


Review

Stormwater Runoff Treatment Using Rain Garden: Performance Monitoring and Development of Deep Learning-Based Water Quality Prediction Models

Minsu Jeon ¹, Heidi B. Guerra ¹ , Hyeseon Choi ¹, Donghyun Kwon ², Hayong Kim ³ and Lee-Hyung Kim ^{1,*}

¹ Civil and Environmental Engineering Department, Kongju National University, Cheonan 31080, Korea; minsu91@kongju.ac.kr (M.J.); heidiguerra@kongju.ac.kr (H.B.G.); hyeseon27@kongju.ac.kr (H.C.)

² Urban System Engineering Department, Kongju National University, Cheonan 31080, Korea; seank4891@gmail.com

³ Department of Construction Environment Research, Land & Housing Institute, Daejeon 34047, Korea; civilkhy@lh.or.kr

* Correspondence: leehyung@kongju.ac.kr

Abstract: Twenty-three rainfall events were monitored to determine the characteristics of the stormwater runoff entering a rain garden facility and evaluate its performance in terms of pollutant removal and volume reduction. Data gathered during the five-year monitoring period were utilized to develop a deep learning-based model that can predict the concentrations of Total Suspended Solids (TSS), Chemical Oxygen Demand (COD), Total Nitrogen (TN), and Total Phosphorus (TP). Findings revealed that the rain garden was capable of effectively reducing solids, organics, nutrients, and heavy metals from stormwater runoff during the five-year period when hydrologic and climate conditions have changed. Volume reduction was also high but can decrease over time due to the accumulation of solids in the facility which reduced the infiltration capacity and increased ponding and overflows especially during heavy rainfalls. A preliminary development of a water quality prediction model based on long short-term memory (LSTM) architecture was also developed to be able to potentially reduce the labor and costs associated with on-site monitoring in the future. The LSTM model predicted pollutant concentrations that are close to the actual values with a mean square error of 0.36 during calibration and a less than 10% difference from the measured values during validation. The study showed the potential of using deep learning architecture for the prediction of stormwater quality parameters entering rain gardens. While this study is still in the preliminary stage, it can potentially be improved for use in performance monitoring, decision-making regarding maintenance, and design of similar technologies in the future.



Citation: Jeon, M.; Guerra, H.B.; Choi, H.; Kwon, D.; Kim, H.; Kim, L.-H. Stormwater Runoff Treatment Using Rain Garden: Performance Monitoring and Development of Deep Learning-Based Water Quality Prediction Models. *Water* **2021**, *13*, 3488. <https://doi.org/10.3390/w13243488>

Academic Editor: Vicenç Puig

Received: 31 August 2021

Accepted: 2 December 2021

Published: 8 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: deep learning; long short-term memory; rain garden; urban stormwater runoff

1. Introduction

As urbanization progresses around the world, impervious areas created by the redevelopment of urban areas and the development of new cities are increasing. As a result, the natural water cycle has been disturbed and the corresponding changes have led to urban flooding, increased diffuse pollution, and the impairment of water quality and ecology [1–3]. In order to alleviate this problem, Korea has been adapting the concept of Low Impact Development (LID) which originated in the US and has similar applications, such as Water Sensitive Urban Design (WSUD) in Australia and Sustainable Drainage System (SUD) in the UK. LIDs are small, decentralized, typically vegetated systems that manage stormwater runoff on-site to keep post-development hydrology close to pre-development conditions. These development techniques aim to manage stormwater at the source through natural processes, such as infiltration, storage, reduction of pollutants during rainfall, as well as establish eco-friendly cities. For example, South Korea is developing urban areas

by designating a water circulation city, China is implementing a variety of projects, and advanced countries, such as France, the UK, Australia, Germany, and the United States have been doing research and applying LID techniques for each local government. Starting with eco-friendly LID techniques in 2004, South Korea has selected several cities for water cycle management and developing cost-effective technologies for stormwater runoff management. Among these technologies, the rain garden is a type of bioretention system that incorporates infiltration functions with vegetation and reduces nonpoint source pollutants by maximizing various treatment mechanisms. Rain gardens do not only use physical treatment mechanisms, such as sedimentation, filtration, and adsorption, but also remove pollutants through chemical and biological treatment mechanisms, such as photosynthesis, respiration, and denitrification [4,5]. Typically, stormwater runoff is allowed to infiltrate the surrounding soil in the rain garden to promote groundwater recharge and mitigate urban flooding. Nitrogen, phosphorus, and organic matter are mainly accumulated due to a large amount of particulate matter, nitrogen, and phosphorus contained in the runoff. Nitrogen and phosphorus in the runoff exist in various forms, such as dissolved organic nitrogen, particulate organic nitrogen, particulate phosphorus, dissolved phosphorus, dissolved organic phosphorus and soluble reactive phosphorus which are removed by filter media, plants, and microorganisms [6–8]. However, sediment deposition in the soil media and the system as a whole due to prolonged usage and aging of the facility causes clogging of the filter media pores. This can reduce the facility's pollutant removal efficiency, infiltration capacity, storage capacity, and its ability to help maintain the natural water cycle in the surrounding area. In order to prevent this, maintenance of the facility must be performed. LID facilities are usually built on a small-scales and distributed within the target area, so maintenance is generally performed all at once. As a result, even the facilities that are still in good condition are receiving maintenance which can increase costs. In addition, because the maintenance cycle is not constant, problems, such as low economic feasibility and the consumption of a lot of manpower can occur.

Generally, nonpoint source pollution and LID monitoring, as well as effectiveness evaluation, include on-site standby and preparation in case of rain. After on-site monitoring, the samples were immediately analyzed accordingly in the laboratory. Field monitoring provides excellent accuracy for data analysis and despite many shortcomings in terms of manpower, time, costs, and maintenance, on-site monitoring is being continuously carried out. Meanwhile, due to the recent fourth industrial revolution, handling of extensive data with the use of deep learning, machine learning, and neural network modeling has become the new trend for data analysis, time-series forecasting, and water quality prediction [9–11]. For example, Baek et al. [12] created a combined deep learning approach by using Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) to simulate the water levels and three water quality parameters namely Total Nitrogen (TN), Total Phosphorus (TP), and Total Organic Carbon (TOC) in the Nakdong River Basin, South Korea. The Nash–Sutcliffe Efficiency (NSE) for the CNN model was 0.933 while that of the LSTM model was 0.75 which was considered as acceptable and very good performance. Different deep learning architectures have also been used for weather forecasting, flooding, and disaster prediction [13,14], as well as in conserving water by developing a decision-making tool for irrigation [15].

Moreover, the interest in the concept of smart cities has been increasing around the world [16–19]. In Korea, the establishment of smart water circulation management systems, such as in Sejong Smart City, Busan Eco Delta Smart City, Daegu National Industrial Complex Smart Water System paved the way to effective management and information dissemination. Information related to water and sewage management, flooding, pipe networks, etc., can be easily available to institutions that perform citizen-tailored administrative tasks and projects through real-time measurement, analysis, and prediction. This way, information regarding the status and operation of facilities can be easily available to citizens anytime and anywhere since the data are provided through automatic measuring sensors, database construction, remote control, and real-time monitoring.

However, there are several limitations to these applications. Since LIDs are small and decentralized and a number of facilities are typically applied over a large area, installing water quality and soil sensors as well as ready-made (off the shelf) products can be quite expensive so the economic feasibility and maintenance of the device can be limited. In addition, LIDs typically consist of an inlet and an outlet, and the treated water is sometimes detained for a period of time in the facility. Thus, water quality sensors are prone to corrosion and damage that can result in measurement errors. Despite these limitations, the idea of incorporating AI-based technologies and developing prediction models for improving LID monitoring and performance evaluation can be promising. Therefore, in this study, a rain garden facility was monitored and the long-term monitoring data collected was utilized in developing pollutant concentration prediction models. This aims to improve the application, monitoring, maintenance, and design of LID technologies by using real-time data and analysis.

