NEW DERIVATION METHOD OF REGION SPECIFIC CURVE NUMBER FOR URBAN RUNOFF PREDICTION AT MELANA WATERSHED IN JOHOR, MALAYSIA

W J Tan¹, L Ling¹*, Z Yusop² and Y F Huang¹

¹Centre of Disaster Risk Reduction, Lee Kong Chian Faculty of Engineering and Science, Civil Engineering Department of Universiti Tunku Abdul Rahman, Jalan Sungai Long, Bandar Sungai Long, 43000 Cheras, Malaysia

²Centre for Environmental Sustainability and Water Security, Faculty of Civil Engineering Department, Universiti Teknologi Malaysia. 81310, Skudai, Johor, Malaysia

*Corresponding Author: linglloyd@utar.edu.my

Abstract. The Soil Conservation Services (SCS) rainfall-runoff model has been incorporated into many types of software and adopted by Malaysian government agencies for design use. However, hydrologists argue the accuracy of the predicted runoff results from this model in recent decades. Malaysia does not have a regional specific curve numbers available for the use in rainfall-runoff modelling, and therefore SCS-CN practitioner has no option but to adopt its guideline and handbook values which are specific to US region. Unlike the conventional SCS-CN technique, the proposed calibration methodology in this paper discarded the use of CN as input to the SCS model, rearranged the model equation and derived statistically significant CN value(s) of a specific region through rainfall-runoff events directly through the use of inferential statistics. The SCS rainfall-runoff framework can now be calibrated quickly to address urban runoff prediction challenge under rapid land use and land cover or climate changes with the proposed methodology. Between July and October of 2004, the derived λ was 0.0005 while λ = 0.20 was found to be statistically insignificant at alpha = 0.01 level at Melana watershed in Johor, Malaysia. Optimum CN of 90.45 was derived to represent the rainfall-runoff characteristic of Melana watershed. Runoff prediction error was reduced by 90% and achieved the Nash-Sutcliffe (NS) index of 0.86. The new method can provide CN adjustment guidelines for SCS practitioners and any software which incorporated the model to predict urban runoff.

Keywords: Bootstrapping, Curve Number, Runoff prediction, Inferential Statistics
1. Introduction

The United States Department of Agriculture (USDA), Soil Conservation Services (SCS) agency developed the curve number (CN) system in 1954. It’s Technical Release 55 (TR-55) classified site conditions with CN. However, researchers reported inconsistent runoff prediction results by using CN and SCS runoff model throughout the world in recent decades [3, 11, 22, 23]. Hawkins (1984) [6] reported that forested watersheds have the highest risk in CN misclassification while inappropriate CN selection was identified as the main cause to produce inconsistent runoff estimates. According to Soulis and Valiantzas (2013) [19], rainfall-runoff (P-Q) dataset should be the base to derive CN. Several researches even reported that CN was a random variable between storm events [9, 20] while watershed with several soil groups could not be represented by a lumped CN value but SCS practitioners seldom attempt to calibrate the model and subjectively selected CN value in their studies. [15].

2. Study Site and Methodology

Instead of selecting CN, this study used a new approach to calibrate SCS runoff predictive model through the use of inferential statistics and P-Q dataset to derive CN for the study site. SCS model was introduced in 1954 as:

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

where

- \( Q \) = Runoff depth (mm)
- \( P \) = Rainfall depth (mm)
- \( I_a \) = The initial abstraction amount (mm)
- \( S \) = Watershed potential water retention (mm)

The initial abstraction is the rainfall retention prior to beginning of runoff process. SCS hypothesized that \( I_a = \lambda S \) where \( \lambda = 0.20 \). Substitution of \( I_a = 0.20S \) simplifies Equation (1) into a form listed below:

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

where \( P > 0.2S \) or else \( Q = 0 \). Worldwide study results in recent decades showed that Equation (2) failed to produce reliable results while \( \lambda = 0.20 \) was also under scrutiny [3, 11, 22, 23]. This study adopted the rainfall-runoff dataset from Chan (2005) [1] which was carried out in Melana watershed to demonstrate the SCS model calibration and CN derivation technique. Inferential statistics was used to derive two key parameters (\( \lambda \) and \( S \) value) from Equation (1) according to the given P-Q dataset. The watershed is located in Johor, Malaysia. The watershed locates between 1º 30’ N to 1º 35’ N and 103º 35’ E to 103º 39’ E (Figure 1). Melana River flows from Gunung Pulai in the northern part through the watershed which covers an area of 21.12 km². It was projected that more than 60% of the area would be populated to become residential area by 2010 [14].
This study used twenty-seven P-Q dataset collected between July and October of 2004 in order to limit the effect due to rapid land use and land cover change. Bootstrapping BCa procedure was selected for its robustness nature while the inferential ability of its confidence interval can used for the assessments of the 1954 SCS proposal through the Null hypothesis stated below[2, 21]:

