are illustrated in Fig. 3.

The output variables follow a similar approach to the first technique, but with slight differences due to specific reasons. Firstly, the CSO rates exhibit a distinct pattern during heavy stormwater events, resembling a striking shape with large values occurring within a short period (1 min to 1 h), while remaining zero the rest of the time. This unique characteristic poses a significant challenge for all machine learning models. Additionally, during the optimization stage, most optimization

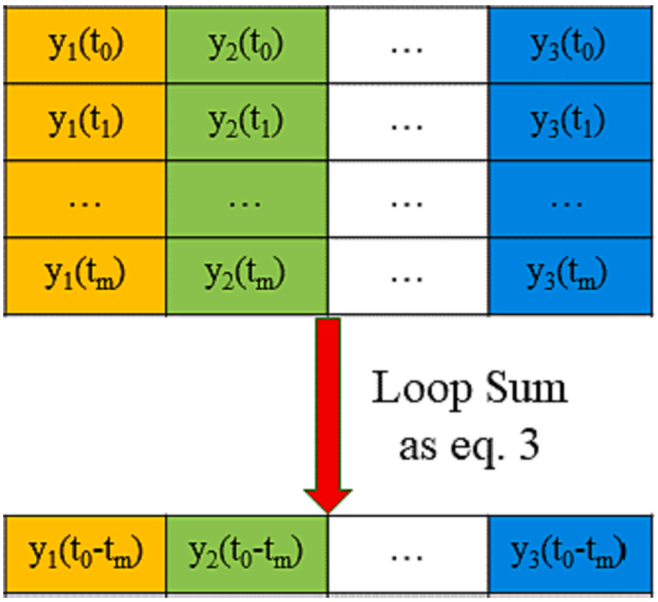


Fig. 4. The preprocessing and reconstruction of output variables.

algorithms are designed to work with a single value as the final reward/fitness/objective value. While the time series pattern of the output may provide more information for human interpretation, it does not necessarily offer additional information that can significantly benefit the optimization algorithms. Thus, it would be ideal for this study to utilize the total CSO volume at each location and total nodal flooding volume at the entire system over the entire event as our output features. As indicated in equation (3), the total CSO volume can be calculated as the integral of the CSO rate over time. In a discrete time-point representation, it can be expressed using a summation format by applying the trapezoidal rule. Since our output sampling time is a constant value, the equation can be simplified to its final form, incorporating only the CSO rates Q_i and Q_{i+1} within the loop summation. In such a case, the details of input variables preprocessing, and reconstruction can be illustrated in Fig. 4.

$$Vol = \int_0^t Q dt \cong \sum_{i=0}^{t-1} \frac{1}{2}(Q_i + Q_{i+1})^* \Delta t_i = \Delta t \sum_{i=0}^{t-1} \frac{1}{2}(Q_i + Q_{i+1}) if\, \Delta t_i is constant \tag{3}$$

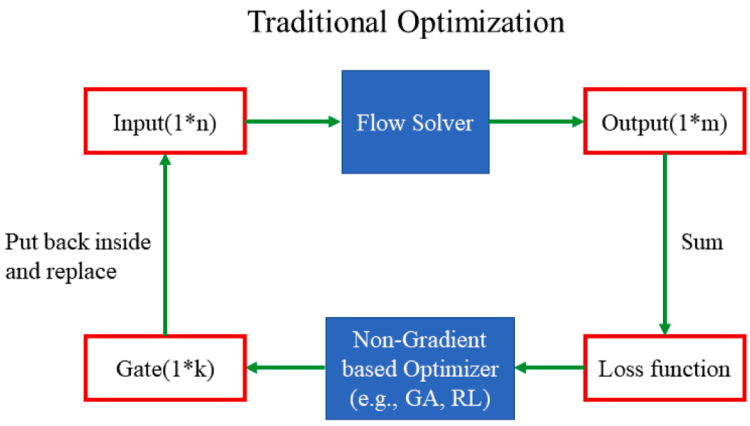
where Q represents for CSO rate or nodal flooding rate, t stands for time.

2.4. Machine learning model architecture

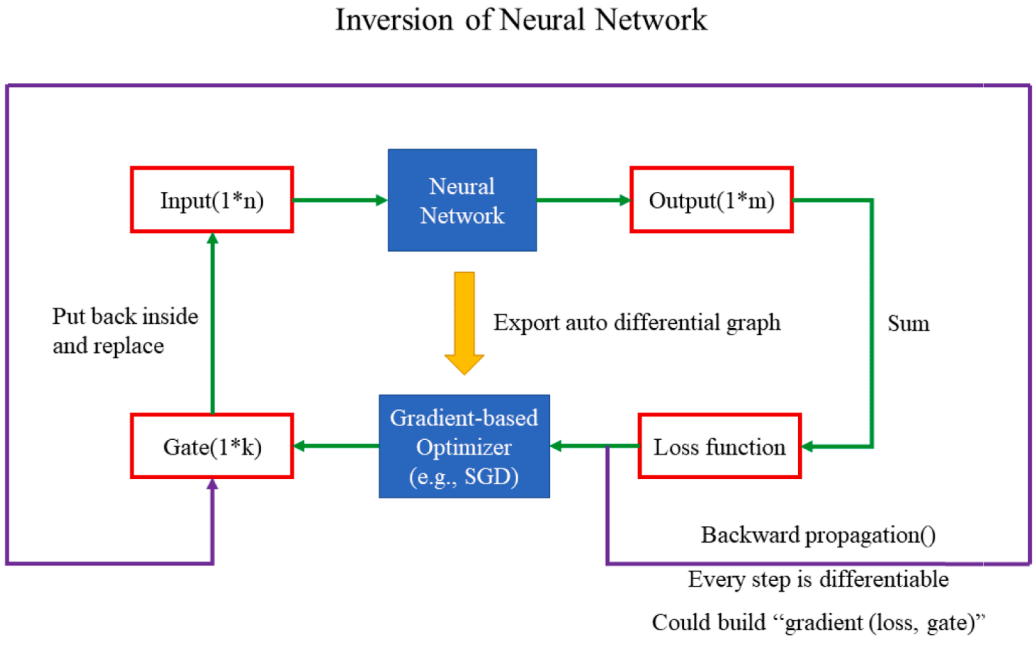
In this paper, we explored a wide range of popular backbone structures, including tree-based models (Xgboost), purely deep neural network models (MLP), recurrent-based models (GRU and LSTM), convolution-based models (CNN and ResNet), and attention-based models (Seq2Seq Attention). The detailed architecture of these machine learning models will be introduced supplementary material.