2. Materials and Methods

2.1. Description of the Facility, Monitoring, and Analysis

The rain garden facility is located inside Kongju National University, Cheonan Campus in Cheonan City, South Korea. It was built in 2014 as shown in Figure 1 and consists of a sedimentation tank for pre-treatment, a filter bed composed of soil, sand, and bottom ash, as well as an overflow channel. The sedimentation tank was provided to allow the settling of larger solids and debris in the stormwater runoff. Perforated drainage pipes were also installed to facilitate infiltration while plants were included to improve pollutant reduction as well the aesthetics of the facility. In addition, an overflow channel was provided on one side of the facility to prevent flooding heavy rainfalls. The rain garden receives runoff from a sidewalk and a bus stop with a parking lot. It has a total area of 476 m² and is 100% impermeable. During the study period, a total of 23 rainfall events were monitored from 2014 to 2018 consisting of rainfall events with 5 mm depth or higher. Water quality samples were collected at the inlet and outlet of the facility at intervals of 0 min, 5 min, 10 min, 15 min, 30 min, 60 min and every succeeding hour until the runoff stops. In addition, water quality analysis was performed immediately after collecting the samples. These are conducted according to the Standard Methods for the Examination of Water and Wastewater [20] and includes particulate matter (Total Suspended Solids, TSS), organic matter (Biochemical Oxygen Demand, BOD), Chemical Oxygen Demand (COD), Total Nitrogen (TN), Total phosphorus (Total Phosphorus, TP) and heavy metals. In addition, sediment samples were also collected at different parts of the facility as shown in Figure 1 to be able to determine the amount of pollutants associated with these accumulated sediments. The description of the rain garden facility and monitoring procedures are summarized in Table 1 while the calculation of even mean concentration (EMC) and pollutant removal efficiency are presented in Equations (1) and (2).

$$EMC(\text{mg/L}) = \frac{\int_0^T C(t) \cdot Q_{TR_u}(t) dt}{\int_0^T Q_{TR_u}(t) dt} \quad (1)$$

$$\text{Removal efficiency}(\%) = \frac{\text{Average influent EMC} - \text{Average of effluent EMC}}{\text{Average influent EMC}} \quad (2)$$

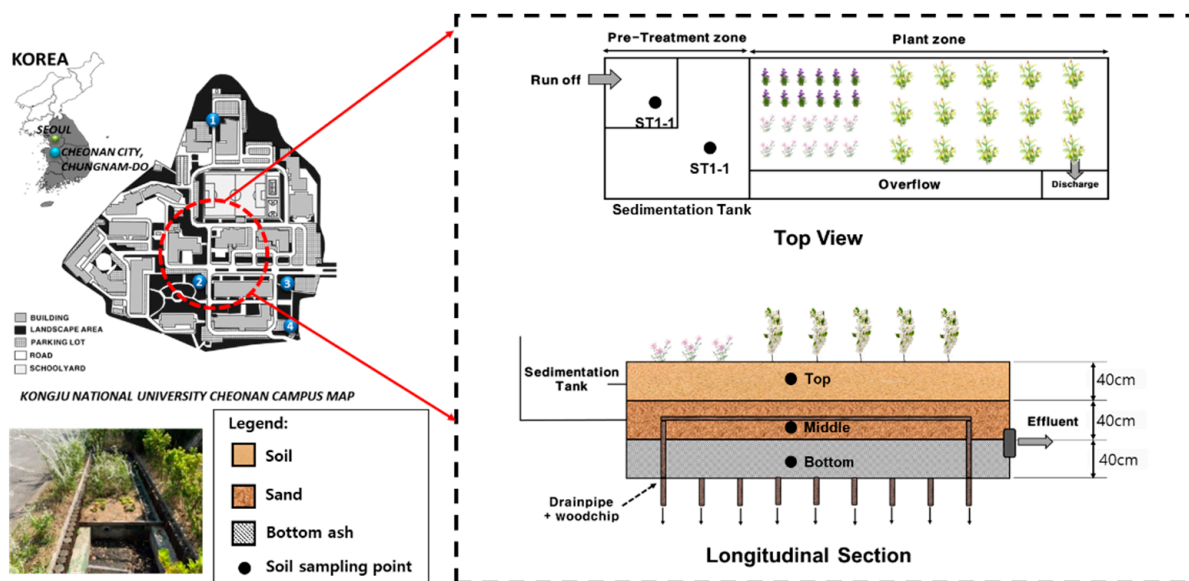


Figure 1. Study site and schematic diagram of the rain garden facility.

Table 1. Characteristics of the rain garden.

	Unit	Characterization/Value
Location	-	KNU-Cheonan campus
Land Use	-	Road and Parking area
Imperviousness rate	%	100
Catchment area	m ²	476
Surface area to catchment area ration(SA/CA ration)	-	1.05
Water Quality	-	Inflow and outflow water samples were collected for water quality analyses on 23 rainfall events monitored from 2014 to 2018.
Water quality sampling time	Min	0, 5, 10, 15, 30, 60, and every succeeding 60 min
Water quality parameters	-	TSS, COD, BOD, TN, TP and Heavy metal
Sediment analysis	-	Particle size, pH, Conductivity, TN, TP, Org-P, Adsorbed-p, NAI-P, Appatite-P, Heavy metal

2.2. Water Quality Prediction Model through LSTM

A recurrent neural network (RNN) model was built and analyzed using PyTorch. In addition, a long short-term memory (LSTM) model, which is s a type of RNN was developed for comparison. Rainfall and atmospheric characteristics namely rainfall depth, antecedent dry days (ADD), air temperature, and fine dust were selected as input parameters as they were found to have a significant impact on pollutant accumulation on paved surfaces during dry days and subsequent wash off during rainy days [21,22]. Rainfall data and ADD were obtained from the weather forecasting website (weather.go.kr (accessed on 23 August 2021)) while air temperature and fine dust data were gathered from Air Korea (airkorea.or.kr/eng (accessed on 23 August 2021)), both created by the Korean Ministry of

Environment, selecting the nearest station from the facility. Nonpoint source pollutants usually appear in particle form [23] which accumulate on urban surfaces and directly affect the concentration of pollutants in the stormwater runoff. The rain garden is considered as a nonpoint source pollution reduction facility and the factors considered influencing this facility include air temperature, solar radiation, temperature, fine dust, vehicle traffic, sidewalk traffic, etc. Since there are many uncertain variables, the items with a high correlation to water quality were selected and analyzed to improve the accuracy of the model. Using the five-year monitoring data, prediction models for TSS, COD, TN, and TP, which are water quality parameters, were developed.

LSTM, a special kind of RNN, was constructed by repeatedly connecting memory blocks (Figures 2 and 3). Each memory cell is composed of one or more repeatedly connected memory cells and three component devices (Input data, Output data, and Forget gate). In this configuration, the Forget gate determines whether the previous data is reflected or not; the Input data puts new data in, and the Output data is reflected and transferred to the next cell [12]. Equations (3)–(8) were used where $C_c^{<t>}$ is the cell state vector, C^{t-1} is the activation function at time step t , x^t is the input at current step t , δ is an element-wise non-linear activation function, τ_i is the input gate, τ_f is the forget gate, τ_o is the output gate, and C_t is a cell state at current step t . The bias and weight matrices are represented as b and W , respectively.

$$C_C^{(t)} = \tanh(W_c[a^t, x^t] + b_i) \tag{3}$$

$$\tau_i = \delta(W_i[C^{t-1}, x^t] + b_i) \tag{4}$$

$$\tau_f = \delta(W_f[C^{t-1}, x^t] + b_f) \tag{5}$$

$$\tau_o = \delta(W_o[C^{t-1}, x^t] + b_o) \tag{6}$$

$$C^t = \tau_i \cdot C_C^{(t)} + \tau_f \cdot C_C^{(t-1)} \tag{7}$$

$$a^t = \tau_o \cdot \tanh C^t \tag{8}$$

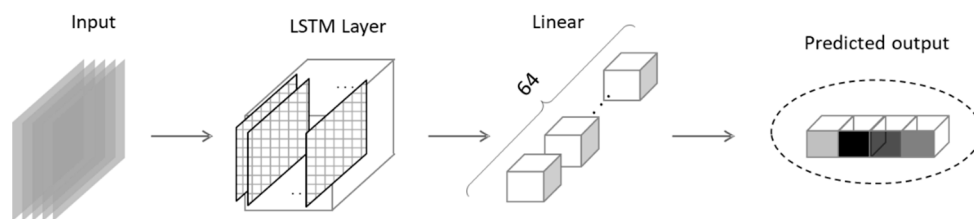


Figure 2. LSTM process.

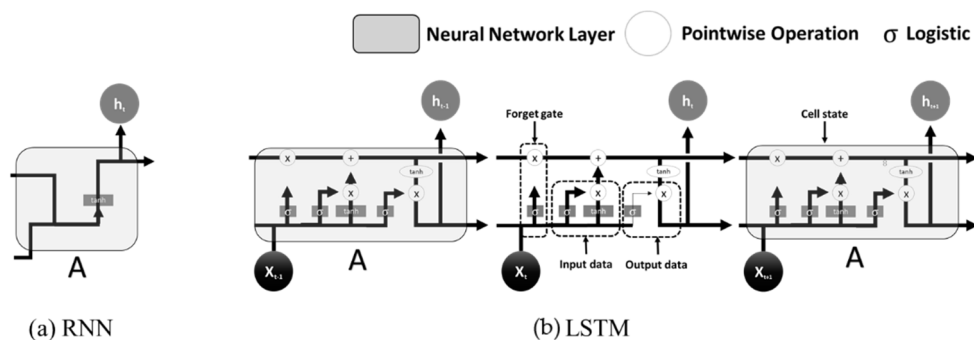


Figure 3. RNN and LSTM architecture.

