Optimal Control of Combined Sewer Systems to Minimize Sewer Overflows by Using Reinforcement Learning

Zeda Yin[1]; Arturo S. Leon[2]; Abbas Sharifi[3]; and M. Hadi Amini[4]

[1]Ph.D. Candidate, Dept. of Civil and Environmental Engineering, Florida International Univ., Miami, FL (corresponding author). Email: zyin005@fiu.edu
[2]Associate Professor, Dept. of Civil and Environmental Engineering, Florida International Univ., Miami, FL. Email: arleon@fiu.edu
[3]Ph.D. Student, Dept. of Civil and Environmental Engineering, Florida International Univ., Miami, FL. Email: asharifi@fiu.edu
[4]Assistant Professor, Knight Foundation School of Computing and Information Sciences, Florida International Univ., Miami, FL. Email: amini@cs.fiu.edu

ABSTRACT

A combined sewer system (CSS) collects rainwater runoff, domestic sewage, and industrial wastewater in the same pipe. The volume of wastewater can sometimes exceed the system capacity during heavy rainfall events. When this occurs, untreated stormwater and wastewater discharge directly to nearby streams, rivers, and other water bodies. This would threaten public health and the environment, contributing to drinking water contamination and other concerns. Minimizing sewer overflows requires an optimization method that can provide an optimal sequence of decision variables at control gates. Conventional strategies use classical optimization algorithms, such as genetic algorithms and pattern search, to find the optimal sequence of decision variables. However, these conventional frameworks are very time-consuming, and it is almost impossible to achieve near real-time optimal control. This paper presents a faster optimization framework by using a new optimal control tool: reinforcement learning. The environment (flow modeler) used in this paper is the numerical model: Environmental Protection Agency's Storm Water Management Model (EPA SWMM) to ensure the accuracy of environment response. The reward function is constructed based on the calculated water depth and overflow rate from SWMM. The process keeps minimizing the reward function to obtain the optimal flow release sequence at each controlled orifice gate. The combined sewer system (CSS) of the Puritan-Fenkell 7-mile facility in Detroit, MI, is chosen as the case study.

INTRODUCTION

A combined sewer system collects rainwater runoff, domestic sewage, and industrial wastewater in the same pipe. The volume of wastewater can sometimes exceed the system capacity during heavy rainfall events. When this occurs, untreated stormwater and wastewater discharge directly to nearby streams, rivers, and other water bodies. The combined sewer overflows would threaten public health and the environment, contributing to drinking water contamination and other concerns. With the increasing frequency and intensity of storms due to climate change (Ezer and Atkinson 2014; Miller and Hutchins 2017; Mounce et al. 2020; NASA 2017), the possibilities and risks of sewer overflow are also increasing. The cost of stormwater flooding in the United States is estimated at billions of dollars per year, including human body damage and property loss (Smith and Katz 2013).

Most current sewer infrastructure does not have the full capability to handle dynamic storm events and rapid changes in runoff volumes (Saliba et al. 2020). To retrofit the storm system, many researchers have shown that real-time active control could provide a promising solution (Mynett and Vojinovic 2009; Kerkez et al. 2016; Mullapudi et al. 2019). Real-time control showed significant performance improvement in flood mitigation compared to passive control (Jin et al. 2020). Using active control to minimize sewer overflows requires an optimization approach to obtain the optimal sequence of decision variables at control units. Conventionally, classical optimization algorithms, for example, genetic algorithms and pattern search, were adapted to minimize the objective function in order to achieve optimal control (Leon et al. 2020; Leon et al. 2021). However, these conventional frameworks are very time-consuming, and it is almost impossible to achieve near real-time optimal control.

Machine learning is a fast-rising strategy that can solve problems in an ultrafast manner. Many studies have adapted machine learning or deep learning to solve the water resource management problem in the past years (Sadler et al. 2018; Sajedi-Hosseini et al. 2018; Zhao et al. 2019; Abdalla et al. 2021; Adnan et al. 2021; Yin et al. 2022; Li et al. 2022). Reinforcement learning (RL), a type of machine learning that learns from trial and error, is an emerging approach to optimization problems. RL has been successfully adapted in many famous fields for years (AlphaGo, Automated driving) (Chen 2016; Sallab et al. 2017; Kiran et al. 2021). The RL does not know the optimal answer but keeps trying different actions. The actions are delivered to the environment, and RL tries to maximize the rewards that are based on the environment feedback. Many researchers have tried this promising approach in water resource engineering. RL has been used in multi-objective reservoirs in many studies, and the results show RL has better performance than conventional optimization algorithms (Pianosi et al. 2013; Delipetrev et al. 2017). Hajgató et al. (2020) adapted deep reinforcement learning (deep Q learning) for real-time pump optimization in water distribution systems. Several pieces of research have also shown that reinforcement learning shows superior performance when compared with passive control (rule-based on control law) in coastal area storm flood mitigation (Saliba et al. 2020; Bowes et al. 2021; Mullapudi and Kerkez 2018).