2.5. Performance metrics

To better measure the performance of our machine learning models, serval metrics are used in this paper. Mean absolute errors (MAE, shown in equation (4) and root mean square error (RMSE, shown in equation (5) are definitely most widely used in terms of comparing two arrays. MAE calculates the average absolute difference between the elements of the two arrays, while RMSE calculates the square root of the average squared difference between the elements of the two arrays. Additionally, R-squared (R^2, shown in equation (6) is also a widely known parameter,

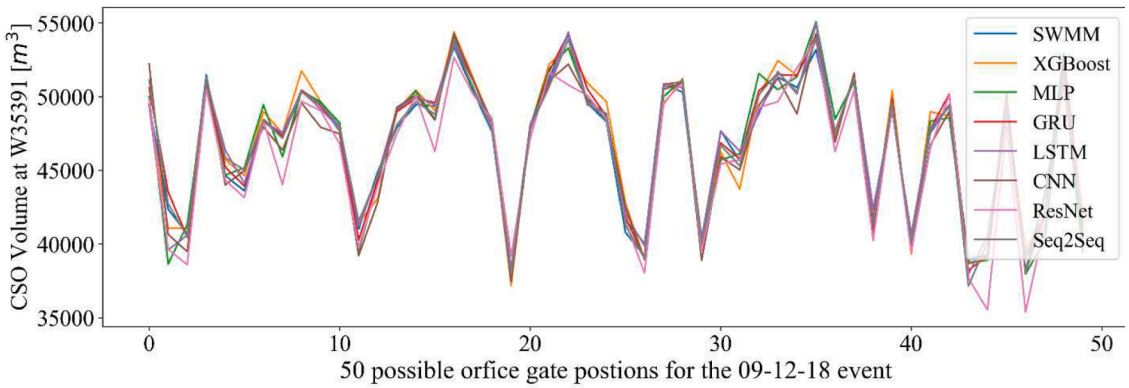


(a)

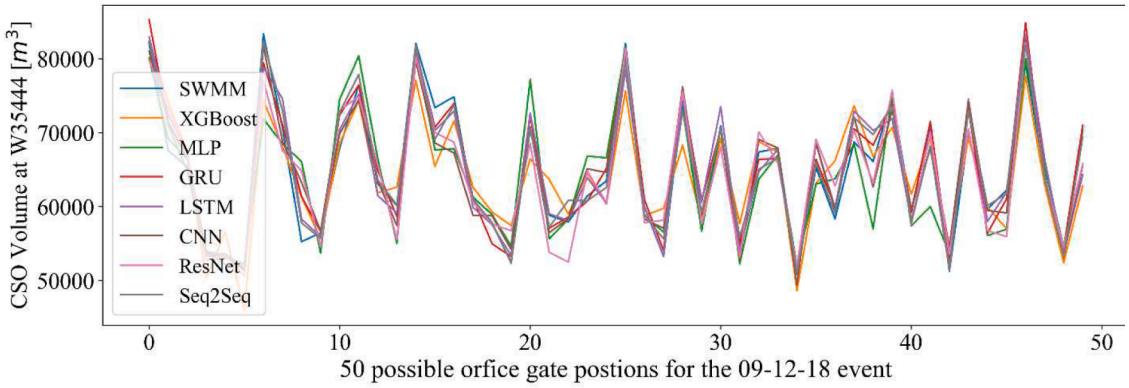


(b)

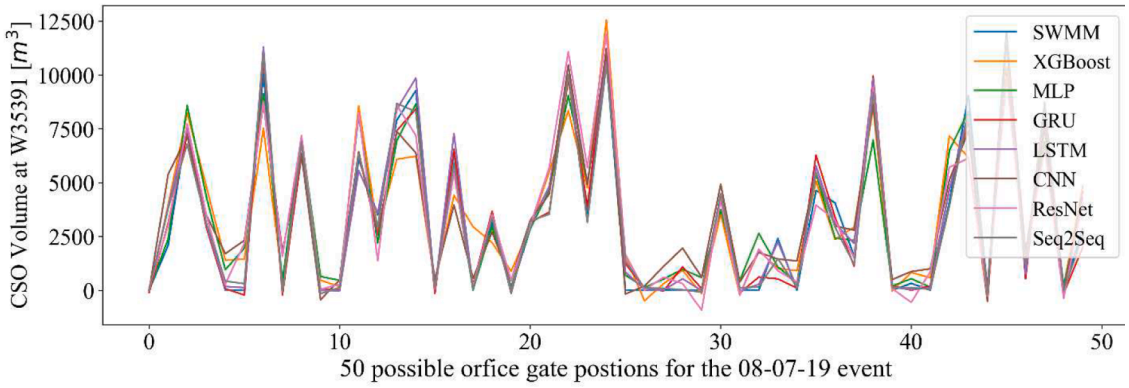
Fig. 5. Optimization Loop. (a) traditional structure; (b) Inversion of Neural network.



(a)



(b)



(c)

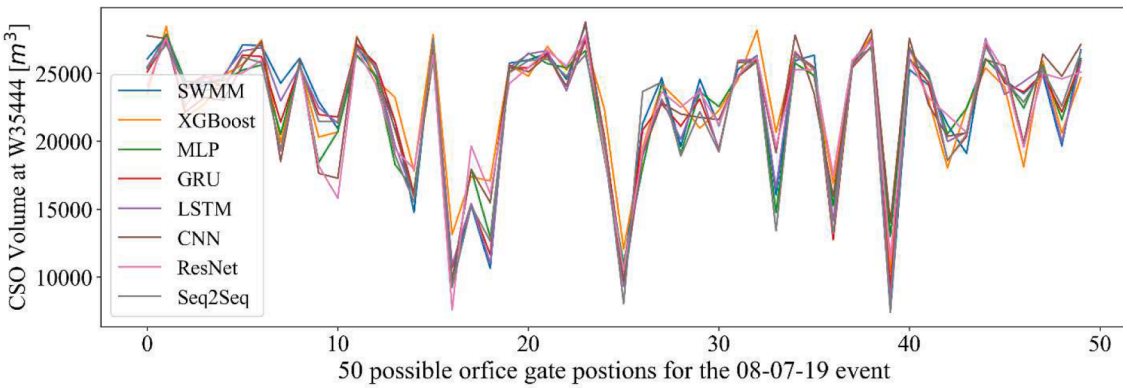
Fig. 6. CSO volume prediction for C-town case: (a) location of W35391 on September 12th, 2018 event; (b) location of W35444 on September 12th, 2018 event; (c) location of W35391 on August 7th, 2019 event; (d) location of W35444 on August 7th, 2019 event.

which measures the proportion of the variance in one array that is predictable from the other array. It provides an indication of how well the two arrays are correlated. Lastly, mean absolute percentage error (MAPE, shown in equation (7) is also a commonly used metrics in forecasting and time series analysis, which is used for measuring the percentage difference between the elements of the two arrays.

$$MAE = \frac{\sum_{i=1}^{n}(Y_{prediction} - Y_{true})}{n} \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{prediction} - Y_{true})^2}{n}} \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_{prediction} - Y_{true})^2}{\sum_{i=1}^{n}(Y_{prediction} - \overline{Y_{prediction}})^2} \tag{6}$$



(d)

Fig. 6. (*continued*).

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_{prediction} - Y_{true}}{Y_{true}}\right| \quad (7)$$

where $Y_{prediction}$ represents the model output, $\overline{Y_{prediction}}$ stands for the mean value of the model outputs, Y_{true} represents ground truth, $\overline{Y_{true}}$ stands for the mean value of the ground truth, and n means the number of total sample points.

