land

MDPI

Article

Development of Flood Damage Regression Models by Rainfall Identification Reflecting Landscape Features in Gangwon Province, the Republic of Korea

Hyun Il Choi

Department of Civil Engineering, Yeungnam University, 280 Daehak-Ro, Gyeongsan 38541, Gyeongbuk, Korea; hichoi@ynu.ac.kr; Tel.: +82-53-810-2413

Abstract: Torrential rainfall events associated with rainstorms and typhoons are the main causes of flood-related economic losses in Gangwon Province, Republic of Korea. The frequency and severity of flood damage have been increasing due to frequent extreme rainfall events as a result of climate change. Rainfall is a major cause of flood damage for the study site, given a strong relationship between the probability of flood damage over the last two decades and the maximum rainfall for 6 and 24 h durations in the 18 administrative districts of Gangwon Province. This study aims to develop flood damage regression models by rainfall identification for use in a simplified and efficient assessment of flood damage risk in ungauged or poorly gauged regions. Optimal simple regression models were selected from four types of non-linear functions with one of five composite predictors averaged for the two rainfall datasets. To identify appropriate predictor rainfall variables indicative of regional landscape features, the relationships between the composite rainfall predictor and landscape characteristics such as district size, topographic features, and urbanization rate were interpreted. The proposed optimal regression models may provide governments and policymakers with an efficient flood damage risk map simply using a regression outcome to design or forecast rainfall data.

Keywords: flood damage; rainfall; landscape; simple regression; damage risk map

Citation: Choi, H.I. Development of Flood Damage Regression Models by Rainfall Identification Reflecting Landscape Features in Gangwon Province, the Republic of Korea. *Land* **2021**, *10*, 123. https://doi.org/10.3390/land10020123

Received: 20 December 2020
Accepted: 26 January 2021
Published: 27 January 2021

1. Introduction

Global warming and climate change have increased the frequency and severity of extreme weather events, which has in turn elevated the risk of severe climate-related natural disasters [1–3]. Natural disasters may directly incur substantial human and economic damage costs, and flood-related disasters are one of the most frequent and deadliest natural disasters worldwide [4]. The Korean Peninsula annually experiences flood damage by the East Asian monsoon, and the flood damage costs caused by rainstorms and typhoons account for the majority of damage losses caused by natural disasters in the Republic of Korea [5]. Climate changes may also have a greater influence on extreme rainfall patterns in Gangwon Province than in other regions of the Korean Peninsula. This is related to the complex geographical landscape of the province associated with the Taebaek Mountain Range and the East Sea. These features divide the province into the western region with a mountainous climate and the eastern region with an oceanic climate. In terms of the historic extreme events, Gangneung City in the eastern province received the highest recorded daily rainfall of 880 mm. This was considered a 200-year event, due to a localized downpour from severe thunderstorms by Typhoon Rusa on 31 August 2002 [5,6]. On 29 August 2018, Cheorwon County in the western province recorded the heaviest downpour, measuring 113.5 mm/h with a return period exceeding 500 years, due to a localized stagnant front created between the cold air mass from the northwest and the hot and humid air mass from the East Sea [5]. A number of severe localized downpours associated with torrential rainstorms and super typhoons frequently occur because of the mountainous and coastal landscapes characteristic to Gangwon Province. The major countermeasures against the

flood damage have focused on supporting recovery costs for flood damaged areas in the Republic of Korea [5]. As such, preemptive flood management measures are required to reduce the human and economic damage costs from recent flood disasters. Assessment of the vulnerability or risk to regional flood hazard is one of the non-structural measures to prepare integrated mitigation measures customized to regional flood damage [7,8]. For proactive approaches to flood risk management strategies, there is a need for a method that can predict future flood damage risk by analyzing the characteristics and trends of regional flood damage records [9].

Flood damage risk or vulnerability assessments should be based on flood hazard and inundation analysis results using hydrologic and hydraulic models. However, the lack of available hydrological data and information of a decent quality introduces a degree of uncertainty in validating model simulation results, particularly the case for ungauged regions. The lack of reliable data is a crucial barrier to flood damage analysis and flood risk assessment [10]. To resolve these issues, regression analysis presents itself as an alternative method that may be an effective tool in predicting hydrological variables through acceptable relationships with influencers to overcome limited hydrological data in spatial and temporal resolutions on target regions to be analyzed [11]. Many studies have shown that rainfall characteristics have a significant impact on flood damage events from complex influencing factors [12–19]. Elucidation of a functional relationship between rainfall and flood damage could relate the amount of flood damage or flood events to rainfall conditions. As such, the risk of flood damage may also be estimated by determining the rainfall–flood damage relationship through regression analysis [12,15,17,19]. Most previous studies have conducted regression analysis using the fixed predictor rainfall variables in a single regression function to develop regional damage regression models. However, the variations in flood damage attributable to rainfall were not high in some rainfall-flood damage regression models. To improve the prediction performance of rainfall-flood damage regression analysis, it is necessary to identify rainfall variables that reflect regional characteristics; these typically have a non-linear relationship with the features of flood damage.

The aim of this study is to provide a methodology to develop rainfall-flood damage regression models for assessing the relative flood damage risk associated with a specific amount of rainfall for designing or forecasting purposes. This paper proposes optimal regression models to estimate regional flood damage. These models were selected from four types of regression functions, with one of the five predictor rainfall variables capable of representing the regional landscape and terrain features. The proposed methodology was implemented through various regression analysis models for Gangwon Province, Republic of Korea. This area characterized by a complex landscape of mountainous and coastal areas and lacks in available and/or reliable hydrological data. Flood damage data caused by rainstorms and typhoons were collected from annual disaster reports [5], provided by the Ministry of the Interior and Safety for the last 20 years from 1999 to 2018. The analysis period over the last two decades was determined by comprehensively considering the amount of data necessary for regression analysis and the consistency in damage features of past data for the study area. Rainfall data were collected from 16 automated surface observing system (ASOS) meteorological stations [20], managed by the Korea Meteorological Agency around the 18 administrative districts of Gangwon Province. The ASOS meteorological gauge stations undertake continuous measurements of hourly rainfall observations for the analysis period of flood damage records. Although there are no generalized guidelines for sample size requirements appropriate for regression analysis, this study has adopted one of the various rules-of-thumb that recommends at least 10 cases per variable [21–23]. Therefore, several non-linear functions were applied to a simple regression analysis with a single composite predictor averaged by different rainfall characteristics. This accounted for the minimum number of 12 damage records for the study site. The identification of a suitable predictor rainfall variable that incorporates regional landscape features may improve the possible applications of rainfall–flood damage regression results.