The data set was divided so that 90% was used for model development and calibration while the remaining 10% was used for validation. In addition, the time series was set at a 30-min interval. The hyperparameters used for the RNN and LSTM models were set as shown in Tables 2 and 3. Both the LSTM and RNN model has a batch size of 30, a time window of 30, and input parameters. Linear was set to 64,4 to be able to analyze TSS, COD, TN, and TP. Mean squared error (MSE) shown in Equation (9) was used as a loss function to evaluate the accuracy of the models [24]. It measures the average of the squared difference between the original and predicted values in the data set.

$$MSE = \frac{1}{N} \sum_{n=1}^N (Y_n - \hat{Y}_n)^2 \quad (9)$$

Table 2. Overview of the LSTM model hyperparameters.

LSTM	Dimensions	Main Parameters	Activation
Input size	(30,30,4)	number of layers = 2	Tanh
LSTM	(2,30,64)	hidden size = 64	
	(2,30,64)	cell state size = 64	
	(30,30,64)	classes = 4	
Linear	(64,4)	Flatten	

Table 3. Overview of the RNN model hyperparameters.

RNN	Dimensions	Main Parameters	Activation
Input size	(30,30,4)	number of layers = 2	Tanh
RNN	(2,30,64)	hidden size = 64	
	(30,30,64)	classes = 4	
Linear	-	Flatten	

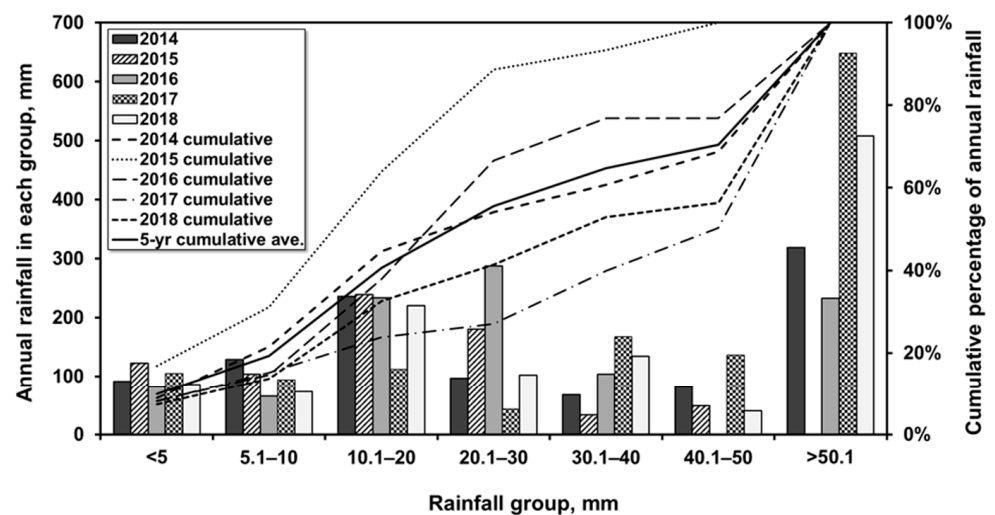
3. Results and Discussion

3.1. Rainfall Characteristics and Air Temperature in the Study Site

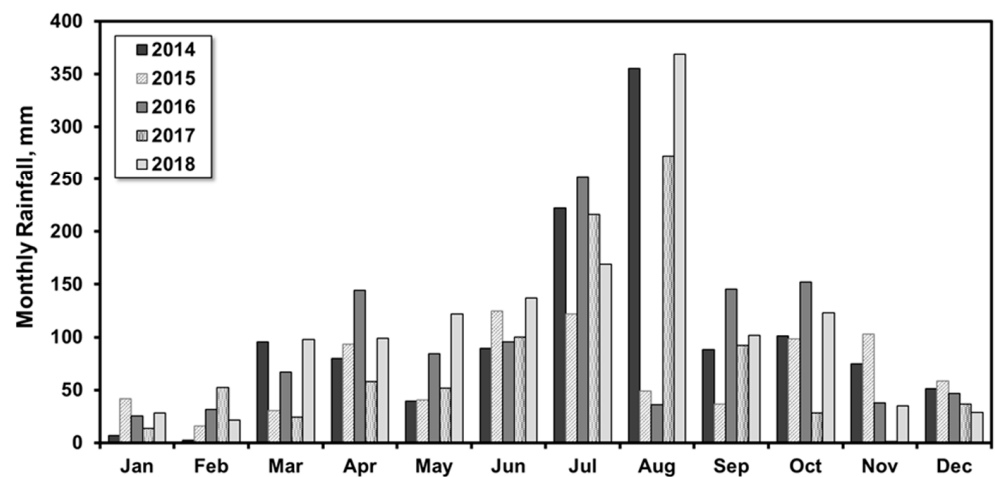
The distribution of rainfall in the rain garden site during the monitoring period is shown in Figure 4 and was divided into different groups in terms of the rainfall depth. The lowest annual rainfall observed was 728 mm in 2015 while the highest was 1305 mm in 2017 with the number of events ranging between 81 and 110 per year. Rainfall that is 10 mm or less constituted 14–31% of the total annual rainfall with the highest cumulative percentage observed during the lowest rainfall year. Meanwhile, the highest annual rainfall observed was due to events that are greater than 50 mm even though these types of events occurred only three to six times a year. Among the rainfall groups, events between 10.1–30 mm constituted the highest percentage of the annual rainfall representing up to 58%. However, in terms of the frequency, the majority of the events are 10 mm or less which occurs up to 87 times per year accounting for 65–77% of the rainfall events per year. Most of these small events occurred during the summer season from June to August averaging 48% of the annual rainfall during the monitored period. Thus, stormwater management technologies in Korea are designed to capture and treat these low but more frequent rainfall events.

In terms of temperature, the average monthly values within a 20-yr period were found to be similar to the temperatures observed in 2020 (Figure 5). However, summer temperatures in 2020 increased by about 2.2 °C as compared to that in 2000 while winter temperatures increased by 1.2 °C. Air temperatures can affect bioretention technologies, such as rain gardens because the microbial biomass of carbon (MBC) and nitrogen (MBN) are important components of ecosystem carbon and nitrogen circulation in the soil and act as sources of carbon and nitrogen for plant absorption [25–27]. However, the variation in temperatures during the summer and winter affects the carbon absorption and respiration

of plants, photosynthesis, and the resistance to low temperatures during the winter [28–31]. Temperature changes promote photosynthesis and activity of nitrate reductase and glutamine synthetase promoting premature aging and death of plants [32]. It also directly affects physiological properties, such as stomatal conductivity and chlorophyll content, and also affects the rate of transpiration in plants, resulting in changes in photosynthesis and changes in growth characteristics [33,34]. In particular, the rise in air temperature has been reported to increase or decrease growth and photosynthesis rates by changing the growth period or photosynthesis activity of seedlings, which may also affect the competitiveness and productivity of individual plants that are in vegetation clusters [35–37]. The physiological activity and growth of trees are also known to be affected by warming temperatures. In addition, the effect of temperature change is different depending on the distribution area of trees. In particular, it is known that forests located in boreal regions respond very sensitively to temperature changes so that the growth and survival of trees are directly affected by climate change [38–42].



(a)



(b)

Figure 4. Distribution of rainfall in the site (a) by rainfall group (b) by month.

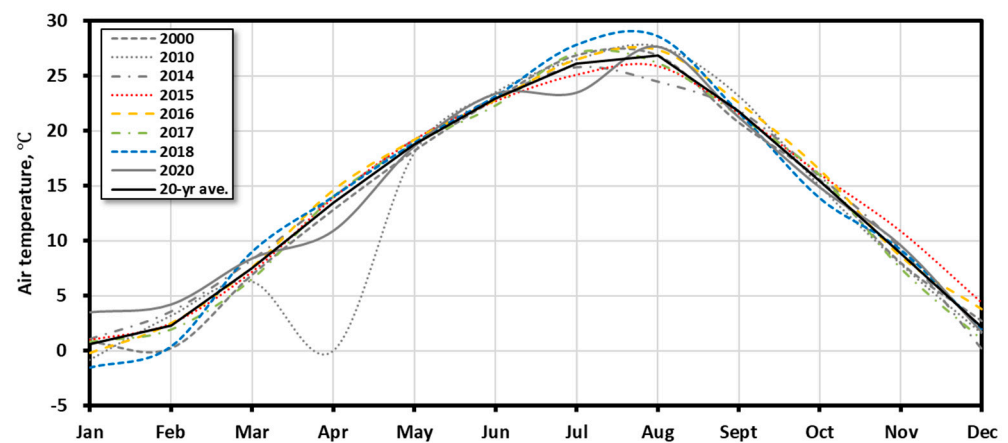


Figure 5. Variation of average monthly air temperature in the study site.

