

Article

Approach for Evaluating LID Measure Layout Scenarios Based on Random Forest: Case of Guangzhou—China

Xianghua Liao ¹, Jiaxuan Zheng ¹, Chengli Huang ² and Guoru Huang ^{1,3,4,*} 

¹ School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China; yuanbodeshizi@163.com (X.L.); zhengjiaxuan0151@163.com (J.Z.)

² Department of Mathematics, the College of William and Mary, Williamsburg, VA 23185, USA; chuang03@email.wm.edu

³ State Key Laboratory of Subtropical Building Science, School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China

⁴ Guangdong Engineering Technology Research Center of Safety and Greenization for Water Conservancy Project, Guangzhou 510640, China

* Correspondence: huanggr@scut.edu.cn; Tel.: +86-186-2016-8422

Received: 24 May 2018; Accepted: 1 July 2018; Published: 4 July 2018



Abstract: Currently, with the rapid development of many cities, water problems, such as water logging and water quality deterioration, occur inevitably. Thus, sponge city construction and low impact development (LID) utilization have become more important worldwide. However, previous works have failed to address the problem of selecting an optimal LID measure layout scenario by simultaneously considering various evaluation indices without subjective factors. In this study, we applied a new and outstanding statistical classifier, random forest, to aid in addressing this conundrum. It was tested on a case study in LiWan district, Guangzhou city. The following conclusions were drawn. (1) To some extent, LID measures are capable of reducing water discharge and generation of pollutants. (2) Excluding subjective factors, random forest can select an optimal LID measure layout scenario when simultaneously considering multiple indices. This study proposed a novel and effective means to evaluate the hydrologic effects of LID measures when constructing sponge cities and provided a guide for optimizing LID layouts.

Keywords: sponge city; optimal low impact development measure layout scenario; random forest

1. Introduction

Studies addressing water problems have a long tradition. During the past decades, people across the world have been threatened by various ecological disasters, causing heavy economic losses and widespread fatalities. For example, Pall et al. [1] thought widespread floods in the United Kingdom such as those during the spring of 1998, autumn of 2000, winter of 2003, and summer of 2007 were related to extreme precipitation. From 6 May at 7:00 p.m. to 7 May at 4:00 a.m., 2010, 128 precipitation stations in Guangzhou recorded rainfall of more than 100 mm. Moreover, rainfall in the urban area was near 130 mm on average. Rapid socio-economic and urban developments have led to climate change and frequent extreme rainstorm events. The increase in urban impervious area has made sustainable development problems in cities increasingly prominent. The increase has resulted in an increase in the runoff coefficient and peak flow, and the peak flow time has advanced, differing from river basin flooding. In response to the increasing water problems, different strategies based on the situations of different countries are indispensable. However, as developed countries in Europe, North America, and other continents have already rapidly industrialized and urbanized, they faced and addressed

water-related problems earlier than China [2–4]. Because of the frequent occurrence of hazards in China and precedents of other countries, it is of great significance for China to put forward new policies suitable to its own realities, known as integrated urban water management [2]. This is the very reason the concept of sponge city in China was proposed. As some projects of sponge city have been launched, low impact development (LID) measures such as green roofs, permeable pavements, and sunken lawns have caught the designer's eye.

Because LID measures can feasibly purify water and reduce water contaminant concentrations, several researchers have studied their hydrologic performance. For instance, Stovin et al. [5] analyzed the performance characteristics of green roofs and employed the green roof moisture flux model to explain runoff behaviors on the foundation of prolonged monitoring data. In another study, Jia et al. [6] evaluated the benefits of low impact development-best management practices (LID-BMPs) in terms of water quantity and quality based on the storm water management model (SWMM) via a case study conducted in Suzhou, China. Ahiablame et al. [7] developed the long-term hydrologic impact assessment (L-THIA)-LID model and found that LID implementation was effective in minimizing the impact of urbanization on the hydrologic process. Chui et al. [8] analyzed the hydrological performance and cost-effectiveness of various LID designs, aiming to identify the optimal design among them. Xing et al. [9] showed that control efficiencies of water quality and quantity were concerned with LID layouts via weighting analysis and SWMM, appealing that further studies on the efficiency and the management of rainfall runoff control are imperative.

Responding to the problem of combining the optimal proportions of each LID facility may not be the optimal solution to the combination of all facilities, studies on LID layouts and optimization have been conducted. By developing different models, Hu et al. [10] tentatively explored the general thinking of LID layouts and optimization methods in a LID demonstration area. Wang et al. [11] attempted to identify the optimal proportion of each LID appliance in an old city area by simulating various proportions of LID appliances in the SWMM. They also explored the optimal LID combination layout. In addition, Young et al. [12] applied the analytic hierarchy process (AHP) to select a single BMP when simultaneously considering several influential criteria.

Nevertheless, given all the previously mentioned studies, further research is still needed. Because some evaluation approaches such as the AHP and the entropy weight method (EWM) have the disadvantages of considering weights and contain subjective factors, it is still of great significance to put forward a novel approach to evaluate combined LID appliances under different rainstorm intensities. In this study, an innovative approach, namely random forest, was applied to this difficult problem.

Random forest is an integrated non-parametric learning machine based on decision trees. It is compatible for classification and regression of variables. With the advantages of simple principles, precise predictions, relative flexibility, and rapid training speed, it has become a new favorite in substantial academic fields. However, its results are not easily interpreted, that is to say, the meaning of the classification model is hard to describe [13].

In this study, we tentatively explored the hydrologic effects of LID measures via a case study in LiWan district, Guangzhou city. Five LID measure layout scenarios were designed, and six categories of contaminants were tested to investigate the influence of the five scenarios on water quality. A model was developed to quantitatively assess the hydrologic effects of LID appliances under six rainstorm design recurrence periods with the help of the LID controls module in the SWMM. Eight indices were selected as assessment criteria for each scenario. Apart from these measures, random forest algorithm based on Python was applied to evaluate the comprehensive performance of each scenario. Finally, some recommendations for sponge city construction are given at the end of this paper based on the simulation results.

2. Materials and Methods

A simple flow chart of the experiment is shown in Figure 1.