2. Materials and Methods

2.1. Study Region

Figure 1 shows that the Gangwon Province is located between 37°02′ N–38°37′ N and 127°05′ E–129°22′ E in the mid-eastern part of the Korean Peninsula. It is located at the eastern end of the Asian continent bordered by the East Sea, a margin of the Western Pacific Ocean. The Gangwon Province comprises 18 administrative districts (7 cities and 11 counties), spanning an area of 16,874 km^2; this makes up 16.8% of the national territory of the Republic of Korea. Figure 2a illustrates that the landscape is dominated by the Taebaek Mountains, which divides the province into two parts—the eastern region has a relatively steep coastline facing the East Sea, whereas the western region is most pronounced in complex mountain terrains containing the headwaters of major rivers in the Republic of Korea, including the Han and Nakdong Rivers. Gangwon Province is generally a mountainous region with a lowland area of less than 100 EL.m, occupying only 5.6% of the total area of the province. The urbanization rate in the province is much lower than the national average due to this rugged mountainous landscape. Figure 2b shows that the urban areas and towns are predominantly located along the coastline in the eastern region. These include (2) Sokcho City, (4) Gangneung City, (5) Donghae City, (6) Samcheok City, and (7) Taebaek City, which are scattered in the Taebaek Mountains. Only two cities are located in the lowland areas of the western region, namely (12) Chuncheon City, the capital, and (16) Wonju City. The climate conditions differ considerably between the eastern and western regions, as the regions are geographically divided by the Taebaek Mountains. The oceanic climate features predominate in the eastern region with steep mountain slopes descending to the coastline. In contrast, the western region predominantly exhibits continental climate features and some highlands around the Taebaek Mountains experience mountain climate characteristics. The climate is characterized by high temperature and humidity due to the temperate climate in summer and low temperature and humidity in winter due to the high continental pressure. The annual average precipitation of Gangwon Province is 1358.9 mm for the last two decades (1999–2018). Approximately 65% of annual precipitation is concentrated in the summer season from June to September [20]. This is mainly due to the East Asian summer monsoon rainfall and the number of typhoon events affecting the Korean Peninsula almost every year. This monsoonal rainfall has caused severe annual flood damage events in nearly all 18 administrative districts of the province (see also Table 1 and Figure 3 for details).

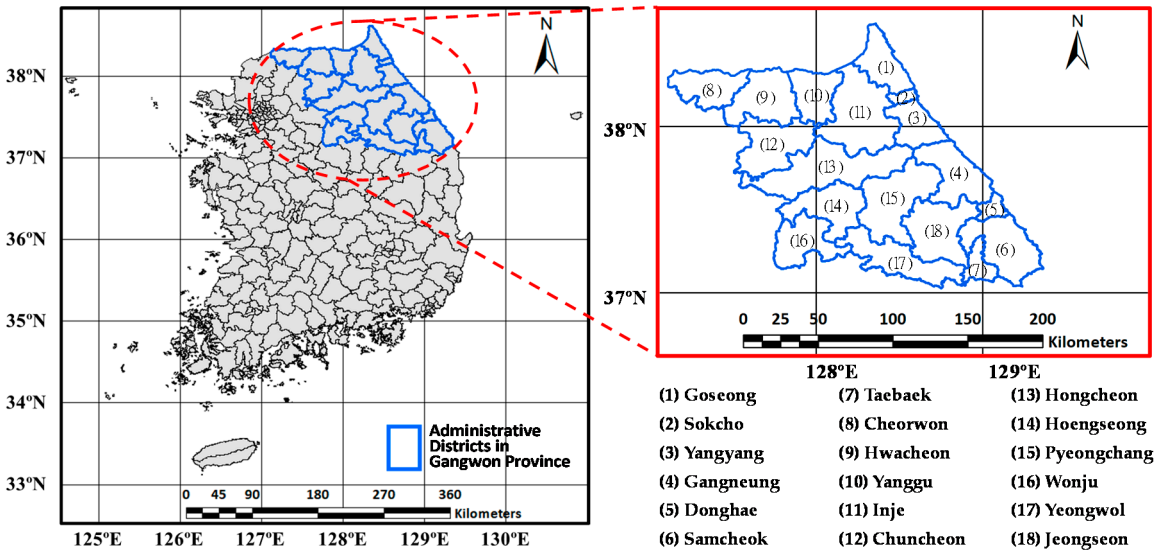

Figure 1. Study area (Gangwon Province, Korea) and the location of the 18 administrative districts in the province.

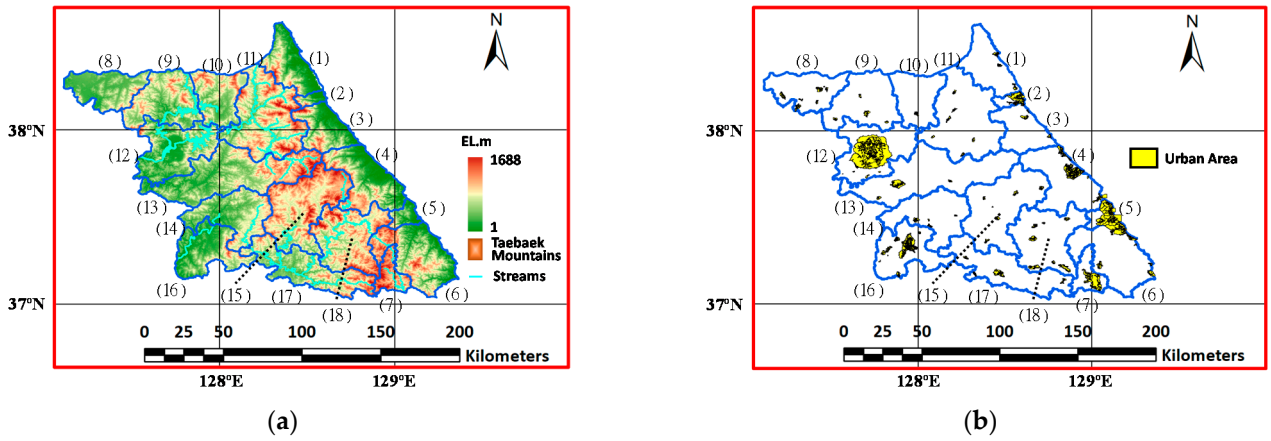

Figure 2. Landscape of the 18 administrative districts of Gangwon Province: (**a**) terrain landscape; (**b**) urban area. Number in parentheses indicates the identification number of each administrative district.

Table 1. Statistics on the economic costs of damage events caused by rainstorms and typhoons in the 18 administrative districts of Gangwon Province from 1999 to 2018.

Causes	Gross Total [1] (US$)	Number of Cases [2]	Annual Average [3] (US$/Year)	Damage Intensity [3] (US$/Event)
Rainstorms	2,088,396,922	280	5,801,103	414,364
Typhoons	3,406,875,471	154	9463.543	1,229,032

[1] Total economic losses to assets including buildings, ships, public facilities, crops, etc., based on the consumer price index and US $ rate in 2018 over the 18 administrative districts. [2] Sum of the total damage events over the 18 administrative districts. [3] Average damage costs per administrative districts.