In this work, reinforcement learning is adapted to the real-world combined sewer system to minimize sewer overflows. The previous flood mitigation studies were only conducted in a very small ideal domain with only a few elements (one or two catchments with one or two junctions and control units). The key factor that the previous literature used to achieve significantly better performance is that they added a non-existing hypothetical huge capacity retention pond into the system. That hypothetical pond can provide a lot more space to temporarily restore the water volume during extreme stormwater events. This work will be the first study to test the RL performance in a real-world combined sewer system with no hypothetical hydraulic structure. Every structure was already there in the real world. Thus, the performance of RL is more realistic and instructive. Furthermore, this work employs a novel and emerging deep reinforcement learning algorithm: Deep Deterministic Policy Gradient (DDPG). The DDPG performance will be compared with other reinforcement learning algorithms in this paper. Lastly, to test and compare its overflow mitigation performance, conventional rule-based control was also conducted and reported in this paper.

METHODOLOGY

Study area

Detroit (42.3314° N, 83.0458° W) is the largest and most populous city in the U.S. state of Michigan and the largest U.S. city on the United States–Canada border. Flood prediction and mitigation for urban areas are crucial. The Puritan Fenkell/Seven-Mile Collection System (PFSMC), is located on the northwest side of Detroit, Michigan. The sewer prototype for this work is the Puritan Fenkell/Seven-Mile Collection System (PFSMC), located on the northwest side of Michigan. The PFSMC system, owned by the Great Lakes Water Authority (Detroit, Michigan), is a relatively isolated system comprising two CSO treatment facilities. These facilities were designed to capture and treat small storms and provide screening, settling, skimming, and disinfection before overflowing into the local river for more significant storms. The schematic of the study region, sub-catchment division, and impervious percentages is shown in Fig. 1a.

The combined sewer system (PFSMC in this paper) simulation was carried out by The U.S. Environmental Protection Agency's Stormwater Management Model (EPA SWMM) version 5.1. The schematic of the SWMM model is shown in Fig. 1b, c, and d. Due to the complexity of the system, the SWMM model consists of 19 sub-catchments, 37 junctions, 6 outfalls, 4 storage units, 38 conduits, 3 pumps, 7 orifices, and 4 weirs. The two most downstream sides (located at the upper right and lower right in Fig. 1c) were defined as outfalls, and most hydraulic structures, including pumps, orifices, and weirs, are located upstream of the outfall. After running the simulation for multiple return periods of rainfall, all the key locations that potentially have nodal flooding are marked in Fig. 1b. The hydraulic information in these locations will be used to construct the reward function in the following sections.

In any combined sewer system, wastewater may discharge to the other water bodies if the volume of water in the system exceeds a certain amount. Under this situation, untreated water may go directly to the other water bodies that cause CSOs. Fig. 1c shows the outfall information at the two most downstream locations. Water can potentially flow both to the wastewater treatment plant and rivers, which caused the CSOs. More specifically, on the upper right downstream side, JCT 3090 is connected to the wastewater treatment plant for the sewer water, and Outfall 65 is connected to the outside river system. Similarly, on the lower right downstream side, JCT 2112 is connected to the wastewater treatment plant for the sewer water, and Outfall 54 is connected to the outside river system. In PFSMC, two water divider systems, consisting of weirs, orifice, pumps, and storage units, are used to control the CSOs. However, the untreated water could still go through outfalls 65 and 54 under heavy stormwater conditions.

Finding the control units is always the first step in any optimal control problems. In our test case: Puritan Fenkell/Seven-Mile Collection System (PFSMC), 7 gated orifices are ideally to serve as the control units. The location of these 7 gated orifices is shown in Fig. 1d. Three gated orifices (O35281B, O1360, O1360B) are majorly used to control the CSOs in the upper half, and the other three (O35291B, O3530, O3530B) are majorly used to control the CSOs in the lower half. The Orifice (O35262) sits on the only pipeline that connects the upper and lower half of the system so that can significantly help to distribute the water amount in the upper and lower half of the system.

