Article

Multiple Linear Regression Models for Predicting Nonpoint-Source Pollutant Discharge from a Highland Agricultural Region

Jae Heon Cho [1,*] and Jong Ho Lee [2]

[1]   Department of Biosystems and Convergence Engineering, Catholic Kwandong University, 24 Beomil-ro, 579 beon-gil, Gangneung-si, Gangwon-do 25601, Korea

[2]   Department of Urban Planning and Real Estate, Cheongju University, 298 Daeseongro, Cheongwon-gu, Cheongju, Chungbuk 28503, Korea; jhlee1013@cju.ac.kr

*   Correspondence: jhcho@cku.ac.kr; Tel.: +82-33-6497534

Abstract: Sediment runoff from dense highland field areas greatly affects the quality of downstream lakes and drinking water sources. In this study, multiple linear regression (MLR) models were built to predict diffuse pollutant discharge using the environmental parameters of a basin. Explanatory variables that influence the sediment and pollutant discharge can be identified with the model, and such research could play an important role in limiting sediment erosion in the dense highland field area. Pollutant load per event, event mean concentration (EMC), and pollutant load per area were estimated from stormwater survey data from the Lake Soyang basin. During the wet season, heavy rains cause large amounts of suspended sediment and the occurrence of such rains is increasing due to climate change. The explanatory variables used in the MLR models are the percentage of fields, subbasin area, and mean slope of subbasin as topographic parameters, and the number of preceding dry days, rainfall intensity, rainfall depth, and rainfall duration as rainfall parameters. In the MLR modeling process, four types of regression equations with and without log transformation of the explanatory and response variables were examined to identify the best performing regression model. The performance of the MLR models was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), coefficient of variation of the root mean square error (CV(RMSE)), the ratio of the RMSE to the standard deviation of the observed data (RSR) and the Nash–Sutcliffe model efficiency (NSE). The performance of the MLR models of pollutant load except total nitrogen (TN) was good under the condition of RSR, and satisfactory for the NSE and R^2. In the EMC and load/area models, the performance for suspended solids (SS) and total phosphorus (TP) was good for the RSR, and satisfactory for the NSE and R^2. The standardized coefficients for the models were analyzed to identify the influential explanatory variables in the models. In the final performance evaluation, the results of jackknife validation indicate that the MLR models are robust.

Keywords: highland agricultural field area; diffuse pollutant discharge; multiple regression model; climate change; jackknife validation

1. Introduction

In Lake Soyang basin of South Korea, large amounts of sediment are discharged from highland agricultural field regions in the wet season. To develop environmental preservation measures that protect water resources from the turbid water problem and diffuse pollution, prediction models are necessary to estimate the amount of pollutants discharged from subbasins. In this study, a multiple linear regression (MLR) model is established to predict the pollutant runoff discharge using environmental parameters, such as land use, rainfall, and topography.

In South Korea, rainfall events of 200 mm or more occurred only once annually, on average, until the end of the 1970s, but increased to a frequency of two per year in the 1980s and thereafter occurred five times in both 1984 and 1998. And, annual precipitation increased by 19% in the past decade, compared to the first half of the 20th century [1]. Conditions in Lake Soyang, located in the upper reaches of the Han River, greatly affect the water quality of the water supply of the capital region of South Korea. Discharged sediments from the highland field area flow into Lake Soyang in the wet season. Consequently, the turbidity of the lake increases to high levels and persists for a long time. In July 2006, a heavy rain event occurred in the Lake Soyang watershed. Overall, 670 mm fell over 8 days, with a maximum intensity of 66 mm per hour. The suspended sediment stayed in Lake Soyang for an extended period of time because of stratification; thus, the turbidity of the lake remained high and was measured at over 20 nephelometric turbidity units (NTU) for 168 days [2,3].

Regression models have been developed to estimate the sediment discharge using the subbasin environmental parameters in many areas. Valtanen et al. [4] applied stepwise multiple linear regression (SMLR) analysis to identify the variables that best explained the variation in event mass loads (EMLs) in each study catchment during cold and warm periods. Runoff duration, peak flow, antecedent dry period, mean runoff intensity, total suspended solids (TSS), TN, TP and total organic carbon (TOC) were used as explanatory variables. Another SMLR analysis was also carried out to assess whether catchment variables explain the EML and EMC values during the cold and warm periods [4]. The catchment variables included total impervious area and land use type. All data were log10-transformed to obtain approximately normal distributions. Bian et al. [5] proposed a procedure combining different statistical methods and a hydrological model to quantify the annual runoff response to spatial and temporal variations in impervious surface areas in an urbanized basin. A hydrological model relating annual runoff depth to precipitation, potential evapotranspiration and spatial metrics of the impervious area for baseline periods and periods of change was built using stepwise multiple regression analysis. Roman et al. [6] developed multivariate regression models to enable the prediction of mean annual suspended sediment discharge on the basis of basin characteristics, which is useful for many ungaged river locations in the eastern United States. The models are based on long-term mean sediment discharge estimates and explanatory variables, such as drainage area, mean elevation, and urban area, obtained from a combined dataset of 1201 US Geological Survey (USGS) stations. Tuset et al. [7] analyzed rainfall, runoff and sediment transport relationships in a meso-scale Mediterranean mountain catchment. The relationships among rainfall, runoff and suspended sediment transport were analyzed with Pearson correlations and multivariate regression analysis. The multivariate regression method was used to analyze the relationship between the independent variables (pre-event conditions, rainfall and runoff) and suspended sediment transport for all flood events. Seasonal relationships between total surface runoff and total sediment transport indicate that the sediment transport magnitude shows a clear seasonality influenced by rainfall intensity and sediment availability.

Buendia et al. [8] attempted to use empirical relationships to assess the relationship between sediment yield and basin scale and to provide an update on the main drivers controlling sediment yield in these particular river systems. Quantile regression analysis was used to assess the correlation between basin area and sediment yield, while additional basin-scale descriptors were related to sediment yield by means of multiple regression analysis. The performance of the model was tested through the jackknife validation method [8–14]. Paule-Mercado et al. [15] used MLRs to identify the significant parameters affecting fecal-indicator bacteria concentrations and to predict the response of bacteria concentrations to changes in land use and land cover. Stormwater temperature, 5-day biochemical oxygen demand (BOD$_5$), turbidity, TSS, and antecedent dry days were the most influential independent variables for the bacteria concentrations at the monitoring sites. Several studies have utilized linear regression techniques to predict bacteria concentrations in rivers [16–19]. Furthermore, regression models have been widely used to predict and characterize rainfall and runoff characteristics and to determine the

relationship between these two variables [20–26]. Process-based erosion prediction models have also been established to predict the intensity of soil erosion in a particular area [27–30].

In this study area, two types of environmental parameters affect the stormwater sediment runoff: meteorological factors, such as rainfall depth, rainfall intensity, rainfall duration; and number of preceding dry days and topographic factors, such as percentage of upland field area, subbasin area, and subbasin slope. In this study, the Pearson correlation test was employed to identify the linear relationship between the explanatory environmental parameters and the observed stormwater discharge. SMLR analysis was applied to identify the best performing regression model. Four types of regression equations were examined to determine the best MLR model. Explanatory and dependent variables with and without log e-transformation were tested. Then, the MLR models were validated via a jackknife validation procedure.