2.6. Inversion for optimization

In the context of the optimization control problem, a closed-loop structure is essential. This allows optimization algorithms to obtain states from the environment/response surface and provide the best actions during each iteration. Fig. 5 (a) illustrates the classic optimization loop for minimizing the CSO problem. In traditional optimization approaches, taking our case as example, the flow solver (SWMM model) takes boundary conditions and all orifice gate positions as inputs to compute the CSO rate, which serves as the output. The objective function is then constructed, typically involving algebraic operations like summing or integrating the CSO rate to create a single value to deliver to the optimization algorithms. The optimization algorithm, such as the genetic algorithm mentioned, takes the calculated value, and provides the optimal gate positions for the flow solver input. It is important to note that the optimization results are usually limited to adjusting gate positions since most of the boundary conditions (e.g., rainfall, downstream water stage) cannot be optimized. Consequently, the value "k" in Fig. 5 (a) is typically smaller than the input dimension "n," and it is necessary to use code to properly place the gate position within the next iteration's input for the flow solver.

However, within the confines of this traditional optimization framework, only non-gradient optimization algorithms can be employed. This limitation arises because the entire process lacks the capability to trace differentiations, especially during the flow solver stage. Notably, one significant advantage of utilizing neural networks as environments or response surfaces to predict the CSO volume or rate, which has been widely overlooked in prior research, is that it renders the entire loop differentiable. By achieving differentiability throughout the loop, we gain the ability to utilize a gradient-based optimizer to replace the non-gradient optimization algorithms. Gradient-based optimizers are typically renowned for their speed and efficiency. This process, referred to as the inversion of the neural network, is represented in our case by the structure depicted in Fig. 5 (b).

The purple line depicted in Fig. 5 (b) represents the traced gradient. To calculate the gradient, it is necessary to track it from the optimizer's input all the way to its output. The process of achieving this through code differs slightly based on the automatic differential computation graph method being used. The dynamic computation graph approach is highly recommended due to its convenience.

Table 1
Metrics for the C-town Case.

	MAE [m^3]	RMSE [m^3]	R^2	MAPE
Xgboost	478.51	1255.38	0.993889	3.64 %
MLP	411.30	1138.87	0.994922	3.32 %
GRU	256.14	756.66	**0.997768**	1.91 %
LSTM	251.79	791.11	0.997595	1.89 %
CNN	419.08	1034.06	0.995804	2.73 %
ResNet	457.44	1120.32	0.994929	3.39 %
Seq2Seq	**242.41**	**761.96**	**0.997768**	**1.77 %**

3. Result and discussion

3.1. Prediction performance analysis

3.1.1. Prediction performance analysis on C-town

We tested the model's performance using two historical rainfall events for C-town: one starting on September 12th, 2018, and the other on August 7th, 2019. Since minimizing CSO is one of our major objectives, our machine learning models were designed to accurately predict the total CSO volume at each potential CSO location. As mentioned above, the C-town case has 2 potential CSO output: W35391 and W35444. To evaluate the model's effectiveness and better visualize the results, we employed random gate positions during testing for each event. To be more specifically, we performed machine learning predictions on 50 random gate positions for the rainfall event starting on September 12th, 2018, and these results are depicted in Fig. 6 (a) and (b). Similarly, for the rainfall event beginning on August 7th, 2019, we conducted predictions on another 50 random gate positions, and the outcomes are displayed in Fig. 6 (c) and (d). Fig. 6 illustrates that all of our models demonstrated commendable accuracy and effectively captured the underlying trends. This also suggests that our models have the capability to assess which set of random gate positions performs better in terms of minimizing CSOs.

To enhance the model performance evaluation, we further generated 2000 different sets of gate positions for each event, resulting in a total of 4000 sets. All four metrics (MAE, RMSE, R^2, and MAPE) were computed based on these 4000 sets. Table 1 illustrates that the models exhibit commendable performance, with MAPE ranging from 3.64 % to 1.77 % and R2 from 0.993889 to 0.997768. Considering the magnitude of total CSO volume as shown in Fig. 6, the MAE values ranging from 478.51 to 242.41 can also be regarded as indicative of good accuracy. In our C-town test case, the Seq2Seq model with attention demonstrates superior

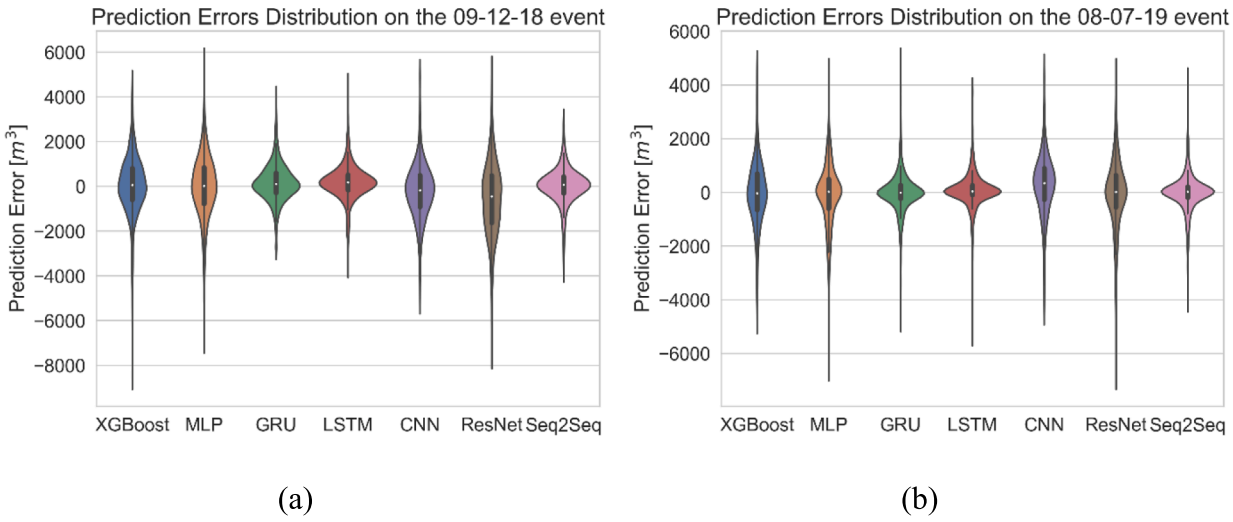


Fig. 7. Prediction error density distribution for C-town: (a) rainfall event starts on 09–12-18; (b) rainfall event starts on 08–07-19.

performance across all four metrics employed.

Violin plots are commonly utilized to illustrate the distribution of errors, showcasing the density distribution of errors alongside key statistics such as the median, peak value, 25th percentile, and 75th percentile (see the top and bottom of the small black box inside the violin shape). Fig. 7 presents the violin plots representing the error density distribution for the C-town model's output. Among the models assessed, namely GRU, LSTM, and Seq2Seq with attention, they stand out due to their better error distribution. These models exhibit a much more concentrated error within the −2000 to 2000 m^3 range, and their peak errors are notably smaller compared to the other models. Considering the total CSO volume in the two test events are ranging from 0 to 90000 m^3, many of our machine learning models, particularly the Seq2Seq with attention model, achieve peak errors between −4000 to 4000 m^3, which can be considered a high level of accuracy.