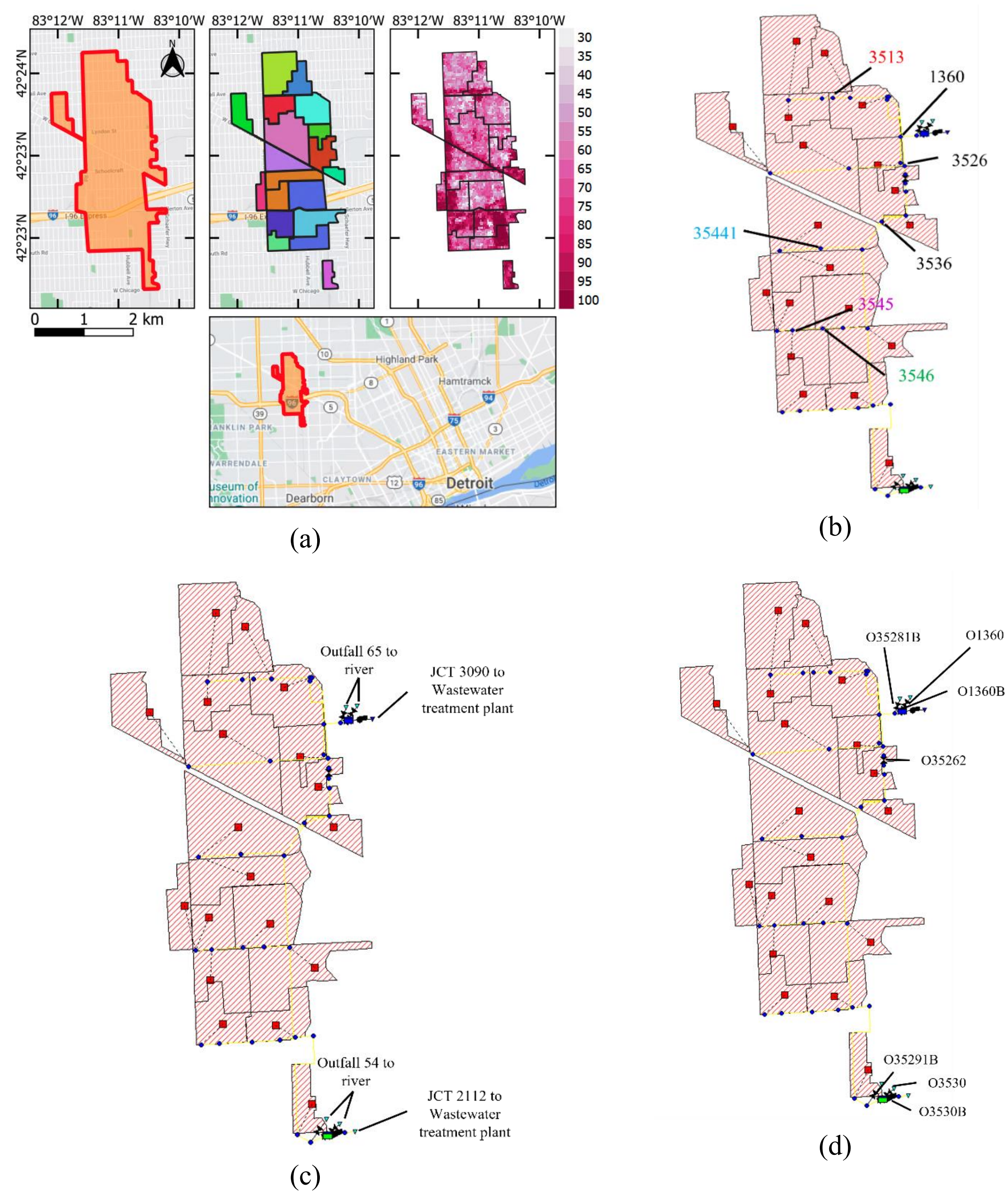


Figure 1. (a) Schematic of sewer prototype study area (top left), sub-catchment division (top middle), and impervious percentages (top right); (b) Key positions that potentially have nodal flooding in the SWMM model; (c) Detailed information on the downstream; (d) The location of the gated orifice.

SWMM model setup and rainfall data

The accuracy of environment response is critical in any optimization problem. Using forecasting machine learning models could extremely speed up the process, however, it is also risky in terms of reliability because the accuracy of environment response is hard to guarantee. Thus, the conventional numerical model, SWMM, is used as the environment (flow modeler) in this paper. To get the most achievable accurate simulation results, our SWMM model is calibrated and validated by ten events occurring between 9/22/2016 and 6/24/2017. To test the performance of reinforcement learning, the SWMM model is run from 6/2/2018 to 6/5/2018, however, the rainfall data used in the model is the hypothetical 25-year return period rainfall rather than the real historical rainfall.

The rainfall data were generated by using the Detroit local Intensity Duration Frequency (IDF) curve. According to the City of Detroit Water and Sewerage Department (2018), the local rainfall characteristics correspond to the NOAA Atlas 14 Precipitation-Frequency Atlas of the United States. Thus, the local IDF is expressed by Eq. (1). Natural Resources Conservation Service (NRCS) Midwest-Southeast (MSE) type 3 was used to get the continuous rainfall distribution based on the instruction of [30].

$$i = \frac{38.4164T^{0.2082}}{(12.3258 + D)^{0.8405}} \tag{1}$$

where i denotes rainfall intensity in inches per hour, T represents the return period in years, and D stands for the rainfall duration in min.

