

Article

# Gridded Flash Flood Risk Index Coupling Statistical Approaches and TOPLATS Land Surface Model for Mountainous Areas

Byong-Ju Lee <sup>1</sup> and Sangil Kim <sup>2,\*</sup>

<sup>1</sup> Department of Research & Development, Hydrology Engineering & Consulting Center, Korea Inc, Seoul 08380, Korea; bjlee0704@gmail.com

<sup>2</sup> Department of Mathematics, Pusan National University, Busan 46241, Korea

Received: 20 February 2019; Accepted: 5 March 2019; Published: 11 March 2019



**Abstract:** This study presents the development of a statistical flash flood risk index model, which is currently operating in research mode for flash flood risk forecasting in ungauged mountainous areas. The grid-based statistical flash flood risk index, with temporal and spatial resolutions of 1 h and 1 km, respectively, has been developed to simulate the flash flood risk index leading to flash flood casualties using hourly rainfall, surface flow, and soil water content in the previous 6 h. The statistical index model employs factor analysis and multi-linear regression to analyze its gridded hydrological components that are obtained from the TOPMODEL-based Land Atmosphere Transfer Scheme (TOPLATS). The performance of the developed index model has been evaluated in estimating flash flooding in ungauged mountain valleys and small streams. Numerical results show that the approach simulated 38 flash flood catastrophes in the Seoul Capital Region with 71% accuracy; therefore, this approach is potentially adequate for flash flood risk forecasting.

**Keywords:** statistical flash flood forecast; factor analysis; distributed hydrological model; ungauged mountainous area

## 1. Introduction

Flash floods are common and widespread, are a leading cause of weather-related deaths worldwide, and remain one of the most difficult weather phenomena to forecast and provide warnings because of their complex and multifaceted characteristics. Localized torrential rainfall that occurred on 27 July 2011 caused flash floods and mountain landslides on the southern flank of an ungauged mountain in Seoul, the capital of Korea, leading to 60 fatalities [1]. Most of those catastrophes have occurred in ungauged mountainous areas, but methods for classifying the flash flood risk index (FFRI) in ungauged mountainous areas remain few. There is some previous literature forecasting and providing warnings of flash floods for ungauged mountainous areas based on a distributed hydrological model [2–8]. For example, Reed et al. [2] studied flash flood forecasting at ungauged locations with a distributed hydrologic model and threshold frequency-based method. Wang et al. [3] established a flash flood warning system based on a distributed hydrological model applied to two ungauged mountainous regions.

Flooding hazards caused by heavy rainfall can be categorized according to the rainfall duration and ground-level damage type. They are divided into floods and flash floods with respect to rainfall characteristics such as intensity and duration, as well as speed of flood onset. Flooding hazards are also divided into urban, mountain, and riverine flooding based on the flood-affected areas, of which the riverine flooding is subdivided into large-, medium-, and small-scale riverine floods depending on river size. Given that flash floods are closely associated with the speed of onset of flooding irrespective of the type of ground-level damage, they can be referred to as urban flash floods or mountain flash

floods. Because a flash flood is a flooding event that occurs rapidly without warning signs, it is difficult to clearly identify the affected area [9–11]. The World Meteorological Organization (WMO) and the US National Weather Service (NWS) define a flash flood as a flooding event that is due to heavy rainfall and shows signs of flooding, such as a rapid rise in river levels above the flood water level and dike discharge, in less than 6 h from dry ground surface conditions (<http://water.noaa.gov/about/nwm>). In addition, efforts of development of hydrologic models to better predict flooding forecasting have been done such as the National Water Model (NWM) by NWS that simulates observed and forecast streamflow over the entire continental United States [12].

Many researchers have sought criteria for predicting flash flood occurrence and conducted many studies to improve the performance of flash flood forecasting [13–22]. Given a forecast flow at a particular location in a catchment, such as the catchment outlet, criteria are required to anticipate subsequent flooding. The simplest approach is a flow comparison between a modeled flow value and a designed flooding threshold. One alternative method is a statistical-distributed modeling approach compares a modeled flow value with the long-term simulated record of the model. Also, multivariate modelling techniques and GIS-based concepts also have been employed to evaluate flood exposure for properties and to develop flood hazard map in urban areas [23–25]. Furthermore, recent advancements in the area of information and communication technologies have resulted in developing hydrologic information systems that allow real-time applications of quantitative precipitation estimation (QPE) and quantitative precipitation forecast (QPF) derived from rain gauge data, radar, and numerical models. Data for real-time flood forecasting with several hours of preceding time can be obtained by modeling the flow rates and inundation areas using the rainfall data in conjunction with appropriate numerical models, such as urban runoff, river flow, inundation, and flood models.

Additionally, flooding risk could be rapidly determined by drawing a flood risk map for each rainfall scenario well ahead of time. Even though medium- and large-scale rivers with catchment areas exceeding hundreds of square kilometers do not lend themselves well to rainfall forecasting because of the uncertainties of QPF values, rough estimates are possible based on observational rainfall data taking into account of a lag time proportional to the size of the catchment area. Kauffeldt et al. [26] provided a technical review of large-scale hydrological models for implementation in operational flood forecasting schemes at the continental scale. This methodology, however, is difficult to apply in mountain valleys and small catchments, as well as for events with very short lag times, because of a lack of observational hydrologic data, difficulties in reflecting detailed hydrogeological information, and limited capacity to analyze lumped hydrologic models, even excluding the uncertainties of rainfall forecasting.

Instead of comparing the flow forecast with flooding flow, one could compare the rainfall required over a specific area to produce flooding at its outlet using the rainfall forecast. This method is commonly known as Flash Flood Guidance (FFG). FFG is defined as the volume of rainfall of a given duration and distributed uniformly over a small catchment [27–31]. However, application of the concepts of threshold runoff and flash flood guidance using the Manning formula, geomorphologic instantaneous unit hydrographs, and lumped continuous rainfall runoff models, as in previous studies, would cause methodological limitations and uncertainties related to the critical flow rate of flash floods (2-year flood discharge, dike discharge), the accuracy of runoff models for streams and valleys, detailed soil moisture modeling, and reflecting the initial flow rate. The report “Flash Flood Early Warning System Reference Guide 2010” [32] noted that “the significant influence of QPE uncertainties and model errors on the small scale of flash flood occurrence have hindered utilization of distributed models for operational forecasting up to now. Nevertheless, distributed models promise to provide additional information and insight regarding hydrologic conditions at locations without sufficient stream flow observations. As the science of distributed modeling advances and the quality of data input improves, the distributed modeling approach will likely replace FFG.” The more recent development of FFG in conjunction with a distributed hydrological model has led to gridded flash flood guidance [23,33]. However, with the currently available techniques, it is still challenging to address the limitations and compute an FFG.

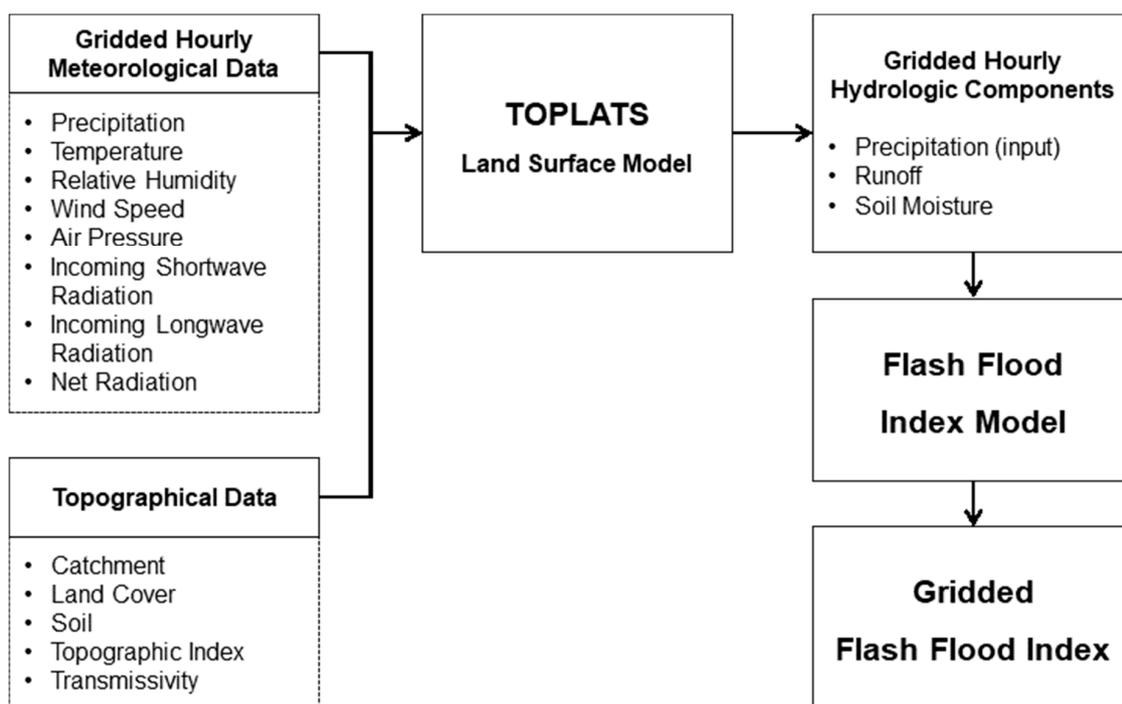
With respect to the hydrological perspective on flash floods, post-flood estimations of peak discharges have been extensively conducted in South Korea to provide effective documentation of flash flood events and enhance understating of the regional behavior of extreme events [34–37]. The Korea flash flood guidance system was developed [38,39]. In addition, ample research endeavors have been focused on capturing apparent trends, temporally and spatially, with respect to the frequency and intensity of extreme rainfall events in the context of climate change, in order to predict future trends. In this paper, the study goal is to develop a statistical flash flood risk index and to evaluate its performance when applied to flash flood events in ungauged mountain basins in the Seoul Capital Region, Korea during the years of 2009–2012. The proposed index model is composed of two procedures: one is a land surface model that provides the hydrological parameters describing different processes occurring on catchments, and the other is the statistical analysis of relationship between the flash flood index and rainfall characteristics. Based on the previous study in Ruiz-Villanueva et al. [8], this study also takes into account non-instrumental data sources such as rescue request times to improve the statistical index in ungauged mountain basins in Seoul Capital Region, Korea.

Therefore, this study focuses on the development and application of a statistical classification method for a flash flood risk index in ungauged mountainous areas to predict potential flood hazards, such as stranding of hikers and campers due to heavy rainfall in mountain valleys and small streams around Seoul in South Korea. That is, the current study focuses on the development and evaluation of such an index in ungauged areas around Seoul. To achieve this goal, the following objectives are pursued in this study: (1) determine the hourly gridded hydrologic components using a land surface model, (2) develop a flash flood index model by analyzing the correlation between simulation results from the land surface model and documented flash flood catastrophes, and (3) develop a high-resolution gridded flash flood index and evaluate its performance. The results of this study allow us to provide FFRI for heavy rain during rainfall seasons and highlight the pressing need to establish measures for efficient flood prevention and damage mitigation.