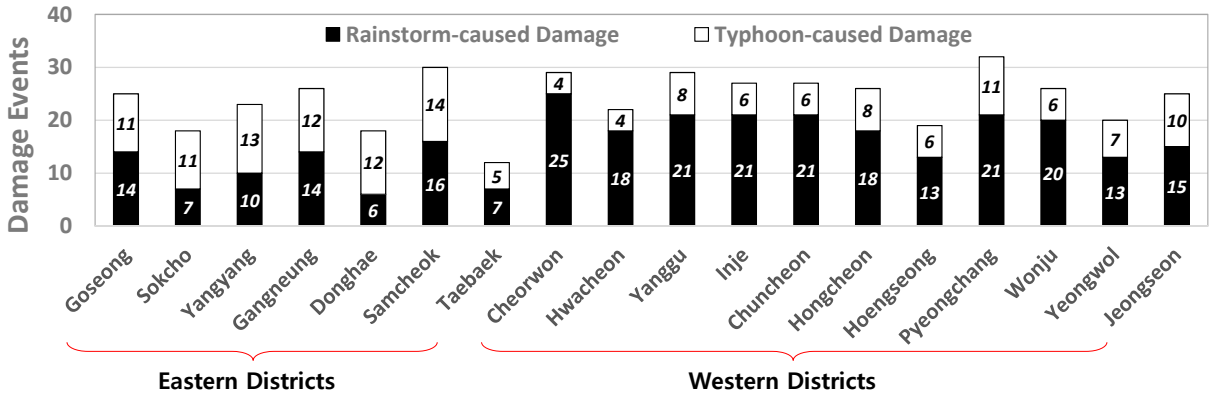

Figure 3. Number of economic damage records of rainstorms and typhoons for each administrative district in Gangwon Province from 1999 to 2018.

2.2. Data

The National Disaster Information Center [5] provides the annual flood damage records of rainstorms and typhoons, along with information on the date and place of each damage event. Flood damage records were collected over the last two decades from 1999 to 2018 for 7 cities and 11 counties in Gangwon Province. This vast amount of data was obtained to secure an adequate amount of historical data necessary for consistency in damage features and also to fulfill the requirements in the rainfall–flood damage regression analysis. Table 1 summarizes the data statistics of economic damage records caused by rainstorms and typhoons from 1999 to 2018 for the 18 administrative districts. There was an annual average of 1.2 flood damage events (434 flood damage events over 20 years per 18 administrative districts) for each administrative district over the last two decades.

In terms of the total occurrence number of economic flood damage events, rainstorm-induced damage events were approximately twice the typhoon-induced events. However, typhoons have incurred much larger cumulative and district average damage costs than rainstorms. The difference is even greater in terms of damage intensity (economic losses per damage event); typhoon-induced damage intensity was approximately three times rainstorm-induced damage intensity for each administrative district. Figure 3 also indicates that these economic damage events caused by typhoons were more frequent than or comparable to those caused by rainstorms in the eastern districts. In contrast, the western districts experienced more frequent economic damage events caused by rainstorms than by typhoons over the last two decades. These distinct regional flood damage patterns are mainly due to the regional complex landscape associated with the Taebaek Mountains and the East Sea. Note that all typhoon-caused damage data were included in the analysis as there are rainfall records for each typhoon-induced damage event in Gangwon Province over the last two decades.

For rainfall–flood damage regression analysis, rainfall observations were also collected from 16 ASOS meteorological stations under the Korea Meteorological Agency [20]. These rainfall data spatially affect the 18 administrative districts; the 16 AOGS stations have been able to continuously secure hourly rainfall data without missing data for the last two decades. Rainfall data considered for this study included those for the analysis period from one day prior to the start date to the end date of each flood damage event for the 18 administrative districts to accommodate for the influential rainfall characteristics that generate each flood damage event. Figure 4 shows that the areal average rainfall was computed based on these rainfall data from the 16 ASOS stations using the Thiessen polygon method [24], which is a spatial interpolation technique commonly used in engineering hydrology. Thiessen polygons are generated from the bisector lines of two neighboring stations, and each polygon that contains a station represents the rainfall for that station. The areal average rainfall was interpolated using the ratios of the Thiessen polygons within a district. Note that ASOS stations are sparsely distributed in the northern districts bordered by the military demarcation line. The rainfall characteristics for regression analysis were tentatively selected as the maximum rainfall recorded during damage events of 1, 2, 3, 6, 12, and 24 h (R_1, R_2, R_3, R_6, R_{12}, and R_{24}, respectively). These durations represent standard durations typically used for the purposes of designing, planning, forecasting, or warning.

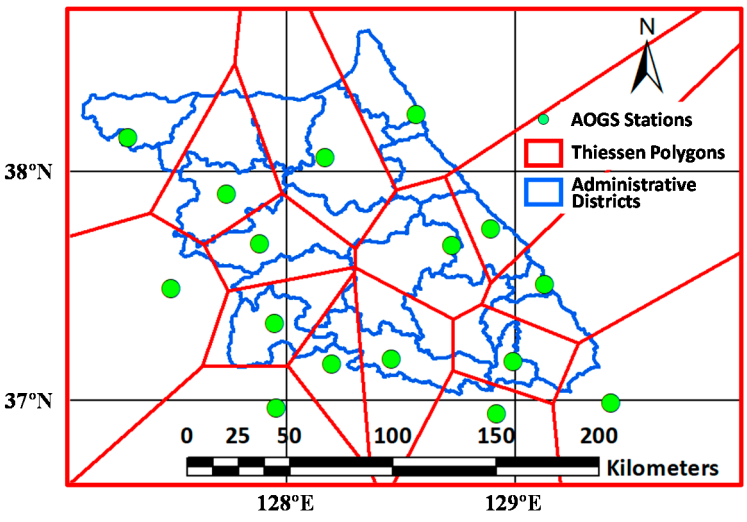

Figure 4. Thiessen polygons overlain with the 16 automated surface observing system (ASOS) meteorological stations spatially reflecting the 18 administrative districts of Gangwon Province.

2.3. Relation Functions

The proposed damage regression models were intended to estimate the relative flood damage risk for a specific amount of rainfall for making design decisions or forecasting, as opposed to predicting the precise cost of flood damage. Previous studies have demonstrated that a significant relationship exists between rainfall characteristics and flood damage features; hence, rainfall data can be used for flood damage risk assessments by utilizing regression functions to estimate the probability of occurrence of flood damage events with respect to a specific amount of rainfall recorded [12,13,15,18]. As the relationship between rainfall and flood damage is also dependent on regional characteristics such as landscape and climate, various regression functions need to be considered while selecting the optimum goodness-of-fit among them, pertaining to each administrative district. Hence, this study used four types of regression functions based on rational functions in Equations (1) and (2) and logistic functions in Equations (3) and (4), as shown below.

$$\frac{D}{1-D} = \alpha + \beta I_w \tag{1}$$

$$\frac{D}{1-D} = \alpha + \beta ln I_w \tag{2}$$

$$ln\left(\frac{D}{1-D}\right) = \alpha + \beta I_w \tag{3}$$

$$ln\left(\frac{D}{1-D}\right) = \alpha + \beta ln I_w \tag{4}$$

where the dependent (response) variable D is the probability of a flood damage event; the independent (predictor) variable I_w is the rainfall amount averaged by a weighting factor; w (see Equation (5) for details); and α and β are regression coefficients. Flood damage records were converted into flood damage density (economic damage costs per district area), which represents the areal density of property damaged by a flood event. A log-normal distribution function was used to compute the occurrence probability of flood damage density; this is generally considered suitable for flood damage data [25].