2. Materials and Methods

2.1. Study Area and Field Data

Lake Soyang formed following the construction of the Soyang River Dam. The dam was built to provide irrigation water, flood control and hydroelectric power. The dam has a height of 123 m, a length of 530 m, and a total storage capacity of 2900 million m^3. The basin area is 2969.3 km^2; the forest occupies 86.4%; and the dry field, paddy field and residential areas occupy 4.4%, 1.58%, and 1.60% of the basin, respectively.

In Lake Soyang basin (Figure 1), sediment discharge occurred mainly from the upper part of the basin. Land use in the tributary watersheds in the dense highland upland field area is shown in Table 1. In the case of Mandaecheon, the percentage of agricultural area is 27.7%, and upland fields represent 75% of the agricultural land. The Jungjohangcheon, Johangcheon and Jauncheon subbasins contain very small paddy field areas. In the highland area, to decrease the damage caused by the continuous cultivation of economic crops, manage pests, maximize crop productivity and improve soil fertility, 30–50 cm of soil dressing has been applied to the top layer of soil. This soil dressing is a significant contributor to the sediment discharge from the highland fields.

During rainfall, water sampling and flow measurements were performed at the same time. Field surveys were conducted to perform flow measurements at most of the survey points. For the remaining points, such as Inbukcheon, Bukcheon, and Soyang River, real-time water level and flow measurements were obtained from the Ministry of Land Infrastructure and Transport and the Korea Water Resources Corporation. The sampled water from the measurement sites was delivered to the laboratory as quickly as possible, and BOD, chemical oxygen demand (COD), SS, TP, TN, and TOC were analyzed using standard methods [2].

Table 1. Land use of the tributary watersheds in the dense highland fields area.

Stream	Subbasin Area (ha)	Land Use				
		Forest (ha)	Upland Field (ha)	Paddy Field (ha)	Others (ha)	Proportion of Agricultural Land (%)
Jungjohangcheon	1022	860	150	0	12	14.6
Johangcheon	4161	3556	489	0.3	117	12.0
Jauncheon	13,641	11,703	1445	0.4	493	11.0
Mandaecheon	6079	1261	1261	420	3137	27.7
Gaahcheon	4732	1104	578	298	2752	18.5

In the statistical analysis of this study, the results of stormwater runoff surveys from 2013 to 2016 at nine points in the Jaun, Mandae and Gaha area [31] were used. Of the 79 rainfall events, nine data were too high or too low for rainfall amount due to runoff load measurements and calculation errors; those data were excluded. The discharge load survey was performed from the beginning of the rainfall to the point where it returned to the normal water level after the end of the rainfall. Runoff discharge

data of 70 storm events were used to build the MLR models to predict the pollutant load, EMC and pollutant load per area. The range of the rainfall depth used in the MLR model construction was from 10 mm to 215 mm.
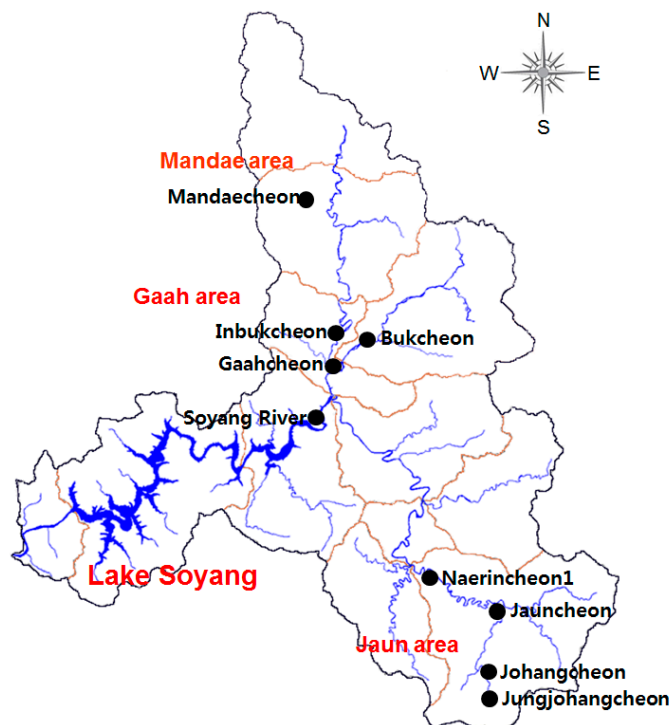

Figure 1. Soyang River basin and stormwater survey sites.

2.2. Data Analysis

Using water quality and runoff flow data from 70 rainfall events in the Lake Soyang basin, pollutant load per event, EMC, and pollutant load per area were estimated for each rainfall event. The total pollutant load during a rainfall event was calculated using Equation (1). The EMC was defined as the pollutant mass contained in the runoff event divided by the total flow volume of the event. The total pollutant load was divided by the subbasin area to estimate the pollutant load per area.

$$Total\ pollution\ load/rainfall\ event = \sum_{i=1}^{n} C_i Q_i \Delta t_i \tag{1}$$

$$EMC = \frac{\sum Q_i C_i \Delta t}{\sum Q_i \Delta t} \tag{2}$$

where *n* represents the number of total measurements, Q_i is the runoff flow at n number of time steps (Δt) and C_i is the concentration of a water quality measurement.

The distribution of the nonpoint pollutant discharge for the 70 rainfall events from 2013 to 2016 is presented in Table 2. Figure 2 shows box plots of pollutant load per event, EMC, and pollutant load per area at the survey points in the Lake Soyang basin. The maximum, minimum and median values of the suspended sediment (SS) load/event were 46,125,100 kg, 613 kg and 263,083 kg, respectively. The maximum, minimum and median values of the TP load/event were 32,406 kg, 1.83 kg and 480 kg, respectively. As shown in Figure 2, all the mean values of the pollutant loads are larger than the values of the third quartile, and the distributions are biased toward the high values. The maximum, minimum and median values of the SS EMC were 1437 mg/L, 3.8 mg/L and 157 mg/L, respectively. The maximum, minimum and median values of the TP EMC were 1.96 mg/L, 0.011 mg/L and

0.27 mg/L, respectively. All the mean values of the EMCs lie between the 50th percentile and the 75th percentile, and the distributions are relatively uniform.

　　The explanatory variables that are considered to explain the nonpoint pollutant discharge in the MLR models are the percentage of fields (% field), subbasin area (SA), and mean slope of subbasin (slope) as topographic parameters, and the number of preceding dry days (Ndry), rainfall intensity (Rint), rainfall depth (Rain), and rainfall duration (Dur) as rainfall parameters (Table 3).

