

# Multiple Regression Model of a Soak-Away Rain Garden in Singapore

Sivarajah Mylevaganam\*, Ting Fong May Chui, Jianguong Hu

Department of Civil and Environmental Engineering, National University of Singapore, Singapore, Singapore  
Email: \*[sivaloga@hushmail.com](mailto:sivaloga@hushmail.com)

Received 2 April 2015; accepted 24 April 2015; published 27 April 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Under the possible hydrological conditions, with a design hyetograph of 3-month average rainfall intensities of Singapore, multiple regression equations on hydrological processes, specifically on overflow volume, average vertical ex-filtration rate and horizontal flow coefficient, of a soak-away rain garden are established based on simulated results of a mathematical model. The model that is based on Richard's equation is developed using COMSOL Multiphysics. The regression equation on overflow volume and the regression equation on log of horizontal flow coefficient show a very strong relationship with the independent variables (saturated hydraulic conductivity of the filter media, saturated hydraulic conductivity of the *in-situ* soil, depth to groundwater table, and surface area of the soak-away rain garden). The coefficients of determination of the fitted equations on overflow volume and log of horizontal flow coefficient were 0.992 and 0.986, respectively. However, the regression equation on average vertical ex-filtration rate has high p-values (p-values > significance level,  $\alpha = 0.01$ ) for saturated hydraulic conductivity of the *in-situ* soil and surface area of the soak-away rain garden. Thus, forward stepwise regression was used to develop the best regression equation on average vertical ex-filtration rate with saturated hydraulic conductivity of the filter media and depth to groundwater table. The coefficient of determination of the fitted equation was found to be 0.911. These easy to use regression equations will be of great utility for local managers in the design of soak-away rain gardens.

## Keywords

COMSOL Multiphysics, Soak-Away Rain Garden, Overflow Volume, Multiple Regression, Ex-Filtration

---

## 1. Introduction

In catchment hydrology, in practice it is not feasible to measure a desired hydrological variable for every possi-

\*Corresponding author.

ble hydrological condition mainly due to financial constraints. This limitation, among others, has promoted the application of mathematical models, whose basic principle is the process of solving physical problems by appropriate simplification of reality, in the field of catchment hydrology, specifically in ground and surface water hydrology, to solve many hydrological problems. Hydrological process, such as overflow volume, average vertical ex-filtration rate and horizontal flow coefficient, of a soak-away rain garden, is one of those hydrological problems.

Soak-away rain gardens, shallow, landscaped depressions commonly located in parking lots or within small pockets in residential areas, receive stormwater runoff, attenuate surface water and enable it to percolate into the surrounding ground. As catchments become urbanized due to population growth, the impervious surfaces created by buildings and pavements cause rainwater to flow quickly over the landscape. To mitigate the adverse impact of urbanization such as increased flooding and depleted groundwater recharge around the world, several best management practices, in other words, green infrastructures, have been practiced [1]-[5], and soak-away rain garden is one of them. Despite the rapid acceptance of soak-away rain gardens throughout the world by water managers and land-use planners, detailed hydrologic performance information and related hydrologic design guidelines of soak-away rain gardens are not currently available for many regions including Singapore [1] [2]. On the other hand, to have a rapid assessment of soak-away rain gardens on a range of potential hydrologic conditions (e.g., size of the soak-away rain garden, saturated hydraulic conductivity of the *in-situ* soil, and saturated hydraulic conductivity of the filter media), detailed hydrologic performance information and related hydrologic design guidelines of soak-away rain gardens from model simulation results need to be rendered into easy to use look-up charts or hydrologic design charts that are specific for local conditions [1] [2]. However, it is not feasible to model for every possible hydrological condition due to resource constraints and the continuous nature of those possible hydrological conditions. Therefore, a need exists to extrapolate and interpolate information from the available model simulation results. In addition, there is also a need to assess the likely hydrological impact of future system response. Therefore, it is the objective of this study to establish regression equations on hydrological processes, specifically on overflow volume, average vertical ex-filtration rate and horizontal flow coefficient, of a soak-away rain garden, based on simulated results of a mathematical model using COMSOL Multiphysics [6] [7] under the possible hydrological conditions.

## 2. Multiple Regression Equation

In many hydrological applications, the relationship proposed by a simple linear regression equation, which has only one “X” variable, called the independent variable, does not sufficiently explain the variation in “Y”, called the dependent variable. This is because the “Y” variable in most hydrological applications depends on more than one independent variable. In such cases, multiple regressions are used to explore the relationship between a dependent variable (e.g., ex-filtration and overflow volume) and a set of independent variables (e.g., size of the soak-away rain garden, saturated hydraulic conductivity of the *in-situ* soil, and saturated hydraulic conductivity of the filter media). Each independent variable is “weighted” differently based on its separate correlation with “Y” and its correlation with the other independent variables in such a way that the combination of independent variables predicts the best “Y” than any single independent variable by itself. In general, the multiple regression equation follows the following format:  $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots$ . The “weight” given to each independent variable in the equation is represented by the letter “b” with the corresponding subscript. In general, these weights are unknown and will have to be estimated from sample data. In this study, the sample data on a range of potential hydrologic conditions (e.g., size of the soak-away rain garden, saturated hydraulic conductivity of the *in-situ* soil, and saturated hydraulic conductivity of the filter media) was generated by simulating a mathematical model based on Richard’s equation [6]-[8] using COMSOL Multiphysics.

## 3. COMSOL Multiphysics

COMSOL Multiphysics is a powerful interactive environment for modeling and solving partial differential equations in scientific and engineering problems. The software provides a powerful integrated desktop environment with a Model Builder where the users get full overview of the model and access to all functionality. With COMSOL Multiphysics the users can easily extend models for one type of physics into multiphysics models that solve coupled physics phenomena. COMSOL Multiphysics also allows the users to perform various types of studies such as stationary and time-dependent studies. When solving the models, COMSOL Multiphysics uses the

proven finite element method. The software runs the finite element analysis together with adaptive meshing and error control using a variety of numerical solvers. A more detailed description of the mathematical and numerical foundation is in the COMSOL Multiphysics Reference Guide [7].