To identify a single predictor variable, I_w, in the regression functions in Equations (1)–(4), rainfall characteristics that were highly correlated with flood damage features over the 18 administrative districts were selected. Generally, the 6 h maximum rainfall R_6 and the 24 h maximum rainfall R_{24} had higher Pearson correlation coefficients than the other rainfall characteristics over the 18 administrative districts. This outcome was based on the correlation between each of the six rainfall values (R_1, R_2, R_3, R_6, R_{12}, and R_{24}) and the probability of flood damage records for each administrative district (see Figure 5).

Accordingly, the two rainfall features R_6 for a short duration and R_{24} for a long duration were selected. To incorporate the effect of the two characteristics of R_6 and R_{24} into a single predictor variable, a composite rainfall variable I_w was proposed as a weighted geometric mean of the two rainfall factors:

$$I_w = R_6^w \cdot R_{24}^{(1-w)} \tag{5}$$

where a weighting factor w representing the relative effect ratio of R_6 to R_{24} was assumed for the five cases; these were 0, 0.25, 0.5, 0.75, and 1. Note that I_o or I_1 indicates that only R_{24} or R_6 represent the regional rainfall characteristics, respectively.

Therefore, each administrative district may have a total of 20 regression models using the four types of regression functions in Equations (1)–(4), with each of the five predictor variables such as I_0 (R_{24}), $I_{0.25}$, $I_{0.5}$, $I_{0.75}$, and I_1 (R_6) in Equation (5). Then, the optimal regression model was selected for each administrative district to compare the significance and variation explained by the 20 regression functions. For robust regression analysis, any outliers in the original dataset were excluded once they were detected by the three

diagnostic methods: Cook's distance [26], Studentized residual, and difference in fits (DFFITS) [27].

	R_1	R_2	R_3	R_6	R_{12}	R_{24}
(1) Goseong	0.455	0.492	0.535	0.640	0.629	0.641
(2) Sokcho	0.416	0.454	0.515	0.654	0.643	0.655
(3) Yangyang	0.622	0.637	0.633	0.746	0.725	0.778
(4) Gangneung	0.746	0.728	0.767	0.884	0.821	0.841
(5) Donghae	0.541	0.538	0.527	0.560	0.519	0.527
(6) Samcheok	0.477	0.506	0.546	0.663	0.589	0.506
(7) Taebaek	0.358	0.529	0.664	0.789	0.559	0.448
(8) Cheorwon	0.255	0.237	0.240	0.300	0.300	0.361
(9) Hwacheon	0.488	0.515	0.480	0.558	0.486	0.516
(10) Yanggu	0.395	0.437	0.407	0.492	0.354	0.461
(11) Inje	0.285	0.260	0.273	0.295	0.192	0.324
(12) Chuncheon	0.560	0.559	0.538	0.545	0.538	0.526
(13) Hongcheon	0.516	0.510	0.483	0.492	0.494	0.547
(14) Hoengseong	0.356	0.341	0.278	0.333	0.337	0.417
(15) Pyeongchang	0.204	0.189	0.231	0.233	0.165	0.218
(16) Wonju	0.509	0.438	0.448	0.574	0.531	0.528
(17) Yeongwol	0.178	0.139	0.229	0.457	0.654	0.784
(18) Jeongseon	0.455	0.492	0.535	0.640	0.629	0.641

Figure 5. Pearson correlation coefficients between flood damage probability and one of the six rainfall values (R_1, R_2, R_3, R_6, R_{12}, and R_{24}) in the 18 administrative districts of Gangwon Province. The darkest red panel indicates the highest correlation coefficient for each administrative district.

3. Results

3.1. Optimal Regression Models

The 20 regression analysis results for each administrative district were compared and evaluated using the R-squared value (the squared correlation coefficient) for prediction stability and the *p*-value of the F-test to indicate significance. Hence, an optimal regression model was proposed based on the highest R-squared value in the regression results, with a *p*-value less than 0.01, for each administrative district. Table 2 summarizes the information on the optimal regression models for the 18 administrative districts of Gangwon Province. As for the type of optimal regression functions, the rational function in Equation (1) was selected for 10 administrative districts, whereas the logistic functions in Equations (3) and (4) were adopted for the remaining 8 districts. There was no clear relationship between the function types and regional characteristics. All *p*-values confirmed the significance of the 18 optimal regression models at a significance level of 1%. Overall, the R-squared values showed a substantial degree of goodness-of-fit based on using the predictor variable I_w as a composite rainfall of R_6 and R_{24} for each administrative district.

To examine the predictor identification results reflecting landscape features, Figure 6a,b compare the weights of R_6 for predictor variables in the optimal regression models with specific landscape features such as average slope, district size, and division type. Table 3 also shows the predictor rainfall variables alongside the landscape features such as area size, average elevation, average slope, and urbanization rate of each of the administrative districts. For the seven administrative districts in the eastern region (locations (1) to (7) in Figure 6), the predictor rainfall variable I_w was identified as $I_{0.5}$, $I_{0.75}$, or I_1. Hence, the relatively short-duration rainfall R_6 can be better at explaining the flood damage variation than the relatively long-duration rainfall R_{24}. This is primarily due to the landscape that characterizes the eastern districts; these districts sit at the interface between the Taebaek

Mountains and the East Sea. In this area, severe torrential rainfall events occur as a result of the high instability of humid air from the East Sea interacting with the Taebaek Mountains. Additionally, this may also be explained by flood damage events that are caused more frequently by typhoons in the eastern region than in the western region where a greater proportion of flood damage events are caused by rainstorms. The rainfall value R_6 also greatly contributes to the composite predictor I_w particularly for five eastern cities that are relatively more urbanized, such as $I_{0.75}$ for (2) Sokcho City and I_1 (R_6 only) for (4) Gangneung City, (5) Donghae City, (6) Samcheok City, and (7) Taebaek City. The effectiveness of R_6 as a predictor rainfall variable was also evident for some western districts; I_1 (R_6 only) was selected for two cities (12) Chuncheon City and (16) Wonju City with a relatively higher urbanization rate as well as for (9) Hwacheon County and (10) Yangyang County with hillslopes in a medium-sized area. For (8) Cheorwon County, with a relatively mildly-sloped inland, and (11) Inje County and (18) Jeongseon County, which have a relatively larger hillslope area adjacent to the eastern region, $I_{0.5}$ and $I_{0.75}$ were selected. In contrast, I_0 (R_{24} only) was selected in four western districts only—(13) Hongcheon County, (14) Hoengseong County, (15) Pyeongchang County, and (17) Yeongwol County in a relatively less urbanized and large-sized area. Overall, R_6 largely affects the flood damage characteristics over most administrative districts that have a relatively small mountainous landscape or a highly urbanized downtown area, as shown in Figure 6 and Table 3.