3.1.2. *Prediction performance analysis on D-town*

Similarly, we conducted tests and visualized the model performance using 50 random gate positions in D-town case. Since D-town case is much larger and more complex, along with 16 potential CSO outputs within the model, led us to present results from a single rainfall event that occurred on November 4th, 2017 in Fig. 8. Due to the length limit of context, we selected four CSO locations randomly to illustrate the results. The CSO volume prediction results for D-town are shown in Fig. 8. Similar to C-town case, all of our models exhibited impressive accuracy and successfully captured the underlying trends. Based on the testing and visualization from both C-town and D-town cases, we can conclude that our machine learning models have the ability to capture complex patterns and predict CSO volume with high precision.

Similar to the pervious C-town case, we employed 2000 distinct sets of gate positions for two rainfall events to evaluate the model accuracy metrics in the D-town scenario. The specific metrics for the D-town case are presented in Table 2. The proposed models consistently deliver outputs with excellent accuracy, as evidenced by the Mean Absolute Percentage Error (MAPE) ranging from 2.84 % to 1.86 % for all models. The performance of each model has undergone significant changes compared to the C-town case. The CNN model exhibits higher accuracy, whereas the LSTM performance drops a lot in our D-town case. The GRU model outperforms all other models in the D-town case, achieving the highest prediction accuracy across all four metrics used. In short, determining which model will provide optimal performance across all existing combined sewer systems is challenging. However, based on this study, it is advisable to begin with a recurrent-based model or a Seq2Seq with attention model, as they have demonstrated better performance.

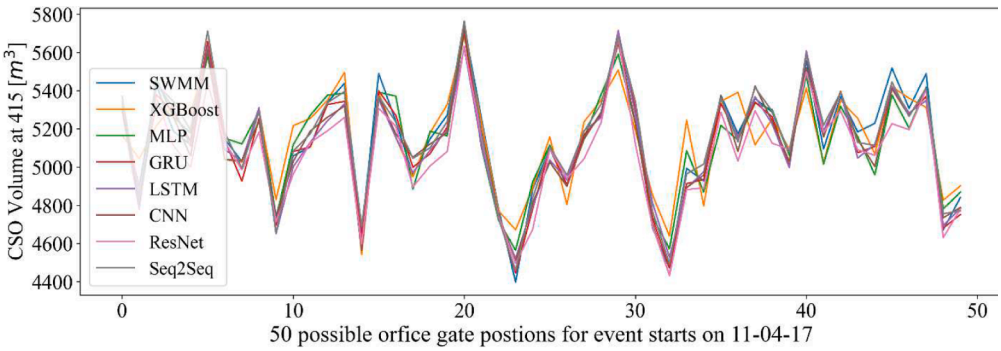
Fig. 9 presents the violin plots representing the error density distribution for the D-town model's output. Compared to the previous case, all models exhibit significantly reduced peak errors in our D-town case. The GRU, Xgboost, and Seq2Seq models emerge as the top three performers in terms of prediction error distribution. However, it appears that the Seq2Seq model has a tendency to slightly overestimate the results. Surprisingly, despite being the most lightweight model, Xgboost outperforms several training-heavy models in our D-town scenario, such as LSTM and ResNet.

Based on all the aforementioned tests and metrics, we can conclude that all of our machine learning models are capable of predicting the total CSO volume throughout the entire events based on the provided boundary and initial conditions. However, when evaluating their performance in C-town and D-town cases, we recommend considering the GRU, LSTM, or Seq2Seq models for higher accuracy. Additionally, the Xgboost model, as our lightweight and fastest, can also provide reliable results in some cases. However, the Xgboost model can only perform the CSO volume prediction because it does not support the time series type of outputs. Despite ResNet model's success in other time series forecasting problems (Chen and Cong, 2022; Yu et al., 2022; Silva et al., 2023), its performance in our two test cases is not ideal.

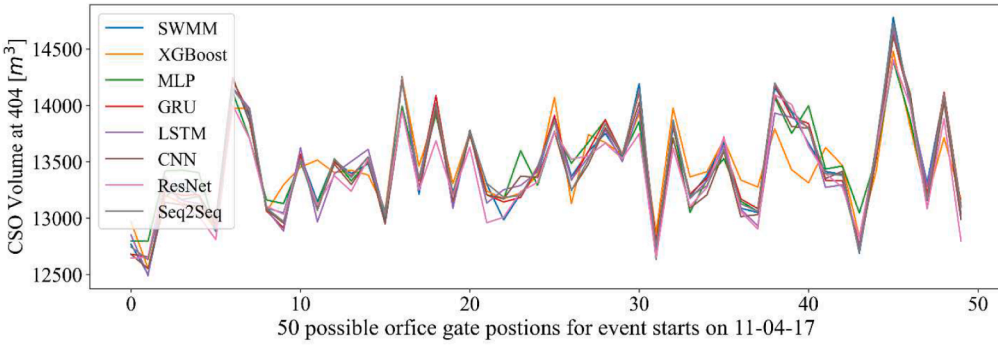
3.2. *Total CSOs volume VS time series CSOs rates prediction*

There has been a limited amount of research on predicting CSO problems. Among this limited research, the majority has been dedicated to the prediction of time series CSO rates (Maltbie et al., 2021; Rosin et al., 2021; Yin et al., 2022). However, predicting CSO rates has two significant reasons compared to predicting the total volume, as done in this paper.

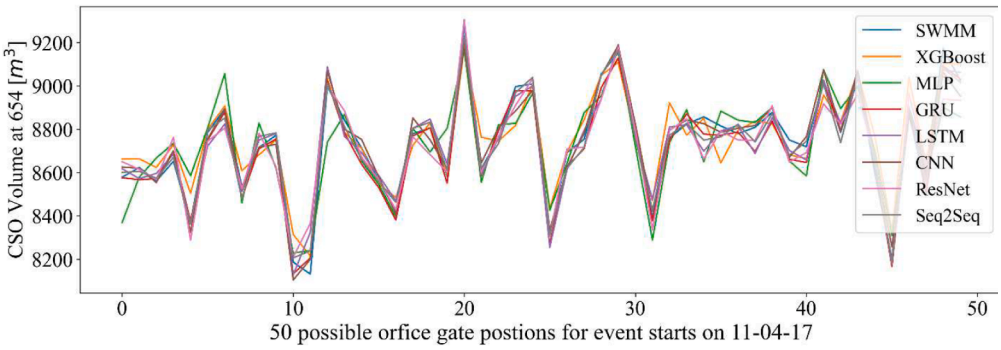
The first major reason is related to accuracy. We experimented and tested all of our proposed models, except Xgboost, for time series CSO rate prediction by using the C-town case. The results, presented in Table 3, demonstrate that most models perform better in predicting CSO volume. This outcome is reasonable because predicting CSO rates involves an additional dimension compared to CSO volume prediction. Furthermore, the nature of CSO rates, being more like spikes or sudden bursts instead of continuous patterns like many other time series problems, adds complexity to the prediction task. As a result, the accuracy of time series CSO rate prediction is lower than that of total CSO volume prediction. The Seq2Seq model stands out as the top performer in time series CSO rate prediction due to its unique outcome method. This approach involves predicting the next output based on the input and the previous output, allowing it to make predictions one element at a time.



