water

Article

The Revised Curve Number Rainfall–Runoff Methodology for an Improved Runoff Prediction

Kenneth Kai Fong Lee [1], Lloyd Ling [1,*] and Zulkifli Yusop [2]

[1] Centre of Disaster Risk Reduction (CDRR), Lee Kong Chian Faculty of Engineering & Science, Civil Engineering Department of Universiti Tunku Abdul Rahman, Jalan Sungai Long, Kajang 43000, Malaysia
[2] Centre for Environmental Sustainability and Water Security, Universiti Teknologi Malaysia, Skudai 81310, Malaysia
* Correspondence: linglloyd@utar.edu.my

Abstract: The Curve Number (CN) rainfall–runoff model is a widely used method for estimating the amount of rainfall and runoff, but its accuracy in predicting runoff has been questioned globally due to its failure to produce precise predictions. The model was developed by the United States Department of Agriculture (USDA) and Soil Conservation Services (SCS) in 1954, but the data and documentation about its development are incomplete, making it difficult to reassess its validity. The model was originally developed using a 1954 dataset plotted by the USDA on a log–log scale graph, with a proposed linear correlation between its two key variables (I_a and S), given by $I_a = 0.2S$. However, instead of using the antilog equation in the power form ($I_a = S^{0.2}$) for simplification, the $I_a = 0.2S$ correlation was used to formulate the current SCS-CN rainfall–runoff model. To date, researchers have not challenged this potential oversight. This study reevaluated the CN model by testing its reliability and performance using data from Malaysia, China, and Greece. The results of this study showed that the CN runoff model can be formulated and improved by using a power correlation in the form of $I_a = S^{\lambda}$. Nash–Sutcliffe model efficiency (E) indexes ranged from 0.786 to 0.919, while Kling–Gupta Efficiency (KGE) indexes ranged from 0.739 to 0.956. The I_a to S ratios (I_a/S) from this study were in the range of [0.009, 0.171], which is in line with worldwide results that have reported that the ratio is mostly 5% or lower and nowhere near the value of 0.2 (20%) originally suggested by the SCS.

Keywords: newly calibrated CN methodology; revised CN model; new runoff predictive model

check for updates

Citation: Lee, K.K.F.; Ling, L.; Yusop, Z. The Revised Curve Number Rainfall–Runoff Methodology for an Improved Runoff Prediction. *Water* **2023**, *15*, 491. https://doi.org/10.3390/w15030491

Academic Editors: Nigus Demelash Melaku, Anoop Valiya Veettil, Di Liu and Aizhong Ye

Received: 19 December 2022
Revised: 8 January 2023
Accepted: 18 January 2023
Published: 26 January 2023

1. Introduction

Floods are natural disasters caused by a large amount of water that overflows a land beyond its normal level. One of the main causes of flooding is the excess rainfall that is not absorbed by the ground or a land surface, which is also known as the runoff amount. When the runoff amount exceeds the capability of the existing flood control infrastructure, overflows of the water will occur, leading to flooding events. Once flooding occurs, it will threaten the lives of the residents, especially those who live in the lower elevated regions. Moreover, the flood will also lead to a huge amount of financial loss, which includes the damage brought by the flood to the victims and financial costs for the government to recover the flood-stricken regions and indemnify the victims of the flood. In addition to impacting human beings, animals will also lose their shelter due to flooding. Thus, a better understanding of the relationship between the rainfall and the runoff amount is crucial to produce a better runoff prediction model. With a better prediction of the runoff amount, government agencies will be able to efficiently plan for the management of flood-related issues. For instance, the government can plan a sufficient amount of flood prevention infrastructure, such as the retention pond, the drainage system, the reservoir, and others, according to the predicted runoff amount with the aim of coping with the possible flood issues.

The change in the runoff depth, or the amount of water that flows over the surface of the land, can be influenced by a number of factors, including climate change, changes in land use, and anthropogenic interventions [1–5]. Some studies have found that both climate change and human activities can contribute to changes in runoff, with climate change having a greater impact in the 1980s and human activities having a greater impact in the 2000s [4]. Other studies have focused specifically on the impact of land use changes, such as urbanization, on runoff depth [3,5]. These studies have found that land use changes, such as converting dryland agriculture to an impervious surface, can increase runoff depth, while converting dryland agriculture or an impervious surface to grass or forest can decrease runoff depth [3]. Additionally, human disturbance such as vegetation restoration projects can also affect the runoff amount [2] while changes in a forest area can also have a significant impact on the change in the runoff regime [1–3].

To estimate the rainfall–runoff amount, the United States Department of Agriculture (USDA) and Soil Conservation Services (SCS) developed the Curve Number (CN) rainfall–runoff model in 1954. The CN model was derived based on the maximum rainfall and runoff data that are collected from less than 200 watershed data in 23 different states in the USA. However, the data and documentation about the initial development of the method are not complete, and many datasets have been lost, which poses a great challenge in reassessing the validity of what was used to formulate the SCS's CN rainfall–runoff model. The SCS developed the method based on data mainly obtained from watersheds monitored with rain and streamflow gauges in the US. [6–10]. With the data and effort from SCS, the SCS's CN model was developed as follows:

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \quad (1)$$

where

Q = Runoff depth (mm).

P = Rainfall depth (mm).

I_a = The initial abstraction amount (mm).

S = Maximum potential water retention of a watershed (mm).

Besides developing the model, the SCS also suggested that $I_a = \lambda S$, in which λ is the initial abstraction coefficient ratio to relate the I_a and S values. As in 1954, there is a limited amount of rainfall data; thus, this equation was flimsily justified based on daily rainfall and runoff data. The only surviving evidence is documented in a different edition of the Natural Resources Conservation Services's (NRCS) National Engineering Handbook, Section 4 (NEH-4) [8–10].

In addition to the model, the SCS also proposed a fixed constant of 0.2 as the value for the λ, for the equation $I_a = \lambda S$. Thus, the linear correlation with the form of $I_a = 0.2S$ was introduced to simplify the SCS's CN model and the calculations. In NRCS NEH-4, there is a note mentioning that the reason for proposing the λ as 0.2 is because 50% of the data points fall between the range of $0.095 < \lambda < 0.38$ [10]. With the substitution of $I_a = 0.2S$ into the SCS's CN model, Equation (1) will eventually be simplified into the conventional SCS runoff forecast model. The conventional (simplified) SCS rainfall–runoff prediction model is as follows:

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S} \quad (2)$$

where the restriction of $P > 0.2S$ must be obeyed, or else there will be no runoff (Q) occurring.

Since the inception of the CN runoff model in 1954, it has gained high acceptance in hydrology studies. The model was also frequently applied to hydrologic problems for which it was not originally designed [11]. The conventional model is commonly introduced in textbooks, certified hydrological manuals, and even incorporated in many software as well as programs since 1954 [12]. The model is also integrated into many USDA SCS systems and hydrology-related models [13]. Since the conventional model is frequently applied to other hydrologic models, the accuracy and the reliability of the model itself can be a major factor that determines the performance of the related software and the reliability of any handbook that is related to the model.

Throughout the years, there are researchers who have questioned the reliability and the accuracy of the hypothesis that proposed the fixed value of 0.2 as the value of λ. Based on the results obtained from different studies based on different regions, different ranges of λ value had been proposed. For example, the US researchers had suggested a range of $0.02 < \lambda < 0.07$ and $\lambda = 0.05$ was the best overall value for American watersheds. On the other hand, Australian studies reported that the range of the λ should be around 0.2 or lower [14,15]. This suggested that the λ value will differ accordingly to a different region; thus, the SCS's hypothesis of fixing the λ value as 0.2 became ambiguous and led to the tendency of the CN runoff model to underpredict or overpredict the runoff amount. Some researchers have overlooked the fact that the λ value is specific to a particular region or watershed and have even directly adopted $\lambda = 0.05$ in their studies, even though it was a result from the US [7]. Therefore, it is crucial to develop a model calibration methodology based on the specific rainfall–runoff characteristics of a watershed for SCS practitioners.

In the past six decades, many studies had been carried out to challenge the SCS's hypothesis, which fixed the λ value at 0.2. There were also studies that tried to determine the optimal λ value for the linear modelling that can improve the accuracy and reliability of the SCS's CN model.

Reviewing the graphs in Figure 1 (graphs obtained from three different sources), the data points are all plotted on a log–log scale graph. However, the linear correlation with the form of $I_a = \lambda S$ was used to formulate and simplify Equation (1) into (2). This can be a mathematical error or an oversight in the past which leads to the inaccuracy of the model's runoff predictive ability. Since the SCS's field data points were plotted on a log–log scale graph, the relation between the two key variables (S and I_a) should be in the power correlation instead of a linear correlation.

