



**Environmental  
Science**  
Water Research & Technology

**Reinforcement Learning-based Real-time Control of Coastal  
Urban Stormwater Systems to Mitigate Flooding and  
Improve Water Quality**

Journal:	<i>Environmental Science: Water Research &amp; Technology</i>
Manuscript ID	EW-ART-08-2021-000582.R2
Article Type:	Paper

SCHOLARONE™  
Manuscripts

1 **Reinforcement Learning-based Real-time Control of Coastal**  
2 **Urban Stormwater Systems to Mitigate Flooding and Improve**  
3 **Water Quality**

4 **Benjamin D. Bowes<sup>1</sup>, Cheng Wang<sup>1</sup>, Mehmet B. Ercan<sup>1,2</sup>, Teresa B. Culver<sup>1</sup>, Peter A.**  
5 **Beling<sup>3</sup>, and Jonathan L. Goodall<sup>1,4\*</sup>**

6 <sup>1</sup>*Dept. of Engineering Systems and Environment, Univ. of Virginia, 151 Engineer's Way, P.O.*  
7 *Box 400747, Charlottesville, VA 22904, USA*

8 <sup>2</sup>*Xylem, South Bend, IN, USA*

9 <sup>3</sup>*Dept. of Industrial and Systems Engineering, Virginia Polytechnic Institute and State*  
10 *University, 250 Perry St, Blacksburg, VA 24061, USA*

11 <sup>4</sup>*Dept. of Computer Science, Univ. of Virginia, Rice Hall, 85 Engineer's Way, PO Box 400740,*  
12 *Charlottesville, VA 22904, USA*

13 *\* Corresponding Author: [goodall@virginia.edu](mailto:goodall@virginia.edu)*

## 14      **ABSTRACT**

15 Real-time control of stormwater systems can reduce flooding and improve water quality. Current  
16 industry real-time control strategies use simple rules based on water quantity parameters at a local  
17 scale. However, system-level control methods that also incorporate observations of water quality  
18 could provide improved control and performance. Therefore, the objective of this research, is to  
19 evaluate the impact of local and system-level control approaches on flooding and sediment-related  
20 water quality in a stormwater system within the flood-prone coastal city of Norfolk, Virginia, USA.  
21 Deep reinforcement learning (RL), an emerging machine learning technique, is used to learn  
22 system-level control policies that attempt to balance flood mitigation and treatment of sediment.  
23 RL is compared to the conventional stormwater system and two methods of local-scale rule-based  
24 control: (i) industry standard predictive rule-based control with a fixed detention time and (ii) rules  
25 based on water quality observations. For the studied system, both methods of rule-based control  
26 improved water quality compared to the passive system, but increased total system flooding due  
27 to uncoordinated releases of stormwater. An RL agent learned controls that maintained target  
28 pond levels while reducing total system flooding by 4% compared to the passive system. When  
29 pre-trained from the RL agent that learned to reduce flooding, another RL agent was able to learn  
30 to decrease TSS export by an average of 52% compared to the passive system and with an  
31 average of 5% less flooding than the rule-based control methods. As the complexity of stormwater  
32 RTC implementations grows and climate change continues, system-level control approaches such  
33 as the RL used here will be needed to help mitigate flooding and protect water quality.

34 **Keywords:** Real-time Control, Reinforcement Learning, Smart Stormwater Systems, Urban  
35 Flooding, Water Quality

## 36      **Water Impact Statement**

37 Advances in smart and connected technologies can reduce flooding and improve water quality  
38 through real-time stormwater system control. Currently, real-time stormwater control operates at  
39 local-scales with fixed rules. We present a method for learning system-level control strategies that  
40 balance competing flood mitigation and pollutant treatment goals. With continued adoption of  
41 stormwater real-time control, these system-level control approaches can improve flood and pollutant  
42 mitigation.

## 43 1 INTRODUCTION

44 Communities rely on stormwater systems to mitigate flooding and treat polluted runoff from urban  
45 areas. However, as urbanization increases and climate change continues to alter precipitation,  
46 temperature, and sea levels, communities will be faced with increased stormwater runoff causing  
47 greater flooding and water pollution<sup>1-4</sup>. Conventional stormwater systems are designed based on  
48 historic data assuming stationarity of future conditions. They are largely static systems, unable to  
49 dynamically adapt to unanticipated conditions. Increasing the resilience of stormwater systems to  
50 these unanticipated and changing land use and climate conditions will require new approaches to  
51 dynamically control both flood mitigation and pollutant treatment.

52 The adoption of smart cities approaches is allowing stormwater managers to begin to monitor  
53 and control individual components of conventional stormwater systems, which are gravity-driven  
54 and behave statically, in real-time<sup>5</sup>. While the use of real-time control (RTC) is fairly established in  
55 combined sewer systems<sup>6-8</sup>, recent research has shown that retro-fitting conventional stormwater  
56 components (e.g., a retention pond) for RTC can allow more efficient local operation, mitigating  
57 flooding from storms<sup>9,10</sup> and preventing erosive, high velocity flows<sup>11</sup>. RTC can also provide  
58 more efficient treatment of pollutants such as sediment and nutrients, primarily through increased  
59 detention time<sup>12,13</sup>. For instance, RTC of a retention pond increased removal of total suspended  
60 solids (TSS) and nitrate (NO<sub>3</sub>) by roughly 40%, compared to passive pond operation<sup>14</sup>.

61 In practice, stormwater RTC is generally performed using local rule-based control (RBC), which  
62 is almost exclusively based on volumetric data (e.g., depth, current and forecast rainfall)<sup>14-16</sup>. For  
63 instance, a rule may open a valve when the water level in a storage pond reaches a certain height or  
64 proactively drain water from a pond based on a rainfall forecast to create additional storage capacity  
65 before a large storm. In most studies using RBC, water quality is not considered or is inferred  
66 through hydraulic retention time, rather than directly observed or used in control rules. However,  
67 pollutant characteristics are highly variable between sites and storms and there is a need for more  
68 generalizable RTC methods for enhancing pollutant treatment. Toward this end, the benefits of using  
69 real-time water quality observations in control rules has recently been explored in simulation. For  
70 example, using the concentration of TSS to trigger a valve controlling outflow from a storage pond  
71 can improve TSS capture in the pond compared to the passive system and other volumetric control

72 rules<sup>17</sup>. Given the effectiveness of RTC-enabled individual infrastructure components to adapt to  
73 different storm events, system-level RTC (i.e., control of multiple infrastructure components based  
74 on information from locations throughout the system) has the potential to more holistically enhance  
75 flood and pollution mitigation through coordinated control of multiple components<sup>18</sup>.

76 As the complexity of controlled stormwater systems increases, the task of creating rules to (i)  
77 mitigate flooding, (ii) protect the quality of receiving waters, or (iii) balance both flooding and water  
78 quality, becomes nontrivial. Further, controlling for flooding and water quality can be competing  
79 goals. For example, draining a stormwater pond is the simplest way to prevent it from flooding.  
80 However, treatment of pollutants can require holding more water; TSS requires still conditions  
81 for settling, but stormwater inflow could resuspend sediment if a pond is drawdown to shallow  
82 depths. Maintaining more submerged (anaerobic) areas can increase denitrification, but reduces  
83 capacity to capture additional stormwater without flooding. Control rules or thresholds can be set to  
84 attempt to balance these goals, but they may only perform well under a limited range of conditions.  
85 Instead of attempting to create rules that cover all possible interactions between stormwater system  
86 components, pollutants, and environmental conditions, recent research has explored system-level  
87 methods of optimizing stormwater RTC. For instance, system-level control of a coastal urban  
88 stormwater system reduced total system flooding, even under sea level rise conditions<sup>9</sup>. In terms of  
89 water quality, flow from a system of ponds to a treatment wetland has been controlled to increase  
90 the efficiency of nitrate removal by 46%<sup>18</sup>. A study using system-level RTC for both water quantity  
91 and quality used linear optimization to control retention basin outflows. However, water quality  
92 control still relied on fixed rules to extend detention time (i.e., hold water after a storm for a set  
93 amount of time) and system control based on either observed or simulated real-time water quality  
94 measurements was not included. Continuing improvements in real-time water quality sensors, could  
95 allow more direct observation and control of water quantity in conjunction with some water quality  
96 parameters<sup>19,20</sup>. Making the best use of these growing sensing capabilities requires new methods of  
97 creating stormwater RTC policies that balance flood mitigation and water quality improvement.

98 Recent advances in Reinforcement Learning (RL), a type of machine learning, provide an  
99 alternate approach to system-level stormwater RTC where control policies can be learned, instead  
100 of using predetermined rules<sup>21</sup>. In RL, an agent (i.e., algorithm) does not have known answers to  
101 learn from, which is the standard supervised machine learning paradigm, but instead is rewarded

102 based on how well its control actions meet specified stormwater system goals (e.g., flood mitigation,  
103 improved water quality). The reward signal is used to guide the agent's learning towards actions  
104 that maximize the return from areward function. Classical tabular RL is closely related to Dynamic  
105 Programming and has been explored for multi-objective reservoir management<sup>22–26</sup>. However,  
106 because tabular RL is limited to systems with relatively small numbers of possible states and actions,  
107 Deep Reinforcement Learning (also widely referred to as RL), which uses neural networks as  
108 function approximators instead of using lookup tables, has been used for control of more complex  
109 systems<sup>27,28</sup>. This approach to learning allows RL increased flexibility to optimize control actions,  
110 balance competing objectives based on the formulation of the reward function, and has the potential  
111 to continually adapt system controls to evolving environmental conditions (e.g., increased runoff  
112 from urbanization or climate change).

113 Initial research with RL for stormwater systems demonstrated control policies that reduced peak  
114 flows could be learned using water quantity observations from a complex system<sup>29</sup>. Flood mitigation  
115 improvements have also been achieved using RL-based RTC to learn system-level policies with water  
116 quantity data<sup>10,30</sup>, while being robust to uncertainty in observations and forecasts<sup>31</sup>. Despite the fact  
117 that many stormwater systems are used for pollutant treatment as well as flood mitigation, previous  
118 RL research has not considered using water quality observations to inform RTC methods. Given  
119 real-time water quality observations, RL may be able to learn to balance competing water quantity  
120 and quality goals throughout a stormwater system over a large range of conditions and could  
121 outperform rule-based methods. This paper is the first to incorporate water quality observations into  
122 RL-based control policies and, therefore, aims to illustrate RL's ability to learn system-level control  
123 policies considering the competing goals of flood mitigation and water quality protection.

## 124 **2 METHODS**

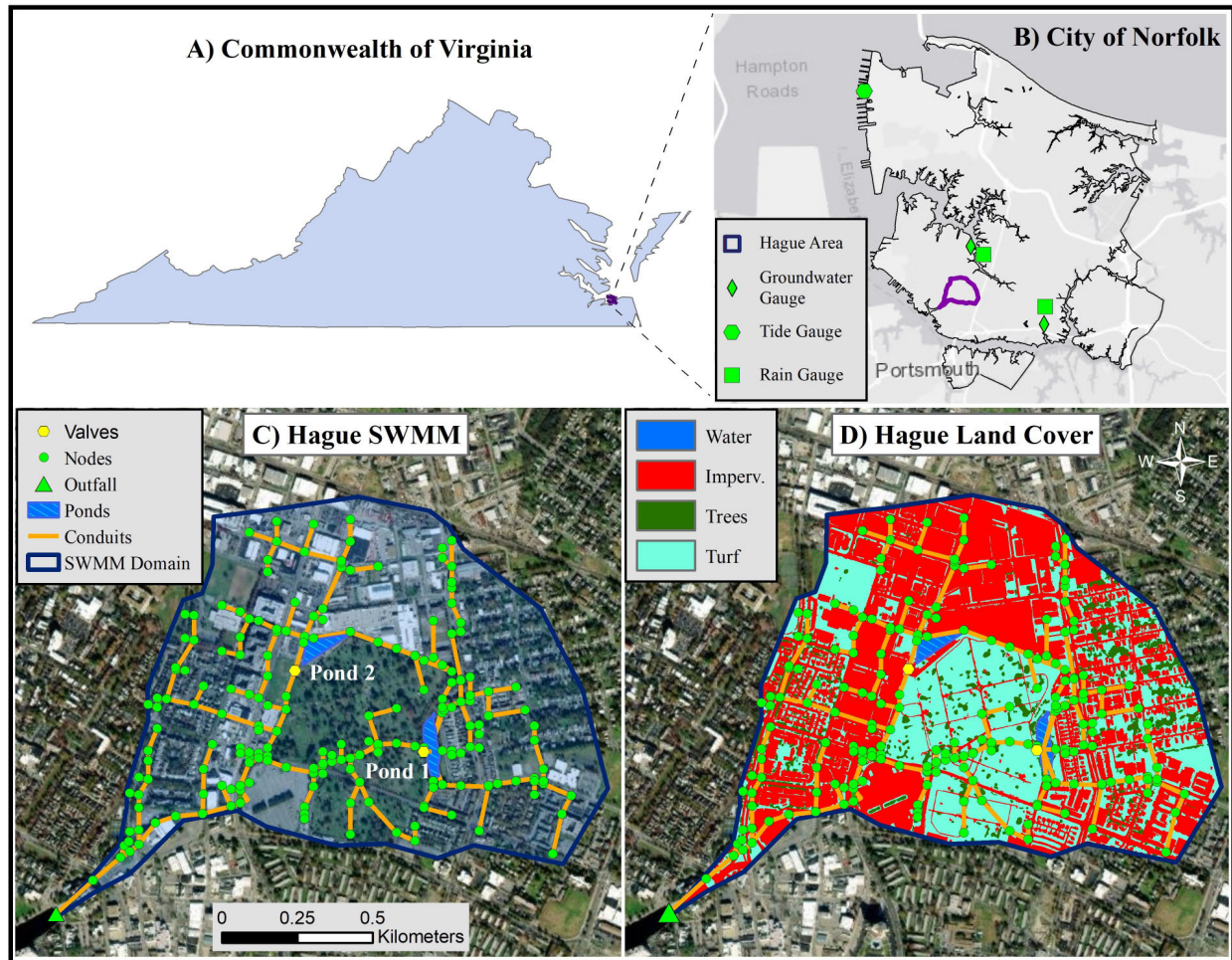
125 This research compares RL and RBC for their ability to both mitigate flooding and improve  
126 water quality compared to conventional static stormwater infrastructure. A simulation of Norfolk,  
127 Virginia's stormwater system including water quantity and quality processes is used as the controlled  
128 system. Two methods of local-scale, rule-based control are implemented: (i) predictive RBC with  
129 a fixed detention time and (ii) RBC based on water quality observations. RL is implemented  
130 for system-level control that incorporates measures of water quality and flood mitigation. After

131 comparing the performance of these methods, their robustness to changes in system behavior is  
132 evaluated by simulating groundwater exchange within the controlled ponds.

## 133 **2.1 Study Area**

134 The City of Norfolk, Virginia, specifically its Hague neighborhood, is used as the study area for  
135 this research. Norfolk is situated near the mouth of the Chesapeake Bay on the eastern coast of  
136 the U.S. (Fig. 1, A and B). The city has a high rate of relative sea level rise partly due to regional  
137 land subsidence<sup>32</sup> and its low elevation, flat topography, and regular hurricane season contribute to  
138 increasingly frequent and severe recurrent flooding<sup>1</sup>. Additionally, Norfolk has a high groundwater  
139 table that responds quickly to storm events<sup>33</sup> and could contribute significant amounts of water to  
140 retention ponds that are being actively controlled<sup>10</sup>. The Hague neighborhood is a historic part  
141 of Norfolk and is adjacent to many city government buildings and the region's main hospital; the  
142 Hague also experiences some of the most frequent flooding in the city<sup>9,34</sup>.

143 The quality of stormwater runoff from the city contributes to the health of the Chesapeake Bay,  
144 which has a long history of impairments such as hypoxia caused by eutrophication<sup>35,36</sup>. Pollutants  
145 carried by the city's stormwater (such as TSS, nitrogen, and phosphorous) are regulated to meet the  
146 Total Maximum Daily Loads (TMDLs) set for the Bay. As outlined in the City's Chesapeake Bay  
147 TMDL Action Plan, the City is required to reduce pollutant loadings by 5.75%, 35%, and 60% by  
148 2021, 2026, and 2031, respectively<sup>37</sup>.



**Figure 1.** Study area - Hague area of Norfolk, Virginia USA with (C) the SWMM model and (D) land cover data.

## 149 2.2 SWMM Model

150 The Hague's recurrent flooding prompted Norfolk to build a model of the existing conventional  
 151 stormwater system using the U.S. Environmental Protection Agency's (EPA) Stormwater Manage-  
 152 ment Model (SWMM) (Fig. 1, C). The city verified that the SWMM model behavior sufficiently  
 153 represented the physical system and calibrated it to match observed flooding in the Hague from  
 154 Hurricane Matthew, which caused wide-spread flooding in October, 2016. The Hague SWMM  
 155 model was updated by Sadler et al.<sup>9</sup> to simulate real-time control infrastructure (i.e., an additional  
 156 retention pond and a valve, pump, and inflatable dam). In the current study, the SWMM simulation  
 157 from Sadler et al.<sup>9</sup> is driven by long-term observed rainfall with a tidal boundary condition and



158 has been enhanced to include groundwater and water quality processes. Land cover within the  
159 Hague SWMM model domain was extracted for each subwatershed from a 1m resolution dataset<sup>38</sup>  
160 and included three pollutant generating land covers: impervious, turf grass, and trees (Fig. 1, D).  
161 Impervious cover represents 56% of the model domain, while turf grass and trees cover 37% and  
162 6%, respectively; the remainder of the land cover is water. SWMM input files with full configuration  
163 details can be found in the open source code repository (see Section 5).