Control strategies

For any optimal control problems, a closed loop with the signal transfer of actions and states is necessary. The schematic of the optimization framework using reinforcement learning is shown in Fig. 2. At the beginning of the optimization, reinforcement learning will first initiate a set of actions (position of control units) and pass it to the environment (pyswmm). The environment will run the model based on the position of control units given from the reinforcement learning, and provide the observation (hydraulic information including depth, flow rate, etc.). The pre-defined reward function will take the observation results to construct a magnitude reward value, and the reward value will be sent back to reinforcement learning. Reinforcement learning will evaluate the reward value and provide other actions in order to maximize the reward. The optimal sequence on the control units will be found after thousands of iterations.

DDPG, a novel and emerging deep reinforcement learning algorithm, is used as the major algorithm in this paper. The DDPG is an actor-critic algorithm, which has two networks: actor and critic. The actor network can provide the mean action to keep trial-and-error. The critic network will take the current step action and the next step observation, and it will use temporal differences error to judge if the quality of the current step action, μ. The observation will be passed to both the actor network and the critic network. The actor network will give an action mean vector after the observation information passes to the actor network. The critic network will take both mean action and the observation value to update the Q value based on the error of the temporal differences. In the training phase, the target of the critic network can be expressed as Eq. (2), and the target of the actor network can be expressed as Eq. (3).

$$y_t = r(s_t, a_t) + \gamma \cdot Q(s_{t+1}, \mu(s_{t+1})) \tag{2}$$

$$\nabla_{\theta^\mu} J \approx \mathbb{E}_{s_t}[\nabla_a Q(s,a) \cdot \nabla_{\theta^\mu}\mu(s)] \tag{3}$$

Where $r(s_t, a_t)$ refers to the reward function of current step states and actions, γ stands for a discount function with the range of [0, 1], Q represents the critic function that is constructed in the critic network, J is the initial distribution of the parameters, and ∇ mean the gradient that can be calculated based on the chain rule.

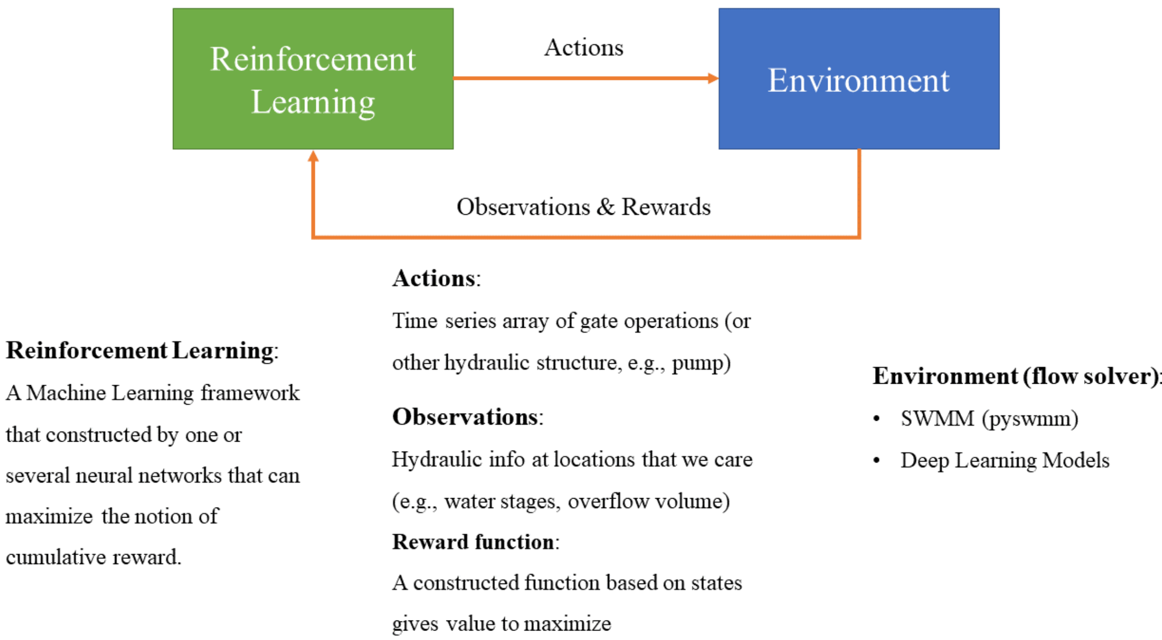


Figure 2. Schematic of the optimization framework

The reward function is critical because it is the only parameter that reinforcement learning evaluates. From the water resource engineering perspective, eliminating CSOs is ideal, however, completely closing the waterway to the river will cause extreme flooding on nodes. To avoid this situation, the water depth on all nodes needs to be controlled within a safety range. This paper considers 2 feet from the maximum allowable depth as the safety limit. Thus, nodal water depth could serve as parameter to control nodal flooding. Two weirs are located upstream of outfall 54 and 65. The CSOs rate is actually equal to the flow rate through two weirs. Thus, the flow rate through the weir could be served as parameter to control the CSOs rate. By combining both points, the reward function is constructed based on the nodal water depth and flow rate through the weir. Since nodal flooding could result in more serious damage than CSOs, we need to give it a higher priority to avoid it. The mathematical expression of the reward function on the single timestep could be written as Eq. (4).