## 4. Methodology and Discussion of Results

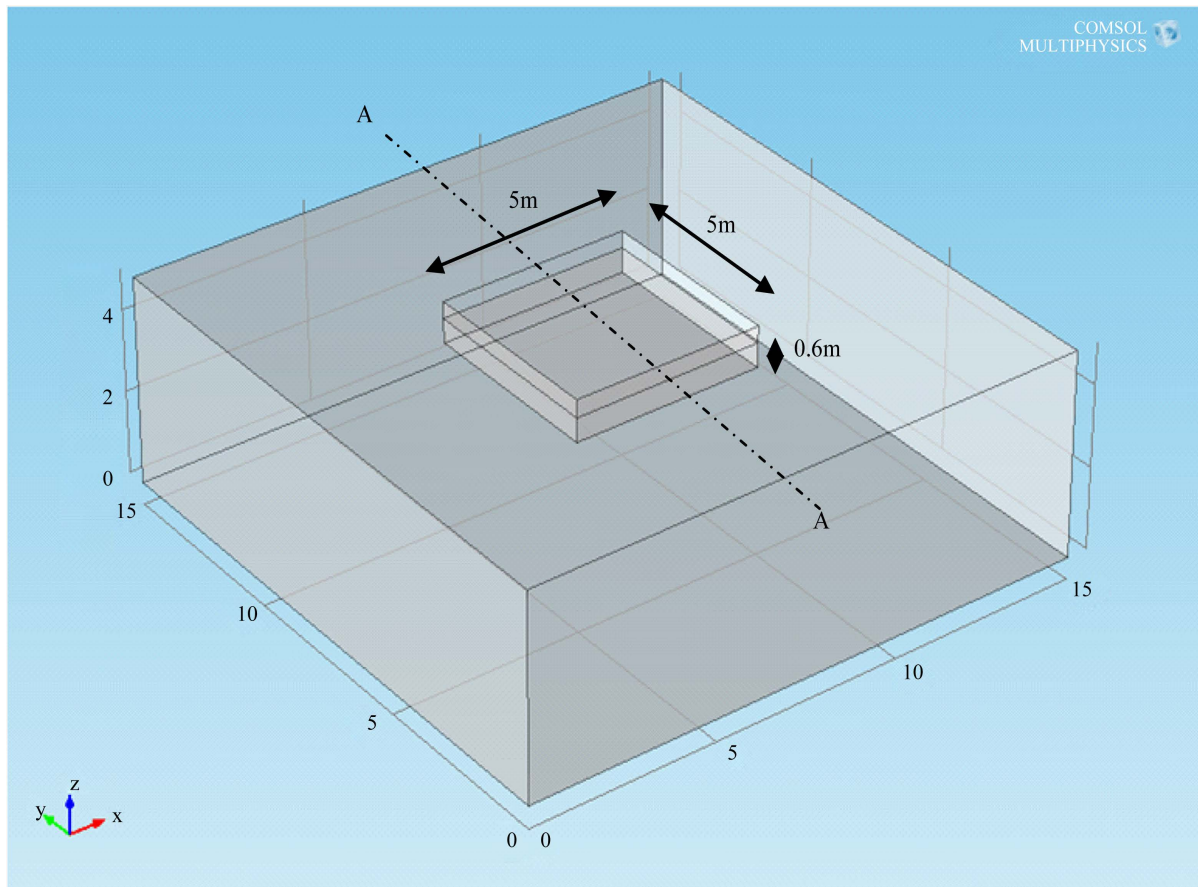
This section of the paper is divided into three sub-sections. Sub-section 4.1 describes the 3D model development in COMSOL Multiphysics, Sub-section 4.2 briefly summarizes the hyetograph to represent the 3-month average rainfall intensities (ARIs), which was used to simulate the developed model, and Sub-section 4.3 discusses the results.

### 4.1. Model Development in COMSOL Multiphysics

Modeling of a soak-away rain garden involves flow in variably saturated porous media. Flow in variably saturated porous media is modeled using the Earth Sciences Module (Subsurface Flow Module) of COMSOL Multiphysics. A review of COMSOL's Earth Sciences Module for simulating flow in variably saturated porous media can be found in [9]. Using the Earth Sciences Module of COMSOL, to develop a model in variably saturated porous media, among other things, it is required to define the model geometry, the mathematical representation of the physical processes of interest, the initial/boundary conditions, and the water balance of the soak-away rain garden.

#### 4.1.1. Model Geometry Using COMSOL Multiphysics

**Figure 1** shows the 3D geometry of a soak-away rain garden that allows stormwater runoff to ex-filtrate into the surrounding soil. For demonstration purpose, a soak-away rain garden with a width of 5 m and a height of 1.0 m



**Figure 1.** Graphical representation of a soak-away rain garden in COMSOL Multiphysics.

is chosen. The length of the soak-away rain garden is 5 m. The shape of the soak-away rain garden is assumed to be of rectangular/square shape. The height of the filter media, the primary soil layer, is 0.6 m. The height of the ponding space, which is defined above the filter media, is limited to 0.2 m. The width and length of the area of the influence is 15 m. Using COMSOL Multiphysics, the 3D geometry of the soak-away rain garden of the above mentioned dimensions was formed by first creating a block feature, whose block name is “blk1”, to represent the outer dimensions. The width and the length of this block were set to “15” and the height was set to “5”. Similar to block feature, “blk1”, another block feature, whose block name is “blk2”, was created. The width and the length of this block were set to “5” and the height was set to “1”, but the block position was set to (5, 5, 4) in (X, Y, Z) direction. To represent the *in-situ* soil, the block feature, “blk2”, was subtracted from the block feature, “blk1”, by creating a Difference feature (“dif1”), a Boolean operation with the ‘input’ properties set to “blk1” and “blk2”. To represent the filter media, similar to block features “blk1” and “blk2”, another block feature, whose block name is “blk3”, was created. The width and the length of this block were set to “5” and the height was set to “0.6”, but the block position was set to (5, 5, 4) in (X, Y, Z) direction. To represent the soak-away rain garden, the block feature “blk3” was unioned with the Difference feature (“dif1”), by creating a Union feature (“uni1”), a Boolean operation with the ‘input’ properties set to “blk3” and “dif1”.

#### 4.1.2. Representation of the Physical Processes Using COMSOL Multiphysics

Richards’ equation models flow in variably saturated porous media. Many efforts to simplify and improve the modeling of flow in variably saturated media have produced a number of variants of Richards’ equation. In this paper, the form of the Richards’ equation adopted in COMSOL Multiphysics is used [6] [7]. The Richards’ equation was applied for both the *in-situ* soil and the filter media of the soak-away rain garden. The adopted equation is presented in Equation (1).

$$\rho \left( \frac{C_m}{\rho g} + S_e S \right) \frac{\partial p}{\partial t} + \nabla \cdot \rho \left( -\frac{k_s}{\mu} k_r (\nabla p + \rho g \nabla D) \right) = Q_m \quad (1)$$

