Development of a Modular Streamflow Model to Quantify Runoff Contributions from Different Land Uses in Tropical Urban Environments Using Genetic Programming

Ali Meshgi[1*], Petra Schmitter[1,3], Ting Fong May Chui[2], Vladan Babovic[1]

[1]*Department of Civil and Environmental Engineering, National University of Singapore, Block E1A, #07-03, No 1 Engineering Drive 2, Singapore*

[2]*Department of Civil Engineering, The University of Hong Kong, Room 6-18A, Haking Wong Building, Pokfulam, Hong Kong*

[3]*Present address: The International Water Management Institute, East Africa & Nile Basin Office, Addis Ababa, Ethiopia*

[*]*Corresponding author; Tel: +65 6516 3650, email: alimeshgi@u.nus.edu*

Abstract

The decrease of pervious areas during urbanization has severely altered the hydrological cycle, diminishing infiltration and therefore sub-surface flows during rainfall events, and further increasing peak discharges in urban drainage infrastructure. Designing appropriate waster sensitive infrastructure that reduces peak discharges requires a better understanding of land use specific contributions towards surface and sub-surface processes. However, to date, such understanding in tropical urban environments is still limited. On the other hand, the rainfall-runoff process in tropical urban systems experiences a high degree of non-linearity and heterogeneity. Therefore, this study used Genetic Programming to establish a physically interpretable modular model consisting of two sub-models: (i) a baseflow module and (ii) a quick flow module to simulate the two hydrograph flow components. The relationship between the input variables in the model (i.e. meteorological data and catchment initial conditions) and its overall structure can be explained in terms of catchment hydrological processes. Therefore, the model is a partial greying of what is often a black-box approach in catchment modelling. The model was further generalized to the sub-catchments of the main catchment, extending the potential for more widespread applications. Subsequently, this study used the modular model to predict both flow components of events as well as time series, and applied optimization techniques to estimate the contributions of various land uses (i.e. impervious, steep grassland, grassland on mild slope, mixed grasses and trees and relatively natural vegetation) towards baseflow and quickflow in tropical urban systems. The sub-catchment containing the highest portion

of impervious surfaces (40% of the area) contributed the least towards the baseflow (6.3%) while the sub-catchment covered with 87% of relatively natural vegetation contributed the most (34.9%). The results from the quickflow module revealed average runoff coefficients between 0.12 and 0.80 for the various land uses and decreased from impervious (0.80), grass on steep slopes (0.56), grass on mild slopes (0.48), mixed grasses and trees (0.42) to relatively natural vegetation (0.12). The established modular model, reflecting the driving hydrological processes, enables the quantification of land use specific contributions towards the baseflow and quickflow components. This quantification facilitates the integration of water sensitive urban infrastructure for the sustainable development of water in tropical megacities.

Keywords: Genetic programming; Modular approach; Baseflow; Quickflow; Land use contribution; Tropical urban environments

1. INTRODUCTION

Increasing urbanization has severely altered the hydrological cycle in many places worldwide, accelerating runoff due to a decrease of pervious areas and therefore infiltration. In order to efficiently drain the increase in surface runoff, intensive drainage networks are often built to prevent flash floods during heavy storm events (Marshall and Shortle, 2005). However, as cities are dynamically expanding, the continuous increase of impervious surfaces and the accompanied excess runoff often exceeds the present channel capacity resulting in local flash floods. To reduce the impact of surface runoff,

water sensitive urban infrastructure (e.g. green roofs, porous pavement, bioretention ponds, swales) retaining rainfall and enhancing infiltration rates in urban cities are being promoted (Burns et al., 2012; Chang, 2010). Water Sensitive Urban Design (WSUD) is an engineering design approach which aims to minimize hydrological and water quality impact of urban development by integrating land use planning with urban water management (Singh and Kandasamy, 2009). The implementation of such technologies requires a detailed understanding of runoff contributions from each specific land use in order to plan the location of these local source control measures. Therefore, a better understanding is needed regarding rainfall-runoff processes in urbanized areas, including an accurate assessment of contributions from different land uses towards quickflow as well as baseflow. This understanding would be essential for integrated management and sustainable development of water resources particularly in tropical megacities which are dependent on water sources that are more vulnerable to inter-annual fluctuations in precipitation.

Land use and land cover affect catchment hydrology primarily through changes in hydrological processes such as infiltration, rainfall interception, and evapotranspiration (Calder, 1993; Calder, 2005; DeFries and Eshleman, 2004; Potter, 1991; Tran and O'Neill, 2013) which may have significant effects on rainfall-runoff processes and catchment water yields (Roa-García et al., 2011). The various contributions from different land uses towards rainfall-runoff processes have attracted worldwide attention, especially in temperate urban regions (e.g. Burns et al., 2005; Diaz-Palacios-Sisternes et al., 2014; Loperfido et al., 2014; Miller et al., 2014). Comparing runoff generation from

different land uses enables us to understand the rainfall-runoff response influenced by particular catchment components and processes and their contribution towards the overall catchment. This understanding contains valuable information with regards to a physical based understanding of rainfall-runoff behaviour when designing appropriate water management infrastructure in tropical megacities. However, it is interesting to note that a review of the literature shows that to date, no detailed investigation has been done to assess the impact of different land uses on rainfall-runoff processes for tropical urban cities.

To evaluate the impact of different land uses on catchment hydrology, rainfall-runoff processes need to be simulated. There are multiple Rainfall-Runoff (R-R) models available that can be applied to simulate rainfall-runoff processes; each one characterized by a different level of complexity, limitations and data requirement (Sorooshian, 2008). Physically-based models usually incorporate simplified forms of physical laws and are generally non-linear, time-varying and deterministic, with parameters that are representative of watershed characteristics. Although these models enhance our understanding towards the physics of hydrological processes, they require significant computational time and large amounts of data (Beven, 2012; Dye and Croke, 2003). Over the past decades, machine learning tools such as Artificial Neural Network (ANN) (e.g. Jeong and Kim, 2005; Kisi et al., 2013; Sudheer et al., 2002; Talei and Chua, 2012) and Genetic Programming (GP) (e.g. Babovic, 2005; Babovic and Keijzer, 2002; Babovic and Keijzer, 2006) have been used to develop rainfall-runoff models. GP offers advantages over other data driven techniques since it is more likely to generate a

function with understandable structure. However, most data driven models are one unit models with adequate input variables that cover all system processes in one input/output structure (Abrahart and See, 1999; Bowden et al., 2005). Such models combine all the various flow components losing valuable information on their specific contributions which is needed when designing local mitigation measures (Corzo and Solomatine, 2007). In addition, covering all the rainfall-runoff processes in one unit without taking into account the different physically interpretable sub-processes may lead to low accuracy in extrapolation. Streamflow is commonly conceptualized to include baseflow and quickflow (also called direct runoff) components. The baseflow component represents the relatively steady contribution to streamflow from groundwater flow, while the quickflow represents the additional streamflow contributed by surface flows (i.e. rapid runoff) and shallow subsurface flows (delayed runoff) (Beven, 2012). One way of retaining as much information as possible is to build separate models for each of the different physically interpretable flow components leading to a modular approach. As such, a modular model for the simulation of streamflow time series consisting of separate modular units for baseflow and stormwater runoff would be suitable in quantifying both flow components in a more flexible manner. The concept of a modular model has been used in modelling tools that use a linear reservoir approach (e.g. unit hydrograph methods) by splitting streamflow into baseflow and quickflow components. However, these models may fail to represent the nonlinear dynamics in the rainfall-runoff process (Rajurkar et al., 2002).

129 Therefore, this paper used GP to develop a physically interpretable modular
130 universally applicable model accounting for baseflow and quickflow. The modular
131 model was applied to address the following research questions in a tropical urbanized
132 system:

133 • What are the contributions of the various land uses towards quickflow?
134 • How does the baseflow contribution change among sub-catchments with different
135 land uses?
136 • How do runoff generation processes vary among the different types of rainfall
137 events?
138 • What are the effects of antecedent catchment conditions on runoff response?

139 In this paper, a description of the study site as well as monitoring network is
140 described in Section 2. Section 3 focuses on the methodology used to develop the
141 modular model consisting of baseflow and quickflow components using GP. The
142 methodology with regards to the quantification of land use specific contributions to the
143 quickflow component is presented in Section 4. The results are discussed in Section 5
144 and lastly, conclusions are summarized in Section 6.

145

146 **2. STUDY AREA AND DATA COLLECTION**

147 **2.1. Description of the study area**

148 Kent Ridge Catchment, a small catchment (0.085 km^2) within the National
149 University of Singapore (NUS), located in the southern part of Singapore was chosen to

setup an intensive monitoring network (Figure 1). This catchment contains all the main land uses of Singapore and hence is representative from a hydrological point of view. Furthermore, the use of a small catchment reduces data uncertainty and inaccuracy with regards to the spatial distribution of precipitation. The overall topography of the catchment is characterized by steep slopes with elevations ranging between 14.0 m and 75.8 m above sea level.

A land use map of the catchment (Figure 2) was created combining the information from Google Earth, NUS campus map and field observations. The identified land uses (Table 1), typically for Singapore, included impervious surfaces (i.e. roof top, road, and paved car parks), grasses on mild (Figure 3a) and steep slopes (Figure 3b), mixed grasses and trees (Figure 3c) and relatively natural vegetation (Figure 3d). Therefore, understanding the behavior and the mechanism of rainfall-runoff processes at Kent Ridge catchment would yield valuable information for tropical urbanized cities such as Singapore.

Water table in unconfined aquifers is often thought to be a subdued replica of the topography and from the elevation of the water table, flow of groundwater can be approximated (Haitjema and Mitchell-Bruker, 2005). Moreover, it is mostly the unconfined aquifer that contributes to the baseflow of a stream (Rumynin, 2011). Therefore, the sub-surface catchment boundary was assumed to coincide with the surface catchment boundary.

