water

MDPI

Article

A CN-Based Ensembled Hydrological Model for Enhanced Watershed Runoff Prediction

Muhammad Ajmal [1,2], Taj Ali Khan [1] and Tae-Woong Kim [3,*]

Received: 2 October 2015; Accepted: 11 January 2016; Published: 15 January 2016
Academic Editor: Kwok-wing Chau

1 Department of Agricultural Engineering, University of Engineering and Technology, Peshawar 25120, Pakistan; engr_ajmal@uetpeshawar.edu.pk or ajmal@hanyang.ac.kr (M.A.); taj_marwat@yahoo.com (T.A.K.)
2 Department of Civil and Environmental Engineering, Hanyang University, Seoul 04763, Korea
3 Department of Civil and Environmental Engineering, Hanyang University, Ansan 15588, Korea
* Correspondence: twkim72@hanyang.ac.kr; Tel.: +82-31-400-5184

Abstract: A major structural inconsistency of the traditional curve number (CN) model is its dependence on an unstable fixed initial abstraction, which normally results in sudden jumps in runoff estimation. Likewise, the lack of pre-storm soil moisture accounting (PSMA) procedure is another inherent limitation of the model. To circumvent those problems, we used a variable initial abstraction after ensembling the traditional CN model and a French four-parameter (GR4J) model to better quantify direct runoff from ungauged watersheds. To mimic the natural rainfall-runoff transformation at the watershed scale, our new parameterization designates intrinsic parameters and uses a simple structure. It exhibited more accurate and consistent results than earlier methods in evaluating data from 39 forest-dominated watersheds, both for small and large watersheds. In addition, based on different performance evaluation indicators, the runoff reproduction results show that the proposed model produced more consistent results for dry, normal, and wet watershed conditions than the other models used in this study.

Keywords: hydrological model; pre-storm soil moisture; runoff prediction; variable initial abstraction

---

1. Introduction

The one-parameter traditional curve number (CN) model (CN model) developed by the U.S. Soil Conservation Service (SCS) now known as Natural Resources Conservation Service (NRCS), has enjoyed a long history of application as a lumped hydrological model. Its simplicity, versatility, and the availability of the necessary data have made it popular worldwide, as reported in [1], and an essential component in various hydrologic models, including water balance and storm routing models. Obviously, the CN model is reputable in the realm of applied hydrology, and CN is a mature concept that will remain in the forefront of engineering design. Its different versions, despite the complicated forms that result from introducing new parameters, cannot apply directly in approximating real situations because of the problem of model closure. Therefore, the CN model has been one of the most appealing and popular models for watershed runoff estimation for more than six decades [2]. The unavailability of any other simple contender has allowed this model to enjoy a long application history [3]. The widespread application of the CN model has led to its inclusion in hydrological software for surface runoff computations, such as CREAMS, CELTHYM, EPIC, HELP, L-THIA, PRZM, SWAT, SWIM [4], AGNPS, EPA-SWMM, GLEAMS, HEC-HMS, NLEAP, WinTR20, and WinTR55 [5]. Furthermore, the absence of hydrologic gauging stations and the high cost of gauging station installations lead hydrologists to estimate the surface direct runoff for ungauged watersheds using various techniques, and the CN model plays the leading role in such a situation [6].

Nonetheless, the CN model is not without its shortcomings. Despite its wide applications, the CN model has some limitations and misapprehensions caused mainly by its basic empirical assumptions [5]. For example, it considers a three-step CN variation based on three antecedent moisture conditions (AMCs), dry, normal, and wet, which permit unreasonable sudden jumps in runoff estimation [7]. Using tabulated CNs also underestimates most of the storm surface direct runoff in steep slope watersheds [8]. Furthermore, application of the CN model for forested watersheds can result in an inaccurate estimate of runoff from a given storm rainfall [9–11]. Thus, despite the long history, the CN model requires further study to develop and introduce a more robust model to better estimate runoff. Among other researchers, [1] and [11] modified the CN model to improve its runoff prediction capability. The modification proposed by [1] introduced storm-to-storm variation by incorporating a new parameter (M_c) to account for soil moisture before rainfall occurrence. However, this parameter sometimes results in negative values, which limits its applications [1]. The simplified one-parameter modified model suggested by [11] claimed improved runoff prediction. Nevertheless, that model has been criticized for not preventing the sudden jumps in runoff estimation [4,5,12]. To avoid sudden jumps in runoff estimation, simulated soil moisture (SM) has been employed using the soil water balance equation [13]. Further the parameter *S* (maximum potential retention) of the CN model has been determined by means of an experimentally derived relationship between *S* and SM as $S = a(1–\theta_e)$, where θ_e is the simulated relative SM at the beginning of the rainfall storm event and *a* is a parameter to be optimized. However, due to limited data in ungauged watersheds, neither θ_e nor *a* can be estimated and hence this reduces the application of the CN model only for gauged watersheds. Recognizing the limitations of the traditional CN model and its modified versions, we set out to develop a new ensembled lumped model that accounts for a continuous initial abstraction and prevents sudden jumps in runoff estimation. In addition, by using a small number of required parameters, we intended for our new conceptualization to be less sensitive to CN variation and structurally more consistent than the traditional CN model and its modified versions.

2. Materials and Methods

2.1. Study Area and Data

To analyze the runoff reproduction, we selected 39 South Korean forest-dominated mountainous watersheds. The selected watersheds represent the overall hydro-meteorological setting of South Korea and their corresponding characteristics are described in [14]. The selected watersheds vary in size from 42.32 to 888.01 km^2 and are characterized by low to high elevations (26 to 911 m above mean sea level) with average slopes from 7.50% to 53.53%. In Table 1, we differentiated the watersheds as small (Area $\leqslant$ 250 km^2) and large (Area > 250 km^2) following [9]. The land cover is mainly forests, followed by agricultural and urbanized land as illustrated in Table 1. Loam and sandy loam are the two major soil textures with some fractions of silt loam.

We collected 30 min time-step rainfall data from the Korea Meteorological Administration (KMA) whereas observed discharge data at the same time step came from the Hydrological Survey Center (HSC) of Korea. We used land cover information from the Ministry of Land, Infrastructure, and Transport (MOLIT). The collected measurements constitute 1804 rainfall-runoff events from 39 watersheds between 2005 and 2012. Figure 1 gives the locations of the large watersheds in the study area; the small watersheds can be seen in [14].

To separate baseflow and direct runoff from discharge, we used the straight-line hydrograph method [15]. We first screened the measured data to exclude small storm rainfall events ($P < 25.4$ mm) to prevent bias in estimating runoff using the traditional CN model [16]. We applied the P_5 (the accumulated prior five days rainfall) criterion to determine the watershed antecedent condition to adjust the CN and its corresponding *S* values from normal to dry and wet conditions [12]. We took the composite CN by applying the procedure documented in the NRCS [17] based on the watersheds' land cover characteristics, which we considered to be the same for all models. According to [18], hydrological models are more reliable when they reproduce satisfactory measurements, especially from watersheds with limited data. To make the models applicability more realistic in ungauged watersheds, we did not calibrate any of the parameters used in this study. Some statistics regarding the measured and predicted data are shown in Table 2.

Table 1. Watersheds and data description.