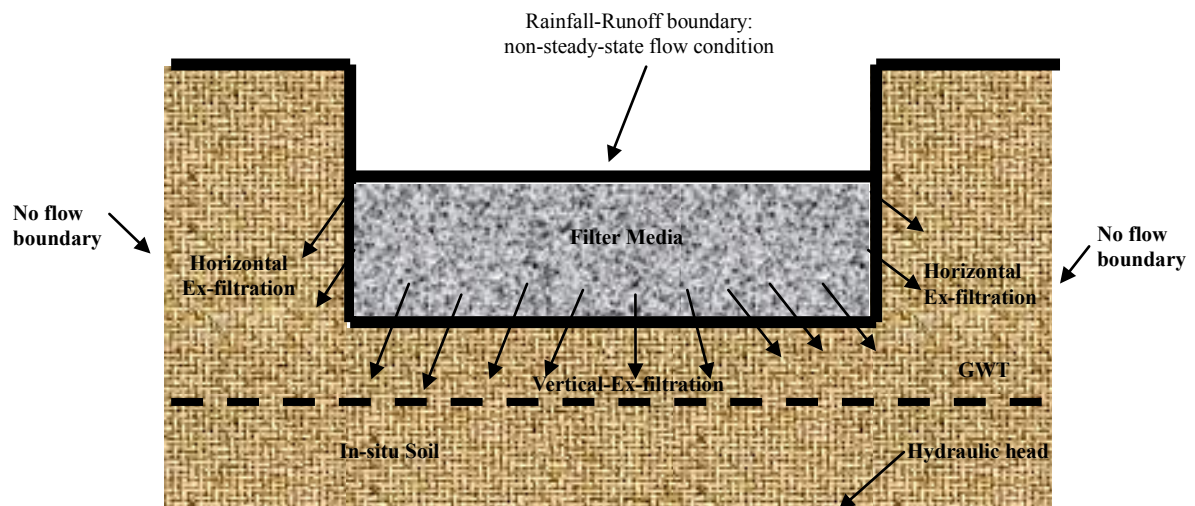
where the pressure,  $p$ , is the dependent variable. In this equation,  $C_m$  represents the specific moisture capacity,  $S_e$  denotes the effective saturation,  $S$  is the storage coefficient,  $\kappa_s$  gives the hydraulic permeability,  $\mu$  is the fluid dynamic viscosity,  $k_r$  denotes the relative permeability,  $\rho$  is the fluid density,  $g$  is acceleration of gravity,  $D$  represents the elevation, and  $Q_m$  is the fluid source (positive) or sink (negative). The fluid velocity across the faces of an infinitesimally small surface is given by Equation (2).

$$u = -\frac{k_s}{\mu} k_r (\nabla p + \rho g \nabla D) \quad (2)$$

where  $u$  is the flux vector. The porous medium consists of pore space, fluids, and solids, but only the liquids move. Equation (2) describes the flux as distributed across a representative surface. To characterize the fluid velocity in the pores, COMSOL Multiphysics also divides  $u$  by the volume liquid fraction,  $\theta_s$ . This interstitial, pore or average linear velocity is  $u_a = u/\theta_s$  [6] [7].

#### 4.1.3. Initial/Boundary Conditions Using COMSOL Multiphysics

To solve flow in variably saturated porous media, it is necessary that appropriate boundary conditions are specified. From a mathematical standpoint, the application of boundary conditions ensures that the solutions to the problems are self-consistent. In this study, the following boundary conditions are identified as appropriate. As shown in **Figure 2**, which represents the frontal view of cut-plane (YZ Plane) A-A of **Figure 1**, the top surface of the rain garden is a rainfall-runoff boundary, a non-steady-state flow condition typical of urban stormwater runoff. The external side boundaries do not allow water to flow in or out of the area of influence, implying that the chosen area is large enough that it does not affect the flow performance around the rain garden. The bottom boundary of the area of influence is specified by a hydraulic head corresponding to an assumed groundwater table level. When water starts to pond, the boundary condition at the top surface of the rain garden becomes a hydraulic boundary. Therefore, there is a need to be able to switch the top surface of the rain garden from flow to hydraulic head. The switching was done using COMSOL Java API, a Java-based interface [6] [7]. The initial condition was set to hydrostatic condition. In other words, above groundwater table, the suction is equal to the distance above groundwater table.



**Figure 2.** Boundary conditions of a soak-away rain garden in COMSOL Multiphysics.

#### 4.1.4. Water Balance of the Soak-Away Rain Garden Using COMSOL Multiphysics

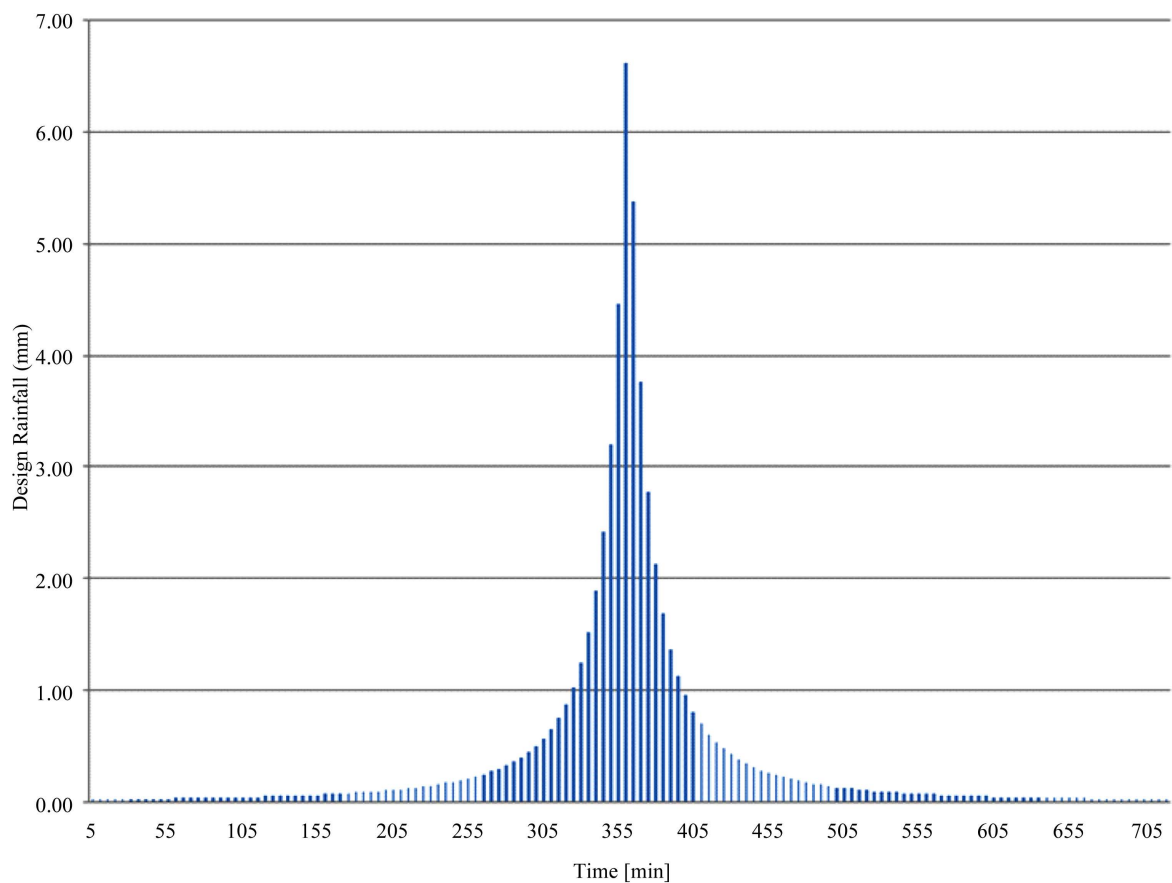
Water Balance of the soak-away rain garden was carried out using a continuity equation whose components are stormwater runoff, overflow (water that is in excess of the ponding space), change of soil moisture within the filter media, change of water storage within the ponding space, vertical ex-filtration, and horizontal ex-filtration. At a given time, horizontal ex-filtration and vertical ex-filtration were computed by integrating the model computed velocity along the four side walls and the bottom surface of the soak-away rain garden, respectively.