3.2. Water Quality Monitoring during Rainfall

3.2.1. Influent Water Quality and Runoff Volume Reduction

The summary of the monitored rainfall events is presented in Table 4. Rainfall depth ranged between 2.50 and 40.3 mm with a mean value of 18.1 whereas the rainfall intensity was between 0.6 mm/h to 27.2 mm/h with a mean value of 7.6 mm/h. Mean runoff duration was observed to be about 3.6 h longer than the rainfall duration which lasted 0.5–13 h.

Table 4. Characteristics of the monitored events.

Parameter	Units	Min	Max	Mean	Std. Dev
ADD *	days	3.00	20.2	6.20	4.50
Rainfall depth	mm	2.50	40.3	18.1	10.5
Rainfall intensity	mm/h	0.6	27.2	7.60	5.10
Rainfall duration	h	0.52	13.1	5.08	3.81
Total runoff duration	h	0.40	136	8.60	27.9
Inflow peak flow	m ³ /min	0.01	0.45	0.13	0.16

* ADD: antecedent dry days.

Figure 6 shows the characteristics of the pollutants flowing into the rain garden. Average concentrations were 98 mg/L for TSS, 133.6 mg/L for COD, 5.77 mg/L for TN, and 0.54 mg/L for TP. TSS concentration was increased to up to 265 mg/L during rainfall events with longer ADDs and intensities of 10 mm/hr or higher. COD also increased with TSS which is believed to be due to the fall of the surrounding vegetation in the rain garden during autumn. Comparison of the TSS concentrations between the first and last year of monitoring showed that suspended solids in the runoff increased from an average of 52 mg/L to 254 mg/L which was almost five times higher. The frequency of heavier rainfalls during the later years combined with warmer summers and an increase in fine dust in the atmosphere in Korea was believed to be the cause of the increase in solid particles flowing into the facility. As a result, the number of particulate matters accumulating in the facility through the sedimentation tank increased and has affected the infiltration and storage capacity of the rain garden. In the beginning, the runoff volume reduction in the facility was at 98% (Figure 7). This was decreased to 88% at the end of the monitoring period which means a 10% decrease in 5 years. The LID facility first treats the stormwater runoff during rainfall before it is discharged to the storm drains. During this time, nonpoint source pollutants contained in the stormwater runoff are removed using filter media, vegetation, and processes involving microorganisms. The filter media removes pollutants, such as heavy metals, nutrients, and polycyclic aromatic hydrocarbons (PAH). Pollutants also include particles that are 10–300µm in size which are adsorbed into

the micropores of the soil until they are blocked [43–46]. This can cause clogging of the filter media decrease in the infiltration capacity of the facility resulting in ponding and a subsequent decrease in the treatment efficiency. If clogging in the facility progresses, contaminants in the runoff will not be removed and can adversely affect the quality and ecosystem of the receiving waters [47].

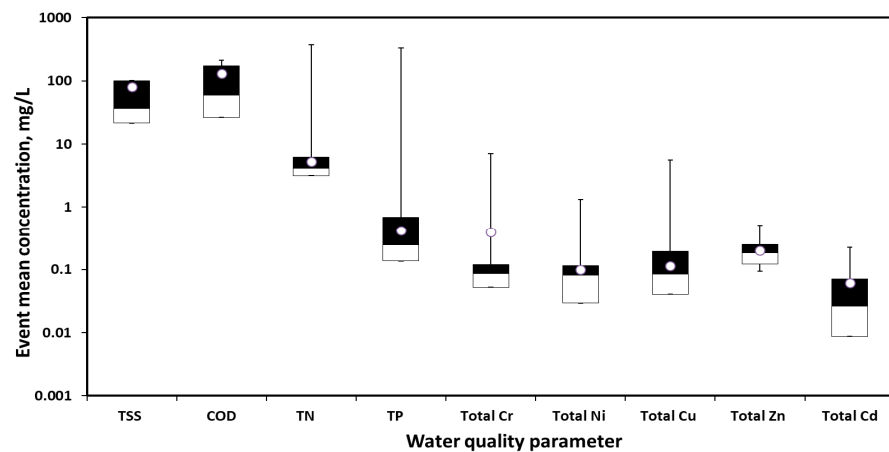


Figure 6. Event mean concentration of the pollutants in the influent.

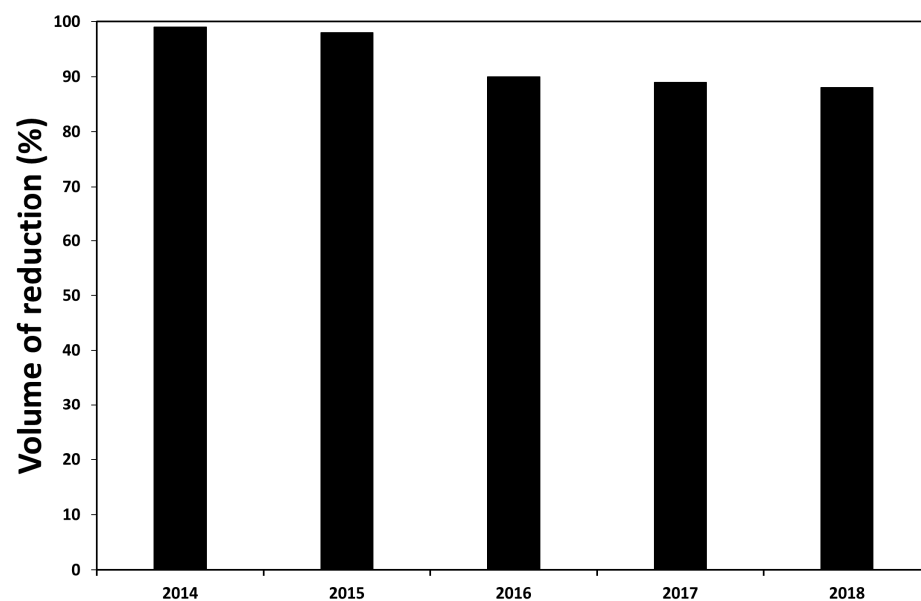


Figure 7. Percent reduction of runoff volume in the rain garden during the monitoring period.

3.2.2. Pollutant Removal Efficiency

The annual average removal efficiency for different diffuse pollutants calculated from the EMCs obtained as described in Section 2.1 is presented in Figure 8. Over the course of the study, the rain garden was able to effectively reduce solids, organics, and nutrients with 85% removal of TSS, 91% for COD, 74% for TN, and 74% for TP. At the start of the operation, the removal of TSS was initially relatively low at 78% due to the washout of particles from the media and the facility itself and not from the inflow. COD and TP are typically found to be well-correlated with TSS ($r = 0.62$ – 0.68) Chow et al. [48] and also had relatively lower removal efficiencies in 2014 and 2015 as compared to the succeeding years. In the case of TN, the removal efficiency in 2015 decreased by about 30% compared to 2014. Then, from 2017 to 2018, it increased by about 45%. The reduction in performance was mainly due to the plants not being able to survive the temperature changes in the summer of 2015. Vegetation removes pollutants by collecting and accumulating not only nitrogen but also inorganic

substances, such as chloride, potassium, calcium, sulfate, and phosphorus [49–53]. In addition, the decrease in infiltration rate and volume reduction within the facility have caused ponding during rainfall. Prolonged ponding affected nitrogen removal by limiting the oxygen transfer on the surface of the facility denitrification mechanisms involving organic materials [54,55].

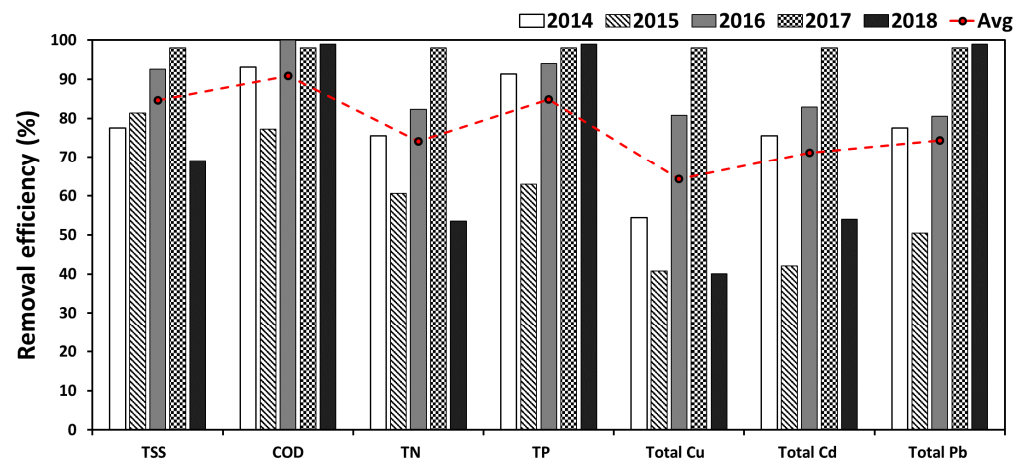


Figure 8. Average annual pollutant removal efficiency.