WS ID	Watershed Name	Major Land Cover Distribution (km^2)				Area	NOE	ME	α	CN
		Forests	Agriculture	Urbanized	Grass	(km^2)		(m)	(%)	
Small watersheds, Area $\leqslant$ 250 km^2										
WS01	Cheonwang	97.05	57.51	30.82	3.86	42.32	29	26	13.40	66
WS02	Daeri	47.31	1.67	0.39	0.25	60.45	39	424	48.13	75
WS03	Janggi	36.76	23.80	2.44	1.11	62.80	42	146	21.50	70
WS04	Dopyeong	106.03	27.75	15.42	3.70	138.36	34	173	28.71	64
WS05	Chunyang	105.06	23.66	2.69	5.20	143.10	40	197	34.30	60
WS06	Cheongju	14.95	14.48	17.23	0.50	161.44	70	202	20.10	69
WS07	Boksu	119.09	30.83	5.02	1.30	161.90	26	343	35.50	60
WS08	Donghyang	111.34	44.76	4.51	2.67	164.66	68	911	35.09	64
WS09	Maeil	152.06	20.79	0.48	0.13	174.86	60	517	39.65	53
WS10	Yulgeuk	42.41	110.95	15.47	3.72	179.95	38	113	7.50	71
WS11	Toigyewon	137.55	37.86	14.06	6.05	200.45	44	285	26.70	64
WS12	Jungrang	94.17	15.54	83.26	2.01	208.41	42	219	17.30	67
WS13	Soochon	79.46	108.11	18.42	2.09	223.19	31	76	15.40	73
WS14	Guryong	163.71	65.18	7.08	3.95	245.50	33	244	26.60	65
WS15	Yoosung	167.64	48.15	17.64	7.29	249.63	65	349	27.30	71
Large watersheds, Area > 250 km^2										
WS01	Kyeongan	153.70	44.73	37.15	11.50	256.91	89	165	22.86	63
WS02	Jeonju	169.78	59.36	39.36	5.03	278.00	72	168	28.33	70
WS03	Cheoncheon	183.91	72.31	8.18	16.85	284.03	89	554	32.23	58
WS04	Gwanchon	217.56	62.11	7.90	7.65	301.26	49	420	33.70	70
WS05	Gapyeong	274.05	18.52	5.28	2.21	305.12	39	490	45.40	69
WS06	Heukcheon	232.82	57.19	13.04	3.75	307.82	19	253	32.70	59
WS07	Heungcheon	124.45	135.52	31.07	7.32	309.08	28	112	13.80	67
WS08	Bookcheon	377.04	123.35	36.04	15.97	330.20	38	733	53.53	52
WS09	Changchon	280.95	35.50	5.64	2.45	335.07	83	523	41.82	68
WS10	Ohsoo	210.60	118.95	11.00	3.92	350.09	61	243	24.60	62
WS11	Wangsungdong	374.02	20.97	1.33	2.12	378.67	24	866	47.80	61
WS12	Sanganmi	350.57	40.55	4.31	2.62	402.45	26	778	39.40	61
WS13	Shinan	301.16	87.52	9.36	3.32	411.96	56	244	31.25	71
WS14	Janghowon	181.36	182.02	23.52	10.92	431.23	26	678	16.80	65
WS15	Youngjung	288.93	101.36	30.87	12.28	445.36	29	268	27.00	61
WS16	Sangyegyo	331.21	134.62	13.42	3.29	496.30	35	268	29.40	71
WS17	Cheongmi	215.69	219.85	29.82	14.12	514.66	30	147	16.70	78
WS18	Hwachon	443.38	50.85	6.45	0.91	523.20	59	499	41.40	57
WS19	Banglim	448.67	56.84	5.81	3.30	527.12	30	763	40.20	63
WS20	Joocheon	449.10	68.49	5.74	1.90	533.23	65	608	38.12	58
WS21	Hoideok	362.89	105.74	94.04	23.68	609.15	41	170	25.70	70
WS22	Songcheon	455.78	131.15	11.86	4.54	612.17	54	386	33.25	64
WS23	Pyeongchang	609.60	79.87	8.92	4.47	697.67	64	734	40.30	64
WS24	Panwoon	757.06	99.67	11.18	6.29	888.01	37	678	40.88	60

WS ID, NOE, ME, and α are the watershed identification, number of events, mean elevation, and mean slope, respectively.

Table 2. Statistics of model outputs from the combined data set from 39 watersheds.

Data Type	Parameter/ Model	Statistics							
		Min	Mean	Median	Max	SD	Skewness	25th Percentile	75th Percentile
Observed data	P (mm)	25.12	78.94	58.32	519.68	60.89	2.32	40.03	94.02
	P_5 (mm)	0.00	58.18	34.80	629.80	75.12	2.63	6.95	79.00
	T (h)	1.50	19.62	16.00	154.00	13.96	2.71	11.00	24.00
	Q_o (mm)	0.17	37.26	19.36	364.38	46.81	2.46	8.10	46.96
Modeled data (Q_c(mm))	CNM	0.00	24.36	6.22	415.63	44.50	3.50	1.49	24.48
	MCM	0.56	44.56	26.08	487.31	55.02	2.63	7.71	56.38
	MRM	0.00	30.84	11.42	432.04	48.60	3.01	1.17	37.64
	AJM	1.62	34.69	20.05	375.63	41.52	2.97	9.95	40.47

Other parameters are defined in the text, T is the storm duration.

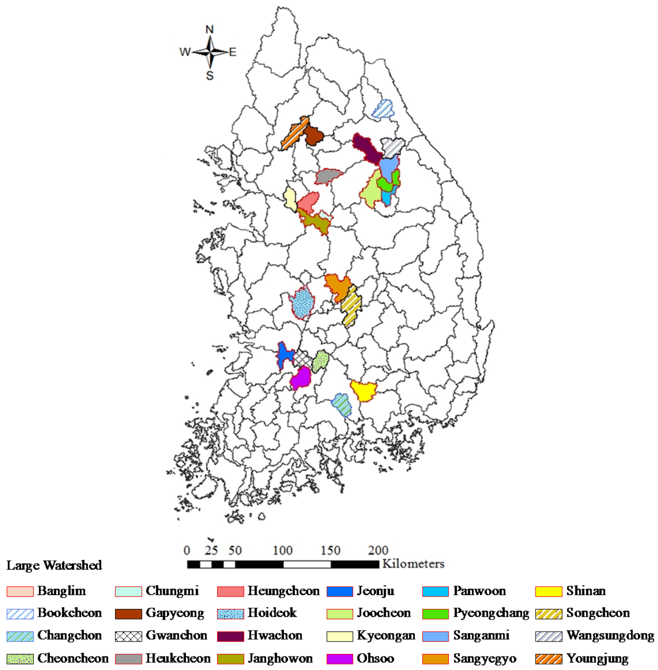

Figure 1. Location of large watersheds (Area > 250 km^2) in the study area. Small watersheds (Area $\leqslant$ 250 km^2) can be seen in [14].

2.2. Development of a New Hydrological Model

2.2.1. The Traditional CN Model (CNM)

For a given P and its corresponding initial abstraction, I_a, the traditional CN model in general is expressed as:

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

where Q, P, I_a, λ and S are surface direct runoff (mm), total rainfall (mm), initial abstraction (mm), initial abstraction coefficient (dimensionless), and potential maximum retention (mm), respectively. The standard $I_a = 0.20S$ [19] and its one modified version for comparable improved runoff prediction is $I_a = 0.05S$ [20]. Different researchers have found structural inconsistencies in the traditional CN model, such as sudden jumps in S with corresponding runoff values for dry, normal, and wet conditions [11,12]. Similarly the standard $I_a = 0.20S$ has been found to be unreasonably high; and researcher across the globe suggested $I_a < 0.2S$ (e.g., [1,14,20–23]). For the CNM, we estimated the weighted composite CN for watershed characteristics and calculated the transformed S using the following Equation:

$$S = 25.4\left(\frac{1000}{\mathrm{CN}} - 10\right) \quad (2)$$

2.2.2. Mishra *et al*. Model (MRM)

To prevent unreasonable sudden jumps in runoff estimation, the relationship for I_a was modified by incorporating a new parameter (M_c) to account for soil moisture content before the rainfall occurrence [1]. The modified I_a is expressed as:

$$I_a = \frac{\lambda S^2}{S + M_c} \quad (3)$$

The new parameter (M_c) can be found from the prior five days accumulated rainfall (P_5), λ and S as $M_C = -0.5\left[(1+\lambda)\,S - \sqrt{(1-\lambda)^2\,S^2 + 4P_5S}\right]$ or with a more simplified version $M_C = \beta\sqrt{P_5S}$ where β = 0.72 is the optimized coefficient. Despite some improvements in runoff prediction, the application of M_c in [1] sometimes causes negative Q values, which is the major drawback of that formulation. Therefore, the model should be revisited to eliminate this complication when estimating runoff. For runoff estimation, the MRM can be expressed as:

$$Q = \frac{(P - I_a)(P - I_a + M_c)}{P - I_a + M_c + S} \tag{4}$$

In Equation (4), $M_C = 0.72\sqrt{P_5S}$ which is one of the best versions presented in [1] and I_a is calculated using Equation (3) (λ = 0.2).