The remainder of this article is organized as follows. Section 2 describes the study area, data, and methodologies of development for the flash flood forecast model. Section 3 presents experimental results including the calibration and validation for the distributed hydrological model, and the statistical forecasting index model. An extensive discussion is given in Section 4, followed by conclusions in Section 5.

## 2. Methodology

The mountain flash flood forecast model proposed in this study is established based on land surface and statistical FFRI models. The gridded FFRI is obtained following the series of integrated processes shown in Figure 1; it is obtained using the simulated gridded hydrological components from the TOPMODEL-based Land Atmosphere Transfer Scheme (TOPLATS) described in [40,41]. The statistical FFRI model is developed by performing a correlation analysis between the gridded hourly hydrologic components of the surveyed region, yielded from the land surface model, and flash flood casualty cases. Numerous researchers have investigated distributed models for hydrologic forecasting applications [42–46].



**Figure 1.** Schematic architecture of the statistical flash flood index forecasting system. If there are multiple panels, they should be listed as: (a) Description of what is contained in the first panel; (b) Description of what is contained in the second panel. Figures should be placed in the main text near to the first time they are cited. A caption on a single line should be centered.

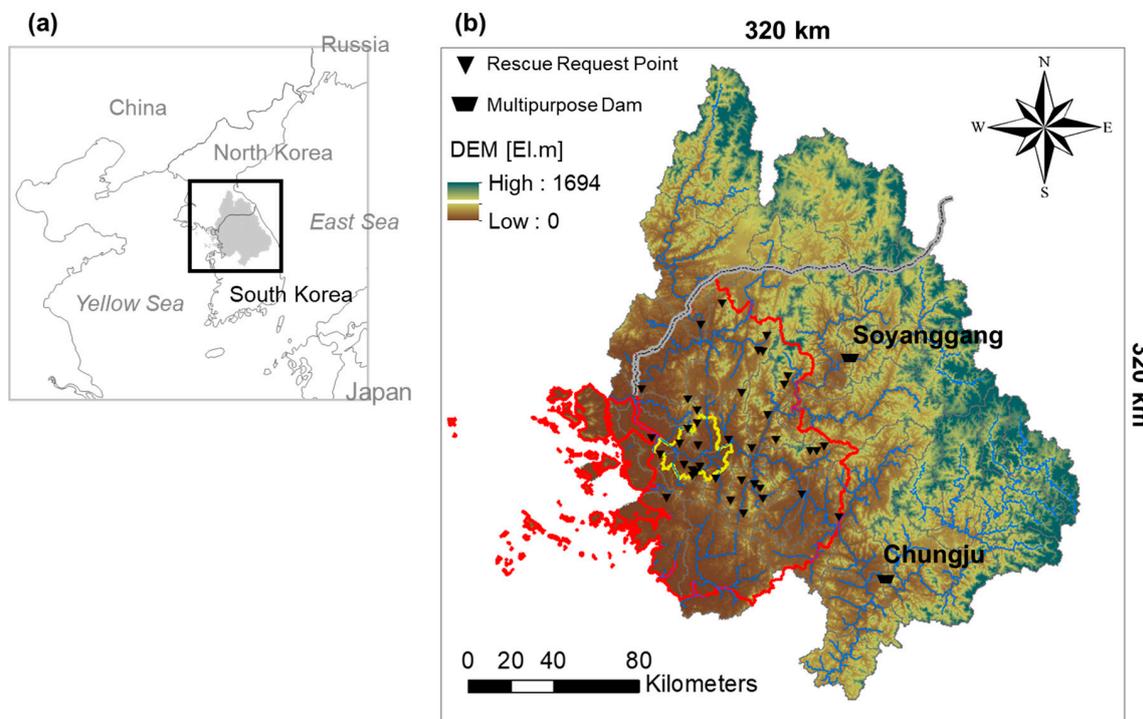
### 2.1. Study Area

Summer monsoonal airflow from the tropics triggers heavy rainfall concentrated over coastal regions of East Asia (see geographic location in Figure 2a). An increasing number of extreme rainfall events have been related to mesoscale convective systems with the synoptic-scale East Asian Summer Monsoon. In most regions of South Korea, such torrential rainfall amounts are more than half the annual precipitation, and have been recently concentrated in urban areas with dense human population and industrial centers. Heavy rainfall results in the isolation of hikers and campers in mountainous areas around Seoul (inside yellow line in Figure 2b), and leads to landslide and flooding that causes human casualties in the region [1].

The Han River basin consists of about 30% lowlands and 70% uplands and mountains (Figure 2b). The Seoul Capital Region (outside red line in Figure 2b) is the metropolitan area centered on Seoul located in northwest South Korea (inside yellow line in Figure 2b), and makes up only 11.8% of the entire land surface with a surface area of 11,930.06 km<sup>2</sup>. However, almost half (48%) the population of South Korean lives in this area. The region is covered with forest area (49.1%), paddy fields (21.6%), and development (12.4%), and the average terrain slope is about 15.7%. Heavy rainfall occurring over short periods in the Seoul Capital Region results in inherent dangers in measuring high peak flows in mountain valleys and small streams; therefore, they remain ungauged. However, emergency rescue requests are available and valuable information, including onset of rainfall and recognition of danger, is provided with the caveat that there are uncertainties in flood peaks, flood amounts, and flooding time.

Over four years (2009–2012), there were 45 rescue request cases related to flooding in the Seoul Capital Region; after excluding ambiguous locations and cases involving negligence, 38 cases were used in this study (black triangles in Figure 2b). Of these, 34 cases (89.5%) involved isolation, death, or inundation due to rapids or rising water levels in valleys or streams, and the remaining cases (10.5%) involved landslides. Isolation and fatalities due to heavy rainfall suggest that the time from the onset

of rainfall to the recognition of danger was very short and that valleys or small streams are at higher risk of incurring such casualties. See Table 1 for more details of each rescue request, such as reported time and catastrophe type provided by the National Disaster Information Center of the Ministry of Public Safety and Security ([www.safekorea.go.kr](http://www.safekorea.go.kr)).



**Figure 2.** The study area in Korea. (a) The black box represents the model domain with grid spacing of 1 km. The shaded area represents the Han River basin. The black line crossing from west to east in the shaded area represents the Demilitarized Zone (DMZ) separating South Korea and North Korea. (b) Numerical model domain for the TOPLATS model (black box in Figure 2a). Color bar indicates the DEM elevation in m, and the blue lines show the river network. The black triangles and trapezoids indicate the rescue request points and major dams, respectively. The red line encircles the Seoul Capital Region and the yellow line encircles Seoul. The thick gray line crossing from west to east indicates the DMZ.

**Table 1.** Cases of flash flood casualties in the capital region from 2009 to 2012.

Case	Date	Report Time	Coordinate (long, lat)	Type	Casualties
1	9 July 2009	13:44	(126.96, 37.44)	valley	isolation: 3
2	9 July 2009	12:46	(126.80, 37.53)	heavy rain	isolation: 11
3	14 July 2009	13:40	(126.80, 37.53)	Stream	isolation: 4
4	10 August 2010	17:21	(126.94, 37.64)	valley	isolation: 32
5	10 August 2010	18:33	(126.90, 37.57)	heavy rain	wet vehicle: 3
6	14 August 2010	22:40	(127.14, 38.21)	heavy rain	rescue
7	15 August 2010	14:05	(127.43, 37.88)	valley	rescue: 52
8	15 August 2010	19:13	(127.39, 38.07)	valley	rescue: 100
9	2 September 2010	15:12	(127.32, 37.96)	stream	isolation: 16
10	10 September 2010	13:55	(127.43, 37.60)	stream	isolation: 2
11	11 September 2010	07:02	(127.44, 37.86)	stream	isolation: 10
12	11 September 2010	19:16	(127.54, 37.33)	stream	isolation: 24
13	12 September 2010	10:46	(126.99, 37.71)	valley	isolation: 20
14	21 September 2010	21:17	(127.09, 37.43)	valley	isolation: 5
15	21 September 2010	13:55	(126.99, 37.57)	valley	isolation: 9
16	21 September 2010	18:47	(127.38, 37.72)	valley	isolation: 4

Table 1. Cont.

Case	Date	Report Time	Coordinate (long, lat)	Type	Casualties
17	25 July 2011	20:07	(126.99, 37.45)	stream	isolation: 12
18	26 July 2011	18:24	(127.25, 37.42)	stream	missing: 2
19	27 July 2011	10:05	(127.26, 37.42)	stream	death: 1
20	27 July 2011	14:03	(127.34, 37.39)	stream	death: 1
21	27 July 2011	14:15	(127.74, 37.28)	stream	death: 1
22	27 July 2011	16:20	(127.58, 37.56)	landslide	burying: 1
23	27 July 2011	09:50	(127.17, 37.32)	stream	missing: 1
24	27 July 2011	16:45	(126.70, 37.80)	stream	missing: 1
25	27 July 2011	16:45	(127.32, 37.96)	landslide	missing: 1
26	27 July 2011	18:32	(127.00, 37.66)	landslide	death: 2, injury: 3
27	27 July 2011	22:29	(127.67, 37.54)	landslide	death: 3, injury: 2, bury: 2
28	31 July 2011	17:25	(126.95, 37.48)	valley	isolation: 4
29	31 July 2011	18:07	(127.15, 37.60)	valley	isolation: 7
30	3 August 2011	17:34	(127.02, 38.09)	valley	isolation: 3
31	12 August 2011	16:33	(127.23, 37.79)	heavy rain	isolation: 7
32	14 August 2011	12:41	(127.30, 37.54)	heavy rain	isolation: 80
33	13 July 2012	20:33	(127.61, 37.58)	stream	isolation: 1
34	15 July 2012	14:10	(127.43, 37.88)	stream	isolation: 3
35	19 July 2012	09:56	(127.44, 37.86)	stream	isolation: 1
36	15 August 2012	13:36	(126.99, 37.71)	valley	isolation: 12
37	15 August 2012	13:59	(126.99, 37.71)	valley	isolation: 18
38	15 August 2012	15:34	(127.00, 37.66)	valley	isolation: 8

## 2.2. Land Surface Model

The TOPLATS model is a multi-scale model for simulating local- to regional-scale catchment water fluxes, and utilizes water balance and energy balance to simulate gridded actual evapotranspiration, soil water content, water table depth, surface runoff, latent heat, sensible heat, ground heat, and net radiation to characterize the redistribution of the water table depth at the sub-catchment scale. The model combines a soil vegetation atmosphere transfer scheme (SVAT) to represent local scale vertical water fluxes with the catchment scale TOPMODEL approach [47] to laterally redistribute the water within a catchment. See Table 2 for the main processes and equations in the TOPLATS model [48].

Table 2. Main processes and equations in the TOPLATS model.