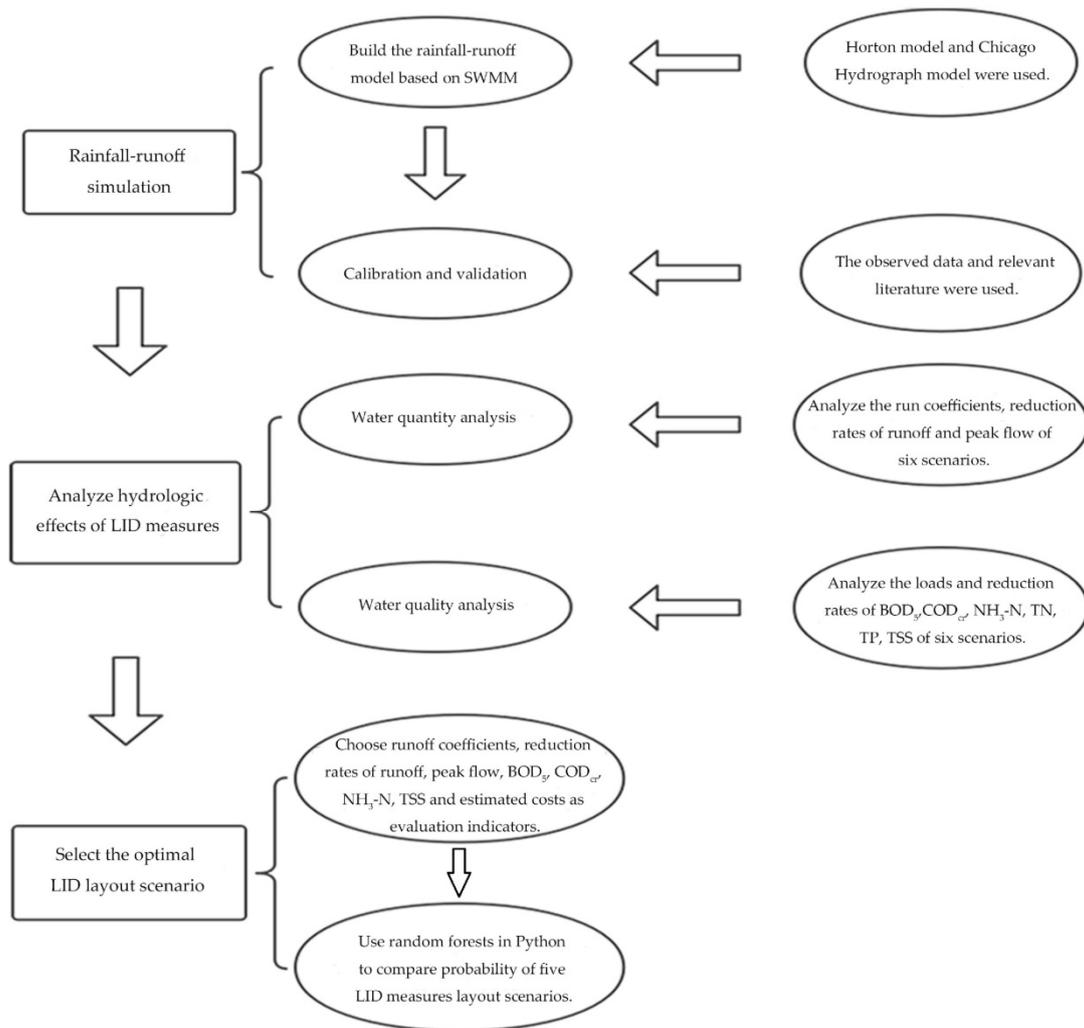


Figure 1. Flow chart of the experiment.

2.1. Research Area

Located in the LiWan district of Guangzhou city and surrounded by rivers on three sides, the research area is residential and covers a total area of 0.4329 km². It has an isolated and closed drainage system with rain and sewage diversion. In this study, roofs, roads, green spaces, and bare lands were chosen to be the underlying surfaces.

Considering the rainwater pipe network, terrain conditions and principles of the SWMM, the model of the area was divided into 128 subcatchments, consisting of 128 nodes, 129 pipes, and 2 outlets (Figure 2).

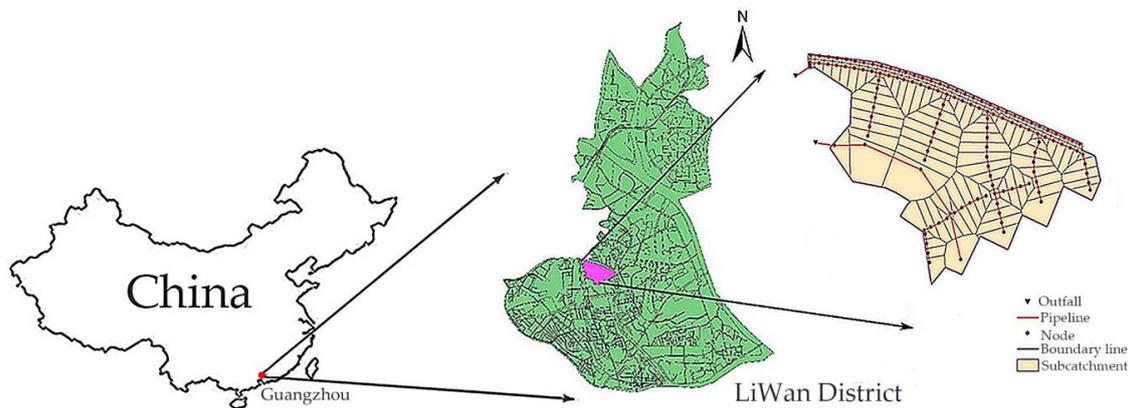


Figure 2. Schematic diagram of the research area.

2.2. SWMM

The SWMM, a dynamic rainfall–runoff simulation model, is available for a single event or long-term simulation of runoff quantity and quality from primarily urban areas [14,15]. First developed by the US Environmental Protection Agency (EPA) in 1971, it is a watershed-scale distributed hydrologic model based on water diffusion. Given the availability of the source code and library functions, it is thought to be the most widely used urban stormwater model throughout the world [16]. Considering the spatial variability and principles of the SWMM, the model of a research area is often simplified into several subcatchments, each of which is generalized as a nonlinear reservoir when completing surface runoff calculations. Since version 5.1.007, the LID controls module has been implemented to simulate the hydrologic effect of rain gardens, permeable pavements, infiltration trenches, etc. Aiming to define the corresponding areal coverage, LID measures can be assigned within the selected subcatchments [17]. According to Ahiablame et al. [18], various combinations of structural layers ought to correspond to various LID practices in the light of permeation theory. In this study, we selected green roofs, permeable pavements, and sunken lawns as LID implementations in each subcatchment and the Horton model was applied to simulate the infiltration of unsaturated soil area under precipitation.

2.3. Random Forest

Put forward by Leo Breiman and Adele Culter, random forest, an integrated non-parametric learning machine, is an ensemble of numerous tree predictors. With the purpose of sample training, considerable classification and regression trees (CART) are often constructed. All trees in the forest are identically distributed and the generation of each tree relies on the values of random vectors [19]. Random forest can fit multiple classification trees to a dataset [20]. In this technique, a small random sample that can interpret variables is chosen and these limited variables are competent to make the best split [21]. Ascribed to this characteristic, this algorithm performs much better than other classifiers based on standard trees [22]. Regarding classification, each tree in random forest votes for the most popular class, which is described as Equation (1). The final result depends on the majority vote of these tree predictors as follows [23]:

$$H(x) = \max_Y \sum_{i=1}^k I(h_i(x) = Y) \quad (1)$$

where $H(x)$ is the classification result of random forest, $h_i(x)$ is the classification result of a single decision tree, Y is the classification object, and $I(\cdot)$ is the indicator function.