$$re = -w_1 \sum_{i=1}^{n} WD_i - (WD_{\lim\_i} - 2) - w_2 \sum_{j=1}^{m} Weir\_flow_j \tag{4}$$

Where re denotes reward value, w_1 denotes the weights to control nodal flooding, WD stands for the SWMM model calculated water depth, WD_{lim} represents maximum allowable nodal water depth, n is the total number of manholes, w_2 denotes the weights to control the CSO, $Weir\_flow_j$ denotes the flow rate through the weir.

RESULT AND DISCUSSION

Reinforcement Learning Performance on 10-year return period rainfall event

In this section, 10-year return period rainfall is used as the precipitation event in the SWMM model. The comparison of CSOs rate between active control by reinforcement learning and rule-based passive control is shown in Fig. 3a. 10-year rainfall is a good choice for the baseline case because it can significantly show the difference between active control and passive control. As shown in Fig. 3a, CSOs occur on both downstream positions, weir 35391 and 35444, based on the rule-based control. However, these CSOs can be eliminated by optimal control. As the orange and red lines show, the CSO rates keep zeros for the entire event period. Also, the accumulated reward value is also zero, which indicates both CSOs and nodal flooding did not occur in the entire period. The optimal orifice gate position is shown in Fig. 3b. The gate position changes occurred in the entire period, and it changes more rapidly during the time periods with high rainfall intensity. From the optimal gate position, we can see that reinforcement learning can control the orifice gate much more precisely. Conventional rule-based control is unable to achieve gate operation with such high precision.

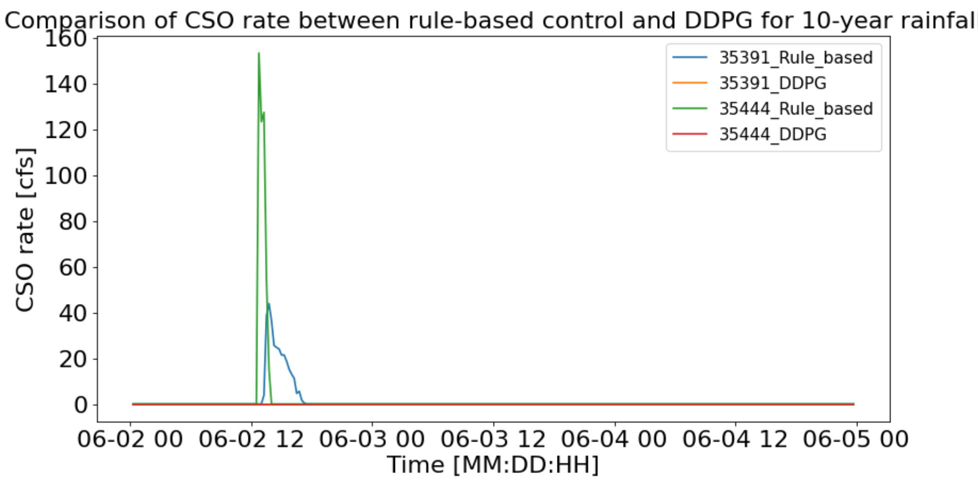
Reinforcement Learning Performance on various return period rainfall events

The comparison of CSOs rate between control by active reinforcement and rule-based passive control for various rainfall events is shown in Fig. 4. As Fig. 4 shows, CSOs are found in all the rainfall events when using rule-based control. It also makes sense that the CSOs rate under rule-based control increases with the increase in rainfall intensity because higher rainfall intensity results in a larger amount of water in the system. For mild to middle levels of rainfall intensity, less than 15-year rainfall, reinforcement learning can eliminate all CSOs without causing any nodal flooding. However, as the intensity increases, optimization results from reinforcement learning cannot always be perfect. The CSOs will occur on the upper half of the system through weir 35444 even when using reinforcement learning. However, the peak value of CSOs rate is minimized by around a half in 20 and 25-year rainfall. Additionally, the total volume of CSOs is significantly minimized to nearly 1/10 of rule-based control results. This could potentially relieve property loss and health issues under extreme events. Lastly, there is no nodal flooding found in the system for all rainfall events because it is constrained by a larger weight.