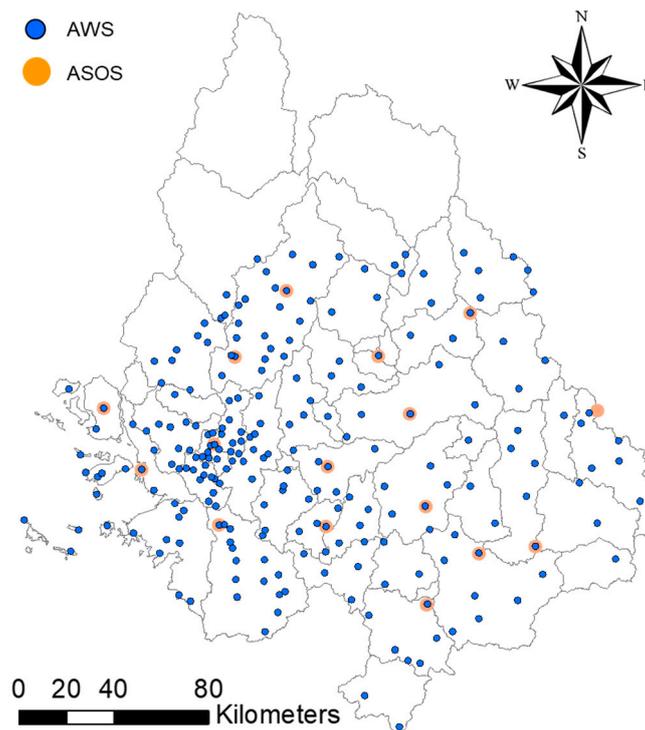
Model Part	Process	Approach
Local SVAT <sup>1</sup>	Interception	Storage approach: storage capacity is proportional to leaf area index
	Potential evapotranspiration (PET)	Penman Monteith equation [49] (plant specific PET)
	Actual evapotranspiration	Reduction of PET by actual soil moisture status (alternative: solving energy balance equation)
	Infiltration	Infiltration capacity after Milly [50] (depending on soil properties and soil water status)
	Infiltration excess runoff	Difference between rainfall rate and infiltration capacity
	Saturation excess runoff	Contributing areas derived from TOPMODEL; approach based on the soils topographic index
	Percolation Capillary rise	Gravity driven drainage
	Lower boundary condition	Capillary rise from local water table on Gardner [51] using Brooks and Corey parameters [52] Top of capillary fringe
TOPMODEL	Spatial distribution of water table depths base flow	Soils-topographic index [53] Exponential decay function; maximum depth flow is base flow at basin saturation

<sup>1</sup> SVAT is soil vegetation atmosphere transfer scheme.

Since the TOPLATS is a grid-based and time-continuous model, and the vertical water fluxes of the grid cells are calculated by the local SVATs, catchment-scale vertical water fluxes are obtained by aggregation of local water fluxes. There is no lateral interaction between the local SVATs accounted for by the model. However, based on the soils topographic index of the TOPMODEL approach [54], a lateral redistribution of water is realized by adaptation of the local ground water levels, which are used as lower boundary conditions of the local SVATs. In addition, the based flow is generated from the integration of local saturated subsurface fluxes along the channel network. A routing routine is not integrated in the model. In the vertical direction, the soil is divided into 2 layers (root zone and transmission zone). According to Sivapalan et al. [53], it is assumed that saturated conductivity exponentially decreases with depth. The percolation is calculated using an approximation for gravity driven drainage, and capillary rise is calculated based on the approach of Gardner [51], both approaches using the Brooks and Corey parameterization of soil retention characteristics [52]. Based on soil texture and porosity, soil parameters are derived using the pedotransfer function of Rawls and Brakensiek [55]. Soil water contents within the root zone and transmission zone are calculated using the soil water balance equations [56]. Plant growth is not directly simulated by TOPLATS, but the seasonal development of plant properties is described by monthly updating the plant parameter sets consisting of e.g., leaf area index, plant height and stomatal resistance. The digital elevation model serves as the basic data set for the calculation of the topographic wetness index [54], which is used for calculation of the soil-topographic index additionally accounting for local differences in transmissivity [53]. For further details about the model, the reader is referred to Famiglietti and Wood [56] and Peters-Lidard et al. [41].

The TOPLATS-based simulation in full distribution mode requires gridded meteorological and topographical data of the studied region [26]. Meteorological data includes datasets of ground-level precipitation (mm), temperature ( $^{\circ}\text{C}$ ), relative humidity (%), wind speed (m/s), air pressure (mm), incoming shortwave and longwave radiation ( $\text{W}/\text{m}^2$ ), and net radiation ( $\text{W}/\text{m}^2$ ). Topographical data includes datasets for catchment, land cover, soil, topographic index, and transmissivity. The TOPLATS model can redistribute ground table depth within a single catchment and should thus be reconstructed by catchment unit (Figure 3), although it is a grid column model. The boundaries of the survey area are set in a square ( $320\text{ km} \times 320\text{ km}$ ) with grid resolution of  $1\text{ km} \times 1\text{ km}$ , and the TM coordinates on the four corners are lower-left (320,000, 80,000), lower-right (320,000, 40,000), upper-right (640,000, 40,000), and upper-left (640,000, 80,000) in transverse Mercator map projection.

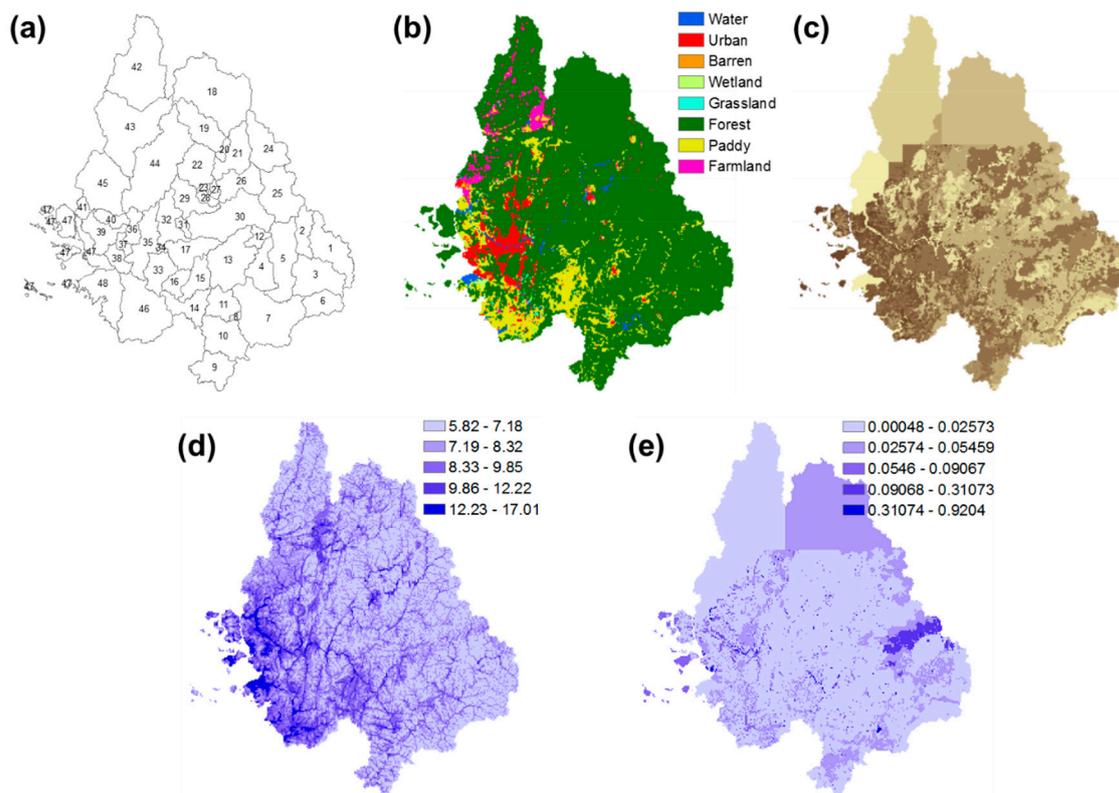
Gridded meteorological data for precipitation, air temperature, relative humidity, wind speed, solar radiation, and air pressure are necessary for simulating catchment water balance with TOPLATS. Given that the purpose of this study is to produce hydrological components per  $1\text{ km} \times 1\text{ km}$  grid, we collected the observations from 600 automated weather systems (AWS) and 72 automated surface observing system (ASOS) stations operated by the Korea Meteorological Administration (KMA). Figure 3 shows the locations of the AWSs and ASOS stations in the Han River basin, which is divided into 48 sub-basins. These data are applied to the model after converting them from point data to grid cells using the inverse distance weighting method.



**Figure 3.** Locations of the automatic weather system (AWS) stations, blue circles, and automated surface observing system (ASOS) stations, orange circles, in the Seoul Capital Region. The catchment is divided into 48 sub-basins, as indicated with the gray lines.

For a smooth simulation of ground-level hydrologic responses to meteorological conditions, an accurate topographical mapping is of vital importance (Figure 4). To construct a TOPLATS model gridded at a resolution of  $1 \text{ km} \times 1 \text{ km}$ , a gridded catchment map with the same resolution, land cover map, topographic index, and transmissivity coefficient data are used. The gridded catchment map is drawn based on a gridded Han River catchment map, dividing the survey area into 48 sub-catchments (Figure 4a). The land cover map, drawn using  $30 \text{ m} \times 30 \text{ m}$  data provided by the Water Resource Management Information System (<http://www.wamis.go.kr>) of the Ministry of Land, Infrastructure and Transport, includes eight land cover types: water, urban, barren, wetland, grasslands, forest, paddy, and farmland (Figure 4b). It is very important to reflect detailed soil properties to successfully simulate spatial water balance using a land surface model. The National Institute of Agricultural Sciences distributes a detailed soil map (1:25,000) classifying the entire South Korean land surface into 1300 soil series in the Soil Information System (<http://soil.rda.go.kr/soil>) (Figure 4c). We use this soil map to construct the soil property distribution and attribute databases. The study area has 379 detailed soil textures. The topographic index and transmissivity are used for redistributing the water table depth within the catchment, as shown in Figure 4d,e, respectively.

Please note that two different types of water excess lead to overland flow: infiltration-excess and saturation-excess. The infiltration-excess is the maximum rate at which water can enter the soil surface [57]. That is, if the rainfall intensity is heavier than the soil surface can absorb, excess water will pond on the surface. Eventually, water starts flowing downslope as overland flow. This overland flow also develops when rain falls on saturated soil; there is no possibility of infiltration and the infiltration capacity is effectively zero. In general, single extreme values of soil transmissivity do not show a homogenous behavior in different catchments [48]. The behavior tends to depend on local soil hydraulic conditions and soil depth.



**Figure 4.** Topographical input data; (a) catchment with 48 sub-basins; (b) land cover classification; (c) soil type; (d) topographic index; and (e) transmissivity. Please note that there is no available information north of the DMZ line (Figure 2b) for soil type and transmissivity.