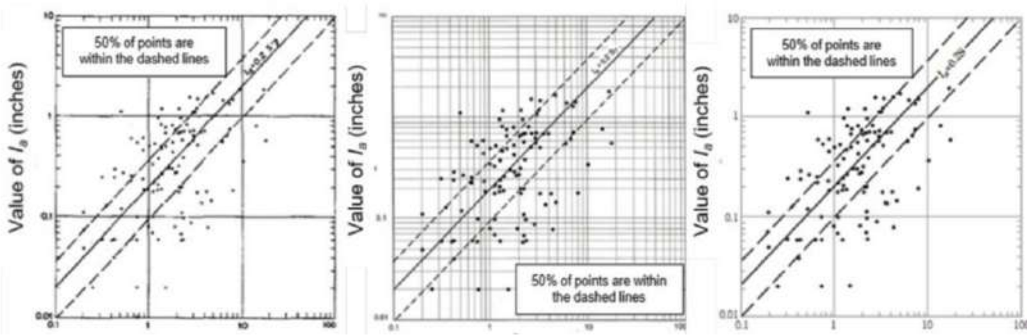

Figure 1. SCS original dataset plotted on log–log-scaled graph. Source: [8–10].

To date, SCS curve number rainfall–runoff-related studies either evolved around its basic model (Equation (1)) calibration or added new variable(s) to create a runoff model variant. No published work has explored other possible correlations between I_a and S and assessed the impact to its rainfall–runoff mechanism yet. Equation (2) exists in all hydrology textbooks; it also led to curve number derivation and has been applied in many designs while many types of software incorporated $I_a = 0.2S$ in their algorithms. If I_a and S correlate otherwise, it will change the whole model and the way SCS curve number theory is taught. If the fundamental core theory of the SCS curve number rainfall–runoff model changed, the model and all its related downstream information will have to be re-derived.

2. Materials and Methods

2.1. Validity of the Linear Correlation ($I_a = 0.2S$) and Introduction of $I_a = S^\lambda$

A previous study reassessed the use of the linear correlation of "$I_a = 0.2S$" based on the 1954 SCS original dataset and rejected the validity of the linear correlation model of the form $I_a = 0.2S$, at alpha level = 0.01. Instead of a constant of $\lambda = 0.2$, the study proved that the best linear model for the 1954 SCS original dataset should be $I_a = 0.112S$ [16], which is also in line with the best regressed linear equation of $I_a = 0.111S$ reported in another study [17]. As the linear correlation of $I_a = 0.2S$ was reported as not statistically significant when reevaluated using the original 1954 SCS field data, the SCS rainfall–runoff predictive model Equation (1) should be recalibrated before it is used.

As supported by the power law, if the independent variables and dependent variables were linearly correlated to each other on a log–log scale graph, then a power regression equation will be able to correlate both the variables, in terms of the normal scale. Since the best-fitted linear equation for the 1954 SCS original dataset was reported as $I_a = 0.112S$ or $I_a = 0.111S$ in the past two studies [16,17], the best fitting model for the same datasets in a normal scale should be in the power model with the form of $I_a = S^{0.112}$ or $I_a = S^{0.111}$ to simplify Equation (1).

Since Equation (2) had been proven to be insignificant at the alpha level of 0.01 by past studies [16,17], rederiving the model will be compulsory to ensure the precision of the model's runoff predictions. There is no new evidence to suggest a different correlation between I_a and S due to the SCS's incomplete documentation and missing datasets. Thus, this study adopted the model calibration technique developed in our past studies and thoroughly assessed the use of the power correlation equation in the general form of $I_a = S^{\lambda}$ to simplify the SCS's CN base model, Equation (1), and recalibrate the runoff prediction model based on the regional rainfall–runoff dataset using inferential statistics.

2.2. Study Site and Models' Performance

In this study, the reliability and performance of the recalibrated rainfall–runoff model were tested using datasets from different locations in Malaysia and datasets from two previous studies in China and Greece. In Malaysia, the Department of Irrigation and Drainage's Hydrological Procedure no. 11 (DID-HP 11) collected 474 rainfall–runoff data pairs from 1964 to 2016 from watersheds throughout Malaysia, while the Hydrological Procedure no. 27 (DID-HP 27) collected 227 rainfall–runoff data pairs from 1970 to 2000 [18,19]. The study also combined the DID-HP 11 and 27 datasets (consisting of 701 data pairs collected from 1970 to 2016) in the analysis. In addition, 72 rainfall–runoff data pairs collected at the Kerayong watershed in Kuala Lumpur, Malaysia, and 93 data pairs from the Kayu Ara watershed in the Petaling Jaya area of Malaysia were included to test the performance of the recalibrated model [20,21]. Data from two journals were also used in this study: 29 rainfall–runoff data pairs (1994 to 1996) from the Wang Jia Qiao watershed in the Zigui County of Hubei Province, China, [7] and 76 rainfall–runoff data pairs from Attica, Greece [22]. Figure 2 shows the locations from which the rainfall–runoff data were collected.

For purposes of the study, the statistically significant total abstraction value (S), the initial abstraction ratio coefficient (λ), and the corresponding 99% Confidence Interval (CI) for both S and λ will be derived via the Bootstrapping method—BCa procedure (with 2000 samples) at alpha = 0.01, by using the IBM Predictive Analytics software (PASW) version 18.0 (commonly known as SPSS)—as the values will be later used for the calculation of the Curve Number (CN) value.

The main reason for applying the bootstrap BCa technique is because of its robustness and data-distribution-free nature, which is suitable for rainfall–runoff data analysis. Additionally, the bootstrap BCa analytical module is available in SPSS, which is conveniently used to carry out the necessary analysis for this study. The BCa technique also has the ability to correct biases when identifying the confidence interval of the variable of interest at a specific alpha level for further statistical assessments [21,23–26]. Two normality tests (Kolmogorov–Smirnov and Shapiro–Wilk) are also available in SPSS with the bootstrap BCa confidence interval to identify whether the mean or median CI should be chosen to identify the optimal S and optimal λ for each dataset. If the dataset is not normally distributed, the median CI will be chosen for variable optimization and vice versa.

Figure 2. Three country locations (seven rainfall–runoff datasets) of this study. (The information of all the catchment locations is summarized in Appendix E.)

To make sure the overall prediction results were not biased against any of the datasets, the calibration of the SCS runoff model in identifying the optimum value of λ and S based on each P–Q data pair was set to fulfil the model prediction bias at zero level. With the zero bias optimization constraint, the supervised numerical algorithm is proceeded to identify the optimum λ value and optimum S value, within the bootstrap BCa 99% confidence interval. The performance and the predictive accuracy of the models were later assessed for each dataset based on the Nash Sutcliffe Index (E). The Nash Sutcliffe Index is a hydrology model efficiency indicator, in which E = 1 means the model is an ideal model for the prediction, while E < 0 shows the model performance is unacceptable [27–29]. The model bias value of 0 indicates that the model is an ideal, error-free model. Both the Nash Sutcliffe Index and Bias can be calculated via Equations (3) and (4) as shown below:

$$\mathrm{E} = 1 - \frac{\sum_{i=1}^{n}\left(\mathrm{Q}_{\mathrm{predicted}} - \mathrm{Q}_{\mathrm{observed}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{Q}_{\mathrm{predicted}} - \mathrm{Q}_{\mathrm{mean}}\right)^2} \tag{3}$$

$$\mathrm{BIAS} = \frac{\sum_{i=1}^{n}\left(\mathrm{Q}_{\mathrm{predicted}} - \mathrm{Q}_{\mathrm{observed}}\right)}{\mathrm{n}} \tag{4}$$

There is another metric, known as the Kling–Gupta Efficiency (KGE), that has been proposed to evaluate the performance of rainfall–runoff models. This metric is derived from the Nash Sutcliffe Index and comprises three components: correlation, variability bias, and mean bias [30]. The KGE is frequently used by hydrologists to calibrate and assess the accuracy of rainfall–runoff models [30]. It can be calculated using the following equation:

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \tag{5}$$

where

r = linear correlation between observations and simulations
σ_{sim} = standard deviation of the simulations
σ_{obs} = standard deviation of the observations
μ_{sim} = mean of simulations

μ_{obs} = mean of observations

Similar to the E Index, a value of 1 for the KGE indicates a perfect match between observed and simulated data [30]. A KGE value greater than -0.41 suggests that the model is performing better than the mean flow benchmark [30].