### 164 **2.2.1 Input Data**

165 Observed rainfall, tide, and groundwater data were collected from gauges in Norfolk for the period  
166 between 1 January, 2010 and 6 November, 2019 (Fig. 1, B). Fifteen minute rainfall data came from  
167 two stations near the Hague that are operated by the Hampton Roads Sanitation District (HRSD).  
168 Rainfall data is processed by first removing any values over the 1000-year 15-minute value for  
169 Norfolk (59.2mm); these large values represented less than 0.01% of the rainfall datasets. Any  
170 missing values from one rain gauge are filled with the value from the other gauge if available; there  
171 were no periods where both rain gauges were missing data. Finally, the mean of the two rain gauges  
172 is taken to create a single time series for the SWMM model. Observed 6-minute tide data came  
173 from the Sewells Point gauge operated by the National Oceanic and Atmospheric Administration  
174 (NOAA). Tide data are referenced to the North American Vertical Datum of 1988 (NAVD88) and  
175 were resampled to an hourly interval for use as a SWMM boundary at the stormwater system outfall.

176 Forecasts for use in the RTC control methods were created from the observed rainfall and tide  
177 data. These forecasts are a rolling window of values over the next  $n$  time steps. In this work, a 24  
178 hour forecast of 15 minute rainfall contains  $n=96$  values. Because the focus of this work is on  
179 comparison of the RTC scenarios, the forecasts were assumed to represent perfect knowledge.

### 180 **2.2.2 Groundwater Exchange Simulation**

181 Because Norfolk has a high groundwater table and is already experiencing impacts from a high  
182 rate of relative sea level rise, the robustness of stormwater RTC methods to groundwater exchange  
183 will be increasingly important. While groundwater interactions with the retention ponds in Norfolk  
184 have not been studied specifically, it has been demonstrated that increased groundwater table levels  
185 due to sea level rise could contribute to retention ponds in coastal areas, decreasing their ability to  
186 appropriately manage consecutive storm events<sup>39</sup>. To address this need, groundwater exchange

187 with controlled ponds is simulated in a number of the scenarios in this research.

188 Groundwater data was collected from two shallow monitoring wells operated by HRSD and  
189 referenced to NAVD88. Outliers from these data were removed with a Hampel filter (as in<sup>33</sup>) to  
190 remove large erroneous values and replace them with the median of a one-day rolling window.  
191 Groundwater observations are then aggregated to an hourly time step. A single time series for the  
192 Hague area was interpolated using inverse distance weighting between Pond 1, the two groundwater  
193 monitoring wells, and the tidal level at the stormwater system outfall (assumed to be equal to the  
194 groundwater table level at the land/water interface). From 2010-2019, the groundwater table is  
195 higher than the water level in Pond 1 and lower than the water level in Pond 2, 93.7% and 73.8%  
196 of the time, respectively. This indicates that Pond 1 may be gaining water from groundwater flow  
197 while Pond 2 may be losing water to groundwater. The groundwater table level is only below the  
198 bottom elevation of either pond 0.09% of the time.

199 The Hague SWMM model provided by the City of Norfolk did not originally simulate ground-  
200 water processes and was not configured to easily allow simulation of groundwater exchange with  
201 the controlled ponds using SWMM's aquifer components. To address this, a conceptual model of  
202 the unconfined aquifer surrounding the existing Hague pond (Pond 1) was developed. Groundwater  
203 exchange was calculated externally from the SWMM simulation using the Dupuit equation and  
204 added (or subtracted, in the case of infiltration) to the pond as an inflow using pyswmm functional-  
205 ity<sup>40</sup>. The Dupuit equation is commonly used to calculate exchange between a water body and an  
206 unconfined aquifer<sup>41</sup> and is written as

$$Q = \frac{K}{2L}(h_1^2 - h_2^2) \cdot A \quad (1)$$

207 where  $Q$  is the seepage rate into or out of the pond,  $K$  is the saturated hydraulic conductivity of soil  
208 surrounding the pond,  $h_1$  and  $h_2$  are the heights above a fixed datum for the pond water level and  
209 groundwater table level, respectively.  $L$  is the horizontal distance between  $h_1$  and  $h_2$ , and  $A$  is the  
210 surface area over which seepage can occur (a function of pond water level).

211 Saturated hydraulic conductivity of the soil surrounding the existing pond (Pond 1) was estimated  
212 from the National Resource Conservation Service (NRCS) Web Soil Survey as 0.60m/day. This  
213 soil is classified as a fine sandy loam with 61% sand, 22% clay, and 17% silt. Values for  $h_1$  were

214 based on SWMM's simulation of pond water level and  $h_2$  was the observed groundwater table level.  
 215 Because  $L$  controls the hydraulic gradient (when the other variables are held constant), smaller  
 216 values of  $L$  should increase exchange between the ponds and the simulated aquifer. The sensitivity  
 217 of pond depth and inflow to the distance between measured water levels ( $L$ ), was tested for  $L = 7.62$ ,  
 218 3.0, 1.5, and 0.3m using the passive (i.e., uncontrolled) SWMM model. A single value of  $L$  was  
 219 chosen and used to demonstrate the impact of groundwater exchange on flooding and water quality  
 220 with the control methods.

221 The impact of groundwater exchange with the controlled ponds was evaluated for the month  
 222 of September, 2016. This month had two hurricanes and one tropical storm, which caused the  
 223 groundwater table level to reach a height of 1.08m (compared to the mean of 0.61m). Because  
 224 groundwater exchange may increase infiltration and reduce flooding and TSS outflow from the  
 225 controlled ponds, a direct comparison of a single RTC method's performance with and without  
 226 groundwater exchange may not be fair. To account for this, the percent difference between the  
 227 passive system and each RTC method's total flooding and TSS loads (with or without groundwater  
 228 exchange) will be compared.

### 229 **2.2.3 Water Quality Simulation**

230 Water quality processes, specifically for TSS, were modelled using SWMM's buildup, washoff,  
 231 and treatment equations<sup>42</sup>. TSS was chosen for this study to allow comparison with previous RTC  
 232 literature, and because it is straight-forward to simulate (through gravitational settling) and known  
 233 to carry other sorbed pollutants<sup>43</sup>. Pollutant buildup within each subcatchment is modelled as a  
 234 power function

$$B = \min(C_1, C_2 \cdot t^{C_3}) \quad (2)$$

235 where  $B$  is the buildup of TSS (mass per unit area),  $C_1$  is the maximum buildup possible,  $C_2$  is the  
 236 buildup rate (buildup per day),  $t$  is the antecedent dry period, and  $C_3$  is a dimensionless buildup  
 237 time exponent. Washoff of accumulated TSS from subcatchments is modelled with an exponential  
 238 function

$$W = E_1 \cdot q^{E_2} \cdot B \quad (3)$$

239 where  $W$  is the washoff rate (mass per area per hr),  $E_1$  is the washoff coefficient (per unit of rain),  
 240  $q$  is the runoff rate (per hr),  $E_2$  is the washoff exponent, and  $B$  is the amount of built-up pollutant  
 241 remaining. Treatment of TSS occurs in the retention ponds and is modelled as a first order decay  
 242 based on a generalized settling velocity (similar to<sup>17</sup>) with resuspension as a factor of depth and  
 243 inflow velocity (inspired by<sup>6</sup>)

$$C = \begin{cases} TSS \cdot \exp(-v_s/DEPTH \cdot DT/3600) & FLOW \leq \tau \\ TSS & FLOW > \tau \\ TSS \cdot (1 - \exp(-v_s/DEPTH \cdot DT/3600)) & FLOW > \tau, DEPTH \leq \delta \end{cases} \quad (4)$$

244 where  $C$  is the TSS concentration (mg/L) in the pond after treatment,  $TSS$  is the inflow concentration,  
 245  $v_s$  is the generalized settling velocity (m/hr),  $DEPTH$  is the pond water depth (m),  $DT$  is the  
 246 SWMM routing time step (seconds),  $FLOW$  is the total inflow rate (m<sup>3</sup>/s) (including groundwater,  
 247 when simulated),  $\tau$  is a flow threshold to distinguish when settling occurs, and  $\delta$  is a depth  
 248 threshold to distinguish when resuspension occurs (one quarter of the maximum pond depth in this  
 249 implementation). The first case in Eq. (4) allows settling over the simulation time step when inflow  
 250 is low and reduces TSS concentration. When the inflow rate is above the threshold, no settling  
 251 occurs (i.e., no TSS treatment). The final case in Eq. (4) simulates resuspension when inflow is  
 252 high and the pond depth is low by increasing the TSS concentration by the amount that would have  
 253 been settled according to  $v_s$ . Resuspension is included because RTC creates the potential for low  
 254 water depths in retention ponds; if a pond is drawdown before high storm inflows, sediment may  
 255 be resuspended and carried downstream.

256 Each land cover category within the SWMM model domain is given individual characteristics  
 257 for the buildup and washoff processes (starting values were taken from<sup>44</sup>). With no observed  
 258 pond water quality data available, the SWMM pollutant processes were calibrated to the annual  
 259 loading and treatment percent of TSS in Pond 1 (the existing pond) (Table 1). TSS loading was  
 260 estimated using the loading rates provided in Norfolk's Virginia Stormwater Management Permit<sup>45</sup>.  
 261 The treatment efficiencies of the passive retention ponds were assumed to be 60% as specified in  
 262 the Virginia Department of Environmental Quality's Chesapeake Bay TMDL Special Condition  
 263 Guidance<sup>46</sup>. The load into Pond 1 was calibrated using the buildup coefficient  $C_2$  so that the mean

264 annual load over 2010-2019 was within 2% of the estimated value. The treatment was calibrated  
 265 using the flow threshold ( $\tau$ ) and the settling velocity ( $v_s$ ) so that the mean annual reduction was  
 266 within 5% of the estimated value for the passive simulation. While calibrating this SWMM model  
 267 to observed values would be desirable, the scope of this paper is on comparison of the RTC methods  
 268 and not exact quantification of TSS.

**Table 1.** Calibrated buildup, washoff, and treatment parameters used in the Hague SWMM model. Note that treatment occurs in the stormwater ponds and is not dependent on land cover.

Land Use	Buildup			Washoff		Treatment	
	$C_1$	$C_2$	$C_3$	$E_1$	$E_2$	$v_s$	$\tau$
Impervious	150.16	0.364	1.54	6.97	1.57		
Turf Grass	62.0	0.325	1.26	5.91	1.46	0.105	5.66
Trees	9.22	0.133	0.87	2.11	1.02		

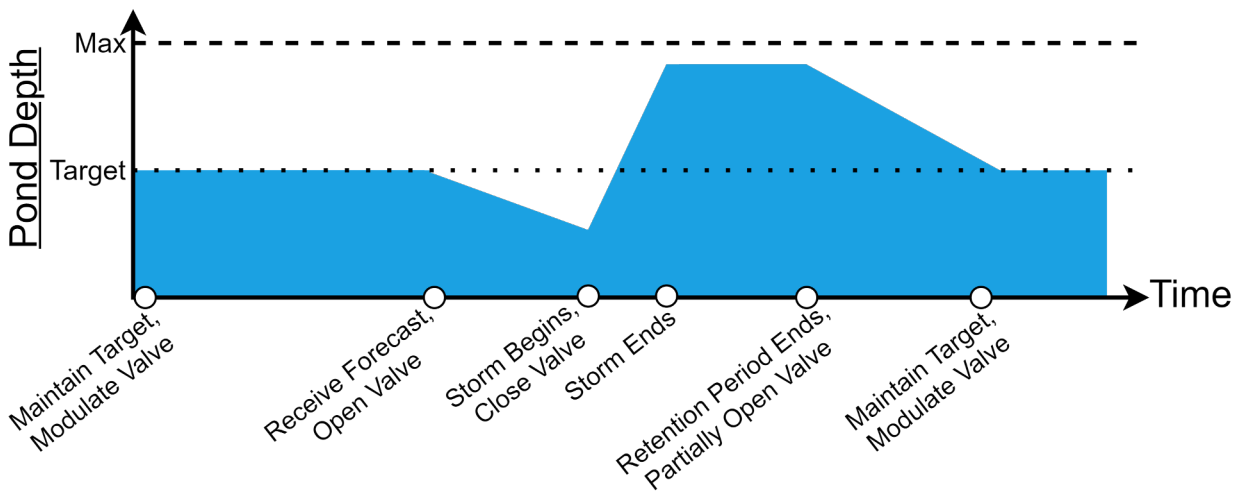
### 269 **2.3 Real-time Control Scenarios**

270 Real-time control of the Hague stormwater system was simulated with three strategies and compared  
 271 to the passive system. The three control strategies are (i) predictive RBC with a fixed detention  
 272 time, (ii) TSS concentration-based RBC, and (iii) RL approaches that includes simulated real-time  
 273 measurement of TSS concentration in the system state and/or reward function. In the passive  
 274 system scenario, weirs control flow out of the retention ponds and maintain a permanent pool of  
 275 approximately half capacity. In the RTC scenarios, the passive weirs are replaced with valves. The  
 276 valve on Pond 1 is at the same elevation of the passive weir (i.e., halfway up the pond's side) due to  
 277 pipe configuration constraints). The valve on Pond 2 is at the bottom of the pond side, which allows  
 278 Pond 2 to be fully emptied or filled. Both RBC scenarios represent local (i.e., individual) control of  
 279 the retention ponds, while RL can coordinate its control actions based on system-level information.  
 280 The pyswmm Python package<sup>40</sup> is used to implement all RTC scenarios on a standard PC with 8  
 281 cores, 16GB RAM, and an NVIDIA Quadro P2000 Graphical Processing Unit (GPU).

#### 282 **2.3.1 Detention Rule-based Control**

283 In this scenario, RBC is based on industry standard methods that use rainfall forecasts for predictive  
 284 control of stored water to mitigate flooding, while controlling water quality with a fixed detention

time<sup>12,14,47</sup>. The general process of this RBC (RBC-DTN) is shown in Figure 2 and detailed in<sup>10</sup>. Briefly, if a forecast storm is expected to flood the pond, the valve will open to drain an equivalent volume of water (plus a safety factor). When the pond is drawdown sufficiently, the valve will close to retain the incoming runoff for a fixed time (24hr in this case). At the end of the retention period, the valve opens to the minimum setting to bring the water level back to the target operating depth within a fixed time (24hr). Outside of storm events, the valve operates based on the observed water level in order to maintain a target depth in the pond. A fail-safe rule overrides any previous rules by completely opening the valve if the pond is flooding. A decision diagram detailing these rules is shown in Appendix A (Fig. 15).

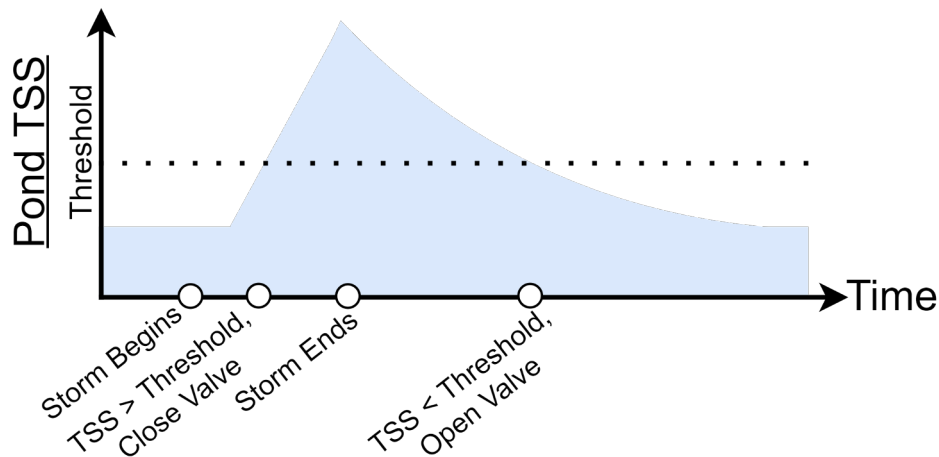


**Figure 2.** General schema of the Detention Rule-based Control (RBC-DTN) scenario. Forecasts allow predictive control of the pond water level to mitigate flooding while a fixed detention time after storm events helps improve water quality.

### 2.3.2 TSS Rule-based Control

The TSS RBC (RBC-TSS) scenario was inspired by Sharior et al.<sup>17</sup>. Instead of using a fixed detention time, this RBC is innovative because it uses the real-time concentration of TSS in a retention pond to trigger valve operation (Fig. 3). For example, when the TSS concentration is above a threshold, the valve can be closed to retain stormwater and allow treatment by settling. Otherwise, the valve is open and acts as a weir to maintain a permanent pool of water. In this study, the TSS threshold was set to 1 mg/L because observed data from the ponds were not available for a

301 more realistic threshold; in Sharior et al.<sup>17</sup>, the threshold is 15 mg/L based on regulatory constraints  
 302 for their study area and calibrated model. A contingency rule limits flooding of the pond by opening  
 303 the valve if a threshold depth is reached. A decision diagram detailing these rules is shown in  
 304 Appendix A (Fig. 16).