(a)

(b)

(c)
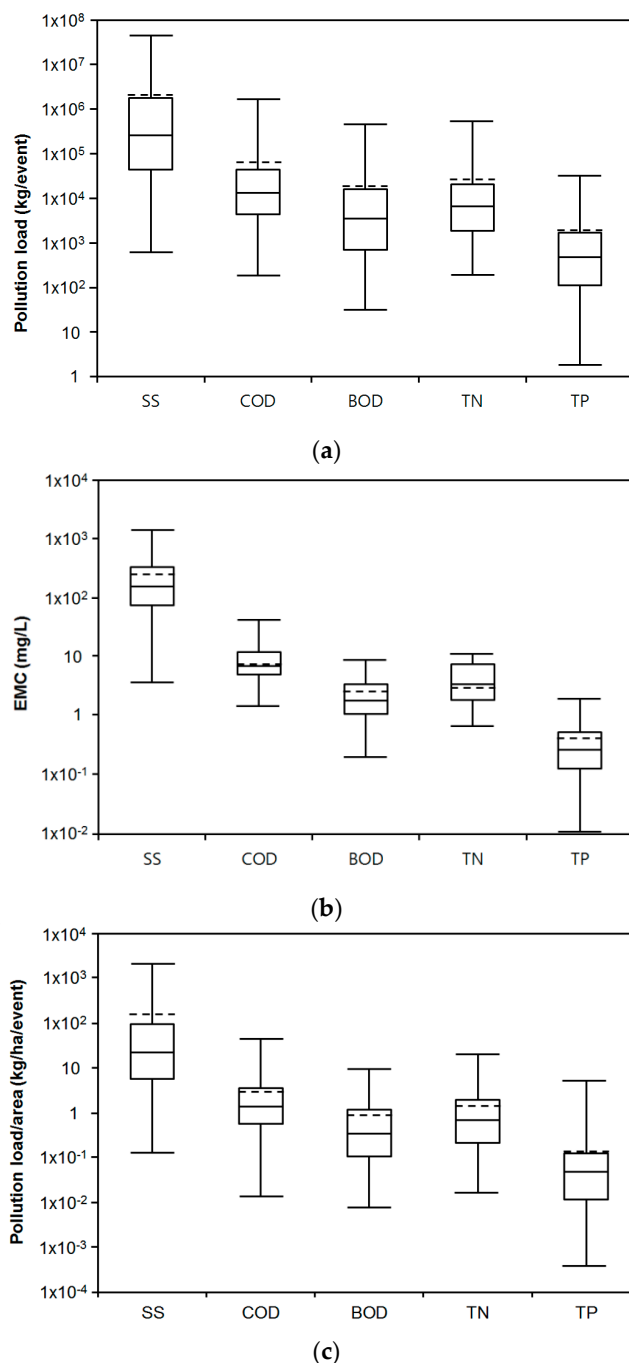
Figure 2. Box plots of nonpoint pollutant discharge in the Lake Soyang basin. The top (**a**) and bottom (**c**) of each box represent the third and first quartiles, the solid line inside the box is the second quartile (**b**), and the dotted line inside the box is the mean. One whisker stretches from the third quartile to the maximum, and the other whisker stretches from the first quartile to the minimum.

Table 2. Distribution of the nonpoint pollutant discharge for the 70 rainfall events in the Lake Soyang basin.

Pollutant	Min.	25th Percentile	50th Percentile	75th Percentile	Max.	Mean
SS load (kg)	613	44,839	263,083	1,802,409	46,125,100	1,843,969
COD load (kg)	186	4410	13,192	44,608	1,686,594	83,565
BOD load (kg)	31	689	3477	15,982	456,773	20,233
TN load (kg)	187	1852	6563	20,745	541,563	28,277
TP load (kg)	1.8	111	480	1685	32,406	2010
SS (EMC) (mg/L)	3.8	75.6	157	338	1437	266
COD (EMC) (mg/L)	1.5	5.13	7.21	12.2	43.6	9.16
BOD (EMC) (mg/L)	0.20	1.1	1.85	3.47	9.0	2.49
TN (EMC) (mg/L)	0.67	1.89	3.48	7.63	11.4	4.75
TP (EMC) (mg/L)	0.011	0.13	0.27	0.54	1.96	0.37
SS (load/area) (kg/ha)	0.129	5.81	22.6	95.37	2118	130
COD (load/area) (kg/ha)	0.0139	0.57	1.38	3.54	44.0	3.12
BOD (load/area) (kg/ha)	0.0078	0.10	0.34	1.18	9.5	0.87
TN (load/area) (kg/ha)	0.0164	0.21	0.69	1.93	20.4	1.54
TP (load/area) (kg/ha)	0.00038	0.012	0.047	0.12	5.19	0.16

Table 3. Explanatory variables considered in the regression models to predict pollutant discharge.

Variables	Description	Units
% field	Percentage of fields	%
SA	Subbasin area	km^2
Ndry	Number of preceding dry days	day
Rint	Rainfall intensity	mm/h
Slope	Mean slope of the subbasin	°
Rain	Rainfall depth	mm
Dur	Rainfall duration	h

A Pearson correlation coefficient matrix was used to identify the correlations among the surveyed pollutant discharge estimates and the explanatory variables. The correlations among natural log-transformed variables were also tested using Pearson correlation.

2.3. MLR Model Building

MLR modeling was performed to predict the pollutant discharge from the subbasins in the Soyang River. The models were built to explain the pollutant discharge using the subbasin topographic and rainfall data. In the MLR modeling, four types of regression equations are examined:

$$\text{Type1}: Y = a_0 + \sum_{i=1}^{n} a_i X_i \tag{3}$$

$$\text{Type2}: Ln(Y) = a_0 + \sum_{i=1}^{n} a_i X_i \tag{4}$$

$$\text{Type3}: Y = e^{a_0} X_1^{a_1} X_2^{a_2} \cdots X_n^{a_n} \text{ or } Ln(Y) = a_0 + \sum_{i=1}^{n} (a_i Ln(X_i)) \tag{5}$$

$$\text{Type4}: Y = e^{a_0} X_1^{a_1} X_2^{a_2} \cdots X_m^{a_m} e^{a_{m+1} X_{m+1}} \cdots e^{a_n X_n} \text{ or } Ln(Y) = a_0 + \sum_{i=1}^{m} (a_i Ln(X_i)) + \sum_{i=m+1}^{n} (a_i X_i) \tag{6}$$

where a_0 is the regression constant and a_i is the regression coefficient of the explanatory variable X_i. In type 1, the original variables are used to build the MLR model. In type 2, dependent variables, such as pollutant load, EMC and load/area, are log e-transformed to reduce skewness. In type 3, all the explanatory and dependent variables are log e-transformed. In type 4, the dependent variables and some of the explanatory variables are log e-transformed. The fitness of the four regression equations was evaluated by the coefficients of determination of the MLR models. The MLR models

were examined in terms of their ability to predict the runoff pollutant discharge for each water quality variable (SS, COD, BOD, TN, and TP).

Collinearity may introduce serious stability problems, such as high mean square errors, in a regression model. Therefore, the collinearity of the predictor variables in the created MLR model were tested by calculating the variance inflation factor (VIF) [32]. Collinearity is present when the largest VIF is greater than 10 or the average VIF value is substantially greater than 1 [32,33]. VIFs were calculated to analyze the multicollinearity in this MLR model.