Unlike in our past studies, the general formula for S was rearranged from Equation (1) with $I_a = \lambda S$ to determine the value of S using the values of P and Q data pairs with a closed-form expression for S [7,14,15,21,29]. However, with the introduction of the power model $I_a = S^\lambda$, the general formula for S does not have a closed-form expression in this study, so a numerical analysis technique will be applied to obtain the value of S. (The derivation and simplification of the general formula for S based on the power regressed model are summarized in Appendix A.) As a result, the simplified S general formula obtained was as follows:

$$(2S^\lambda + Q - 2P)^2 = 4SQ + Q^2 \tag{6}$$

The numerical analysis technique will need to be carried out on Equation (6) to identify the optimal λ and S value, denoted by S_λ, according to the P and Q data pairs of the respective study sites. Besides the S_λ values obtained from Equation (6) via the numerical analysis technique, the $S_{0.2}$ values for each of the P–Q data pairs for all the study sites were also determined based on the S general equation that was introduced in the previous study, where the linear correlation was applied, with the form of $I_a = 0.2S$ to simplify and formulate the current SCS runoff predictive model [6,7,14–17,21,29]. (The S general Equation that was used to obtain the $S_{0.2}$ values was summarized in Appendix B.) The S general equation that was used to obtain the $S_{0.2}$ values was shown as follows:

$$S_{0.2} = 5\left(P + 2Q - \sqrt{4Q^2 + 5PQ}\right) \tag{7}$$

Once the S_λ and $S_{0.2}$ values were obtained with Equations (6) and (7) according to the corresponding P and Q data pairs, the best correlation between S_λ and $S_{0.2}$ and the corresponding best correlation equation can be mapped via the SPSS model regression module for a study site. When the optimal λ value is different from the conventional value of λ = 0.2, a correlation between the newly found λ value and 0.2 must be used to calculate the curve number again [7,11,14,16,21,29]. Therefore, it is important to identify the best correlation between S_λ and $S_{0.2}$ and then convert the S_λ back to the equivalent $S_{0.2}$ value to derive the conventional SCS curve number ($CN_{0.2}$) value that SCS practitioners are familiar with for routine application. The best correlation between S_λ and $S_{0.2}$ is determined by the highest adjusted R square value ($R^2{}_{adj}$), the lowest standard error of the estimate, and a model with a significant *p* value, which can be determined in SPSS through the regression module [31,32]. (The approach is summarized in Appendix C.)

For example, the best correlation equation from SPSS between S_λ and $S_{0.2}$ for the Kerayong watershed, Malaysia, was identified in SPSS as follows:

$$S_{0.2} = 3.223\ S_{0.199}{}^{0.355} \tag{8}$$

As introduced by the SCS, the equation to calculate the SCS curve number value is shown below:

$$CN_{0.2} = \frac{25,400}{254 + S_{0.2}} \tag{9}$$

$CN_{0.2}$ = Conventional Curve Number, Curve number of λ = 0.2.

$S_{0.2}$ = Total abstraction amount of λ = 0.2 (mm).

At the same time, Equation (9) can be rearranged into Equation (10):

$$S_{0.2} = 254\left(\frac{100}{CN_{0.2}} - 1\right) \tag{10}$$

By substituting the best correlation equation (Equation (8)) back to the SCS curve number (Equation (9)), the conventional curve number for the study site can be calculated accordingly:

$$CN_{0.2,\ Kerayong} = \frac{25,400}{254 + \left[3.223\ (S_{0.199})^{0.355}\right]} \quad (11)$$

The correlation equation that correlates both S_λ and $S_{0.2}$ will be very crucial as it will directly affect the overall performance of the rainfall–runoff model derived from the power regressed model. The correlation equation is also the key to convert S_λ back to its equivalent $S_{0.2}$ value for the calculation of $CN_{0.2}$ for SCS practitioners [29,32].

3. Results

3.1. Study Site and Models' Performance

In this study, the S and λ values were calculated for every dataset based on each corresponding P and Q value. Then, both the Kolmogorov–Smirnov and Shapiro–Wilk tests were carried out to test the normality of each dataset. Both λ and S were not normally distributed ($\rho < 0.001$) for all the datasets except for Wang Jia Qiao watershed in China. The Kolmogorov–Smirnov and Shapiro–Wilk tests on the λ values showed a significance level of 0.319 and 0.2, which means that the λ values were normally distributed while S values were not normally distributed for Wang Jia Qiao's dataset ($\rho < 0.01$). This normality analysis in this study later helps in determining the optimal S value and λ value for each dataset based on the power regression correlation ($I_a = S^\lambda$) within the 99% Confidence Interval of both S and λ values, accordingly [7,21,29,33,34]. Once the optimal S value and λ value were obtained for each corresponding dataset, the runoff prediction model can be formulated (refer to Appendix C section for model formulation guide) and the performance for each model can be analyzed. Table 1 shows the inferential statistics, which include the Nash Sutcliffe index (E), optimal S (S), optimal λ, initial absorption (I_a), and the ratio of I_a to S, for each corresponding dataset.

Table 1. Inferential Statistics of all the datasets, including DID-HP 27, DID-HP 11, the combination of DID-HP 27 and DID-HP 11, Kerayong, Kayu Ara, Attica, and Wang Jia Qiao.

Dataset & Location	E	N	Bias (mm)	KGE	S_λ (mm)	$I_a = S^\lambda$ (mm)	I_a/S
DID HP 11, Malaysia	0.823	474	2.35	0.805	139.48	3.71	0.027
DID HP 27, Malaysia	0.919	227	0.65	0.956	165.94	2.33	0.014
DID HP 11+27, Malaysia	0.875	701	2.535	0.863	143.99	3.19	0.022
Kayu Ara, Malaysia	0.810	94	0	0.865	24.79	2.20	0.089
Kerayong, Malaysia	0.888	73	0	0.818	9.08	1.55	0.171
Attica, Greece	0.786	77	0	0.798	23.73	1.85	0.078
Wang Jia Qiao, China	0.795	29	0	0.739	386.57	3.57	0.009

All models were optimized under a bias value = 0 constraint, except for DID HP 11, DID HP 27, and DID HP 11+27, as those models were unable to obtain bias = 0 within their 99% BCa CI. Therefore, those models were optimized to obtain the maximum E Index. The E and KGE indexes showed good results for all datasets, with the E index ranging from 0.786 to 0.919 and the KGE index ranging from 0.739 to 0.956. The initial absorption (I_a) for each corresponding dataset was determined using the newly proposed power correlation of $I_a = S^\lambda$. The ratios of I_a to S values were calculated to benchmark against past study results. The I_a to S ratios (I_a/S) for each study site were in the range of [0.009, 0.171], which is in line with past results that reported that the ratio of I_a to S was mostly 5% or lower and nowhere near the value of 0.2 (20%) as initially suggested by the SCS to simplify the SCS runoff predictive framework into Equation (2) [6,7,11–17,21,22,29,32–34].

The best-fitted correlation equation that correlates S_λ to $S_{0.2}$ for each dataset was modelled via SPSS. To measure the fitness of the correlation equation to the datasets, the adjusted coefficient of determination (R^2_{adj}) for each corresponding equation was also identified, together with the ρ-value. The SCS practitioner(s) was taught to choose the CN value from the NEH handbook according to the land use or land cover condition of the watershed and to calculate the S value from Equation (9) or (10) directly; however, it is no longer applicable as the linear correlation of $I_a = 0.2S$ was proven invalid [7,16,17,21,29,33,34]. Thus, the best-fitted correlation equation between S_λ and $S_{0.2}$ proposed herewith should be used by the SCS practitioners, in which the optimal S value of each dataset must be converted back to the equivalent $S_{0.2}$ value via the mapped correlation equation (Table 2) prior to calculate the $CN_{0.2}$ value of a watershed [7,21,29,32–34] from Equation (9) or (10), as shown in the aforementioned example to derive Equations (8) and (11). SCS practitioners no longer have to decide and choose the CN value from any handbook; they can derive the CN value of a specific watershed according to the regional rainfall–runoff conditions.

Table 2. Adjusted coefficient of determination (R^2_{adj}) and ρ-value of all the datasets' correlation equations, including DID-HP 27, DID-HP 11, the combination of DID-HP 27 and DID-HP 11, Kerayong, Kayu Ara in Malaysia, Attica in Greece, and Wang Jia Qiao in China. Power correlation models were identified as the best correlation model in SPSS with the lowest standard error of estimate.