The generalization error is given by Equation (2) as follows:

$$PE^* = P_{X,Y}(mg(X,Y) > 0) \quad (2)$$

where P is the probability, E^* is the mathematical expectation, and $mg(X, Y)$ is the margin function. The convergence of the generalization error is given by Equation (3) as follows:

$$\lim_{k \rightarrow \infty} PE^* = P_{X,Y}(P_{\theta}(h(X, \theta_k) = Y) - \max_{j \neq Y} P_{\theta}(h(X, \theta_k) = j) > 0) \quad (3)$$

The results show that random forest will not overfit as more decision trees are added.

To clarify, the implementation process of this eminent algorithm can be summarized as follows.

Step I: Create an original training set termed N , and then, apply the bootstrap method to randomly extract some self-help sample sets in the number of k . Samples that are not extracted each time will be a component of data out of the bag.

Step II: Assume a total of m variables and x variables are randomly selected at each node of each tree, and then a variable with the most classification ability is selected in the x variable. The threshold of the variable classification is determined by checking each classification point.

Step III: Each tree is fully grown without any pruning.

Step IV: A random forest is finally composed of multiple trees, and the new data are classified by a random forest classifier. The classification results are determined according to the vote of the tree classifier.

A strengths, weaknesses, opportunities, and threats (SWOT) analysis of random forest is shown in Figure 3.

<p>Strengths</p> <ol style="list-style-type: none"> 1. Feature subsets are randomly selected. 2. The training speed is fast. 3. The implementation of the algorithm is simple and effective. 4. A classifier with high accuracy can be generated for varieties of data. 	<p>Weakness</p> <ol style="list-style-type: none"> 1. It will be overfitting when dealing with some classification and regression problems with loud noise. 2. The property weights produced by random forests on the data with different value properties are not credible.
<p>Opportunities</p> <ol style="list-style-type: none"> 1. It is useful in data mining, outliers detection and data visualization. 2. It provides an experimental method to detect variable interactions. 3. It can be applied to unlabeled data. 	<p>Threats</p> <p>The texonomic performance of the support vector machine (SVM) can be compared with random forests.</p>

Figure 3. A SWOT analysis of random forest.

2.4. Methodology

2.4.1. Model Building

Aiming to simplify the research area, the equal angle line method was applied to define the boundary lines of the catchment. Subsequently, the Tyson polygon rule was used to divide the catchment into several subcatchments. According to the principle of water nearby discharge,

the research area was finally simplified into 128 subcatchments. In the research conducted by Zhu et al. [24] in the adjacent area, they used rainfall–runoff events during 2013 and 2014 to calibrate the parameters and applied ten other events to validate the model, which promised sufficient precision. Thus, the water quantity parameters in this study, as shown in Table 1, were determined by referring to both the model manual [14] and relevant research mentioned previously. Water quality simulation in SWMM is related to different underlying surfaces. The street sweeping parameters shown in Table 2 were used for water quality simulation in the SWMM.

Table 1. Water quantity parameters of each subcatchment.

Parameter	Physical Meaning	Value
N-Imperv	Manning coefficient of impervious area	0.015
N-Perv	Manning coefficient of permeable area	0.4
Destore-Imperv	Depth of water storage in impervious area (mm)	1
Destore-Perv	Depth of water storage in permeable area (mm)	10
Percent Zero-Imperv	Proportion of no depressions and impervious zones in impervious area	50
MaxRate	Maximum infiltration rate (mm·h ⁻¹)	103.81
MinRate	Minimum infiltration rate (mm·h ⁻¹)	11.44
Decay	Osmotic attenuation coefficient	2.75

Table 2. Street sweeping parameters of various underlying surfaces.

Parameters	Underlying Surfaces		
	Road	Green Roof	Green Space
Sweeping interval/d	1	1	0
Fraction availability	1	0.2	0
Last swept time/d	1	1	0

To calibrate parameters and validate the built model, data of three rainfall events during September 2016 were collected. Five-day biological oxygen demand (BOD₅), chemical oxygen demand (COD_{Cr}), ammoniacal nitrogen (NH₃-N), total nitrogen (TN), total phosphorous (TP), and total suspended solids (TSS) are water quality indices typically used to analyze water pollution of various underlying surfaces. Specifically, data of the rainfall event that occurred on 7 September were used to validate the water quality and quantity parameters of green spaces and roads. Data collected on 2 September were used to validate the same parameters for roofs. The relative errors (*E*) between the simulation results and the observed data of the six contaminants were determined to ensure the precision of the model. The calculation formula is shown as Equation (4):

$$E = \frac{(C_{s,max} - C_{o,max})^2 + (\bar{C}_s - \bar{C}_o)^2 + (C_{s,same} - C_{o,same})^2}{(C_{o,max} - \bar{C})^2 + (\bar{C}_o - \bar{C})^2 + (C_{o,same} - \bar{C})^2} \times 100\% \quad (4)$$

where *E* is the relative error, %; *C_{s,max}* is the maximum simulation value; *C_{o,max}* is the maximum observed value; \bar{C}_s is the average value of simulation; \bar{C}_o is the average value of observations; *C_{s,same}* is an instantaneous value of simulation; *C_{o,same}* is an instantaneous value of observations at the same time of the simulation; and \bar{C} is the average value of *C_{o,max}*, \bar{C}_o , and *C_{o,same}*.

The relative errors (*E*) of the three previously mentioned rainfall events were analyzed. Based on the results shown in Figure 4, the relative errors (*E*) were overwhelmingly less than 15%, which illustrates that the simulation results conform to the precision requirements.

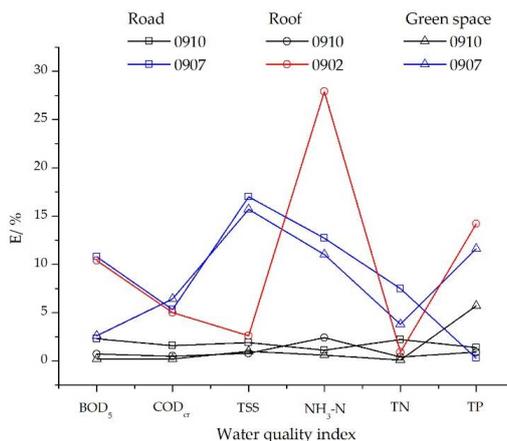


Figure 4. The relative errors (E) between the simulation and monitoring data.