Four soil textural classes can be identified within the catchment: loamy sand (55%), clay loam (33%), sandy loam (9%) and silt loam (2.7%) (Meshgi and Chui, 2014). In

172 total, six sub-catchments were delineated based on the Digital Elevation Map (DEM) as
173 well as their drainage location (see Section 2.2) within the network (Figure 1).

174 The pattern of rainfall varies over the year due to the two monsoons: the northeast
175 (mid-November to early March) and the southwest monsoon (mid-June to September).
176 Moderate to heavily rainfall events to intense thunderstorm activity are typically
177 observed in the monsoon period while long shower events interrupted by thunderstorms
178 occur in the inter-monsoon period. According to the weather station maintained by the
179 Department of Geography at NUS, located nearby the study area, the mean annual
180 precipitation from 2004 until 2013 is 2500 mm and the mean daily temperature varies
181 between 23.9°C and 32.3°C.

182

183 **2.2 Hydro-meteorological data**

184 One rainfall monitoring station was installed within the Kent Ridge catchment
185 (Figure 1) and operated from September 2011 to August 2012 and January to June 2013
186 recording precipitation at a one-minute resolution with an accuracy of 0.2 mm.
187 Meteorological data from the NUS Geography weather station were used to estimate
188 potential evapotranspiration from physically-based Penman-Monteith.

189 Five water level measurement stations (Figure 1) recorded at the same temporal
190 resolution during the same period as the rain gage. Measured water levels were
191 converted into discharge using standard stage-discharge relationships for the control
192 structures (Bos, 1989). Drainage areas of the five discharge monitoring locations are
193 presented in Table 1. Sub-catchment 1 and 2 drain into Stations A and B, respectively,

while Station C measured discharges from sub-catchment 3. Stations A and B together with the discharge draining from sub-catchment 4 are recorded by Station D. The outlet (Station E) receives the flows from the upstream Stations C and D as well as those from sub-catchment 5. Gap filling and data quality assessments were performed for all five discharge stations using the Aquarius software (Aquatic Informatics Inc., 2009).

The pressure head for this study was obtained by using the validated numerical HYDRUS-3D groundwater model (Meshgi et al., 2014) for the entire period (September 2011 to June 2013). For more information on the calibration and validation procedure of HYDRUS-3D for the Kent Ridge Catchment, readers are referred to Meshgi et al. (2014).

3. DEVELOPMENT OF A MODULAR MODEL USING GP

A modular model for simulating streamflow can be defined as:

$$Q_{Total(t)} = Q_{BF(t)} + Q_{QF(t)} \quad (1)$$

where, $Q_{Total(t)}$ is streamflow (L^3/T), $Q_{BF(t)}$ is baseflow (L^3/T), $Q_{QF(t)}$ is quickflow (L^3/T). In Meshgi et al., (2014) a generalized empirical equation has been derived to estimate baseflow time series using GP with minimal data requirements and preservation of physical catchment information. According to Meshgi et al. (2014), differences between baseflow time series simulated by a groundwater numerical model (i.e. HYDRUS-3D) and the baseflow module were minimal, indicating that the baseflow module can accurately estimate baseflow time series. For detailed information about the baseflow module, reference is made to Meshgi et al. (2014). In the present study, the second

modular unit was developed to simulate quickflow while the baseflow module was used as the first unit. Subtracting the predicted baseflow from the measured discharge resulted in the quickflow component. This component was taken as the target variable (i.e. output) in the GP software called GPKERNEL (Babovic and Keijzer, 2000) to develop the second modular unit. For a detailed description of GP, reference is made to Meshgi et al. (2014). The modular approach was further generalized to approximate streamflow in other sub-catchments within Kent Ridge Catchment.

3.1 Separating baseflow from total discharge

The quickflow time series, needed for the derivation of the empirical equation using GP, was obtained by separating baseflow from total discharge using the baseflow separation technique proposed by Meshgi et al. (2014) as follows:

$$Q_{BF(t)} = Q_{Bmin} + \sqrt{b\,A}\,\Delta h^2_{p(t+k)} \quad (2)$$

where $Q_{BF(t)}$ is the daily baseflow volume (m^3), Q_{Bmin} is minimum daily baseflow of the entire period (m^3), A is the total un-paved surface area in the catchment (m^2), $\Delta h_{p(t)}$ presents the normalized daily average of pressure head within a piezometer ($\Delta h_{p(t)} = h_{(t)} - h_{min}$ in which $h_{(t)}$ is the daily average of pressure head (m) and h_{min} is the minimum daily average of the pressure head (m) over the entire period), b is dimensionless coefficient (-) and k is the lag time between the rainfall events and groundwater table responses (day). For this catchment, the lag time was found to be 0 (Meshgi et al., 2014).

3.2 Approximating the quickflow component

The subtraction of the estimated baseflow (Section 3.1) from the total discharge at Station E resulted in the quickflow component. GPKERNEL was set up to relate the quickflow component with hydro-meteorological and catchment variables including five-minute rainfall intensity, daily evapotranspiration data and simulated pressure head (Meshgi et al., 2014), and area of the catchment

Data from September 2011 until August 2012 was used for model development, while data from January to June 2013 was used for model validation. Moreover, to evaluate the performance of the modular model, six events with different rainfall characteristics (i.e. seasonal variability, total rainfall depth, duration and shape of hydrograph) and antecedent catchment conditions were selected within the period of September 2011 to June 2013 (Table 2). Performance of the established equation was evaluated using three statistical error functions: Relative Root Mean Squared Error (RRMSE), Correlation Coefficient (CC) and the Nash–Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970).

3.3 Generalization of Modular Model

The modular model derived for the outlet (Station E, Figure 1) needed to be generalized in order to simulate both hydrograph flow components for the four sub-catchments (Stations A to D) in Section 4.2. For the baseflow component the generalized equation developed by Meshgi et al. (2014) was used. The generalization of the quickflow

component was done through the combination of GA with an Interior Point Algorithm (IPA) (a local search method) in MATLAB. Combining both methods can improve performance by using the good global property of random searching and the convergence rate of a local method (Grosan and Abraham, 2007). In this optimization procedure, GA was first used to obtain a global optimization providing the global optimal solution. The global optimal solution was subsequently fed into IPA for a local search to achieve the improved results. The parameter settings for the procedure are given in Table 3.

4 QUANTIFICATION OF QUICKFLOW CONTRIBUTIONS FROM SPECIFIC LAND USES

4.1 Clustering analysis

In a tropical area, catchment responses to the rainfall events are expected to vary significantly from event to event due to different types of rainfall events and antecedent catchment conditions (Peng and Wang, 2012). Therefore, rainfall events were divided into clusters and sub-clusters based on types of rainfall events and antecedent catchment conditions using a statistical hierarchical clustering technique proposed by Ward (1963).

The following variables were used: total precipitation in the event, maximum 30-min intensity and duration. This resulted in a total of 150 events grouped into four clusters (Table 4). Rainfall Cluster I represents rainfall events which are less intensive than other clusters. Rainfall Cluster II includes rainfall events with moderate rainfall depth, intensity and duration while Rainfall Cluster III consists of storms that have high

rainfall depth, intensity and duration. Rainfall Cluster IV represents extreme rainfall storms with very high rainfall depth and intensity. Most rainfall events were categorized into Rainfall Cluster I with 102 events while Rainfall Cluster IV only contained 10 events. In addition, events in Rainfall Cluster III and II occurred 21 and 17 times, respectively.

The sub-clusters contained the various antecedent catchment conditions. As the spatio-temporal variations of the antecedent soil moisture data are often not available, the antecedent baseflow derived using the baseflow module (Equation 2) was used to present the catchment state prior to the event for the entire period, resulting in three sub-clusters (Table 4). Sub-cluster one contained events with low antecedent baseflow between 0.98 and 2.4 L/s, events with moderate antecedent baseflow between 2.41 and 3.83 L/s were grouped in Sub-Cluster-2 while events with high antecedent baseflow between 3.84 and 5.26 L/s were classified in Sub-Cluster-3.

4.2 Land use specific runoff coefficient

This section derived an approach to estimate land use specific runoff coefficients (i.e. the portion of rainfall contributing to quickflow) during an event. For each station, the relation between the weighted average runoff coefficient and runoff coefficient of each particular land use was derived for a given event:

$$\begin{bmatrix} C_{T_{St1_i}}.A_{T_{Stn\,1}} \\ C_{T_{St2_i}}.A_{T_{Stn\,2}} \\ C_{T_{St3_i}}.A_{T_{Stn\,3}} \\ C_{T_{St4_i}}.A_{T_{Stn\,4}} \\ C_{T_{St5_i}}.A_{T_{Stn\,5}} \end{bmatrix} = [C_{IMPi}\ C_{GMi}\ C_{GSi}\ C_{GUTi}\ C_{Bi}] \begin{bmatrix} A_{IMP_{Stn\,1}} & A_{IMP_{Stn\,2}} & A_{IMP_{Stn\,3}} & A_{IMP_{Stn\,4}} & A_{IMP_{Stn\,5}} \\ A_{GM_{Stn\,1}} & A_{GM_{Stn\,2}} & A_{GM_{Stn\,3}} & A_{GM_{Stn\,4}} & A_{GM_{Stn\,5}} \\ A_{GS_{Stn\,1}} & A_{GS_{Stn\,2}} & A_{GS_{Stn\,3}} & A_{GS_{Stn\,4}} & A_{GS_{Stn\,5}} \\ A_{GT_{Stn\,1}} & A_{GT_{Stn\,2}} & A_{GT_{Stn\,3}} & A_{GT_{Stn\,4}} & A_{GT_{Stn\,5}} \\ A_{B_{Stn\,1}} & A_{B_{Stn\,2}} & A_{B_{Stn\,3}} & A_{B_{Stn\,4}} & A_{B_{Stn\,5}} \end{bmatrix} \quad (3)$$

where, i is an event, C_T is the weighted average runoff coefficient (-), C_{IMP}, C_{GM}, C_{GS}, C_{GUT}, C_B are the runoff coefficients of total impervious, grass on mild slope, grass on steep slope, mixed grasses and trees and relatively natural vegetation areas, A_T is the total area (m^2), A_{IMP}, A_{GM}, A_{GS}, A_{GUT}, A_B are the areas of impervious, grass on mild slope, grass on steep slope, mixed grasses and trees and relatively natural vegetation (m^2), respectively (Table 1).