Table 2. Information on the optimal rainfall–flood damage regression models for the 18 administrative districts of Gangwon Province.

No.	District	Predictor	α	β	R-Squared	p-Value	Function
(1)	Goseong	$I_{0.5}$	−14.696	0.195	0.476	0.000	Equation 1
(2)	Sokcho	$I_{0.75}$	−26.009	0.456	0.474	0.001	Equation 1
(3)	Yangyang	$I_{0.5}$	−20.232	0.291	0.626	0.000	Equation 1
(4)	Gangneung	I_1 (R_6)	−1.084	0.020	0.846	0.000	Equation 3
(5)	Donghae	I_1 (R_6)	−1.309	0.027	0.368	0.005	Equation 4
(6)	Samcheok	I_1 (R_6)	−17.898	0.393	0.462	0.000	Equation 1
(7)	Taebaek	I_1 (R_6)	−17.354	0.288	0.794	0.001	Equation 1
(8)	Cheorwon	$I_{0.5}$	0.118	0.009	0.344	0.003	Equation 1
(9)	Hwacheon	I_1 (R_6)	−1.343	0.019	0.403	0.003	Equation 3
(10)	Yanggu	I_1 (R_6)	−4.490	1.092	0.417	0.001	Equation 4
(11)	Inje	$I_{0.5}$	−4.593	1.062	0.426	0.001	Equation 4
(12)	Chuncheon	I_1 (R_6)	−0.750	0.038	0.526	0.000	Equation 1
(13)	Hongcheon	I_0 (R_{24})	−8.673	1.861	0.728	0.000	Equation 4
(14)	Hoengseong	I_0 (R_{24})	−1.515	0.029	0.482	0.001	Equation 1
(15)	Pyeongchang	I_0 (R_{24})	−6.575	1.352	0.324	0.004	Equation 4
(16)	Wonju	I_1 (R_6)	−2.644	0.025	0.459	0.000	Equation 3
(17)	Yeongwol	I_0 (R_{24})	−2.316	0.043	0.551	0.000	Equation 1
(18)	Jeongseon	$I_{0.75}$	−17.587	0.270	0.830	0.000	Equation 1

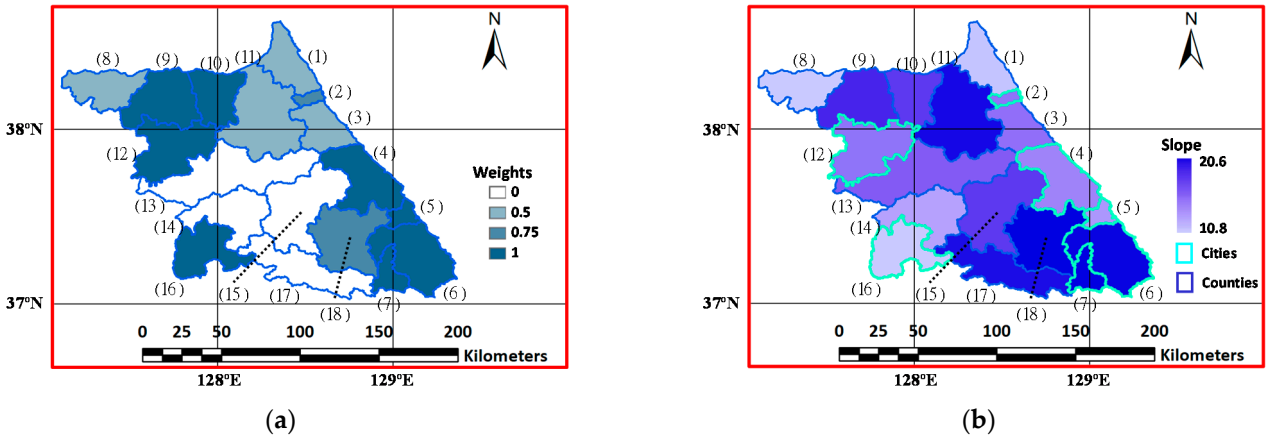

Figure 6. Comparison of (**a**) the weights of R_6 for the predictor variable in the optimal regression model; and (**b**) landscape features such as average slope, district size, and division type over the 18 administrative districts of Gangwon Province. Number in parentheses indicates the identification number of each administrative district.

Table 3. Landscape features affecting the predictor rainfall variables of optimal regression models over the 18 administrative districts of Gangwon Province.

No.	District	Predictor	Area Size (km^2)	Average Elevation (EL. m)	Average Slope (%)	Urbanized Rate (%)	Division Type
(1)	Goseong	$I_{0.5}$	660.9	307.4	13.6	1.9	County
(2)	Sokcho	$I_{0.75}$	105.6	353.0	15.9	39.8	City
(3)	Yangyang	$I_{0.5}$	629.9	380.1	16.0	1.7	County
(4)	Gangneung	I_1 (R_6)	1040.4	392.8	15.3	8.1	City
(5)	Donghae	I_1 (R_6)	180.2	270.8	14.8	68.8	City
(6)	Samcheok	I_1 (R_6)	1186.6	486.3	19.2	6.7	City
(7)	Taebaek	I_1 (R_6)	303.4	901.7	19.3	22.0	City
(8)	Cheorwon	$I_{0.5}$	889.4	334.5	10.8	3.0	County
(9)	Hwacheon	I_1 (R_6)	909.0	467.8	18.2	1.5	County
(10)	Yanggu	I_1 (R_6)	661.9	505.9	17.6	0.6	County
(11)	Inje	$I_{0.5}$	1645.2	643.0	19.1	0.9	County
(12)	Chuncheon	I_1 (R_6)	1116.4	304.7	15.9	30.0	City
(13)	Hongcheon	I_0 (R_{24})	1819.7	497.8	16.5	1.7	County
(14)	Hoengseong	I_0 (R_{24})	997.8	437.9	14.5	1.0	County
(15)	Pyeongchang	I_0 (R_{24})	1463.7	755.9	17.6	1.1	County
(16)	Wonju	I_1 (R_6)	872.4	314.5	13.4	10.1	City
(17)	Yeongwol	I_0 (R_{24})	1127.6	533.7	19.0	2.3	County
(18)	Jeongseon	$I_{0.75}$	1219.7	709.6	20.6	3.5	County

3.2. Damage Risk Map

The proposed regression models were applied to flood damage risk assessment for the 100-year design rainfall in the 18 administrative districts. First, 100-year frequency rainfall data were collected for the same 16 ASOS meteorological stations previously selected in the regression analysis from the rainfall frequency analysis results by the Water Management Information System [28]. The areal average rainfall amount in each administrative district was also computed based on the 100-year frequency amounts for R_6 and R_{24} from the 16 ASOS stations using the Thiessen polygon method [24]. Then, the flood damage risk was estimated based on the probability of damage from the optimal regression models for the predictor rainfall variable I_w relating to a 100-year frequency event in the 18 administrative districts.