### 2.3. Statistical Model

Flash flood indices such as the Flash Flood Index (FFI), Flash Flood Potential Index (FFPI), and Flash flood severity index (FFSI) have been previously proposed. The FFI is a quantitative index that calculates the differences between the average basin rainfall and predetermined FFG produced by the NWS River Forecast Centers [58]. As a result of the data assimilated into the FFG product, the FFI is limited to areas containing relatively large gauged rivers [59]. The FFPI accounts for watershed physiographic characteristics and combines them with the forecast and observed rainfall to determine the likelihood of flash flood occurrence. The method was shown to have poor skill in forecasting flash flooding when applied operationally for flash flood forecasting in western United States [60]. The FFSI is a damage-based post-event assessment tool with five categories ranging from 1 to 5, from least to most destructive, and still additional refinement is needed [61].

In this study, variables sensitive to the flash flood occurrence are determined from factor analysis, which is a useful tool for investigating variable relationships for complex concepts that are not easily measured directly [62]. The method collapses a large number of variables into a few interpretable underlying factors. Here, factor analysis reveals that flash flood occurrence is sensitive to precipitation, surface runoff, and soil water content, among the hydrological components produced by the land surface model simulation (see Section 3.1). The statistical FFRI model is then expressed as the first-order linear function of the factor analysis results for surface runoff and rainfall during the six hours preceding the target time. That is, the time-series of precipitation, surface runoff, and soil water content 6 h prior to the current time are expressed using a first-order linear function based on the factor analysis and regression analysis results.

The basic formula for the statistical FFRI model is expressed as follows:

$$FFI = F_{PCP} \times w_1 + F_{RUN} \times w_2 + F_{SWS} \times w_3 \quad (1)$$

where the factor scores  $F_{PCP}$ ,  $F_{RUN}$ , and  $F_{SWS}$  respectively represent the components of the flash flood index for precipitation (PCP), i.e., the amount of rainfall; surface runoff (RUN), i.e., the rate of rainfall exceeding that infiltration capacity of the soil, or additional rainfall after saturated soil water content; and soil water content (SWS), i.e., quantity of water contained in a given mass of the soil. The coefficients  $w_1$ ,  $w_2$ , and  $w_3$  are the weights of the factor scores. The factor scores and their weights are calculated based on all data available over the four years between 2009 and 2012, and the FFRI model is applied for all flooding occurrences over those years.

Please note that in Equation (1), there are two different kinds of water excesses that provide overland flow: infiltration-excess and saturation-excess. The infiltration-excess is excess water that cannot infiltrate ponds on the soil surface, and eventually flows downslope across the soil surface as overland flow. Saturation-excess is subsurface flow returning to the surface when the capacity of the soil to transmit flow is exceeded. Therefore, any rain falling on the saturated zone adds to the overland flow [57]. In the TOPLATS model, the transmissivity parameter of the soils controls the infiltration-excess; therefore, excess rainfall directly adds to the overland flow [48], and is represented in the RUN variable in the statistical model in Equation (1). Rainfall directly onto the already saturated soil adds to the return flow, this overland flow is represented in the SWS parameter in Equation (1).

### 3. Numerical Results

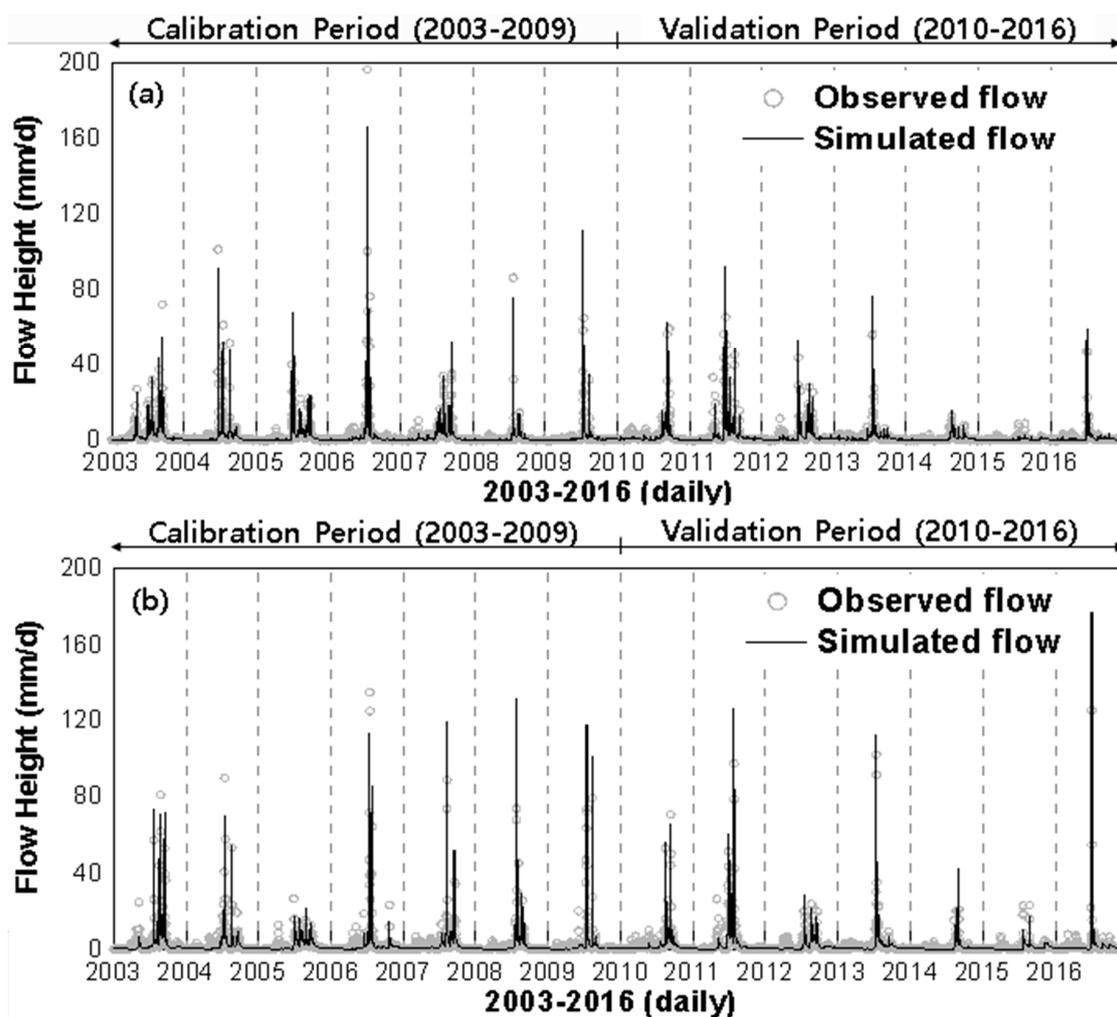
#### 3.1. TOPLATS Results

The most important component of mountain flash flood prediction using the gridded components obtained from the land surface model is the accuracy of the simulated hydrological components in describing hydrological responses to flash flood risk areas. For model calibration and validation, this study considers a 7-year calibration period (2003–2009), and 7-year validation period (2010–2016). The parameter estimation procedures in the TOPLATS model and its streamflow simulation values obtained here are similar to those reported in [45]. The accuracy of hydrologic components is evaluated using observed inflow data from two dams: Chungju Dam, the largest multipurpose concrete dam in Korea, and Soyanggang Dam, the world's fourth largest rock-fill dam containing 29 million tons of water (Figure 2b). It is because in the study area there were no available gauges so, for simplification and to facilitate catchment behavior comparison, we compared simulated catchment outflow with the daily amount of dam inflow. Please note that, except for this validation, all other simulation results are performed and compared with hourly model results.

Please note that in TOPLATS model a channel routing module is unavailable, so the stream flow itself cannot be simulated in the model. In general, the stream flow in the model is underestimated in the beginning of the rainy season and overestimated in the high rainy season and in the end of the rainy season. Thus, during the rainfall, the instant peaks of precipitation generally do not correspond to the peaks of dam inflow in time partly because water from remote places to the dam takes times to arrive the dam. However, peak flows in the model are well simulated, except that the highest peak is simulated with some delay, as no routing is integrated in the model. Therefore, the simulated flow is defined as the sum of runoff and base-flow, and inflow from the simulated results for daily water balance are compared with observed dam inflows.

Figure 5 shows hydrographs of upstream catchments for Chungju Dam and Soyanggang Dam. During the calibration periods, the parameterization of the TOPLATS model is conducted by deriving or directly using as many parameters as possible from standard databases, such as topographic image data and their attribute database. Calibration can be reduced to an adjustment of plant-specific stomatal resistances using a constant factor to satisfy with the long-term water balance and calibrate the parameters for the base flow recession curve. For the validation periods, the accuracy of the simulation is satisfactory for both Chungju Dam (Figure 5a) and Soyanggang Dam (Figure 5b) in terms of total water balance. The relative volume error for the total water budget in each analysis period is satisfactory, in the range of  $\pm 10\%$ . This results in ratios of relative volume error for the total water

budget (%) for calibration and validation of  $-7.81\%$  and  $-7.55\%$ , respectively, for Chungju Dam, and  $-4.51\%$  and  $-4.84\%$ , respectively, for Soyanggang Dam. Although the statistical skill scores for the validation period are only slightly worse than for the calibration periods, the statistical values show that the model replicated the observed outcomes well, i.e., with statistical significance.



**Figure 5.** Hydrographs of upstream catchments for (a) Chungju Dam and (b) Soyanggang Dam. The gray circles are observations of daily dam inflow and black solid lines are simulated daily flow between 2003 and 2016. The years 2003–2009 are used for calibration, and 2010–2016 is the validation period.

Figure 6 shows that in daily discharge, the root mean squared error (RMSE) and coefficient of determination ( $R^2$ ) values for the averaged inflows show moderate quality. The RMSE and  $R^2$  for Chungju Dam are, respectively, 16.35 mm and 0.67 for calibration, and 9.53 mm and 0.60 for validation. The RMSE and  $R^2$  for Soyanggang Dam are, respectively, 15.97 mm and 0.74 for calibration and 7.39 mm and 0.82 for validation. The RMSE values for the calibration period are the lower than that for validation because there was more precipitation during the calibration period than during the validation. That is, the average annual precipitation at Chungju Dam is 1389.7 mm in the calibration period and 1152.4 mm in the validation period, while it is 1480.2 mm in the calibration period and 1242.3 mm in the validation period at Soyanggang Dam.

Figure 7 shows hourly spatial distribution maps for precipitation (PCP), surface runoff (RUN), and soil water content (SWS) over the sub-basin No. 31 between 1300 UTC 27 July 2011 and 1600 UTC 27 July 2011. The maximum rainfall is 64.1 mm, and the basin mean rainfall is 19.30 mm at 1300 UTC 27 July 2011. Subsequently, the rainfall amounts gradually decrease. The surface runoff is less than

rainfall amount due to the infiltration capacity of the soil, and the spatial distribution of the surface runoff is similar to the precipitation. The soil water content shows that the soil along streams has higher water quantity than other regions, and soils in the upper stream areas also show higher water content than those in downstream regions. Table 3 shows the basin means for rainfall, surface runoff, and soil water content during the flood event.