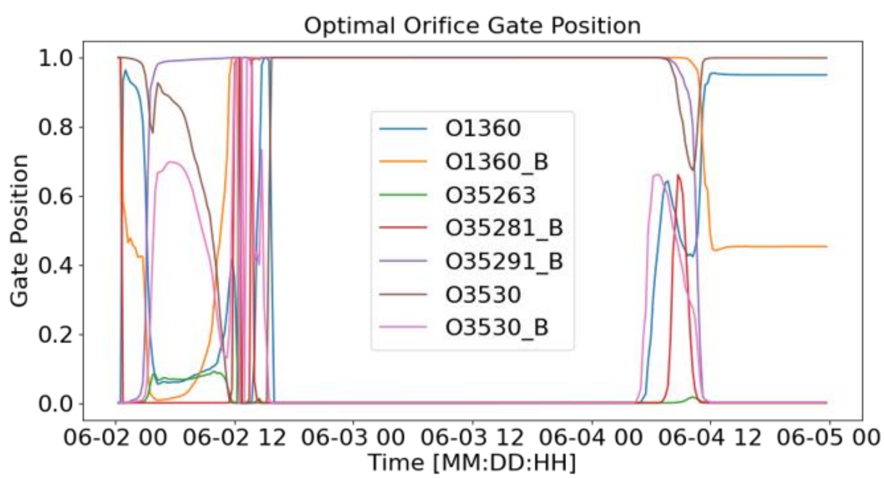
Performance comparison among various reinforcement learning algorithms

Reinforcement learning, as one of three basic machine learning paradigms besides supervised and unsupervised learning, has multiple algorithms to choose from in this project. Since both actions and observations are classified as the continuous numbers, three more acceptable algorithms can be used to test the performance: Proximal Policy Optimization (PPO), Advantage

Actor Critic (A2C), Soft Actor Critic (SAC). To make the comparison fair, the weight on nodal flooding terms is large enough so that there is no nodal flooding found in this section. The comparison of CSOs rate among four different reinforcement learning algorithms are shown in Fig. 5. All reinforcement learning algorithms can solve the CSOs issue for the lower half of the system (CSOs through weir 35391), however, they showed major differences in the upper half system (CSOs through weir 35444). Although PPO algorithm can solve the CSOs issue through weir 35391, it provided an even worse CSOs rate than rule-based control in the upper half system. DDPG, one of the novel and emerging algorithms, provided the best performance, and it can solve the CSOs in the entire computational domain. A2C and SAC showed much better performance than the rule-based control, even though they cannot completely eliminate the CSOs in the system.

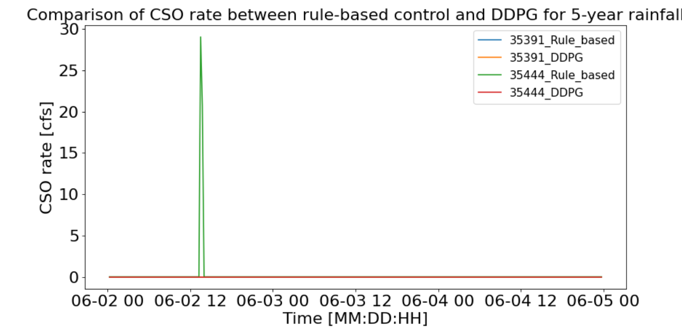


(a)

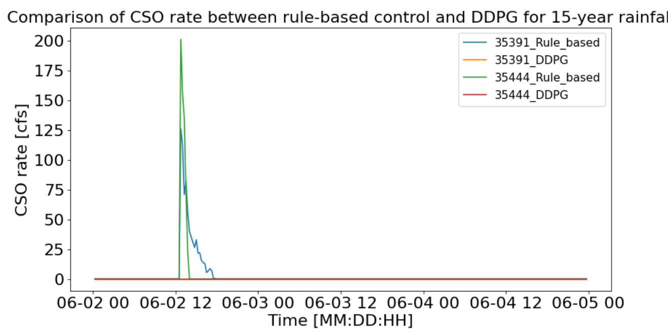


(b)

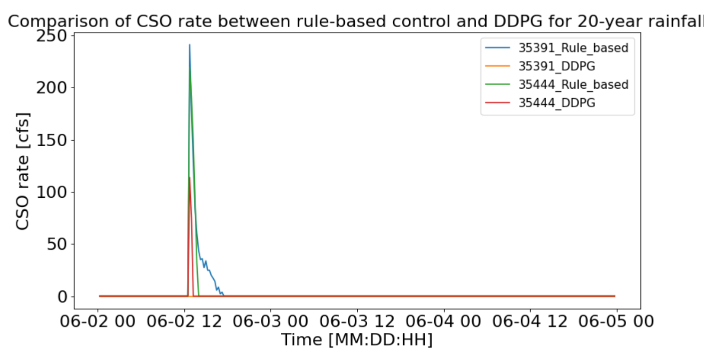
Figure 3. Optimal Control on 10-year return period stormwater event. (a) Comparison of CSOs rate between RL and rule-based control; (b) Optimal gate Position