Table 4 and Figure 7 show that the flood damage risk was forecasted to be higher for all seven eastern districts, whereas it was predicted that most western districts would experience low flood damage risk. The exception to this is (18) Jeongseon County, which is adjacent to the eastern region. This flood damage risk assessment is based on flood damage density (economic damage costs per district area); the highest damage risk was estimated for (5) Donghae City, a small-sized city with the highest urbanization rate in the province. The lowest damage risk was estimated for (8) Cheorwon County, a relatively less urbanized and low-lying county that is inland and farthest from the sea. Table 4 shows that there were similar differences between R_6 and R_{24} for the 100-year frequency across the 18 administrative districts, suggesting that flood damage risk outcomes may be only slightly influenced by the different predictor rainfall variables in the optimal regression model. In terms of the flood damage risk versus the amount of rainfall (R_6 or R_{24}) for the 100-year frequency in each district, the risk rank of flood damage was not directly related to the rank of the 100-year rainfall event. There were notable inverse links between flood damage risk and the amount of rainfall amount; (6) Samcheok City and (9) Taebaek City had higher risk of flood damage from lower amounts of rainfall, indicating that these cities are more vulnerable to flood damage than others. In contrast, (8) Cheorwon County and (15) Pyeongchang County had lower risk of flood damage from higher amounts of rainfall, which indicates the lower vulnerability of these counties to flood damage.

Table 4. Flood damage risk assessment from the optimal regression models for the 100-year design rainfall over the 18 administrative districts of Gangwon Province.

No.	District	R_6	R_{24}	Predictor		Damage Estimate [1]	Risk Rank [2]
				Variable	Amount		
(5)	Donghae	208.80	398.20	I_1 (R_6)	255.00	0.996	1
(2)	Sokcho	234.00	414.80	$I_{0.75}$	326.58	0.992	2
(3)	Yangyang	230.27	424.98	$I_{0.5}$	373.60	0.990	3
(4)	Gangneung	229.91	447.36	I_1 (R_6)	280.83	0.989	4
(6)	Samcheok	178.43	351.11	I_1 (R_6)	206.72	0.984	5
(1)	Goseong	233.68	414.31	$I_{0.5}$	372.60	0.983	6
(18)	Jeongseon	189.04	382.70	$I_{0.75}$	262.49	0.982	7
(7)	Taebaek	166.90	336.60	I_1 (R_6)	186.49	0.973	8
(16)	Wonju	195.76	341.62	I_1 (R_6)	215.39	0.944	9
(9)	Hwacheon	189.26	347.84	I_1 (R_6)	206.58	0.931	10
(17)	Yeongwol	161.18	312.12	I_0 (R_{24})	348.60	0.927	11
(13)	Hongcheon	184.43	355.77	I_0 (R_{24})	399.43	0.922	12
(14)	Hoengseong	186.41	336.47	I_0 (R_{24})	374.07	0.904	13
(12)	Chuncheon	173.26	332.49	I_1 (R_6)	189.17	0.864	14
(15)	Pyeongchang	207.50	425.84	I_0 (R_{24})	493.97	0.859	15
(11)	Inje	178.09	328.85	$I_{0.5}$	269.77	0.794	16
(10)	Yanggu	169.82	316.11	I_1 (R_6)	185.42	0.771	17
(8)	Cheorwon	211.70	371.30	$I_{0.5}$	306.73	0.734	18

[1] The cumulative probability of flood damage estimated for the design rainfall amount. [2] The rank order of the 18 administrative districts by flood damage estimate results.

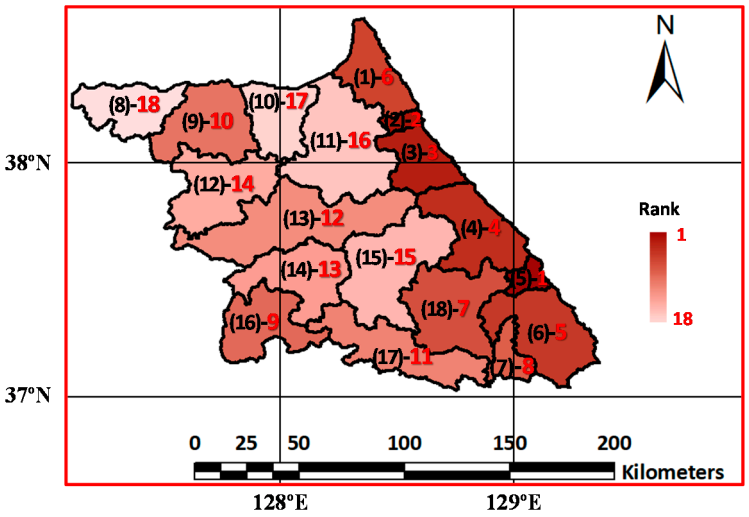

Figure 7. Flood damage risk map indicating risk rankings for the 100-year design rainfall over the 18 administrative districts of Gangwon Province. Number in parentheses indicates identification number of each administrative district.

4. Discussion

Frequent and severe flood damage events caused by rainstorms and typhoons occur in Gangwon Province. Most administrative districts in the province have experienced flood damage events on an annual basis (an annual average of 1.2 events per district) for the last two decades from 1999 to 2018, as shown in Table 1 and Figure 3. As most of the economic damage losses from natural disasters are caused by rainstorms and typhoons in Gangwon Province [5], a capable and efficient assessment methodology is required to estimate flood damage risk for each of the administrative districts in the province when runoff observations are not available or reliable. Regression analysis is a common statistical technique applied for hydrological analysis of regions with limited data availability [11]. The aim of this study was to evaluate the regional trends of potential flood damage

for ungauged or poorly gauged regions; this has previously been attempted in other studies [13,15,18]. The characteristics of rainfall were assumed to be major factors that had significant influence on the probability of flood damage events in Gangwon Province; hence, various rainfall–flood damage regression analyses were conducted for the 18 administrative districts in the province.

Economic flood damage records were converted into flood damage density (economic damage costs per district area) to reflect the flood damage risk to assets concentrated in different sized-areas for each administrative district. The response variable was given as the probability of occurrence for flood damage density by a log-normal distribution function. This was because the relative amount of flood damage was sufficient to measure risk with the response variable that also requires normalization for robust regression analysis. In terms of the predictor rainfall variable, the areal average rainfall data were first constructed one day prior to the start date of the event till the end date of the event to consider potential rainfall characteristics that impact flood damage records for each administrative district. Note that the point-wise rainfall data were generally transformed into areal average rainfall as rainfall measurements from gauge stations are insufficient to provide the necessary spatial information, particularly for study sites with a sparse rain gauge network. From the correlation analysis between rainfall characteristics and flood damage probability, most administrative districts exhibited higher correlation coefficients for R_6 and R_{24}, the maximum rainfall amount at 6 h and 24 h, respectively, throughout the flood damage records. Hence, R_6 and R_{24} were selected as the representative short-duration and long-duration rainfall characteristics, respectively. A simple regression analysis was adopted for the rainfall–flood damage regression models due to the minimum 12 flood damage records available in (7) Taebaek City. This was based on the rule of thumb that at least 10 sample cases per variable have been recommended [21–23]. The use of outdated data for larger sample sizes may lead to inconsistencies in flood damage characteristics due to changes in watershed hydrologic responses. Accordingly, the composite rainfall variable I_w was assigned as the single predictor variable in the simple regression analysis. This variable was a weighted geometric mean of R_6 and R_{24} with the five weighting factors of 0, 0.25, 0.5, 0.75, and 1 to integrate the various relative effects of the characteristics of R_6 and R_{24}. The optimal regression model was selected for each administrative district among the total 20 regression analysis outcomes from the four types of regression functions, with one of the five composite rainfall variables. This strategy was expected to improve the predictability of rainfall–flood damage regression models used in flood damage risk assessment based on the estimation of flood damage probability for a specific rainfall event.