Dataset & Location	Correlation Equation	R^2_{adj}	*p*-Value
DID HP 11, Malaysia	$S_{0.2} = S_{0.265}{}^{0.878}$	0.977	<0.001
DID HP 27, Malaysia	$S_{0.2} = S_{0.166}{}^{0.874}$	0.997	<0.001
DID HP 11+27, Malaysia	$S_{0.2} = S_{0.234}{}^{0.880}$	0.998	<0.001
Kayu Ara, Malaysia	$S_{0.2} = 4.263\ S_{0.246}{}^{0.438}$	0.893	<0.001
Kerayong, Malaysia	$S_{0.2} = 3.223\ S_{0.199}{}^{0.355}$	0.849	<0.001
Attica, Greece	$S_{0.2} = 5.057\ S_{0.194}{}^{0.357}$	0.846	<0.001
Wang Jia Qiao, China	$S_{0.2} = S_{0.214}{}^{0.702}$	0.975	<0.001

3.2. Derivation of Curve Number and Its Confidence Interval

The Bootstrapping BCa 99% CI for the S value of each dataset and the correlation between S_λ and $S_{0.2}$ were determined using SPSS in this study. The BCa 99% CI for the corresponding $CN_{0.2}$ value of each dataset can then be calculated once the respective correlation equation between S_λ and $S_{0.2}$ from Table 2 is substituted into Equation (9) or (10) to derive the equivalent $CN_{0.2}$ range (refer to derivation steps in Appendix C), as shown in Table 3. Table 2 lists the best regressed correlation equation identified in SPSS while Table 3 summarizes the Bootstrapping BCa 99% CI for the S value and the calculated $CN_{0.2}$ value range of each dataset in this study.

Table 3. Bootstrapping BCa 99% CI for the S value and $CN_{0.2}$ value of all the datasets' correlation equations, including DID-HP 27, DID-HP 11, the combination of DID-HP 27 and DID-HP 11, Kayu Ara, Kerayong, Attica, and Wang Jia Qiao.

Datasets	**BCa 99% Confidence Interval**			
S & Curve Numbers	**$S_{0.2}$ (mm)**		**$CN_{0.2}$**	
Dataset & Location	**Lower**	**Upper**	**Lower**	**Upper**
DID HP 11, Malaysia	100.99	139.48	76.88	81.54
DID HP 27, Malaysia	118.65	165.94	74.46	79.62
DID HP 11+27, Malaysia	115.01	143.98	76.21	79.60
Kayu Ara, Malaysia	20.75	331.07	82.43	94.04
Kerayong, Malaysia	8.77	397.55	90.40	97.33
Attica, Greece	23.73	314.15	86.57	94.19
Wang Jia Qiao, China	289.86	553.69	75.08	82.60

4. Discussion

4.1. Validity of the Linear Correlation ($I_a = 0.2S$) and Introduction of Power Regression ($I_a = S^\lambda$)

The Soil Conservation Services Curve Number (SCS CN) method is a widely used method for predicting runoff from precipitation in hydrology studies. It is based on the assumption of a linear relationship between infiltration and soil moisture storage, which is used to estimate the amount of runoff that will occur during a storm event. However, the accuracy and consistency of the SCS CN method have been questioned by many researchers due to concerns about the underlying assumptions and the limited range of conditions under which it has been tested.

Two past studies assessed the linear correlation proposed by the SCS and concluded that $I_a = 0.2S$ was statistically invalid even to its own dataset [16,17]. In 1954, the SCS introduced the linear correlation ($I_a = 0.2S$) to simplify the SCS rainfall–runoff model into the conventional SCS runoff prediction Equation (2) that is in use until today. However, worldwide studies reported runoff prediction accuracy and consistency issues with the model [6,7,11–17,21,22,29–34]. This simplification may have introduced an inherent risk of Type 2 error into the SCS CN model. (A Type 2 error, also known as a false negative, occurs when a statistical test fails to reject the null hypothesis when it is actually false.) In the context of the SCS CN model, this means that the model may incorrectly predict the runoff amount. The risk of Type 2 error in the SCS CN model has potential implications for downstream derivations and extended applications of the model in hydrology. Researchers and practitioners who use the SCS CN model should be aware of this risk and carefully evaluate the assumptions and limitations of the model in order to ensure the accuracy and reliability of their results.

The 1954 SCS original dataset was plotted on a log–log scale graph, but the SCS adopted a linear correlation to correlate both S and I_a directly to simplify its model framework without taking the antilog form [8–10]. In the authors' opinion, it is possible that the antilog correlation form was not used in the past due to technical limitations in simplifying Equation (1) with $I_a = S^{0.2}$ without the use of computers, or the fundamental mathematical mistake was overlooked. Two past studies reassessed the correlation between S and I_a according to the 1954 SCS original dataset and reported that the best correlation equation should be $I_a = 0.111S$ or $I_a = 0.112S$ on the log–log scale graph [16,17]. As such, the antilog form or the power regression equations of $I_a = S^{0.111}$ or $I_a = S^{0.112}$ will be better correlation equations to the SCS original dataset as compared to what the SCS had proposed ($I_a = 0.2S$). This finding necessitated the development of a revised form of the SCS rainfall–runoff model that incorporates the power correlation between I_a and S in the general form of $I_a = S^{\lambda}$ to simplify Equation 1 and rederive a new SCS runoff predictive model. The calibration of the SCS rainfall–runoff model (Equation (1)) can be performed based on the newly proposed power regressed model according to the regional or watershed specific rainfall–runoff dataset to improve the runoff prediction accuracy.

The SCS CN basic model is in the form of $Q = f(P, \lambda, S)$, so there are only two key parameters for optimization. In contrast to previous studies that calibrated the SCS CN model with the $I_a = \lambda S$ equation according to watershed-specific rainfall–runoff datasets using inferential statistics and reported that $\lambda = 0.2$ was not statistically significant at the alpha = 0.05 level for modeling runoff [7,16,21,29,33,34], the novelty of this study is to calibrate the SCS CN model with $I_a = S^{\lambda}$ and formulate new CN runoff prediction models. To date, no other researchers have attempted to do this.

This study preserved its fundamental framework without adding any new parameters in order to study the possibility of using $I_a = S^{\lambda}$ to calibrate and formulate the runoff predictive model. It is noteworthy to highlight that although Equation (2) was able to achieve attractive E and KGE index values (Appendix D, Table A2) in certain datasets, the λ value of 0.2 was not significant even at $\alpha = 0.05$ level for any dataset in this study. As such, the conventional SCS simplified model (Equation (2)) is not statistically significant for modelling any rainfall–runoff in this study. All SCS models have a tendency to underpredict runoff in this study. Additionally, the I_a value of all SCS models violated a requirement from the SCS CN theory which states that there will be no runoff when the rainfall depth (P) is less than the I_a value. Table A2 tabulates the total count of rainfall–runoff data pairs where the recorded P value is less than the SCS model's I_a value. In contrast, none of the newly formulated power models have this issue.

4.2. Application of the Power Regression Model at Different Study Sites

In this study, datasets were collected from Malaysian watersheds (DID-HP 11 and DID-HP 27, Kerayong and Kayu Ara watershed in Kuala Lumpur area) while two past study datasets were adopted from the Attica watershed in Greece and the Wang Jia Qiao watershed in China. The analytical results showed that the power regression model is able to obtain high runoff predictive ability at all study sites consistently (Appendix D, Table A1) when compared to the conventional SCS runoff predictive model (Equation (2)). The calculated ratio values of I_a to S (I_a/S) for all the study sites were in the range of [0.009, 0.171] (Table 1) which is in line with past results whereby the ratios were mostly 5% or lower and nowhere near to the value of 0.2 (20%) as initially suggested by the SCS [6,7,11–17,21,22,29,32–34]. It is in line with the results of the largest scale of field tests on the SCS CN model, which analyzed more than half a million rainfall events across 24 states in the USA and reported that I_a/S values were mostly below 5% to achieve better runoff modeling results in American watersheds [17,29,32]. Using the revised CN model may allow for more accurate and reliable predictions of runoff in different types of soils and hydrological conditions. This can be particularly useful for water resource management, as it allows for more informed decision making about the availability of water for various uses and the potential for flood hazards. It is important for researchers and practitioners to be aware of the revised form of the SCS rainfall–runoff model and to consider its use in their work.

4.3. Application of Machin-Learning Techniques to the Rainfall Runoff Model

In recent decades, there has been a significant increase in the use of machine learning in hydrology studies in Latin America and the Caribbean (LAC) [35–38]. Researchers have introduced various machine-learning techniques, such as long short-term memory (LSTM), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multilayer Perceptron (MLP), Wavelet Neural Network (WNN), Ensemble Prediction Systems (EPS), Support Vector Machine (SVM) and Support Vector Regression (SVR), and Artificial Neural Networks (ANN), to improve the accuracy and performance of predictive models [35]. Hybrid machine-learning approaches, such as ANFIS and WNN, have also been found to have improved accuracy and performance for long-term and short-term rainfall–runoff models [35]. Machine learning can also be used to overcome the issue of lacking time series data for flow hydrographs due to the absence of gauged stations by utilizing flood modelling methods such as the reverse flood routing model, HEC-RAS, and GIS flood maps [38]. It is possible that machine-learning techniques can be combined with the newly proposed rainfall–runoff model from this study.