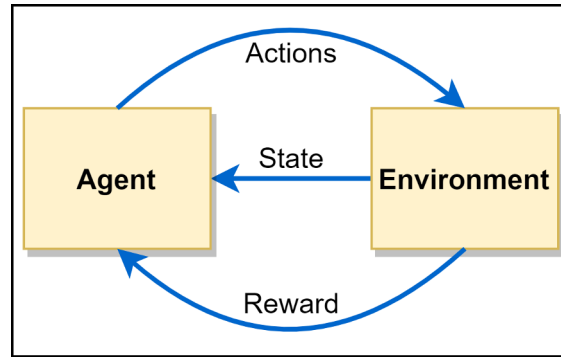
**Figure 3.** General schema of the TSS Rule-based Control (RBC-TSS) scenario. Detention is based on observed TSS concentration, not a fixed length of time, making it adaptive to individual storm events.

### 305 2.3.3 Reinforcement Learning

306 Reinforcement learning can be visualized as an agent that interacts with an environment (Fig. 4).  
 307 The RL agent learns through sequential interactions with the environment. At each step in the  
 308 learning process, the RL agent receives information about the state ( $s$ ) of the environment and  
 309 can take actions ( $a$ ). The next state ( $s'$ ), therefore, depends on the agent's actions and the agent is  
 310 rewarded (positively or negatively) based on how well its actions meet user-specified objectives  
 311 in a reward function ( $r$ ). The agent's ultimate goal is to find a policy ( $\pi(a|s)$ ) that maximizes the  
 312 expected return

$$G_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k} \quad (5)$$

313 where  $r_t = r(s_t, a_t, s_{t+1})$  and  $\gamma \in [0, 1]$  is a discount factor weighting the importance of short-term  
 314 and long-term reward.



**Figure 4.** Reinforcement learning paradigm.

315 In this case the environment is the SWMM model described in section 2.2, which provides state  
 316 information at a 15-minute simulation time step. The state space ( $S$ ) is defined as: the current depth  
 317 (m) and outflow ( $\text{m}^3/\text{s}$ ) of the two retention ponds, the load of TSS (mg) in each pond's outflow,  
 318 the current position of each controllable valve, the sum of the 24 hr rainfall forecast (mm), and the  
 319 mean value of the 24 hr tide forecast (m). The action space ( $A$ ) of the agent is to open or close either  
 320 valve to any degree. The reward ( $r$ ) is based on how well the agent meets user-specified objectives  
 321 such as flood and pollutant reduction.

322 The deep reinforcement learning algorithm used in this research, Deep Deterministic Policy  
 323 Gradients (DDPG), is an actor-critic RL agent using deep neural networks as function approxima-  
 324 tors<sup>28</sup>. DDPG allows controls (i.e., valve positions) over a continuous action state and has been  
 325 used in previous research to learn control policies that mitigate flooding<sup>10,30,31</sup>. The actor in DDPG  
 326 is a deep feed-forward neural network that learns a policy ( $\pi(a|s)$ ); the critic is a deep feed-forward  
 327 neural network that approximates the value of being in a specific state and taking specific actions  
 328 called the Q-value

$$Q^\pi(s, a) = r(s, a, s') + \gamma \sum_{s' \in S} P_{s, s'}^a \sum_{a' \in A} \pi(a' | s') Q^\pi(s', a') \quad (6)$$

329 where  $P_{s, s'}^a$  is the probability of transitioning between two states. This equation is known as the  
 330 Bellman equation and is a key component of RL<sup>21</sup>. By approximating the Q-value, the critic can  
 331 reduce the variance of policy gradients from the actor, which helps speed the learning process.  
 332 During training, the actor receives the state of the stormwater system and outputs the actions to  
 333 be taken based on its learned policy. The critic then receives the actions and states and outputs an



334 estimated Q-value. The actions and Q-value estimates output from the critic are used to update the  
 335 agent. An in-depth description of the DDPG algorithm can be found in Lillicrap et al.<sup>28</sup>.

336 When training RL agents, more explicit reward functions can improve the ability to learn  
 337 appropriate policies<sup>29</sup>. The reward functions used here have a conditional format based on the  
 338 rainfall forecast that guide agent learning under different conditions; this structure has been used to  
 339 improve RL agent policies for flood mitigation<sup>10</sup>. The rainfall forecast can signal to the agent that  
 340 flooding may occur and control actions are rewarded differently because the pond level may need  
 341 to be altered from the target. When no rainfall is forecast, a different set of rewards is triggered  
 342 that incentivize actions for goals like maintaining target depths or increasing retention time for  
 343 additional TSS treatment.

344 In this research, three RL agents are trained and tested. The first agent (RL-FD) is rewarded for  
 345 reducing total flooding throughout the stormwater system and maintaining target pond depths

$$r = \begin{cases} -\Sigma Flooding[system, Pond1 * 1000, Pond2] & F \geq \delta \\ -(|Pond1_{depth} - \tau| + |Pond2_{depth} - \tau|) & F < \delta \end{cases} \quad (7)$$

346 where  $Flooding[system]$  is the incremental system flood volume,  $Flooding[Pond1]$  is the flooding  
 347 rate at Pond 1, and  $Flooding[Pond2]$  is a binary reward (0 or 1000).  $F$  is the sum of rainfall in a  
 348 24hr forecast,  $\delta$  is the rainfall threshold (12.7mm in this research), and  $\tau$  is the target depth (1.8m  
 349 and 1.1m for Ponds 1 and 2, respectively). Several of the nodes upstream of Pond 2 are at lower  
 350 elevations than the top of the pond and flood before Pond 2 will; therefore instead of the Pond 2  
 351 flooding rate, the binary reward acts as a penalty in cases where the pond is above the depth that  
 352 causes flooding upstream (1.75m).

353 The second RL agent (RL-FDTSS) is rewarded for reducing total flooding throughout the  
 354 stormwater system, maintaining target pond depths, and minimizing the export of TSS from the  
 355 ponds

$$r = \begin{cases} -\Sigma \text{Flooding}[\text{system}, \text{Pond1} * 1000, \text{Pond2}] \\ \quad + \text{TSS}[\text{Valve1}, \text{Valve2}] & F \geq \delta \\ -(|\text{Pond1}_{\text{depth}} - \tau| + |\text{Pond2}_{\text{depth}} - \tau| \\ \quad + \text{TSS}[\text{Valve1}, \text{Valve2}] + \text{Flooding}[\text{system}/35000]) & F < \delta \end{cases} \quad (8)$$

356 where  $\text{TSS}[\text{Valve1}, \text{Valve2}]$  is the incremental TSS load of the controlled valves.

357 The third RL agent (RL-FD+FDTSS) aims to balance RL-FD and RL-FDTSS by initializing  
 358 the trained neural network weights and memory from RL-FD and training for 50,000 additional  
 359 time steps using the reward for RL-FDTSS (Eq. 8). This can be considered as pre-training for RL-  
 360 FD+FDTSS, a common practice in deep machine learning to provide appropriate initial conditions  
 361 and reduce computational time (for examples in hydrology see<sup>48</sup> or<sup>49</sup>).

362 The RL agents are trained on one month of data (August, 2019), which has the fifth highest  
 363 monthly total rainfall (256.5mm) of the dataset distributed across 7 storm events. The mean tide  
 364 level in this month is 0.16m, with a maximum value of 1.0m from Tropical Storm Erin late in  
 365 the month. In previous research, this month of data was found to provide a representative range  
 366 of state information that allowed an RL agent to learn effective flood mitigation policies<sup>10</sup>, while  
 367 also keeping computational costs reasonable. A visualization of the rainfall and sea level training  
 368 data, as well as the TSS concentration in each pond is given in Figure 5. RL-FD is trained for  
 369 100,000 steps of the training data with a discount factor (weighting of current and future rewards)  
 370 of 0.5. RL-FDTSS and RL-FD+FDTSS are both trained for 150,000 steps, when the pre-training  
 371 from RL-FD is considered for RL-FD+FDTSS, with a discount factor of 0.99. RL agents are  
 372 tested on the remaining data (2010-2019). Each RL agent has the same neural neural network  
 373 architecture; these and the shared RL hyperparameters are documented in the open source code  
 374 repository linked in section 5. The DDPG algorithm is implemented with the keras-rl<sup>50</sup>, openai  
 375 gym<sup>51</sup>, and Tensorflow<sup>52</sup> python packages; the wandb<sup>53</sup> python package was used for tracking  
 376 training progress and comparing agents during the hyperparameter tuning process.

### 377 **2.3.4 RTC Comparisons**

378 The RTC scenarios are evaluated in three main comparisons as shown in Table 2; each comparison  
 379 has a different control scale and prioritization of flooding and water quality improvement. First, a

380 baseline for flood mitigation is established by comparing the passive system and RL-FD, which  
 381 does not consider water quality in its control policy. While the design of the passive system does  
 382 consider pollutant treatment, the main focus is on flood mitigation. Second, trade-offs between the  
 383 RBC methods, which focus on flooding and TSS at the pond scale, are compared to the passive  
 384 system. In this comparison, the RBC controls prioritize enhancing pollutant treatment as this is one  
 385 of the largest benefits these systems have had in practice. Third, system-level control trade-offs  
 386 with RL-FDTSS and RL-FD+FDTSS, which considered both flooding and TSS in their training, are  
 387 compared to the passive system and RL-FD. These three comparisons are made without simulating  
 388 groundwater exchange to keep the focus on control actions and reduce computational expense. The  
 389 impact of groundwater exchange is then examined on a subset of the data to evaluate its potential  
 390 impact on RTC of the stormwater system.

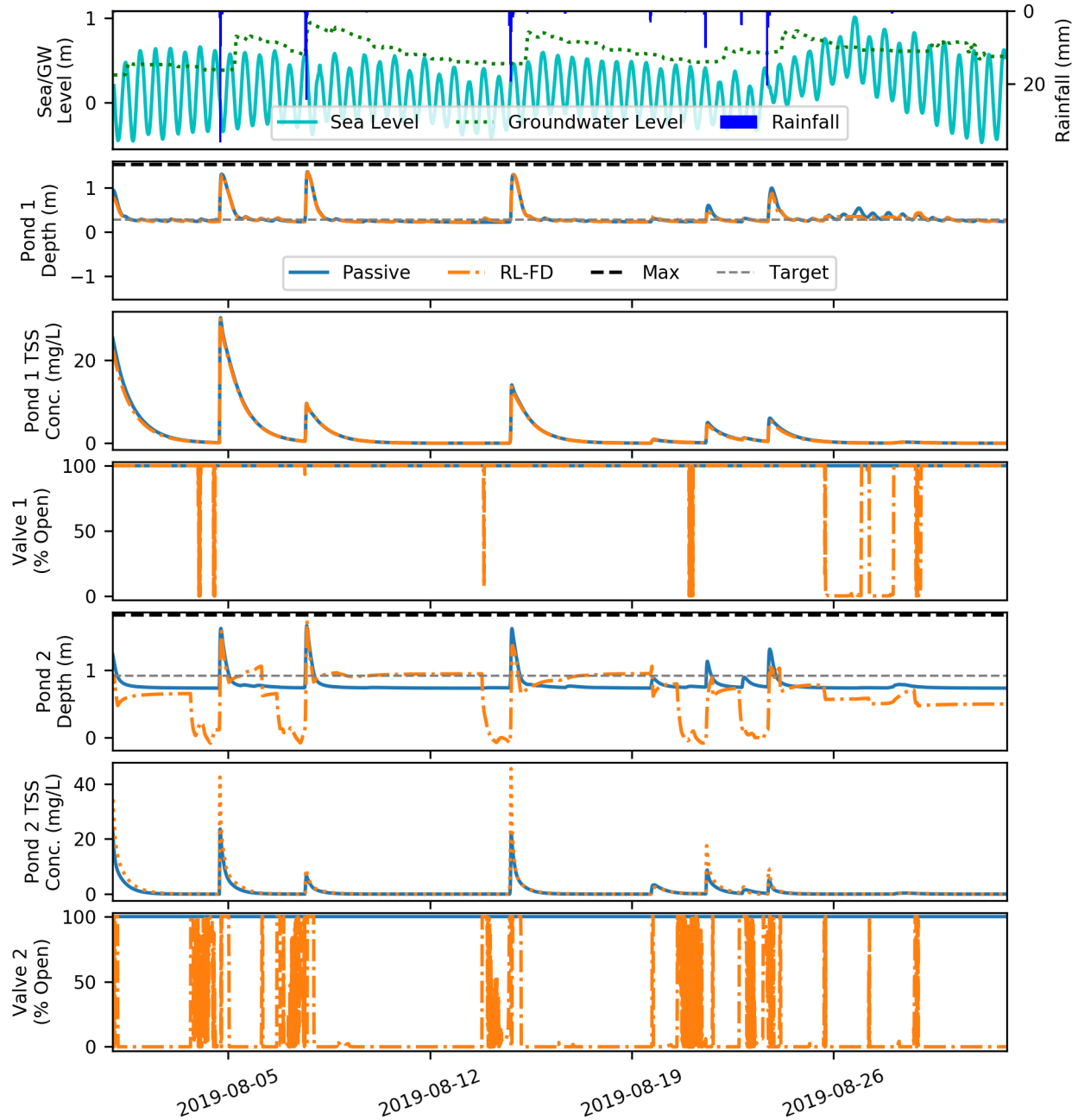
**Table 2.** Comparisons of stormwater control scenarios

Comparison	Control Method	
Baseline	Passive	RL-FD
Local	RBC-DTN	RBC-TSS
System	RL-FDTSS	RL-FD+FDTSS

## 391 **3 RESULTS**

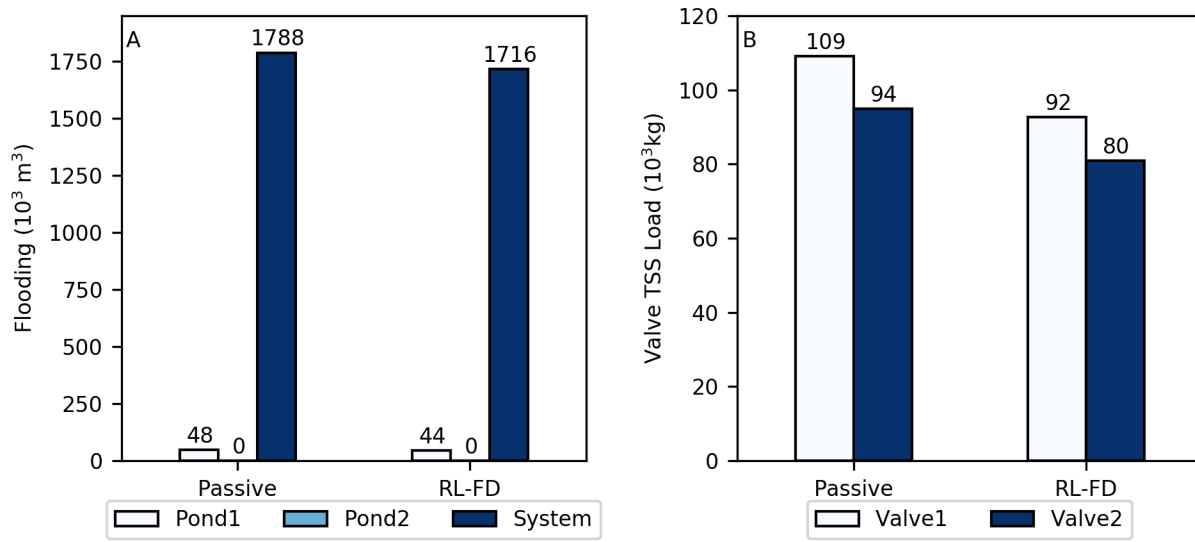
### 392 **3.1 Baseline Flood and TSS Control**

393 Figure 5 illustrates how the passive system and RL-FD respond to the storm events in August, 2019.  
 394 Operation of Pond 1 is similar between these two methods because the controllable valve is at the  
 395 same elevation as the fixed weir; water is released as soon as depth increases from a storm event.  
 396 However, RL-FD learned to close the valve when high tide levels caused backflow into the pond to  
 397 prevent water level fluctuations (e.g., Aug. 26-27). RL-FD learned to lower Pond 2's depth, which  
 398 is fully controllable, before certain storm events (e.g., the Aug. 4 storm) while remaining close to  
 399 the target depth during dry periods.



**Figure 5.** Comparison of passive and RL-FD system operation for August, 2019. From top to bottom, these plots illustrate the hydrological model drivers (rainfall, sea level, and groundwater level) and the depth, TSS concentration, and valve position for Ponds 1 and 2, respectively. In this case, the passive system cannot alter its behavior, while RL-FD can control the valves in response to observed and forecast water quantity conditions.