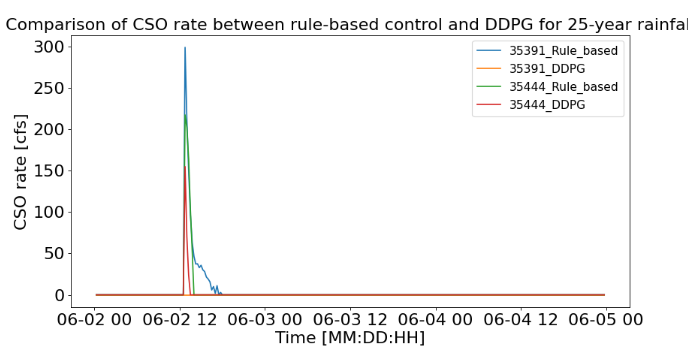
(a)



(b)

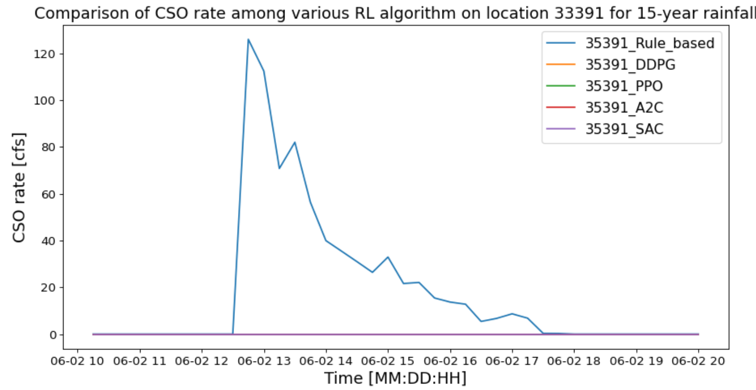


(c)

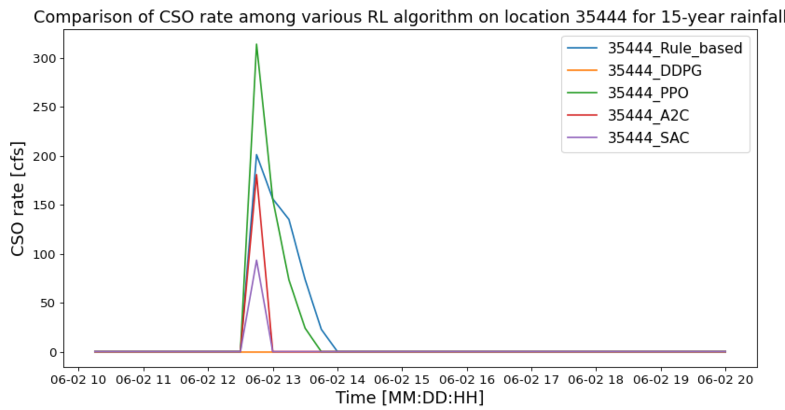


(d)

Figure 4. Comparison of CSOs rate between RL and rule-based control on various return period rainfalls: (a) 5-year return period; (b) 15-year return period; (c) 20-year return period; (d) 25-year return period.



(a)



(b)

Figure 5. Comparison of CSOs among different reinforcement learning algorithms for 15-year return period rainfall: (a) CSOs through weir 35391; (b) CSOs through weir 35444.

CONCLUSION

In conclusion, reinforcement learning was applied to the real-world combined sewer system, Puritan Fenkell/Seven-Mile Collection System, to test the active control performance. Seven orifice gates in the system were chosen to serve as the control unit. The key conclusions are as follows:

- Reinforcement learning shows the ability to actively control the gate at a high precision and complexity level, while rule-based control is hard to achieve. These results showed that reinforcement learning is able to handle and eliminate CSOs in the more intense rainfall event.
- Reinforcement learning can fully eliminate the CSOs for the rainfall event with less than 20-year return period. As the rainfall intensity increases, reinforcement can only minimize the CSOs and reduce the loss.
- By comparing multiple reinforcement learning algorithms, we found that DDPG shows the best performance in our cases. That also matches the expectation of many researchers, which stated that the DDPG algorithm is recommended for time-series optimal control problems. The performance rank from best to worse is DDPG, SAC, A2C, and PPO.

REFERENCES

Abdalla, E. M. H., Pons, V., Stovin, V., De-Ville, S., Fassman-Beck, E., Alfredsen, K., and Muthanna, T. M. (2021). Evaluating different machine learning methods to simulate runoff from extensive green roofs. *Hydrology and Earth System Sciences Discussions*, 1-24.

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., and Kisi, O. (2021). Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic Environmental Research and Risk Assessment*, 35(3), 597-616.