5. Conclusions

This study proposed a new correlation model to improve the accuracy of the SCS (Soil Conservation Service) rainfall–runoff predictive model, using inferential statistics. The key findings of the study are as follows:

1. The linear correlation hypothesis (I_a = 0.2S) proposed by the SCS was found to be statistically invalid, even in the 1954 original dataset. Therefore, it is important to revise the current conventional SCS runoff prediction model to better prepare for the challenges posed by climate change conditions. The use of I_a = 0.2S to simplify equation (1) into the conventional SCS runoff model (equation 2) may have been an oversight in 1954. The power correlation equation ($I_a = S^{0.111}$ or $I_a = S^{0.112}$) should be used to simplify the 1954 SCS runoff model equation (1) and derive the runoff predictive model, as the original SCS dataset was plotted on a log–log graph.
2. The newly proposed power regression model ($I_a = S^{\lambda}$) demonstrates good runoff prediction ability at different study sites, including those in Malaysia, Greece, and China. The calibrated SCS rainfall–runoff model based on the proposed power regression model is promising for modelling the rainfall–runoff characteristics of different watersheds in different countries. The ratio of I_a to S for all study sites is mostly 5% or

lower, which is in line with past worldwide study results and much lower than the value of 0.2 (20%) as suggested by the SCS.

3. There is concern about the use of the original form of the SCS CN model in education, as it may teach students an oversimplified and potentially inaccurate model for predicting runoff, which could have serious consequences for fields such as water resource management, environmental science, and civil engineering. There is also concern about the widespread use of the original form of the model in educational materials, such as textbooks, software, and government agency handbooks and trainings, which may perpetuate the use of an oversimplified and potentially inaccurate model for predicting runoff.
4. The proposed methodology has several limitations, including the need for a minimum sample size of at least 20 data pairs to obtain meaningful inferential results and the reliance on the bootstrap BCa method to produce confidence intervals for key variable optimization and the formulation of a new runoff predictive model. The statistical software used must also include the bootstrap BCa method as an option. There are several areas of research that have not been explored in this manuscript due to financial and time constraints. Future studies will examine the potential impacts of the proposed model on flood risk, financial losses caused by flooding, and its potential for downstream development and wider application.

Author Contributions: Conceptualization, L.L. and Z.Y.; methodology, L.L.; software, K.K.F.L. and L.L.; validation, L.L. and Z.Y.; formal analysis, K.K.F.L. and L.L.; investigation, K.K.F.L., L.L. and Z.Y.; resources, L.L. and Z.Y.; data curation, L.L.; writing—original draft preparation, K.K.F.L. and L.L.; writing—review and editing, K.K.F.L., L.L. and Z.Y.; visualization, L.L.; supervision, L.L.; project administration, L.L. and Z.Y.; funding acquisition, L.L. and Z.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Ministry of Higher Education (MoHE) through the Fundamental Research Grant Scheme (FRGS/1/2021/WAB07/UTAR/02/1) and was partly supported by the Brunsfield Engineering Sdn. Bhd., Malaysia (Brunsfield 8013/0002 & 8126/0001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Please email requests to Ir. Dr. Ling Lloyd at linglloyd@utar.edu.my.

Acknowledgments: The authors appreciate the guidance from R.H. Hawkins at The University of Arizona, Tucson, AZ, USA.

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

Appendix A

Referring to Section 1, the SCS-CN rainfall–runoff model was successfully rearranged into the S general formula, as the S general formula will be the main approach to obtain the S values, based on the available P-Q datasets.

The SCS-CN rainfall–runoff model is defined as followed:

$$Q = \frac{(P - I_a)^2}{(P - I_a + S)}$$

where

Q = Runoff depth (mm).
P = Rainfall depth (mm).
I_a = The initial abstraction amount (mm).
S = Maximum potential water retention of a watershed (mm).

The S general formula was derived based on the conventional SCS CN rainfall–runoff model, where the hypothesis of the linear correlation between I_a and S was applied with the form of I_a = 0.2S. The conventional SCS CN rainfall–runoff model was defined as followed:

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S}$$

where the restriction of P > 0.2S must be obeyed, or else there will be no runoff (Q) occurring.

As shown in Figure 1 in Section 1, the original datasets were plotted in a log–log scaled graph; thus, instead of a linear regressed model, a power regressed model is more suitable to be applied to the SCS CN model to capture the original datasets, according to the power law. In this study, the power regressed model ($I_a = S^\lambda$) had substituted the linear regressed model ($I_a = \lambda S$); thus, the simplification of the SCS CN model will be carried out, as the attempt to obtain a closed form for the S general formula, similar to the previous study. In this derivation, the power regressed model is included in the SCS CN rainfall–runoff model and followed by the simplification of the model to obtain the closed-form S general formula for the new calibrated runoff model:

$$\begin{gathered}
\text{New found correlation in Power form}: \ I_a = S^\lambda \\
Q = \frac{(P - S^\lambda)^2}{(P - S^\lambda + S)} \\
Q\left(P - S^\lambda + S\right) = (P - S^\lambda)^2 \\
QP - QS^\lambda + SQ = P^2 - 2PS^\lambda + S^{2\lambda} \\
QP - P^2 = S^{2\lambda} - 2PS^\lambda + QS^\lambda - SQ \\
QP - P^2 = S^{2\lambda} + (Q - 2P)S^\lambda - SQ \\
QP - P^2 = S^{2\lambda} + (Q - 2P)S^\lambda + \left(\frac{Q-2P}{2}\right)^2 - SQ - \left(\frac{Q-2P}{2}\right)^2 \\
QP - P^2 = (S^\lambda + \frac{Q-2P}{2})^2 - SQ - \left(\frac{Q-2P}{2}\right)^2 \\
(S^\lambda + \frac{Q-2P}{2})^2 = SQ + QP + \left(\frac{Q-2P}{2}\right)^2 - P^2 \\
(S^\lambda + \frac{Q-2P}{2})^2 = \frac{1}{4}\left(4SQ + 4PQ + Q^2 - 4PQ + 4P^2 - 4P^2\right) \\
(S^\lambda + \frac{Q-2P}{2})^2 = \frac{1}{4}\left(4SQ + Q^2\right) \\
(2S^\lambda + Q - 2P)^2 = 4SQ + Q^2
\end{gathered} \tag{A1}$$

Since there is no further simplification of the model that can be proceeded, the simplest form for the S general formula will be Equation (A1).

As for now, there is no closed form for the S general formula obtained from the new calibrated model; therefore, the numerical analysis techniques will be used to obtain the S value based on S general formula. (Equation (A1) is also known as Equation (6) in this article.)

Appendix B

Several attempts had been made by the researchers in the previous research studies to obtain the S general formula, where the linear correlation with the form of $I_a = \lambda S$ is still intact with the SCS CN rainfall–runoff model. The rearrangement of the SCS CN rainfall runoff model, eventually the S general formula, was as follows [7]:

$$S_\lambda = \frac{\left[P - \frac{(\lambda-1)Q}{2\lambda}\right] - \sqrt{PQ - P^2 + \left[P - \frac{(\lambda-1)Q}{2\lambda}\right]^2}}{\lambda} \tag{A2}$$

To obtain the $S_{0.2}$ values, the λ for Equation (A2) was set to be 0.2. When λ = 0.2, eventually the Equation (A2) can be simplified as follows:

$$S_{0.2} = \frac{\left[P - \frac{(0.2-1)Q}{2(0.2)}\right] - \sqrt{PQ - P^2 + \left[P - \frac{(0.2-1)Q}{2(0.2)}\right]^2}}{0.2}$$

$$S_{0.2} = 5\left(\left[P + \frac{(0.8)Q}{0.4}\right] - \sqrt{PQ - P^2 + \left[P + \frac{(0.8)Q}{0.4}\right]^2}\right)$$

$$S_{0.2} = 5\left([P + 2Q] - \sqrt{PQ - P^2 + [P + 2Q]^2}\right)$$

$$S_{0.2} = 5\left([P + 2Q] - \sqrt{PQ - P^2 + P^2 + 4PQ + 4Q^2}\right)$$

$$S_{0.2} = 5\left(P + 2Q - \sqrt{4Q^2 + 5PQ}\right) \quad \text{(A3)}$$

Equation (A3) represents the S general equation where the λ = 0.2, which helps to obtain the $S_{0.2}$ values based on the P and Q values accordingly. Equation (A3) was also the same as the previous research study from Hawkins [6,7,17].

Thus, with Equation (A3), the value for $S_{0.2}$ can be obtained for each P–Q data pair from all the study sites. Then, only the correlation between the $S_{0.2}$ and S_λ for each study site can be determined via SPSS, and, later, the correlation equation will be substituted back into the SCS CN curve number model to derive the $CN_{0.2}$ values for each dataset. (Equation (A3) is also known as Equation (7) in this article.)