(a)



(b)



(c)

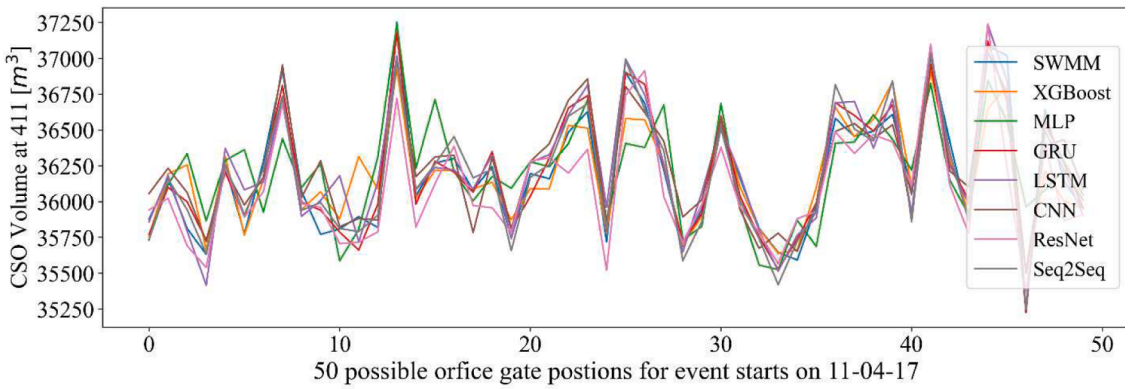
Fig. 8. CSO volume prediction for D-town case: (a) location of CSO weir 415; (b) location of CSO weir 404; (c) location of CSO weir 654; (d) location of CSO weir 411.

In contrast, other models provide all results in one go, contributing to the high dimensionality problem. The Seq2Seq model's ability to handle predictions step-by-step partially addresses this challenge.

The second reason is that higher dimensionality output is not advantageous for most popular genetic algorithms, including genetic algorithm, pattern search, and reinforcement learning. While time series CSO rates can offer additional information to human observers, such as identifying peak flow occurrences, it does not provide the same level of intuitiveness to optimization algorithms. All these popular optimization algorithms are typically designed to maximize or minimize a single value derived from specific loss, objective, or reward functions. Using CSO volume values instead of CSO rate arrays barely has an impact on the accuracy during the optimization stage. However, using total CSO volumes as the output significantly enhances the computational efficiency of the entire optimization loop. This improvement occurs because providing CSO volumes as output allows the computer to avoid handling the complexities of time series arrays within the user-defined loss, objective, or reward functions.

3.3. Comparison with conventional optimization framework and verification for inversion

Applying the inversion that is introduced in section 2.6, we can obtain the optimal series of gate positions and corresponding optimal CSO results at potential CSO locations. In every test conducted, we consistently assigned a weight of 10 to total flooding and a weight of 1 to each CSO location. This decision was made because the potential for untreated water flowing through the nodes (manholes) may cause more



(d)

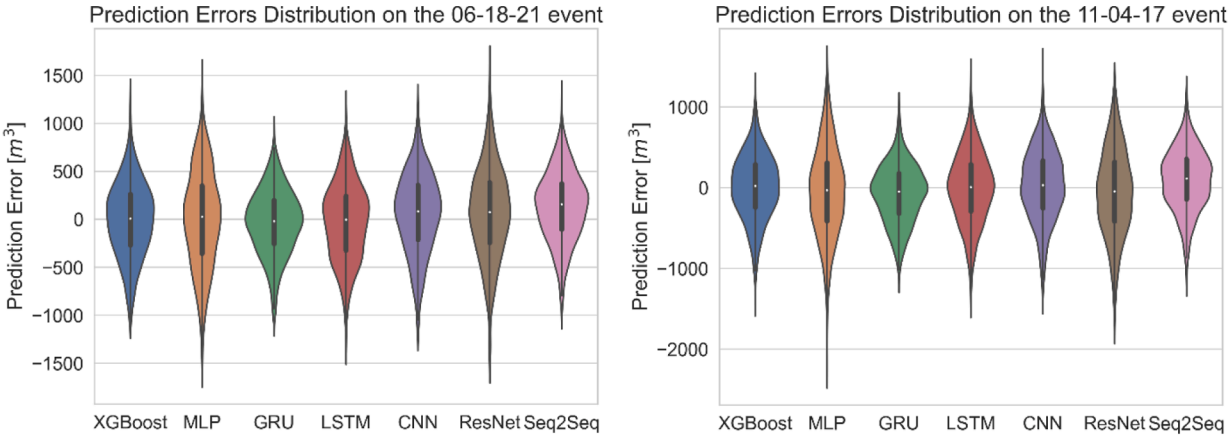
Fig. 8. (*continued*).

Table 2
Metrics for the D-town Case.

	MAE [m^3]	RMSE [m^3]	R^2	MAPE
Xgboost	116.52	196.18	0.941258	1.90 %
MLP	192.35	310.54	0.840966	3.55 %
GRU	**103.87**	**188.67**	**0.945844**	**1.86 %**
LSTM	131.24	224.43	0.925754	2.31 %
CNN	134.07	229.86	0.922002	2.39 %
ResNet	161.70	265.02	0.899441	2.84 %
Seq2Seq	110.97	198.98	0.940232	1.98 %

significant damage. As a result, we aimed to avoid this nodal flooding scenario during the optimization stage.

Some researchers and engineers might not trust the results since all the computation process is completed in a single neural network. Thus, we have employed the most conservative verification method to confirm the effectiveness of our inversion approach. The optimal series of gate positions obtained through the inversion method were utilized as gate inputs in the SWMM model. By employing the SWMM model, we were able to calculate the corresponding CSO output at each location. Furthermore, we established a baseline case for comparison purposes. In this case, we employed the widely-used traditional optimization algorithm, genetic algorithm, and integrated it with the SWMM model as the optimization environment. This allowed us to assess the performance of



(a) (b)

Fig. 9. Prediction error density distribution for D-town: (a) rainfall event starts on 06–18-21; (b) rainfall event starts on 11–04-17.

Table 3
Comparison of total CSOs volume prediction and time series CSOs rates prediction.