In the case of heavy metals, the concentrations were high during the first year due to the frequent operation of large buses in the parking lot near the rain garden facility. The removal efficiency of heavy metals did not show a clear annual trend. In the case of Total Pb, the removal efficiency of heavy metals was found to be similar to that of TSS except in 2015, and both Total Cu and Total Cd were analyzed to be as high as 98% in 2017.

Pollutant removal in different rainfall groups was also analyzed as shown in Figure 9. The highest pollutant removal was observed during rainfalls less than 10 mm which was mainly due to the 100% capture of runoff at rainfalls less than 5 mm. In those cases, the pollutant removal efficiency was assumed to be 100% since all the runoff was treated by the facility. For TSS, the efficiency decreased to about 78% during heavier rainfalls. For all the pollutants, the removal efficiency was found to decrease as the rainfall depth increased. In general, the higher the rainfall depth and intensity, the shorter the amount of time needed to reach peak flow within the facility which resulted in flooding and overflow. When overflow occurs, the treatment efficiency decreases since the contaminants were not able to pass through the facility and are discharged untreated. Moreover, since the removal efficiency is calculated based on the collected samples from the inlet and outlet, determining an accurate removal efficiency during overflow and flooding can be tricky. During these instances, it is believed that the actual removal efficiencies are higher than the calculated values based on the monitoring. These can be assumed for rainfalls that are 11 mm or higher. Since performance is determined through water analysis that is carried out in a laboratory, relatively accurate data can be obtained. However, it is challenging to consider all the aspects of the facility in the calculated removal efficiency because the actual conditions in the site during a specific rainfall event may not be reflected in the samples. Therefore, it is necessary to obtain real-time data from the facility in addition to monitoring regular rainfalls (≤ 10 mm) to determine the facility's internal diagnosis and function.

3.3. Analysis of the Sediments in the Facility

The particle size analysis of the sediments was collected from a total of six locations, namely the inlet (CA), sedimentation tank (ST1-1, ST1-2), as well as in the top, middle, and bottom of the filter media shown in Figure 10. The sediment in the inlet was categorized as sandy clay loam and contains a high amount of heavy metals with 0.36 mg/kg Total Cr, 5.17 mg/kg Total Cu, and 6.04 mg/kg Total Pb. Fine particles of less than 10 μm appeared in the order Bottom > CA > Middle > Top > ST1-1 > ST1-2. In the sedimentation

tank, the average size of the sediment particles ranged at 49–113 μm and that 60% of the sediment generated at the inlet was removed and accumulated in this part of the facility. Particles bigger than 50 μm were filtered, adsorbed, and deposited and more than 40% were removed. Particles smaller than 50 μm were removed from the middle and bottom part of the facility due to soil and vegetation.

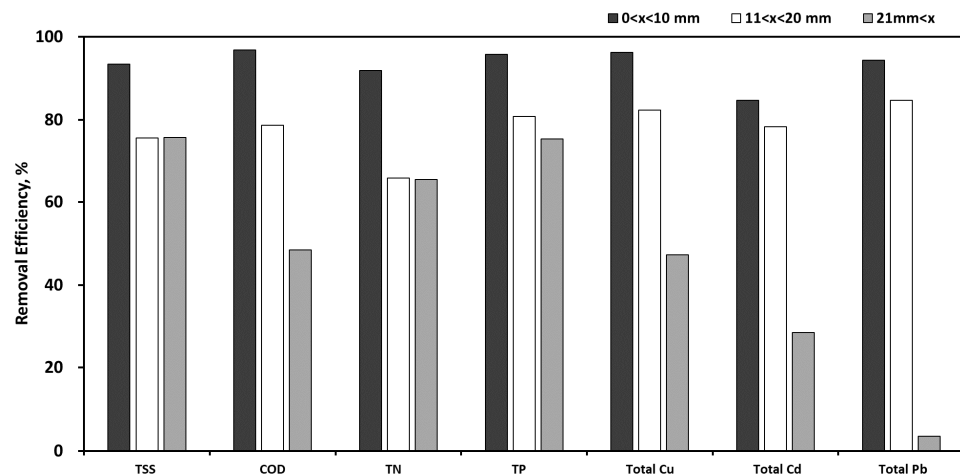


Figure 9. Analysis of the pollutant efficiency by rainfall depth.

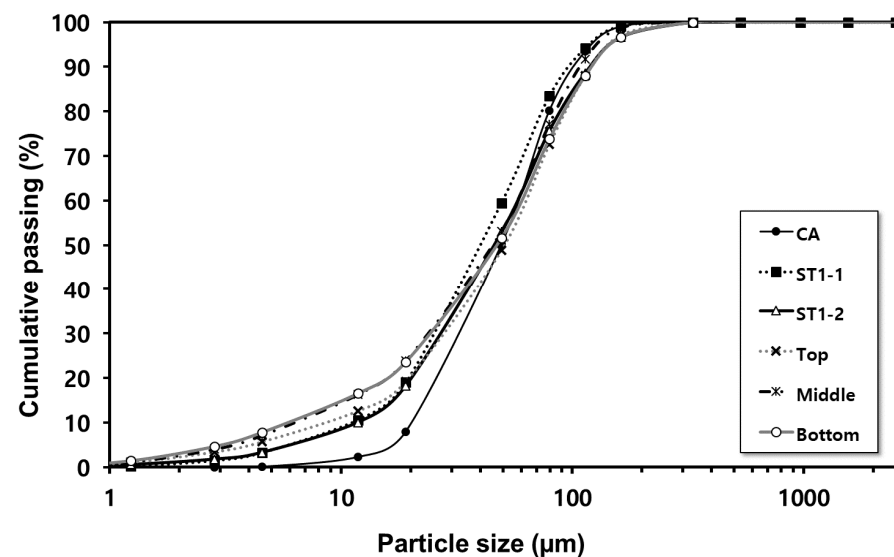


Figure 10. Sediment particle size analysis.

Results of the analysis of pollutants found in the accumulated sediments are shown in Figure 11. The average concentration of TN in the inlet was 829 mg/kg which is 2.6 times higher than the average concentration in the facility (309 mg/kg). Meanwhile, the average concentration of TP was 757 mg/kg, 0.27 times lower than the average value of 1028 mg/kg. The concentration of nitrogen is considered to be high due to the decomposition of the leaves from the *Spirea prunifolia*, a shrub inside and outside the facility, and the hot weather during the summer [56–58]. In the case of phosphorus, NAI-P was composed of 22%, Appatite-P 1%, and Adsorbed-P 10%. Most of the phosphorus introduced into the rain garden is NAI-P, which forms a complex, and Adsorbed-P, which is phosphorus adsorbed to fine particles and is dissolved depending on the changes in the soil pH. When the pH in the soil becomes higher than 4, phosphorus dissolves in the soil and stabilization proceeds from the pH of 6 or higher [59–62].

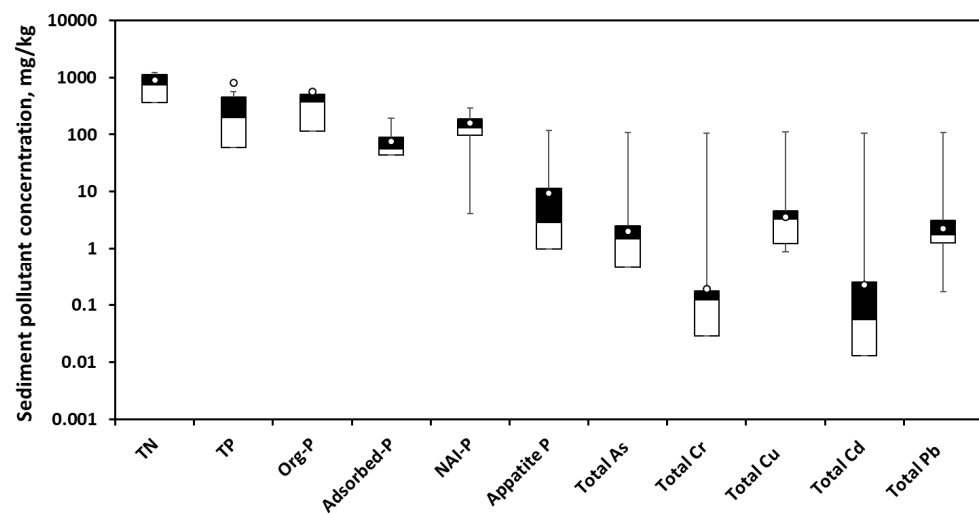


Figure 11. Analysis of the contaminants in the sediments that accumulated in the rain garden.