2.2.3. Michel *et al*. Model (MCM)

For λ = 0.2, the simplified forms of the MCM for dry (AMC-I), normal (AMC-II), and wet conditions (AMC-III), respectively, are as follows:

$$Q = \frac{P^2}{P + S} \tag{5}$$

$$Q = P\left(\frac{0.48S + 0.72P}{S + 0.72P}\right) \tag{6}$$

$$Q = P\left(\frac{0.79S + 0.46P}{S + 0.46P}\right) \tag{7}$$

Similarly, a simplification of the model parameters reported in [11] also considers a three-step S parameter and cannot eradicate the problem of sudden jumps in runoff prediction.

2.2.4. The Proposed Model (AJM)

To circumvent the sudden jumps and other inconsistencies, we here derived a new model ensembling the traditional CN and GR4J models [23] with the pre-storm moisture content concept [11] and further introduced a continuous S formulation. The GR4J (which stands for modèle du génie rural à 4 paramé tres journalier) is a French daily lumped four-parameter rainfall-runoff model that belongs to the family of soil moisture accounting models. It determines four parameters for daily runoff prediction: maximum storage capacity, groundwater exchange coefficient, one day prior maximum routing storage capacity, and time base of unit hydrograph. More details on the GR4J model are available in [23].

We replaced the M_c in Equation (3) with V_o adopted from [11], to make the following relationship:

$$I_a = \frac{\lambda S^2}{S + V_o} \tag{8}$$

where V_o is the soil moisture content before a storm event. Different values of V_o are found for different AMCs [11] which are differentiated based on P_5 [14]. Using the available CN adjustment formulae from normal (AMC-II) to dry (AMC-I) and wet (AMC-III) conditions, the expression for V_o becomes [4]:

$$V_o = \gamma S \quad \text{for AMC-II} \tag{9}$$

$$V_o = \gamma\left[\frac{S}{2.3}\right] = 0.435\gamma S \quad \text{for AMC-I} \tag{10}$$

$$V_o = \gamma \left[\frac{S + 289.56}{0.43} \right] \quad \text{for AMC-III} \tag{11}$$

where γ is the fraction to link V_o and S that can be optimized from measured data. After testing the measured rainfall-runoff data from 140 French watersheds with different surface and variable moisture conditions, V_o as some fraction of S was simplified for dry, normal and wet conditions as; $V_o = 0.33S$, $V_o = 0.61S$, $V_o = 0.87S$, respectively. Taking $\lambda = 0.20$, as generally assumed for practical applications in the traditional CN model, and substituting V_o from Equations (12)–(14), respectively into Equation (8) gives:

$$I_a = \frac{0.2S^2}{S + 0.33S} = 0.150S \quad \text{for AMC-I} \tag{12}$$

$$I_a = \frac{0.2S^2}{S + 0.61S} = 0.124S \quad \text{for AMC-II} \tag{13}$$

$$I_a = \frac{0.2S^2}{S + 0.87S} = 0.107S \quad \text{for AMC-III} \tag{14}$$

To incorporate the new initial abstraction formulation, we conceptualized a new model by combining the CN model, the PSMA procedure [11], and the event-based empirical GR4J model. To do this, we initialized V_o as the soil moisture store level at the onset of an event and V as the moisture level corresponding to the total rainfall P [11].

$$V = V_o + P - Q \tag{15}$$

Substituting Q from Equation (1) into Equation (15) yields

$$V = V_o + P - \frac{(P - I_a)^2}{P - I_a + S} = V_o + \frac{(S + I_a)\, P - {I_a}^2}{(P - I_a + S)} \tag{16}$$

Now the simplified form of the GR4J runoff model described in [11] can be expressed in its cumulative form as:

$$Q = (P - PE) \times \left(\frac{V}{S + S_a} \right)^2 \quad P > PE \tag{17}$$

Here PE is the daily potential evapotranspiration and is assumed negligible because the runoff from rainfall usually lasts for an event of sufficiently limited duration [24]. Hence, the above equation becomes:

$$Q = P \times \left(\frac{V}{S + S_a} \right)^2 \tag{18}$$

Equation (18) yields $Q = P$ for $V = S + S_a$ as a maximum capacity of V, where S_a is an intrinsic parameter equal to; $S_a = V_o + I_a$ [11]. Substituting the expression for V from Equation (16) into Equation (18) and simplifying yields

$$Q = P \times \left[\frac{V_o\,(P - I_a + S) + (S + I_a)\, P - {I_a}^2}{(P - I_a + S)\,(S + V_o + I_a)} \right]^2 \tag{19}$$

Now, substituting $I_a = 0.15S$ from Equation (12) and $V_o = 0.33S$ (assuming the watershed dry condition), after simplification Equation (19) becomes:

$$Q = P \times \left[\frac{1.48PS + 0.258S^2}{1.48PS + 1.258S^2} \right]^2 \quad \text{or} \quad Q = P \times \left[\frac{P + 0.174324S}{P + 0.85S} \right]^2 \tag{20}$$

It is interesting that Equation (20) is a simple, one-parameter model like the traditional CN model. Using the traditional CN model, as in Equation (1), after substituting I_a from Equation (13) could improve runoff prediction if S is taken as the transformed value from the dimensionless CN assessed as a function of land use/cover, soil types, and AMCs (differentiated from P_5). However, many researchers of various Mediterranean watersheds reported limitations and drawbacks to that approach, and it inconsistently displays pre-storm soil moisture temporal variations (e.g., [19,25–27]). Some researchers (e.g., [26,28]) have specifically mentioned that the cumulative rainfall in the previous five days is not a good proxy of the wetness conditions of the catchment before a flood event. Similarly, the use of a fixed I_a in the traditional CN model causes unusual sudden jumps in runoff estimation [1,4,11,12]. To circumvent this issue, we present a continuous spatio-temporal expression for the initial abstraction by replacing the S value in the proposed model with a new expression; $S \times \left(\frac{P}{P+P_5}\right)$. Assuming a watershed in dry condition, Equation (12) for the new I_a becomes

$$I_a = 0.15S\left(\frac{P}{P+P_5}\right) \tag{21}$$

Substituting Equation (21) into Equation (19), and retaining $V_o = 0.33S$ for dry watershed condition and simplifying gives:

$$Q = P \times \left(\frac{P_5 + P + 0.174324S}{P_5 + P + 0.85S}\right)^2 \quad \text{for} \quad P > 0.15S\left(\frac{P}{P+P_5}\right) \tag{22}$$

For $P_5 \to \infty$ then $Q \to P$ which means the watershed is fully saturated by pre-storm soil moisture content and it has no more capacity to retain water. Hence all rainfall will be converted to runoff. In the case of $S \to 0$, then $Q = P$, whereas, if $S \to P$ and $P_5 = 0$ (dry condition), then $Q = 0.40P$. Using this new model (Equation (22)) does not require any adjustment for S values from normal to dry or wet conditions, because the P_5 variation accounts for the I_a adjustment associated to different pre-storm soil moisture responses in a watershed.

From this point, we investigated four models in this study: the traditional CN model (CNM), the model proposed in [11] as Michel model (MCM) and in [1] Mishra model (MRM), and our proposed model (AJM). Details of the different models are given in Table 3.

Table 3. Models and their corresponding parameters identification.

Model ID	λ	CN	Model Expression	Remarks
CNM	0.20	NEH-4 Tables	Equations (1) and (2)	Original CN model in [17]
MCM	-do-	-do-	Equations (2), (5), (6) and (7)	Modification in [11]
MRM	-do-	-do-	Equations (2)–(4)	Modification in [1]
AJM	-do-	-do-	Equations (2) and (22)	Proposed Model

-do- means as above.