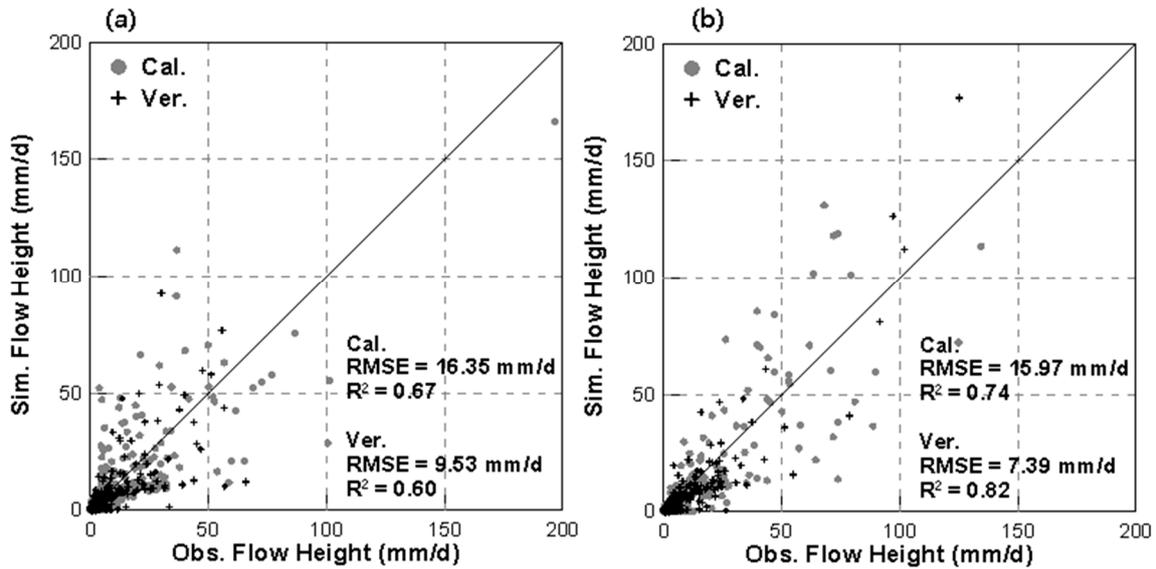


Figure 6. Scatter plots of observed flow height (x-axis) and model results (y-axis) for (a) Chungju Dam and (b) Soyanggang Dam. The flow height data for the calibration period are indicated with gray circles and verification period with black plus signs.

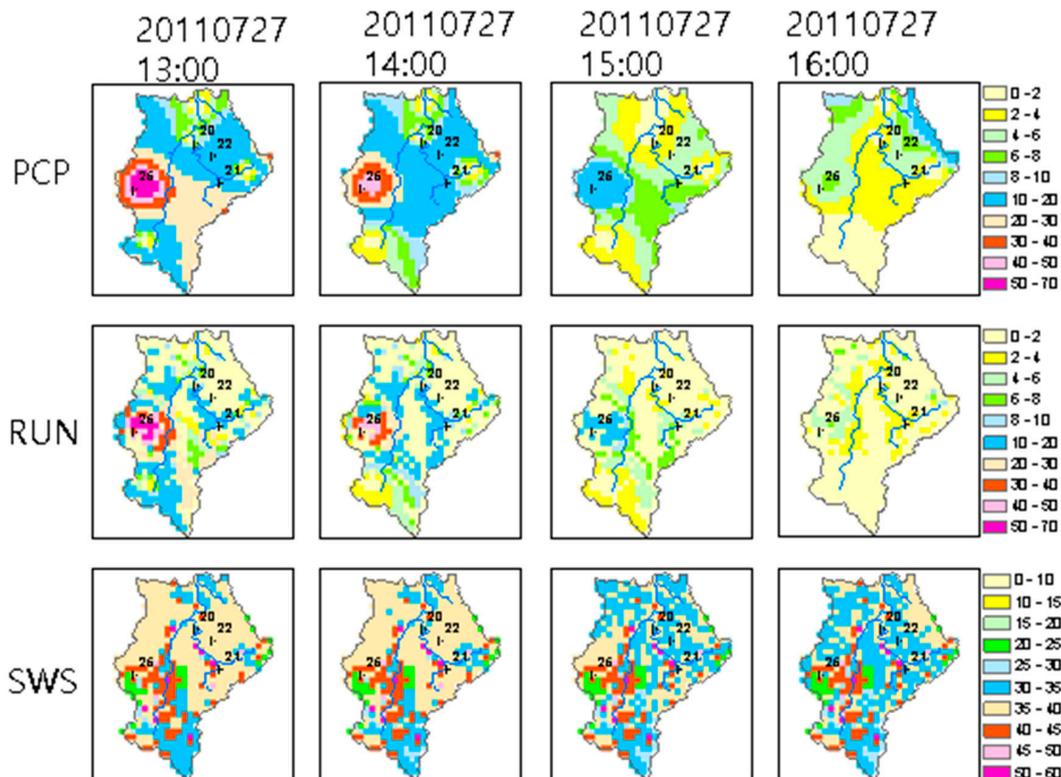


Figure 7. Spatial distribution map of rainfall (PCP), overland flow (RUN), and root zone soil water (SWS) at 1 km resolution in sub-basin No. 31 (see Figure 4a) for flash flood event (13:00–6:00 27 July 2011).

**Table 3.** Basin mean for rainfall (PCP), overland flow (RUN), and soil water in root zone (SWS) in sub-basin No. 31 for the flood event on 21 July 2011.

Time (UTC)	PCP (mm/h)	RUN (mm/h)	SWS (%)
21 July 2011 13:00	19.39	10.25	35.77
21 July 2011 14:00	13.45	6.14	35.58
21 July 2011 15:00	5.57	2.22	34.85
21 July 2011 16:00	4.0	1.07	34.60

Overall, the surface runoff, soil water content, and water table depth generally show satisfactory simulation results with respect to the precipitation time-series in the simulated grids, which suggests that the water balance analysis from the TOPLATS distributed land surface model works with moderate quality at the grid level. Because previous similar studies have shown similar results [43,45] and further improvement of the TOPLATS model seems beyond of the scope of this study, these default simulation parameters in the model are employed to evaluate the performance of the statistical flash flood risk index model.

### 3.2. Statical Flash Flood Risk Index Model

As mentioned before, WMO and US NWS define a flash flood as a rapid and extreme flow of high water into a normally dry area, or a rapid water level rise in a stream or creek above a predetermined flood level beginning within six hours of the causative event, e.g., intense rainfall, dam failure, or ice jam. However, the actual time threshold may vary regionally. Ongoing flooding can intensify flash flooding in cases where intense rainfall results in a rapid surge of rising flood waters (<http://w1.weather.gov/glossary/index.php?letter=f>). In Korea, the KMA issues a heavy rain special report, flood warning, on the basis of 6 h or 12 h total rainfall. Specifically, a watch level is issued when the predicted total rainfall exceeds 70 mm in 6 h or 110 mm in 12 h, and a warning level is issued when the predicted total rainfall exceeds 100 mm in 6 h or 180 mm in 12 h. In the following, two advisory thresholds for watch and warning are selected as flood risk criteria, and the parameters in the statistical FFRI model are determined on the basis of 6 h total rainfall.

#### 3.2.1. Factor Analysis

Factor analysis is a statistical method used to describe variability between observed, correlated variables based on a potentially lower number of unobserved variables, termed factors [62]. Each factor captures a certain amount of the overall variance in the observed variables, and factors are always listed in order of the amount of variation they explain. The eigenvalue is a measure of the amount of variance in the observed variables a factor explains. Any factor with an eigenvalue  $\geq 1$  explains more variance than a single observed variable [63].

Table 4 shows the factor analysis results for surface runoff and rainfall data for the 6 h prior to the flash flood. In general, the number of factors is determined at the point where the eigenvalue is  $\geq 1$  or the cumulative explanatory power for the total variance is within the range of 60–70% or higher. Rainfall and surface runoff are explained using three factors. Categorizing the factors using the values of the rotated component matrix in Table 4, the three rainfall factors explain 84.3% of the total variance. In more detail, factors 1, 2, and 3 are short-term (1 h, 2 h, 3 h), mid-term (4 h), and long-term (5 h, 6 h) rainfall and have explanatory powers of 35.0%, 23.3%, and 26.0%, respectively. The three factors for surface runoff are short-term (1 h, 2 h), mid-term (3 h), and long-term (4 h, 5 h, 6 h) and have explanatory powers of 31.0%, 17.7%, and 42.3%, respectively.

**Table 4.** Factor analysis results for (left column) rainfall and (right column) runoff. The shading indicates the statistically significant factors with high factor loading numbers.

	Rainfall			Runoff		
	Short	Mid	Long	Short	Mid	Long
1 h	0.90	0.26	−0.06	0.86	0.40	0.06
2 h	0.93	−0.03	0.13	0.94	0.11	0.19
3 h	0.64	0.59	0.01	0.34	0.89	0.26
4 h	0.1	0.82	0.43	0.34	0.08	0.84
5 h	0.15	0.54	0.71	0.08	0.29	0.92
6 h	−0.04	0.13	0.93	0.02	0.10	0.94
Eigenvalue	2.1	1.4	1.56	1.86	1.06	2.54
Variance explained (%)	35.0%	23.3%	26.0%	31.0%	17.7%	42.3%
Total variance explained	84.3%			84.3%		