#### 4.2. Development of Design Hyetograph

The design of soak-away rain gardens involves water quality. Thus, the establishment of a design hyetograph for the design of soak-away rain gardens, specifically, requires data on intensity-duration-frequency (IDF) values for relatively frequent storms such as 3-month ARIs that carry up to 90% of the total load on annual basis. As underscored in the literature, to date, there are few methods available for the establishment of design hyetographs using IDF data [10]. In this paper, the alternating block method [10], which represents an event of a selected return period both for the selected duration of the event and for any period within this selected duration, is used in developing a design hyetograph from an IDF relationship of Singapore. Storm duration of 720 min was considered. Considering an event of 720 min of the 3-month ARIs, a design hyetograph for 3-month ARIs built-up using this method represents a 3-month ARI event both for the 720 min total duration and for any period (*i.e.*, 5 min, 10 min, 15 min, 30 min, 60 min, ..., 360 min) within this duration centered on the maximum block [10]. The design hyetograph produced by this method specifies the rainfall depth occurring in  $n$  successive time intervals of duration  $\Delta t$  over a total duration of 720 min =  $n\Delta t$ . Duration  $\Delta t$  is often determined by the finest resolution of the hydrological model that is used to generate the design hydrograph, the time distribution of discharge. The hyetograph to represent a 3-month ARI event of 720 min duration is shown in **Figure 3** for a duration ( $\Delta t$ ) of 6 min which is the finest resolution of MUSIC (Model for urban stormwater improvement conceptualization) model which was used to generate the hydrographs for different urbanized (impervious percentage of 90% was assumed) catchment sizes varied from 100 to 250 m<sup>2</sup>.

#### 4.3. Discussion of Results

In this study, to obtain the values for dependent variables (overflow volume, horizontal flow coefficient, and average vertical ex-filtration rate), the simulation was carried out for different values of independent variables (saturated hydraulic conductivity of the *in-situ* soil, surface area of the soak-away rain garden as a % of catchment area, saturated hydraulic conductivity of the filter media, and depth to groundwater table measured from bottom of the filter media). The *in-situ* hydraulic conductivity was varied from 10 mm/hr to 50 mm/hr, typical range in Singapore. The surface area of the soak-away rain garden (as a % of catchment area) was varied from 6% to 15%. The width and the length of the soak-away rain garden were assumed to be of the same size. The saturated



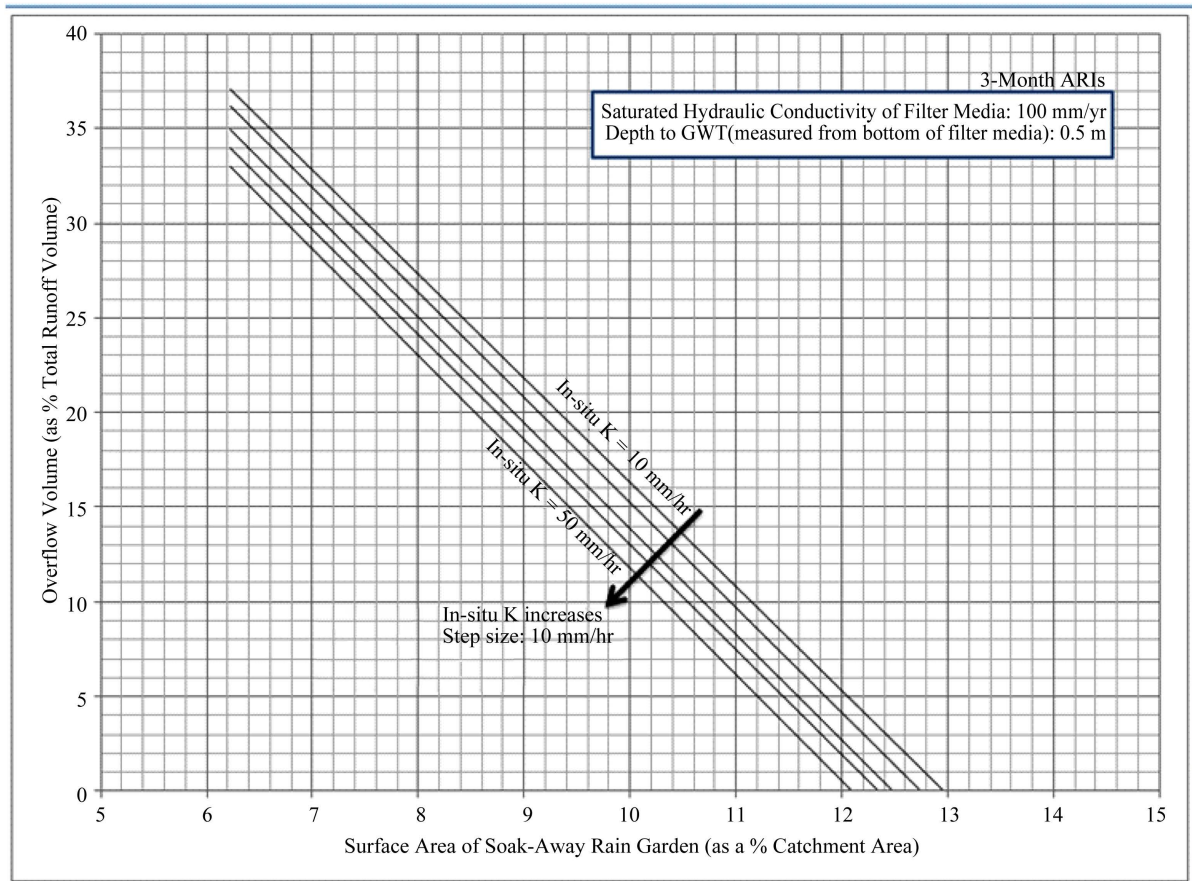
**Figure 3.** Design hyetograph for 3-month ARIs.

hydraulic conductivity of the filter media was varied from 100 mm/hr to 200 mm/hr, typical range in Singapore. The depth to groundwater table measured from bottom of the filter media was varied from 0.5 m to 1.5 m. Having obtained the values for dependent variables for different values of independent variables, the multiple regression equations were fitted for each of the dependent variables.