The MLR model performance was evaluated using the R^2, RMSE, CV(RMSE), RSR and the NSE.

$$R^2 = \frac{\left[\sum_{i=1}^{n} \left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)\right]^2}{\sum_{i=1}^{n} \left(P_i - \overline{P}\right)^2 \sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2} \tag{7}$$

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^{n} \left(P_i - O_i\right)^2\right]^{1/2} \tag{8}$$

$$CV(RMSE) = \frac{\left[\sum_{j=1}^{n} \left(P_{ij} - O_{ij}\right)^2 / n\right]^{1/2}}{\left(\sum_{i=1}^{n} O_{ij} / n\right)} \tag{9}$$

$$RSR = \frac{\left[\sqrt{\sum_{i=1}^{n} \left(O_i - P_i\right)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2}\right]} \tag{10}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n} \left(O_i - P_i\right)^2}{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2} \tag{11}$$

where O_i is the observed daily load, $\overline{O}$ is the mean of the observed daily load, p_i is the calculated daily load, and n is the number of data values. The R^2 index describes the ability of the model to explain variability among the data. RSR incorporates the benefits of error index statistics and includes a scaling/normalization factor; the lower the RSR is, the better the model simulation performance. The performance ratings for stream flow proposed by Moriasi et al. [33] were 'very good' ($0.00 \leq$ RSR ≤ 0.50), 'good' ($0.50 <$ RSR ≤ 0.60), or 'satisfactory' ($0.60 <$ RSR ≤ 0.70). NSE is a normalized statistic that reflects the relative magnitude of the residual variance compared with the variance in the observed data (good (NSE > 0.7), satisfactory ($0.4 <$ NSE ≤ 0.7) and unsatisfactory (NSE ≤ 0.4)) [30,34].

Finally, the performance of the MLR model was tested using the jackknife validation method [8,11,14]. This method consists of deleting one site and carrying out the multiple regression analysis with the same dependent variables and the remaining sites. The pollutant discharge of the deleted site is calculated with the equation resulting from the multiple regression associated with the remaining sites. This process is repeated, deleting one site each time.

3. Results and Discussion

3.1. Correlation Analysis between Nonpoint Pollutant Discharge and Explanatory Variables

Table 4 shows the Pearson correlation between runoff discharge and subbasin characteristics, without log transformation of the variables. In Table 5, the Pearson correlation matrix between log-transformed variables is introduced. As the values of the correlation coefficients between log-transformed variables were slightly higher (r < 0.69; Table 5) than those of the non-log-transformed variables (r < 0.65; Table 4), we used log-transformed variables as explanatory variables in the MLR models. Compared to other environmental parameters, the rainfall depth and subbasin area showed a relatively significant correlation with most response discharge variables. The rainfall intensity had a relatively significant positive correlation with the response variables of EMC and load/area because

the rainfall intensity directly affects the EMC and load/area of each storm event. On the other hand, the subbasin slope had a negative correlation with the response variables of EMC and load/area.

Table 4. Pearson correlation matrix between stormwater runoff discharge and subbasin characteristics.

Variables	% field	SA	Rain	Dur	Ndry	Rint	Slope
SS (load)	−0.187	0.102	**0.524**	0.236	−0.135	0.260	0.088
COD (load)	−0.211	**0.563**	**0.317**	0.297	0.015	0.022	0.239
BOD (load)	−0.217	**0.458**	**0.352**	**0.353**	−0.080	0.011	0.213
TN (load)	−0.214	**0.488**	**0.374**	**0.316**	−0.053	0.057	0.220
TP (load)	−0.205	**0.417**	**0.514**	**0.356**	−0.150	0.160	0.178
SS (EMC)	**0.498**	−0.293	**0.387**	0.181	−0.115	0.251	**−0.585**
COD (EMC)	0.125	−0.199	0.166	−0.098	−0.099	**0.397**	−0.150
BOD (EMC)	0.120	**−0.317**	0.187	0.095	−0.049	0.147	−0.227
TN (EMC)	0.196	−0.065	0.157	−0.022	0.227	0.195	−0.207
TP (EMC)	**0.355**	**−0.366**	0.223	−0.115	−0.216	**0.404**	**−0.391**
SS (load/area)	0.198	−0.166	**0.632**	**0.313**	−0.185	0.283	−0.240
COD (load/area)	0.102	−0.095	**0.599**	**0.367**	−0.175	0.224	−0.108
BOD (load/area)	0.077	−0.147	**0.652**	**0.476**	−0.212	0.172	−0.124
TN (load/area)	0.132	−0.180	**0.583**	**0.348**	−0.172	0.210	−0.137
TP (load/area)	0.108	−0.114	**0.441**	0.190	−0.148	0.212	−0.106

Note: Bold marked correlations are significant at $p < 0.01$.

Table 5. Pearson correlation matrix between natural log-transformed stormwater runoff discharge and subbasin characteristics.

Variables	ln(% field)	ln(SA)	ln(Rain)	ln(Dur)	ln(Ndry)	ln(Rint)	Slope
ln(SS(load))	−0.12	**0.42**	**0.58**	**0.44**	−0.12	0.29	−0.16
ln(COD(load))	**−0.38**	**0.69**	**0.43**	**0.49**	−0.12	0.08	0.25
ln(BOD(load))	**−0.34**	**0.60**	**0.51**	**0.49**	−0.11	0.18	0.14
ln(TN(load))	**−0.32**	**0.61**	**0.47**	**0.48**	−0.12	0.14	0.17
ln(TP(load))	−0.16	**0.48**	**0.59**	**0.43**	−0.16	0.30	−0.06
ln(SS(EMC))	**0.40**	**−0.37**	**0.47**	0.03	−0.11	**0.48**	**−0.63**
ln(COD(EMC))	0.19	**−0.40**	0.23	−0.12	−0.06	**0.33**	−0.14
ln(BOD(EMC))	0.20	**−0.44**	0.37	−0.06	−0.02	**0.44**	−0.36
ln(TN(EMC))	**0.33**	**−0.49**	0.23	−0.13	0.09	**0.35**	−0.38
ln(TP(EMC))	**0.45**	**−0.56**	0.36	−0.17	−0.08	**0.52**	−0.59
ln(SS(load/area))	**0.34**	−0.30	**0.64**	0.28	−0.16	**0.47**	**−0.54**
ln(COD (load/area))	0.17	−0.18	**0.62**	**0.39**	−0.22	**0.37**	−0.23
ln(BOD (load/area))	0.18	−0.24	**0.65**	**0.35**	−0.18	**0.43**	−0.33
ln(TN(load/area))	0.30	**−0.37**	**0.59**	0.29	−0.19	**0.41**	−0.36
ln(TP(load/area))	**0.37**	**−0.37**	**0.65**	0.24	−0.21	**0.52**	**−0.52**

Note: Bold marked correlations are significant at $p < 0.01$.

3.2. MLR Analysis

Four types of MLR models corresponding to Equations (3)–(6) were tested to identify the most suitable models (Table 6). The R^2 values for SS, COD, BOD, TN, and TP in the type 1 MLR of pollutant load ranged from 0.275 to 0.447. The R^2 values for SS, COD, BOD, TN, and TP in the type 1 MLR of EMC and load/area were also low, indicating poor performance of the regression models. The R^2 values for SS, COD, BOD, TN, and TP in the type 2 MLR of pollutant load were 0.76, 0.67, 0.64, 0.65, and 0.80, respectively. The R^2 values of the type 2 MLR were quite high, but most of the VIF values were larger than 5, with a few values greater than 10. Thus, the VIF showed that multicollinearity was observed in the established models and that the type 2 MLR was not adapted. Although the R^2 values of the type 2 MLR for load/area were acceptable, the VIF values were high, indicating multicolinearity. VIF values and other statistics of MLRs were presented only for the selected model. The results of the MLR model employing the type 4 equation are listed in Tables 7–9.