3. Models' Goodness-of-Fit Evaluation

A vital step in evaluating and inferring model results is creating graphs to display them and determining goodness-of-fit indicator values to quantify model performance in terms of prediction accuracy [29]. No formal standards evaluate goodness-of-fit to data visually or numerically through quantitative assessment. This leads to considerable variability in the techniques used, which are frequently decided based on applications in previous similar studies. According to [30], model evaluation (validation) and sometimes performance are judged by comparing calculated values with measured benchmark data. Commonly, evaluations of model performance are based on several statistics and techniques. To assess models' prediction accuracy quantitatively, the most commonly used methods are the root mean squares error (RMSE) [15,31–33], Nash-Sutcliffe efficiency (NSE) [32–34], and Percent Relative Error (PRE) [35] expressed as:

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(Q_o - Q_c)_i^2\right]^{0.5} \tag{23}$$

$$NSE = \left[1 - \left(\frac{RMSE}{SD}\right)^2\right] \tag{24}$$

$$PRE = \left[\frac{\sum_{i=1}^{N}(Q_o - Q_c)_i}{\sum_{i=1}^{N}(Q_o)_i}\right] \times 100 \tag{25}$$

where Q_{oi} and Q_{ci} are the observed and estimated runoff values for storm events i to N and SD is the standard deviation of observed data. Smaller RMSE values (optimum = 0.0) depict better model runoff estimation. Here, $\overline{Q}_o$ represents the mean of observed runoff values for storm events i to N. A model is deemed satisfactory if its NSE > 0.50 [35]. Another recent study considered a hydrologic model satisfactory if NSE > 0.65 [36].

We further tested the appropriateness of each model based on the best-fit distribution and its corresponding statistic for the cumulative observed and modeled runoff from the 39 watersheds. To do so, we analyzed the observed and modeled data using three different tests: Kolmogorov-Smirnov test, Anderson Darling test, and chi-squared test [37]. We used those tests to identify that the data belong to a specific distribution. We identified the best-to-worst fit distribution for those three tests from rank 1 to rank 18 after analyzing for 18 different distribution fittings. The model depicting the same distribution fit as exhibited by the measured runoff would be the best alternative for runoff prediction in the study area. Significant agreement between the measured and modeled runoff can be further validated in comparing models based on closeness of statistical values from any of the aforementioned three tests.

Beside the above evaluation indicators, to examine the model performance, the time series data for measured and modeled runoff were compared for two example watersheds selected from the study area. In addition, percent errors on peak discharge (E_{Q_P}) [13] were also calculated to identify the physical behaviors of the models in estimating runoff.

$$E_{Q_P} = \left[\frac{\max(Q_o) - \max(Q_c)}{\max(Q_o)}\right] \times 100 \tag{26}$$

where max (Q_o) and max (Q_c) are the maximum peak discharge values for the observed and modeled runoff.

4. Results and Discussion

To select an appropriate physical relationship between P_5 and I_a we assessed the watersheds' runoff-producing response for the common discrete (fixed) I_a = 0.2S and the proposed variable (continuous) formulation. This assessment was accomplished once the pre-storm soil moisture was adjusted in I_a based on the recommendation from [11] and was subsequently adopted by other researchers in their studies (e.g., [4,5,12]). In our analysis, we kept the CN as a major parameter obtained from watershed characteristics under normal conditions fixed for all models. The measured rainfall-runoff events we used in the model assessments were characterized by a complex pattern to represent a diverse response. Considering P and CN as the fundamental and common parameters and other physiographic characteristics as shown in Table 1, all the models compared in this study both structurally differ from each other and lie within the parameter space of the traditional CN model.

Figure 2 illustrates the variation of fixed and variable I_a used in the CNM and AJM respectively. For the CNM, the evident high I_a and sudden jumps in S were the two major reasons for underestimating runoff from steep slope watersheds like in South Korea. As illustrated [9], the traditional CN model should be employed cautiously in watersheds larger than 250 km^2. Keeping in mind this criteria, the model's analysis was carried out for small (Area $\leqslant$ 250 km^2) and large (Area > 250 km^2) watersheds as well as based on pre-storm watershed hydrological conditions (dry, normal and wet). As shown in Figure 3a, for small watersheds, the MCM showed no evident improvement in terms of mean (median) RMSE (mm) values of 22.66 (23.25) compared to the CNM values of 22.07 (20.15). However, the spread of RMSE values in those watersheds between the 25th and 75th percentiles depicted improvement followed by the MRM with mean (median) RMSE of 20.79 (18.78). Contrarily, the AJM indicated more accurate mean (median) RMSE of 13.13 (12.34). The improved results by the AJM indicate the reliability of the proposed variable I_a in estimating runoff from mountainous watersheds compared to the CNM and the models presented in [1] and [11].

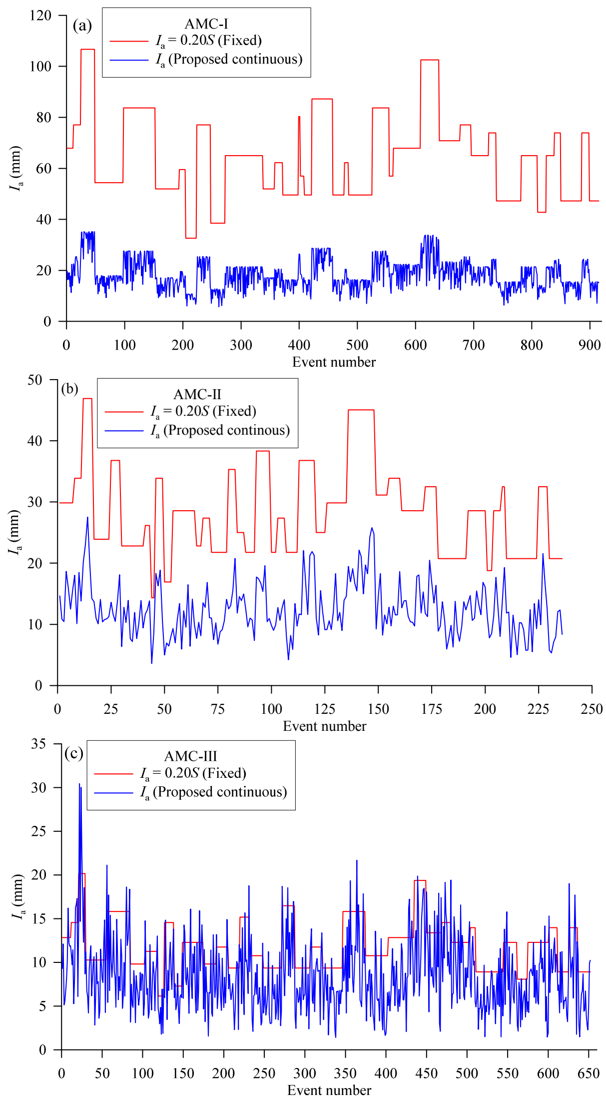

Figure 2. Variation of fixed and variable initial abstraction (I_a) for rainfall events in (**a**) dry condition; (**b**) normal condition; and (**c**) wet condition.

To assess the model's efficacy in large watersheds, as shown in Figure 3b, it is similar to that for small watersheds, as depicted in Figure 3a. The mean (median) RMSE for the CNM, MCM, MRM, and AJM were 27.94 (24.85), 27.00 (25.86), 25.50 (23.48) and 17.40 (16.16), respectively. The MCM showed almost no improvement over the CNM and the MRM exhibited modest improvement, whereas the

AJM indicated significantly lower RMSE values in all watersheds. The cumulative RMSE values for all types of events, shown in Figure 3c, support that the AJM has more reliable results than the other contenders. The consistently lower RMSE values evidenced by the AJM for all kinds of events (combined dry, normal, and wet conditions from 39 watersheds) indicate its superiority.

Analyzing the runoff prediction efficacy illustrated by different models using NSE as the performance indicator demonstrated modest improvement by the MCM followed by the MRM, as shown in Figure 4a. The mean NSE (0.57) for the MCM showed almost no improvement compared to the CNM mean NSE (0.57), but the spread of the NSE from this figure indicates good model performance by the MCM. Moreover, the MRM mean NSE (0.63) indicated some increases in the model's overall efficacy compared to the CNM and MCM. Nevertheless, the AJM had the highest NSE values (mean 0.85) indicating its reliability for accurate runoff estimation in small watersheds. Figure 4b evidences a similar pattern of performance for large watersheds. Here, the mean NSE for the CNM, MCM, MRM, and AJM were 0.64, 0.65, 0.71, and 0.86, respectively. The higher NSE values in both small and large watersheds confirmed the AJM enhanced runoff prediction. According to the NSE criteria in [35], the CNM, MCM, MRM, and AJM had respectively, good to very good results 46.67%, 33.33%, 60.00%, and 100% of small watersheds. Similarly, the models respectively fulfilled the same performance rating in 41.67%, 54.17%, 58.33%, and 100% of large watersheds. Considering the performance rating criteria in [36], the good to very good performance by the CNM, MCM, MRM, and AJM respectively, was in 26.67%, 6.67%, 26.67%, and 66.67% of small watersheds. Likewise, those ratings respectively were met in 25.00%, 16.67%, 37.50%, and 87.50% of large watersheds.