2.4.2. Rainfall Simulation

In this study, the formula for storm intensity (Equation (5)) was applied to calculate rainfall in Guangzhou. On the foundation of storm rainfall data over the last 30 years, this formula was derived using the multi sample method and is as follows [25,26]:

$$q = \frac{3618.427(1 + 0.438LgP)}{(t + 11.259)^{0.75}} \tag{5}$$

where q is the design storm intensity (L/s·ha), t is the duration of a rainfall event (min), and P is the rainstorm design recurrence period (year).

Additionally, a rainfall pattern that is near the reality of the research area using the Chicago Hydrograph Model [27] was selected to describe the procedures of rainfall events under six rainstorm design recurrence periods (i.e., $P = 0.5, 1, 2, 5, 10,$ and 20 years). The duration of each rainstorm design recurrence period in this manuscript was determined to be 120 min. The Chicago rainfall process lines of different rainstorm design recurrence periods are shown in Figure 5.

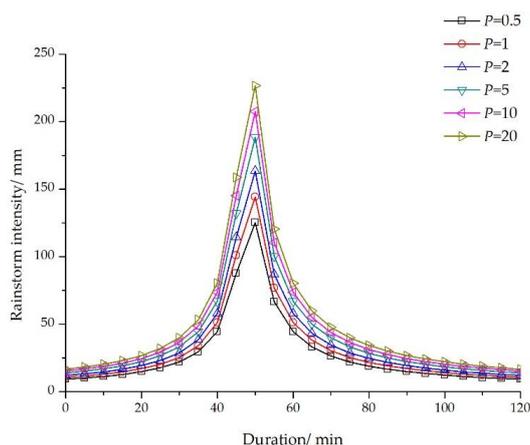


Figure 5. Chicago rainfall process lines of different rainstorm design recurrence periods.

2.4.3. Computational Principles of LID Measures

LID rainwater utilization technology mainly utilizes different functions of different facilities to decrease total runoff volume and peak flow, purify and utilize rainwater, and improve the eco-environment. A single LID measure such as a sunken lawn may have multiple functions.

LID measures are formed using different layers. For example, a permeable pavement is composed of pavers, sub-base, and a reservoir course; a green roof is composed of a vegetation layer, growing medium, drainage filter, and drainage layer; and sunken lawns and bio-retention cells are composed of storage layers, cultivated soils, and local soils. Based on the principles of various LID measures, the SWMM has incorporated seven LID control measures: bio-retention cell, infiltration trench, permeable pavement, rain barrel, vegetative swale, rain garden, and green roof. In this model, these measures are also represented by combinations of different vertical layers, including the surface and a soil, pavement, storage, and drain layer. The specific layers that each LID measure contains are shown in Table 3.

Table 3. Structural layers of each LID measure.

LID type	Surface	Soil	Pavement	Storage	Drain
Bio-retention cell	✓	✓	✗	○	○
Infiltration trench	✓	✗	✗	✓	○
Permeable pavement	✓	○	✓	✓	○
Rain barrel	✗	✗	✗	✓	✓
Vegetative swale	✓	✗	✗	✗	✗
Rain garden	✓	✓	✗	✗	✗
Green roof	✓	✓	✗	✗	✓ (drainage mat)

Note: ✓ means necessary, ✗ means unnecessary, ○ means optional.

Among these seven LID measures, the bio-retention cell is introduced as an example for LID computational principles. The SWMM has generalized each LID measure as a reservoir that contains fillers with pores (Figure 6). Direct rainwater and rainwater from other impervious areas flows into the surface of the reservoir, in which vegetation has been planted. Different parameters are also set for the surface, including berm height, vegetation volume fraction, Manning coefficient, and slope. When the surface capacity is saturated, runoff discharges from the top. The soil layer needs to consist of coefficients of thickness, porosity, field capacity, wilting point, conductivity, conductivity slope, and suction head. Specifically, porosity determines the volume of the pores, which is the volume of the water that can be stored; wilting point measures the water content of the soil when it is the driest; and field capacity indicates the maximal water content when there is no exchange of water between the soil layer and other layers such that it must be greater than the wilting point, otherwise the model would report errors. The storage layer is similar to a reservoir, and needs a thickness, porosity, and seepage rate. Rainwater can seep via the storage layer to the local soil, and if the drainage layer is set, the water would discharge through it. The drainage layer must contain a flow coefficient, flow exponent, and offset height.

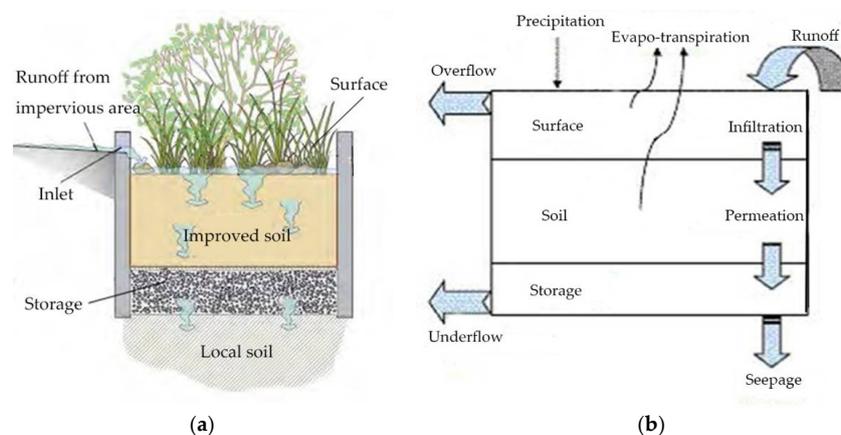


Figure 6. Schematic diagram of a bio-retention cell: (a) components; and (b) hydrologic process.

2.4.4. LID Measure Layout Approaches

In the SWMM, LID control measures are established in the subcatchment properties, and either different or identical LID measures can be established in the same subcatchment. There are two means to establish LID measures in the subcatchments: adding a single LID measure or multiple LID measures into a subcatchment with no existing LID measures (Figure 7a) or adding only one LID measure into one subcatchment (Figure 7b). The first approach allows mixing of different LID control measures in the same subcatchment and all function simultaneously. Each LID measure addresses the water flow data in its corresponding area, and the outflow data from one LID measure cannot be taken as the inflow data for another LID measure.

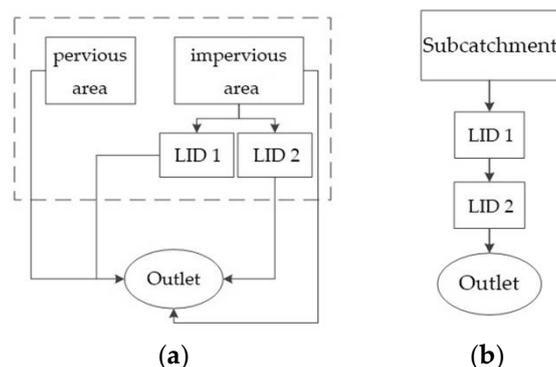


Figure 7. Setup of LID measures in the SWMM: (a) first LID measures layout approach; and (b) second LID measures layout approach.