#### 4.3.1. Multiple Regression Equation of Overflow Volume (as a % Total Runoff Volume)

To understand the relationship between the dependent variable, overflow volume, and the independent variables, as shown in **Figure 4**, the graph of overflow volume (total volume of water for the simulation period = 720 min in excess of ponding space) versus the surface area of the soak-away rain garden (as a % of catchment area) was plotted. The overflow volume is expressed as a % of total runoff volume. For this graph, the saturated hydraulic conductivity of the filter media and the depth to groundwater table were set to 100 mm/hr and 0.5 m, respectively. The graph also shows the variation with the saturated hydraulic conductivities of the *in-situ* soil. The saturated hydraulic conductivity of the *in-situ* soil varies from 10 mm/hr to 50 mm/hr. As can be observed from the graph, there exists a linear relationship between the dependent variable and the independent variables. This same behaviour was observed with the other graphs, which account for the variation in depth to groundwater table and the saturated hydraulic conductivity of the filter media, as well. Therefore, it was decided to fit a multiple linear regression. In the subsequent paragraphs, the statistical measures of the regression are discussed.

As shown in **Table 1**, before fitting a multiple regression, the simulated results for all the considered ranges of potential hydrologic conditions or independent variables (e.g., size of the soak-away rain garden, saturated hydraulic conductivity of the *in-situ* soil, saturated hydraulic conductivity of the filter media) were analyzed to develop a correlation matrix to detect if the problem of multicollinearity exists. The problem of multicollinearity, which can have impact on the quality and stability of the fitted regression model, occurs when some of the independent variables are highly correlated.



**Figure 4.** Overflow volume (as a % total runoff volume) for a depth to groundwater table of 0.5 m and a saturated hydraulic conductivity of filter media of 100 mm/hr.

**Table 1.** Correlation matrix of the regression for overflow volume data.

	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
Y	1.00E+00				
X <sub>1</sub>	-2.49E-01	1.00E+00			
X <sub>2</sub>	-1.25E-01	1.37E-17	1.00E+00		
X <sub>3</sub>	2.84E-01	8.63E-03	1.63E-17	1.00E+00	
X <sub>4</sub>	-7.71E-01	-1.73E-01	-1.61E-17	2.16E-01	1.00E+00

In **Table 1**, dependent variable, “Y”, is the overflow volume which is expressed as a % of total runoff volume. X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, and X<sub>4</sub>, which are the independent variables of the regressed equation, are saturated hydraulic conductivity of the filter media, saturated hydraulic conductivity of the *in-situ* soil, depth to groundwater table measured from bottom of the filter media, and size of the soak-away rain garden (as a % of catchment area), respectively. The correlation matrix shows that there is essentially no correlation between the independent variables and dependent variable, and between the independent variables themselves. Thus, as the problem of multicollinearity does not exist, the regression relationship was further analyzed. As placed in **Table 2**, the sample multiple correlation is 0.996. It has a very strong positive correlation. **Table 2** also shows the coefficient of determination, which in this case is 0.992, indicating that the proportion of the variation in Y that is due to independent variables in the sample is 0.992. In other words, there is very little unexplained (residual) variation whose mean square error is only 1.182, as shown in **Table 3**.

**Table 2.** Regression statistics of the regression for overflow volume data.

Regression Statistics	
Multiple R	0.996
R Square	0.992
Adjusted R Square	0.992
Standard Error	1.087
Observations	230

**Table 3.** ANOVA of the regression for overflow volume data.

	df	SS	MS	F	Significance F
Regression	4	32536.598	8134.150	6881.361	6.3702E-234
Residual	225	265.962	1.182		
Total	229	32802.561			

On the other hand, the test of significance is placed in **Table 3**. If the independent variables explain very little of the Y, dependent variable, the F-value should have been 1 or less. However, the fitted regression is very significant as the F-value is 6881.361. In the fitted regression, the explained variation due to regression and the independent variables is 6881.361 times greater than the unexplained variation. Furthermore, if the selected independent variables and Y are unrelated, the probability of getting the sample evidence is more extreme and is 6.3702E-234. In other words, it is almost impossible to get this kind of data as a result of chance and thus the null hypothesis is rejected. Thus, it is concluded that there is a relationship between these variables in the population.

The “weight” given to each independent variable in the fitted regression is presented in **Table 4**. In a multiple linear regression, the value of the coefficient for each independent variable gives the size of the effect that variable is having on the dependent variable, and the sign on the coefficient (positive or negative) gives the direction of the effect. In the fitted regression, except the independent variable,  $X_3$ , depth to groundwater table measured from bottom of the filter media, the considered independent variables have a negative effect on dependent variable, overflow volume which is expressed as a % of total runoff volume. When these values of coefficients are placed in the regression equation, the equation for predicting the overflow volume, which is expressed as a % of total runoff volume, from the considered independent variables, is  $Y = 74.441 - 0.122X_1 - 0.105X_2 + 14.672X_3 - 5.224X_4$ .

The t-test results to test the regression for significance are also presented in **Table 4**. As can be observed from **Table 4**, the t-values for independent variables are very significant. The probabilities of getting of these magnitudes if the null hypothesis for the test is true are very small as given by p-values. Thus, it is almost impossible to get this kind of data as a result of chance and thus the null hypothesis is rejected. Therefore, it can be concluded that fitted regression between these variables is significant.