	MAE of CSO Volume Prediction [m^3]	MAE of CSO rate Prediction [m^3]	Difference	RMSE of CSO Volume Prediction [m^3]	RMSE of CSO rate Prediction [m^3]	Difference
MLP	411.30	443.48	+7.82 %	1138.87	1162.45	+2.07 %
GRU	256.14	265.32	+3.58 %	756.66	830.89	+9.81 %
LSTM	251.79	278.11	+10.45 %	791.11	841.56	+6.38 %
CNN	419.08	443.65	+5.86 %	1034.06	1043.16	+0.88 %
ResNet	457.44	649.88	+42.07 %	1120.32	1506.35	+34.45 %
Seq2Seq	242.41	259.84	+7.19 %	761.96	744.87	−2.24 %

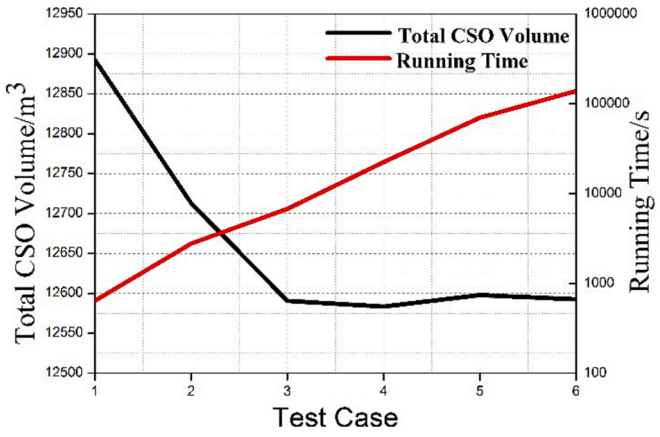


Fig. 10. Convergence test for genetic algorithm.

our inversion method against the conventional approach.

The performance of the genetic algorithm is influenced by its hyperparameters. To evaluate this impact, we conducted a hyperparameter convergence test as shown in Fig. 10. The convergence test was carried out on the C-town case at August 7th, 2019 event. The number of generations are set as 100, 200, 400, 800, 1500, 2000, in the test case 1 to 6, respectively. The number of parents mating are set as 4, 4, 8, 8, 16, 32, and the number of solution are set as 8, 16, 32, 32, 64, 128, in the test case 1 to 6, respectively. By examining the convergence behavior under these different settings, we found that the settings in test case 4 is a good choice in terms of both accuracy and computation efficiency.

The comparison of conventional optimization (GA + SWMM) and inversion of neural network on the C-town case at August 7th, 2019 event is shown in Table 4. As illustrated in Fig. 10, the rainfall event on August 7th, 2019, was moderately heavy in comparison to the event on September 12th, 2018. Consequently, all the methods listed in Table 4 were successful in minimizing total CSO to zero during the event. However, the notable distinction lies in the computation time, where the inversion method significantly outperforms the conventional optimization framework. The computation time for the inversion method ranged from 0.65 to 61.53 s, making real-time optimization control feasible. This enables the optimal gate positions to be promptly updated based on the latest weather forecast in real-time scenarios.

Table 5 presents a comparison between the conventional optimization (GA + SWMM) and the neural network inversion approach applied to the D-town case on November 4th, 2017. In this scenario, D-town has a much larger catchment than C-town and comprises 16 potential CSO outlets directly discharging into the Ohio River. By using the conventional optimization framework, the total CSO volume in the entire system was minimized to 135,306 m^3. The optimization also significantly minimized CSO volumes at each location, compared to the visualization shown in Fig. 6. Alternatively, the inversion methods achieved a total CSO volume within the range of 131,226 to 140,520 m^3, which shows a similar performance to the conventional approach. However, it's important to note that the computation time required by the conventional optimization framework was exceedingly long, which takes up to 13.8 days. Such a long processing time makes the entire optimization framework meaningless in the real-world applications. The slowness was mainly attributed to the numerical model's time-consuming responses, as it needed to run thousands of times to compute numerous gate position setups and to update CSO states. The numerical model needs to take around four to five minutes to compute this complex D-town cases, making it is unfeasible for practical optimization usage. In contrast, our inversion method successfully addressed this problem, completing the entire optimization process with a much short time. The computation time ranged from 3.24 s to 794.27 s, depending on the complexity of machine learning models. Such a significant improvement in computation time even makes real-time optimal control feasible, as the inversion method can continuously update the best solution based on weather forecast updates.

4. Conclusion

In conclusion, we developed and evaluated seven mainstream machine learning models for rainfall induced CSOs prediction in two

Table 4
Comparison of conventional optimization and inversion of neural network on the C-town case at August 7th, 2019 event.

	GA	MLP	GRU	LSTM	CNN	ResNet	Seq2Seq
CSOs on W35391 [m^3]	0	0	0	0	0	0	0
CSOs on W35444 [m^3]	0	0	0	0	0	0	0
Total CSOs Volume [m^3]	0	0	0	0	0	0	0
Time [s]	12628.61	0.65	19.41	40.87	0.71	2.28	61.53

Table 5
Comparison of conventional optimization and inversion of neural network on the D-town case at November 4th, 2017 event.

	GA	MLP	GRU	LSTM	CNN	ResNet	Seq2Seq
CSO on 412 [m^3]	6592	3905.	4269	8060	2894	2436	3758
CSO on 415 [m^3]	4042	3463	3957	2698	6034	6235	3246
CSO on 404 [m^3]	10,970	8726	10,420	8932	8437	9874	8251
CSO on 403 [m^3]	755	792	936	214	832	135	642
CSO on 402 [m^3]	2052	600	2481	490	5327	627	3028
CSO on 406 [m^3]	7545	6353	8851	8075	6748	9112	8907
CSO on 654 [m^3]	7919	7906	7796	8487	8812	9322	9584
CSO on 408 [m^3]	21,190	23,742	24,286	21,540	23,069	20,411	24,172
CSO on 541 [m^3]	1700	1463	1721	1286	2565	1425	1503
CSO on 223 [m^3]	10,593	8996	8611	9327	7930	10,621	8187
CSO on 411 [m^3]	28,613	32,116	31,328	28,022	32,560	33,198	27,443
CSO on 414 [m^3]	1814	1768	2004	2197	1854	1714	2021
CSO on 415B [m^3]	13,325	12,163	13,745	11,697	13,840	17,676	15,654
CSO on 413 [m^3]	2215	1976	1738	1826	2810	2698	1993
CSO on 416 [m^3]	15,973	17,249	16,460	18,572	16,807	13,926	14,972
Total CSOs Volume [m^3]	135,306	131,226	138,605	131,423	140,520	139,411	133,361
Time [s]	1,196,444	74.77	392.32	559.91	3.24	5.17	794.27

existing combined sewer systems. All the models have demonstrated high accuracy when evaluated on our test set. Additionally, we introduced and successfully implemented the inversion of neural networks method to minimize the CSO. The key findings can be summarized as follows:

- Our machine learning models demonstrated consistently high accuracy in predicting CSOs. The mean absolute percentage error ranged from 1.77 % to 3.64 % for Puritan Fenkell/Seven-Mile Collection System and from 1.86 % to 2.84 % for Metropolitan Sewer District of Greater Cincinnati.
- The study revealed that the prediction of total CSOs volume is more accurate than the prediction of time series CSO rates across all tested machine learning models. Moreover, the total CSOs volume prediction exhibited a better fit in the optimization process, making it a more suitable choice for optimization purposes.
- The introduced inversion of neural networks method approach can provide the high-quality optimal results comparable to the conventional GA optimization method. Furthermore, the inversion process using neural networks exhibited a substantial performance advantage over the traditional GA optimization framework in terms of speed. The completion times for the entire optimization process ranged from as little as 3.24 s to a maximum of 794.27 s, whereas the GA approach required a lengthy 13.5 days to complete. This remarkable speed makes real-time optimal control feasible, as the inversion method can continuously update the best solution based on weather forecast updates.