The second approach allows a LID measure to occupy the entire subcatchment, and it can take the outflow from its upstream subcatchment as its inflow, thus satisfying flow direction settings. However, this approach requires building new subcatchments and establishing the corresponding LID properties.

To utilize either the first or the second LID measure layout approach, the impervious rates of the subcatchments have to be adjusted according to the respective measures. For instance, the impervious rate needs to be re-calculated if an originally impervious area becomes pervious, and for the second approach specifically, the areas for the old and new subcatchments need to be adjusted.

Referring to a prior study in Shenzhen [10] and considering the actual situation of the research area, five LID measure layout scenarios were designed to determine the hydrologic effects of LID measures on the research area. The content of each scenario can be summarized as follows:

- Scenario A: 35% of the roofs are transformed into green roofs, through which rainwater directly discharges into the drainage system. In addition, 50% of the green spaces are transformed into sunken lawns, which are used to collect rain generated from roofs without transformation.
- Scenario B: 50% of the roofs are converted into green roofs, through which rainwater directly discharges into the drainage system. In addition, 50% of the green spaces are converted into sunken lawns, which are used to collect the rain generated from the roofs without transformation.
- Scenario C: 35% of the roads are transformed into permeable pavements, through which rainwater directly discharges into the drainage system. In addition, 50% of the green spaces are transformed into sunken lawns, which are responsible for the purification of rainwater generated from the roofs.
- Scenario D: 50% of the roads are transformed into permeable pavement, through which rainwater directly discharges into the drainage system. In addition, 50% of the green spaces are converted into sunken lawns, which are responsible for the purification of rainwater generated from roofs.
- Scenario E: A combination of scenarios A and C; 35% of the roofs are transformed into green roofs and 35% of the roads are converted into permeable pavement. Rainwater generated from both discharges directly into the drainage system. In addition, 50% of the green spaces

are retrofitted into sunken lawns, which are used to purify rainwater produced by the roofs without transformation.

Referring to similar studies [10,28,29] and the model manual [14], the parameters of the LID measures previously mentioned are listed in Tables 4–6.

Table 4. Parameters of green roofs in the SWMM.

Surface	Berm Height/mm	Vegetation Volume Fraction	Surface Roughness	Surface Slope/%	-	-	-
	50	0.2	0.15	1	-	-	-
Soil	Thickness/mm	Porosity	Field capacity	Wilting point	Conductivity/mm·h ⁻¹	Conductivity slope	Suction head/mm
	250	0.18	0.15	0.03	18	10	90
Drainage mat	Thickness/mm	Void fraction	Roughness	-	-	-	-
	30	0.43	0.3	-	-	-	-

Table 5. Parameters of sunken lawns in the SWMM.

Surface	Berm Height/mm	Vegetation Volume Fraction	Surface Roughness	Surface Slope/%	-	-	-
	200	0.2	0.15	1	-	-	-
Soil	Thickness/mm	Porosity	Field capacity	Wilting point	Conductivity/mm·h ⁻¹	Conductivity slope	Suction head/mm
	500	0.18	0.1	0.03	3.6	10	90
Storage	Thickness/mm	Void fraction	Seepage rate/mm·h ⁻¹	Clogging factor	-	-	-
	200	0.2	500	0	-	-	-
Drain	Flow coefficient/mm·h ⁻¹	Flow exponent	Offset height/mm	-	-	-	-
	0	0.5	50	-	-	-	-

Table 6. Parameters of permeable pavements in the SWMM.

Surface	Berm Height/mm	Vegetation Volume Fraction	Surface Roughness	Surface Slope/%	-
	20	0	0.11	1	-
Pavement	Thickness/mm	Void ratio	Impervious surface fraction	Permeability/mm·h ⁻¹	Clogging factor
	60	0.1	0	200	0
Storage	Thickness/mm	Void ratio	Seepage rate/mm·h ⁻¹	Clogging factor	-
	250	0.43	500	0	-
Drain	Flow coefficient/mm·h ⁻¹	Flow exponent	Offset height/mm	-	-
	0	0.5	6	-	-

2.4.5. Calculation of Estimated Cost for Each LID Measure Layout Scenario

Based on the construction area of each LID measure for the five scenarios (Figure 8) and the unit price of different types of LID measures based on the investment of more than 40 domestic projects, the estimated cost of each scenario can be calculated as shown in Table 7.

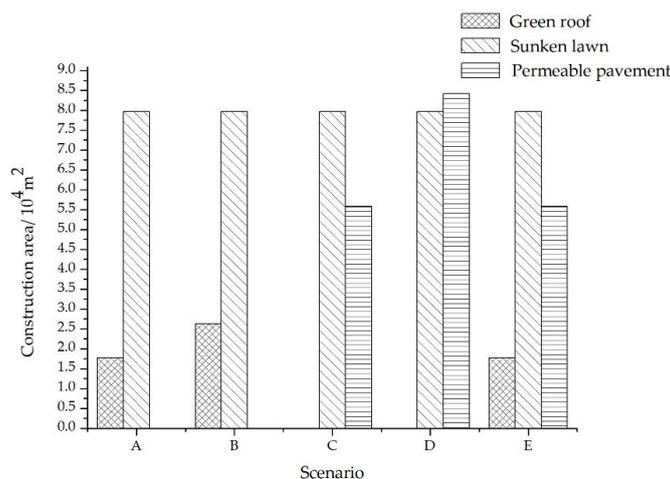


Figure 8. LID measure construction area of each scenario.

Table 7. Estimated costs of different LID measure layout scenarios.

Scenario	LID Measure	Unit Price/USD·m ⁻²	Cost/10 ⁴ USD	Total Cost/10 ⁴ ·USD
A	Green roof	31.18	55.18	129.72
	Sunken lawn	9.35	74.54	
B	Green roof	31.18	82.00	156.54
	Sunken lawn	9.35	74.54	
C	Permeable pavement	35.07	195.71	270.25
	Sunken lawn	9.35	74.54	
D	Permeable pavement	35.07	295.32	369.86
	Sunken lawn	9.35	74.54	
E	Green roof	31.18	55.18	325.43
	Permeable pavement	35.07	195.71	
	Sunken lawn	9.35	74.54	

2.4.6. Selection of the Optimal LID Measure Layout Scenario

Aiming to determine the most reasonable layout among the five scenarios, runoff coefficient, reduction rates of runoff, peak flow, BOD₅, COD_{cr}, NH₃-N, TSS, and estimated cost of each scenario under three rainstorm design recurrence periods (i.e., $P = 1, 5,$ and 10 years) were chosen for comprehensive performance indices. Random forest classification of Scikit-Learn machine learning in Python was applied to analyze the dataset. Values of the partial parameters are shown in Table 8, while default values were used for parameters not shown.