Table 6. Coefficients of determination from four types of MLR analysis.

Runoff Discharge Type	MLR Type	SS	COD	BOD	TN	TP
Load	Type 1	0.275	0.425	0.340	0.386	0.447
	Type 2	0.764	0.672	0.641	0.654	0.801
	Type 3	0.720	0.687	0.688	0.614	0.689
	Type 4	0.736	0.687	0.694	0.614	0.741
EMC	Type 1	0.477	0.157	0.100	0.254	0.33
	Type 2	0.646	0.123	0.273	0.321	0.584
	Type 3	0.536	0.226	0.324	0.539	0.592
	Type 4	0.655	0.226	0.324	0.539	0.662
Load/Area	Type 1	0.448	0.359	0.460	0.340	0.195
	Type 2	0.734	0.503	0.526	0.497	0.686
	Type 3	0.640	0.424	0.496	0.471	0.651
	Type 4	0.695	0.427	0.509	0.471	0.675

The R^2 values for SS, COD, BOD, TN, and TP in the type 3 MLR of the pollutant load were also fairly high, but all VIF values were less than 5. Among the type 3 MLR models, the SS, TN, and TP in the MLR of EMC and the SS and TP in the MLR of load/area showed acceptable R^2 values. The values of R^2 for SS, COD, BOD, TN, and TP in the type 4 MLR of pollutant load were 0.74, 0.69, 0.69, 0.61, and 0.74 respectively. The R^2 values of the type 4 MLR were a little better than those of the type 3 MLR, and all VIF values were less than 5. Thus, we selected the type 4 equation as the MLR model to predict the runoff pollutant discharge in the study area. However, the COD and BOD in the MLR of EMC and COD and TN in the MLR of load/area could not explain the variance in the pollutant discharge properly.

Using the stepwise variable selection method, two to five variables were retained in the pollutant load model, as shown in Table 7. In the case of the SS model, given the R^2 value, 73.6% of the variability of the dependent variable ln(SS load) is explained by the four explanatory variables. The MLR models indicated in Tables 7–9 are statistically significant at $p < 0.0001$ except for the ln(COD EMC) model ($p = 0.00019$). The R^2 values for SS, COD, BOD, TN, and TP in the type 4 MLR of pollutant load were fairly high ($0.614 < R^2 < 0.741$), as indicated in Table 7. The performance evaluation by CV(RMSE) [35] shows that the SS model was the best and that the other models of the water quality variables were also acceptable. The range of RSR for SS, COD, BOD, and TP in the MLR models of pollutant load (Table 7) was from 0.509 to 0.559, and the performance of the MLR for these variables was good [34]. The RSR for TN was 0.622, and the performance of the TN model was satisfactory. The NSE values for the MLR models of pollutant load ranged from 0.61 to 0.74, and the MLR models of the pollutant load had good performance. As a special case, in linear regression forecasting models like this study, NSE is equal to the coefficient of determination, R^2 [36]. Overall, all the MLR models of the pollutant load had good prediction performance.

All VIF values in Table 7 are lower than 5, and the mean VIF values are not large. These results suggest that the coefficient of regression for the explanatory variables could be statistically acceptable and that multicollinearity was not present in the established models.

Standardized coefficients refer to how many standard deviations a dependent variable will change in response to an increase of one standard deviation in the predictor variable. This statistic allows us to compare the relative contribution of each independent variable in the prediction of the dependent variable. The higher the absolute value of a coefficient is, the more important the weight of the corresponding variable. Standardized coefficients are useful for comparing effects across different measures. The standardized regression coefficients of Table 7 indicate that subbasin area ($0.576 < \beta_i < 0.709$) and rainfall depth ($0.453 < \beta_i < 0.563$) are important influential parameters for all the load predictions. In addition, % field has relatively small effects on the SS, BOD and TP models.

Table 7. MLR models for pollutant load discharge during stormwater events.

Response Variable	Explanatory Variables	a_0	a_i	β_i	VIF	DW	R^2	RMSE	CV(RMSE)	RSR	NSE
ln(SS load)	Intercept	12.50				1.773	0.736	1.230	0.099	0.514	0.736
	ln(% field)		−2.31	−0.40	4.017						
	ln(SA)		0.81	0.58	1.610						
	ln(Rain)		1.78	0.56	1.008						
	Slope		−0.25	−0.75	3.377						
ln(COD load)	Intercept	0.44				1.746	0.687	1.135	0.119	0.559	0.687
	ln(SA)		0.86	0.71	1.001						
	ln(Rain)		1.24	0.45	1.001						
ln(BOD load)	Intercept	4.92				1.938	0.694	1.172	0.145	0.553	0.694
	ln(% field)		−1.53	−0.30	4.017						
	ln(SA)		0.80	0.64	1.610						
	ln(Rain)		1.44	0.51	1.008						
	Slope		−0.12	−0.41	3.377						
ln(TN load)	Intercept	0.72				1.873	0.614	1.121	0.127	0.622	0.614
	ln(SA)		0.67	0.63	1.001						
	ln(Rain)		1.19	0.49	1.001						
ln(TP load)	Intercept	3.44				1.539	0.741	1.047	0.174	0.509	0.741
	ln(% field)		−1.36	−0.28	4.025						
	ln(SA)		0.77	0.64	1.620						
	ln(Rain)		1.50	0.56	1.039						
	ln(Ndry)		−0.24	−0.14	1.049						
	Slope (°)		−0.17	−0.60	3.423						

Notes: a_0 is the regression constant; a_i is the regression coefficient of the explanatory variable X_i; β_i is the standardized regression coefficient.

Table 8. MLR models for EMCs of stormwater events.

Response Variables	Explanatory Variables	a_0	a_i	β_i	VIF	DW	R^2	RMSE	CV(RMSE)	RSR	NSE
ln(SSEMC)	Intercept	10.94				1.654	0.655	0.875	0.180	0.587	0.655
	ln(% field)		−1.79	−0.50	4.017						
	ln(SA)		−0.17	−0.19	1.610						
	ln(Rain)		0.83	0.42	1.008						
	Slope		−0.19	−0.93	3.377						
ln(CODEMC)	Intercept	2.46				1.338	0.226	0.515	0.252	0.880	0.226
	ln(SA)		−0.12	−0.35	1.054						
	ln(Rint)		0.22	0.26	1.054						
ln(BODEMC)	Intercept	−0.07				1.129	0.324	0.723	1.223	0.822	0.324
	ln(SA)		−0.23	−0.43	1.001						
	ln(Rain)		0.42	0.36	1.001						
ln(TNEMC)	Intercept	2.47				0.843	0.539	0.504	0.384	0.679	0.539
	ln(SA)		−0.29	−0.65	1.054						
	ln(Rint)		0.25	0.23	1.054						
ln(TPEMC)	Intercept	4.62				0.919	0.662	0.723	−0.481	0.581	0.662
	ln(% field)		−1.14	−0.38	4.017						
	ln(SA)		−0.25	−0.34	1.693						
	ln(Ndry)		−0.19	−0.19	1.057						
	ln(Rint)		0.75	0.43	1.100						
	Slope		−0.12	−0.69	3.396						

Table 9. MLR models for pollutant load per area during stormwater events.