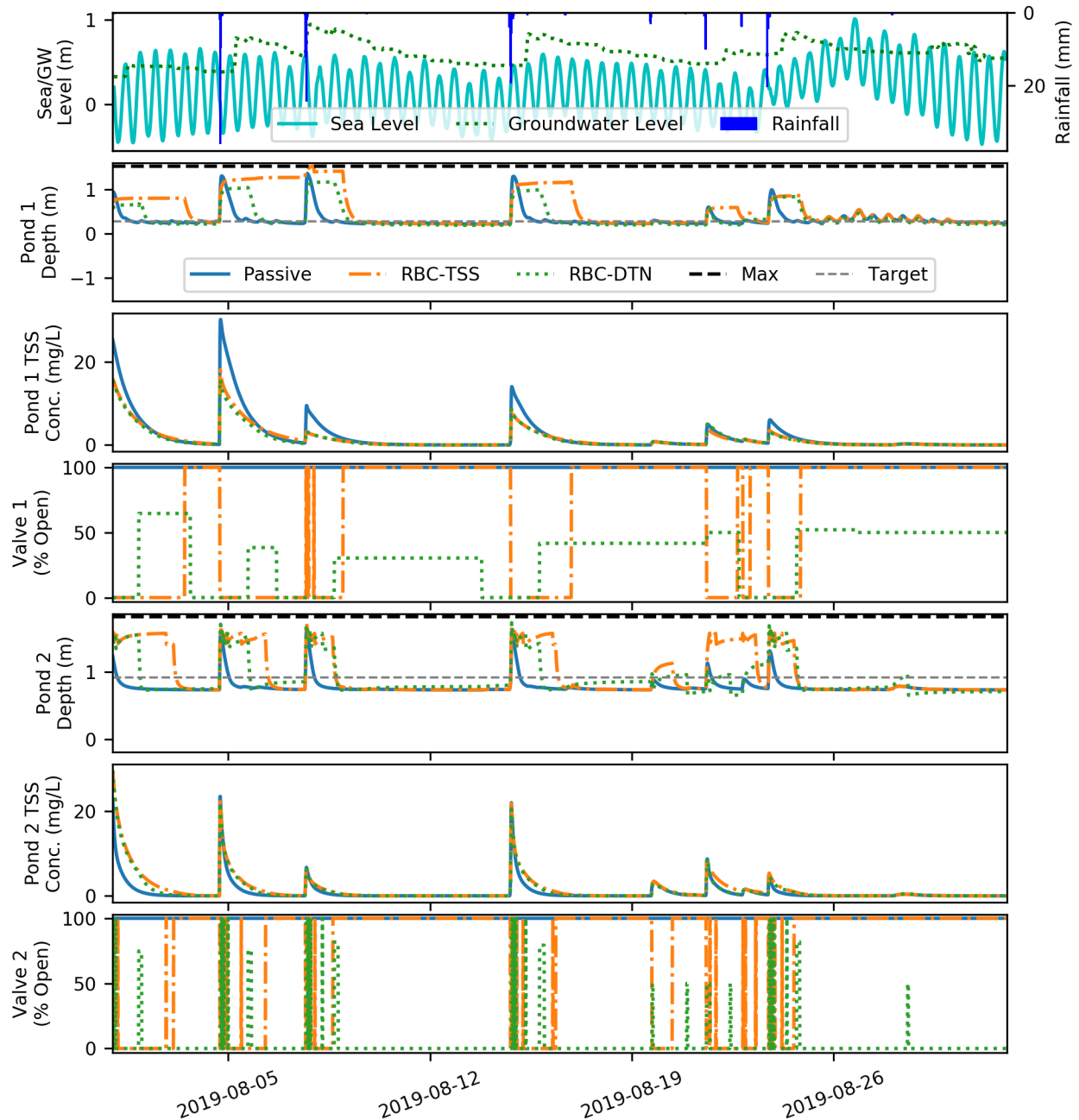
400 The system-level control policy learned by RL-FD allowed it to reduce the total volume of  
 401 flooding by 4.0% (72301m<sup>3</sup>) compared to the passive system (Fig. 6, A). While RL-FD's training  
 402 did not include any water quality information, it's policy does provide improved TSS capture at  
 403 both ponds (i.e., lower loads at the valves). Compared to the passive system, RL-FD reduced TSS  
 404 by 15.1% (16436kg) and 14.8% (14074kg) at Valves 1 and 2, respectively (Fig. 6, B).



**Figure 6.** Total flood volumes (A) and TSS loads (B) for the passive and RL-FD baseline scenarios, 2010-2019.

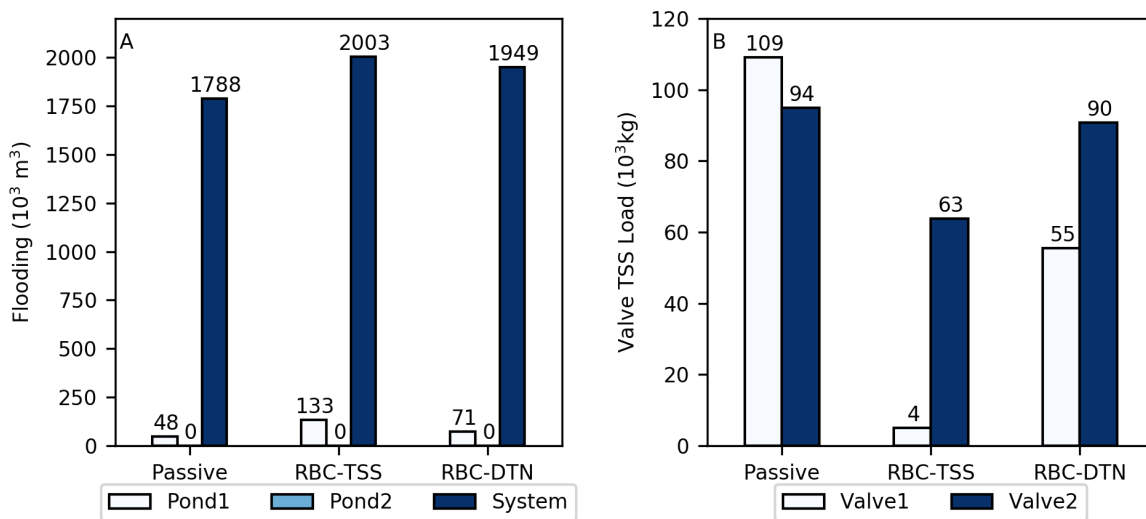
### 405 3.2 Local Control with RBC

406 An example of the RBC methodologies compared to the passive system is shown in Figure 7. Both  
 407 RBC methods operate the ponds individually (i.e., rules are not coordinated between the ponds) to  
 408 mitigate flooding of the pond by releasing water or to improve water quality by retaining runoff after  
 409 a storm event. RBC-DTN has a fixed detention time, while RBC-TSS adapts detention time based  
 410 on the concentration of TSS in the pond. For example, after the Aug. 8 storm both methods retain  
 411 water for similar amounts of time. This indicates that the fixed 24hr retention time of RBC-DTN  
 412 was adequate to treat the TSS washed into the ponds after the short buildup period following the  
 413 Aug. 4 storm. However, after longer periods of TSS buildup, RBC-TSS retains stormwater longer  
 414 than RBC-DTN until TSS is sufficiently treated and the concentrations drop below the threshold  
 415 (e.g., following the Aug. 15 and 22 storms).



**Figure 7.** Comparison of local RTC methods (RBC-TSS, RBC-DTN) and passive system operation for August, 2019. From top to bottom, these plots illustrate the hydrological model drivers (rainfall, sea level, and groundwater level) and the depth, TSS concentration, and valve position for Ponds 1 and 2, respectively. In this case, the passive system cannot alter its behavior, RBC-DTN retains water for a fixed period after storm events to allow settling of TSS, and RBC-TSS adaptively retains water when TSS concentrations are above a threshold.

416 Both rule-based control methods provide reductions in TSS export from the controlled ponds  
 417 compared to the passive system. However, this is at the expense of increased flooding because  
 418 operation of the two valves is not coordinated and does not consider flooding in other parts of the  
 419 stormwater system (Fig. 8). Compared to the passive system, RBC-TSS increased total system flood  
 420 volume by 12.0% (215011m<sup>3</sup>), while decreasing TSS by 95.5% (104222kg) and 32.8% (31116kg)  
 421 at Valves 1 and 2, respectively. RBC-DTN increased flooding by 9.0% (161259kg) and decreased  
 422 TSS for Valves 1 and 2 by 49.2% (53710kg) and 4.5% (4227kg) compared to the passive system.  
 423 RBC for Pond 2 does not treat TSS as efficiently as Pond 1 due in part to the system configuration.  
 424 Water needs to be released if the Pond 2 depth exceeds 1.75m; this is necessary to alleviate upstream  
 425 flooding due to this SWMM model's specific pipe configuration. Further, valve 1 is approximately  
 426 halfway up the side of Pond 1 (i.e., Pond 1 cannot be fully drained), which increases the detention  
 427 time compared to Pond 2.



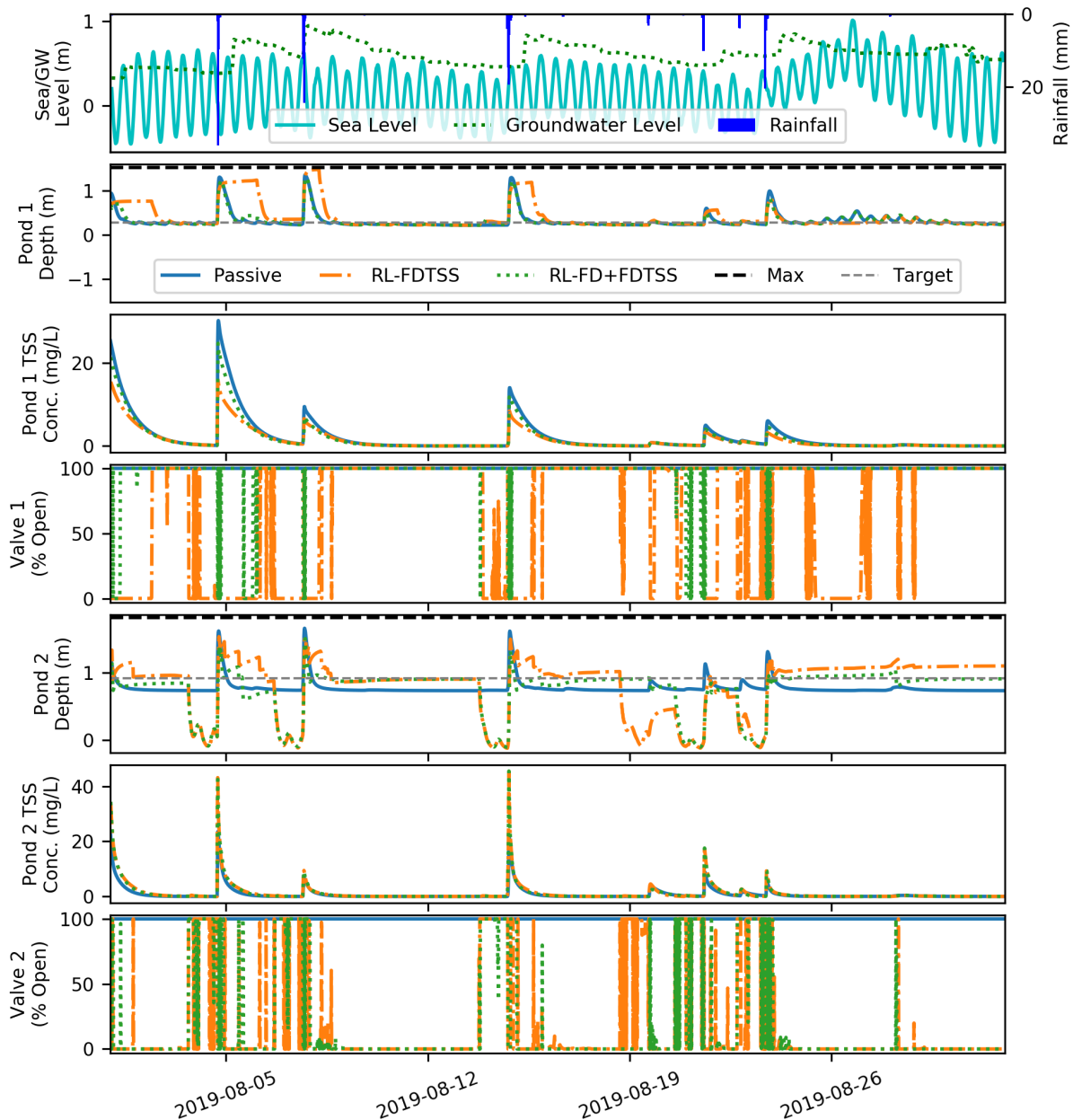
**Figure 8.** Total flood volumes (A) and TSS loads (B) for local RTC methods (RBC-TSS, RBC-DTN) and passive system operation, 2010-2019.

### 428 3.3 System-level Control with RL

429 Both RL-FDTSS and RL-FD+FDTSS learned policies with multiple objectives of flood mitigation,  
 430 TSS reduction, and target pond depths. When tested on the training data (Fig. 9), these agents  
 431 generally kept valve 1 open to maintain the target depth before and between storms (neither agent

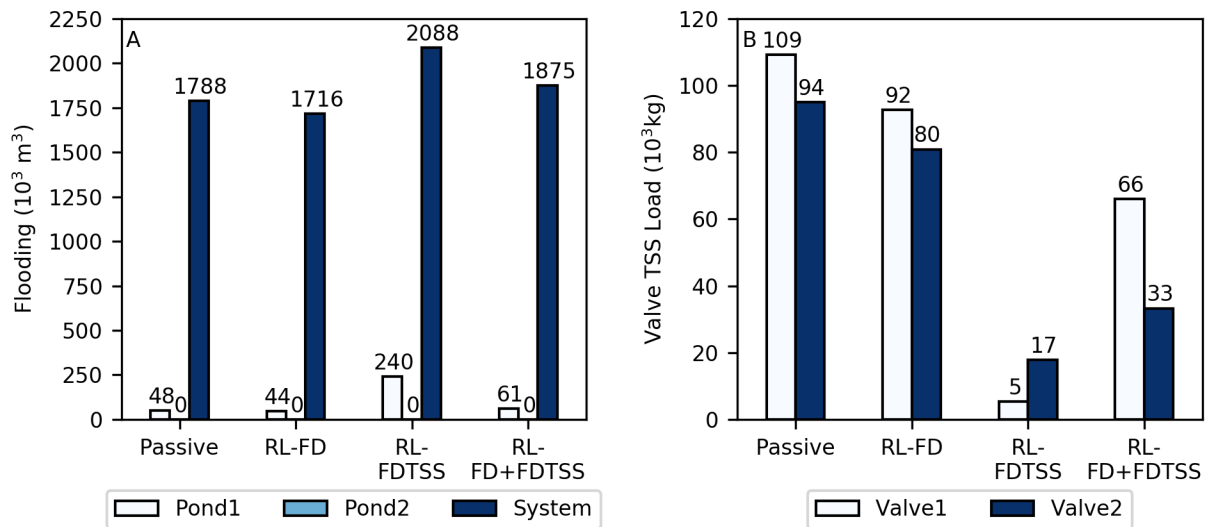
432 can lower the water level in Pond 1 below the target, due to the valve placement) and closed valve 1  
433 during storms to capture TSS. After storm events, RL-FDTSS held water to improve TSS treatment;  
434 in contrast RL-FD+FDTSS closed valve 1 long enough to capture initial TSS inflow, but quickly  
435 released water to return the pond to its target depth. The agents have similar policies for valve 2 that  
436 favor holding water above the target depth to treat TSS while draining the pond before storm events  
437 to prevent flooding. However, RL-FDTSS tended to release water more gradually and hold it at  
438 high levels between storms than RL-FD+FDTSS, increasing TSS treatment.





**Figure 9.** Comparison of RL-FDTSS, RL-FD+FDTSS, and passive system operation for August, 2019. From top to bottom, these plots illustrate the hydrological model drivers (rainfall, sea level, and groundwater level) and the depth, TSS concentration, and valve position for Ponds 1 and 2, respectively. The passive system cannot alter its behavior; RL-FDTSS and RL-FD+FDTSS use water quantity and quality (i.e., TSS observations) information to make control decisions. RL-FD+FDTSS was pre-trained from RL-FD and learned a different balance of flood and TSS control than RL-FDTSS.

439 On the test dataset (2010-2019), RL-FDTSS had 11.3% (212740m<sup>3</sup>) more total system flooding  
 440 and 74.6% (179429m<sup>3</sup>) more flooding at Pond 1 than RL-FD+FDTSS (Fig. 10). Both RL-  
 441 FDTSS and RL-FD+FDTSS increased system-wide flooding compared to the passive system by  
 442 16.8% (300183m<sup>3</sup>) and 4.9% (87443m<sup>3</sup>), respectively. In terms of TSS reduction, both of these  
 443 agents provide improvements compared to the passive system. RL-FDTSS reduced TSS by 95.1%  
 444 (103816kg) and 81.3% (77185kg) at valves 1 and 2, while RL-FD+FDTSS reduced TSS by 39.5%  
 445 (43129kg) and 65.0% (61701kg).