Table 2 shows that all optimal regression models indicate a significant correlation (p-value < 0.01) and a sufficient degree of goodness-of-fit (R-squared of 32.4–84.6%). This means that these models are adequate to meet the purpose of the proposed regression models, i.e., to estimate the risk of relative damage rather than the exact damage-related costs for a flood event. This implies that the regional composite rainfall of R_6 and R_{24} is able to satisfactorily explain the variation in relative flood damage over most of the 18 administrative districts of Gangwon Province. There were some administrative districts that had a relatively low R-squared value, such as (8) Cheorwon County, (9) Hwacheon County, (10) Yanggu County, (14) Hoengseong County, and (15) Pyeongchang County, where rainfall data were not observed at local gauge stations due to a sparsely distributed ASOS network. This study utilized the Thiessen polygon method commonly used in engineering hydrology for areal rainfall estimation. This method may be easily applied for flood damage risk assessments by governments and policymakers. However, this traditional interpolation method may be limited in providing reliable estimates of areal rainfall quantity regions with complex terrain and a sparse network of rainfall gauge stations. For better regression results in these counties, spatial rainfall variability needs further consideration in regression analysis through the use of high-density rainfall data from weather radar measurements; this can only occur once radar rainfall analysis improves in accuracy. In addition to rainfall features identified as predictors of flood damage in

this study site, there may be other hazard factors that influence the characteristics of flood damage in different landscapes. These can typically include the wind speed for typhoon- or tornado-impacted areas, the surge tide or wave height for coastal areas, and the torrent flow velocity or depth for hill slope mountain areas. To improve regression model performance and develop robust regression models for differing landscapes, future research should identify predictor variables from various hazard factors that cause floods.

Table 3 and Figure 6 indicate that the relatively short-duration rainfall R_6 plays an important role in identifying the predictor variable. These were identified as $I_{0.5}$, $I_{0.75}$, or I_1 (equally or more weighted for R_6) in the optimal regression models for all seven eastern coastal districts interfaced between mountains and sea and for some urbanized western districts or those characterized by steep mountain slopes. The relatively long-duration rainfall R_{24} only contributes to the predictor variable I_0 (not weighted for R_6) in the optimal regression model for the four western districts that are relatively large and are less urban. There was no optimal regression model that identified the predictor $I_{0.25}$ (less weighted for R_6 and more weighted for R_{24}) in Gangwon Province. This may imply that the marginal effect of R_6 may be neglected when R_{24} primarily affects the flood damage in some administrative districts. Overall, the optimal regression models identified a composite predictor I_w, with a weighting factor that exceeded 0.5, in most administrative districts, thus indicating that R_6 may be a key driver affecting the nature of flood damage in Gangwon Province with a relatively small and mountainous landscape. This is because peak flow is largely influenced by relatively short-duration rainfall in regions that are relatively small in area with a relatively short concentration time. It was inferred that R_6 was close to concentration times over most administrative districts in Gangwon Province. Meanwhile, the relatively long-duration rainfall R_{24} can increase flood volume, causing inundation in the lowland or riverside areas. Flood damage may be more influenced by R_{24} than R_6 in some relatively large administrative districts in Gangwon Province, which have a relatively long concentration time and a slow drainage rate than relatively small administrative districts. As such, the predictor rainfall variables for the flood damage regression models were found to be greatly influenced by regional landscape features, such as district size, elevation, slope, and urban area. In contrast, there was no apparent relationship between the type of regression functions and the characteristics of regional landscapes in the administrative districts of Gangwon Province. Further quantitative interpretations need to be explored in future research to clearly investigate the detailed relationship between the damage estimate regression models and the regional comprehensive context, including regional resilience and socioeconomic traits.

The flood damage risk map may also be presented by applying a design or forecasting rainfall event to the proposed regression models. Table 4 and Figure 7 indicate that all seven eastern districts were more vulnerable whereas most of the western districts were less vulnerable to flood damage as estimated from the optimal regression models for the 100-year design rainfall. Such damage risk outcomes based on flood damage records are greatly influenced by the nature of the flood damage caused by typhoons that are more frequent in the eastern region than the western region. Although more accurate and detailed flood risk assessments may be estimated using flood hazard and inundation analysis with sophisticated modeling and simulation techniques based on highly detailed geomorphological and hydrometeorological information, this comes at a greater computational cost. In contrast, regression models are simple and effective tools that overcome the limitations or complexity of the hydrological modeling approach. Flood risk assessment normally aims to estimate the probability or potential impact of a flood or consequential loss in conditions of uncertainty. In this context, the proposed method can readily provide information on flood damage estimates using a simple regression model, particularly for ungauged regions. Taking into account the nature of uncertainty in risk analysis, future research needs to comprehensively review the flood damage risk outcomes by analyzing the multidimensional aspects affecting the characteristics of flood damage for each administrative district.

5. Conclusions

This study proposed a methodology to develop an optimal rainfall–flood damage regression model by selecting a predictor rainfall variable that is indicative of regional landscape features. A flood damage risk map was also provided using a regression outcome for a design rainfall condition. To implement the proposed method, this study investigated relationships between rainfall characteristics and economic damage records caused by rainstorms and typhoons over the last two decades from 1999 to 2018 over 18 administrative districts of Gangwon Province, Republic of Korea. The probability of occurrence for the economic damage density was highly correlated to the two maximum rainfall data; this was 6 h for a short-duration and 24 h for a long-duration rainfall event in most administrative districts. This study introduced and employed a single predictor variable from each composite value averaged by the five weights of the two rainfall characteristics in the four types of nonlinear functions; this provided a simple regression analysis that takes into account the sample size for the study site. Landscape characteristics such as district size, topographic features, and urbanization rate had a significant influence on the predictor rainfall variable of the optimal regression model selected from the total 20 regression results for each administrative district. A substantial portion of the variation (32.4–84.6% approximated by the R-squared statistic) was explained by the regional rainfall factor at a significance level of 1% (p-value < 0.01 for the F-statistic) in the optimal regression models. The models were regarded to have a reasonable degree of goodness-of-fit to estimate the relative flood damage risk. They may be improved by incorporating radar data to represent the spatial variability impact of rainfall and wind data to explain typhoon-related impacts on flood damage. As the proposed optimal regression models produced a flood damage risk map for a 100-year frequency rainfall event, other flood damage risk maps may also be developed for weather forecasting information or climate change scenarios. This flood damage risk map for design rainfall was estimated using a regression model based on historical flood damage records. Given that the primary goal of flood risk assessment is to estimate the degree of potential flood effects on a system, the proposed flood damage risk map offers governments and policymakers with basic information on the preliminary financial and risk management of each administrative district against flood damage. For a gauged watershed with sufficient available data for modeling applications, a model-based flood damage estimate may produce flood damage risk maps with greater accuracy and detail at higher complexity and cost in hydrologic/hydraulic modeling and flood inundation analysis. However, when sophisticated models have limited applicability to areas where available data are scarce and uncertain in the required spatio-temporal resolutions, a practicable regression model in statistical hydrology is useful to readily provide flood-related information, particularly for ungauged regions. The proposed flood damage risk map provides a reasonable district-level assessment outcome of potential flood damage at the study site. For reliable validations and extensible applications of the proposed approach based on the statistical relationship between historical characteristics of rainfall and floods, further analysis is required for other regions that preserve the long-term time series of rainfall and flood damage records.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1A2B4005232).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. Flood damage records can be found in the National Disaster Information Center, the Republic of Korea [http://www.safekorea.go.kr]; rainfall observations are openly available in the Korea Meteorological Agency [http://www.kma.go.kr]; frequency analysis rainfall data was obtained from the Water Management Information System [http://wamis.go.kr/WSC/WSC_FREQUENCSY.aspx].