Response Variables	Explanatory Variables	a_0	a_i	β_i	VIF	DW	R^2	RMSE	CV(RMSE)	RSR	NSE
ln(SS load/area)	Intercept	5.70				1.728	0.695	1.246	0.408	0.552	0.695
	ln(% field)		−1.81	−0.33	3.365						
	ln(Rain)		1.79	0.60	1.007						
	Slope		−0.25	−0.79	3.376						
ln(COD load/area)	Intercept	−3.88				1.779	0.427	1.123	4.710	0.757	0.427
	ln(Rain)		1.22	0.61	1.003						
	Slope		−0.04	−0.19	1.003						
ln(BOD load/area)	Intercept	−5.65				1.927	0.509	1.206	−0.995	0.701	0.509
	ln(Rain)		1.47	0.63	1.003						
	Slope		−0.07	−0.29	1.003						
ln(TN load/area)	Intercept	−3.88				1.873	0.471	1.121	−2.108	0.727	0.471
	ln(SA)		−0.33	−0.36	1.001						
	ln(Rain)		1.19	0.58	1.001						
ln(TP load/area)	Intercept	−6.57				1.519	0.675	1.090	−0.328	0.570	0.675
	ln(Rain)		1.53	0.60	1.033						
	ln(Ndry)		−0.25	−0.15	1.036						
	Slope		−0.13	−0.49	1.011						

The area with the high density of highland fields in Lake Soyang basin has steeper slopes than the other areas. However, Lake Soyang basin also contains highly mountainous terrain; thus, the mean slopes of the dense highland field subbasins are lower than the average slope of the entire Lake Soyang basin. Therefore, the standardized regression coefficients of mean slope for the SS and TP load models have (−) signs, and the mean slope has a negative influence on the SS and TP loads.

The explanatory variables for the SS and TP models explained 65.5% and 66.2% of variation in the response variables of EMC. The R^2 values were fairly high, as indicated in Table 8. The R^2 value for the TN model of EMC was 0.539, and the TN model was acceptable [34]. The CV(RMSE) value of the BOD model was quite high, and the model was not acceptable. The RSR values for SS and TP in the MLR models of EMC (Table 8) were 0.587 and 0.581, respectively, and the performances of these models were good. The RSR for the TN model was 0.679, and the performance of the TN model was satisfactory. However, the RSR values for the COD and BOD models were high, and these models were unsatisfactory. The NSE values for the MLR models of the EMC show that the SS, TP, and TN models were satisfactory but that the COD and BOD models were not satisfactory. The VIF values for the EMC models were lower than 5, and the mean VIF values were not large. Overall, the MLR models for SS and TP have good prediction performance, and the TN model has acceptable performance.

The standardized regression coefficients in Table 8 indicate that rainfall intensity and rainfall depth are influential explanatory variables for the EMC response variables. Rainfall intensity ($0.234 < \beta_i < 0.426$) is an important factor for the TP, TN, and COD models, and rainfall depth is important for the SS and BOD models. In the pollutant load model, rainfall depth is a very important parameter, whereas rainfall intensity is not an important explanatory variable. However, rainfall intensity is an influential parameter for the EMC of a storm event. From the Pearson correlation matrix between natural log-transformed stormwater runoff discharge and subbasin characteristics in Table 5, we also can see that EMCs are better correlated to rainfall intensity than rainfall depth, and pollutant loads are much better correlated to rainfall depth than rainfall intensity. In agricultural areas such as the study area, the larger the rainfall intensity, the more nutrients are released from fertilizer and vegetation roots. The standardized regression coefficients of the mean slope for the SS and TP load models have (−) sign, and the mean slope has a large negative influence on the SS and TP EMC. Additionally, % field also has a negative impact on the SS and TP EMC.

The explanatory variables for the SS and TP models explained 69.5% and 67.5% of the variation in the load/area response variables, and the R^2 values were fairly high, as indicated in Table 9. The R^2 value for the BOD model of load/area was 0.51; thus, the BOD model was acceptable. The RSR values for SS and TP in the MLR models of load/area (Table 9) were 0.55 and 0.57, respectively, and the performances of these models were good. The RSR for the BOD model was 0.70, and the performance of the TN model was satisfactory. The NSE values in the MLR models of the load/area show that the SS, TP, and BOD models were satisfactory. The VIF values for the load/area models were less than 5, and the mean VIF values were not large. Overall, the MLR models of load/area for SS and TP have good performance, and the BOD model has acceptable performance.

The standardized regression coefficients in Table 9 indicate that rainfall depth ($0.576 < \beta_i < 0.634$) is a highly influential parameter for all response variables in the load/area prediction. The β coefficients of the mean slope for the SS and TP load/area models are −0.79 and −0.49, respectively, and the absolute values of the coefficients are comparable to the coefficients of rainfall depth, indicating that the mean slope is a remarkable negative parameter on the SS and TP load/area results.

3.3. Jackknife Validation of the MLR Model

The performance of the jackknife validation was evaluated using R^2, RSR and NSE (Table 10). The R^2 values were calculated by the linear regression between observed and jackknife validation values, and RSR and NSE were also calculated. The R^2 (Figure 3) and NSE values associated with the jackknife procedure were slightly lower than the results of the MLR models, whereas the RSR values were slightly higher than the MLR models. Therefore, the performance of the jackknife validation was