The weighted average runoff coefficient in Equation 3 can be calculated for each station according to:

$$C_{T_i} = \frac{Q_R}{P.A} \quad (4)$$

where C_{T_i} presents the weighted average runoff coefficient for event i, Q_R and P are total runoff volume obtained from the quickflow module (m^3) and total precipitation depth (m) of a given event, respectively, and A is the area of a catchment/sub-catchment (m^2).

Hybrid GA was used to optimize the parameter values for land use specific runoff coefficients (C_{IMP}, C_{GM}, C_{GS}, C_{GUT}, C_B) in Equation 3 using the Optimization Tool in MATLAB. The objective function of the optimization processes was defined as reducing the Root Mean Squared Error (RMSE). The following constraints were set based on the physical meaning of runoff coefficient:

$0 \leq C_{GM}, C_{GS}, C_{GUT}, C_B \leq C_{IMP} \leq 1.$

4.3 Estimating total contribution of different land uses towards the quickflow component

To evaluate the contribution of various land uses towards the quickflow component, the runoff volume generated by each land use was calculated at catchment level. Total contributions of each land use were normalized as follows:

$$C_{N_j} = \frac{\frac{C_{T_j}}{A_j}}{S}, \qquad j = IMP, GS, GM, GT, B \tag{5}$$

and,

$$S = \frac{C_{T_{IMP}}}{A_{IMP}} + \frac{C_{T_{GS}}}{A_{GS}} + \frac{C_{T_{GM}}}{A_{GM}} + \frac{C_{T_{GT}}}{A_{GT}} + \frac{C_{T_B}}{A_B} \tag{6}$$

where, C_N, is the normalized contribution and C_T is the total contribution of each land use, IMP, GS, GM, B, and GT represent impervious, grass on steep slope, grass on mild slope, relatively natural vegetation, mixed grasses and trees, respectively.

5 RESULTS AND DISCUSSION

5.1 Development of a Modular Model

5.1.1 Approximating quickflow time series

The following empirical equation, based on the quickflow time series filtered from the observed streamflow data using Equation (1), with a minimal RMSE was selected:

$$\begin{aligned} \mathrm{Q}_{QF(t)} &= 0.5 \times 10^{-3} \mathrm{A} \left(\Delta \mathrm{h}_p \times \mathrm{P}_T\right)^{0.25} \left[0.65\mathrm{P}_{(t-5)} + 0.35\mathrm{P}_{(t-10)}\right] \\ &+ 0.17 \times 10^{-3} \mathrm{A} \left(\Delta \mathrm{h}_p \times \mathrm{P}_T\right)^{0.25} \begin{bmatrix} 0.07\mathrm{P}_{(t-15)} + 0.09\mathrm{P}_{(t-20)} + 0.11\mathrm{P}_{(t-25)} + \\ 0.61\mathrm{P}_{(t-30)} + 0.12\mathrm{P}_{(t-35)} \end{bmatrix}^{0.5} \end{aligned} \tag{7}$$

where $\mathrm{Q_{QF(t)}}$ presents the quickflow (L/s), $\mathrm{P_{(t-L)}}$ is the rainfall intensity (mm/min) with L being minutes of lag time, $\mathrm{P_T}$ is the total rainfall depth during the event (mm), $\Delta \mathrm{h_p}$ is

the normalized daily averaged pressure head prior to the event and A is the total area of the catchment (m^2).

In both training and testing periods, model performance parameters (i.e. error criteria) suggested there are only minimal differences between the filtered quickflow from the observed discharge data and those obtained by Equation 7, at Station E (Table 5).

From all the input variables and parameters, the total rainfall depth, rainfall intensity, pressure head and the total catchment area were selected by GP. The first term of the empirical equation is the rapid runoff component corresponding to quickflow, while the second term approximates the delayed runoff component as the lag time increases. Both terms include the total catchment area (A), rainfall (e.g. $0.65P_{(t-5)} + 0.35P_{(t-10)}$) and antecedent catchment condition (Δh_p). The effect of evapotranspiration on antecedent catchment condition can also be captured by groundwater table fluctuations (Δh_p) (Carlson Mazur et al., 2014) which is likely the reason why evapotranspiration itself was not selected by GP. In this equation, the term $\Delta h_p . P_T$ allows for variations in the portion of rainfall that contributes to the runoff component for various events; higher Δh_p (i.e. relatively saturated conditions) and higher P_T (i.e. heavy rainfall events) yield higher runoff volume. The structure of empirical equation also indicates quickflow is almost equally responsive to changes in the total rainfall depth during the event (P_T) or antecedent catchment condition (Δh_p).

As such, a modular model for simulating streamflow at the outlet (i.e., Station E) is the combination of the baseflow module from Meshgi et al (2014) and the above derived Q_{QF} component:

$$Q_{Total(t)} = Q_{BF(t)} + Q_{QF(t)}$$

and:

$$\begin{cases} Q_{BF(t)} = Q_{Bmin} + \sqrt{0.29\,A}\,\Delta h_{p(t)}^{2} \\ Q_{QF(t)} = 0.5 \times 10^{-3} A \left(\Delta h_p \times P_T\right)^{0.25} \left[0.65 P_{(t-5)} + 0.35 P_{(t-10)}\right] \\ \quad + 0.17 \times 10^{-3} A \left(\Delta h_p \times P_T\right)^{0.25} \left[\begin{matrix} 0.07 P_{(t-15)} + 0.09 P_{(t-20)} + 0.11 P_{(t-25)} \\ +0.61 P_{(t-30)} + 0.12 P_{(t-35)} \end{matrix}\right]^{0.5} \end{cases} \quad (8)$$

Based on the performance indicators the quickflow module is able to predict the rapid and delayed runoff at the outlet with a correlation coefficient of 0.98, an NSE of 0.96 and a RRMSE of 0.65 (Table 5).

Combining the baseflow and quickflow module resulted in the overall prediction of the discharge at the outlet (Station E). A correlation coefficient of 0.97 was obtained between the observed streamflow and those simulated by the modular model (Figure 4). The RRMSE and NSE were 0.7 and 0.94, respectively. Six events were analyzed in detail for the outlet station (Table 2). According to these results, differences between observed and estimated streamflow were found to be minimal, confirming that the modular model can successfully estimate streamflow in rainfall events with different characteristics. The rapid runoff contributed on average 66% for the six events while the contribution of the delayed runoff and baseflow components were on average 26% and 8%, respectively. Figure 5 presents the different hydrograph components (i.e. rapid

runoff, delayed runoff and baseflow) estimated by the modular model for the 21/11/2013 event (Table 2). The time of concentration is very short and the rapid runoff component has a steep rising and falling limb, dominating the total runoff hydrograph during the rainfall event. This reflects the hydrological characteristics of the basin: slopes are steep, infiltration is low through which the rainfall results in a rapid quick flow component, being unable to recharge the groundwater in such a short-time period.

5.1.2 Generalization of the modular model

The modular model derived for the catchment outlet was generalized to simulate the hydrograph flow components at sub-catchment level through the generalization of the quickflow module:

$$\begin{aligned} Q_{QF(t)} &= 10^{-3} A\, a\left(\Delta h_p \times P_T\right)^r \left[j_1 P_{(t)} + \cdots + j_n P_{(t-n)}\right] \\ &\quad + 10^{-3} A\, c\left(\Delta h_p \times P_T\right)^r \left[d_1 P_{(t)} + \cdots + d_n P_{(t-n)}\right]^{0.5} \end{aligned} \tag{9}$$

where "a" , "c", "j_n", "d_n" and "r" are the dimensionless parameters.

As such, a generalization of the modular model is the combination of the generalized baseflow module (Meshgi et al. 2014) and Equation 9:

$$\begin{cases} Q_{BF(t)} = Q_{Bmin} + \sqrt{b\,A}\,\Delta h^2_{p(t+k)} \\ Q_{QF(t)} = 10^{-3} A \left[\begin{array}{l} a\left(\Delta h_p \times P_T\right)^r \left[\sum_{i=1}^{n} j_i P_{(t-i+1)}\right] \\ + c\left(\Delta h_p \times P_T\right)^r \left[\sum_{i=1}^{n} d_i P_{(t-i+1)}\right]^{0.5} \end{array} \right] \end{cases} \tag{10}$$

and $Q_{Total(t)} = Q_{BF(t)} + Q_{QF(t)}$.

387 Details regarding the estimation of model parameters for the baseflow module can
388 be found in Meshgi et al. (2014).

389 The estimation of model parameters for the generalized Q_{QF} module was first
390 performed for Station E using the hybrid GA. In order to prevent parameters from
391 taking unrealistic values, the following constraints, based on the physical meaning of
392 the model's parameters, were set:

393 $\sum_{i=1}^{n} \mathrm{j}_i = \sum_{i=1}^{n} \mathrm{d}_i = 1$,

394 $a, c, \mathrm{j}_i, \mathrm{d}_i \geq 0$.