3.4. LSTM Development Using the Five-Year Monitoring Data

Using meteorological and water quality monitoring data, the concentration of pollutants in the stormwater runoff during rainfall was predicted. The data set consisted of four weather data (rainfall depth, ADD, air temperature, and fine dust) and four water quality data (TSS, COD, TN, and TP). Atmospheric influences, such as the concentration of fine dust in the air affect the accumulation of particulate pollutants during the dry season. Fine dust accumulated during the dry season is washed off in the runoff during rainfall, increasing the concentrations of COD and TN. order to fit the data into the deep learning structure, the collected data set is interpolated in time series, and through this process, missing values were provided and the learning process in the model became smooth.

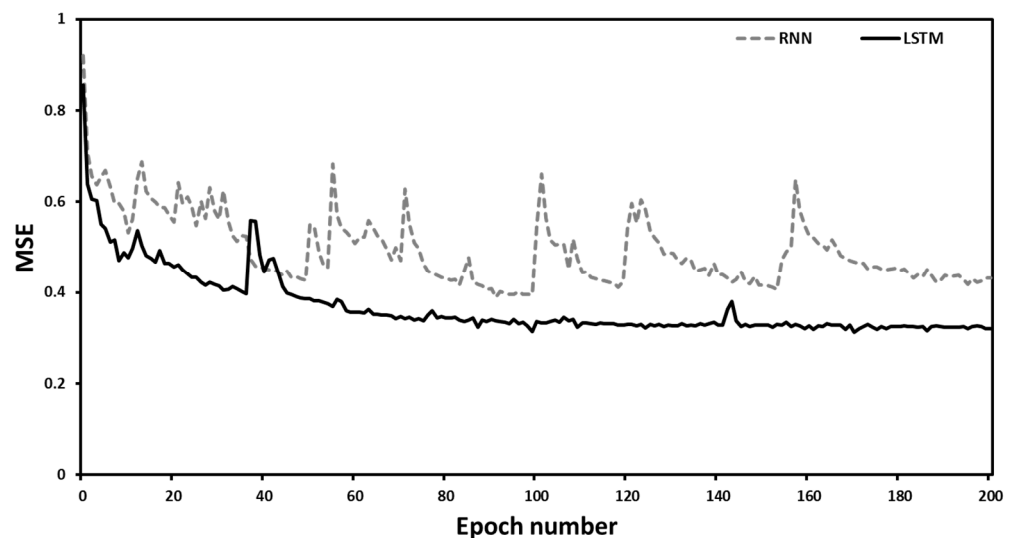
Rainfall monitoring data which includes measured flow rates and water quality from the collected samples from the beginning to the end of the rainfall-runoff were analyzed and the data sets were listed in time series. Prediction of pollutant concentration was done at a 30-min interval for accuracy. However, since the actual monitoring data was at 0, 5, 10, 15, 30, 60 min and every hour thereafter as previously discussed, values halfway within each hour were interpolated using the existing hourly data. The interpolated values constituted 14% of the total data set.

Table 5 shows the details of the data set that was used for developing the LSTM model. The data set was divided so that 90% was used for model development and calibration while the remaining 10% was used for validation. Experimental data were arranged in time series and the timestamp of each instance in the contamination (CT) dataset allows us to order experiments chronologically. The data sets contain sequential information where adjacent timestamps demonstrate a higher similarity. Then, it lists experiments in a time series through adjacent timestamps and uses data with high similarity to identify patterns that repeat at a specific frequency. However, nonpoint pollutants are generated during rainfall and were analyzed to have lower accuracy than data performed in the laboratory at uncertain intervals. Likewise, CT data usually has a certain periodic pattern, which repeats with a certain frequency. However, the experiment result showed high dynamic variance the experiments have taken manually with uncertain intervals.

Table 5. Description of the contamination dataset.

Parameter	Contamination Data
Collected intervals	0, 5, 10, 15, 30, 60, 120 min
Total Experiments	18
Start	2014.05.11
End	2018.06.26
Data set	3692
Data shape	30 Timesteps (3692, 30, 4)
Training set	90% of data set
Test set	10% of data set

Figure 12 shows that the comparison of the MSE values during model training were found to be 0.49 for the RNN model and 0.36 for the LSTM model. LSTM showed consistently lower error values than the RNN model, thus, the LSTM model was ultimately used as the water quality prediction model. The actual and predicted values of the pollutant concentrations are shown in Figure 13 representing two rainfall events and showing that the predicted values are close to the actual values aside from several peak concentrations. For example, the average inflow TSS concentration obtained from the model was 852 mg/L with only a 6% difference from the actual TSS average concentration. Meanwhile, the average COD, TN, and TP concentrations from the model were 97.7 mg/L, 2.62 mg/L, and 0.27 mg/L with 9%, 1%, and 8% difference from the actual values.

**Figure 12.** Result of RNN and LSTM model training.

The good performance of the LSTM model can be attributed to the high concentration of particulate matter that flows into the facility at the beginning of rainfall. In the case of TSS, COD, and TP, which are highly correlated with the concentration of particulate matter, the difference between the pollutant concentration in the initial 0–30 min and the pollutant concentration after 30 min decreases by about 60–80% on average. In addition, nonpoint source pollutants have many variables, such as air temperature during dry season, microdust, solar radiation, plant types, vehicle traffic, human traffic frequency, and land use types. However, monitoring from the past has only observed atmospheric data, and continuous time series data of water quality cannot be obtained in the case of LID facilities in urban areas, and it has non-linearity and data instability. However, it is believed that various predictive models will be developed when data, such as soil humidity, soil temperature, and atmospheric measurement in the area are secured in the LID facility. In addition, it is considered that various studies using various AI technologies are necessary.

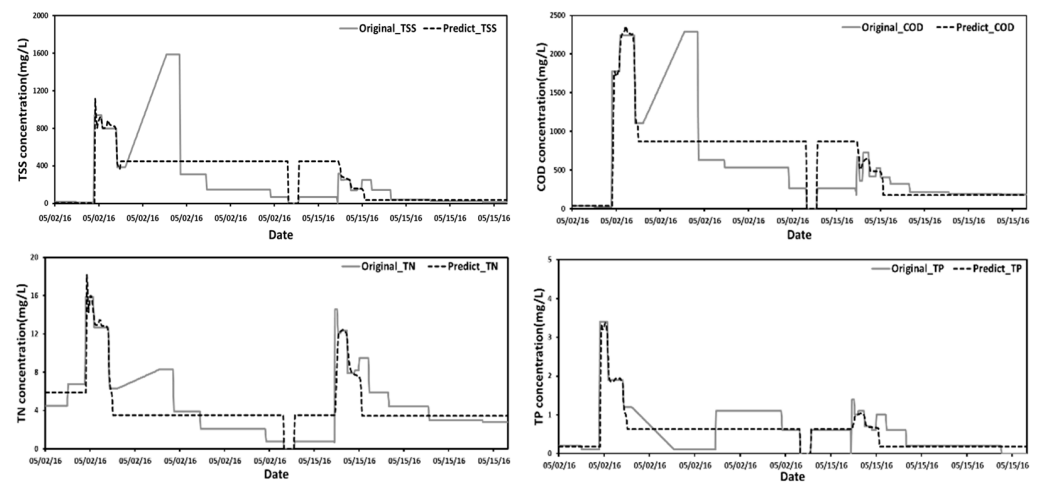


Figure 13. Actual and predicted values of water quality parameters.

4. Conclusions

In this study, 5-year monitoring of a rain garden facility was conducted to be able to evaluate its five-year performance. Moreover, a simple AI-based prediction model for pollutant concentration were developed using the data gathered during the monitoring period. The results revealed the following conclusions:

- (1) The majority of the annual rainfall occurred during the summer and was mostly 10 mm or less, with rainfall of 10.1–30 mm constituting a large percentage of the annual rainfall. This indicates that rain garden facilities in Korea can be designed based on these events. In addition, temperature variation throughout each year and within the monitoring period was observed to have affected plant growth which suggests selecting plants that are more resistant to temperature changes for optimal performance less maintenance.
- (2) During the summer season, the occurrence frequency of rainfall and the high-temperature and short dry periods in between increased the concentration of particles and in the stormwater runoff resulting in an increase in suspended solids in the influent of the facility. During the autumn season, the increase in plant debris in the rain garden, as well as in the surrounding area resulted in an increase in COD and TN concentrations in the inflow.
- (3) The rain garden was able to effectively remove suspended solids, organics, and nutrients from the runoff within the 5-year period. Relatively lower pollutant removal efficiencies were observed during the initial period of monitoring especially for TSS, COD, and TP due to the washout of particles from within the facility. Dead plant materials contributed to the decrease in the removal efficiency of nitrogen especially during periods where temperature variation is prominent. Moreover, it is believed that actual pollutants during heavy rainfalls could be higher than the calculated values based on inflow and outflow samples due to the occurrence of overflow and untreated runoff.
- (4) The sedimentation tank after the inlet was able to remove 60% of the sediments with a particle size of 49–113 μm . Solids smaller than 50 μm were then removed in the middle and bottom part of the rain garden facility.
- (5) The results of the training for RNN and LSTM produced MSE values of 0.49 and 0.36, respectively indicating that LSTM is a more appropriate model for water quality prediction. It also shows that longer dry seasons with high concentrations of fine dust leads to higher initial pollutant concentration during rainfall.