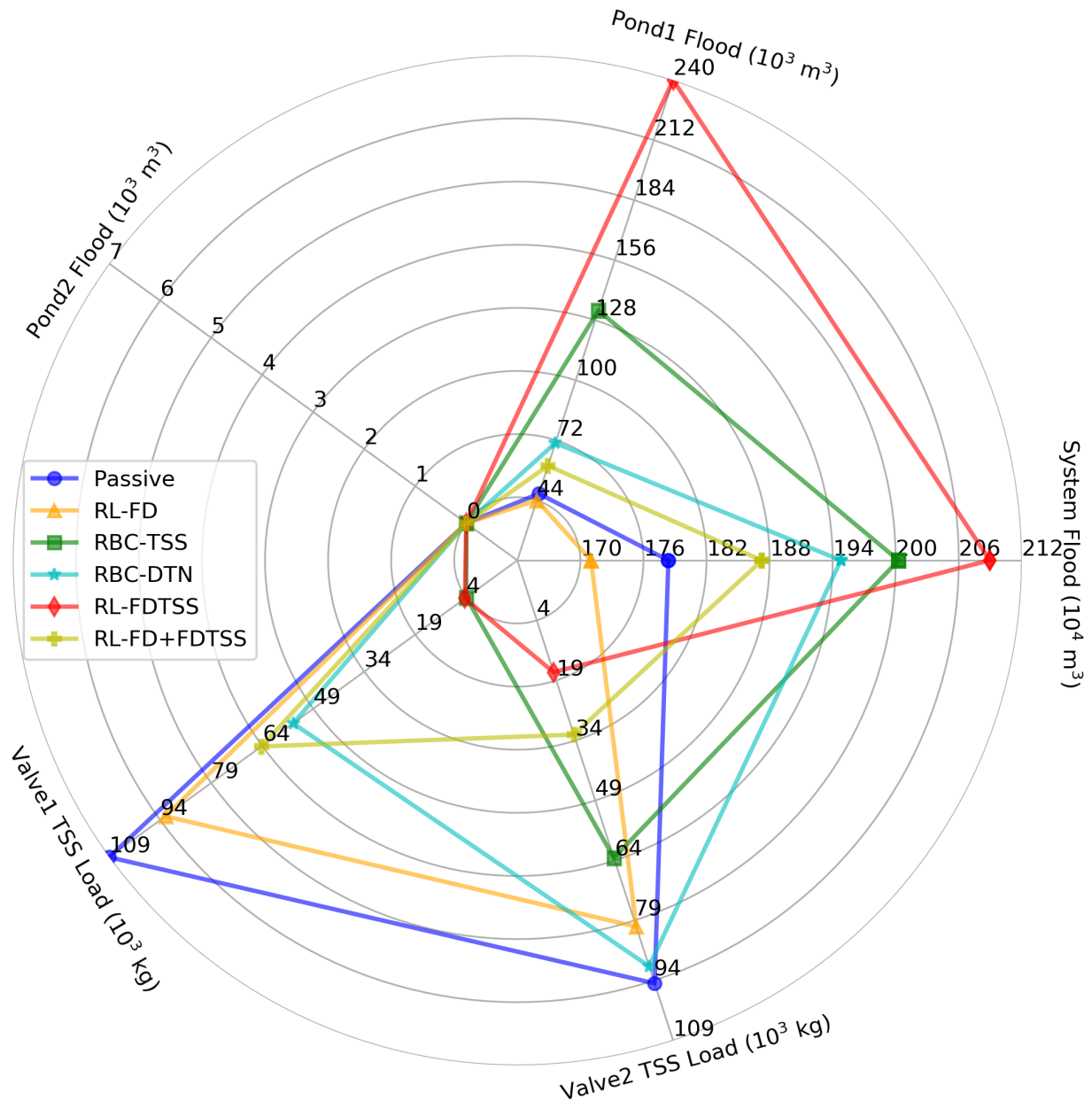


**Figure 10.** Total flood volumes (A) and TSS loads (B) for RL-FDTSS and RL-FD+FDTSS, 2010-2019.

### 446 3.4 Multi-objective Comparison of RTC Methods

447 A comparison of performance trade-offs for each stormwater control method is shown in Figure  
 448 11. In terms of flood volume, only RL-FD reduced flooding compared to the passive system at  
 449 both the system-level and at Pond 1. RL-FD+FDTSS outperformed the local-scale RBC methods  
 450 and RL-FDTSS. Pond 2 did not flood in any of the scenarios because of the configuration of this  
 451 SWMM model; several nodes upstream of Pond 2 have lower maximum depths and flood with any  
 452 rainfall when the pond is above a certain level.

453 All RTC methods reduced TSS loads at both valves compared to the passive system. TSS load  
 454 reduction at valve 1 was greatest for RBC-TSS and RL-FDTSS; RBC-TSS used water quality



**Figure 11.** Comparison of flood volume and TSS load trade-offs for each control method, 2010-2019.

455 observations to inform control, while RL-FDTSS learned a control policy from scratch that included  
 456 penalties for high TSS loads. At valve 2, the local-scale RBC methods had fixed rules to release  
 457 water when Pond 2's depth reached the threshold for upstream flooding. This limited their ability to  
 458 capture the first flush of TSS during large storm events. The system-level RL agents outperformed  
 459 the passive system and had similar trends in performance for both valves. RL-FD did not consider

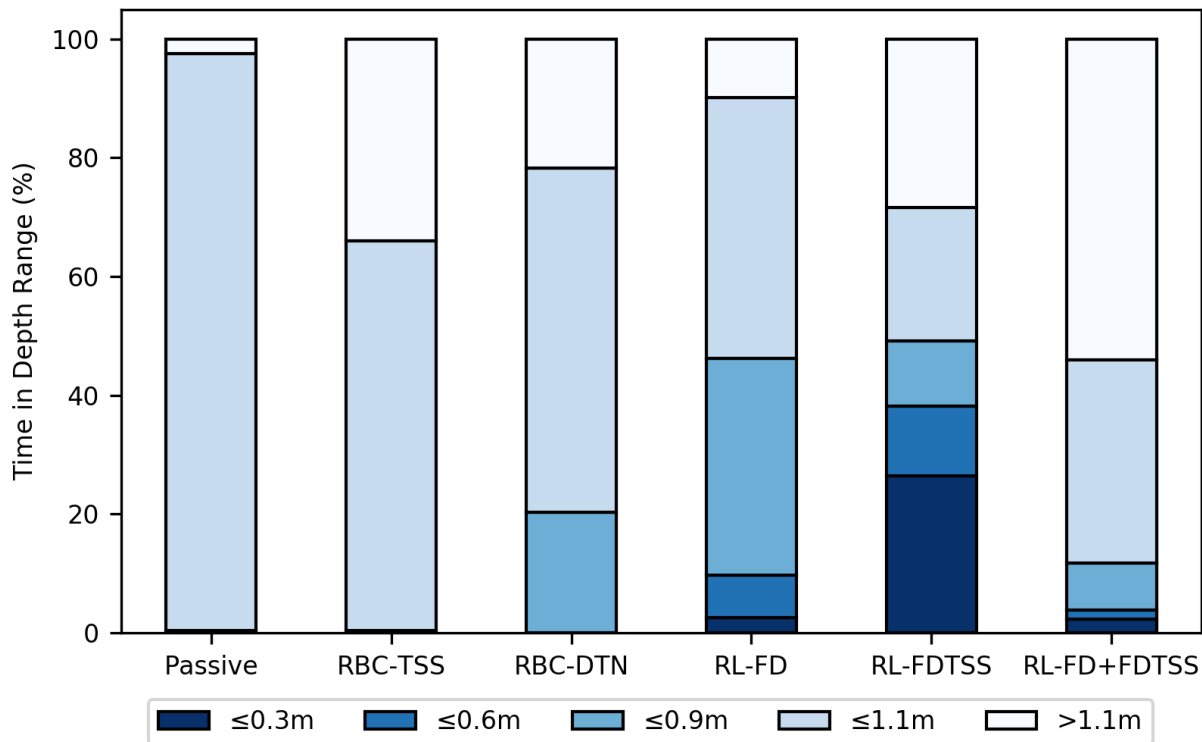
460 TSS in its policy and had the smallest reduction; RL-FD+FDTSS, which had some training with  
 461 the reward function for RL-FDTSS, had more TSS reduction than RL-FD. RL-FDTSS was trained  
 462 from start to finish with a reward function that penalized TSS export from the ponds and had the  
 463 greatest reductions in TSS.

464 The RTC methods made varying degrees of progress towards meeting the city's TMDL TSS  
 465 reduction goals of 5.75, 35, and 60% by 2021, 2026, and 2031, respectively. The percent reductions  
 466 achieved by the RTC methods compared to the passive system are given in Table 3. All RTC  
 467 methods exceeded the 5.75% reduction goal. RL-FD+FDTSS exceeded the 35% goal and both  
 468 RBC-TSS and RL-FDTSS exceeded the 60% goal.

**Table 3.** Percent reduction in total pond TSS export for each RTC method compared to the Passive system.

Control Method	RL-FD	RBC-DTN	RBC-TSS	RL-FDTSS	RL-FD+FDTSS
Reduction (%)	14.94	28.38	66.29	88.66	51.35

469 In terms of maintaining the target depth at Pond 2, RBC-TSS was most similar to the passive  
 470 system because the valve was at the same height as the target depth (Fig. 12). However, RBC-TSS  
 471 was able to close the valve to treat TSS and therefore had a greater percentage of time above the  
 472 target compared to the passive system. RBC-DTN and the RL agents could fully drain or fill Pond 2  
 473 and had a greater percentage of time at lower depths. This helped prevent the pond from flooding,  
 474 but long periods of time at low depths are undesirable in reality. The target depth comparison also  
 475 illustrates differences in policy learned by RL-FDTSS and RL-FD+FDTSS. Across the entire test set,  
 476 RL-FDTSS had a tendency to keep Pond 2 at very low water levels. In contrast, RL-FD+FDTSS's  
 477 policy kept the water level at or above the target depth approximately 90% of the time, indicating  
 478 that it learned a policy to only drain the ponds when needed (a benefit of pretraining RL-FD+FDTSS  
 479 from RL-FD).



**Figure 12.** Comparison of time below or above the Pond 2 target depth (1.1m) for each control method, 2010-2019.

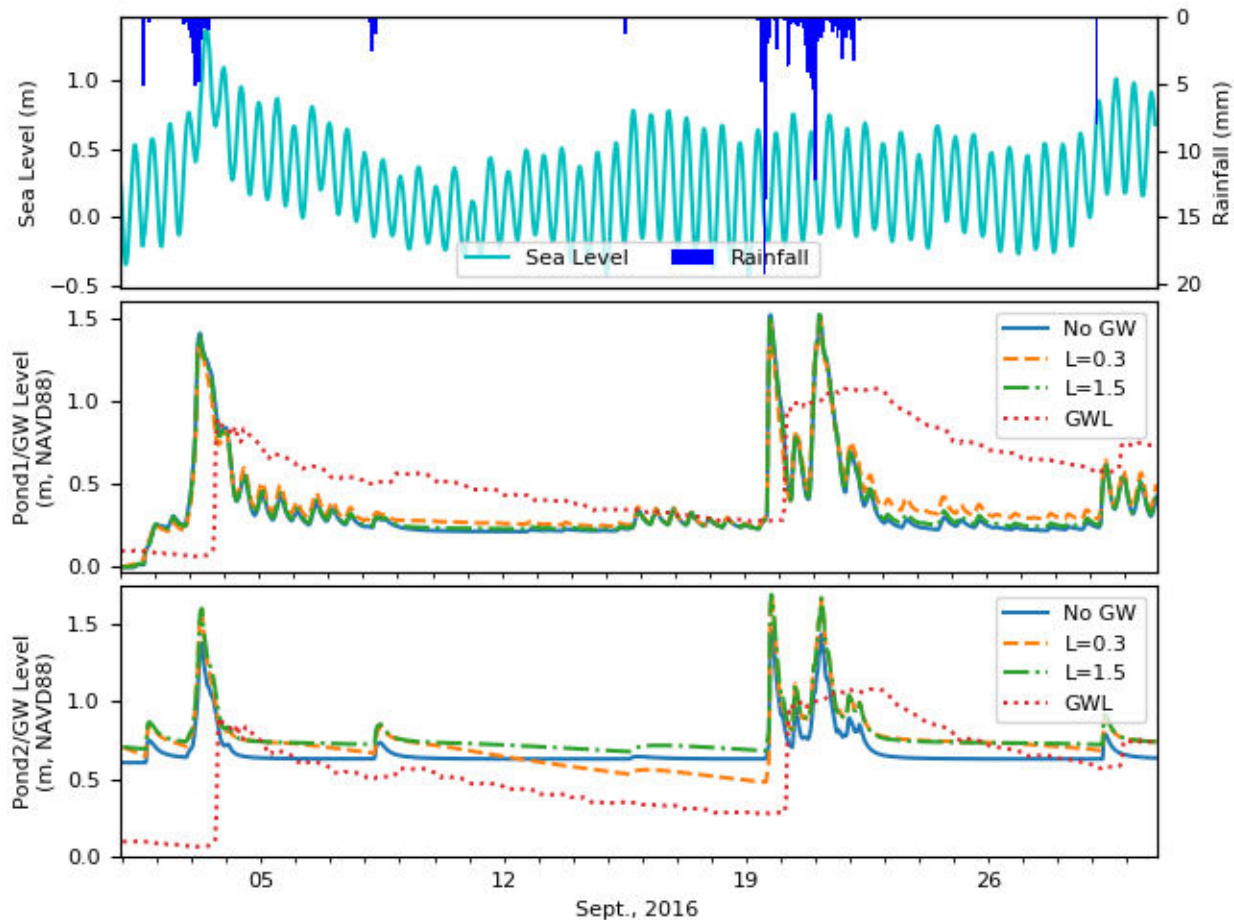
### 480 3.5 Impact of Groundwater Exchange on RTC Methods

481 In comparing the sensitivity of pond-aquifer flow to the Dupuit fitting parameter  $L$ , it was found that  
 482  $L=7.6\text{m}$  and  $L=3.0\text{m}$  had no noticeable impact on the mean depth of Pond 1, while the mean depth  
 483 at Pond 2 increased by 14% (Table 4). When  $L=1.5\text{m}$ , Pond 1 tends to gain a small amount of water,  
 484 while Pond 2 gains slightly less water compared to the larger values of  $L$ . As an example, during  
 485 the dry period without groundwater exchange between Sept. 9 and 19, the water level at both ponds  
 486 is slightly elevated compared to the simulation without groundwater exchange (Fig. 13, No GW).  
 487 When  $L=0.3\text{m}$ , the mean depth at Ponds 1 and 2 increased by 10% and 7%, respectively. This value  
 488 of  $L$  caused total monthly inflow volume to increase at Pond 1 by 24%. At Pond 2, however, total  
 489 monthly inflow volume decreased by 1% and the pond lost water between Sept. 9 and 19 (Fig. 13).  
 490 Because  $L=0.3\text{m}$  had the largest increase in depth at Pond 1 and altered Pond 2's behavior during  
 491 dry weather, it was chosen for use in the RTC simulation with groundwater exchange.

492 Because groundwater exchange also allows increased infiltration, all of the RTC methods have a

**Table 4.** Percent difference in mean pond depth for groundwater exchange simulated with varying values of  $L$  compared to the simulation without groundwater exchange over the month of Sept. 2016.

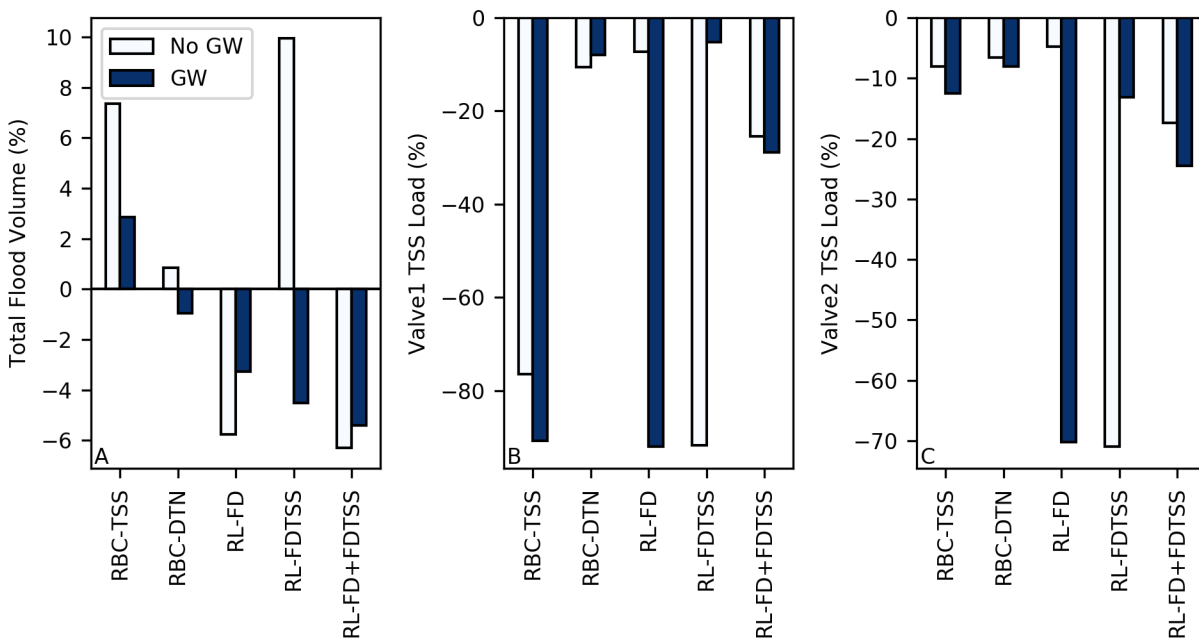
$L$ (m)	7.6	3.0	1.5	0.3
Pond 1	0.0	0.0	+2.8	+10.3
Pond 2	+13.9	+13.9	+12.8	+6.9



**Figure 13.** Comparison of passive pond operation for simulations without groundwater exchange (No GW) and with  $L = 1.5\text{m}$  or  $L = 0.3\text{m}$  in the Dupuit equation, September, 2016.