Appendix C

Since the power regressed model was introduced into the SCS CN rainfall runoff model in this study, the calibration and the derivation of the $CN_{0.2}$ value has to be adjusted as well as shown below:

$$CN_{0.2} \text{ formula}: \; CN_{0.2} = \frac{25,400}{254 + S_{0.2}}$$

The $CN_{0.2}$ value derivation and the calibration steps were summarized as follows:

1. Given that the effective rainfall (P_e) = P − I_a and I_a = λS, the SCS runoff model ($Q = \frac{P-I_a}{P-I_a+S}$) can be rearranged as follows:

   $$Q = \frac{{P_e}^2}{P_e + S}$$

   where $S = \frac{{P_e}^2}{Q} - P_e$ and $\lambda = \frac{I_a}{S}$.
2. For each P–Q data pair (P_i, Q_i), calculate corresponding λ_i and S_i values via numerical analysis techniques.
3. Perform bootstrap BCa procedure, Shapiro–Wilk, and Kolmogorov–Smirnov tests in SPSS (version 18.0 or an equivalent statistics software) for (λ_i, S_i) to check on the normality of both λ_i and S_i. Check the Shapiro–Wilk and Kolmogorov–Smirnov test results of S_i and λ_i to see whether it is normally distributed or not:
   (a) If yes, refer to the mean BCa confidence interval for S_i and λ_i optimization.
   (b) Otherwise, refer to the median BCa confidence interval for S_i and λ_i optimization.
4. Based on the normality of λ_i and S_i, calculate the optimal value for λ_i and S_i, denoted by $\lambda_{optimum}$ and $S_{optimum}$.

5. Substitute the $\lambda_{optimum}$ and $S_{optimum}$ value into $Q = \frac{P-I_a}{P-I_a+S}$, where $I_a = S^{\lambda}$, to formulate the new SCS CN rainfall–runoff model and calibrate the model according to the given P–Q datasets.
6. Given (P_i, Q_i) data pairs, compute S_i values and $\lambda_{optimum}$ with $(2S_i{}^{\lambda} + Q - 2P)^2 = 4S_iQ + Q^2$ (Equation (A1) or (6)) via Excel's numerical iteration. To date, there is no closed form for the S general equation formula.
7. Given (P_i, Q_i) data pairs and λ = 0.2, compute $S_{0.2i}$ values with Equation (A3) or (7).
8. Correlate $S_{0.2i}$ and S_i to obtain a correlation equation between $S_{0.2i}$ and S_i via SPSS (or an equivalent statistics software).
9. Substitute the S correlation equation from step 9 into the SCS curve number formula $CN_{0.2} = \frac{25{,}400}{S_{0.2}+254}$ to derive $CN_{0.2}$ value for the watershed of interest.

Note: Refer to the example discussion from Equations (8)–(11) in the article.

Appendix D

In Section 3.1, the performance of the newly calibrated predictive models was tabulated in Table 1 for each of the study sites. Tables A1 and A2 below show the comparison of runoff predictions for those recalibrated models and the conventional SCS CN rainfall runoff models (Equation (2)).

Table A1. Power models' runoff prediction comparison for each study site.

Datasets	New SCS (Power) Model			Remark
	E	BIAS	KGE	
DID HP 11, Malaysia	0.823	2.347	0.805	I_a < min rainfall data
DID HP 27, Malaysia	0.919	0.647	0.956	I_a < min rainfall data
DID HP 11+27, Malaysia	0.875	2.531	0.863	I_a < min rainfall data
Kayu Ara, Malaysia	0.810	0	0.865	I_a < min rainfall data
Kerayong, Malaysia	0.888	0	0.818	I_a < min rainfall data
Attica, Greece	0.786	0	0.798	I_a < min rainfall data
Wang Jia Qiao, China	0.795	0	0.739	I_a < min rainfall data

Table A2. Conventional SCS CN models' runoff prediction comparison for each study site.

Datasets	SCS Model (Equation (2))			Remark
	E	BIAS	KGE	
DID HP 11, Malaysia	0.839	−6.104	0.759	I_a > 60 rainfall data (12.7%)
DID HP 27, Malaysia	0.910	−4.085	0.895	I_a > 6 rainfall data (2.6%)
DID HP 11+27, Malaysia	0.879	−4.451	0.863	I_a > 68 rainfall data (9.7%)
Kayu Ara, Malaysia	0.805	−1.162	0.831	I_a > 6 rainfall data (6.5%)
Kerayong, Malaysia	0.891	−0.885	0.843	I_a > 2 rainfall data (2.8%)
Attica, Greece	0.780	−1.846	0.784	I_a > 5 rainfall data (6.6%)
Wang Jia Qiao, China	0.482	1.586	0.418	I_a > 4 rainfall data (13.8%)

Note: λ = 0.2 is not significant at α = 0.05 level for any dataset. SCS CN theory requires I_a < rainfall depth (P).

It is noteworthy to highlight that Equation (2) was not significant even at α = 0.05 level for any dataset in this study. As such, it is not statistically significant to model any rainfall–runoff in this study. The I_a value of all SCS models violated a requirement from the SCS CN theory that I_a < the rainfall depth (P). Table A2 tabulates the total count of rainfall–runoff data pairs where the recorded P value is less than the SCS model's I_a value. In contrast, none of the newly formulated power models have this issue. The conventional SCS runoff model (Equation (2)) has a model bias ranging from -6.104 to 1.586 in this study, indicating inconsistency in its runoff predictions with an underprediction concern. In the context of climate change, the conventional SCS runoff predictive model is not reliable for accurately estimating runoff amounts.

Appendix E

In Section 2.2, this study had mentioned that there will be six study sites from three countries in the study, which are from Malaysia, Greece, and China. The geographic coordinates of each catchment area that was included in those study sites had been obtained and tabulated in Table A3.

Table A3 shows the geographic coordinates of each catchment area that was included in the study sites from three different countries.

Table A3. The geographic coordinates of each catchment area that was included in the study sites from three different countries.

No.	Catchment Locations	Latitude	Longitude
1	Attica, Greece	38° 4′ N	23° 50′ E
2	Wangjiaqiao, China	31° 8′ N	111° 41′ E
	Peninsula Malaysian Catchments		
1	Kayu Ara, Malaysia	3° 8′ N	101° 37′ E
2	Kerayong, Malaysia	3° 6′ N	101° 42′ E
3	Parit Madirono di Weir	1° 41′ N	103° 16′ E
4	Sg. Johor di Rantau Panjang	1° 46′ N	103° 44′ E
5	Sg. Sayong di Jam. Johor Tenggara	1° 48′ N	103° 40′ E
6	Sg. Kahang di Bt.26 Jln. Kluang	2° 15′ N	103° 35′ E
7	Sg. Lenggor di Bt.42 Kluang/Mersing	2° 15′ N	103° 44′ E
8	Sg. Muar di Buloh Kasap	2° 33′ N	102° 45′ E
9	Sg. Serting di Jam.Padang Gudang	3° 06′ N	102° 28′ E
10	Sg. Triang di Jam. Keretapi	3° 14′ N	102° 24′ E
11	Sg. Bentong di Kuala Marong	3° 30′ N	101° 54′ E
12	Sg. Lepar di Jam. Gelugor	3° 41′ N	102° 58′ E
13	Sg. Kuantan di Bukit Kenau	3° 55′ N	103° 03′ E
14	Sg. Lipis di Benta	4° 01′ N	101° 57′ E
15	Sg. Cherul di Ban Ho	4° 08′ N	103° 10′ E
16	Sg. Kemaman di Rantau Panjang	4° 16′ N	103° 15′ E
17	Sg. Dungun di Jam. Jerangau	4° 50′ N	103° 12′ E
18	Sg. Berang di Menerong	4° 56′ N	103° 03′ E
19	Sg. Telemong di Paya Rapat	5° 10′ N	102° 54′ E
20	Sg. Lebir di Kg. Tualang	5° 16′ N	102° 16′ E
21	Sg. Nerus di Kg. Bukit	5° 17′ N	102° 55′ E
22	Sg. Chalok di Jam. Chalok	5° 26′ N	102° 50′ E
23	Sg. Lanas di Air Lanas	5° 47′ N	101° 53′ E
24	Sg. Besut di Jambatan Jerteh	5° 44′ N	102° 29′ E
25	Sg. Pelarit di Titi Baru	6° 35′ N	100° 12′ E
26	Sg. Buloh di Kg. Batu Tangkup	6° 33′ N	100° 17′ E
27	Sg. Kulim di Ara Kuda	5° 26′ N	100° 30′ E
28	Sg. Kerian di Selama	5° 13′ N	100° 41′ E
29	Sg. Plus di Kg. Lintang	4° 56′ N	101° 06′ E
30	Sg. Raia di Keramat Pulai	4° 32′ N	101° 08′ E
31	Sg. Kampar di Kg. Lanjut	4° 20′ N	101° 06′ E
32	Sg. Bidor di Bidor Malayan Tin Bhd.	4° 04′ N	101° 14′ E
33	Sg. Sungkai di Sungkai	3° 59′ N	101° 18′ E
34	Sg. Slim At Slim River	3° 49′ N	101° 24′ E
35	Sg. Bernam di Tanjung Malim	3° 40′ N	101° 31′ E
36	Sg. Selangor di Rasa	3° 30′ N	101° 38′ E
37	Sg. Gombak di Damsite	3° 14′ N	101° 42′ E
38	Sg. Batu di Kg. Sg. Tua	3° 16′ N	101° 41′ E
39	Sg. Lui di Kg. Lui	3° 10′ N	101° 52′ E
40	Sg. Langat di Dengkil	2° 59′ N	101° 47′ E
41	Sg. Linggi di Sua Betong	2° 40′ N	101° 55′ E
42	Sg. Melaka di Pantai Belimbing	2° 20′ N	102° 15′ E
43	Sg. Kesang di Chin Chin	2° 17′ N	102° 29′ E
44	Sg Sembrong di Bt 2 Air Hitam, Yong Peng	2° 4′ N	103° 22′ E