The factor scores of surface runoff and rainfall obtained through factor analysis are expressed in Equations (2)–(7), where the weights are the values of the rotated component matrix (shaded values in Table 2).  $P_x$  and  $Q_x$  are the hourly rainfall (mm/h) and surface runoff (mm/h) at the preceding time  $x$ . The factor scores  $F_{PCP,S}$ ,  $F_{PCP,M}$ , and  $F_{PCP,L}$  indicate short-term, mid-term, and long-term rainfall, respectively, and  $F_{RUN,S}$ ,  $F_{RUN,M}$ , and  $F_{RUN,L}$  respectively indicate the factor scores of short-term, mid-term, and long-term surface runoff, such that

$$F_{PCP,S} = \frac{0.90 \times P_1 + 0.93 \times P_2 + 0.64 \times P_3}{0.90 + 0.93 + 0.64} \quad (2)$$

$$F_{PCP,M} = P_4 \quad (3)$$

$$F_{PCP,L} = \frac{0.71 \times P_5 + 0.93 \times P_6}{0.71 + 0.93} \quad (4)$$

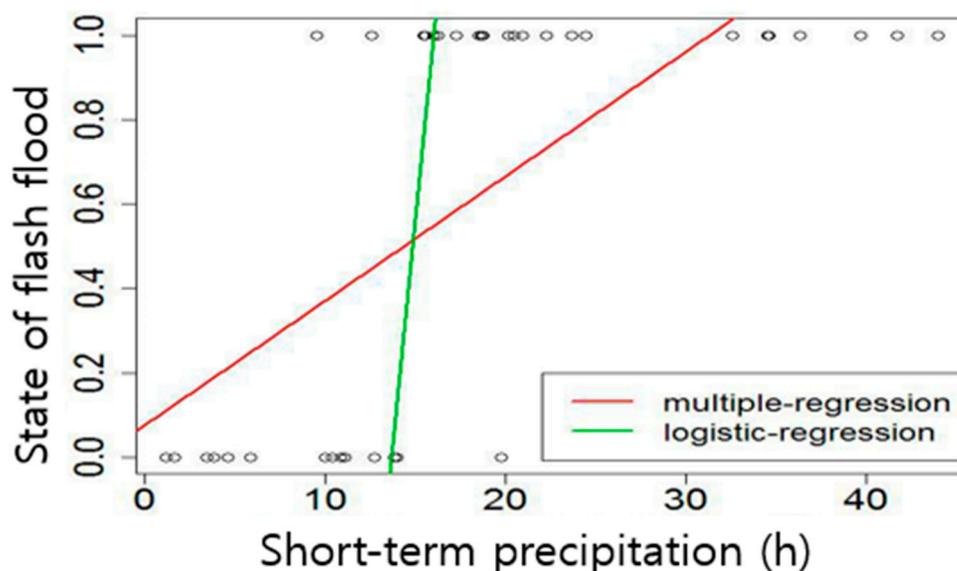
$$F_{RUN,S} = \frac{0.86 \times Q_1 + 0.94 \times Q_2}{0.86 + 0.94} \quad (5)$$

$$F_{RUN,M} = Q_3 \quad (6)$$

$$F_{RUN,L} = \frac{0.84 \times Q_4 + 0.92 \times Q_5 + 0.94 \times Q_6}{0.84 + 0.92 + 0.94} \quad (7)$$

### 3.2.2. Linear Regression Analysis

Because the presence (rescue reports) or absence (no rescue reports) of a flash flood is a binary data point, a logistic regression model is generally accepted to be an appropriate method. However, in the scatter plot for short-term rainfall, which is the factor that exerts the greatest influence on flash flood occurrence (Figure 8), the slope of the linear regression model is less than that of the logistic regression model. A sharper slope of the logistic regression model can better differentiate the presence or absence of flash flood. However, it is inappropriate as a qualitative index for calculating risk levels because of the flash flood potential interval (0%–100%). Therefore, the logistic regression model is inadequate for use in the forecasting and warning system, and the linear regression model is more suitable for smoothly expressing the flash flood index.



**Figure 8.** Scatter plot of short-term precipitation versus flash flood status. The circles indicate the 38 flooding occurrences. The red and green lines are respectively the results of the multiple linear regression and logistic regression.

Table 5 shows the results of the linear regression analysis. The parameters B and SE indicate the regression coefficient and its standard error, which represents the uncertainty of the regression coefficient. Statistical significance of the estimated B in Table 3 is tested using the test statistics with a t value that measures the degree of agreement between a data sample and the null hypothesis of the normal distribution for the ratio of B and SE. Then, F statistics are used in combination with the p-value to determine the significance of the estimated regression analysis. See [64] for more details. The analysis of flash flood occurrence using the linear regression model revealed that surface runoff ( $F = 18.5, p < 0.001$ ) and rainfall ( $F = 59.52, p < 0.001$ ) are statistically significant. Factors exerting important influence on flash flood occurrence are short-term surface runoff ( $t = 5.502, p < 0.001$ ), short-term rainfall ( $t = 7.356, p < 0.001$ ), and mid-term rainfall ( $t = 2.077, p = 0.045$ ). The  $R^2$  values for runoff and rainfall are 61.3% and 83.6%, respectively (Table 5). The unstandardized coefficients and model explanatory power derived here for each factor are applied to the flash flood index model.

**Table 5.** Results of the regression analysis.

	Run			
	B	SE	t	P
short	0.037	0.007	5.502	0.001
mid	−0.011	0.013	−0.873	0.389
long	0.014	0.014	1.012	0.318
$R^2 = 0.613, F = 18.5, p < 0.001$				
	Rainfall			
	B	SE	t	P
short	0.025	0.003	7.356	0.001
mid	0.022	0.011	2.077	0.045
long	0.007	0.009	0.812	0.422
$R^2 = 0.836, F = 59.52, p < 0.001$				

Parameters B and SE represent the regression coefficient and standard error, respectively. The statistical values t and P are statistical values for hypothesis testing, and significance values, respectively.  $R^2$  values are the coefficients of determination.

### 3.2.3. Flash Flood Risk Index Analysis

The statistical FFRI model is then obtained using the first-order linear function of the results from the factor analysis of surface runoff and rainfall during the 6 h preceding the target time; the state variable soil water content is expressed as in Equations (8)–(10).  $F_{PCP}$ ,  $F_{RUN}$ , and  $F_{SWS}$  indicate the factor scores of precipitation, surface runoff, and soil water content, respectively. Then, the weights of three factor scores for each factor are determined using the non-standard coefficient B in Table 3.

$$F_{PCP} = 0.025F_{PCP,S} + 0.022F_{PCP,M} + 0.007F_{PCP,L} \quad (8)$$

$$F_{RUN} = 0.0037F_{RUN,S} - 0.011F_{RUN,M} + 0.014F_{RUN,L} \quad (9)$$

$$F_{SWS} = \frac{SW}{SW_{SAT}} \quad (10)$$

Here,  $SW$  and  $SW_{SAT}$  are soil water content at the target time (%) and saturated soil water content (%), respectively, and the factor score  $F_{SWS}$  is expressed as the current soil water content divided by the saturated soil water content (i.e., soil-to-water ratio).

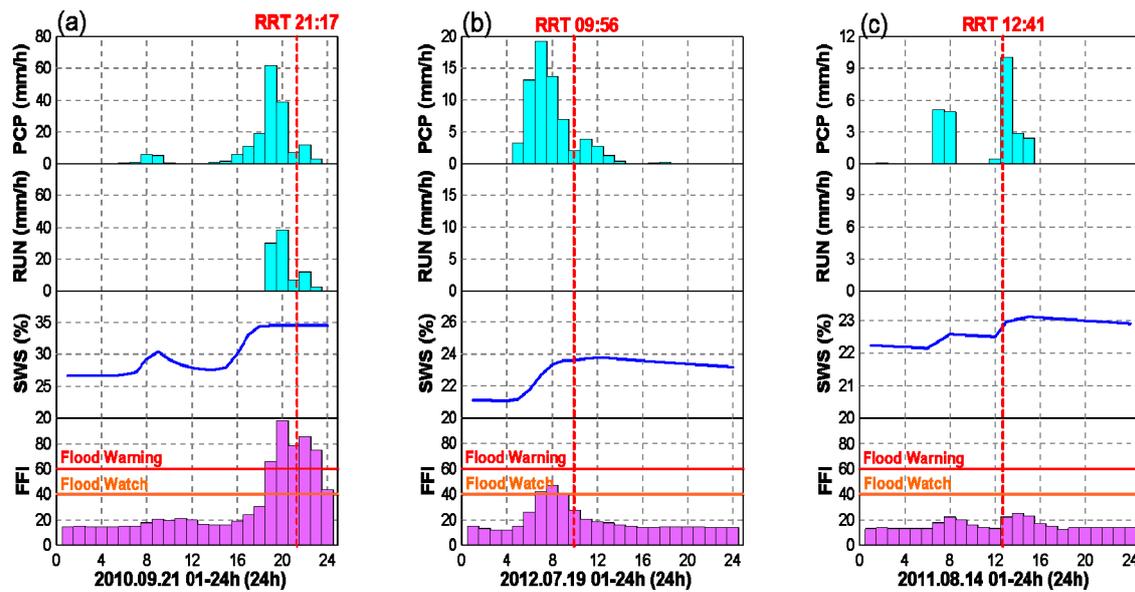
The statistical measure  $R^2$  values in Table 5 are used as the weights for the precipitation and soil runoff in the FFRI model in Equation (1) because the  $R^2$  value indicates that the model explains the percentage of variability in the response data around its mean. Then, the FFRI is obtained from the gridded hydrological components as follows:

$$FFRI = w_1 \times \frac{0.836F_{PCP} + 0.613F_{RUN}}{1.449} + w_2 \times F_{SWS} \quad (11)$$

To obtain the weights in Equation (11), we first set the sum of weights in FFRI to 80. The index at 40 (or 60) corresponding to half (or three-fourths) of the range of the flash flood level is set as the threshold for a warning level to be issued. We optimized the value  $w_1$  that minimizes the flooding advisory of 40 and maximizes the FFRI values subject to the available data such that we obtained  $w_1 = 61.5$  and  $w_2 = 18.5$ .

Catastrophes caused by ungauged mountain flash floods mostly occur when rapids that form in streams and valleys isolate hikers or campers. Given the difficulty of accurate parameterization of the land surface hydrological dynamics in remote mountainous areas, predicting the occurrence or non-occurrence of flash floods for each hour preceding the target time is a more plausible and practicable approach than simulating river level or inundation area in detail. We obtained FFRI values for the survey area using the statistical FFRI model with temporal and spatial resolutions of 1 h and 1 km, respectively.

Figure 9 presents the results of computing the time-series flash flood index according to various conditions based on cell-based hydrological components corresponding to the rescue request times and places. Figure 9a describes a case in which heavy rainfall exceeding 50 mm/h resulted in surface runoff over 30 mm and soil saturation, causing the FFRI to exceed the warning level (60). Moreover, the rainfall and surface runoff fell below 10 mm at the RRT, but the flash flood index exceeded 60 in the simulation, reflecting the values of the hydrological components during the preceding 6 h. In Figure 9b, surface runoff did not occur due to low soil water content, but the flash flood index exceeded the watch level (40) because of cumulative rainfall. Figure 9c shows a case in which gridded rainfall is low, resulting in little change in soil water content and no surface runoff, and the FFRI remains below 40. This suggests that insufficient rainfall did not trigger any changes in the simulated hydrological components, which may be due to the uncertainty in gridded rainfall.



**Figure 9.** Time series of simulated hydrological components for the flash flooding cases on (a) 21 September 2010, (b) 19 July 2012, and (c) 14 August 2011. The x-axis is time (h), and the panels from top to bottom indicate Precipitation (PCP, mm/h), Surface Runoff (RUN, mm/h), Soil Water Content (SWS, %), and statistical flash flooding risk index (FFRI). The vertical dashed red lines are Rescue Request Time (RRT), and the horizontal red and orange lines are the predetermined flood warning and flood watch, respectively.