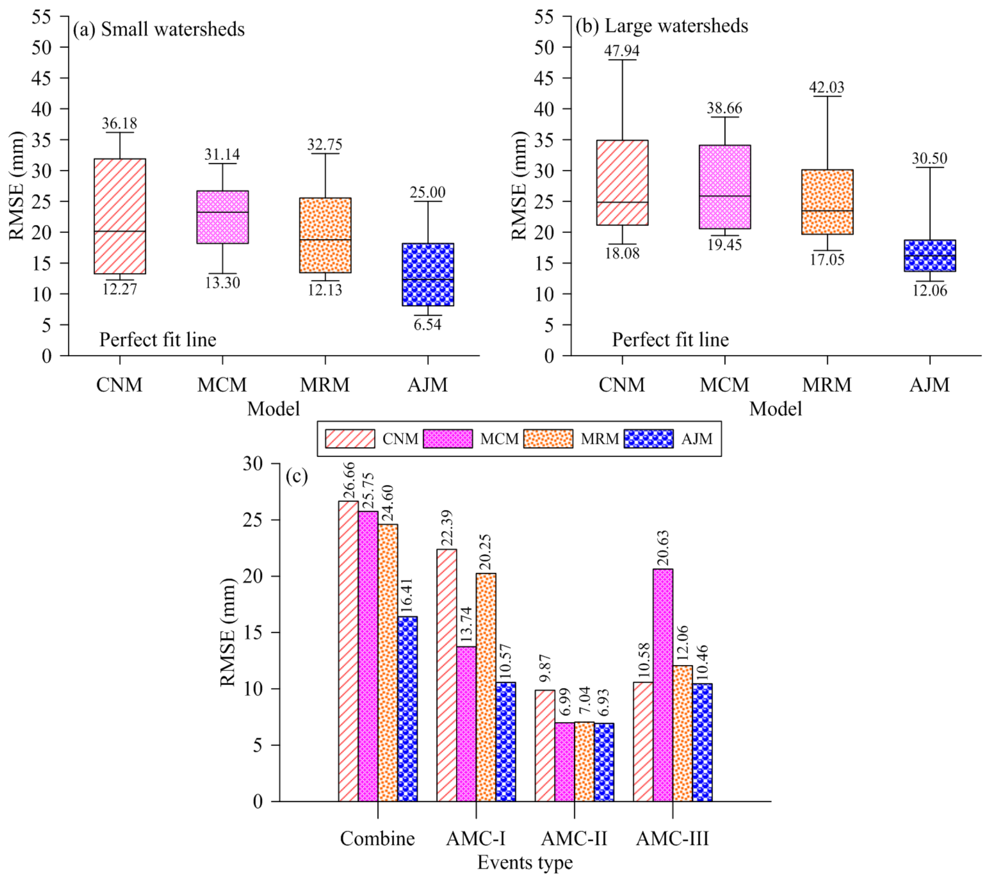

Figure 3. Model performance based on root mean squares error (RMSE) (**a**) in small watersheds (Area < 250 km^2); (**b**) in larger watersheds (Area > 250 km^2) and (**c**) for different types of events (dry, normal, and wet).

For cumulative events, all AMC-I, AMC-II and AMC-III events from 39 watersheds, the CNM showed improvement from AMC-I to AMC-II and AMC-III. However, the MCM exhibited improvement in AMC-II events compared to AMC-I but worse results in AMC-III. The MRM increased in efficiency from AMC-I to AMC-II and AMC-III events. Nonetheless, the higher and consistent NSE values exhibited by the AJM indicated superior performance in all watersheds for all types of events. This performance can be attributed to the incorporation of the variable I_a in the CN model in the AJM.

The RMSE and NSE cannot indicate over- or underestimation of the model runoff. Therefore, we assessed the models' performances using the PRE statistic. Figure 5a,b show that the CNM most underestimated all events, followed by the MRM. Contrarily, the MCM overestimated runoff for the majority of the events. Both the CNM and MCM were inferior because their PRE values are in the unsatisfactory range (PRE $\geqslant$ +25 or PRE $\leqslant$ −25) [14,35]. The mean PRE (%) in small (large) watersheds for CNM, MCM, MRM and AJM respectively were 36.55 (34.64), −23.01 (−20.83), 13.04 (16.59), and 5.66 (8.26). Evidently, both in small and large watersheds, the mean PRE values for the AJM were in the very good performance range [14,35]. Using the PRE statistics, all models show better performance in large watersheds than in small watersheds in the mountainous parts of the study area.

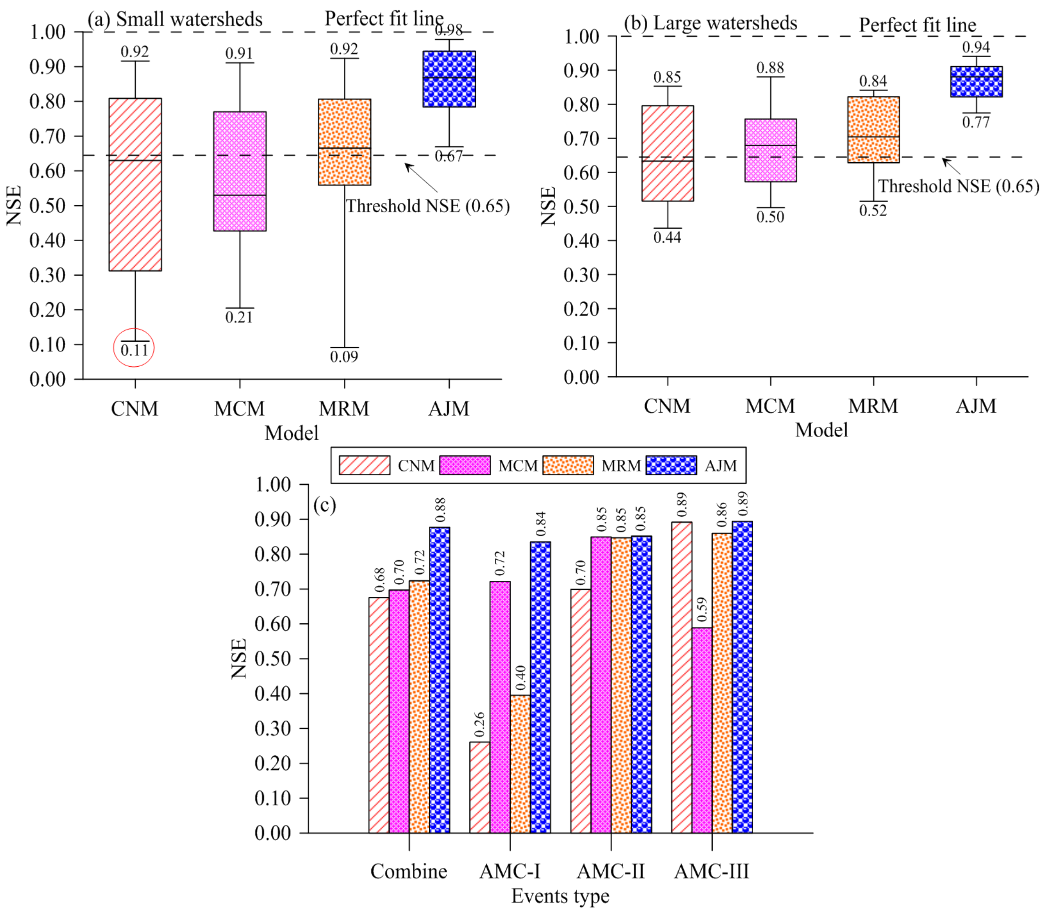

Figure 4. Model performance based on Nash-Sutcliffe efficiency (NSE) (**a**) in small watersheds (Area $\leqslant$ 250 km^2); (**b**) in larger watersheds (Area > 250 km^2) and (**c**) for different types of events (dry, normal and wet). The circled NSE value for the CNM shows the second overall value (the minimum was −0.05, which is not shown here).