395 The model parameters $(r, a, c, \mathrm{j}_i, \mathrm{d}_i)$ optimized with hybrid GA for Station E were
396 same as those obtained with GP in Equation (8), confirming that the proposed hybrid
397 GA is an appropriate method for estimating model parameters of the quickflow module.
398 The model performance indicators (i.e. NSE, CC and RRMSE) between observed
399 quickflow in Station A to D and those estimated by the empirical equation showed a
400 similar accuracy as those obtained for Station E (Table 5). These results demonstrate the
401 successful prediction of the rapid and delayed runoff using the quickflow module
402 independent on sub-catchment characteristics and area. The model parameters for the
403 baseflow module (b and k) as well as quickflow module $(r, a, c, \mathrm{j}_i, \mathrm{d}_i)$ for Station A to D
404 are also presented in Table 6. Similar model parameters for the quickflow module were
405 found for Stations C, D, and E likely due to the relative similar land use distributions in
406 those stations (Table 1). The larger parameter 'a' was found for Station A likely due to
407 the larger fraction of impervious surface of the sub-catchment (Table 1). In contrast, the
408 lowest value was observed at Station B where relatively natural vegetation areas were

409 the dominant land use. Parameter ‘d’ represents the starting and ending point of delayed
410 runoff in quickflow module. As can be seen in Table 6, delayed runoff began between
411 15 to 20 minutes after the start of a rainfall event in all the monitoring stations.
412
413 **5.2 Quantifying quickflow contributions from different land use types**
414 **5.2.1 Event-based land use specific runoff coefficients**
415 The average runoff coefficients of different land uses towards the predicted
416 quickflow for each cluster and sub-cluster were obtained with hybrid GA using the
417 runoff module (Table 7). Comparison of the average runoff coefficient for all events
418 belonging to one sub-cluster using Equation (3) and those estimated by Equation (4)
419 (Table 7) demonstrates the successful estimation of land use specific runoff coefficients.
420 The small standard deviation of relative absolute errors (Equation 4 estimates relative to
421 Equation 3 values) suggests that the average runoff coefficients were estimated with low
422 uncertainty.
423 Results indicated that land uses exert a major influence on runoff coefficients of an
424 urban tropical environment. Similar results have been also reported for urban temperate
425 systems, indicating that there is a strong positive correlation between the amount of
426 quickflow and the level of urbanization (e.g.Sun et al., 2013). The average runoff
427 coefficient of different land uses decreased from impervious surface (0.8), grass on
428 steep slope (0.56), grass on mild slope (0.48), mixed grasses and trees (0.42) and to
429 relatively natural vegetation (0.12). As expected, impervious surfaces contributed the
430 most to the rapid and delayed runoff among all land uses. In contrast, the lowest runoff

coefficient was found for relatively natural vegetation ranging from 0.04 to 0.24 due to canopy interception and evapotranspiration (Sriwongsitanon and Taesombat, 2011). In addition, larger infiltration in relatively natural vegetation area occurs as a result of extensive root zone development which increases the porosity. Human activities, resulting in soil compaction and subsequently reducing soil porosity and infiltration capacity, in recreational grass areas, play an important role in generating surface runoff (Dadkhah and Gifford, 1980). Additionally, runoff increases with increasing slope gradients, due to decreased infiltration rates (Huang et al., 2013). As such, higher runoff coefficients were observed for the grass areas with steep slopes than those areas with mild slopes that included trees.

With regards to the effect of antecedent catchment conditions, the antecedent soil moisture content had a larger effect on the pervious land uses. Normalized variation in runoff coefficients (with respect to their minimum value within each land use) of different land uses from Cluster-I/Sub-Cluster-1 to Cluster-IV/Sub-Cluster-3 listed in Table 7 are shown in Figure 6. Table 7 shows the increasing trend of runoff coefficients from Cluster-I/Sub-Cluster-1 to Cluster-IV/Sub-Cluster-3 for all the land uses. In addition, Figure 6 shows the largest variation for the runoff coefficients associated with the relatively natural vegetation followed by grass based land uses and impervious surfaces. As the types of rainfall events had the largest effect on relatively natural vegetation areas compared to other land uses, runoff coefficients for relatively natural vegetation fluctuated about 2 to 4 times more compared to those for grass based and impervious surfaces, respectively (Figure 6). This is because rainfall loss due to

evapotranspiration, canopy interception and infiltration, especially during small rainfall events, is typically higher for natural vegetation areas than for non- natural vegetation areas (Sriwongsitanon and Taesombat, 2011). In addition, canopy interception may reduce with increasing rainfall intensity due to splashing of larger raindrops from vegetation (Calder, 2005). This could cause a large variation in runoff coefficient for relatively natural vegetation area from Cluster-I/Sub-Cluster-1 to Cluster-IV/Sub-Cluster-3. On average, the runoff coefficients of all the land uses increased gradually from sub-cluster-1 (relatively un-saturated condition) to sub-cluster-3 (relatively saturated condition) by 17% (Table 7). With regards to the pervious surfaces, this can be explained by the catchment initial conditions. In fact, higher levels of groundwater table and initial soil moisture would reduce the soil water suction and potential (Hawke et al., 2006) which reduces infiltration rate (Philip, 1957) and consequently increases the runoff volume. However, with regards to the impervious surfaces, the runoff coefficients increased slightly probably due to the antecedent precipitation which could increase the initial storage and subsequently lead to the greater runoff coefficient.

The suitability of land use specific runoff coefficients derived in this section for the assessment of runoff generated by an extreme rainfall event (e.g., 10 year ARI) was investigated. According to the Rainfall Intensity-Duration-Frequency (IDF) established for Singapore by Public Utilities Board (PUB) (Code of Practice-Drainage Design and Considerations, 2011), an event with 10 year ARI (128 mm) was monitored during 2010-2011. It should be mentioned that this event was not used during the optimization procedure for quantifying land use contributions towards rapid and delayed runoff

component. Assessment of the runoff generated by this event which would be categorized in Cluster-IV/Sub-Cluster-3 showed that with less than 5% error, the runoff coefficient of Cluster-IV/Sub-Cluster-3 can be used to estimate the total runoff for an extreme rainfall event. This indicated that even for such a rainfall event, the contribution of relatively natural vegetation area is about 4 times smaller than that of impervious surfaces. As such, increasing urban pressure and the related conversion of pervious surfaces to impervious areas clearly influences not only hydrological processes at watershed scale but also increases flood risks tremendously. However, land use conversion due to demographic pressure, frequently inhibits the conservation of forests and natural vegetation. Therefore, it is of uttermost importance to account for water sensitive features in urban cities that have similar properties to natural vegetation in order to restore hydrological processes in tropical urban environments. This could eventually ensure dry season baseflow sustenance as well as modulation of quickflow responses to the extreme rainfall events.

5.2.2 Average runoff coefficients at catchment scale

Average runoff coefficients varied between 0.09 and 0.61 for the various sub-catchments (Figure 7). As expected, the average runoff coefficients among the various types of rainfall events differed significantly ($p < 0.001,\ \alpha = 0.05$) and were in decreasing order of Rainfall Cluster IV>III>II>I. These results showed a consistent positive relationship between types of rainfall events and runoff coefficient (i.e. increasing runoff volume with increasing rainfall depth, duration and intensity). Sub-

Cluster-3 (relatively saturated condition) contributed the most towards the quickflow during rainfall events. Rainfall events in Sub-clusters 3 had a shorter dry antecedent weather period (0.8 days) when compared to sub-clusters 1 and 2 (2.3 and 1.8 days, respectively). As evapotranspiration losses increases with increasing dry weather period, higher antecedent soil moisture was expected in Sub-Cluster-3 as compared to other sub-classes. Therefore, a reduction in the infiltration and thus the water buffering capacity of the soil results in a larger quickflow fraction.

When analyzing the various sub-catchments, larger average runoff coefficients were found for Station A due to the larger fraction of impervious surface of the sub-catchment (Figure 7, Table 1). In contrast, the lowest runoff coefficient was observed at Station B where relatively natural vegetation areas were the dominating. In fact, quickflow was very small for most rainfall events, and large runoff in this station could only be generated by rainfall storms larger than 55 mm (Cluster-IV/Sub-Cluster-3). No significant differences in the average catchment runoff coefficients were observed among Stations C, D, and E ($p = 0.4$, $\alpha = 0.05$) due to the relative similar land use distributions in those stations (Table 1).

5.2.3 Contribution of different types of land use towards overall stormwater runoff

The mean total quickflow of five land uses descended in an order of impervious surfaces, grass on mild slope, relatively natural vegetation, mixed grasses and trees, grass on steep slope (Figure 8a). Although the percentage of area covered by relatively

519 natural vegetation is about 1.7 times larger than that covered by impervious surfaces, the
520 mean total quickflow from impervious surface is approximately 3.4 times greater than
521 from the relatively natural vegetation. As can be seen in Table 1, the areas of different
522 land uses vary largely. Hence, in order to provide a fair comparison, total contributions
523 of land uses on equivalent area basis (i.e. area of each land use is equal) is presented in
524 Figure 8b. The amount of total quickflow on equivalent area basis change to:
525 impervious surfaces > grass on steep slope > grass on mild slope > mixed grasses and
526 trees > relatively natural vegetation. These results showed that impervious surfaces
527 exhibited the greatest quickflow while the average contribution of relatively natural
528 vegetation areas was as low as about 5.4% which was 5.8 times smaller than that of
529 impervious surfaces. The total quickflow on equivalent area basis were similar among
530 the grass based land uses with grass areas on steep slopes being the second largest
531 contributor (23.5%), followed by grass on mild slope (21%) and grass with trees
532 (18.7%). Due to the urbanization effect such as soil compaction, the contribution of
533 impervious surfaces was in average only 1.4 times greater than the grass based land uses
534 (i.e. steep slope, mild slope and underneath trees) contributions. The buffer capacity of
535 the relatively natural vegetation area is large enough to even buffer heavy rainfall
536 events, reducing the quickflow in an urban environment.
537
538 **5.3 Baseflow contributions at catchment scale**
539 Comparison was made between the average baseflow contributions towards the
540 overall discharge at the various stations. The lowest baseflow contribution (6.3%) was

observed at Station A whose drainage area contained 40% of impervious surfaces (Figure 9). In contrast, the highest proportion of baseflow contribution to the streamflow generation (34.9%) was detected at Station B with relatively natural vegetation being the main land use (87% of the total area). As natural vegetation can both increase baseflow and reduce runoff, it plays an important role in catchment water yields, streamflow dynamics and sustainable development of water resources. Similar contributions of baseflow (about 18%) were observed for Stations C, D, and E due to the similar land use composition (Table 1). These results showed a negative relationship between the amount of impervious surfaces and baseflow contributions (i.e. decreasing baseflow contributions with increasing impervious surfaces). Similar results have also been found in some studies indicating that increasing urbanization (i.e. impervious surface) might result in significant loss of groundwater flow contribution in streams due to reduced infiltration (Chang, 2007; Leopold and Geological, 1968; Price, 2011; Rose and Peters, 2001; Simmons and Reynolds, 1982).