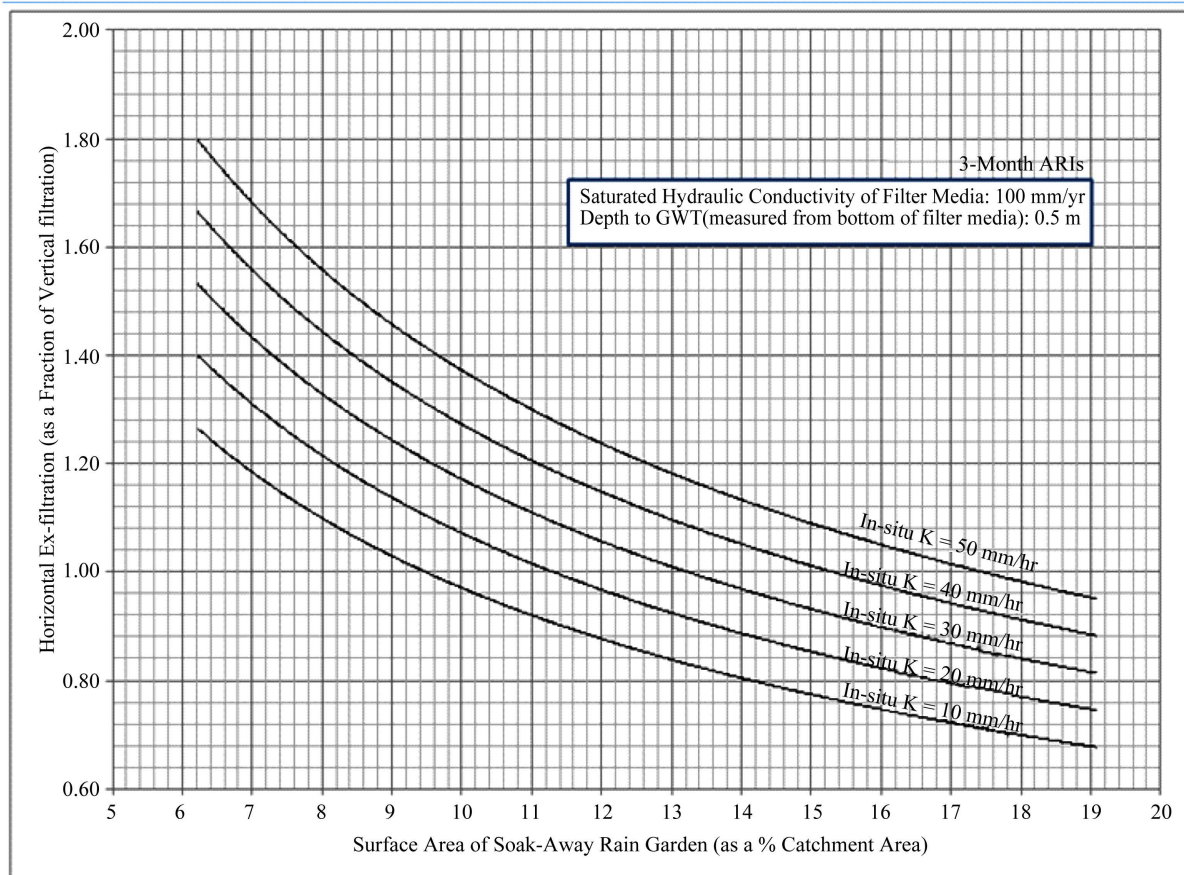
#### 4.3.2. Multiple Regression Equation of Horizontal Flow Coefficient

To understand the relationship between the dependent variable, horizontal flow coefficient, and the independent variables, as shown in **Figure 5**, the graph of horizontal flow coefficient versus the surface area of the soak-away rain garden (as a % of catchment area) was plotted. Horizontal flow coefficient is defined as the ratio between total horizontal ex-filtration (drained through sides of the soak-away rain garden, summed over the simulation period = 720 min, and expressed in  $m^3$ ) and total vertical ex-filtration (drained through bottom of the soak-away rain garden, summed over the simulation period, and expressed in  $m^3$ ). For this graph, the saturated hydraulic conductivity of the filter media was set to 100 mm/hr and the depth to groundwater table measured from bottom of the filter media was set to 0.5 m. The graph also shows the variation with the *in-situ* hydraulic conductivities. As can be observed from the graph, there exists a non-linear (power:  $Y = aX^b$ ) relationship between the dependent variable and the independent variables. This same behaviour was observed with the other



**Table 4.** Coefficients of the regression for overflow volume data.

	Coefficients	Standard Error	t Stat	p-Value
Intercept	74.441	0.466	159.720	1.46E-233
X <sub>1</sub>	-0.122	0.002	-68.358	1.70E-152
X <sub>2</sub>	-0.105	0.005	-20.807	2.47E-54
X <sub>3</sub>	14.672	0.183	80.119	2.14E-167
X <sub>4</sub>	-5.224	0.034	-152.012	8.95E-229

**Figure 5.** Horizontal flow coefficient for a depth to groundwater table of 0.5 m and a saturated hydraulic conductivity of filter media of 100 mm/hr.

graphs, which account for the variation in depth to groundwater table and the saturated hydraulic conductivity of the filter media, as well. Thus, it was required to transform the data using a log-log transform. Having transformed the data, a multiple linear regression to find the best-fit line for the transformed data was obtained. In the subsequent paragraphs, the statistical measures of the regression are discussed.

**Table 5** shows the correlation matrix. As shown in **Table 5**, the dependent variable, “Y”, is the log of horizontal flow coefficient. X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, and X<sub>4</sub>, which are the independent variables of the regressed equation, are the logs of saturated hydraulic conductivity of the filter media, saturated hydraulic conductivity of the *in-situ* soil, depth to groundwater table measured from bottom of the filter media, and size of the soak-away rain garden (as % of catchment area), respectively. The correlation matrix shows that there is essentially no correlation between the independent variables and dependent variable, and between the independent variables themselves. Thus, as the problem of multicollinearity does not exist, the regression relationship was further analyzed.

As shown in **Table 6**, the coefficient of determination is 0.986 and thus there is a very little unexplained variation. Furthermore, as shown in **Table 7**, the fitted regression is very significant and the explained variation due to regression and the independent variables is 5498.676 times greater than the unexplained variation. Furthermore, if the selected independent variables and Y are unrelated, the probability of getting the sample evidence is more extreme and is 2.2218E-286. Thus, it was concluded that there is a relationship between these variables in the population.

The “weight” given to each independent variable in the fitted regression is presented in **Table 8**. When these values are placed in the regression equation, the equation for predicting log of horizontal flow coefficient, from logs of the considered dependent variables, is  $Y = 1.318 - 0.236X_1 + 0.160X_2 + 1.569X_3 - 0.552X_4$ . The t-test results to test the regression for significance are also presented in **Table 8**. As can be observed from **Table 8**, the t-values for independent variables are very significant. The probabilities of getting of these magnitudes if the null hypothesis for the test is true are very small as given by p-values. Thus, it is almost impossible to get this

**Table 5.** Correlation matrix of the regression for horizontal flow coefficient.

	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
Y	1.00E+00				
X <sub>1</sub>	-8.86E-02	1.00E+00			
X <sub>2</sub>	1.20E-01	-9.68E-17	1.00E+00		
X <sub>3</sub>	9.44E-01	-7.11E-16	-5.07E-17	1.00E+00	
X <sub>4</sub>	-2.78E-01	-3.33E-18	2.22E-18	-7.97E-03	1.00E+00

**Table 6.** Regression statistics of the regression for horizontal flow coefficient.

Regression Statistics	
Multiple R	0.993
R Square	0.986
Adjusted R Square	0.986
Standard Error	0.039
Observations	315

**Table 7.** ANOVA of the regression for horizontal flow coefficient.

	df	SS	MS	F	Significance F
Regression	4	33.487	8.372	5498.676	2.2218E-286
Residual	310	0.472	0.002		
Total	314	33.959			

**Table 8.** Coefficients of the regression for horizontal flow coefficient.

	Coefficients	Standard Error	t Stat	p-Value
Intercept	1.318	0.043	30.824	2.07E-96
X <sub>1</sub>	-0.236	0.018	-13.232	5.43E-32
X <sub>2</sub>	0.160	0.009	17.946	7.00E-50
X <sub>3</sub>	1.569	0.011	140.620	7.55E-283
X <sub>4</sub>	-0.552	0.014	-40.392	1.56E-125