slightly worse than that of the MLR models. The results of jackknife validation indicate that the MLR models are robust.
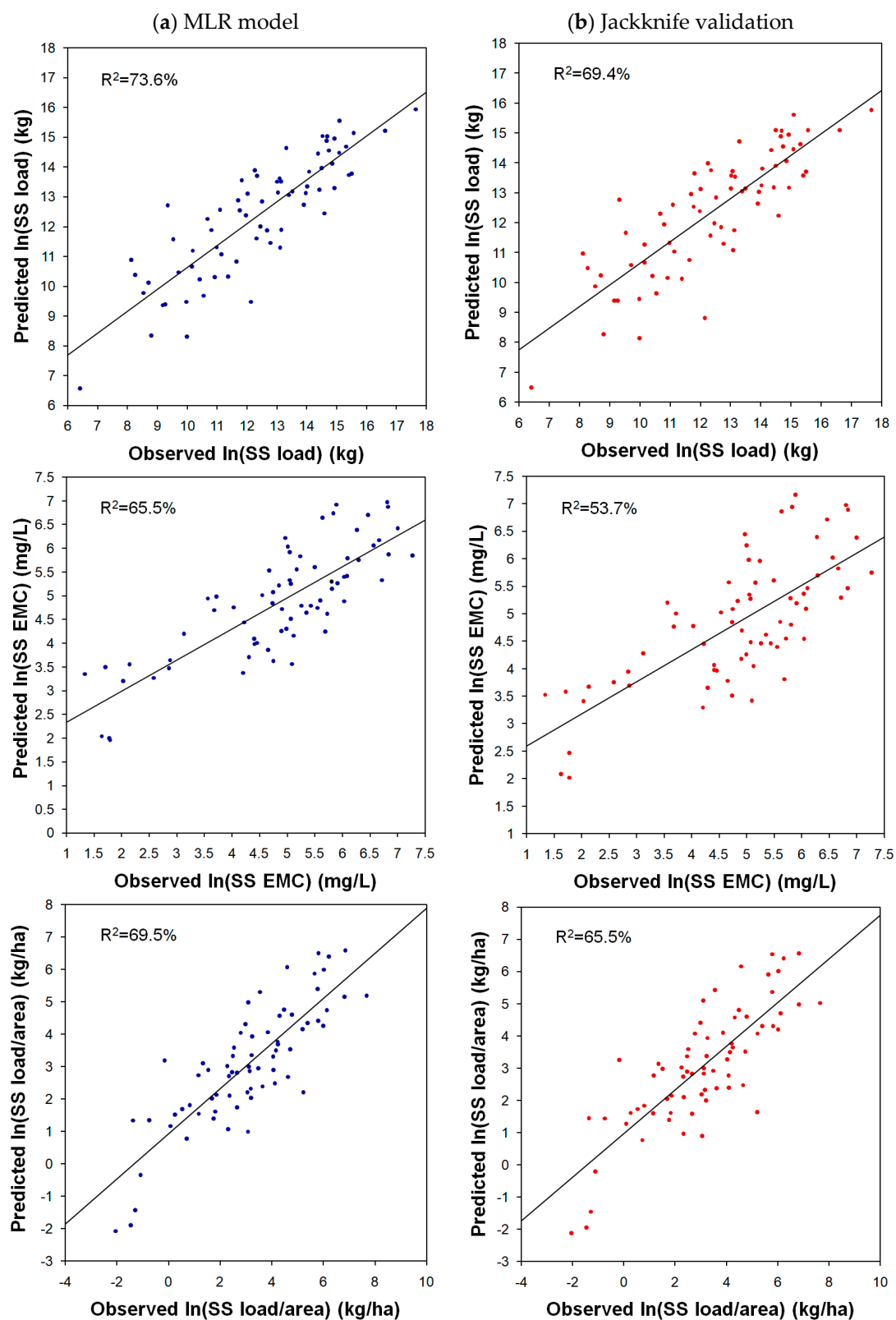
Figure 3. Comparison between the observed and predicted values during storm events based on (**a**) MLR models and (**b**) the results of jackknife validation.

Table 10. Three performance indicators for the stormwater runoff discharge values based on jackknife validation.

Response Variable (Jackknife)	R^2	RSR	NSE
ln(SS load)	0.694	0.554	0.693
ln(COD load)	0.630	0.611	0.627
ln(BOD load)	0.607	0.630	0.603
ln(TN load)	0.550	0.674	0.545
ln(TP load)	0.609	0.629	0.605
ln(SS EMC)	0.537	0.684	0.533
ln(COD EMC)	0.155	0.924	0.147
ln(BOD EMC)	0.211	0.894	0.202
ln(TN EMC)	0.503	0.730	0.468
ln(TP EMC)	0.601	0.633	0.599
ln(SS load/area)	0.655	0.588	0.654
ln(COD load/area)	0.305	0.845	0.287
ln(BOD load/area)	0.478	0.723	0.477
ln(TN load/area)	0.413	0.768	0.410
ln(TP load/area)	0.602	0.632	0.600

4. Conclusions

MLR models were built to predict the nonpoint-source pollutant discharge in the highland field area in the wet season using environmental parameters as explanatory variables. Runoff discharge data from 70 storm events were used to build the MLR models to predict the pollutant load, EMC and pollutant load per area. Pearson correlation tests were employed to identify the linear relationships between subbasin environmental parameters and the observed stormwater discharge. As the values of correlation coefficients between log-transformed variables were slightly higher than those of variables that had not been log transformed, the log-transformed variables were selected as explanatory variables in the MLR models.

The R^2 values for SS, COD, BOD, TN, and TP in the type 4 MLR of pollutant load were quite high (the best among the four examined MLR types), and all VIF values were less than 5. Thus, the type 4 equation was chosen as the MLR model to predict the runoff pollutant discharge.

The R^2 values for the five water quality variables in the MLR of pollutant load were fairly high ($0.614 < R^2 < 0.741$), and the RSR values for SS, COD, BOD, and TP in the MLR models of pollutant load ranged from 0.509 to 0.559. Hence, the performance of the MLR for these variables was good [34]. The RSR for TN was 0.622, and the performance of the TN model was satisfactory. The NSE values for the MLR models of the pollutant load indicated good performance. Hence, most of the MLR models of the pollutant load have good prediction performance.

The MLR models of EMC for SS and TP also have good prediction performance, and TN model has acceptable performance. The MLR models of load/area for SS and TP have relatively good performance, and the BOD model has acceptable performance. Based on the R^2, RSR and NSE values, the performance of the jackknife validation was slightly worse than that of the MLR models. Thus, the results of jackknife validation indicate that the MLR models are robust.

The results of the standardized coefficients for the MLR models indicate that subbasin area and rainfall depth are important influential parameters for all the load predictions. The mean slope exerts a negative influence on the SS and TP loads on account of topographic characteristics, as previously explained. For the pollutant load model, rainfall depth is a very important parameter, whereas rainfall intensity was not chosen as an explanatory variable. However, rainfall intensity has an influence on the EMC of the storm event. The mean slope has a large negative influence on the SS and TP EMC, and % field has a negative impact on the SS and TP EMC. Additionally, the rainfall depth is a highly influential parameter for all response variables of the load/area predictions, similar to the pollutant load models. The mean slope has a large negative influence on the SS and TP load/area.

The average slope of fields, rather than the average slope of the whole sub-basin, can be an important explanatory variable for the pollutant discharge load of each subbasin. Similar or even better MLR results for EMC could have been obtained using peak rainfall intensity as explanatory variables. Therefore, future studies on MLR need to consider this.