CRediT authorship contribution statement

Zeda Yin: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Yasaman Saadati:** Methodology, Software, Formal analysis, Writing – original draft. **Arturo S. Leon:** Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **M. Hadi Amini:** Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Linlong Bian:** Validation, Visualization, Writing – review & editing. **Beichao Hu:** Conceptualization, Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. CBET-2203292. Also, the authors are grateful to the anonymous reviewers for their constructive comments, which helped to significantly improve the quality of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2023.130515.

References

Albo-Salih, H., Mays, L.W., Che, D., 2022. Application of an optimization/simulation model for the real-time flood operation of river-reservoir systems with one-and two-dimensional unsteady flow modeling. Water 14 (1), 87.

Autixier, L., Mailhot, A., Bolduc, S., Madoux-Humery, A.S., Galarneau, M., Prévost, M., Dorner, S., 2014. Evaluating rain gardens as a method to reduce the impact of sewer overflows in sources of drinking water. Sci. Total Environ. 499, 238–247.

Bachmann-Machnik, A., Brüning, Y., Ebrahim Bakhshipour, A., Krauss, M., Dittmer, U., 2021. Evaluation of combined sewer system operation strategies based on highly resolved online data. Water 13 (6), 751.

Bakhshipour, A. E., Koochali, A., Dittmer, U., Haghighi, A., Ahmad, S., & Dengel, A. (2023). A Bayesian Generative Adversarial Network (GAN) to Generate Synthetic Time-Series Data, Application in Combined Sewer Flow Prediction. arXiv preprint arXiv:2301.13733.

Balla, K.M., Bendtsen, J.D., Schou, C., Kallesøe, C.S., Ocampo-Martinez, C., 2022. A learning-based approach towards the data-driven predictive control of combined wastewater networks–An experimental study. Water Res. 221, 118782.

Botturi, A., Ozbayram, E.G., Tondera, K., Gilbert, N.I., Rouault, P., Caradot, N., Fatone, F., 2021. Combined sewer overflows: A critical review on best practice and innovative solutions to mitigate impacts on environment and human health. Crit. Rev. Environ. Sci. Technol. 51 (15), 1585–1618.

Box, F., 1978. A heuristic technique for assigning frequencies to mobile radio nets. IEEE Trans. Veh. Technol. 27 (2), 57–64.

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

Chen, X., Cong, D., 2022. Application of improved algorithm based on four-dimensional ResNet in rural tourism passenger flow prediction. Journal of Sensors 2022, 1–8.

Cohen, J.P., Field, R., Tafuri, A.N., Ports, M.A., 2012. Cost comparison of conventional gray combined sewer overflow control infrastructure versus a green/gray combination. J. Irrig. Drain. Eng. 138 (6), 534–540.

El Ghazouli, K., El Khatabi, J., Soulhi, A., Shahrour, I., 2022. Model predictive control based on artificial intelligence and EPA-SWMM model to reduce CSOs impacts in sewer systems. Water Sci. Technol. 85 (1), 398–408.

Fu, X., Goddard, H., Wang, X., Hopton, M.E., 2019. Development of a scenario-based stormwater management planning support system for reducing combined sewer overflows (CSOs). J. Environ. Manage. 236, 571–580.

Fuchs, L., Beeneken, T., 2005. Development and implementation of a real-time control strategy for the sewer system of the city of Vienna. Water Sci. Technol. 52 (5), 187–194.

Fuchs, L., Beeneken, T., Spönemann, P., Scheffer, C., 1997. Model based real-time control of sewer system using fuzzy-logic. Water Sci. Technol. 36 (8–9), 343–347.

García, J.T., Espín-Leal, P., Vigueras-Rodriguez, A., Castillo, L.G., Carrillo, J.M., Martínez-Solano, P.D., Nevado-Santos, S., 2017. Urban runoff characteristics in combined sewer overflows (CSOs): Analysis of storm events in southeastern Spain. Water 9 (5), 303.

Gasperi, J., Zgheib, S., Cladière, M., Rocher, V., Moilleron, R., Chebbo, G., 2012. Priority pollutants in urban stormwater: Part 2–Case of combined sewers. Water Res. 46 (20), 6693–6703.

Gavrilas, M., 2010. October). Heuristic and metaheuristic optimization techniques with application to power systems. In: In Proceedings of the 12th WSEAS International Conference on Mathematical Methods and Computational Techniques in Electrical Engineering, p. (p. 9)..

Gooré Bi, E., Monette, F., Gasperi, J., Perrodin, Y., 2015. Assessment of the ecotoxicological risk of combined sewer overflows for an aquatic system using a coupled "substance and bioassay" approach. Environ. Sci. Pollut. Res. 22, 4460–4474.

Gu, X., Liao, Z., Zhang, G., Xie, J., Zhang, J., 2017. Modelling the effects of water diversion and combined sewer overflow on urban inland river quality. Environ. Sci. Pollut. Res. 24, 21038–21049.

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

Jean, M. È., Morin, C., Duchesne, S., Pelletier, G., & Pleau, M. (2021). Optimization of Real-Time Control With Green and Gray Infrastructure Design for a Cost-Effective Mitigation of Combined Sewer Overflows. Water Resources Research, 57(12), e2021WR030282.

Kroll, S., Weemaes, M., Van Impe, J., Willems, P., 2018. A methodology for the design of RTC strategies for combined sewer networks. Water 10 (11), 1675.

Leon, A.S., Tang, Y., Qin, L., Chen, D., 2020. A MATLAB framework for forecasting optimal flow releases in a multi-storage system for flood control. Environ. Model. Softw. 125, 104618.

Leon, A.S., Bian, L., Tang, Y., 2021. Comparison of the genetic algorithm and pattern search methods for forecasting optimal flow releases in a multi-storage system for flood control. Environ. Model. Softw. 145, 105198.

Li, X., Hou, J., Chai, J., Du, Y.E., Han, H., Fan, C., Qiao, M., 2022. Multisurrogate Assisted Evolutionary Algorithm-Based Optimal Operation of Drainage Facilities in Urban Storm Drainage Systems for Flood Mitigation. J. Hydrol. Eng. 27 (11), 04022025.

Liu, T., Su, X., Prigiobbe, V., 2018. Groundwater-sewer interaction in urban coastal areas. Water 10 (12), 1774.

Liu, T., Ramirez-Marquez, J.E., Jagupilla, S.C., Prigiobbe, V., 2021. Combining a statistical model with machine learning to predict groundwater flooding (or infiltration) into sewer networks. J. Hydrol. 603, 126916.

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), e2019WR026272.