The study showed the potential of using deep learning architecture for the prediction of stormwater quality parameters entering rain gardens or other low impact development facilities using selected atmospheric data. By obtaining these data without manual sam-

pling, on-site measurements, and subsequent laboratory analysis, can reduce the time, labor, and money that is typically necessary for performance monitoring and maintenance while possibly providing a larger set of data that can be used for further analyses and studies. While this study is still in the preliminary stage, it can potentially be improved to achieve a more accurate prediction of pollutant concentrations for use in performance monitoring, decision-making regarding maintenance, and the design of similar technologies in the future.

Author Contributions: M.J.: conceptualization, investigation, formal analysis, software, writing—original draft preparation, writing—review & editing; H.B.G.: investigation, data curation, formal analysis, writing—original draft preparation, writing—review & editing; H.C.: conceptualization, investigation, data curation, formal analysis, supervision; D.K.: investigation, formal analysis, software; H.K.: writing—review & editing, funding acquisition, supervision, resources; L.-H.K.: project administration, methodology, funding acquisition, supervision, validation, conceptualization, resources. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Korea Environment Industry & Technology Institute (KEITI) through Intelligent Management Program for Urban Water Resources Project, funded by Korea Ministry of Environment (MOE) (2019002950003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jacobson, C.R. Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review. *J. Environ. Manag.* **2011**, *92*, 1438–1448. [[CrossRef](#)]
- Paule-Mercado, M.A.; Lee, B.Y.; Memon, S.A.; Umer, S.R.; Salim, I.; Lee, C.H. Influence of land development on stormwater runoff from a mixed land use and land cover catchment. *Sci. Total Environ.* **2017**, *599*, 2142–2155. [[CrossRef](#)] [[PubMed](#)]
- Zhou, Q. A review of sustainable urban drainage systems considering the climate change and urbanization impacts. *Water* **2014**, *6*, 976–992. [[CrossRef](#)]
- Davis, A.P. Green engineering principles promote low-impact development. *Sci. Total Environ.* **2005**, *39*, 338–344. [[CrossRef](#)]
- Gurung, S.B.; Geronimo, F.K.; Hong, J.S.; Kim, L.H. Application of indices to evaluate LID facilities for sediment and heavy metal removal. *Chemosphere* **2018**, *206*, 693–700. [[CrossRef](#)] [[PubMed](#)]
- Taylor, G.D.; Fletcher, T.D.; Wong, T.H.F.; Breen, P.F.; Duncan, H.P. Nitrogen composition in urban runoff—Implications for stormwater management. *Water Res.* **2005**, *39*, 1982–1989. [[CrossRef](#)]
- Flint, K.R.; Davis, A.P. Pollutant mass flushing characterization of highway stormwater runoff from an ultra-urban area. *J. Environ. Eng.* **2007**, *133*, 616–626. [[CrossRef](#)]
- Li, H.; Davis, A.P. Urban particle capture in bioretention media. I: Laboratory and field studies. *J. Environ. Eng.* **2008**, *134*, 409–418. [[CrossRef](#)]
- Torres, J.F.; Hadjout, D.; Sebaa, A.; Martínez-Álvarez, F.; Tronsoco, A. Deep Learning for Time Series Forecasting: A survey. *Big Data* **2021**, *9*, 3–21. [[CrossRef](#)]
- Chen, Y.; Song, L.; Liu, Y.; Yang, L.; Li, D. A review of the artificial neural network models for water quality prediction. *Appl. Sci.* **2020**, *10*, 5776. [[CrossRef](#)]
- Ahmed, A.N.; Othman, F.B.; Afan, H.A.; Ibrahim, R.K.; Fai, C.M.; Hossain, M.S.; Ehteram, M.; Elshafie, A. Machine learning methods for better water quality prediction. *J. Hydrol.* **2019**, *578*, 124084. [[CrossRef](#)]
- Baek, S.-S.; Pyo, J.; Chun, J.A. Prediction of Water Level and Water Quality Using a CNN-LSTM Combined Deep Learning Approach. *Water* **2020**, *12*, 3399. [[CrossRef](#)]
- Satwik, P.M.; Sundram, M. An integrated approach for weather forecasting and disaster prediction using deep learning architecture based on memory Augmented Neural Network's (MANN's). *Mater. Today Proc.* **2021**, in press. [[CrossRef](#)]
- Raj, J.R.; Charless, I.; Latheef, M.A.; Srinivasulu, S. Identifying the Flooded Area Using Deep Learning Model. In Proceedings of the 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 28–30 April 2021; pp. 582–586.
- Chen, M.; Cui, Y.; Wang, X.; Xie, H.; Liu, F.; Luo, T.; Zheng, S.; Luo, Y. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. *Agric. Water Manag.* **2021**, *250*, 106838. [[CrossRef](#)]
- Hollands, R.G. Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City* **2008**, *12*, 303–320. [[CrossRef](#)]