Table A3. *Cont.*

No.	Catchment Locations	Latitude	Longitude
45	Sg Segamat di Segamat	2° 31′ N	102° 51′ E
46	Sg Sayong di Johor Tenggara	1° 48′ N	103° 35′ E
47	Sg Muar di Bt 57 Jln GemasRompin	2° 25′ N	102° 30′ E
48	Sg Lepar di Jam Gelugor	3° 43′ N	102° 56′ E
49	Sg Lenggor di Bt 42 KluangMersing	2° 12′ N	103° 41′ E
50	Sg Kuantan di Bkt Kenau	3° 53′ N	103° 8′ E
51	Sg Kepis di Jam Kayu Lama	2° 41′ N	102° 20′ E
52	Sg Kemaman di Rantau Panjang	4° 15′ N	103° 16′ E
53	Sg Kecau di Kg Dusun	4° 22′ N	102° 6′ E
54	Sg Kahang di Bt 26 Jln Kluang	2° 10′ N	103° 31′ E
55	Sg Johor di Rantau Panjang	1° 37′ N	103° 54′ E
56	Sg Cherul di Ban Ho	4° 10′ N	103° 8′ E
57	Sg Berang di Menerong	4° 57′ N	103° 0′ E
58	Sg Bentong di Jam K Marong	3° 31′ N	101° 55′ E
59	Sg Bekok di Bt 77 Jln Yong Peng/Labis	2° 7′ N	103° 6′ E
60	Sg Sungkai di Sungkai	4° 2′ N	101° 18′ E
61	Sg Raia di Keramat Pulai	4° 35′ N	101° 14′ E
62	Sg Pari di Jln Silibin, Ipoh	4° 36′ N	101° 4′ E
63	Sg Semenyih di Sg Rinching	2° 56′ N	101° 50′ E
64	Sg Selangor di Rasa	3° 27′ N	101° 27′ E
65	Sg Plus di Kg Lintang	4° 56′ N	101° 9′ E
66	Sg Pelarit di Wang Mu	6° 34′ N	100° 13′ E
67	Sg Melaka di Pantai Belimbing	2° 20′ N	102° 14′ E
68	Sg Lui di Kg Lui	3° 9′ N	101° 54′ E
69	Sg Linggi di Sua Betong	2° 37′ N	101° 60′ E
70	Sg Langat di Dengkil	2° 58′ N	101° 38′ E
71	Sg Kurau di Pondok Tg	4° 59′ N	100° 32′ E
72	Sg Kulim di Ara Kuda	5° 23′ N	100° 32′ E
73	Sg Kerian di Selama	5° 12′ N	100° 38′ E
74	Sg Kinta di Weir G, Tg Tualang	4° 21′ N	101° 3′ E
75	Sg Kesang di Chin Chin	2° 17′ N	102° 31′ E
76	Sg Durian Tunggal di Bt 11 Air Resam	2° 19′ N	102° 17′ E
77	Sg Cenderiang di Bt 32 Jln Tapah	4° 15′ N	101° 10′ E
78	Sg Bidor di Bidor Malayan Tin Bhd	4° 2′ N	101° 11′ E
79	Sg Bernam di Tg Malim	3° 46′ N	101° 3′ E
80	Sg Selangor di Rantau Panjan	3° 27′ N	101° 27′ E
	Sabah State, Malaysian Catchments		
1	Sg Tawau di Kuhara	4° 16′ N	117° 53′ E
2	Sg Kalabakan di Kalabakan	4° 27′ N	117° 23′ E
3	Sg Kalumpang di Mostyn Bridge	4° 38′ N	118° 9′ E
4	Sg Talangkai di Lotong	4° 43′ N	116° 26′ E
5	Sg Mengalong di Sindumin	4° 59′ N	115° 34′ E
6	Sg Kuamut di Ulu Kuamut	5° 4′ N	117° 26′ E
7	Sg Lakutan di Mesapol Quarry	5° 7′ N	115° 37′ E
8	Sg Segama di Limkabong	5° 7′ N	118° 7′ E
9	Sg Sook di Biah	5° 15′ N	116° 8′ E
10	Sg Baiayo di Bandukan	5° 26′ N	116° 8′ E
11	Sg Apin-Apin di Waterworks	5° 29′ N	116° 15′ E
12	Sg Kegibangan di Tampias P.H.	5° 41′ N	116° 22′ E
13	Sg Papar di Kaiduan	5° 46′ N	116° 5′ E
14	Sg Papar di Kogopon	5° 42′ N	116° 2′ E
15	Sg Labuk di Tampias	5° 43′ N	116° 51′ E
16	Sg Moyog di Penampang	5° 54′ N	116° 6′ E
17	Sg Tungud di Basai	6° 3′ N	117° 18′ E
18	Sg Tuaran di Pump House 1	6° 9′ N	116° 14′ E
19	Sg Sugut di Bukit Mondou	6° 11′ N	117° 14′ E
20	Sg Kadamaian di Tamu Darat	6° 15′ N	116° 27′ E

Table A3. *Cont.*

No.	Catchment Locations	Latitude	Longitude
21	Sg Wariu di Bridge No.2	6° 19′ N	116° 29′ E
22	Sg Bongan di Timbang Batu Sabah	6° 26′ N	116° 48′ E
23	Sg Bengkoka di Kobon	6° 37′ N	117° 2′ E
	Sarawak State, Malaysian Catchments		
1	Sg Kayan di Krusen	1° 4′ N	110° 29′ E
2	Sg Kedup di New Meringgu	1° 3′ N	110° 33′ E
3	Sg Entebar di Entebar	1° 0′ N	111° 32′ E
4	Sg Ai di Lubok Antu	1° 2′ N	111° 49′ E
5	Sg Sabal Kruin di Sabal Kruin	1° 8′ N	110° 53′ E
6	Sg Sarawak Kanan di Pk Buan Bidi	1° 23′ N	110° 6′ E
7	Sg Sarawak Kiri di Kg Git	1° 21′ N	110° 15′ E
8	Sg Tuang di Kg Batu Gong	1° 20′ N	110° 26′ E
9	Sg Sekerang di Entaban	1° 19′ N	111° 37′ E
10	Sg Layar di Ng Lubau	1° 29′ N	111° 35′ E
11	Sg Sebatan di Sebatan	1° 48′ N	111° 20′ E
12	Sg Katibas di Ng Mukeh	1° 50′ N	112° 37′ E
13	Sg Sarikei di Ambas	1° 58′ N	111° 30′ E
14	Btg Rajang di Ng Ayam	1° 56′ N	111° 53′ E
15	Sg Oya di Setapang	3° 1′ N	112° 35′ E
16	Btg Mukah di Selangau	2° 54′ N	112° 5′ E
17	Sg Sibiu di Sibiu (Atc)	3° 13′ N	113° 9′ E
18	Sg Limbang di Insungai	4° 44′ N	114° 59′ E
19	Sg Trusan di Long Tengoa D	4° 35′ N	115° 20′ E