Lund, N.S.V., Falk, A.K.V., Borup, M., Madsen, H., Steen Mikkelsen, P., 2018. Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management. Crit. Rev. Environ. Sci. Technol. 48 (3), 279–339.

Maltbie, N., Niu, N., Van Doren, M., & Johnson, R. (2021, August). XAI tools in the public sector: A case study on predicting combined sewer overflows. In Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (pp. 1032-1044).

Mancipe Muñoz, N.A., 2015. Detention-based Green/Gray Infrastructure Framework to Control Combined Sewer Overflows. University of Cincinnati). Doctoral dissertation,.

Matthews, S., Bingham, D.R., Greenland, F., 2000. Combined Sewer Overflow Facilities Plan for the Cleveland Westerly District. In: Collection Systems Conference 2000. Water Environment Federation, pp. 544–555.

McGarity, A. E., Szalay, S., & Cohen, J. (2017). StormWISE model using green infrastructure to achieve Philadelphia's CSO volume reductions at minimum cost. In World Environmental and Water Resources Congress 2017 (pp. 334-344).

Moffa, P.E. (Ed.), 1997. The Control and Treatment of Combined Sewer Overflows. John Wiley & Sons.

Mollerup, A.L., Mikkelsen, P.S., Sin, G., 2016. A methodological approach to the design of optimising control strategies for sewer systems. Environ. Model. Softw. 83, 103–115.

Mullapudi, A., Lewis, M.J., Gruden, C.L., Kerkez, B., 2020. Deep reinforcement learning for the real time control of stormwater systems. Adv. Water Resour. 140, 103600.

Peng, Z., Jin, X., Sang, W., Zhang, X., 2021. Optimal Design of Combined Sewer Overflows Interception Facilities Based on the NSGA-III Algorithm. Water 13 (23), 3440.

Pleau, M., Colas, H., Lavallée, P., Pelletier, G., Bonin, R., 2005. Global optimal real-time control of the Quebec urban drainage system. Environ. Model. Softw. 20 (4), 401–413.

Rathnayake, U., Anwar, A.F., 2019. Dynamic control of urban sewer systems to reduce combined sewer overflows and their adverse impacts. J. Hydrol. 579, 124150.

Rosin, T.R., Romano, M., Keedwell, E., Kapelan, Z., 2021. A committee evolutionary neural network for the prediction of combined sewer overflows. Water Resour. Manag. 35, 1273–1289.

Sadeghi, S., Samani, J., Samani, H., 2022. Optimal Design of Storm Sewer Network Based on Risk Analysis by Combining Genetic Algorithm and SWMM Model. Amirkabir Journal of Civil Engineering 54 (5), 1903–1924.

Sadler, J.M., Goodall, J.L., Behl, M., Morsy, M.M., Culver, T.B., Bowes, B.D., 2019. Leveraging open source software and parallel computing for model predictive control of urban drainage systems using EPA-SWMM5. Environ. Model. Softw. 120, 104484.

Schütze, M., Campisano, A., Colas, H., Schilling, W., & Vanrolleghem, P. A. (2002). Real-time control of urban wastewater systems-where do we stand today?. In Global Solutions for Urban Drainage (pp. 1-17).

Shi, J., Yin, Z., Myana, R., Ishtiaq, K., John, A., Obeysekera, J., ... & Narasimhan, G. (2023). Deep Learning Models for Water Stage Predictions in South Florida. arXiv preprint arXiv:2306.15907.

Silva, A.Q.B., Gonçalves, W.N., Matsubara, E.T., 2023. DESCINet: A hierarchical deep convolutional neural network with skip connection for long time series forecasting. Expert Syst. Appl. 120246.

Su, X., Liu, T., Beheshti, M., Prigiobbe, V., 2020. Relationship between infiltration, sewer rehabilitation, and groundwater flooding in coastal urban areas. Environ. Sci. Pollut. Res. 27, 14288–14298.

Tang, Y., Leon, A.S., Kavvas, M.L., 2020. Impact of dynamic storage management of wetlands and shallow ponds on watershed-scale flood control. Water Resour. Manag. 34, 1305–1318.

Tao, J., Li, Z., Peng, X., Ying, G., 2017. Quantitative analysis of impact of green stormwater infrastructures on combined sewer overflow control and urban flooding control. Front. Environ. Sci. Eng. 11, 1–12.

Tavakol-Davani, H., Burian, S.J., Devkota, J., Apul, D., 2016. Performance and cost-based comparison of green and gray infrastructure to control combined sewer overflows. Journal of Sustainable Water in the Built Environment 2 (2), 04015009.

Ten Veldhuis, J.A.E., Clemens, F.H.L.R., Sterk, G., Berends, B.R., 2010. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water Res. 44 (9), 2910–2918.

Tian, W., Liao, Z., Zhi, G., Zhang, Z., & Wang, X. (2022a). Combined Sewer Overflow and Flooding Mitigation Through a Reliable Real-Time Control Based on Multi-Reinforcement Learning and Model Predictive Control. Water Resources Research, 58(7), e2021WR030703.

Tian, W., Liao, Z., Zhang, Z., Wu, H., & Xin, K. (2022b). Flooding and Overflow Mitigation Using Deep Reinforcement Learning Based on Koopman Operator of Urban Drainage Systems. Water Resources Research, 58(7), e2021WR030939.

Van Der Werf, J.A., Kapelan, Z., Langeveld, J., 2022. Towards the long term implementation of real time control of combined sewer systems: a review of performance and influencing factors. Water Sci. Technol. 85 (4), 1295–1320.

Van Der Werf, J.A., Kapelan, Z., Langeveld, J., 2023. Real-time control of combined sewer systems: Risks associated with uncertainties. J. Hydrol. 617, 128900.

Yazdi, J., 2019. Optimal operation of urban storm detention ponds for flood management. Water Resour. Manag. 33, 2109–2121.

Yin, Z., Zahedi, L., Leon, A. S., Amini, M. H., & Bian, L. A Machine Learning Framework for Overflow Prediction in Combined Sewer Systems. In World Environmental and Water Resources Congress 2022 (pp. 194-205).

Yin, Z., Leon, A. S., Sharifi, A., & Amini, M. H. Optimal Control of Combined Sewer Systems to Minimize Sewer Overflows by Using Reinforcement Learning. In World Environmental and Water Resources Congress 2023 (pp. 711-722).

Yu, X., Liu, Y., Sun, Z., Qin, P., 2022. Wavelet-based ResNet: A deep-learning model for prediction of significant wave height. IEEE Access 10, 110026–110033.

Zhang, D., Lindholm, G., Ratnaweera, H., 2018. Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. J. Hydrol. 556, 409–418.

Zhang, Z., Tian, W., Liao, Z., 2023. Towards coordinated and robust real-time control: a decentralized approach for combined sewer overflow and urban flooding reduction based on multi-agent reinforcement learning. Water Res. 229, 119498.

Zhao, W., Beach, T.H., Rezgui, Y., 2017. 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.

Zukovs, G., Marsalek, J., 2004. Planning and design of combined sewer overflow treatment. Water Quality Research Journal 39 (4), 439–448.