**Null Hypothesis 1 (H₀):** Equation (2) is valid at the Melana watershed where \( \lambda = 0.20 \).

H₀ must be accepted in order for Equation (2) to be used to model runoff conditions at Melana watershed. Rejection of H₀ makes it necessary to derive new regional \( \lambda \) value because the initial SCS proposal of \( \lambda = 0.2 \) fails. The P-Q data pairs were used to derive \( S \) and \( \lambda \) values from Equation (1). The effective rainfall depth \( (P_e) \) is the difference between the rainfall depth \( (P) \) and initial abstraction \( (I_a) \) thus \( P - I_a = P_e \) [7-8]. Substitute \( P_e \) into Equation (1), it can be re-arranged to calculate corresponding \( S \) and \( \lambda \) values for respective P-Q data pair. Non-parametric Bootstrapping technique, BCa procedure with 2,000 random sampling [4] was conducted to produce the 99% confidence interval (CI) of the derived \( \lambda \) and \( S \) pairs [16]. Supervised numerical optimisation technique was used to select the best \( \lambda \) and \( S \) value within the 99% BCa CI [5] in order formulate a new calibrated SCS runoff prediction model for Melana watershed. The runoff prediction accuracy of the new calibrated SCS runoff prediction model will be assessed and benchmarked against Equation (2) in order to quantify the runoff prediction improvement due to the model calibration and CN derivation.

![Melana watershed and P-Q graph](image_url)

**Figure 1.** Melana watershed [1] and its P-Q graph.

Equation (1) can be rearranged to solve for \( S = f(P, Q, \lambda) \). Substitute \( I_a = \lambda S \) into Equation (1) to solve for \( S \) will produce the \( S \) general formula. If new derived \( \lambda \) value is not equal to 0.20, \( S \) general formula can be used to calculate the \( S_{1} \) value which is different from \( S_{0.2} \) (where \( \lambda = 0.2 \)). Any other \( \lambda \) value will result in \( S_{1} \) which cannot be used to calculate CN. A correlation must be identified between...
Sₖ and S₀.₂ in order to derive the conventional CN₀.₂ value with the CN formula (CN = \(\frac{25,400}{S + 254}\)) which was proposed by SCS [7-8, 10]. The general Sₖ formula which we solved is listed below:

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

(3)

With the given P-Q data pairs and \(\lambda = 0.2\), corresponding series of S₀.₂ values were calculated with Equation (3). Bootstrapping (BCa) procedure was used again to generate a 99% CI of the S₀.₂ values while the supervised numerical optimisation technique identified the optimum S₀.₂ value as the best collective representation for the entire school of S₀.₂. The substitution of the optimum S₀.₂ value into \(CN = \frac{25,400}{S + 254}\) produced the SCS CN₀.₂ to represent Melana watershed.

2.1. Runoff Models Assessment

Nash-Sutcliffe model’s prediction efficiency (NS), model residual sum of squares (RSS) and model BIAS were calculated in order to draw comparison between the calibrated and SCS runoff models with following formulas:

\[
NS = 1 - \frac{RSS}{\sum_{i=1}^{n}(Q_{predicted} - Q_{mean})^2}
\]

(4)

\[
RSS = \sum_{i=1}^{n}(Q_{predicted} - Q_{observed})^2
\]

(5)

\[
BIAS = \frac{\sum_{i=1}^{n}(Q_{predicted} - Q_{observed})}{n}
\]

(6)

where

\(n\) = Total number of data pairs

RSS value indicates the model prediction error. Nash-Sutcliffe value (NS) ranges from minus 1.0 to 1.0 while value of 1.0 indicates a perfect model. In the event when NS < 0, using the average value of the observed data can predict the dataset better than the model. Zero BIAS value indicates a perfect model prediction with no error while the positive value indicates the runoff over-prediction tendency of a model and vice versa.

3. Results and Discussion

3.1. Null Hypothesis Assessment

The descriptive statistics of all derived \(\lambda\) values was tabulated in Table 1. Due to the skewed \(\lambda\) dataset nature, the supervised \(\lambda\) optimization study was conducted within the median confidence interval [0.0002, 0.0005] (Table 1). The best \(\lambda\) value was identified to be 0.0005 to formulate the optimum rainfall-runoff predictive model according to the given P-Q dataset from Melana watershed.