Acknowledgments: This work was supported by the 2020 Yeungnam University Research Grant.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*; McCarthy, J.J., Canziani, O.F., Leary, N.A., Dokken, D.J., White, K.S., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2001.
2. Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Pachauri, R.K., Reisinger, A., Eds.; IPCC: Geneva, Switzerland, 2007.
3. Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Pachauri, R.K., Meyer, L.A., Eds.; IPCC: Geneva, Switzerland, 2014.
4. Centre for Research on the Epidemiology of Disasters (CRED) & United Nations International Strategy for Disaster Reduction (UNISDR). Economic Losses, Poverty & Disasters, 1998–2017. 2018. Available online: https://www.unisdr.org/we/inform/publications/61119 (accessed on 1 January 2021).
5. National Disaster Information Center. Available online: http://www.safekorea.go.kr (accessed on 1 January 2021).
6. Kim, N.W.; Won, Y.S.; Chung, I.M. The scale of typhoon RUSA. *Hydrol. Earth Syst. Sci. Discuss.* **2006**, *3*, 3147–3182.
7. Associated Programme on Flood Management (APFM) Technical Support Unit. *Integrated Flood Management. Concept Paper*, 2nd ed.; APFM Technical Document No.1; APFM: Geneva, Switzerland, 2004.
8. Van Der Most, H.; Marchand, M. *Selecting Measures and Designing Strategies for Integrated Flood Management, a Guidance Document*; World Meteorological Organization (WMO): Geneva, Switzerland, 2017.
9. Choi, H.I. Spatial assessment of damage vulnerability to storms based on the analysis of historical damage cost data in the Korean Peninsula. *Sustainability* **2019**, *11*, 6051. [CrossRef]
10. Albano, R.; Crăciun, I.; Mancusi, L.; Sole, A.; Ozunu, A. Flood damage assessment and uncertainty analysis: The case study of 2006 flood in Ilisua basin in Romania. *Carpathian J. Earth. Environ. Sci.* **2017**, *12*, 335–346.
11. Stedinger, J.R.; Vogel, R.M.; Foufoula-Georgiou, E. Frequency analysis of extreme events. In *Handbook of Hydrology*; Maidment, D.R., Ed.; McGraw-Hill: New York, NY, USA, 1993; pp. 17.1–17.66.
12. Kim, Y.-O.; Seo, S.B.; Jang, O.-J. Flood risk assessment using regional regression analysis. *Nat. Hazards* **2012**, *63*, 1203–1217. [CrossRef]
13. Spekkers, M.H.; Kok, M.; Clemens, F.H.L.R.; Ten Veldhuis, J.A.E. A statistical analysis of insurance damage claims related to rainfall extremes. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 913–922. [CrossRef]
14. Sampson, C.C.; Fewtrell, T.J.; O'Loughlin, F.; Pappenberger, F.; Bates, P.B.; Freer, J.E.; Cloke, H.L. The impact of uncertain precipitation data on insurance loss estimates using a flood catastrophe model. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2305–2324. [CrossRef]
15. Wobus, C.; Lawson, M.; Jones, R.; Smith, J.; Martinich, J. Estimating monetary damages from flooding in the United States under a changing climate. *J. Flood Risk Manag.* **2014**, *7*, 217–229. [CrossRef]
16. Torgersen, G.; Bjerkholt, J.T.; Kvaal, K.; Lindholm, O.G. Correlation between extreme rainfall and insurance claims due to urban flooding—Case study Fredrikstad, Norway. *J. Urban Environ. Eng.* **2015**, *9*, 127–138. [CrossRef]
17. Supriya, P.; Krishnaveni, M.; Subbulakshmi, M. Regression analysis of annual maximum daily rainfall and stream flow for flood forecasting in Vellar River Basin. *Aquat. Procedia* **2015**, *4*, 957–963. [CrossRef]
18. Cortès, M.; Turco, M.; Llasat-Botija, M.; Llasat, M.C. The relationship between precipitation and insurance data for floods in a Mediterranean region (northeast Spain). *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 857–868. [CrossRef]
19. Jato-Espino, D.; Sillanpää, N.; Andrés-Doménech, I.; Rodriguez-Hernandez, J. Flood risk assessment in urban catchments using multiple regression analysis. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04017085. [CrossRef]
20. Korea Meteorological Administration. Available online: http://www.kma.go.kr (accessed on 1 January 2021).
21. Halinski, R.S.; Feldt, L.S. The selection of variables in multiple regression analysis. *J. Educ. Meas.* **1970**, *7*, 151–157. [CrossRef]
22. Miller, D.E.; Kunce, J.T. Prediction and statistical overkill revisited. *Meas. Eval. Guid.* **1973**, *6*, 157–163. [CrossRef]
23. Nunnally, J.C. *Psychometric Theory*, 2nd ed.; McGraw Hill: New York, NY, USA, 1978.
24. Thiessen, A.H. Precipitation averages for large areas. *Mon. Weather Rev.* **1911**, *39*, 1082–1084. [CrossRef]
25. Zhai, G.; Fukuzono, T.; Ikeda, S. Modeling flood damage: Case of Tokai Flood 2000. *J. Am. Water Resour. As.* **2005**, *41*, 77–92. [CrossRef]
26. Cook, R.D. Detection of influential observation in linear regression. *Technometrics* **1977**, *19*, 15–18.
27. Belsley, D.A.; Kuh, E.; Welsch, R.E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*; John Wiley & Sons, Ltd.: New York, NY, USA, 1980; pp. 6–84.
28. Water Management Information System. Available online: http://wamis.go.kr/WSC/WSC_FREQUENCSY.aspx (accessed on 1 January 2021).