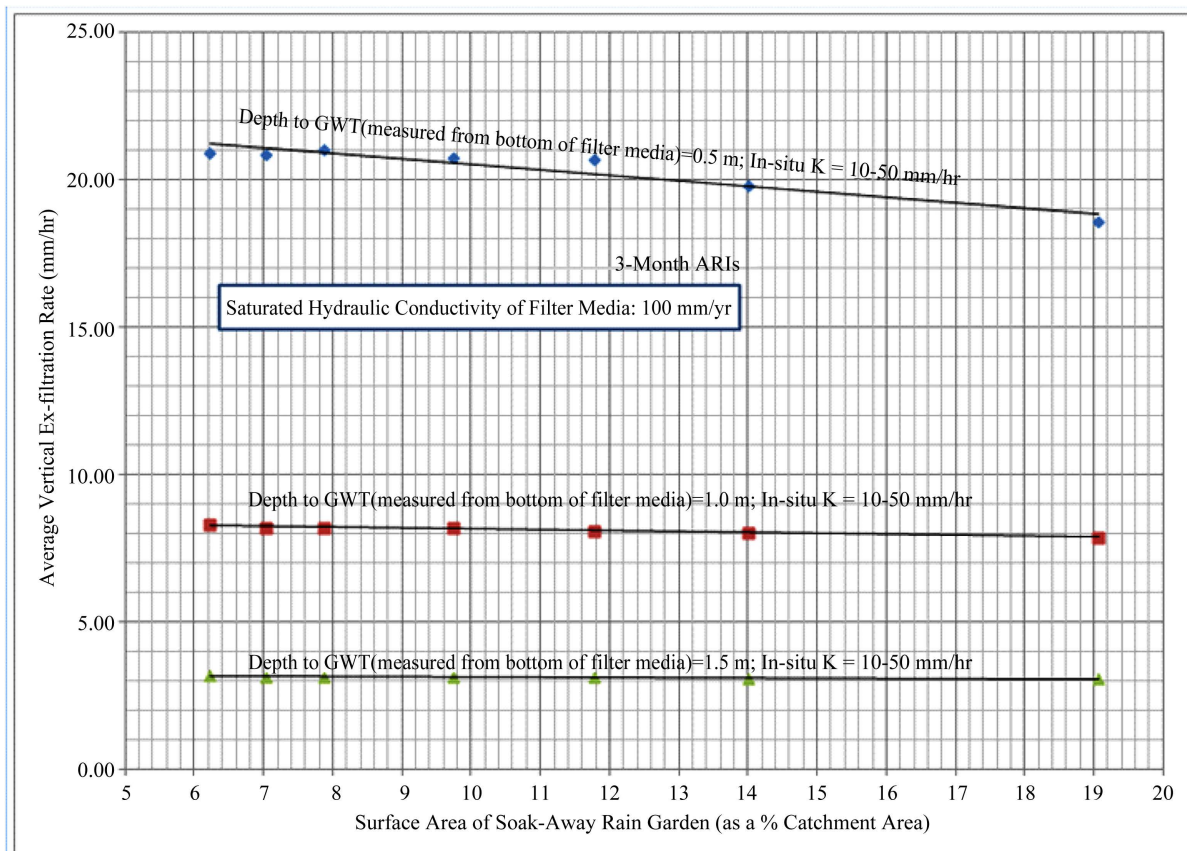
kind of data as a result of chance and thus the null hypothesis is rejected. Therefore, it can be concluded that fitted regression between these variables is significant.

#### 4.3.3. Multiple Regression Equation of Average Vertical Ex-Filtration Rate

To understand the relationship between the dependent variable, average vertical ex-filtration rate, and the independent variables, as shown in **Figure 6**, the graph of average vertical ex-filtration rate was plotted. The average vertical ex-filtration rate is obtained by dividing the average vertical ex-filtration (drained through bottom of the soak-away rain garden, averaged over the simulation period, and expressed in  $\text{m}^3$ ) by the surface area of the soak-away rain garden and the simulation time step. For this graph, the saturated hydraulic conductivity of the filter media was set to 100 mm/hr. The graph also shows the variation with the *in-situ* hydraulic conductivities. The saturated hydraulic conductivity of the *in-situ* soil varies from 10 mm/hr to 50 mm/hr.

As can be observed from the graph, there exists a linear relationship between the dependent variable and the independent variables. This same behaviour was observed with the other graphs, which account for the variation in saturated hydraulic conductivity of the filter media. Therefore, it was decided to fit a multiple linear regression. In the subsequent paragraphs, the statistical measures of the regression are discussed. As shown in **Table 9**, by performing a normal multiple regression using all the independent variables, it was found that  $X_2$ , saturated hydraulic conductivity of the *in-situ* soil, and  $X_4$ , surface area of the soak-away rain garden (as % of catchment area) have high p-values (p-values > significance level,  $\alpha = 0.01$ ) even though the “Significance F” is very low. Thus, forward stepwise regression was used to develop the best model that contains some of the independent variables.

With forward stepwise regression, as shown in **Table 10**, the correlation matrix shows that there is essentially no correlation between the independent variables and dependent variable, and between the independent variables themselves. Thus, as the problem of multicollinearity does not exist, the regression relationship was further analyzed.



**Figure 6.** Average vertical ex-filtration rate for a saturated hydraulic conductivity of filter media of 100 mm/hr.

As shown in **Table 11**, the coefficient of determination is 0.911 and thus there is a very little unexplained variation. Furthermore, as shown in **Table 12**, the fitted regression is very significant and the explained variation due to regression and the independent variables is 1591.399 times greater than the unexplained. If the selected independent variables and Y are unrelated, the probability of getting the sample evidence is more extreme and is 2.062E-164. Thus, it was concluded that there is a relationship between these variables in the population.

The “weight” given to each independent variable in the fitted regression is presented in **Table 13**. When these values are placed in the regression equation, the equation for predicting the average vertical ex-filtration rate from the considered independent variables, is  $Y = 26.682 + 0.077X_1 - 23.483X_3$ .