6 SUMMARY AND CONCLUSION

Meteorological, physiographic, hydrologic and land use data was used to derive a physically interpretable modular model consisting of a baseflow module and a quickflow module. The structure of the derived modular model, using GP, was simple and physically interpretable. The quickflow module contained a rapid and delayed streamflow generation component which corresponds to the overland flow and shallow sub-surface flow, respectively.

563 The modular model was generalized to predict rapid and delayed runoff at sub-
564 catchment and catchment scales, revealing its potential application for other catchments
565 independent from the prevailing meteorological and catchment condition. In a latter step
566 the model was further validated on its representation of catchment processes through the
567 quantification of land use specific overland flow, shallow sub-surface and baseflow
568 contributions in the tropical urban context. Results from the modular model showed that
569 baseflow contributions decrease with the increase of impervious surfaces, and runoff
570 volume increases with the increase in rainfall depth, duration and intensity. The model
571 results also suggested that both very large and small rainfall events may cause runoff
572 generation processes to be significantly different among different land uses. Even for an
573 extreme rainfall event, the quickflow contribution of relatively natural vegetation areas
574 was about four times less than that of impervious surfaces. As such, the modular model
575 is able to quantify the various hydrograph components in the landscape and could
576 potentially be used in other catchments to simulate the rainfall-runoff processes and also
577 to quantify runoff contributions from different land uses.
578
579 **ACKNOWLEDGMENTS**
580 The authors gratefully acknowledge the support and contributions of Singapore-
581 Delft Water Alliance (SDWA). The research presented in this work is carried out as part
582 of the SDWA's Multi-objective Multiple Reservoir Management research program (R-
583 303-001-005-272). The authors also would like to thank Dr. Abhay Anand for his

assistance in MATLAB coding. They are also very grateful to the reviewers whose insightful comments guided improvements on the original manuscript.

REFERENCES

Abrahart, R.J., See, L., 1999. Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. Hydrology and Earth System Sciences, 6(4): 655-670.

Aquatic Informatics Inc., 2009. Aquarius Hydrologic Workstation Software, Vancouver, Canada.

Babovic, V., 2005. Data mining in hydrology. Hydrological Processes, 19(7): 1511-1515.

Babovic, V., Keijzer, M., 2000. Genetic programming as a model induction engine. Journal of Hydroinformatics, 2(1): 35-60.

Babovic, V., Keijzer, M., 2002. Rainfall runoff modelling based on genetic programming. Nordic Hydrology, 33(5): 331-346.

Babovic, V., Keijzer, M., 2006. Rainfall-Runoff Modeling Based on Genetic Programming, Encyclopedia of Hydrological Sciences. John Wiley & Sons, Ltd.

Beven, K.J., 2012. Rainfall-runoff modelling: the primer. Wiley-Blackwell, Chichester, West Sussex; Hoboken, NJ.

Bos, M.G., 1989. Discharge measurement structures, 20. International Institute for Land Reclamation and Improvement, Wageningen, Netherlands.

Bowden, G.J., Dandy, G.C., Maier, H.R., 2005. Input determination for neural network models in water resources applications. Part 1—background and methodology. Journal of Hydrology, 301(1–4): 75-92.

Burns, D. et al., 2005. Effects of suburban development on runoff generation in the Croton River basin, New York, USA. Journal of Hydrology, 311(1): 266-281.

Burns, M.J., Fletcher, T.D., Walsh, C.J., Ladson, A.R., Hatt, B.E., 2012. Hydrologic shortcomings of conventional urban stormwater management and opportunities for reform. Landscape and Urban Planning, 105(3): 230-240.

Calder, I.R., 1993. Hydrologic Effects of Land-use Change. In: Maidment, D.R. (Ed.), Handbook of Hydrology, McGraw Hill, New York, pp. 13.1-13.5.

Calder, I.R., 2005. Blue revolution: integrated land and water resources management. Earthscan, Sterling, VA.

Carlson Mazur, M.L., Wiley, M.J., Wilcox, D.A., 2014. Estimating evapotranspiration and groundwater flow from water-table fluctuations for a general wetland scenario. Ecohydrology, 7(2): 378-390.

Chang, H., 2007. Comparative streamflow characteristics in urbanizing basins in the Portland Metropolitan Area, Oregon, USA. Hydrological Processes, 21(2): 211-222.

Chang, N.-B., 2010. Hydrological Connections between Low-Impact Development, Watershed Best Management Practices, and Sustainable Development. Journal of Hydrologic Engineering, 15(6): 384-385.

Code of Practice-Drainage Design and Considerations, 2011. http://www.pub.gov.sg/general/code/Pages/SurfaceDrainagePart2-7.aspx.

Corzo, G., Solomatine, D., 2007. Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. Hydrological Sciences Journal, 52(3): 491-507.

Dadkhah, M., Gifford, G.F., 1980. Influence of Vegetation, Rock Cover, and Trampling on Infiltration Rates and Sediment Production. Water Resources Bulletin, 16(6): 979-979.

DeFries, R., Eshleman, K.N., 2004. Land-use change and hydrologic processes: a major focus for the future. Hydrological Processes, 18(11): 2183-2186.

630 Diaz-Palacios-Sisternes, S., Ayuga, F., Garcia, A.I., 2014. A method for detecting and describing land use
631 transformations: An examination of Madrid's southern urban-rural gradient between 1990 and
632 2006. Cities, 40: 99-110.
633 Dye, P.J., Croke, B.F.W., 2003. Evaluation of streamflow predictions by the IHACRES rainfall-runoff
634 model in two South African catchments. Environmental Modelling & Software, 18(8–9): 705-
635 712.
636 Grosan, C., Abraham, A., 2007. Hybrid Evolutionary Algorithms: Methodologies, Architectures, and
637 Reviews. In: Abraham, A., Grosan, C., Ishibuchi, H. (Eds.), Hybrid Evolutionary Algorithms.
638 Studies in Computational Intelligence. Springer Berlin Heidelberg, pp. 1-17.
639 Haitjema, H.M., Mitchell-Bruker, S., 2005. Are Water Tables a Subdued Replica of the Topography?
640 Ground Water, 43(6): 781-786.
641 Hawke, R.M., Price, A.G., Bryan, R.B., 2006. The effect of initial soil water content and rainfall intensity
642 on near-surface soil hydrologic conductivity: A laboratory investigation. CATENA, 65(3): 237-
643 246.
644 Huang, J., Wu, P., Zhao, X., 2013. Effects of rainfall intensity, underlying surface and slope gradient on
645 soil infiltration under simulated rainfall experiments. CATENA, 104: 93-102.
646 Jeong, D.-I., Kim, Y.-O., 2005. Rainfall-runoff models using artificial neural networks for ensemble
647 streamflow prediction. Hydrological Processes, 19(19): 3819-3835.
648 Kisi, O., Shiri, J., Tombul, M., 2013. Modeling rainfall-runoff process using soft computing techniques.
649 Computers & Geosciences, 51: 108-117.
650 Leopold, L.B., Geological, S., 1968. Hydrology for urban land planning - a guidebook on the hydrologic
651 effects of urban land use, 554. United States Government Printing Office, Washington.
652 Loperfido, J.V., Noe, G.B., Jarnagin, S.T., Hogan, D.M., 2014. Effects of distributed and centralized
653 stormwater best management practices and land cover on urban stream hydrology at the
654 catchment scale. Journal of Hydrology, 519: 2584-2595.
655 Marshall, E., Shortle, J., 2005. Urban development impacts on ecosystems. In: Goetz, S., Shortle, J.,
656 Bergstrom, J. (Eds.), In: *Land Use Problems and Conflicts: Causes, Consequences and*
657 *Solutions*. Routledge, Taylor and Francis group, New york.
658 Meshgi, A., Chui, T.F.M., 2014. Analysing tension infiltrometer data from sloped surface using two-
659 dimensional approximation. Hydrological Processes, 28(3): 744-752.
660 Meshgi, A., Schmitter, P., Babovic, V., Chui, T.F.M., 2014. An empirical method for approximating
661 stream baseflow time series using groundwater table fluctuations. Journal of Hydrology, 519,
662 Part A: 1031-1041.
663 Miller, J.D. et al., 2014. Assessing the impact of urbanization on storm runoff in a peri-urban catchment
664 using historical change in impervious cover. Journal of Hydrology, 515: 59-70.
665 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion
666 of principles. Journal of Hydrology, 10(3): 282-290.
667 Peng, T., Wang, S.-j., 2012. Effects of land use, land cover and rainfall regimes on the surface runoff and
668 soil loss on karst slopes in southwest China. CATENA, 90: 53-62.
669 Philip, J.R., 1957. The theory of infiltration: 5. The influence of the initial moisture content. Soil Science,
670 84(4): 329-340.
671 Potter, K.W., 1991. Hydrological impacts of changing land management practices in a moderate-sized
672 agricultural catchment. Water Resources Research, 27(5): 845-855.
673 Price, K., 2011. Effects of watershed topography, soils, land use, and climate on baseflow hydrology in
674 humid regions: A review. Progress in Physical Geography, 35(4): 465-492.
675 Rajurkar, M.P., Kothyari, U.C., Chaube, U.C., 2002. Artificial neural networks for daily rainfall—runoff
676 modelling. Hydrological Sciences Journal, 47(6): 865-877.
677 Roa-García, M.C., Brown, S., Schreier, H., Lavkulich, L.M., 2011. The role of land use and soils in
678 regulating water flow in small headwater catchments of the Andes. Water Resources Research,
679 47(5): W05510.
680 Rose, S., Peters, N.E., 2001. Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a
681 comparative hydrological approach. Hydrological Processes, 15(8): 1441-1457.