Table 8. Partial parameters of random forest in Python.

Parameter	Value
max_features	sqrt
n_estimators	100
Min_sample_leaf	1
n_jobs	4

3. Results and Discussion

The water quantity simulation of the five LID measure layout scenarios under six rainstorm design recurrence periods in the SWMM are shown in Table 9.

Table 9. Water quantity simulation of each LID measure layout scenario under different rainstorm design recurrence periods.

Scenario	Rainstorm Design Recurrence Period/Year	Runoff Coefficient	Reduction Rate/%	
			Runoff	Peak Flow
Without LID measures	$P = 0.5$	0.458	-	-
	$P = 1$	0.483	-	-
	$P = 2$	0.522	-	-
	$P = 5$	0.570	-	-
	$P = 10$	0.602	-	-
	$P = 20$	0.603	-	-
A	$P = 0.5$	0.306	38.00	31.05
	$P = 1$	0.329	36.67	31.00
	$P = 2$	0.358	36.05	30.42
	$P = 5$	0.394	35.04	29.38
	$P = 10$	0.418	35.00	28.52
	$P = 20$	0.441	26.74	27.73
B	$P = 0.5$	0.308	37.54	30.42
	$P = 1$	0.330	36.27	30.39
	$P = 2$	0.359	35.72	29.87
	$P = 5$	0.394	34.87	28.89
	$P = 10$	0.417	34.87	28.11
	$P = 20$	0.440	26.68	27.42
C	$P = 0.5$	0.234	49.67	48.76
	$P = 1$	0.256	47.45	48.59
	$P = 2$	0.288	44.98	47.56
	$P = 5$	0.329	41.98	45.77
	$P = 10$	0.357	40.56	44.40
	$P = 20$	0.381	32.27	43.18
D	$P = 0.5$	0.198	57.59	56.51
	$P = 1$	0.220	55.03	56.33
	$P = 2$	0.251	50.99	55.15
	$P = 5$	0.291	48.77	53.13
	$P = 10$	0.318	47.04	51.57
	$P = 20$	0.341	39.31	50.20
E	$P = 0.5$	0.232	53.64	47.43
	$P = 1$	0.253	51.92	47.35
	$P = 2$	0.279	50.58	46.51
	$P = 5$	0.311	49.03	45.01
	$P = 10$	0.333	48.30	43.78
	$P = 20$	0.355	41.09	42.68

According to the data listed in Table 9, the following conclusions can be readily drawn. In response to the increase in the rainstorm design recurrence period, the runoff and its coefficient, peak flow, as well as contaminant concentration of all the scenarios increases while the reduction rates decrease. In addition, comparing the five LID measure layout scenarios to the scenario without any LID appliance from multiple indices, LID measures have enormous effects on reducing water quantity and contaminant concentrations. Under three of the rainstorm design recurrence periods (i.e., $P = 0.5$, 1, and 2 years), the five scenarios with runoff coefficients ranking from low to high are D, E, C, A, and B, and the order of the reduction rates of the runoff ranking from low to high is the exact opposite. In terms of the reduction rates of the peak flow under the same rainstorm intensities, the five scenarios ranking from high to low are D, C, E, A, and B. Under the other three rainstorm design recurrence periods (i.e., $P = 5$, 10, and 20 years), the five scenarios with runoff coefficients ranking from low to high are E, D, C, A, and B, whose order is the opposite to the reduction rates of runoff. In terms of the reduction rates of the peak flow under the previously mentioned three rainstorm intensities, the five scenarios ranking from high to low are D, C, E, A, and B.

Based on the water quality simulation of the model (Table 10), reduction rates of the six pollutants under different rainstorm design recurrence periods can be calculated (Figure 9).

Table 10. Pollutant loads of each LID measure layout scenario under different rainstorm design recurrence periods.

Scenario	Rainstorm Design Recurrence Period/Year	Pollutant Loads/kg					
		BOD ₅	COD _{cr}	NH ₃ -N	TN	TP	TSS
Without LID measures	P = 0.5	653.30	1833.35	25.77	56.48	2.31	974.00
	P = 1	711.35	2061.94	27.55	61.27	2.84	1051.12
	P = 2	763.44	2280.90	29.44	66.16	3.49	1118.65
	P = 5	807.41	2501.75	31.30	71.18	4.43	1170.87
	P = 10	824.09	2616.74	33.02	73.96	5.17	1191.32
	P = 20	830.84	2675.02	33.69	75.15	5.60	1198.63
A	P = 0.5	347.83	984.69	13.20	27.99	1.08	471.86
	P = 1	382.13	1127.40	14.18	30.59	1.34	521.48
	P = 2	415.27	1267.52	15.21	33.38	1.66	569.59
	P = 5	439.18	1409.41	16.07	35.89	2.15	604.45
	P = 10	469.20	1523.38	17.32	38.72	2.54	640.43
	P = 20	479.99	1597.27	18.11	40.40	2.95	655.01
B	P = 0.5	342.03	961.21	13.02	27.59	1.04	465.71
	P = 1	377.15	1104.80	14.00	30.19	1.30	516.07
	P = 2	411.04	1247.15	15.03	32.98	1.61	565.08
	P = 5	435.73	1390.17	15.89	35.48	2.09	600.60
	P = 10	466.04	1507.82	17.15	38.33	2.47	637.45
	P = 20	477.15	1583.66	17.93	40.02	2.88	652.47
C	P = 0.5	218.38	603.60	8.69	18.38	0.66	306.30
	P = 1	247.62	713.64	9.48	20.53	0.84	348.12
	P = 2	278.69	836.61	10.41	23.08	1.09	393.93
	P = 5	308.17	979.77	11.35	25.80	1.48	434.93
	P = 10	330.25	1075.83	12.30	28.02	1.80	461.45
	P = 20	340.65	1146.21	12.99	29.53	2.14	474.60
D	P = 0.5	163.26	440.89	6.80	14.26	0.48	234.64
	P = 1	190.29	533.86	7.51	16.20	0.62	272.87
	P = 2	220.38	643.80	8.36	18.57	0.81	316.74
	P = 5	252.15	780.82	9.26	21.23	1.14	359.63
	P = 10	273.23	875.13	10.10	23.25	1.41	385.39
	P = 20	284.37	947.57	10.73	24.68	1.70	399.59
E	P = 0.5	205.09	544.19	8.32	17.41	0.56	287.13
	P = 1	235.51	651.38	9.07	19.47	0.72	329.06
	P = 2	267.71	770.57	9.95	21.88	0.92	374.57
	P = 5	299.98	914.03	10.81	24.46	1.25	418.21
	P = 10	323.49	1014.63	11.72	26.62	1.51	447.62
	P = 20	336.13	1094.08	12.39	28.17	1.81	464.59