For not only the cumulative events from all 39 watersheds, but also the cumulative AMC-I, II, and III events, the AJM more accurately and consistently predicted the runoff. On the contrary, the other models were inconsistent in their runoff predictions. The performance ratings of different models for runoff prediction are given in Table 4. Using different statistical indicators, the proposed model performed significantly better than the other models investigated. One reason for the CNM's inferior results might be the watersheds slope difference. The CNM was derived from rainfall-runoff data from watersheds with up to 5% slope [15], whereas the watersheds in our study area range between 7.50% and 53.53% slope (refer to Table 1). In addition, no statistical indicator showed any significant overall performance difference in predicting runoff from small or large watersheds, which indicates that the size of watershed is not a major concern for any models in this study.

Table 5 shows the best-fit distribution, ranking 1 out of 18 different continuous distributions associated with each test for observed runoff and that predicted by models. It is evident that the observed and AJM-based estimated runoff not only followed the Log-Pearson 3 (LP3) distribution with rank 1 from all three tests but also signified their association from very close test statistic values. Contrarily, the runoff modelled with the CNM and MRM were fitted to other types of distributions. The MCM output runoff values match only the LP3 test based on the Anderson Darling test with comparatively high test statistics. These distributions and their associated test statistics also showed that the proposed model (AJM) exhibited more reliable runoff prediction for steep slope watersheds in the study area than did the other models.

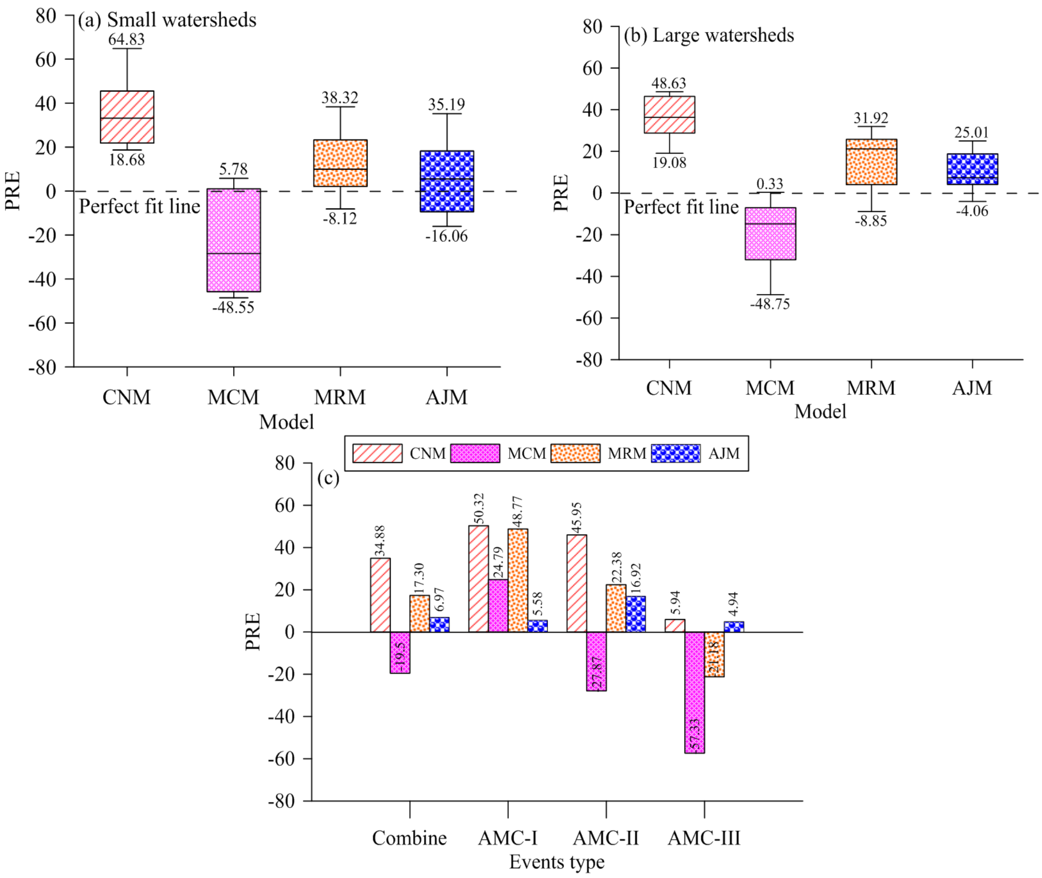

Figure 5. Model performance based on Percent Relative Error (PRE) (**a**) in small watersheds (Area $\leqslant$ 250 km^2); (**b**) in larger watersheds (Area > 250 km^2); and (**c**) for different types of events (dry, normal and wet).

We also evaluated the models' performances using scatter plots (Figure 6) between the observed and modeled runoff values for all events from 39 watersheds. It is evident that the CNM and MRM underestimated the runoff and made inferior predictions for most events. The runoff estimated with the MRM exhibited modest improvement compared to the CNM. The MCM depicted better results than the CNM and MRM. However, the proposed (AJM) model produced more consistent and statistically significant results for runoff estimation. The close agreement between the measured runoff values and those modeled by the AJM can be seen from the best-fit line approaching the 1:1 line with a high coefficient of determination (R^2 = 0.88) compared to that for the CNM (R^2 = 0.75), MRM (0.76), and MCM (R^2 = 0.82).

A time series plot (Figure 7a,b) compares the measured and modeled runoff values for the selected events from Cheoncheon and Donghyang watersheds highlighting the performance of different models used in this study. It was found that the AJM is comparatively more consistent than the other models.

Figure 8 shows the model performance in individual watersheds and represents their cumulative frequency distribution for the total 39 watersheds in this study. The maximum peak discharge values were calculated after estimating the peak discharge for the individual storm events in the respective watersheds. The evaluation based on the percentage error in peak discharge, E_{Q_P} (Figure 8a) as well as the percentage absolute error in peak discharge, $|E_{Q_P}|$ (Figure 8b), indicated that the runoff estimated using the proposed model is in better agreement with the measured data when compared to other models. Similarly, the CNM, MCM, MRM, and AJM depicted the overall mean (median) percentage absolute peak discharge $|E_{Q_P}|$ for the 39 watersheds as 16.82 (19.19), 26.15 (25.66), 18.30 (15.99), and 14.33 (12.81), respectively. From these statistical scores the proposed model can be encouraged for application in the study area as well as other hydrological similar areas.

Table 4. Model performance ratings in 39 watersheds based on different statistical indicators.

	According to [35]				
Model	Performance Index Range	0.75 < NSE ⩽ 1.00	0.65 < NSE ⩽ 0.75	0.50 < NSE ⩽ 0.65	NSE ⩽ 0.50
	Performance Rating	Very good	Good	Satisfactory	Unsatisfactory
	CNM	14	2	12	11
	MCM	10	8	13	8
	MRM	15	8	12	4
	AJM	36	3	0	0
	According to [36]				
Model	Performance index range	NSE ⩾ 0.90	0.80 ⩽ NSE < 0.90	0.65 ⩽ NSE < 0.80	NSE < 0.65
	Performance rating	Very good	Good	Satisfactory	Unsatisfactory
	CNM	2	6	8	23
	MCM	1	3	14	21
	MRM	1	10	12	16
	AJM	13	17	9	0
	According to [4,35]				
Model	Performance index range	PRE < 10	10 ⩽ PRE < 15	15 ⩽ PRE < 25	PRE ⩾ 25
	Performance rating	Very good	Good	Satisfactory	Unsatisfactory
	CNM	1	0	10	28
	MCM	14	25	21	17
	MRM	14	2	13	10
	AJM	19	8	11	5

Table 5. Comparison of observed and modeled runoff based on best-fit distributions (statistic).