Rumynin, V.G., 2011. Subsurface solute transport models and case histories: with applications to radionuclide migration, 25. Springer, Dordrecht; New York.
Simmons, D.L., Reynolds, R.J., 1982. Effects of urbanization on base flow of selected south-shore streams, Long Island, New York. JAWRA Journal of the American Water Resources Association, 18(5): 797-805.
Singh, G., Kandasamy, J., 2009. Evaluating performance and effectiveness of water sensitive urban design. Desalination and Water Treatment, 11(1-3): 144-150.
Sorooshian, S., Hsu, K., Coppola, E.,Tomassetti, B., Verdecchia, M., Visconti, G., 2008. Hydrological Modelling and the Water Cycle Coupling the Atmospheric and Hydrological Models. In: V.P. Singh (Ed.). Springer, Texas A&M University, College Station, U.S.A.
Sriwongsitanon, N., Taesombat, W., 2011. Effects of land cover on runoff coefficient. Journal of Hydrology, 410(3-4): 226-238.
Sudheer, K.P., Gosain, A.K., Ramasastri, K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrological Processes, 16(6): 1325-1330.
Sun, Z., Li, X., Fu, W., Li, Y., Tang, D., 2013. Long-term effects of land use/land cover change on surface runoff in urban areas of Beijing, China. APPRES, 8(1): 084596-084596.
Talei, A., Chua, L.H.C., 2012. Influence of lag time on event-based rainfall–runoff modeling using the data driven approach. Journal of Hydrology, 438–439: 223-233.
Tran, L.T., O'Neill, R.V., 2013. Detecting the effects of land use/land cover on mean annual streamflow in the Upper Mississippi River Basin, USA. Journal of Hydrology, 499: 82-90.
Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 58(301): 236-244.

706 Table 1: Drainage areas and their respective land uses as delineated by the monitoring locations

Station	IDs of the contributing sub-catchment areas[1]	Total drainage area (m^2)	Impervious surfaces (%)	Grass on steep slope (%)	Grass on mild slope (%)	Mixed grasses and trees (%)	Relatively natural vegetation (%)
A	1	13576	40	0	15	9	36
B	2	18721	5	0	6	2	87
C	3	21862	27	0	32	2	39
D	1,2,4	53904	20	4	13	17	46
E	1,2,3,4,5	85000	25	6	16	11	42

707 [1] The sub-catchment delineation is represented in Figure 1

708

Table 2: Main characteristics of selected rainfall events and error criteria between observed streamflow time series and those estimated by the modular model

Event	Date	Number of dry hours before beginning of event	Normalized daily average of pressure head (m)	Total Rainfall (mm)	Rainfall duration (hour)	Maximum 30-minutes rainfall intensity (mm/hour)	Shape of hydrograph	Error criteria		
								RRMSE	NSE	CC
1	21/11/2011	15.8	0.94	39.8	2.3	69.2	Single	0.51	0.97	0.98
2	09/03/2012	87.2	0.59	32.6	3.6	46.1	Single	0.46	0.97	0.98
3	25/03/2012	74.7	0.65	64.2	4.8	89.5	Multiple	0.87	0.96	0.98
4	08/02/2013	16.7	1.18	82.2	2.7	104.1	Single	0.42	0.98	0.99
5	09/03/2013	165.4	0.81	34.0	3.6	58.7	Multiple	0.58	0.97	0.98
6	27/04/2013	25.5	0.78	12.0	2.7	22.5	Single	0.38	0.98	0.99

Table 3: Parameter settings of Genetic Algorithm (GA) and Interior Point Algorithm (IPA)

Algorithm	Parameter	Setting
GA	Population size	50
	Selection function	Stochastic uniform
	Mutation function	Adaptive feasible
	Crossover function	Scattered
	Hybridization	IPA
	Number of generations	100
	Function tolerance	1e-10
	Nonlinear constraint tolerance	1e-10
IPA	Start point	Optimal values from GA
	Maximum iterations	1000
	Maximum function evaluations	3000
	Function tolerance	1e-10
	Nonlinear constraint tolerance	1e-10
	X tolerance	1e-10
	Hessian	BFGS
	Derivative type	Central differences

735 Table 4: Statistical features of the rainfall events

Rainfall Events	Parameter	Mean	StDev	Number of occurrences		
				Sub-Cluster 1	Sub-Cluster 2	Sub-Cluster 3
Cluster I	P	3.8	2.6	35	24	43
	I_{30}	5.5	4.6			
	R_D	1.5	0.4			
Cluster II	P	16.2	3.9	4	10	3
	I_{30}	22.6	4.6			
	R_D	2.7	0.6			
Cluster III	P	31.2	4.3	4	7	10
	I_{30}	42.5	9.2			
	R_D	3.5	0.3			
Cluster IV	P	59.7	10.9	3	4	3
	I_{30}	67.6	20.0			
	R_D	5.0	1.0			

736 P: Rainfall depth (mm)
737 I_{30}: Maximum 30-min intensity (mm/hr)
738 R_D: Rainfall duration (hr)
739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754 Table 5: Error criteria between filtered quickflow time series from observed discharge
755 data and those estimated by the quickflow module of modular model

Station	Data Set	Error criteria		
		RRMSE	**NSE**	**CC**
A	Train	0.69	0.94	0.96
	Test	0.73	0.95	0.97
B	Train	0.67	0.95	0.96
	Test	0.71	0.94	0.97
C	Train	0.65	0.95	0.97
	Test	0.66	0.95	0.97
D	Train	0.51	0.97	0.99
	Test	0.60	0.96	0.98
E	Train	0.54	0.97	0.99
	Test	0.65	0.96	0.98

756 RRMSE: Relative Root Mean Squared Error
757 NSE: Nash–Sutcliffe Efficiency
758 CC: Correlation Coefficient
759
760
761
762
763
764
765
766
767
768

Table 6: The validated model parameters for the quickflow and baseflow modules

Station	Baseflow module		Quickflow module											
			Rapid runoff				Delayed runoff							
	b	k	a	r	j_1	j_2	c	d_1	d_2	d_3	d_4	d_5	d_6	d_7
A	0.30	0	0.58	0.23	0.73	0.27	0.14	0.00	0.00	0.05	0.06	0.10	0.60	0.19
B	0.32	0	0.21	0.27	0.58	0.42	0.11	0.00	0.00	0.00	0.07	0.19	0.65	0.09
C	0.27	0	0.47	0.25	0.70	0.30	0.15	0.00	0.00	0.13	0.15	0.20	0.50	0.02
D	0.31	0	0.46	0.24	0.67	0.33	0.16	0.00	0.00	0.08	0.11	0.13	0.58	0.10
E	0.29	0	0.49	0.25	0.65	0.35	0.17	0.00	0.00	0.07	0.09	0.11	0.61	0.12

780 Table 7: Average runoff coefficient for each land use within clusters and sub-clusters

Rainfall Events	Sub-Cluster	Runoff coefficient (-)					Relative absolute error[1] (%)	
		Impervious surfaces	Grass on steep slope	Grass on mild slope	Mixed grasses and trees	Relatively natural vegetation	Mean	Std.dev
Cluster I	1	0.66	0.30	0.25	0.20	0.04	4.6	0.7
	2	0.67	0.31	0.27	0.21	0.05	3.9	1.7
	3	0.69	0.38	0.34	0.29	0.06	4.4	1.0
Cluster II	1	0.73	0.52	0.43	0.36	0.07	3.8	1.6
	2	0.75	0.53	0.45	0.38	0.07	3.6	1.2
	3	0.75	0.55	0.45	0.38	0.10	4.2	1.9
Cluster III	1	0.82	0.57	0.46	0.39	0.11	3.9	0.6
	2	0.83	0.63	0.47	0.40	0.12	3.2	0.9
	3	0.84	0.66	0.55	0.52	0.17	3.7	0.3
Cluster IV	1	0.94	0.71	0.63	0.55	0.17	3.4	1.4
	2	0.95	0.75	0.68	0.61	0.18	4.0	0.4
	3	0.96	0.81	0.77	0.73	0.24	3.8	2.0

781 [1] The relative absolute error was calculated according to the absolute error of Equation 4 estimates relative to values obtained from Equation 3