493 smaller change in total flood volume compared to the passive system when groundwater exchange is  
 494 included (with the exception of RBC-DTN, which had a larger percent change and reduced flooding,  
 495 instead of increasing it) (Fig. 14, A). All RTC methods were still effective at reducing TSS loads

496 for valves 1 and 2 (Fig. 14, B and C, respectively). Of the RBC methods, RBC-DTN had a smaller  
 497 decrease in Valve 1 TSS load with groundwater exchange than without. RL-FDTSS was the only  
 498 RL method to perform worse for TSS reduction when groundwater exchange was added to the  
 499 simulation. This may indicate overfitting to the training data (which did not include groundwater  
 500 exchange), limiting RL-FDTSS's ability to control new pond behaviors. An example time series  
 501 visualization and statistics of valve operation by the RTC methods is available in Appendix A, Figs.  
 502 17 and 18).



**Figure 14.** Comparison of percent difference from the passive system for each RTC method's total flood volume (A) and TSS loads (B and C) for simulations with and without groundwater (GW) exchange for September, 2016.

## 503 4 DISCUSSION

### 504 4.1 Towards System-level Control

505 As the complexity of an environment and control objectives increases, it becomes much harder  
 506 for a single RL agent to learn an effective control policy. This can be seen in the performance  
 507 of RL-FD and RL-FDTSS. RL-FD had fewer goals and a simpler reward function that allowed  
 508 it to learn an effective policy. In contrast, RL-FDTSS had a more complicated reward function

509 and more goals. While it learned an effective policy for minimizing TSS, that was at the expense  
510 of both increasing system flooding and allowing Pond 2 to remain at undesirably low depths for  
511 long periods of time. As demonstrated by RL-FD+FDTSS, pretraining from an agent that performs  
512 well on simpler, but related, goals is one way to approach this challenge. This pretraining allowed  
513 RL-FD+FDTSS to outperform RL-FDTSS for flood mitigation, but at the expense of somewhat  
514 reduced TSS treatment. Other methods such as Multi Agent RL (MARL), Multi-Objective RL  
515 (MORL), and boosting/ensemble methods may also be beneficial. In MARL, each pond could  
516 be controlled by an individual agent tuned to that pond's specific goals, while also operating  
517 cooperatively towards system-level goals<sup>54,55</sup>. In MORL, sets of policies are learned to approximate  
518 a Pareto frontier<sup>56</sup>; this is especially valuable for comparing trade-offs among agents. Similar  
519 multi-objective optimization is well studied for reservoir operation and could provide an alternative  
520 to MORL<sup>57</sup>. Boosting and other ensemble methods attempt to combine agent policies or neural  
521 network outputs to increase performance<sup>58,59</sup>. In the context of RL for stormwater systems, this  
522 maybe beneficial for combining agents that are trained for different purposes (e.g., an agent for  
523 extreme events, an agent for average events, an agent for dry periods).

524 Of the RTC methods implemented here, both RBC-DTN and the RL agents use current observa-  
525 tions and forecasts to inform control decisions ahead of storm events. Perfect forecast data were  
526 used in this research to keep the focus on the control methodology, however, forecasts can contain a  
527 significant amount of uncertainty in reality. As an example specific to coastal systems, tide forecasts  
528 are based on the astronomical tide cycle which does not account for storm tides. In practice, RBC  
529 implementations have handled forecast uncertainty by using a probability threshold (e.g., take an ac-  
530 tion if the rainfall forecast probability is greater than 50%), as well as other fail-safes<sup>14</sup>. Stormwater  
531 RTC research using linear optimization and water quality control rules found that errors in rainfall  
532 prediction (i.e., an unforeseen storm event) could cause flooding of stormwater ponds, but that the  
533 system-level control could quickly adapt and recover based on observations of current conditions<sup>60</sup>.  
534 Recent work with RL (specifically the DDPG algorithm used in this study) for stormwater RTC  
535 has indicated that this algorithm is robust to uncertainty in sensed and forecast data in both training  
536 and testing<sup>31</sup>. While quality-controlled observations could be used in off-line training, doing so  
537 could limit an agent's performance when deployed and using noisy data to inform control actions.  
538 In the current research, the RL agents were robust to altered pond behavior when groundwater



539 exchange was simulated (groundwater exchange was not included in the RL training process).  
540 However, as stormwater RTC continues to move towards system-level control to accommodate the  
541 increasing density of controlled infrastructure components, changing environmental conditions, and  
542 more stringent environmental regulations, understanding the impact of sensed and forecast data  
543 uncertainty on RTC methods will be essential.

544 While sensed and forecast data can be a source of uncertainty, the formulation of RTC methods  
545 can also introduce uncertainty in their performance. For example, both RBC methods use thresholds  
546 to trigger control actions. RBC-DTN uses a time threshold (24 hours) for retaining runoff after  
547 storm events. RBC-TSS uses a TSS concentration threshold to either retain or release water from  
548 the ponds. Changing these thresholds would change the performance of the RBC methods (e.g.,  
549 increased detention time can be expected to increase TSS treatment to a certain extent), however  
550 the exact impact on the performance of the RBC methods used in this study is uncertain. The RL  
551 implementations in this research also include user defined thresholds in their reward functions and  
552 the agents' performance can be very sensitive to these values. In addition, the RL agents benefit  
553 from system-level information when learning their control policies. In practice, sensor networks are  
554 subject to accuracy limitations, communication interruptions, and cyber-cognitive vulnerabilities  
555 (i.e., automated control algorithms, like the RL agents trained here, being used in unexpected  
556 situations that they may not have been trained or tested for)<sup>61</sup>, to name a few sources of uncertainty.

557 RL is known to suffer from issues including reward gaming, where the agent learns to exploit  
558 its environment in unintended ways to gain reward<sup>62</sup>. In the context of stormwater RTC, reward  
559 gaming was observed in early attempts at training RL agents related to simulation processes within  
560 the SWMM model. For example, flood water in the Hague SWMM model does not pond and  
561 reenter the stormwater system as it would in reality, but is simply recorded as flooding and lost  
562 from the simulation. One consequence of this model process is that any TSS within flood water is  
563 also lost from the system. If rewards are poorly shaped (i.e., TSS much more heavily weighted than  
564 flooding), the RL agent can learn policies that induce flooding because the rewards gained by the  
565 corresponding TSS reduction outweigh penalties for flooding. This highlights the need for domain  
566 specific knowledge when crafting reward functions and careful consideration of simplifications  
567 within simulated environments.

568 Beyond water level, flooding, and water quality, more direct monetary costs could be included

569 in the RL reward functions. Some costs of RTC are long term (e.g., the purchase and installation  
570 of sensors and valves, as well as their maintenance), and may not be useful when learning real-  
571 time control policies. A small cost could be incurred for every valve adjustment, which could be  
572 considered in optimization. However, a city may not want to limit valve movements based on a  
573 small cost if it also limits system efficiency. As with trade-offs between flooding and TSS capture,  
574 the balance between limiting and allowing valve movements could be difficult to find. Another  
575 potential cost that could be included is that of dredging retention ponds to remove accumulated  
576 sediment and maintain appropriate capture volumes. Such maintenance may have to become more  
577 frequent with RTC methods that capture sediment, but may still be too long of a time scale when  
578 developing sub-hourly control policies. Additionally, improved water quality and flood mitigation  
579 could offset costs associated with RTC (see, for example,<sup>14,47</sup>).

## 580 **4.2 Trade-offs of Local-scale RBC**

581 Both RBC methods used in this research performed RTC at the local-scale (i.e., operating each  
582 pond individually) and reduced TSS loads, but at the expense of increased system-level flooding.  
583 RBC-DTN showed similar TSS reductions for Pond 1 (49%) as previous studies in other locations  
584 (approximately 40% reported by Marchese et al.<sup>12</sup>. However, as water quality sensor technology  
585 becomes less expensive and more robust<sup>63–66</sup>, control based on water quality observations, such as  
586 the RBC-TSS implemented here, may provide a more adaptive solution. RBC-TSS reduced TSS  
587 by 96% for Pond 1 compared to the passive system, similar to the value found by Sharior et al.<sup>17</sup>  
588 for a different site. The RBC methods did not perform as well for Pond 2 in this study due to the  
589 configuration of the upstream pipes. Specifically, when water reached 1.75m (which is less than the  
590 maximum depth), the contingency rules to prevent upstream flooding would open valve 2. Without  
591 this rule, the RBC methods greatly increased upstream flooding, but it also releases stormwater with  
592 high concentrations of TSS during large storm events.

593 The results of RBC demonstrate that fixed rules, like those used in RBC-DTN, may not provide  
594 the most efficient treatment because pollutants are highly variable between sites and storms<sup>19</sup>. One  
595 solution could be the combination of the two RBC methods used here (e.g., the predictive drawdown  
596 capability of RBC-DTN coupled with adaptive detention time based on observed water quality as  
597 in RBC-TSS), but this is still limited as a local-control scheme. While adapting rules based on

598 water quality may be fairly straight-forward for a single pollutant at a single site, controlling a  
599 stormwater system for multiple pollutants with different treatment processes (e.g., nitrogen species)  
600 will require system-level control<sup>18</sup>. As an example, consider two ponds in series; the upstream  
601 pond is controlled to optimize TSS removal, while the downstream pond is controlled to maintain  
602 anaerobic conditions for denitrification. If the upstream pond retains water to allow settling of  
603 sediment, it might deprive the downstream pond of inflow needed to maintain anaerobic conditions  
604 unless these ponds are operated as a system.

### 605 **4.3 Groundwater Exchange Limitations and Impact**

606 Due to the specific configuration of the studied SWMM model, groundwater exchange was calcu-  
607 lated externally from the SWMM model and added (or subtracted in the case of infiltration) to the  
608 ponds' inflow at each control time step. While this process is based on in-situ soil properties for  
609 Pond 1 in Norfolk's Hague region, the Dupuit equation (which is intended for systems at a steady  
610 state) may not provide the most accurate representation of groundwater exchange. Under real-time  
611 control, ponds can be rapidly drained and refilled before and during a storm event. The Boussinesq  
612 equation for transient unconfined aquifer flow would provide a more realistic representation and  
613 is commonly implemented as a simpler alternative to Richards equation (see, for example,<sup>67</sup>).  
614 Coupling such a model with the SWMM model used here would allow for more precision, but as an  
615 initial demonstration of groundwater impact on ponds controlled in real-time, the Dupuit equation  
616 was quick to implement and run.

617 In the simulated RTC scenarios set up in this research, groundwater exchange with controlled  
618 ponds decreased flooding through infiltration; TSS loads were also reduced because less water  
619 was exiting the ponds through the valves. It should be noted that neither of the RBC methods  
620 were recalibrated to account for groundwater exchange nor were the RL agents retrained with  
621 groundwater exchange being simulated. Retraining with groundwater exchange simulated and  
622 including groundwater observations or forecasts as part of the RL agents' state may allow the agents  
623 to learn more effective policies. However, with the limited impact of groundwater exchange in this  
624 specific simulation, it was not necessary; the RL agents' learned policies and the local scale RBC  
625 methods were robust toward altered pond behavior when groundwater exchange was simulated.

## 626 **5 CONCLUSIONS**

627 In this research, real-time control (RTC) methods are applied to a coastal stormwater infrastructure  
628 system and evaluated on their ability to mitigate flooding and improve water quality by capturing  
629 TSS in controlled retention ponds. The RTC methods used include local control with rules (RBC)  
630 and system-level control with deep reinforcement learning (RL). The impact of groundwater  
631 exchange on the performance of the controlled ponds was evaluated as a condition that may be  
632 important in coastal areas. This research contributes to the growing field of stormwater RTC  
633 by being the first to evaluate the ability of RL to learn system-level control policies considering  
634 both water quantity and water quality goals, as well as being the first to consider the impact of  
635 groundwater on the performance on controlled ponds in a coastal city.

636 Two methods of RBC were used (i) RBC-DTN, which is based on industry standard stormwater  
637 RTC and predictively manages ponds to prevent flooding while retaining runoff for a fixed detention  
638 time to improve water quality and (ii) RBC-TSS, which uses observations of water quality to inform  
639 valve operation in order to improve TSS capture. Both RBC methods are transparent and provide  
640 water quality benefits compared to the passive system. RBC-TSS provided more adaptive operation  
641 and demonstrates the potential for water quality observations to be incorporated with RTC as sensor  
642 technology improves. However, the local operation of both RBC methods caused increased total  
643 system flooding.

644 Three RL agents were trained and tested for their ability to learn effective system-level control  
645 policies. The goal of the first agent (RL-FD) was to mitigate flooding and maintain target pond  
646 depths; it reduced flooding compared to the passive system, but did not consider water quality in  
647 its control policy. The second and third RL agents (RL-FDTSS and RL-FD+FDTSS) attempted  
648 to learn policies for more objectives: mitigate flooding, maintain target pond depths, and reduce  
649 TSS load at the controlled valves. RL-FDTSS learned a policy from scratch, while RL-FD+FDTSS  
650 was pretrained by using the neural network weights and memory from RL-FD, but was trained to  
651 consider water quality as well using the reward function from RL-FDTSS. Both RL-FDTSS and  
652 RL-FD+FDTSS provided water quality benefits but increased flooding compared to the passive  
653 system. RL-FDTSS decreased TSS loads by an average of 88%, but increased system-wide flooding  
654 by 17%. RL-FD+FDTSS's pretraining was effective at reducing training time and allowed it to learn

655 a policy that reduced TSS by an average of 52%, with only a 5% increase in total flood volume,  
656 compared to the passive system.

657 Given the growing adoption of rule-based stormwater RTC and the ability of RL to learn system-  
658 level control policies, future work could investigate control of more complex stormwater systems and  
659 integrations of RL and RBC. More complex stormwater systems could include retention ponds in  
660 series, pollutants that are treated through chemical and biological processes (e.g., nitrogen)/multiple  
661 pollutants, and different controllable assets such as pumps. Integration of RL and RBC could  
662 include using RL to better parameterize variables within an existing control rule (see Likmeta, et  
663 al.,<sup>68</sup> for an example in autonomous vehicles), as well as adding or removing rules from a set of  
664 rules. These avenues for future research could allow stormwater RTC providers to increase the  
665 complexity of controlled networks, improving flood mitigation and water quality, while maintaining  
666 the operational transparency needed for critical stormwater infrastructure systems.

## 667 **DATA, MODEL, AND CODE AVAILABILITY**

668 The data, models, and code used in this study are available on GitHub at [https://github.com/](https://github.com/UVAdMIST/swmm_wq_rl)  
669 [UVAdMIST/swmm\\_wq\\_rl](https://github.com/UVAdMIST/swmm_wq_rl).

## 670 **CONFLICTS OF INTEREST**

671 There are no conflicts of interest to declare.

## 672 **ACKNOWLEDGMENTS**

673 This work was funded as part of two National Science Foundation grants: Award #1735587 (CRISP-  
674 Critical, Resilient Interdependent Infrastructure Systems and Processes) and Award #1737432  
675 (SCC-IRG Track 1: Overcoming Social and Technical Barriers for the Broad Adoption of Smart  
676 Stormwater Systems). We acknowledge HRSD for continued access to their high quality data and  
677 the City of Norfolk for information regarding their stormwater retention ponds.