		Test Type		
Data Type/Model		**Kolmogorov-Smirnov**	**Anderson Darling**	**Chi-Squared**
Runoff		Best-fit distribution (Statistic)		
Observed (Q_o)		LP3 (0.023)	LP3 (1.273)	LP3 (21.821)
Modeled (Q_c)	CNM	PBW (0.033)	W (39.865)	PBW (30.628)
	MCM	W (0.021)	LP3 (8.696)	GP (48.923)
	MRM	GEV (0.145)	W (37.654)	W (89.477)
	AJM	LP3 (0.024)	LP3 (1.523)	LP3 (15.949)

LP3 = Log Pearson 3; W = Wakeby; GP = General Pareto; PBW = Phased Bi Weibull; GEV = General Extreme Value.

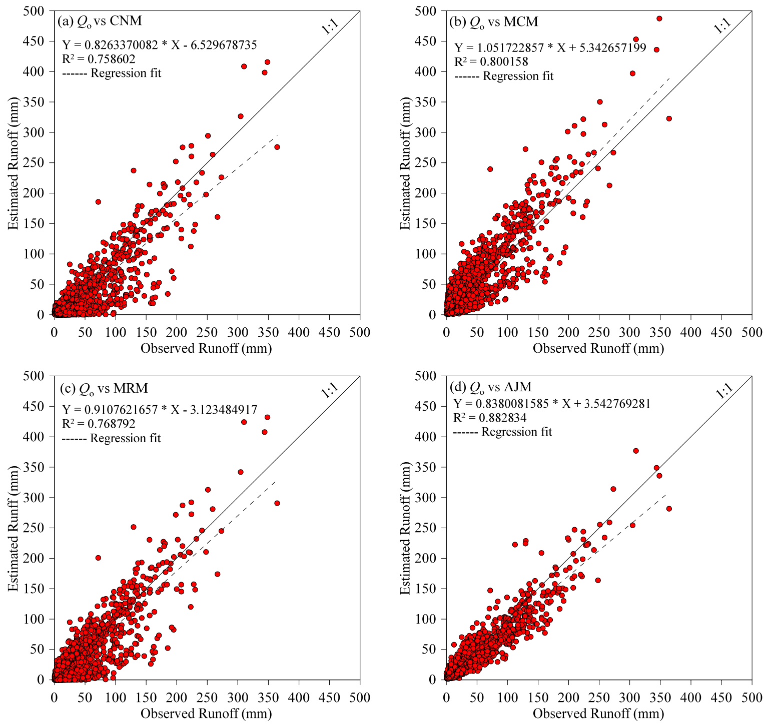

Figure 6. Comparison of events from all watersheds for (**a**) observed runoff (Q_o) *vs.* predicted with CNM; (**b**) Q_o *vs.* predicted with MCM; (**c**) Q_o *vs.* predicted with MRM; and (**d**) Q_o *vs.* predicted with AJM.

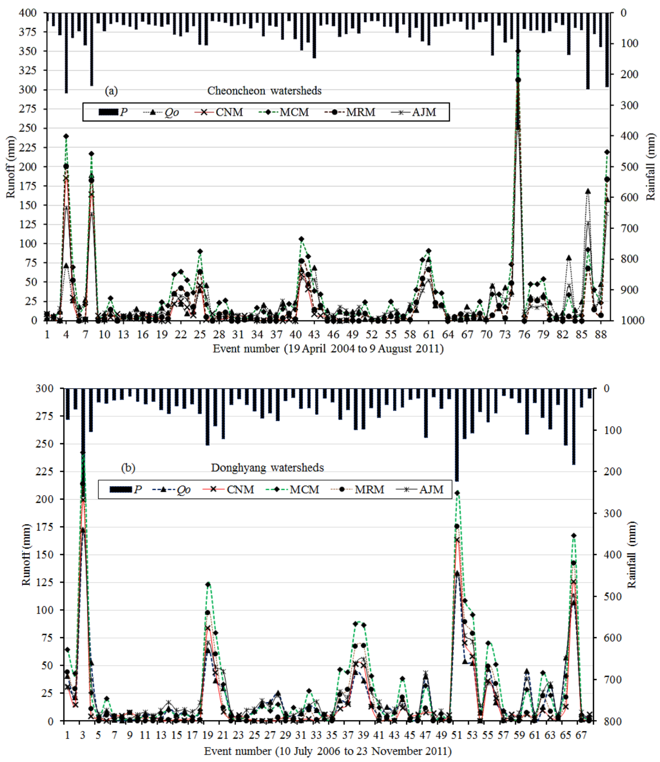

Figure 7. Comparison between the measured runoff time series values and those obtained from application of four models for (**a**) Cheoncheon watershed; and (**b**) Donghyang watershed.

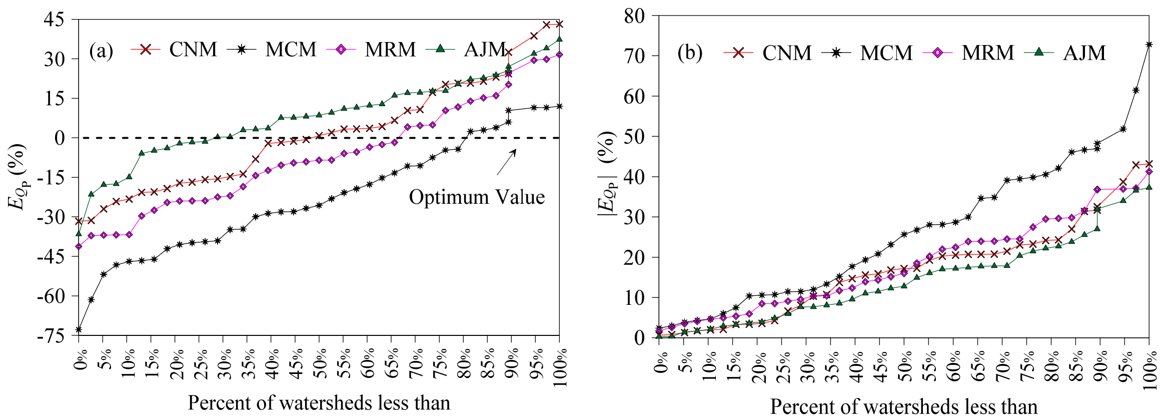

Figure 8. Model evaluation for 39 watersheds based on: (**a**) cumulative distribution of percentage error in peak discharge (E_{Q_P}), and (**b**) cumulative distribution of absolute percentage error in peak discharge $|E_{Q_P}|$.

5. Conclusions

We tested the rainfall and runoff data collected for 1804 selected events from 39 mountainous watersheds for runoff reproduction using our proposed model and three earlier contenders. Because it underestimated the runoff for the majority of events, the CNM was found inferior to the other models. This might be attributable to the fixed initial abstraction and the absence of a credible parameter for pre-storm soil moisture, which results in sudden jumps in runoff estimation. After testing data from 140 French watersheds [11], researchers conceptualized an intrinsic parameter and incorporated it into the CN model to enhance its runoff prediction capability. However, their simplified model was unable to circumvent sudden jumps in runoff prediction and hence overestimated the majority of events. The model's amendment presented in [1] was also unable to handle this inconsistency and provided only modest improvement on the CNM.

In the proposed model, we conceptualized a CN-based ensembled approach by amending the previously suggested formulation [1] for initial abstraction after incorporating the pre-storm soil moisture adjustment [11]. To circumvent sudden jumps in runoff estimation, we used a continuous and variable initial abstraction that depends not only on S, but also takes into account the storm magnitude and the prior rainfall effects irrespective of the three AMCs. Our modification resulted in a structurally more consistent model. Our proposed model has shown a high degree of reliability in predicting runoff from a majority of watersheds in the study area irrespective of the watershed size. Because of its simplicity and statistically reliable performance, the proposed model can be incorporated fruitfully in continuous hydrologic modeling. Nevertheless, a thorough investigation is required for watersheds of other biomes. In addition, the current study is limited to the prediction of runoff from large storm rainfall events ($P \geqslant 25.4$ mm). The application of the proposed model can be tested in future studies by estimating runoff from small storm rainfall events ($P < 25.4$ mm). It could be more interesting to compare runoff estimation based on watershed scale and the partial source areas.

Acknowledgments: Research for this paper was supported by a grant [MPSS-NH-2015-79] from the Natural Hazard Mitigation Research Group funded by the Ministry of Public Safety and Security, Korea. Special thanks to Jae-Hyun Ahn at Seokyeong University for motivating this research.