| Table 1. 99% CI Bootstrapping BCa results of all derived \(\lambda\) values |
|----------------------------------|------------------|------------------|
| Melana dataset Descriptive Statistics (\(\lambda\)) 99% BCa | Descriptive Statistics (S) 99% BCa |
|----------------------------------|------------------|------------------|
The skewness of $S$ dataset can be considered as normally distributed because its value is near to zero and therefore, the $S$ optimisation search was conducted within its mean confidence interval from 41.8377 to 86.2105 mm. The best collective $S$ value was identified as 86.2105 mm. As $I_a = \lambda S$ the optimum $\lambda$ and $S$ value yields $I_a = 0.0431$ mm. The calibrated SCS rainfall-runoff prediction model for Melana watershed was formulated as:

$$Q_{0.0005} = \frac{(P - 0.0431)^2}{P + 86.167} \quad (7)$$

where

$Q_{0.0005}$ = Runoff depth (mm) of $\lambda = 0.0005$

As Equation (7) was formed with the optimum $\lambda$ and $S$ value, it will have the same significant level at alpha = 0.01. Not only the standard deviation of $\lambda$ dataset is not equal to zero but also with high fluctuation percentages hence, $\lambda$ cannot be a constant as proposed by SCS. None of the $\lambda$ BCa CI span includes value of 0.2 (Table 1) thus $\lambda$ cannot be equal to 0.2 for this site. $H_01$ can be rejected at alpha = 0.01 level leading to the conclusion that Equation (2) is statistically insignificant and invalid to model runoff at this site.

3.2. The Correlation between $S_{0.0005}$ and $S_{0.2}$ for Melana Watershed

$S_{0.0005}$ and $S_{0.2}$ can be calculated for the $P-Q$ dataset using equation (3) through the substitution of respective $\lambda = 0.0005$ and 0.20 corresponding to the same $P-Q$ dataset for Melana watershed. Equation (8) has an adjusted $R^2$ of 0.924, low standard error of 0.457 and $p < 0.001$. Through the substitution of the $S_{0.2}$ parameter in SCS-CN formula ($CN = \frac{25 \cdot 400}{S_{0.2} + 254}$), the best representative CN value of 90.45 was derived to represent the given rainfall-runoff condition within 99% CN CI range from 90.45 to 95.12. A statistical significant correlation between $S_{0.0005}$ and $S_{0.2}$ was identified as:

$$S_{0.2} = 0.314S_{0.0005}^{0.998} \quad (8)$$

where

$S_{0.015}$ = Total abstraction amount (mm) of $\lambda = 0.015$
$S_{0.2}$ = Total abstraction amount (mm) when $\lambda = 0.2$
As shown in Figure 3, Optimum CN of 90.45 formulated a runoff predictive model with lowest RSS and the highest NS value at α = 0.01 level. When λ = 0.2, the best collective $S_{0.2} = 14.855$ mm for SCS model. The calculated SCS CN value = 94.47. The substitution of $S_{0.2}$ value into Equation (2) formulated the conventional SCS runoff model.

### 3.3. Runoff Predictive Models Comparison

Runoff models’ prediction results were compared in Table 2. Equation (7) has higher NS value and lower RSS compared to Equation (2). Both model’s residual were skewed thus the median residual value was used to assess predictive model’s accuracy. Equation (7) has the model residual value nearest to zero and the smaller 99% BCa confidence interval range to indicate that the model is capable of predicting near to perfect (zero residual) runoff prediction.

In contrast, Equation (2) has the tendency to over predict runoff amount (median residual confidence interval range spans within positive figures only). Equation (7) also has lower standard deviation and variance in its model’s residual. It also has smaller residual confidence interval range and therefore, it is more stable and reliable while Non-parametric correlation statistics (Kendall’s tau b and Spearman’s rho) between the predictive model’s runoff prediction results and the observed dataset of Equation (7) are also larger than Equation (2) at alpha = 0.01 level to infer that Equation (7) is capable to produce better runoff prediction results than Equation (2).

| Predictive model’s     | Equation (7) | SCS: Equation (2) |
|------------------------|--------------|-------------------|
| NS                     | 0.86         | -0.36             |
| RSS                    | 198.15       | 1970.12           |
| BIAS                   | 0.19         | 4.93              |
| CN_{0.2}               | 90.45        | 94.47             |
| λ                       | 0.0005       | 0.20              |
| Residual Skewness      | -1.73        | 1.29              |
| Residual Kurtosis      | 9.52         | 0.76              |
| Median Residual        | -0.05        | 0.47              |
| Median Residual: 99% BCa CI | [-0.231, -0.008] | [-0.060, 7.83] |
| Model Residual Standard Deviation | 2.75 | 7.11 |
| Residual Variance      | 7.59         | 50.56             |
| Correlation to Observed dataset: | | |
| Kendall’s tau b        | 0.897**      | 0.622**           |
| Spearman’s rho         | 0.974**      | 0.811**           |

**Correlation is significant at the 0.01 level (2-tailed)**

On average, Equation (2) over-predicted runoff depth from 7 mm in this study and further magnified toward higher rainfall depths.