## 678 A ADDITIONAL FIGURES

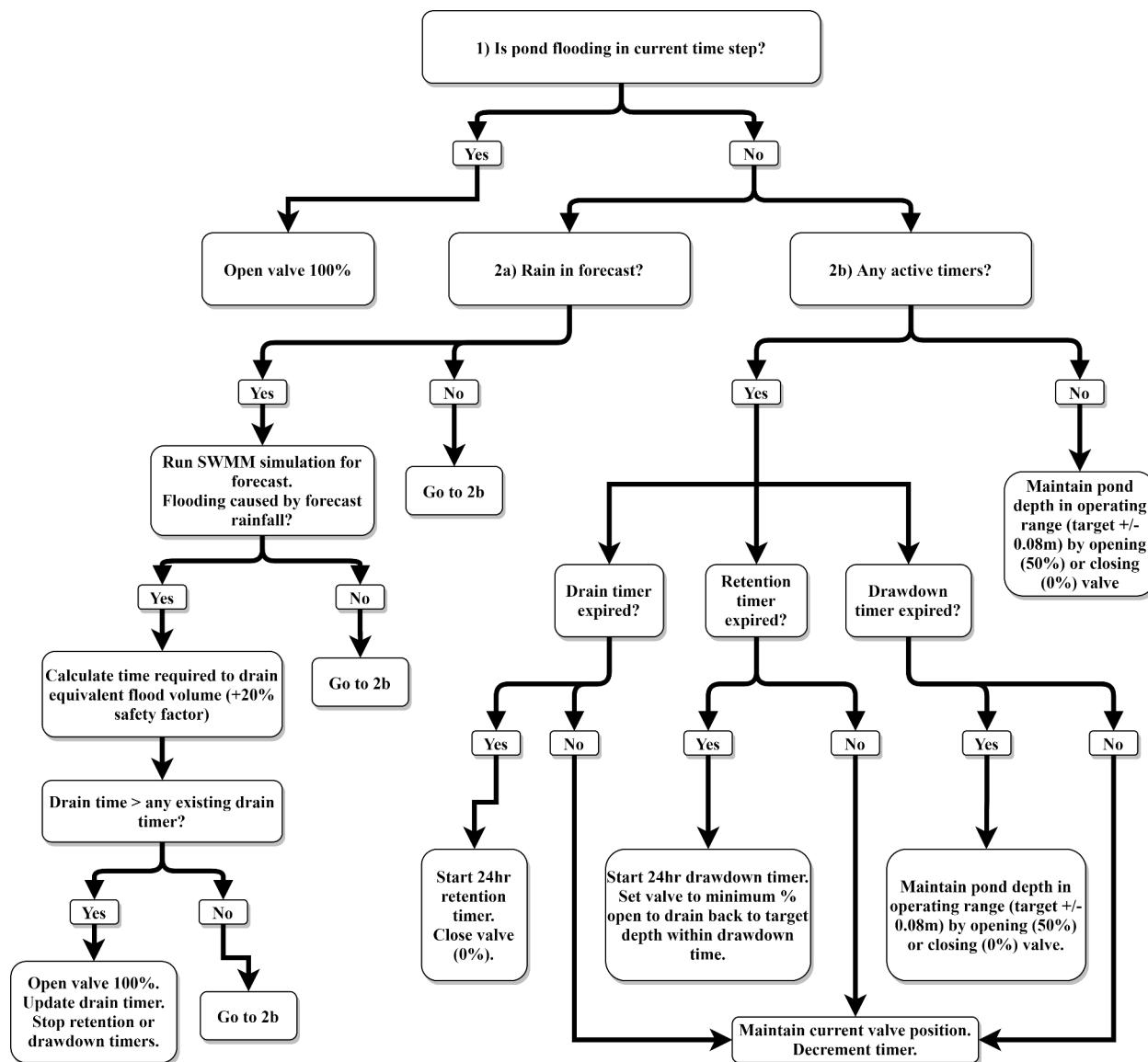
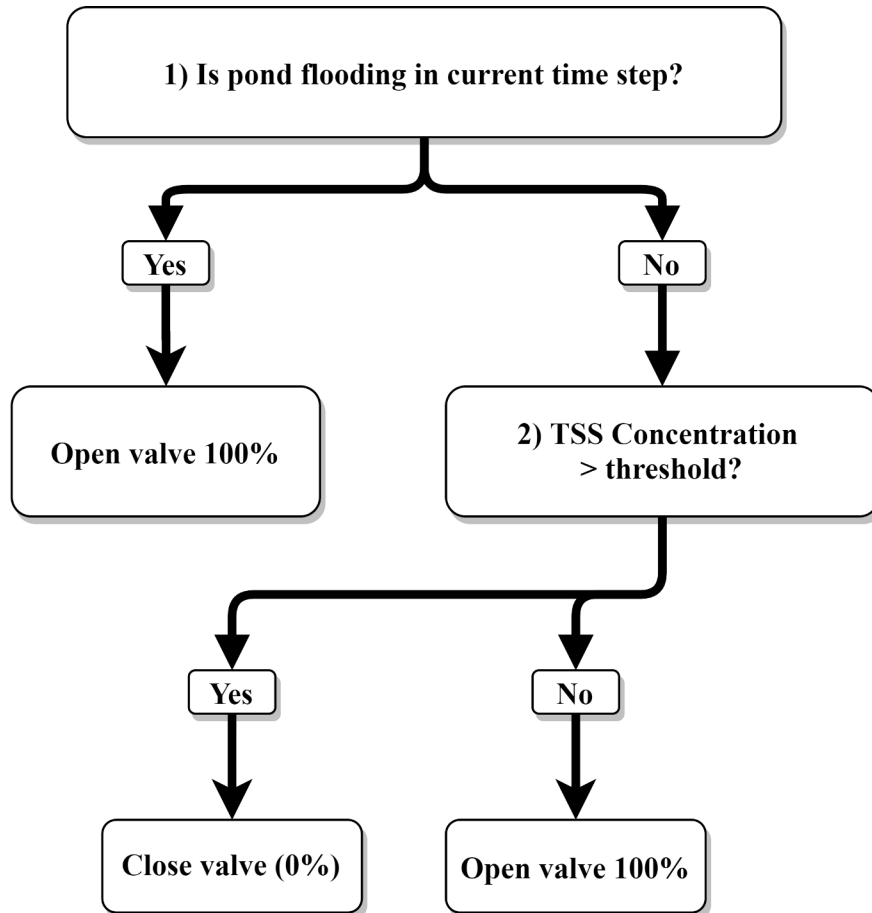
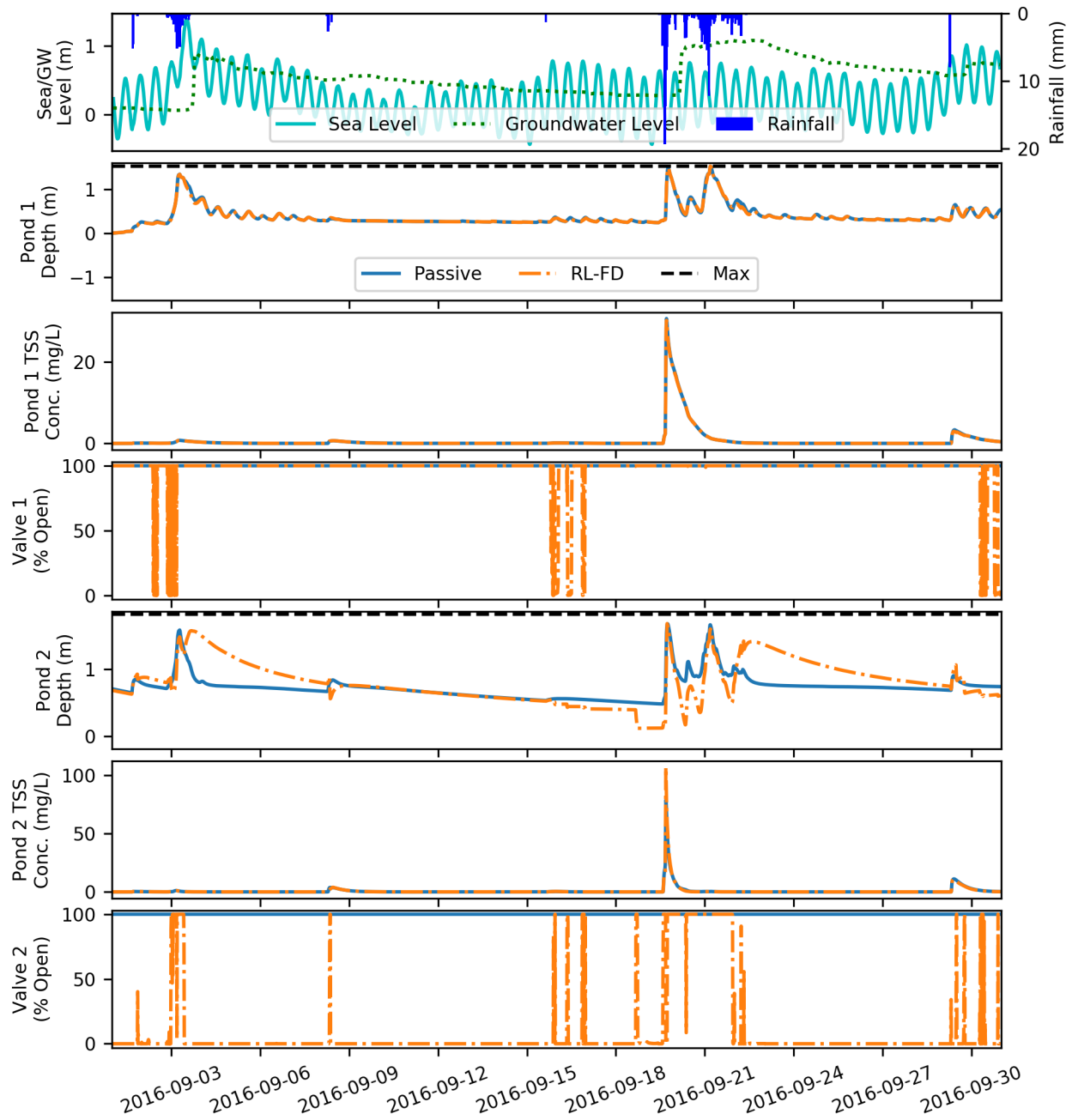


Figure 15. RBC-DTN decision tree (adapted from<sup>10</sup>).

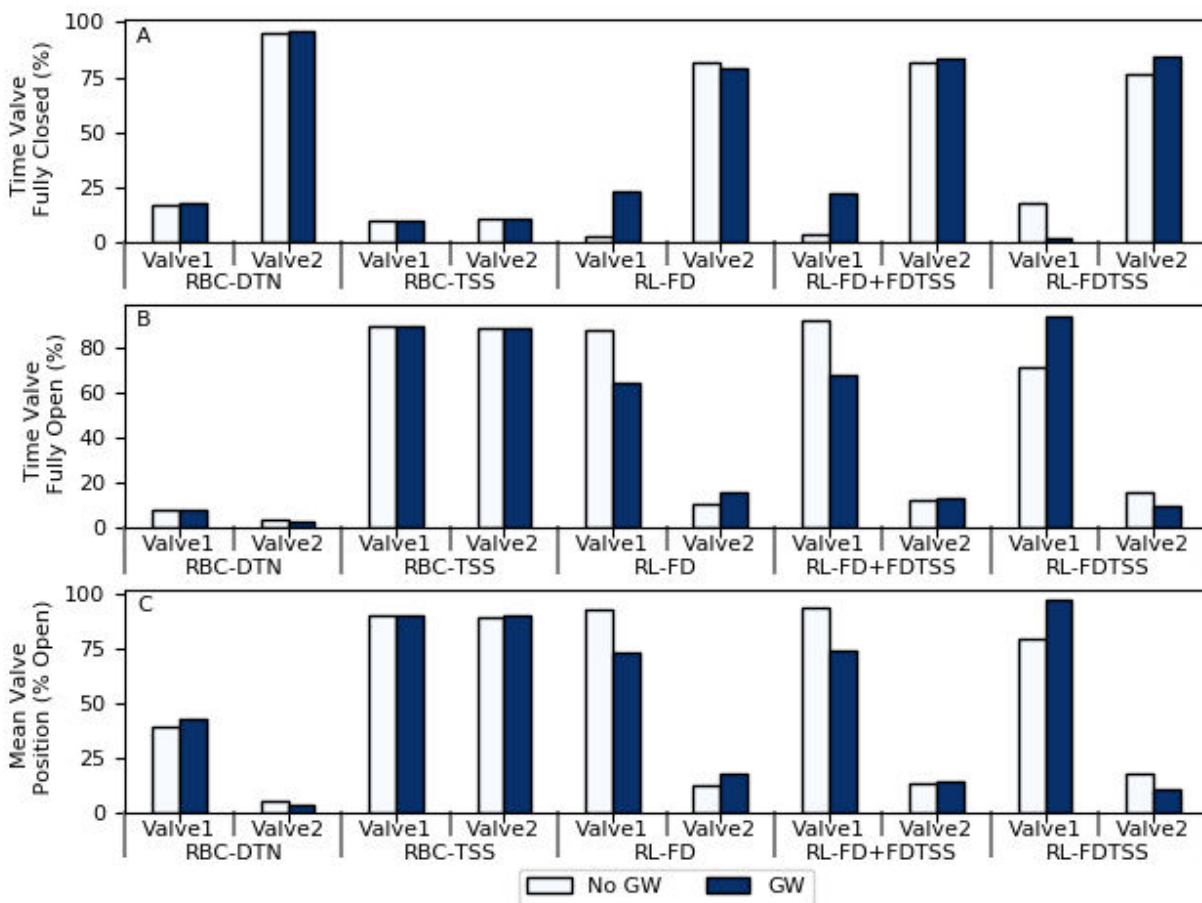


**Figure 16.** RBC-TSS decision tree.



**Figure 17.** Comparison of RL-FD and passive system operation for September, 2016 with groundwater exchange at the controlled ponds.





**Figure 18.** Comparison of control policies (% of time a valve is fully closed (A), fully open (B), and the mean valve position (C)) for simulations with and without groundwater (GW) exchange for September, 2016.

**REFERENCES**

- 679
- 680 1 Sweet WV, Park J. From the extreme to the mean: Acceleration and tipping points of coastal  
681 inundation from sea level rise. *Earth's Future*. 2014 12;2(12):579-600. Available from: <http://doi.wiley.com/10.1002/2014EF000272>.  
682
  - 683 2 Moftakhari HR, AghaKouchak A, Sanders BF, Feldman DL, Sweet W, Matthew RA, et al.  
684 Increased nuisance flooding along the coasts of the United States due to sea level rise: Past  
685 and future. *Geophysical Research Letters*. 2015 11;42(22):9846-52. Available from: <http://doi.wiley.com/10.1002/2015GL066072>.  
686
  - 687 3 Moftakhari HR, AghaKouchak A, Sanders BF, Matthew RA. Cumulative hazard: The case of  
688 nuisance flooding. *Earth's Future*. 2017 2;5(2):214-23. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/2016EF000494>.  
689
  - 690 4 Alamdari N, Sample DJ, Ross AC, Easton ZM. Evaluating the Impact of Climate Change on  
691 Water Quality and Quantity in an Urban Watershed Using an Ensemble Approach. *Estuaries  
692 and Coasts*. 2020 1;43(1):56-72. Available from: <https://doi.org/10.1007/s12237-019-00649-4>.
  - 693 5 Kerkez B, Gruden C, Lewis M, Montestruque L, Quigley M, Wong B, et al. Smarter Stormwater  
694 Systems. *Environmental Science and Technology*. 2016;50:72677273. Available from: <https://pubs.acs.org/doi/10.1021/acs.est.5b05870>.  
695
  - 696 6 Troutman SC, Love NG, Kerkez B. Balancing water quality and flows in combined sewer  
697 systems using real-time control. *Environmental Science: Water Research and Technology*.  
698 2020 5;6(5):1357-69. Available from: [https://pubs.rsc.org/en/content/articlehtml/2020/ew/  
699 c9ew00882a](https://pubs.rsc.org/en/content/articlehtml/2020/ew/c9ew00882a).
  - 700 7 Kroll S, Weemaes M, Van Impe J, Willems P. A Methodology for the Design of RTC Strategies  
701 for Combined Sewer Networks. *Water*. 2018 11;10(11):1675. Available from: [http://www.mdpi.  
702 com/2073-4441/10/11/1675](http://www.mdpi.com/2073-4441/10/11/1675).
  - 703 8 Montestruque L, Lemmon MD. Globally Coordinated Distributed Storm Water Management  
704 System. In: *Proceedings of the 1st ACM International Workshop on Cyber-Physical Systems  
705 for Smart Water Networks*. New York, NY, USA: ACM; 2015. p. 1-6. Available from: <https://dl.acm.org/doi/10.1145/2738935.2738948>.  
706
  - 707 9 Sadler JM, Goodall JL, Behl M, Bowes BD, Morsy MM. Exploring real-time control of

- 708 stormwater systems for mitigating flood risk due to sea level rise. *Journal of Hydrology*.  
709 2020 4;583(124571):124571. Available from: [https://linkinghub.elsevier.com/retrieve/pii/  
710 S0022169420300317](https://linkinghub.elsevier.com/retrieve/pii/S0022169420300317).
- 711 10 Bowes BD, Tavakoli A, Wang C, Heydarian A, Behl M, Beling PA, et al. Flood mitigation in  
712 coastal urban catchments using real-time stormwater infrastructure control and reinforcement  
713 learning. *Journal of Hydroinformatics*. 2021 5;23(3):529-47. Available from: [https://iwaponline.  
714 com/jh/article/23/3/529/77759/Flood-mitigation-in-coastal-urban-catchments-using](https://iwaponline.com/jh/article/23/3/529/77759/Flood-mitigation-in-coastal-urban-catchments-using).
- 715 11 Wong BP, Kerkez B. Real-Time Control of Urban Headwater Catchments Through Linear Feed-  
716 back: Performance, Analysis, and Site Selection. *Water Resources Research*. 2018;54(10):7309-  
717 30. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR022657>.
- 718 12 Marchese D, Johnson J, Akers N, Huffman M, Hlas V. Quantitative Comparison of Ac-  
719 tive and Passive Stormwater Infrastructure: Case Study in Beckley, West Virginia. *Pro-  
720 ceedings of the Water Environment Federation*. 2018 1;2018(9):4298-311. Available from:  
721 [https://accesswater.org/publications/-300096/quantitative-comparison-of-active-and-passive-  
722 stormwater-infrastructure--case-study-in-beckley--west-virginia](https://accesswater.org/publications/-300096/quantitative-comparison-of-active-and-passive-stormwater-infrastructure--case-study-in-beckley--west-virginia).
- 723 13 Shishegar S, Duchesne S, Pelletier G. An integrated optimization and rule-based approach for  
724 predictive real time control of urban stormwater management systems. *Journal of Hydrology*.  
725 2019;577:124000.
- 726 14 OptiRTC, Geosyntec Consultants Inc . Water Quality Summary Report National Fish and  
727 Wildlife Foundation Smart, Integrated Stormwater Management Systems Anacostia River  
728 Watershed Water Quality Study; 2017. Available from: [www.optirtc.com](http://www.optirtc.com).
- 729 15 Muschalla D, Vallet B, Anctil F, Lessard P, Pelletier G, Vanrolleghem PA. Ecohydraulic-driven  
730 real-time control of stormwater basins. *Journal of Hydrology*. 2014 4;511:82-91.
- 731 16 Gaborit E, Muschalla D, Vallet B, Vanrolleghem PA, Anctil F. Improving the performance of  
732 stormwater detention basins by real-time control using rainfall forecasts. *Urban Water Journal*.  
733 2013 8;10(4):230-46.
- 734 17 Sharior S, McDonald W, Parolari AJ. Improved reliability of stormwater detention basin  
735 performance through water quality data-informed real-time control. *Journal of Hydrology*.  
736 2019;573:422-31. Available from: [https://www.sciencedirect.com/science/article/pii/  
737 S0022169419302598](https://www.sciencedirect.com/science/article/pii/S0022169419302598).