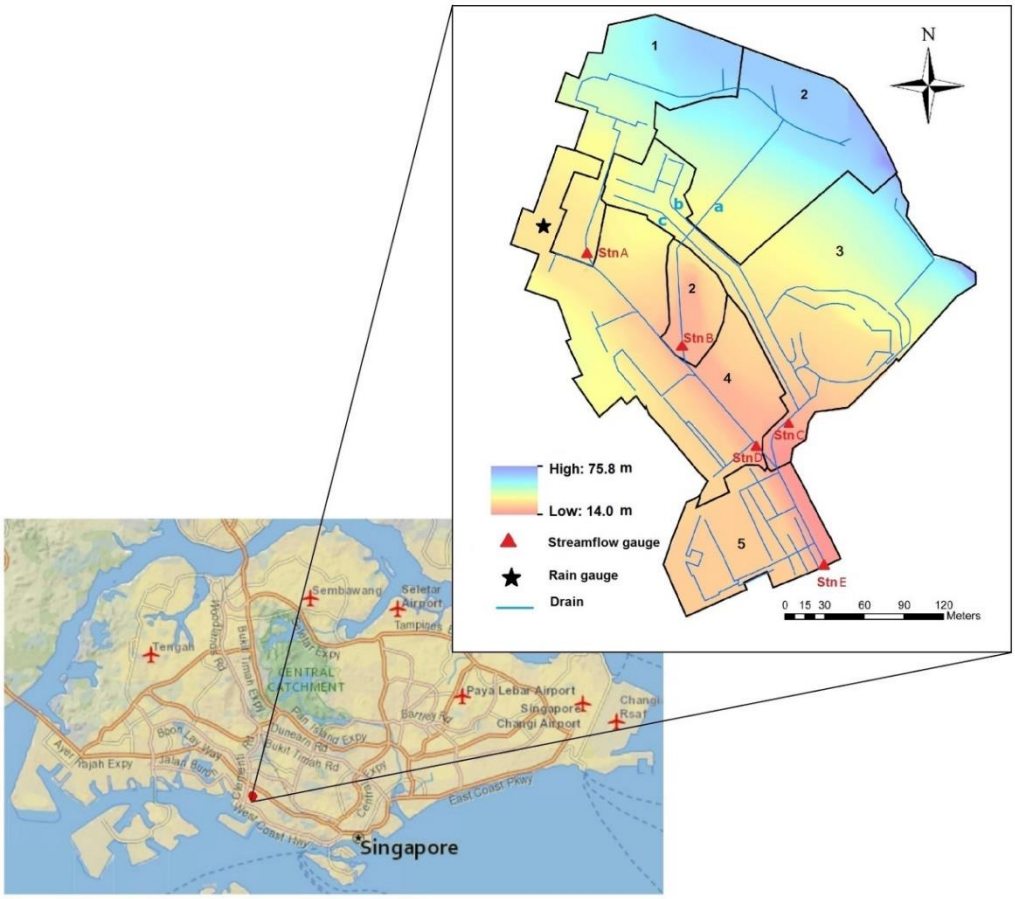
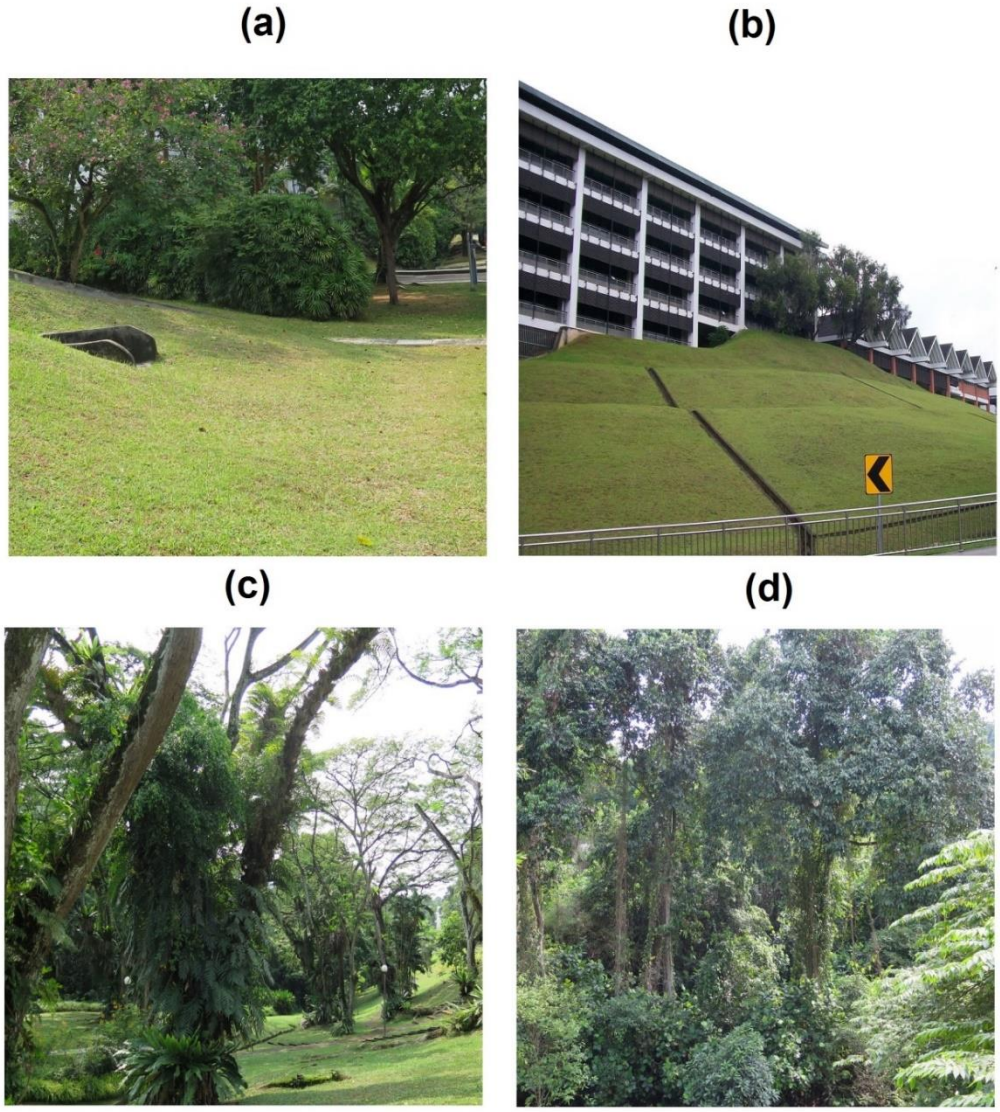


Figure 1: Location of the study area (Kent Ridge Catchment, Singapore) and its respective topography, monitoring stations (Stn A-E), sub-catchments and drainage infrastructure (drain ‘a’ crosses drains ‘b’ and ‘c’ on the map, but they actually do not intersect on site)

790
791 **Figure 2:** Land uses of study area

792

793



794

795 **Figure 3:** Land uses of Kent Ridge Catchment including a) grass on mild slope, b) grass
796 on steep slope, c) mixed grasses and trees and d) relatively natural vegetation

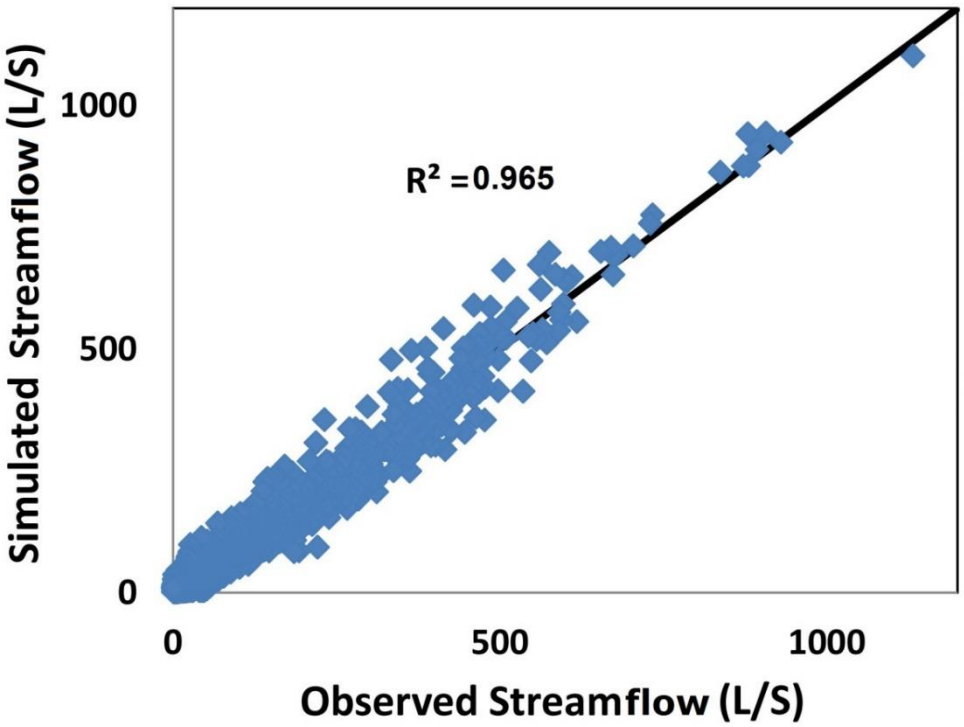


Figure 4: Scatter plot between observed streamflow and those estimated by modular model at Station E which situates at catchment outlet

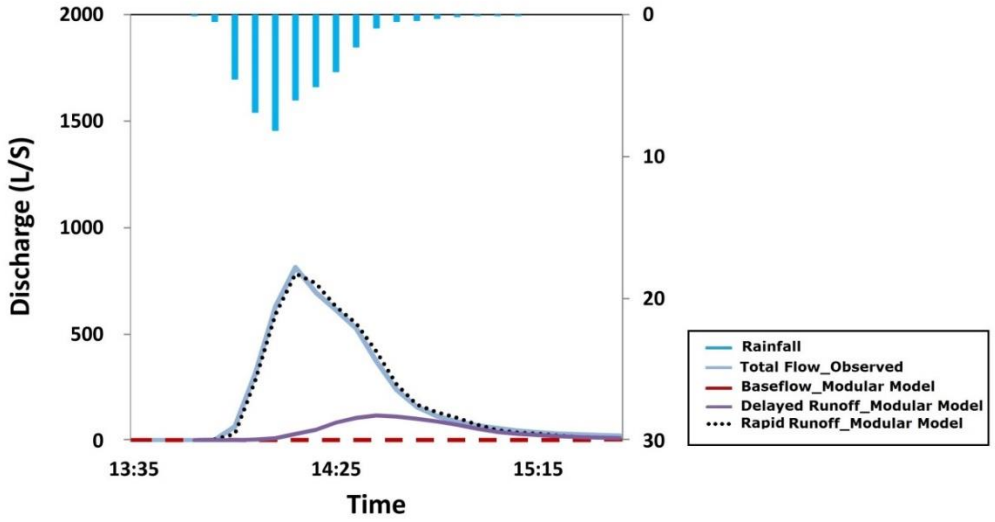


Figure 5: Separation of observed streamflow data into its respective flow components using the modular model for a selected rainfall event occurred on 21/11/2011