**Table 9.** Coefficients of the regression for average vertical Ex-filtration rate with all the independent variables.

	df	SS	MS	F	Significance F
Regression	4	32103.210	8025.802	810.448	9.8951E-163
Residual	310	3069.905	9.903		
Total	314	35173.115			

	Coefficients	Standard Error	t Stat	p-Value
Intercept	28.170	0.993	28.370	3.71E-88
X <sub>1</sub>	0.077	0.004	17.640	1.04E-48
X <sub>2</sub>	-0.013	0.013	-1.041	0.298
X <sub>3</sub>	-23.483	0.434	-54.070	6.90E-160
X <sub>4</sub>	-0.103	0.042	-2.450	0.015

**Table 10.** Correlation matrix of the regression for average vertical ex-filtration rate.

	Y	X <sub>1</sub>	X <sub>3</sub>
Y	1.00E+00		
X <sub>1</sub>	2.96E-01	1.00E+00	
X <sub>3</sub>	-9.07E-01	0.00E+00	1.00E+00

**Table 11.** Regression statistics of the regression for average vertical ex-filtration rate.

Regression Statistics	
Multiple R	0.954
R Square	0.911
Adjusted R Square	0.910
Standard Error	3.172
Observations	315

**Table 12.** ANOVA of the regression for average vertical ex-filtration rate.

	df	SS	MS	F	Significance F
Regression	2	32033.015	16016.508	1591.399	2.062E-164
Residual	312	3140.100	10.064		
Total	314	35173.115			

**Table 13.** Coefficients of the regression for average vertical ex-filtration rate.

	Coefficients	Standard Error	t Stat	p-Value
Intercept	26.682	0.809	32.969	1.1246E-103
X <sub>1</sub>	0.077	0.004	17.498	3.01733E-48
X <sub>3</sub>	-23.483	0.438	-53.634	1.5956E-159

The t-test results to test the regression for significance are also presented in **Table 13**. As can be observed from **Table 13**, the t-values for independent variables are very significant. The probabilities of getting of these magnitudes if the null hypothesis for the test is true are very small as given by p-values. Thus, it is almost impossible to get this kind of data as a result of chance and thus the null hypothesis is rejected. Therefore, it can be concluded that fitted regression between these variables is significant.

## 5. Conclusions

To have a rapid assessment of soak-away rain gardens on a range of potential hydrologic conditions, with a design hyetograph of 3-month average rainfall intensities of Singapore, this study establishes easy to use multiple regression equations of soak-away rain gardens, which are specific for local conditions, by developing a mathematical model based on Richard's equation using COMSOL Multiphysics, a finite element analysis and solver software package for various physics and engineering applications. The outcome of this study, among other things, shows the followings:

1) The regression equation on overflow volume (total volume of water for the simulation period = 720 min in excess of ponding space), which is expressed as a % of total runoff volume, with the independent variables—saturated hydraulic conductivity of the filter media (X<sub>1</sub>), saturated hydraulic conductivity of the *in-situ* soil (X<sub>2</sub>), depth to groundwater table measured from bottom of the filter media (X<sub>3</sub>), and surface area of the soak-away rain garden as a % of catchment area (X<sub>4</sub>), shows a very strong relationship (coefficient of determination = 0.992). The fitted regression is  $Y = 74.441 - 0.122X_1 - 0.105X_2 + 14.672X_3 - 5.224X_4$ .

2) The regression equation on log of horizontal flow coefficient which is defined as the ratio between total horizontal ex-filtration (drained through sides of the soak-away rain garden, summed over the simulation period = 720 min, and expressed in m<sup>3</sup>) and total vertical ex-filtration (drained through bottom of the soak-away rain garden, summed over the simulation period, and expressed in m<sup>3</sup>), with the independent variables (logs of X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, and X<sub>4</sub>), shows a very strong relationship (coefficient of determination = 0.986). The fitted regression is  $Y = 1.318 - 0.236X_1 + 0.160X_2 + 1.569X_3 - 0.552X_4$ .

3) The regression equation on average vertical ex-filtration rate that is obtained by dividing the average vertical ex-filtration (drained through bottom of the soak-away rain garden, averaged over the simulation period, and expressed in m<sup>3</sup>) by the surface area of the soak-away rain garden and the simulation time step has high p-values (p-values > significance level,  $\alpha = 0.01$ ) for saturated hydraulic conductivity of the *in-situ* soil (X<sub>2</sub>) and surface area of the soak-away rain garden as a % of catchment area (X<sub>4</sub>).

4) The regression equation on average vertical ex-filtration rate with saturated hydraulic conductivity of the filter media (X<sub>1</sub>) and depth to groundwater table (X<sub>3</sub>) shows a very strong relationship (coefficient of determination = 0.911). The fitted regression is  $Y = 26.682 + 0.077X_1 - 23.483X_3$ .

## Acknowledgements

The authors would like to thank PUB, Singapore's national water agency, for providing the financial support to conduct this research. The first author, who was a researcher at National University of Singapore, is currently attached to South Eastern University of Sri Lanka.

## References

- [1] Allan, P.D., Robert, G.T. and William, F.H. (2010) Improving Urban Stormwater Quality: Applying Fundamental Principles. *Journal of Contemporary Water Research and Education*, **146**, 3-10.  
<http://dx.doi.org/10.1111/j.1936-704X.2010.00387.x>

- [2] Jia, L., David, J.S., Cameron, B. and Yuntao, G. (2014) Review and Research Needs of Bioretention Used for the Treatment of Urban Stormwater. *Water*, **6**, 1069-1099. <http://dx.doi.org/10.3390/w6041069>
- [3] Hunt, W.F., Jarrett, A.R., Smith, J.T. and Sharkey, L.J. (2006) Evaluating Bioretention Hydrology and Nutrient Removal at Three Field Sites in North Carolina. *Journal of Irrigation and Drainage Engineering*, **132**, 600-608. [http://dx.doi.org/10.1061/\(ASCE\)0733-9437\(2006\)132:6\(600\)](http://dx.doi.org/10.1061/(ASCE)0733-9437(2006)132:6(600))
- [4] Jones, M.P. and Hunt, W.F. (2009) Bioretention Impact on Runoff Temperature in Trout Sensitive Waters. *ASCE Journal of Environmental Engineering*, **135**, 577-585. [http://dx.doi.org/10.1061/\(ASCE\)EE.1943-7870.0000022](http://dx.doi.org/10.1061/(ASCE)EE.1943-7870.0000022)
- [5] Li, H., Sharkey, L.J., Hunt, W.F. and Davis, A.P. (2009) Mitigation of Impervious Surface Hydrology Using Bioretention in North Carolina and Maryland. *ASCE Journal of Hydrologic Engineering*, **14**, 407-415. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2009\)14:4\(407\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2009)14:4(407))
- [6] COMSOL AB (2012) COMSOL Multiphysics User's Guide (Version 4.3). Stockholm.
- [7] COMSOL AB (2012) COMSOL Multiphysics Reference Guide (Version 4.3). Stockholm.
- [8] Richards, L.A. (1931) Capillary Conduction of Liquids through Porous Mediums. *Journal of Applied Physics*, **1**, 318-333. <http://dx.doi.org/10.1063/1.1745010>
- [9] Li, Q., Ito, K., Wu, Z., Lowry, C.S. and Loheide II, S.P. (2009) COMSOL Multiphysics: A Novel Approach to Ground Water Modeling. *Groundwater*, **47**, 480-487. <http://dx.doi.org/10.1111/j.1745-6584.2009.00584.x>
- [10] Chow, V.T., Maidment, D.R. and Mays, L.W. (1988) Applied Hydrology. McGraw Hill, New York.