Bowes, B. D., Tavakoli, A., Wang, C., Heydarian, A., Behl, M., Beling, P. A., and Goodall, J. L. (2021). Flood mitigation in coastal urban catchments using real-time stormwater infrastructure control and reinforcement learning. *Journal of Hydroinformatics*, 23(3), 529-547.

Chen, J. X. (2016). The evolution of computing: AlphaGo. *Computing in Science & Engineering*, 18(4), 4-7.

Delipetrev, B., Jonoski, A., and Solomatine, D. P. (2017). A novel nested stochastic dynamic programming (nSDP) and nested reinforcement learning (nRL) algorithm for multipurpose reservoir optimization. *Journal of Hydroinformatics*, 19(1), 47-61.

Ezer, T., and Atkinson, L. P. (2014). Accelerated flooding along the US East Coast: On the impact of sea-level rise, tides, storms, the Gulf Stream, and the North Atlantic Oscillations. *Earth's Future*, 2(8), 362-382.

Hajgató, G., Paál, G., and Gyires-Tóth, B. (2020). Deep reinforcement learning for real-time optimization of pumps in water distribution systems. arXiv preprint arXiv:2010.06460.

Jin, A., Fox-Lent, C., and Linkov, I. (2020). Resilience for smart water systems. *J. Water Resour. Plann. Manage*, 146(1), 02519002.

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., and Pak, C. (2016). *Smarter stormwater systems*.

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., and Pérez, P. (2021). Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*.

Leon, A. S., Tang, Y., Qin, L., and Chen, D. (2020). A MATLAB framework for forecasting optimal flow releases in a multi-storage system for flood control. *Environmental Modelling & Software*, 125, 104618.

Leon, A. S., Bian, L., and 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. *Environmental Modelling & Software*, 145, 105198.

Li, P., Zhang, J., and Krebs, P. (2022). Prediction of Flow Based on a CNN-LSTM Combined Deep Learning Approach. *Water*, 14(6), 993.

Miller, J. D., and Hutchins, M. (2017). The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *Journal of Hydrology: Regional Studies*, 12, 345-362.

Mounce, S. R., Shepherd, W., Ostojin, S., Abdel-Aal, M., Schellart, A. N. A., Shucksmith, J. D., and Tait, S. J. (2020). Optimisation of a fuzzy logic-based local real-time control system for mitigation of sewer flooding using genetic algorithms. *Journal of Hydroinformatics*, 22(2), 281-295.

Mullapudi, A., Wong, B. P., and Kerkez, B. (2017). Emerging investigators series: building a theory for smart stormwater systems. *Environmental Science: Water Research & Technology*, 3(1), 66-77.

Mullapudi, A., and Kerkez, B. (2018). Autonomous control of urban storm water networks using reinforcement learning. *EPiC Series in Engineering*, 3, 1465-1469.

Mynett, A. E., and Vojinovic, Z. (2009). Hydroinformatics in multi-colours—Part red: Urban flood and disaster management. *Journal of Hydroinformatics*, 11(3-4), 166-180.

NASA. (2017). Earth observatory: how will global warming change earth? URL: https://earthobservatory.nasa.gov/Features/GlobalWarming/page6.php (2017).

Pianosi, F., Castelletti, A., and Restelli, M. (2013). Tree-based fitted Q-iteration for multi-objective Markov decision processes in water resource management. *Journal of Hydroinformatics*, 15(2), 258-270.

Sadler, J. M., Goodall, J. L., Morsy, M. M., and Spencer, K. (2018). Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest. *Journal of hydrology*, 559, 43-55.

Sajedi-Hosseini, F., Choubin, B., Solaimani, K., Cerdà, A., and Kavian, A. (2018). Spatial prediction of soil erosion susceptibility using a fuzzy analytical network process: Application of the fuzzy decision making trial and evaluation laboratory approach. *Land degradation & development*, 29(9), 3092-3103.

Saliba, S. M., Bowes, B. D., Adams, S., Beling, P. A., and Goodall, J. L. (2020). Deep reinforcement learning with uncertain data for real-time stormwater system control and flood mitigation. *Water*, 12(11), 3222.

Sallab, A. E., Abdou, M., Perot, E., and Yogamani, S. (2017). Deep reinforcement learning framework for autonomous driving. *Electronic Imaging*, 2017(19), 70-76.

Smith, A. B., and Katz, R. W. (2013). US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Natural hazards*, 67(2), 387-410.

The city of Detroit water and sewerage department. (2018). Stormwater Management Design Manual. Retrieved form https://detroitmi.gov/sites/detroitmi.localhost/files/2018-11/Stormwater%20Management%20Design%20Manual%202018-07-26.pdf.

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

Zhao, G., Pang, B., Xu, Z., Peng, D., and Xu, L. "Assessment of urban flood susceptibility using semi-supervised machine learning model." *Science of The Total Environment* 659 (2019): 940-949.