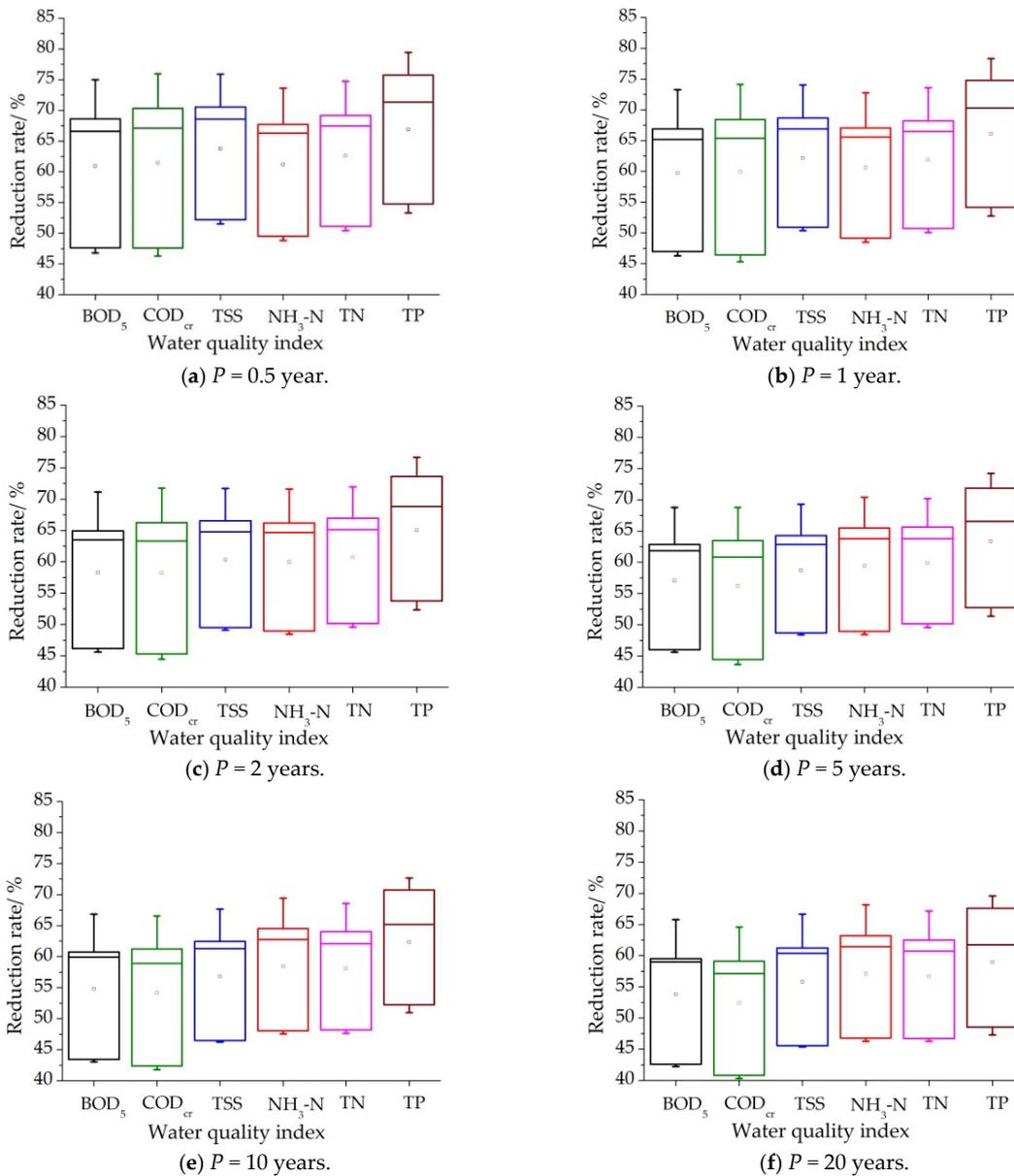


Figure 9. Pollutant reductions from different LID measure layout scenarios.

Deducing from the data shown in Table 10 and the tendencies shown in Figure 9, the scenarios with reduction rates for the six pollutants ranking from high to low are D, E, C, B, and A under all six rainstorm design recurrence periods. Although all contaminant concentrations increase with the increase in rainstorm intensity, the reduction rates among all the six contaminants decrease.

In virtue of the random forest algorithm in Python, the probabilities of each LID measure layout scenario to be selected as the best scheme are listed in Table 11.

Table 11. Probability of five LID measure layout scenarios to be the optimal under rainstorm design recurrence periods of $P = 1, 5,$ and 10 years.

Scenario	Probability		
	$P = 1$ Year	$P = 5$ Years	$P = 10$ Years
A	0.02	0.05	0.10
B	0.07	0.04	0.07
C	0.17	0.18	0.17
D	0.54	0.44	0.38
E	0.20	0.29	0.28

From a short review of the data shown in Table 11, the following findings are evident. Comparatively speaking, Scenario D is the optimal scenario when considering runoff coefficient, reduction rates of runoff, peak flow, BOD₅, COD_{cr}, NH₃-N, TSS, and estimated cost at the same time under rainstorm design recurrence periods of $P = 1, 5,$ and 10 years. The probability of Scenario D being selected as optimal decreases with an increase in rainstorm intensity.

4. Conclusions

At present, LID utilization is increasingly important for the avoidance of substantial ecological disasters throughout the world. To explore the effects of LID measures on water quality and quantity in a residential area in Guangzhou, a model in the SWMM was developed and five LID measure layout scenarios were analyzed. Focusing on the most reasonable scenario among the five, random forest algorithm in Python was applied simultaneously considering eight indices. Overall, it was demonstrated that there are far-reaching effects resulting from implemented LID measures, which play an important role in purifying water, reducing flood risk, and providing a better environment for citizens. Moreover, via the case study in a residential area in the LiWan district in Guangzhou, the hypothesis that the random forest algorithm can select the most reasonable LID measure layout scenario was tested. Despite the limitations, these conclusions are valuable for sponge city construction, which is an important measure for sustainable urban development and storm water management. For the future research, a systematic study based on this methodology will be carried out in accordance with the actual planning and design scenarios of the sponge city.

Based on the results of this study, several recommendations can be made as follows. First, original roofs can be retrofitted into green roofs and original roads can be retrofitted into permeable pavements to decrease the level of imperviousness in residential areas, and sunken lawns are qualified to directly purify rainwater discharged into them. Second, because the comprehensive reduction effects of LID have a correlation with the combination of different LID measures as well as the proportion of the identical LID measure, selecting the optimal LID layout in a more objective manner is important. Third, because the random forest algorithm can calculate the probability of each LID layout scenario to be the optimal under specific rainstorm design recurrence periods without considering the weight of each selected index, it is a more reasonable and effective method than the AHP and the EWM.