17. Cretu, L.G. Smart cities design using event-driven paradigm and semantic web. *Inform. Econ.* **2012**, *16*, 57.
18. Nguyen, H.Q.; Radhakrishnan, M.; Bui, T.K.N.; Tran, D.D.; Ho, L.P.; Tong, V.T.; Huynh, L.T.P.; Chau, N.X.Q.; Ngo, T.T.T.; Pathirana, A.; et al. Evaluation of retrofitting responses to urban flood risk in Ho Chi Minh City using the motivation and ability (MOTA) framework. *Sustain. Cities Soc.* **2019**, *47*, 101465. [[CrossRef](#)]
19. Yigitcanlar, T.; Han, H.; Kamruzzaman, M.; Ioppolo, G.; Sabatini-Marques, J. The making of smart cities: Are Songdo, Masdar, Amsterdam, San Francisco and Brisbane the best we could build? *Land Use Policy* **2019**, *88*, 104187. [[CrossRef](#)]
20. APHA; AWWA; WEF. *Standard Methods for the Examination of Water and Wastewater*, 19th ed.; American Public Health Association/American Water Works Association/Water Environment Federation: Washington, DC, USA, 1995.
21. Yuan, Q.; Guerra, H.B.; Kim, Y. An Investigation of the Relationships between Rainfall Conditions and Pollutant Wash-Off from the Paved Road. *Water* **2017**, *9*, 232. [[CrossRef](#)]
22. Lee, J.Y.; Kim, H.J.; Kim, Y.J.; Han, M.Y. Characteristics of the event mean concentration (EMC) from rainfall runoff on an urban highway. *Environ. Pollut.* **2011**, *159*, 884–888. [[CrossRef](#)]
23. Shaw, S.B.; Walter, M.T.; Steenhuis, T.S. A physical model of particulate wash-off from rough impervious surfaces. *J. Hydrol.* **2006**, *327*, 618–626. [[CrossRef](#)]
24. Heinermann, J.; Kramer, O. Machine learning ensembles for wind power prediction. *Renew. Energy* **2016**, *89*, 671–679. [[CrossRef](#)]
25. Bai, E.; Li, S.; Xu, W.; Li, W.; Dai, W.; Jiang, P. A meta-analysis of experimental warming effects on terrestrial nitrogen pools and dynamics. *New Phytol.* **2013**, *199*, 441–451. [[CrossRef](#)]
26. Lu, M.; Zhou, X.; Yang, Q.; Li, H.; Luo, Y.; Fang, C.; Chen, J.; Yang, X.; Li, B. Responses of ecosystem carbon cycle to experimental warming: A meta-analysis. *Ecology* **2013**, *94*, 726–738. [[CrossRef](#)]
27. Zhang, X.Z.; Shen, Z.X.; Fu, G. A meta-analysis of the effects of experimental warming on soil carbon and nitrogen dynamics on the Tibetan Plateau. *Appl. Soil Ecol.* **2015**, *87*, 32–38. [[CrossRef](#)]
28. Adams, W.W.; Demmig-Adams, B.; Rosenstiel, T.N.; Brightwell, A.K.; Ebbert, V. Photosynthesis and photoprotection in overwintering plants. *Plant Biol.* **2002**, *4*, 545–557. [[CrossRef](#)]
29. de Faÿ, E.; Vacher, V.; Humbert, F. Water-related phenomena in winter buds and twigs of *Picea abies* L. (Karst.) until bud-burst: A biological, histological and NMR study. *Ann. Bot.* **2000**, *86*, 1097–1107. [[CrossRef](#)]
30. Jouve, L.; Jacques, D.; Douglas, G.C.; Hoffmann, L.; Hausman, J.F. Biochemical characterization of early and late bud flushing in common ash (*Fraxinus excelsior* L.). *Plant Sci.* **2007**, *172*, 962–969. [[CrossRef](#)]
31. Bokhorst, S.; Bjerke, J.W.; Davey, M.P.; Taulavuori, K.; Taulavuori, E.; Laine, K.; Callaghan, T.V.; Phoenix, G.K. Impacts of extreme winter warming events on plant physiology in a sub-Arctic heath community. *Physiol. Plant.* **2010**, *140*, 128–140. [[CrossRef](#)]
32. Argente Martínez, L.; Garatuza Payán, J.; Arredondo Moreno, J.T.; Yezpez González, E.A. Effects of experimental warming on peroxidase, nitrate reductase and glutamine synthetase activities in wheat. *Agron. Res.* **2019**, *17*, 22.
33. Saxe, H.; Cannell, M.G.R.; Johnsen, Ø.; Ryan, M.G.; Vourlitis, G. Tree and forest functioning in response to global warming. *New Phytol.* **2001**, *149*, 369–399. [[CrossRef](#)]
34. Lewis, J.D.; Lucash, M.; Olszyk, D.; Tingey, D.T. Seasonal patterns of photosynthesis in Douglas fir seedlings during the third and fourth year of exposure to elevated CO₂ and temperature. *Plant. Cell Environ.* **2001**, *24*, 539–548. [[CrossRef](#)]
35. Arend, M.; Kuster, T.; Gunthardt-Goerg, M.S.; Dobbertin, M. Provenance-specific growth responses to drought and air. *Tree Physiol.* **2011**, *31*, 287–297. [[CrossRef](#)]
36. Xu, Z.; Hu, T.; Zhang, Y. Effects of experimental warming on phenology, growth and gas exchange of treeline birch. *Eur. J. For. Res.* **2012**, *131*, 811–819. [[CrossRef](#)]
37. Wu, Z.; Dijkstra, P.; Koch, G.W.; Peñuelas, J.; Hungate, B.A. Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. *Glob. Chang. Biol.* **2011**, *17*, 927–942. [[CrossRef](#)]
38. An, J.; Han, S.H.; Chang, H.; Park, M.J.; Son, Y. Effects of Experimental Warming on Physiological Responses of Four Deciduous Tree Species Seedlings. *J. Korean Soc. For. Sci.* **2016**, *105*, 510–516. [[CrossRef](#)]
39. Chung, H.; Muraoka, H.; Nakamura, M.; Han, S.; Muller, O.; Son, Y. Experimental warming studies on trees species and forest ecosystems: A literature review. *J. Plant Res.* **2013**, *126*, 447–460. [[CrossRef](#)]
40. Fisichelli, N.A.; Frelich, L.E.; Reich, P.B. Sapling growth responses to warmer temperatures ‘cooled’ by browse pressure. *Glob. Chang. Biol.* **2012**, *18*, 3455–3463. [[CrossRef](#)]
41. Matala, J.; Ojansuu, R.; Peltola, H.; Raitio, H.; Kellomäki, S. Modelling the response of tree growth to temperature and CO₂ elevation as related to the fertility and current temperature sum of a site. *Ecol. Model.* **2006**, *199*, 39–52. [[CrossRef](#)]
42. Scholze, M.; Knorr, W.; Arnell, N.W.; Prentice, I.C. A climate-change risk analysis for world ecosystems. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13116–13120. [[CrossRef](#)]
43. El-Mufleh, A.; Béchet, B.; Ruban, V.; Legret, M.; Clozel, B.; Barraud, S.; Gonzalez-Merchan, C.; Bedell, J.P.; Delolme, C. Review on physical and chemical characterizations of contaminated sediments. *Environ. Sci. Pollut. Res.* **2014**, *21*, 5329–5346. [[CrossRef](#)] [[PubMed](#)]
44. Badin, A.L.; Méderel, G.; Béchet, B.; Borschneck, D.; Delolme, C. Study of the aggregation of the surface layer of Technosols from stormwater infiltration basins using grain size analyses with laser diffractometry. *Geoderma* **2009**, *153*, 163–171. [[CrossRef](#)]
45. Clozel, B.; Ruban, V.; Durand, C.; Conil, P. Origin and mobility of heavy metals in. *Appl. Geochem.* **2006**, *21*, 1781–1798. [[CrossRef](#)]
46. Schuh, W.N. Seasonal variation of clogging of an artificial recharge basin in a northern climate. *J. Hydrol.* **1990**, *121*, 193–215. [[CrossRef](#)]

47. Cannavo, P.; Coulon, A.; Charpentier, S.; Béchet, B.; Vidal-Beaudet, L. Water balance prediction in stormwater infiltration basins using 2-D modeling: An application to evaluate the clogging process. *Int. J. Sediment Res.* **2018**, *33*, 371–384.
48. Chow, M.F.; Yusop, Z. Characterization and source identification of stormwater runoff in tropical urban catchments. *Water Sci. Technol.* **2014**, *69*, 244–252. [[CrossRef](#)]
49. Bastviken, S.K.; Weisner, S.E.B.; Thiere, G.; Svensson, J.M.; Ehde, P.M.; Tonderski, K.S. Effects of vegetation and hydraulic load on seasonal nitrate removal in treatment wetlands. *Ecol. Eng.* **2009**, *35*, 946–952. [[CrossRef](#)]
50. Anjana Umar, S.; Iqbal, M.; Abrol, Y.P. Are nitrate concentrations in leafy vegetables within safe limits? *Curr. Sci.* **2007**, *92*, 355–360.
51. Renseigné, N.; Umar, S.; Iqbal, M. Nitrate accumulation in plants, factors affecting the process, and human health implications. *Agron. Sustain. Dev.* **2007**, *27*, 45–57.
52. Santamaria, P.; Elia, A.; Serio, F.; Todaro, E. A survey of nitrate and oxalate content in fresh vegetables. *J. Sci. Food Agric.* **1999**, *79*, 1882–1888. [[CrossRef](#)]
53. Cantliffe, D.J. Nitrate accumulation in table beets and spinach as affected by nitrogen, phosphorus, and potassium nutrition and light intensity. *Agron. J.* **1973**, *65*, 563–565. [[CrossRef](#)]
54. Spieles, D.J.; Mitsch, W.J. The effects of season and hydrologic and chemical loading on nitrate retention in constructed wetlands: A comparison of low- and high-nutrient riverine systems. *Ecol. Eng.* **1999**, *14*, 77–91. [[CrossRef](#)]
55. Raisin, G.W.; Mitchell, D.S. The use of wetlands for the control of non-point source pollution. *Water Sci. Technol.* **1995**, *32*, 177–186. [[CrossRef](#)]
56. Compton, J.E.; Boone, R.D. Long-term impacts of agriculture on soil carbon and nitrogen in New England forests. *Ecology* **2000**, *81*, 2314–2330. [[CrossRef](#)]
57. Wang, Y.; Zhang, X.; Huang, C. Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the Loess Plateau, China. *Geoderma* **2009**, *150*, 141–149. [[CrossRef](#)]
58. Wang, H.; Hall, C.A.; Cornell, J.D.; Hall, M.H. Spatial dependence and the relationship of soil organic carbon and soil moisture in the Luquillo Experimental Forest, Puerto Rico. *Landsc. Ecol.* **2002**, *17*, 671–684. [[CrossRef](#)]
59. Xu, Y.; Hu, H.; Liu, J.; Luo, J.; Qian, G.; Wang, A. PH dependent phosphorus release from waste activated sludge: Contributions of phosphorus speciation. *Chem. Eng. J.* **2015**, *267*, 260–265. [[CrossRef](#)]
60. Min, Y.H.; Hyun, D.Y.; Eum, C.H.; Lee, S.H. A study on the relationship between concentration of phosphorus, turbidity, and pH in water and soil. *Anal. Sci. Technol.* **2011**, *24*, 304–309. [[CrossRef](#)]
61. Wang, S.; Jin, X.; Zhao, H.; Wu, F. Phosphorus fractions and its release in the sediments from the shallow lakes in the middle and lower reaches of Yangtze River area in China. *Colloids Surf. A Physicochem. Eng. Asp.* **2006**, *273*, 109–116. [[CrossRef](#)]
62. Pardo, P.; Rauret, G.; López-Sánchez, J.F. Shortened screening method for phosphorus fractionation in sediments: A complementary approach to the standards, measurements and testing harmonised protocol. *Anal. Chim. Acta* **2004**, *508*, 201–206. [[CrossRef](#)]