References

1. Kim, H.J.; Lee, K.K. A comparison of the water environment policy of Europe and South Korea in response to climate change. *Sustainability* **2018**, *10*, 384. [CrossRef]
2. Cho, J.H.; Lee, J.H. Stormwater runoff characteristics and effective management of nonpoint source pollutants from a highland agricultural region in the Lake Soyang watershed. *Water* **2017**, *9*, 784. [CrossRef]
3. Kim, B.; Jung, S. Turbid storm runoffs in Lake Soyang and their environmental effect. *J. Korean Soc. Environ. Eng.* **2007**, *29*, 1185–1190.
4. Valtanen, M.; Sillanpää, N.; Setälä, H. Key factors affecting urban runoff pollution under cold climatic conditions. *J. Hydrol.* **2015**, *529*, 1578–1589. [CrossRef]
5. Bian, G.D.; Du, J.K.; Song, M.M.; Xu, Y.P.; Xie, S.P.; Zheng, W.L.; Xu, C.Y. A procedure for quantifying runoff response to spatial and temporal changes of impervious surface in Qinhuai River basin of southeastern China. *Catena* **2017**, *157*, 268–278. [CrossRef]
6. Roman, D.C.; Vogel, R.M.; Schwarz, G.E. Regional regression models of watershed suspended-sediment discharge for the eastern United States. *J. Hydrol.* **2012**, *472–473*, 53–62. [CrossRef]
7. Tuset, J.; Vericat, D.; Batalla, R.J. Rainfall, runoff and sediment transport in a Mediterranean mountainous catchment. *Sci. Total Environ.* **2016**, *540*, 114–132. [CrossRef] [PubMed]
8. Buendia, C.; Herrero, A.; Sabater, S.; Batalla, R.J. An appraisal of the sediment yield in western Mediterranean river basins. *Sci. Total Environ.* **2016**, *572*, 538–553. [CrossRef] [PubMed]
9. Castiglioni, S.; Lombardi, L.; Toth, E.; Castellarin, A.; Montanari, A. Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach. *Adv. Water Resour.* **2010**, *33*, 1235–1242. [CrossRef]
10. Tramblay, Y.; Saint-Hilaire, A.; Ouarda, T.B.M.J.; Moatar, F.; Hecht, B. Estimation of local extreme suspended sediment concentrations in California Rivers. *Sci. Total Environ.* **2010**, *408*, 4221–4229. [CrossRef] [PubMed]
11. Lombardi, L.; Toth, E.; Castellarin, A.; Montanari, A.; Brath, A. Calibration of a rainfall-runoff model at regional scale by optimising river discharge statistics: Performance analysis for the average/low flow regime. *Phys. Chem. Earth* **2012**, *42–44*, 77–84. [CrossRef]
12. Ali, M.; Seeger, M.; Sterk, G.; Moore, D. A unit stream power based sediment transport function for overland flow. *Catena* **2013**, *101*, 197–204. [CrossRef]
13. Heng, S.; Suetsugi, T. Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction. *J. Hydrol.* **2014**, *512*, 240–253. [CrossRef]
14. Zhao, J.; Vanmaercke, M.; Chen, L.; Govers, G. Vegetation cover and topography rather than human disturbance control gully density and sediment production on the Chinese Loess Plateau. *Geomorphology* **2016**, *274*, 92–105. [CrossRef]
15. Paule-Mercado, M.A.; Ventura, J.S.; Memon, S.A.; Jahng, D.; Kang, J.H.; Lee, C.H. Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff. *Sci. Total Environ.* **2016**, *550*, 1171–1181. [CrossRef] [PubMed]
16. Eleria, A.; Vogel, R.M. Predicting fecal coliform bacteria levels in the Charles River, Massachusetts, USA. *J. Am. Water Resour. Assoc.* **2005**, *41*, 1195–1209. [CrossRef]

17. David, M.M.; Haggard, B.E. Development of regression-based models to predict fecal bacteria numbers at select sites within the Illinois River Watershed, Arkansas and Oklahoma, USA. *Water Air Soil Pollut.* **2011**, *215*, 525–547. [CrossRef]

18. Motamarri, S.; Boccelli, D.L. Development of a neural-based forecasting tool to classify recreational water quality using fecal indicator organisms. *Water Res.* **2012**, *46*, 4508–4520. [CrossRef] [PubMed]

19. Herrig, I.M.; Böer, S.I.; Brennholt, N.; Manz, W. Development of multiple linear regression models as predictive tools for fecal indicator concentrations in a stretch of the lower Lahn River, Germany. *Water Res.* **2015**, *85*, 148–157. [CrossRef] [PubMed]

20. Khan, S.; Lau, S.-L.; Kayhanian, M.; Stenstrom, M.K. Oil and grease measurement in highway runoff—Sampling time and event mean concentrations. *J. Environ. Eng.* **2006**, *132*, 415–422. [CrossRef]

21. Kayhanian, M.; Suverkropp, C.; Ruby, A.; Tsay, K. Characterization and prediction of highway runoff constituent event mean concentration. *J. Environ. Manag.* **2007**, *85*, 279–295. [CrossRef] [PubMed]

22. Ha, S.J.; Stenstrom, M.K. Predictive modeling of storm-water runoff quantity and quality for a large urban watershed. *J. Environ. Eng.* **2008**, *134*, 703–711. [CrossRef]

23. Maniquiz, M.C.; Lee, S.; Kim, L.H. Multiple linear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables. *J. Environ. Sci.* **2010**, *22*, 946–952. [CrossRef]

24. Madarang, K.J.; Kang, J.-H. Evaluation of accuracy of linear regression models in predicting urban stormwater discharge characteristics. *J. Environ. Sci.* **2014**, *26*, 1313–1320. [CrossRef]

25. Feng, X.; Cheng, W.; Fu, B.; Lü, Y. The role of climatic and anthropogenic stresses on long-term runoff reduction from the Loess Plateau, China. *Sci. Total Environ.* **2016**, *571*, 688–698. [CrossRef] [PubMed]

26. Hou, X.; Zhou, F.; Leip, A.; Fu, B.; Yang, H.; Chen, Y.; Gao, S.; Shang, Z.; Ma, L. Spatial patterns of nitrogen runoff from Chinese paddy fields. *Agric. Ecosyst. Environ.* **2016**, *231*, 246–254. [CrossRef]

27. Smith, R.E.; Goodrich, D.C.; Quinton, J.N. Dynamic, distributed simulation of watershed erosion—The KINEROS2 and EUROSEM models. *Trans. ASAE* **1995**, *50*, 517–520.

28. De Roo, A.P.J.; Offermans, R.J.E.; Cremers, N.H.D.T. LISEM: A single-event, physically based hydrological and soil erosion model for drainage basins. II: Sensitivity analysis, validation and application. *Hydrol. Process.* **1996**, *10*, 1119–1126. [CrossRef]

29. Morgan, R.P.C.; Quinton, J.N.; Smith, R.E.; Govers, G.; Poesen, J.W.A.; Auerswald, K.; Chisci, G.; Torri, D.; Styczen, M.E.; Folly, A.J. The European soil erosion model (EUROSEM): Documentation and user guide. *Earth Surf. Process. Landf.* **1998**, *23*, 527–544. [CrossRef]

30. Wu, B.; Wang, Z.; Shen, N.; Wang, S. Modelling sediment transport capacity of rill flow for loess sediments on steep slopes. *Catena* **2016**, *147*, 453–462. [CrossRef]

31. Wonju Regional Environmental Office. *Monitoring and Assessment for the Nonpoint Source Pollution Management Area of Mandae, Gaah and Jaun Region*; Ministry of Environment: Wonju, Korea, 2016.

32. Cho, K.H.; Kang, J.H.; Ki, S.J.; Park, Y.; Cha, S.M.; Kim, J.H. Determination of the optimal parameters in regression models for the prediction of chlorophyll-a: A case study of the Yeongsan Reservoir, Korea. *Sci. Total Environ.* **2009**, *407*, 2536–2545. [CrossRef] [PubMed]

33. Gonzalez, R.A.; Noble, R.T. Comparisons of statistical models to predict fecal indicator bacteria concentrations enumerated by qPCR- and culture-based methods. *Water Res.* **2014**, *48*, 296–305. [CrossRef] [PubMed]

34. 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]

35. Chong, A.; Lam, K.P.; Pozzi, M.; Yang, J. Bayesian calibration of building energy models with large datasets. *Energy Build.* **2017**, *154*, 343–355. [CrossRef]

36. Hwang, S.H.; Ham, D.H.; Kim, J.H. A new measure for assessing the efficiency of hydrological data-driven forecasting models. *Hydrol. Sci. J.* **2012**, *57*, 1257–1274. [CrossRef]