Author Contributions: X.L. and J.Z. developed the model and analyzed the simulation results. C.H., X.L. and J.Z. discussed the results, drafted the manuscript, and critically revised the paper. G.H. designed and supervised the study.

Funding: This study was financially supported by the National Natural Science Fund of China (51739011); the Guangzhou Science and Technology Planning Project (201707020020); the Science and Technology Planning Project of Guangdong Province, China (2016A020223003); and the National Training Program of Innovation and Entrepreneurship for Undergraduates (201710561185).

Acknowledgments: The authors are grateful to the editor and the three anonymous reviewers for their insightful comments and constructive advice, which greatly improved the manuscript.

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

References

- Pall, P.; Aina, T.; Stone, D.A.; Stott, P.A.; Nozawa, T.; Hilberts, A.G.; Lohmann, D.; Allen, M.R. Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* **2011**, *470*, 382–385. [[CrossRef](#)] [[PubMed](#)]
- Hao, W.; Chao, M.; Liu, J.H.; Shao, W.W. A new strategy for integrated urban water management in China: Sponge city. *Sci. China Technol. Sci.* **2018**, *61*, 1–13.
- Loubet, P.; Roux, P.; Loiseau, E.; Bellonmaurel, V. Life cycle assessments of urban water systems: A comparative analysis of selected peer-reviewed literature. *Water Res.* **2014**, *67*, 187–202. [[CrossRef](#)] [[PubMed](#)]
- Saraswat, C.; Kumar, P.; Mishra, B.K. Assessment of stormwater runoff management practices and governance under climate change and urbanization: An analysis of Bangkok, Hanoi and Tokyo. *Environ. Sci. Policy* **2016**, *64*, 101–117. [[CrossRef](#)]
- Stovin, V.; Vesuviano, G.; Kasmin, H. The hydrological performance of a green roof test bed under UK climatic conditions. *J. Hydrol.* **2012**, *414*, 148–161. [[CrossRef](#)]
- Jia, H.; Ma, H.; Sun, Z.; Yu, S.; Ding, Y.; Liang, Y. A closed urban scenic river system using stormwater treated with LID-BMP technology in a revitalized historical district in China. *Ecol. Eng.* **2014**, *71*, 448–457. [[CrossRef](#)]
- Ahiablame, L.M.; Engel, B.A.; Chaubey, I. Effectiveness of low impact development practices in two urbanized watersheds: Retrofitting with rain barrel/cistern and porous pavement. *J. Environ. Manag.* **2013**, *119*, 151–161. [[CrossRef](#)] [[PubMed](#)]
- Chui, T.F.M.; Liu, X.; Zhan, W. Assessing cost-effectiveness of specific LID practice designs in response to large storm events. *J. Hydrol.* **2016**, *533*, 353–364. [[CrossRef](#)]
- Xing, W.; Li, P.; Cao, S.B.; Gan, L.L.; Liu, F.L.; Zuo, J.E. Layout effects and optimization of runoff storage and filtration facilities based on SWMM simulation in a demonstration area. *Water Sci. Eng.* **2016**, *9*, 115–124. [[CrossRef](#)]
- Hu, A.; Ren, X.; Ding, N.; Tang, W. LID Facilities Layout and optimization in an area in Shenzhen based on SWMM. *China Water Wastewater* **2015**, *31*, 96–100. (In Chinese)
- Wang, T.; Diao, X.M.; Liu, J.; Luan, M.; Ding, N. Optimization of LID layout proportions in old city area based on SWMM. *South-to-North Water Transf. Water Sci. Technol.* **2017**, *15*, 29–43, 128. (In Chinese)
- Young, K.D.; Younos, T.; Dymond, R.L.; Kibler, D.F.; Lee, D.H. Application of the analytic hierarchy process for selecting and modeling stormwater best management practices. *J. Contemp. Water Res. Educ.* **2010**, *146*, 50–63. [[CrossRef](#)]
- Jake VanderPlas. *Python Data Science Handbook*, 1st ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2016; p. 433, ISBN 9781491912058.
- Rossmann, L.A. *Storm Water Management Model User's Manual, Version 5.0*; National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency: Cincinnati, OH, USA, 2010.
- Cipolla, S.S.; Maglionico, M.; Stojkov, I. A long-term hydrological modelling of an extensive green roof by means of SWMM. *Ecol. Eng.* **2016**, *95*, 876–887. [[CrossRef](#)]
- Shao, Z.; Zhang, X.; Li, S.; Deng, S.; Chai, H.; Shao, Z.; Zhang, X.; Li, S.; Deng, S.; Chai, H. A novel SWMM based algorithm application to storm sewer network design. *Water* **2017**, *9*, 747. [[CrossRef](#)]
- Palla, A.; Gnecco, I. Hydrologic modeling of low impact development systems at the urban catchment scale. *J. Hydrol.* **2015**, *528*, 361–368. [[CrossRef](#)]
- Ahiablame, L.; Shakya, R. Modeling flood reduction effects of low impact development at a watershed scale. *J. Environ. Manag.* **2016**, *171*, 81–91. [[CrossRef](#)] [[PubMed](#)]
- Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
- Cutler, D.R.; Edwards, E.T., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)] [[PubMed](#)]
- Breiman, L.L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees (CART). *Encycl. Ecol.* **1984**, *40*, 582–588.
- Liaw, A.; Wiener, M.; Liaw, A. Classification and regression by random forests. *R News* **2002**, *2/3*, 18–22.
- Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]

24. Zhu, Z.; Chen, Z.; Chen, X.; He, P. Approach for evaluating inundation risks in urban drainage systems. *Sci. Total Environ.* **2016**, *553*, 1–12. [[CrossRef](#)] [[PubMed](#)]
25. Ministry of Housing and Urban-Rural Development, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. *Code for Design of Outdoor Wastewater Engineering (GB 50014-2006)*; Ministry of Housing and Urban-Rural Development, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China: Beijing, China, 2011. (In Chinese)
26. Wu, H.; Huang, G.; Meng, Q.; Zhang, M.; Li, L. Deep tunnel for regulating combined sewer overflow pollution and flood disaster: A case study in Guangzhou city, China. *Water* **2016**, *8*, 329. [[CrossRef](#)]
27. Keifer, C.J.; Chu, H.H. Synthetic storm pattern for drainage design. *J. Hydraul. Div.* **1957**, *83*, 1–25.
28. Cai, Q.; Chen, Z.; Chen, X.; Chen, X.; Zhang, D. Simulation of control efficiency of low impact development measures for urban stormwater. *Water Resour. Prot.* **2017**, *33*, 31–36. (In Chinese)
29. Su, Y.; Wang, S.; Che, W.; Wei, Y.; Dong, Y. Optimization design of sunken greenbelt based on the concept of “sponge city”. *South Archit.* **2014**, *3*, 39–43. (In Chinese)



© 2018 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 (<http://creativecommons.org/licenses/by/4.0/>).