#### 4. Discussion

To investigate the performance of the statistical approach to flash flood forecasting in ungauged mountainous areas, we should quantify and validate its accuracy against the observed data. There are individual differences in flash flood Rescue Request Time (RRT) even for the same flood event, so it is important to run accurate simulations of signs of flash flood in the pre-RRT hours in addition to calculating the FFRI at the rescue request time.

To calculate the accuracy of the cases seen in Table 1 for issued flash flood watch or warning for each of the six pre-RRT hours, i.e., times between RRT-6 and RRT, we first define the accuracy calculation (ACC) for the flash flooding index model as:

$$ACC_k(\%) = \frac{\sum_{i=1}^N HC_k^i}{N} \times 100 \tag{12}$$

Here,  $N$  is the number of the total flash flood catastrophe.  $HC_k^i$  has a value of 1 if the flash flood watch or warning is issued once or more between pre-RRT hour  $k$  and RRT, irrespective of the time of the actual flash flood occurrence, and has a value of 0 if not. In Equation (12), the indices  $i$  and  $k$  indicate the case number and preceding time (or forecast lead duration time) shown in the first and third columns in Table 4, respectively, such that  $i = 1, 2, \dots, N$ , and  $k = 0, 1, 2, \dots, 6$ .

Table 6 shows that a flash flood warning would have been issued in all documented cases of flash flooding. That is, Table 6 presents the results of calculating the flash flood index for each of the 38 cases of flash flood rescue requests during the six hours preceding the RRT. “Cell No.” in the second column refers to the sequential number of the grid cell counted from the upper-left corner of the grid frame for the survey area. The numbers ‘1’ and ‘2’ indicate flash flood watch (with yellow shading) and warning (with red shading), respectively. For example, RRT- $k$  ( $k = 0, 1, 2, \dots, 6$ ) is the flash flood index calculated based on the gridded rainfall and surface runoff during from pre-RRT hour  $k + 6$  to hour  $k + 1$  and the soil water content at pre-RRT hour  $k$ .

**Table 6.** Flash flood risk index for cells during preceding time based on rescue request time (RRT) corresponding to rescue request locations, where 1 is Flash Flood Watch and 2 is Flash Flood Warning. The first column lists the flash flooding cases shown in Table 1 and the second column lists the cells corresponding to the cases.

Case	Cell No.	Preceding Time						RRT
		−6	−5	−4	−3	−2	−1	0
1	65399	1		1		1	1	2
2	61222			1	1	1	2	2
3	63154							1
4	57716						1	2
5	59952							
6	38852			1				1
7	50725							
8	43991							
9	50725							
10	45909			1	2	1		
11	59036					1	1	
12	51043							
13	51043							
14	68008			2	2	2	2	2
15	55158					1	2	2
16	65089				1	2	2	2
17	60281							
18	55834					1	2	2
19	64440	2	2	2	2	2	1	
20	66386			1	1	1	1	
21	67991			1	1	1	1	
22	67029	1	1	1				
23	71227				1	1	1	
24	60972		1	1	1	2	2	1
25	68296	2	2	2	2	1		
26	51614					2	2	2
27	45911	1	2	2	2	2	2	1
28	55158					1	1	1
29	57081					1	1	
30	57081					1		
31	60659							
32	63798							
33	59335							
34	42042							
35	52301			1	1			
36	60306					2	2	2
37	51043					1	2	1
38	60656				1	2	2	1

For the cases in which the model shows a time lapse between rainfall onset and rescue request time or no indication of watch/warning problems, the changes in the hydrological components are not shown, because of the absolute shortage of precipitation data in the context of reference preceding time and rescue request time. The cases marked in white and dark gray are the cases with cumulative rainfall below 5 mm and below 10 mm, respectively, during the 6 h pre-RRT (Table 6), and cannot be easily improved in the water balance process within the land surface model. This is likely due to QPE error based on AWS precipitation observations. Uncertainty is potentially generated in the gridded precipitation data from interpolating the AWS precipitation data using inverse distance weighting and the possibility of incorrect flash flood rescue request time as a result of personal judgment. However, the lack of stream gauges in mountain areas combined with these factors suggests that the rainfall-related uncertainty is the main source of error, but the rescue request time is correctly forecasted.

We also classify the flash flooding cases into four different scenarios to obtain the accuracy of the statistical FFRI model: Scenario 1 includes all 38 flash flood catastrophes; Scenario 2 includes 36 catastrophes, excluding those without rainfall in the corresponding grid cells (Cases 7 and 8); Scenario 3 includes 34 catastrophes, excluding those with cumulative rainfall during the six pre-RRT hours below 5 mm (Cases 7, 8, 9, and 31); Scenario 4 includes 31 catastrophes, excluding those with cumulative rainfall below 10 mm during the 6 hours pre-RRT (Cases 7, 8, 9, 13, 17, 31, and 33).

Figure 10 shows the results of analyzing the accuracy of the four subsets of catastrophe scenarios. The accuracy is highest when the 6 hours pre-RRT is used, and decreases as the pre-RRT time approaches the RRT; the accuracy from RRT between 6 h and 4 h is 71–87%, which decreases to 42–52% at 0 h. The accuracy of a flash flood watch or warning within 6 h pre-RRT reflects the corrected predicted result irrespective of the actual time of occurrence. Therefore, the accuracy decreases as the preceding time decreases because there are fewer probabilities in the frequency of the rescue request calls.

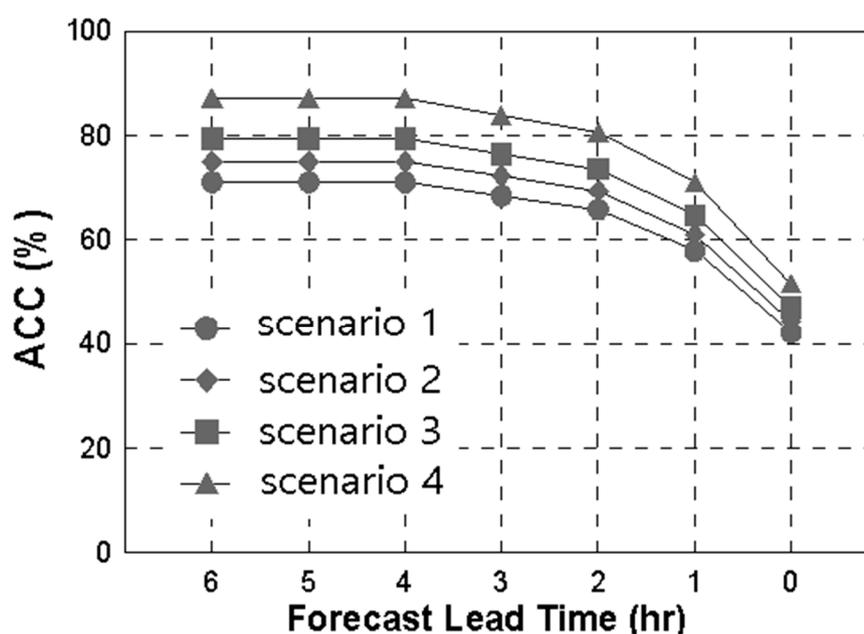


Figure 10. Accuracy of flash flood occurrence with respect to preceding time.

Overall, the results indicate that FFRI with temporal and spatial resolutions of 1 h and 1 km, respectively established for the Han River are more than 70% accurate within the 6 h pre-RRT. Assuming that the accuracies of the QPE and QPF data simulated via radar or numerical models coincide with that of the QPE obtained from interpolating the AWS point rainfall data in the Capital Region, it is expected that potential flash flood catastrophe areas can be predicted with relatively high accuracy. Thus, the FFRI, which was obtained from analyzing the correlations between the gridded hydrological components simulated with the TOPLATS land surface model and flash flood catastrophe, can be used for predicting areas that may be affected by flash flooding.

It is worth noting that experimental results might exhibit case-by-case variability. Therefore, further evaluation of the performance of the statistical FFRI model should be tested to understand the coefficient sensitivity for determining different factors. Future work should also analyze the correlations between simulated results and undocumented mountain valley/stream runoff, i.e., flash flood cases with no rescue requests and no catastrophe reports. In addition, there would be value in applying the FFRI in a forecasting mode, e.g., based on QPF, which has a large uncertainty compared with observations; this could potentially provide more scientific and operational values to the flash flood forecasting community. Nonetheless, this study provides an important basis to begin providing

accurate forecasts to ungauged mountainous areas, and these improvements will be iterations in future work.

## 5. Conclusions

This study analyzed the gridded flash flood detection ability obtained from land surface and statistical FFRI models. We produced gridded hydrological components using the TOPLATS land surface model to predict flash flooding caused by high water and rapids in mountain valleys and small streams. We developed a statistical FFRI model to derive a grid-based flash flood index and evaluated its applicability. After establishing a TOPLATS land surface model in the Han River basin, we simulated high-resolution hydrological components with temporal and spatial resolutions of 1 h and 1 km, respectively. We analyzed the performance of the gridded hydrological components in simulating the land surface conditions leading to flash flood casualties and the accuracy of the grid-based FFRI.

Analysis of the gridded hydrological components of the TOPLATS land surface model for the 38 flash flood catastrophes confirmed that rainfall, surface runoff, and soil water content were adequately simulated in 27 cases (71%), excluding 11 cases with rainfall below 10 mm in the 6 h pre-RRT. The simulated results of rainfall onsets are also compared with RRTs shown in Table 1. The gridded hydrological components reasonably expressed RRT in cases where (1) surface runoff resulted from precipitation exceeding the absorption capacity of soil in an unsaturated state, and (2) rainfall increased the infiltration rate up to soil saturation and continuing precipitation equaled the surface runoff amount (figure not shown). The model also represents RRT prediction with precipitation alone, even in the absence of surface runoff. In this case, the entirety of rainfall infiltrated the soil at conditions of very low initial water table depth, resulting in an increase in water table depth and soil water content but not in surface runoff.

A flash flood watch or a flash flood warning is issued when the flash flood risk index ranges from 40–60 or is greater than 60, respectively. After calculating the flash flood index of each of the 38 cases during the six hours preceding the RRT, it was found that flash flood watch or warning was issued in 27 cases. Accuracy analyses for each pre-RRT hour revealed that the prediction accuracy with respect to the rainfall conditions in the corresponding grid cells was highest between 6 h and 4 h pre-RRT (71–87%) and lowest at 0 h (42–52%). The accuracy decreases as the preceding time decreases, because the error between the RRT and the watch or warning time of the simulated flash flood index results from relatively restricted conditions. A distributed hydrological or land surface model has the advantage of performing spatially refined simulation of hydrological components over a large area. However, this presupposes the availability of accurate input data for soil and land cover, indicating land surface properties and accuracy of gridded rainfall. These problems are expected to be resolved as more accurate GIS data become available and as radar and numerical models improve.

**Author Contributions:** B.-J.L., S.K. conceived and designed the experiments; B.-J.L. performed the experiments; B.-J.L. and S.K. analyzed the data and wrote the paper.