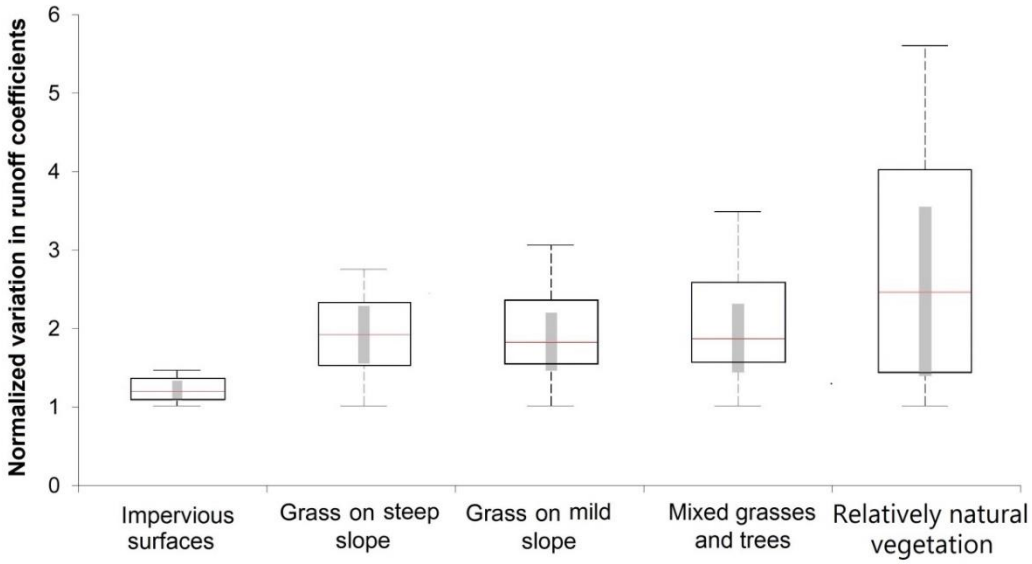
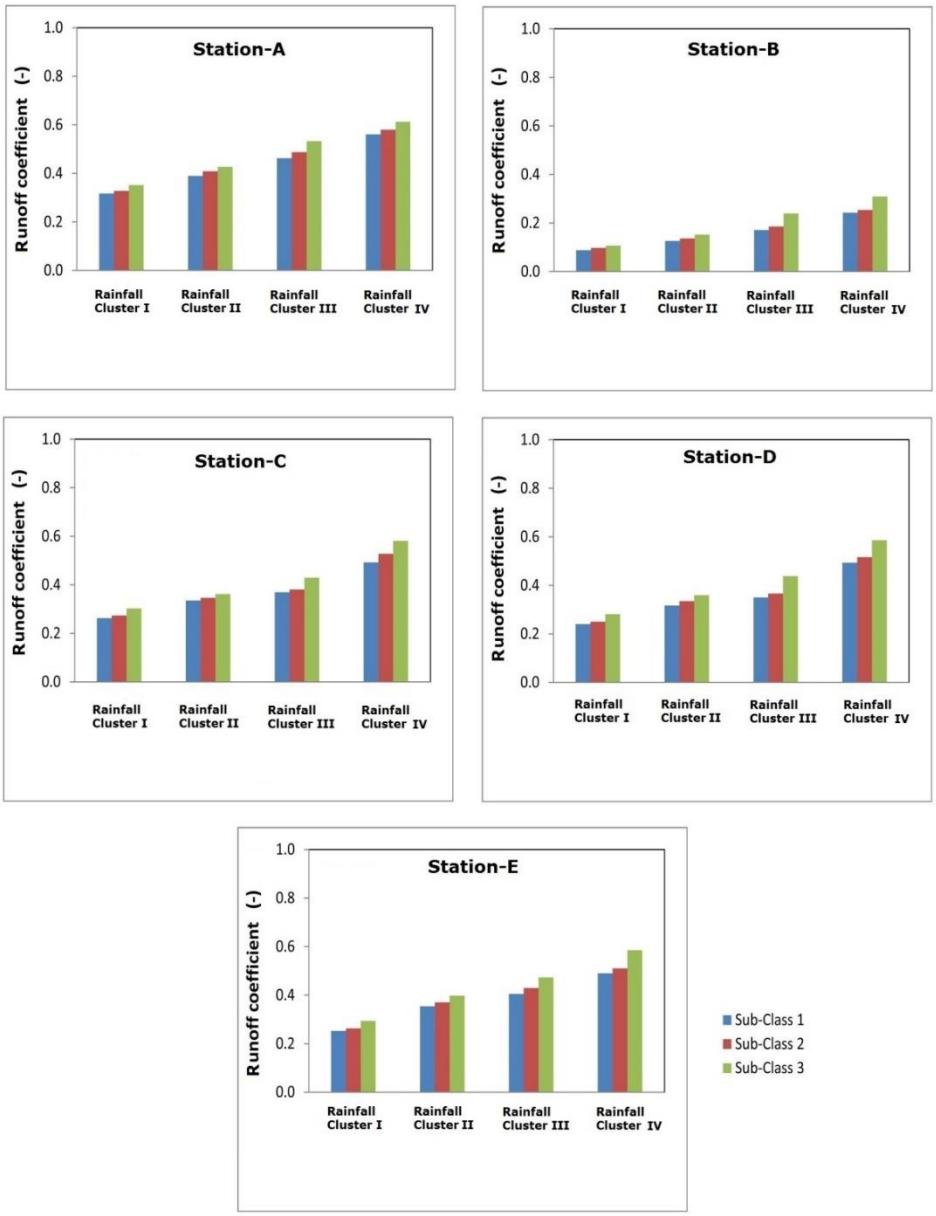


Figure 6: Normalized variation in runoff coefficients (with respect to their minimum value within each land use) of different land uses from Cluster-I/Sub-Cluster-1 to Cluster-IV/Sub-Cluster-3 (grey bars represents the expected range of variability of the median)

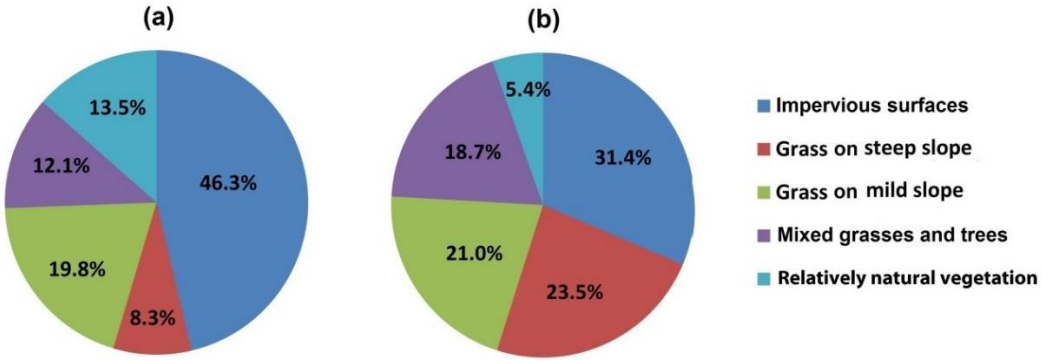


833

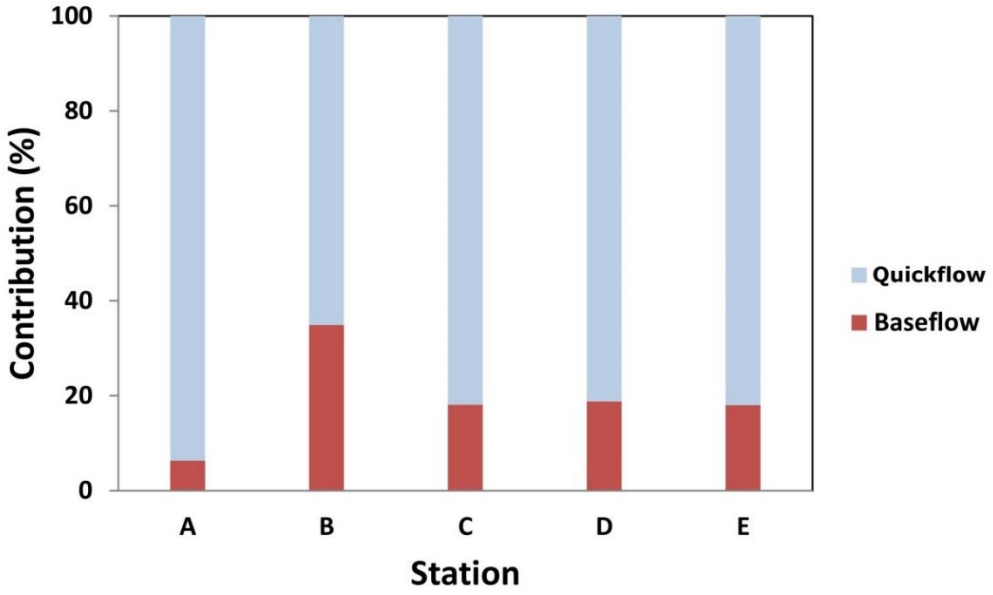
834 **Figure 7**: Average runoff coefficients for all discharge monitoring stations for the
835 various rainfall clusters and sub-clusters

836

837



838
839 **Figure 8**: Total land use specific quickflow contributions towards Station E from
840 September 2011 until August 2012 for: a) absolute amount basis and b)
841 equivalent area basis
842
843
844
845
846
847
848
849



850

851 **Figure 9**: Average contribution (%) of baseflow and quickflow from 150 rainfall events
852 towards the discharge measured at sub-catchment (Stations A-D) and
853 catchment (Station E) level

854