References

1. Danáčová, M.; Földes, G.; Labat, M.M.; Kohnová, S.; Hlavčová, K. Estimating the Effect of Deforestation on Runoff in Small Mountainous Basins in Slovakia. *Water* **2020**, *12*, 3113. [CrossRef]
2. Li, Y.; Liu, C.; Zhang, D.; Liang, K.; Li, X.; Dong, G. Reduced Runoff due to Anthropogenic Intervention in the Loess Plateau, China. *Water* **2016**, *8*, 458. [CrossRef]
3. Zhang, Y.; Xia, J.; Yu, J.; Randall, M.; Zhang, Y.; Zhao, T.; Pan, X.; Zhai, X.; Shao, Q. Simulation and Assessment of Urbanization Impacts on Runoff Metrics: Insights from Landuse Changes. *J. Hydrol.* **2018**, *560*, 247–258. [CrossRef]
4. Zhai, R.; Tao, F. Contributions of Climate Change and Human Activities to Runoff Change in Seven Typical Catchments across China. *Sci. Total Environ.* **2017**, *605–606*, 219–229. [CrossRef] [PubMed]
5. Li, C.; Liu, M.; Hu, Y.; Shi, T.; Qu, X.; Walter, M.T. Effects of Urbanization on Direct Runoff Characteristics in Urban Functional Zones. *Sci. Total Environ.* **2018**, *643*, 301–311. [CrossRef]
6. Tedela, N.H.; McCutcheon, S.C.; Rasmussen, T.C.; Hawkins, R.H.; Swank, W.T.; Campbell, J.L.; Adams, M.B.; Jackson, C.R.; Tollner, E.W. Runoff Curve Numbers for 10 Small Forested Watersheds in the Mountains of the Eastern United States. *J. Hydrol. Eng.* **2012**, *17*, 1188–1198. [CrossRef]
7. Ling, L.; Yusop, Z.; Yap, W.-S.; Tan, W.L.; Chow, M.F.; Ling, J.L. A Calibrated, Watershed-Specific SCS-CN Method: Application to Wangjiaqiao Watershed in the Three Gorges Area, China. *Water* **2019**, *12*, 60. [CrossRef]
8. N.E.D.C. Engineering Hydrology Training Series. In *Module 205-SCS Runoff Equation*; N.E.D.C.: London, UK, 1997; Available online: https://d32ogoqmya1dw8.cloudfront.net/files/geoinformatics/steps/nrcs_module_runoff_estimation.pdf (accessed on 8 July 2022).
9. Soil Conservation Service (S.C.S.). *National Engineering Handbook*; US Soil Conservation Service: Washington, DC, USA, 1964; Chapter 10; Section 4. Available online: https://directives.sc.egov.usda.gov/RollupViewer.aspx?hid=17092 (accessed on 10 July 2022).
10. USDA; NRCS. *National Engineering Handbook, Part 630 Hydrology*; US Soil Conservation Service: Washington, DC, USA, 1964; Chapter 10.
11. Mishra, S.K.; Babu, P.S.; Singh, V.P. SCS-CN Method Revisited. In *Advances in Hydraulics and Hydrology*; Water Resources Publications: Littleton, CO, USA, 2007.
12. Tan, W.J.; Ling, L.; Yusop, Z.; Huang, Y.F. New Derivation Method of Region Specific Curve Number for Urban Runoff Prediction at Melana Watershed in Johor, Malaysia. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2018; Volume 401, p. 012008. [CrossRef]
13. Yuan, L.; Sinshaw, T.; Forshay, K.J. Review of Watershed-Scale Water Quality and Nonpoint Source Pollution Models. *Geosciences* **2020**, *10*, 25. [CrossRef]
14. Hawkins, R.H.; Yu, B.; Mishra, S.K.; Singh, V.P. Another Look at SCS-CN Method. *J. Hydrol. Eng.* **2001**, *6*, 451–452. [CrossRef]

15. Hawkins, R.; Ward, T.J.; Woodward, E.; van Mullem, J.A. Continuing Evolution of Rainfall-Runoff and the Curve Number Precedent. In Proceedings of the 2nd Joint Federal Interagency Conference, Las Vegas, NV, USA, 27 June–1 July 2010; pp. 1–12.
16. Tan, W.J.; Ling, L.; Yusop, Z.; Huang, Y.F. Claim Assessment of a Rainfall Runoff Model with Bootstrap. In Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017), Langkawi, Malaysia, 7–8 November 2017; Springer Nature: Singapore, 2019.
17. Hawkins, R.H.; Khojeini, A.V. Initial Abstraction and Loss in the Curve Number Method. *Hydrology and Water Resources in Arizona and the Southwest* **2000**, *30*, 29–35.
18. DID. *Hydrological Procedure 27 Design Flood Hydrograph Estimation for Rural Catchments in Malaysia*; JPS, DID: Kuala Lumpur, 2010. Available online: https://www.water.gov.my/jps/resources/PDF/Hydrology%20Publication/Hydrological_Procedure_No_27_(HP_27).pdf (accessed on 5 August 2022).
19. DID. *Hydrological Procedure 11 Design Flood Hydrograph Estimation for Rural Catchments in Malaysia*; JPS, DID: Kuala Lumpur, 2018. Available online: http://h2o.water.gov.my/man_hp1/HP11.pdf (accessed on 20 July 2022).
20. Abustan, I.; Sulaiman, A.H.; Wahid, N.A.; Baharudin, F. Determination of Rainfall-Runoff Characteristics in An Urban Area: Sungai Kerayong Catchment, Kuala Lumpur. In Proceedings of the 11th International Conference on Urban Drainage, Palermo, Italy, 23–26 September 2018; Springer International Publishing: New York, NY, USA, 2018.
21. Ling, L.; Chow, M.F.; Tan, W.L.; Tan, W.J.; Tan, C.Y.; Yusop, Z. New Regional-Specific Urban Runoff Prediction Model of Sungai Kayu Ara Catchment in Malaysia. In *Lecture Notes in Civil Engineering*; Springer Nature: Singapore, 2020; Volume 59, pp. 161–168.
22. Baltas, E.A.; Dervos, N.A.; Mimikou, M.A. Technical Note: Determination of the SCS Initial Abstraction Ratio in an Experimental Watershed in Greece. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1825–1829. [CrossRef]
23. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman and Hall/CRC: New York, NY, USA, 1993; ISBN 978-0-412-04231-7.
24. Davison, A.C. *Cambridge Series in Statistical and Probabilistic*; Cambridge University Press: New York, NY, USA, 2003; ISBN 978-0-511-67299-6.
25. Efron, B. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*; Cambridge University Press: New York, NY, USA, 2013; ISBN 978-1-107-61967-8.
26. Rochowicz, J.A.J. Bootstrapping Analysis, Inferential Statistics and EXCEL. *Spredsheets Educ. (Ejsie)* **2010**, *4*, 1–23.
27. Moriasi, D.N.; Arnold, J.G.; van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]
28. Nash, J.E.; Sutcliffe, J.V. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
29. Ling, L.; Yusop, Z.; Ling, J.L. Statistical and Type II Error Assessment of a Runoff Predictive Model in Peninsula Malaysia. *Mathematics* **2021**, *9*, 812. [CrossRef]
30. Knoben, W.J.M.; Freer, J.E.; Woods, R.A. Technical Note: Inherent Benchmark or Not? Comparing Nash—Sutcliffe and Kling—Gupta Efficiency Scores. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4323–4331. [CrossRef]
31. Miles, J. R Squared, Adjusted R Squared. In *Wiley StatsRef: Statistics Reference Online*; Wiley: Hoboken, NJ, USA, 2014.
32. Hawkins, R.H.; Theurer, F.D.; Rezaeianzadeh, M. Understanding the Basis of the Curve Number Method for Watershed Models and TMDLs. *J. Hydrol. Eng.* **2019**, *24*, 06019003. [CrossRef]
33. Ling, L.; Yusop, Z. Derivation of Region-Specific Curve Number for an Improved Runoff Prediction. In *Improving Flood Management, Prediction and Monitoring: Case Studies in Asia*; Community, Environment and Disaster Risk Management; Emerald Publishing Limited: Bingley, UK, 2018; Volume 20, pp. 37–48. ISBN 978-1-78756-552-4.
34. Ling, L.; Yusop, Z.; Chow, M.F. Urban Flood Depth Estimate with a New Calibrated Curve Number Runoff Prediction Model. *IEEE Access* **2020**, *8*, 10915–10923. [CrossRef]
35. Mosavi, A.; Ozturk, P.; Chau, K. Flood Prediction Using Machine Learning Models: Literature Review. *Water* **2018**, *10*, 1536. [CrossRef]
36. Pinos, J.; Quesada-Román, A. Flood Risk-Related Research Trends in Latin America and the Caribbean. *Water* **2021**, *14*, 10. [CrossRef]
37. Fu, M.; Fan, T.; Ding, Z.; Salih, S.Q.; Al-Ansari, N.; Yaseen, Z.M. Deep Learning Data-Intelligence Model Based on Adjusted Forecasting Window Scale: Application in Daily Streamflow Simulation. *IEEE Access* **2020**, *8*, 32632–32651. [CrossRef]
38. Kaya, C.M.; Tayfur, G.; Gungor, O. Predicting Flood Plain Inundation for Natural Channels Having No Upstream Gauged Stations. *J. Water Clim. Change* **2019**, *10*, 360–372. [CrossRef]