### 4. Conclusions
This study agreed with previous research conclusion that the SCS runoff prediction model can be calibrated with regional specific P-Q characteristics. Statistical assessment rejected $H_0$ with 99% confident level. As such, Equation (2) became invalid to model runoff conditions in this study while adoption of Equation (2) will commit a type II error.

This study presented a new CN derivation approach with supervised numerical optimization technique through the guide of inferential statistics. The new calibrated runoff predictive model outperformed SCS runoff model. Residual sum of square (RSS) was reduced by 90%. In comparison to SCS runoff predictive model, it also has higher model efficiency ($NS$), lower $RSS$ and $BIAS$ to produce the smallest runoff prediction error at $\alpha = 0.01$ level. CN value of 90.45 was derived to represent the P-Q conditions at Melana watershed (between July and October of 2004) with the inherent statistical significance at $\alpha = 0.01$ level.

Direct CN derivation with rainfall-runoff conditions is a swift and economical solution to calibrate SCS runoff model to reflect the latest P-Q condition of any watershed. In the event of rapid urbanisation and climate change, SCS practitioners are able to conduct regional specific calibration for SCS model as proposed to derive region and time period specific CN values for Malaysian watershed.

5. References

[1] Chan C H 2005 *Modelling Watershed Cumulative Effects On Runoff Based On Major Land Uses*

[2] Cox D R 2006 *Principles of Statistical Inference* Cambridge University Press

[3] D’Asaro F, Grillone G and Hawkins R H 2014 *J. Hydrol. Eng.* 19 04014035 (1–13)

[4] Efron B 2010 *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction.* (Institute of Mathematical Statistics Monographs/Cambridge University Press) ISBN 9780521192491

[5] Fattorini H O 1999 *Infinite Dimensional Optimization and Control Theory* (Cambridge Univ. Press 1999)

[6] Hawkins R H 1984 *A comparison of predicted and observed runoff curve numbers* In: Proceeding of ASCE Irrigation and Drainage Division special conference Flagstaff AZ: ASCE pp 702–709.

[7] Hawkins R H, Ward T, Woodward D E and Van Mullem J 2009 *Curve Number Hydrology: State of the Practice.* Reston, Virginia, ASCE

[8] Hawkins R H 2014 e-mail communication

[9] Hjelmfelt A T 1991 *J. Hydraul. Eng.* 117 725-737
[10] Jiang R 2001 *Investigation of Runoff Curve Number, Initial Abstraction Ratio* (University of Arizona)

[11] Ling L and Yusop Z 2013 *A Micro Focus with Macro Impact: Exploration of Initial Abstraction Coefficient Ratio (λ) in Soil Conservation Curve Number (CN) Methodology* IOP Conf.Series: Earth and Environmental Science 18 p 012121

[12] Ling L and Yusop Z 2014 *Inferential Statistics of Claim Assessment* AIP Conference Proceedings 1635 p 805

[13] Ling L and Yusop Z 2014b *Inferential Statistics Modelling and Claim Re-assessment* ICCEMS Conference Proceedings pp 835-840

[14] MPJBT 2001 *Rancangan Tempatan Skudai (1993-2010)* Johor Bahru, Malaysia Pihak Berkuasa Perancangan Tempatan Daerah Johor Bahru.

[15] Ponce V M and Hawkins R H 1996 *J. Hydrol.Eng.* 1 11-19

[16] Rochoxicz John A Jr 2011 *eJSiE* 4

[17] USDA-SCS 1964 *National engineering handbook Sect. 4 Hydrology* Washington DC

[18] USDA-SCS 1972 *National engineering handbook Sect. 4 Hydrology* Washington DC

[19] Soulis K X and Valiantzas J D 2013 *Water Resour. Manage.* 27 1737-49

[20] Van Mullem J A, Woodward D E, Hawkins R H and Hjelmfelt A T 2012 *Runoff Curve Number Method: Beyond the Handbook* U.S. Geological Survey Advisory Committee on Water Information - Second Federal Interagency Hydrologic Modeling Conference Las Vegas, Nevada

[21] Young G A and Smith R L 2005 *Essentials of Statistical Inference* ed Gill R, Ripley B D, Ross S, Silverman B W and Stein M Cambridge University Press

[22] Yuan Y, Nie J, McCutcheon S C, Taguas E V 2014 *Hydrol. Process.* 28 774-783

[23] Zhou S M, Lei T W 2011 *Calibration of SCS-CN Initial Abstraction Ratio of a Typical Small Watershed in the Loess Hilly-Gully Region* China Agric. Sci. 44 pp 4240-47