- 738 18 Mullapudi A, Wong BP, Kerkez B. Emerging investigators series: building a theory for smart  
739 stormwater systems. *Environmental Science: Water Research & Technology*. 2017 1;3(1):66-77.  
740 Available from: <http://xlink.rsc.org/?DOI=C6EW00211K>.
- 741 19 Wong BP, Kerkez B. Adaptive measurements of urban runoff quality. *Water Resources Research*.  
742 2016 11;52(11):8986-9000. Available from: <http://doi.wiley.com/10.1002/2015WR018013>.
- 743 20 Chen Y, Han D. Water quality monitoring in smart city: A pilot project. *Automation in*  
744 *Construction*. 2018 5;89:307-16.
- 745 21 Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, Mas-  
746 sachusetts: The MIT Press; 2018.
- 747 22 Lee JH, Labadie JW. Stochastic optimization of multireservoir systems via reinforcement  
748 learning. *Water Resources Research*. 2007;43(11). Available from: <http://doi.wiley.com/10.1029/2006WR005627>.
- 750 23 Castelletti A, Yajima H, Giuliani M, Soncini-Sessa R, Weber E. Planning the Optimal Operation  
751 of a Multioutlet Water Reservoir with Water Quality and Quantity Targets. *Journal of Water*  
752 *Resources Planning and Management*. 2014;140(4):496-510. Available from: <https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29WR.1943-5452.0000348>.
- 754 24 Castelletti A, Pianosi F, Restelli M. A multiobjective reinforcement learning approach to water  
755 resources systems operation: Pareto frontier approximation in a single run. *Water Resources*  
756 *Research*. 2013;49:3476-86. Available from: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/wrcr.20295>.
- 758 25 Pianosi F, Castelletti A, Restelli M. Tree-based fitted Q-iteration for multi-objective Markov  
759 decision processes in water resource management. *Journal of Hydroinformatics*. 2013;15(2):258-  
760 70. Available from: <https://iwaponline.com/jh/article-pdf/15/2/258/386917/258.pdf>.
- 761 26 Delipetrev B, Jonoski A, Solomatine DP. A novel nested stochastic dynamic program-  
762 ming (nSDP) and nested reinforcement learning (nRL) algorithm for multipurpose reser-  
763 voir optimization. *Journal of Hydroinformatics*. 2017;19(1):47-61. Available from: <https://iwaponline.com/jh/article-pdf/19/1/47/390803/jh0190047.pdf>.
- 765 27 Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level  
766 control through deep reinforcement learning. *Nature*. 2015;518:529-43. Available from:  
767 <https://www.nature.com/articles/nature14236.pdf>.

- 768 28 Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, et al. Continuous control with  
769 deep reinforcement learning. *International Conference on Learning Representations*. 2015 9:14.  
770 Available from: <https://goo.gl/J4PIAzhttp://arxiv.org/abs/1509.02971>.
- 771 29 Mullapudi A, Lewis MJ, Gruden CL, Kerkez B. Deep reinforcement learning for the real time  
772 control of stormwater systems. *Advances in Water Resources*. 2020 6;140:103600. Available  
773 from: <https://linkinghub.elsevier.com/retrieve/pii/S0309170820302499>.
- 774 30 Wang C, Bowes B, Tavakoli A, Adams S, Goodall J, Beling P. Smart Stormwater Control  
775 Systems: A Reinforcement Learning Approach. In: Hughes AL, McNeill F, Zobel C, editors.  
776 ISCRAM 2020 Conference Proceedings - 17th International Conference on Information Systems  
777 for Crisis Response and Management. Blacksburg, VA; 2020. p. 2-13.
- 778 31 Saliba SM, Bowes BD, Adams S, Beling PA, Goodall JL. Deep Reinforcement Learning with  
779 Uncertain Data for Real-Time Stormwater System Control and Flood Mitigation. *Water*. 2020  
780 11;12(11):3222. Available from: <https://www.mdpi.com/2073-4441/12/11/3222>.
- 781 32 Eggleston J, Pope J. Land Subsidence and Relative Sea-Level Rise in the Southern Chesapeake  
782 Bay Region. Reston, Virginia: U.S. Geological Survey; 2013. Available from: <https://pubs.usgs.gov/circ/1392/pdf/circ1392.pdf>.
- 783
- 784 33 Bowes BD, Sadler JM, Morsy MM, Behl M, Goodall JL. Forecasting Groundwater Table in  
785 a Flood Prone Coastal City with Long Short-term Memory and Recurrent Neural Networks.  
786 *Water*. 2019;11(5):1098. Available from: <https://www.mdpi.com/2073-4441/11/5/1098>.
- 787 34 Sadler JM, Goodall JL, Morsy MM, Spencer K. Modeling urban coastal flood severity from  
788 crowd-sourced flood reports using Poisson regression and Random Forest. *Journal of Hy-*  
789 *drology*. 2018 4;559:43-55. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0022169418300519>.
- 790
- 791 35 Chesapeake Bay Foundation. State of the Bay. Chesapeake Bay Foundation; 2018. Available  
792 from: <https://www.cbf.org/document-library/cbf-reports/2018-state-of-the-bay-report.pdf>.
- 793 36 Murphy RR, Kemp WM, Ball WP. Long-Term Trends in Chesapeake Bay Seasonal Hypoxia,  
794 Stratification, and Nutrient Loading. *Estuaries and Coasts*. 2011 11;34(6):1293-309. Available  
795 from: <https://link.springer.com/article/10.1007/s12237-011-9413-7>.
- 796 37 CHESAPEAKE BAY TMDL ACTION PLAN VSMP MS4 Permit No. VA0088650. Norfolk:  
797 City of Norfolk; 2018. Available from: <https://www.norfolk.gov/DocumentCenter/View/38025/>

- 798 Final-Report---Chesapeake-Bay-TMDL-Action-Plan---06\_28\_2018\_FINAL?bidId=.
- 799 38 Virginia Geographic Information Network. Virginia Land Cover Dataset; 2016. Available from:  
800 <https://vgin.maps.arcgis.com/home/item.html?id=d3d51bb5431a4d26a313f586c7c2c848>.
- 801 39 Davtalab R, Mirchi A, Harris RJ, Troilo MX, Madani K. Sea Level Rise Effect on Groundwater  
802 Rise and Stormwater Retention Pond Reliability. *Water*. 2020 4;12(4):1129. Available from:  
803 <https://www.mdpi.com/2073-4441/12/4/1129>.
- 804 40 McDonnell B, Ratliff K, Tryby M, Wu J, Mullapudi A. PySWMM: The Python Interface  
805 to Stormwater Management Model (SWMM). *Journal of Open Source Software*. 2020  
806 8;5(52):2292. Available from: <https://joss.theoj.org/papers/10.21105/joss.02292>.
- 807 41 Pells SE, N Pells PJ. Application of Dupuit's Equation in SWMM to Simulate Baseflow. *Journal*  
808 *of Hydrologic Engineering*. 2016 1;21(1):06015009. Available from: [https://ascelibrary.org/doi/](https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0001245)  
809 [abs/10.1061/%28ASCE%29HE.1943-5584.0001245](https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0001245).
- 810 42 Rossman LA, Huber WC. Storm Water Management Model Reference Manual Volume III –  
811 Water Quality. Cincinnati: USEPA; 2016.
- 812 43 Guan M, Ahilan S, Yu D, Peng Y, Wright N. Numerical modelling of hydro-morphological  
813 processes dominated by fine suspended sediment in a stormwater pond. *Journal of Hydrology*.  
814 2018 1;556:87-99.
- 815 44 Tetra Tech. Stormwater Best Management Practices (BMP) Performance Analysis. USEPA;  
816 2010. Available from: [https://www3.epa.gov/region1/npdes/stormwater/assets/pdfs/BMP-](https://www3.epa.gov/region1/npdes/stormwater/assets/pdfs/BMP-Performance-Analysis-Report.pdf)  
817 [Performance-Analysis-Report.pdf](https://www3.epa.gov/region1/npdes/stormwater/assets/pdfs/BMP-Performance-Analysis-Report.pdf).
- 818 45 of Norfolk V. CHESAPEAKE BAY TMDL ACTION PLAN VSMP MS4 Permit No.  
819 VA0088650. Norfolk; 2018. Available from: [https://www.norfolk.gov/DocumentCenter/View/](https://www.norfolk.gov/DocumentCenter/View/38025/Final-Report---Chesapeake-Bay-TMDL-Action-Plan---06_28_2018_FINAL?bidId=)  
820 [38025/Final-Report---Chesapeake-Bay-TMDL-Action-Plan---06\\_28\\_2018\\_FINAL?bidId=](https://www.norfolk.gov/DocumentCenter/View/38025/Final-Report---Chesapeake-Bay-TMDL-Action-Plan---06_28_2018_FINAL?bidId=).
- 821 46 Virginia Department of Environmental Quality. Chesapeake Bay TMDL Action Plan Guidance;  
822 2015.
- 823 47 Wright J, Marchese D. Briefing: Continuous monitoring and adaptive control: the 'smart' storm  
824 water management solution. *Proceedings of the Institution of Civil Engineers - Smart Infras-*  
825 *tructure and Construction*. 2017 12;170(4):86-9. Available from: [https://www.icevirtuallibrary.](https://www.icevirtuallibrary.com/doi/10.1680/jsmic.17.00017)  
826 [com/doi/10.1680/jsmic.17.00017](https://www.icevirtuallibrary.com/doi/10.1680/jsmic.17.00017).
- 827 48 Read JS, Jia X, Willard J, Appling AP, Zwart JA, Oliver SK, et al. Process-Guided Deep Learning

- 828 Predictions of Lake Water Temperature. *Water Resources Research*. 2019 11;55(11):9173-90.  
829 Available from: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019WR024922>.
- 830 49 Jia X, Willard J, Karpatne A, Read J, Zwart J, Steinbach M, et al. Physics Guided RNNs  
831 for Modeling Dynamical Systems: A Case Study in Simulating Lake Temperature Profiles.  
832 In: *Proceedings of the 2019 SIAM International Conference on Data Mining*. Philadelphia,  
833 PA: Society for Industrial and Applied Mathematics; 2019. p. 558-66. Available from: <https://epubs.siam.org/doi/10.1137/1.9781611975673.63>.
- 834  
835 50 Plappert M. *keras-rl*; 2016. Available from: <https://github.com/keras-rl/keras-rl>.
- 836 51 Brockman G, Cheung V, Pettersson L, Schneider J, Schulman J, Tang J, et al. *OpenAI Gym*.  
837 *arXiv*; 2016. Available from: <http://arxiv.org/abs/1606.01540>.
- 838 52 Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. *TensorFlow: Large-Scale*  
839 *Machine Learning on Heterogeneous Distributed Systems*. *arXiv preprint arXiv:160304467*.  
840 2016. Available from: <https://arxiv.org/pdf/1603.04467.pdf>.
- 841 53 Biewald L. *Experiment Tracking with Weights and Biases*; 2020. Available from: <https://www.wandb.com/>.
- 842  
843 54 Su J, Adams SC, Beling PA. *Value-Decomposition Multi-Agent Actor-Critics*. *CoRR*.  
844 2020;abs/2007.1. Available from: <https://arxiv.org/abs/2007.12306>.
- 845 55 Baldazo D, Parras J, Zazo S. *Decentralized Multi-Agent Deep Reinforcement Learning in*  
846 *Swarms of Drones for Flood Monitoring*. In: *27th European Signal Processing Conference*  
847 *(EUSIPCO)*; 2019. Available from: [https://www.eurasip.org/Proceedings/Eusipco/eusipco2019/](https://www.eurasip.org/Proceedings/Eusipco/eusipco2019/Proceedings/papers/1570533953.pdf)  
848 [Proceedings/papers/1570533953.pdf](https://www.eurasip.org/Proceedings/Eusipco/eusipco2019/Proceedings/papers/1570533953.pdf).
- 849 56 Parisi S, Pirotta M, Restelli M. *Multi-objective reinforcement learning through continuous*  
850 *pareto manifold approximation*. *Journal of Artificial Intelligence Research*. 2016 10;57:187-227.  
851 Available from: <https://jair.org/index.php/jair/article/view/11026>.
- 852 57 Quinn JD, Reed PM, Giuliani M, Castelletti A. *What Is Controlling Our Control Rules? Opening*  
853 *the Black Box of Multireservoir Operating Policies Using Time-Varying Sensitivity Analysis*.  
854 *Water Resources Research*. 2019 7;55(7):5962-84. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024177>.
- 855  
856 58 Wiering MA, van Hasselt H. *Ensemble algorithms in reinforcement learning*. *IEEE Transactions*  
857 *on Systems, Man, and Cybernetics, Part B: Cybernetics*. 2008 8;38(4):930-6. Available from:

- 858 <https://ieeexplore.ieee.org/document/4509588>.
- 859 59 Wang Y, Jin H. A Boosting-based Deep Neural Networks Algorithm for Reinforcement Learning.  
860 In: 2018 Annual American Control Conference (ACC). IEEE; 2018. p. 1065-71. Available  
861 from: <https://ieeexplore.ieee.org/document/8431647/>.
- 862 60 Shishegar S, Duchesne S, Pelletier G, Ghorbani R. A smart predictive framework for system-  
863 level stormwater management optimization. *Journal of Environmental Management*. 2021  
864 1;278:111505.
- 865 61 Marchese D, Jin A, Fox-Lent C, Linkov I. Resilience for Smart Water Systems. *Journal of*  
866 *Water Resources Planning and Management*. 2020;146(1):2519002. Available from: <https://doi.org/10.1061/http://ascelibrary.org/doi/10.1061/%28ASCE%29WR.1943-5452.0001130>.
- 867 [//doi.org/10.1061/http://ascelibrary.org/doi/10.1061/%28ASCE%29WR.1943-5452.0001130](https://doi.org/10.1061/http://ascelibrary.org/doi/10.1061/%28ASCE%29WR.1943-5452.0001130).
- 868 62 Amodei D, Olah C, Steinhardt J, Christiano P, Schulman J, Dan M. Concrete Problems in AI  
869 Safety. *CoRR*. 2016. Available from: <http://arxiv.org/abs/1606.06565>.
- 870 63 Miller MP, Tesoriero AJ, Capel PD, Pellerin BA, Hyer KE, Burns DA. Quantifying watershed-  
871 scale groundwater loading and in-stream fate of nitrate using high-frequency water quality  
872 data. *Water Resources Research*. 2016 1;52(1):330-47. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/2015WR017753><https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017753><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015WR017753>.
- 873 <https://onlinelibrary.wiley.com/doi/full/10.1002/2015WR017753><https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017753><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015WR017753>.
- 874 <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015WR017753>.
- 875 64 Hensley RT, Cohen MJ, Korhnaek LV. Hydraulic effects on nitrogen removal in a tidal spring-fed  
876 river. *Water Resources Research*. 2015 3;51(3):1443-56. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/2014WR016178><https://onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016178><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014WR016178>.
- 877 <https://onlinelibrary.wiley.com/doi/full/10.1002/2014WR016178><https://onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016178><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014WR016178>.
- 878 <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014WR016178>.
- 879 65 United States Geological Survey. Next Generation Water Observing System (NGWOS);. Avail-  
880 able from: [https://www.usgs.gov/mission-areas/water-resources/science/next-generation-water-observing-system-ngwos?qt-science\\_center\\_objects=0#qt-science\\_center\\_objects](https://www.usgs.gov/mission-areas/water-resources/science/next-generation-water-observing-system-ngwos?qt-science_center_objects=0#qt-science_center_objects).
- 881 [https://www.usgs.gov/mission-areas/water-resources/science/next-generation-water-observing-system-ngwos?qt-science\\_center\\_objects=0#qt-science\\_center\\_objects](https://www.usgs.gov/mission-areas/water-resources/science/next-generation-water-observing-system-ngwos?qt-science_center_objects=0#qt-science_center_objects).
- 882 66 United States Geological Survey. WaterQualityWatch – Continuous Real-Time Water Quality  
883 of Surface Water in the United; . Available from: [https://waterwatch.usgs.gov/wqwatch/faq?faq\\_id=1](https://waterwatch.usgs.gov/wqwatch/faq?faq_id=1).
- 884 [https://waterwatch.usgs.gov/wqwatch/faq?faq\\_id=1](https://waterwatch.usgs.gov/wqwatch/faq?faq_id=1).
- 885 67 Litwin D, Tucker G, Barnhart K, Harman C. GroundwaterDupuitPercolator: A Landlab compo-  
886 nent for groundwater flow. *Journal of Open Source Software*. 2020 2;5(46):1935. Available  
887 from: <https://joss.theoj.org/papers/10.21105/joss.01935>.



888 68 Likmeta A, Metelli AM, Tirinzoni A, Giol R, Restelli M, Romano D. Combining reinforcement  
889 learning with rule-based controllers for transparent and general decision-making in autonomous  
890 driving. *Robotics and Autonomous Systems*. 2020;131:103568.