Author Contributions: Tae-Woong Kim supervised the research by managing the research results and data collected from different organizations; Muhammad Ajmal had the original idea, performed the experimental work and statistical analysis, and prepared the manuscript; Taj Ali Khan discussed and reviewed the manuscript.

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

References

1. Mishra, S.K.; Sahu, R.K.; Eldho, T.I.; Jain, M.K. An improved I_a-S relation incorporating antecedent moisture in SCS-CN methodology. *Water Resour. Manag.* **2006**, *20*, 643–660. [CrossRef]
2. Chung, W.H.; Wang, I.T.; Wang, R.Y. Theory-based SCS-CN method and its applications. *J. Hydrol. Eng.* **2010**, *15*, 1045–1058. [CrossRef]
3. Garen, D.C.; Moore, D.S. Curve number hydrology in water quality modeling: Uses, abuses, and future directions. *J. Am. Water Resour. Assoc.* **2005**, *41*, 377–388. [CrossRef]
4. Durbude, D.G.; Jain, M.K.; Mishra, S.K. Long-term hydrologic simulation using SCS-CN-based improved soil moisture accounting procedure. *Hydrol. Process.* **2011**, *25*, 561–579. [CrossRef]
5. Grimaldi, S.; Petroselli, A.; Romano, N. Green-Ampt curve number mixed procedure as an empirical tool for rainfall-runoff modelling in small and ungauged basins. *Hydrol. Process.* **2013**, *27*, 1253–1264. [CrossRef]
6. Patil, J.P.; Sarangi, A.; Singh, A.K.; Ahmad, T. Evaluation of modified CN methods for watershed runoff estimation using a GIS-based interface. *Biosyst. Eng.* **2008**, *100*, 137–146. [CrossRef]
7. Wang, X.; Liu, T.; Yang, W. Development of a robust runoff-prediction model by fusing the rational equation and a modified SCS-CN method. *Hydrol. Sci. J.* **2012**, *57*, 1118–1140. [CrossRef]
8. Garg, V.; Nikam, B.R.; Thakur, P.K.; Aggarwal, S.P. Assessment of the effect of slope on runoff potential of a watershed using NRCS-CN method. *Int. J. Hydrol. Sci. Technol.* **2013**, *3*, 141–159. [CrossRef]
9. Ponce, V.M.; Hawkins, R.H. Runoff curve number: Has it reached maturity? *J. Hydrol. Eng.* **1996**, *1*, 11–19. [CrossRef]
10. Schneider, L.E.; McCuen, R.H. Statistical guidelines for curve number generation. *J. Irrig. Drain. Eng.* **2005**, *131*, 282–290. [CrossRef]
11. Michel, C.; Andréassian, V.; Perrin, C. Soil conservation service curve number method: How to mend a wrong soil moisture accounting procedure. *Water Resour. Res.* **2005**, *41*, 1–6. [CrossRef]
12. Sahu, R.K.; Mishra, S.K.; Eldho, T.I.; Jain, M.K. An advanced soil moisture accounting procedure for SCS curve number method. *Hydrol. Process.* **2007**, *21*, 2872–2881. [CrossRef]
13. Massari, C.; Brocca, L.; Barbetta, S.; Papathanasiou, C.; Mimikou, M.; Moramarco, T. Using globally available soil moisture indicators for flood modelling in Mediterranean catchments. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 839–853. [CrossRef]
14. Ajmal, M.; Kim, T.-W. Quantifying excess stormwater using SCS-CN-based rainfall runoff models and different curve number determination methods. *J. Irrig. Drain. Eng.* **2015**, *141*, 04014058. [CrossRef]
15. Deshmukh, D.S.; Chaube, U.C.; Hailu, A.E.; Gudeta, D.A.; Kassa, M.T. Estimation and comparison of curve numbers based on dynamic land use land cover change, observed rainfall-runoff data and land slope. *J. Hydrol.* **2013**, *492*, 89–101. [CrossRef]
16. Hawkins, R.H.; Ward, T.J.; Woodward, D.E.; van Mullem, J.A. *Curve Number Hydrology-State of Practice*; The American Society of Civil Engineers (ASCE): Reston, VA, USA, 2009.
17. Natural Resources Conservation Service. Chapter 10: Hydrology. *National Engineering Handbook, Supplement A, Section 4*; The United States Department of Agriculture: Washington, DC, USA, 2004.
18. Seibert, J. On the need for benchmarks in hydrological modelling. *Hydrol. Process.* **2001**, *15*, 1063–1064. [CrossRef]
19. Tramblay, Y.; Bouaicha, R.; Brocca, L.; Dorigo, W.; Bouvier, C.; Camici, S.; Servat, E. Estimation of antecedent wetness conditions for flood modelling in northern Morocco. *Hydrol. Earth Syst. Sci.* **2001**, *16*, 75–86.
20. Woodward, D.E.; Hawkins, R.H.; Jiang, R.; Hjelmfelt, A.T., Jr.; van Mullem, J.A.; Quan, D.Q. Runoff curve number method: Examination of the initial abstraction ratio. In Proceedings of the World Water and Environmental Resources Congress, Philadelphia, PA, USA, 23–26 June 2003.
21. 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]
22. D'Asaro, F.; Grillone, G. Empirical investigation of curve number method parameters in the Mediterranean area. *J. Hydrol. Eng.* **2012**, *17*, 1141–1152. [CrossRef]
23. Perrin, C.; Michel, C.; Andréassian, V. Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.* **2003**, *279*, 275–289. [CrossRef]
24. Yuan, Y.; Nie, W.; McCutcheon, S.C.; Taguas, E.V. Initial abstraction and curve numbers for semiarid watersheds in southeastern Arizona. *Hydrol. Process.* **2014**, *28*, 774–783. [CrossRef]

25. Yuan, Y.; Mitchell, J.K.; Hirschi, M.C.; Cooke, R.A.C. Modified SCS curve number method for predicting subsurface drainage flow. *Trans. Am. Soc. Agric. Eng.* **2001**, *44*, 1673–1682. [CrossRef]
26. Brocca, L.; Melone, F.; Moramarco, T.; Singh, V.P. Assimilation of observed soil moisture data in storm rainfall-runoff modeling. *J. Hydrol. Eng.* **2009**, *14*, 153–165. [CrossRef]
27. Tramblay, Y.; Bouvier, C.; Martin, C.; Didon-Lescot, J.F.; Todorovik, D.; Domergue, J.M. Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling. *J. Hydrol.* **2010**, *387*, 176–187. [CrossRef]
28. Beck, H.E.; de Jeu, R.A.M.; Schellekens, J.; van Dijk, A.I.J.M.; Bruijnzeel, L.A. Improving curve number based storm runoff estimates using soil moisture proxies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 250–259. [CrossRef]
29. Massari, L.; Brocca, L.; Moramarco, T.; Tramblay, Y.; Lescot, J.-F.D. Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall. *Adv. Water Res.* **2014**, *74*, 44–53. [CrossRef]
30. Harmel, R.D.; Smith, P.K.; Migliaccio, K.W.; Chaubey, I.; Douglas-Mankin, K.R.; Benham, B.; Shukla, S.; Muñoz-Carpena, R.; Robson, B.J. Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. *Environ. Model. Softw.* **2014**, *57*, 40–51. [CrossRef]
31. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
32. Chen, X.Y.; Chau, K.W.; Busari, A.O. A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Eng. Appl. Artif. Intell.* **2015**, *46*, 258–268. [CrossRef]
33. Wu, C.L.; Chau, K.W.; Li, Y.S. Methods to improve neural network performance in daily flows prediction. *J. Hydrol.* **2009**, *372*, 80–93. [CrossRef]
34. Nash, J.; Sutcliffe, J. River flow forecasting through conceptual models part I-A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
35. Moriasi, D.N.; Arnold, J.G.; van Liew, M.W.; Binger, R.L.; Harmel, R.D.; Veith, T. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. Am. Soc. Agric. Biosyst. Eng.* **2013**, *50*, 885–900.
36. Ritter, A.; Muñoz-Carpena, R. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *J. Hydrol.* **2013**, *480*, 33–45. [CrossRef]
37. EasyFit Professional (Version 5.5). Mathwave Technologies, 2004–2010. Available online: www.mathwave.com (accessed on 15 July 2015).

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).