**Funding:** This work was supported by the Research Fund Program of Research Institute for Basic Sciences, Pusan National University, Korea, 2018, Project No. RIBS-PNU-2018-001, and the National Research Foundation of Korea(NRF) Grant funded by the Korean Government(MSIP) (NRF-2017R1A5A1015722).

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

## References

1. Korea Meteorological Administration. *Heavy Rainfall Events Top 10*; KMA Registered Pub. No. 11-136000-0008333-01; KMA (Korea Meteorological Administration): Seoul, Korea, 2011. (In Korean)
2. Reed, S.; Schaake, J.; Zhang, Z. A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *J. Hydrol.* **2007**, *337*, 402–420. [[CrossRef](#)]
3. Wang, Y.; Liu, R.; Guo, L.; Tian, J.; Zhang, X.; Ding, L.; Wang, C.; Shang, Y. Forecasting and Providing Warnings of Flash Floods for Ungauged Mountainous Areas Based on a Distributed Hydrological Model. *Water* **2017**, *9*, 776. [[CrossRef](#)]

4. Sivapalan, M.; Takeuchi, K.; Franks, S.W.; Gupta, V.K.; Karambiri, H.; Lakshmi, V.; Liang, X.; Mendiondo, E.M.; O'connell, P.E.; Oki, T.; et al. IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **2003**, *48*, 857–880. [[CrossRef](#)]
5. Hrachowitz, M.; Savenije, H.H.G.; Blöschl, G.; McDonnell, J.J.; Sivapalan, M.; Pomeroy, J.W.; Arheimer, B.; Blume, T.; Clark, M.P.; Ehret, U.; et al. A decade of Predictions in Ungauged Basins (PUB)—A review. *Hydrol. Sci. J.* **2013**, *58*, 1198–1255. [[CrossRef](#)]
6. Razavi, T.; Coulibaly, P. An evaluation of regionalization and watershed classification schemes for continuous daily streamflow prediction in ungauged watersheds. *Can. Water Resour. J. Rev. Can. Des. Ressour. Hydr.* **2017**, *42*, 2–20. [[CrossRef](#)]
7. Song, X.; Kong, F.; Zhu, Z. Application of Muskingum routing method with variable parameters in ungauged basin. *Water Sci. Eng.* **2011**, *4*, 1–12.
8. Ruiz-Villanueva, V.; Díez-Herrero, A.; Bodoque, J.M.; Ballesteros Cánovas, J.A.; Stoffel, M. Characterization of flash floods in small ungauged mountain basins of Central Spain using an integrated approach. *Catena* **2013**, *110*, 32–43. [[CrossRef](#)]
9. Bathrellos, G.D.; Skilodimou, H.D.; Chousianitis, K.; Youssef, A.M.; Pradhan, B. Suitability estimation for urban development using multi-hazard assessment map. *Sci. Total Environ.* **2017**, *575*, 119–134. [[CrossRef](#)] [[PubMed](#)]
10. Bathrellos, G.D.; Gaki-Papanastassiou, K.; Skilodimou, H.D.; Papanastassiou, D.; Chousianitis, K.G. Potential suitability for urban planning and industry development by using natural hazard maps and geological-geomorphological parameters. *Env. Earth Sci.* **2012**, *66*, 537–548. [[CrossRef](#)]
11. Xia, J.; Falconer, R.A.; Lin, B.; Tan, G. Numerical assessment of flood hazard risk to people and vehicles in flash floods. *Environ. Model. Softw.* **2011**, *26*, 987–998. [[CrossRef](#)]
12. Maidment, D.R. Conceptual Framework for the National Flood Interoperability Experiment. *JAWRA J. Am. Water Resour. Assoc.* **2017**, *53*, 245–257. [[CrossRef](#)]
13. Alfieri, L.; Velasco, D.; Thielen, J. Flash flood detection through a multi-stage probabilistic warning system for heavy precipitation events. *Adv. Geosci.* **2011**, *29*, 69–75. [[CrossRef](#)]
14. Alfieri, L.; Pappenberger, F.; Wetterhall, F. The extreme runoff index for flood early warning in Europe. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1505–1515. [[CrossRef](#)]
15. Balica, S.F.; Popescu, I.; Beevers, L.; Wright, N.G. Parametric and physically based modeling techniques for flood risk and vulnerability assessment: A comparison. *Environ. Model. Softw.* **2013**, *41*, 84–92. [[CrossRef](#)]
16. Davis, R.S. Flash flood forecast and detection methods. In *Severe Convective Storms*; Doswell, C.A., III, Ed.; American Meteorological Society: Boston, MA, USA, 2001; pp. 481–525.
17. Doswell, C.A., III; Brooks, H.E.; Maddox, R.A. Flash flood forecasting: An ingredients-based methodology. *Weather* **1996**, *11*, 560–581. [[CrossRef](#)]
18. Gichamo, Z.G.; Popescu, I.; Jonoski, A.; Solomatine, D.P. River cross section extraction from ASTER global DEM for flood modeling. *Environ. Model. Softw.* **2012**, *31*, 37–46. [[CrossRef](#)]
19. Hartanto, I.M.; Beevers, L.; Popescu, L.; Wright, N.G. Application of a coastal modeling code in fluvial environments. *Environ. Model. Softw.* **2011**, *26*, 1685–1695. [[CrossRef](#)]
20. Maddox, R.A.; Chappell, C.F.; Hoxit, L.R. Synoptic and meso-alpha aspects of flash flood events. *Bull. Am. Meteorol. Soc.* **1979**, *60*, 115–123. [[CrossRef](#)]
21. Van, S.N.; Ronsyn, J.; Willems, P. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. *Environ. Model. Softw.* **2012**, *33*, 92–105.
22. Georakakos, K.P. Real-time flash flood prediction. *J. Geophys. Res.* **1987**, *92*, 9615–9629. [[CrossRef](#)]
23. Carpenter, T.M.; Georakakos, K.P. *GIS based Procedures in Support of Flash Flood Guidance*; IHR Report No. 366; Iowa Institute of Hydraulic Research, The University of Iowa: Iowa City, IA, USA, 1993; pp. 1–27.
24. Torgersen, G.; Rød, J.K.; Kvaal, K.; Bjerkholt, J.T.; Lindholm, O.G. Evaluating Flood Exposure for Properties in Urban Areas Using a Multivariate Modelling Technique. *Water* **2017**, *9*, 318. [[CrossRef](#)]
25. Gigović, L.; Pamučar, D.; Bajić, Z.; Drobnjak, S. Application of GIS-Interval Rough AHP Methodology for Flood Hazard Mapping in Urban Areas. *Water* **2017**, *9*, 360. [[CrossRef](#)]
26. Kauffeldt, A.; Wetterhall, F.; Pappenberger, F.; Salamon, P.; Thielen, J. Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environ. Model. Softw.* **2016**, *75*, 68–76. [[CrossRef](#)]

27. Blöschl, G.; Reszler, C.; Komma, J. A spatially distributed flash flood forecasting model. *Environ. Model. Softw.* **2008**, *23*, 464–478. [[CrossRef](#)]
28. Sweeney, T.L. *Modernized Areal Flash Flood Guidance*; NOAA Technical Memorandum NWS HYDRO44; Hydrologic Research Laboratory, National Weather Service, NOAA: Silver Spring, MD, USA, 1992; p. 21.
29. Mogil, H.M.; Monroe, J.C.; Groper, H.S. NWS's flash flood warning and disaster preparedness programs. *Bull. Am. Meteorol. Soc.* **1978**, *59*, 60–69. [[CrossRef](#)]
30. Skilodimou, H.; Livaditis, G.; Bathrellos, G.; Verikiou-Papaspiridakou, E. Investigating the flooding events of the urban regions of Glyfada and Voula, Attica, Greece: A contribution to Urban Geomorphology. *Geogr. Ann. A* **2003**, *85*, 197–204. [[CrossRef](#)]
31. Bathrellos, G.D.; Karymbalis, E.; Skilodimou, H.D.; Gaki-Papanastassiou, K.; Baltas, E.A. Urban flood hazard assessment in the basin of Athens Metropolitan city, Greece. *Environ. Earth Sci.* **2016**, *75*, 319. [[CrossRef](#)]
32. Hill, C.; Verjee, F.; Barrett, C. *Flash Flood Early Warning System Reference Guide 2010*; University Corporation for Atmospheric Research: Boulder, CO, USA, 2010; ISBN 978-0-615-37421-5.
33. Smith, S.B.; Filiaggi, M.T.; Churma, M.; Roe, J.; Glaudemans, M.; Erb, R.; Xin, L. Flash flood monitoring and prediction in AWIPS Build 5 and beyond. In Proceedings of the 15th Conference on Hydrology, Long Beach, CA, USA, 9–14 January 2000.
34. Bae, D.H.; Jung, I.W.; Chang, H. Long-term trend of precipitation and runoff in Korean river basins. *Hydrol. Process.* **2008**, *22*, 2644–2656. [[CrossRef](#)]
35. Chang, H.; Kwon, W.T. Spatial variations of summer precipitation trends in South Korea, 1973–2005. *Environ. Res. Lett.* **2007**, *2*, 1–9. [[CrossRef](#)]
36. Jung, I.W.; Bae, D.H.; Kim, G. Recent trends of mean and extreme precipitation in Korea. *Int. J. Clim.* **2011**, *31*, 359–370. [[CrossRef](#)]
37. Lee, S.; Kwon, W.T. A variation of summer rainfall in Korea. *J. Korean Geogr. Soc.* **2004**, *39*, 819–832.
38. Bae, D.H.; Kim, J.H. Development of Korea flash flood guidance system: (I) theory and system design. *J. Korean Soc. Civ. Eng.* **2007**, *27*, 237–243.
39. Bae, D.H.; Kim, J.H. Development of Korea flash flood guidance system: (II) Component Development and System Analysis. *J. Korean Soc. Civ. Eng.* **2007**, *27*, 245–254.
40. Famiglietti, J.S.; Wood, E.F. Application of multiscale water and energy balance models on a tallgrass prairie. *Water Resour. Res.* **1994**, *30*, 3061–3078. [[CrossRef](#)]
41. Peters-Lidard, C.D.; Zion, M.S.; Wood, E.F. A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes. *J. Geophys. Res.* **1997**, *102*, 4303–4324. [[CrossRef](#)]
42. Bell, V.A.; Moore, R.J. A grid-based distributed flood forecasting model for use with weather radar data: Part 2. Case studies. *Hydrol. Earth Syst. Sci. Discuss.* **1988**, *2*, 308–318. [[CrossRef](#)]
43. Ji, H.S.; Lee, B.J.; Nam, K.Y.; Lee, C.K.; Jung, H.S. Evaluation of hydrological components simulated by water and energy balance. *J. Korea Water Resour. Assoc.* **2014**, *47*, 25–35. [[CrossRef](#)]
44. Refsgaard, J.C.; Knudsen, J. Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **1996**, *32*, 2189–2202. [[CrossRef](#)]
45. Lee, B.J.; Choi, Y.J. Evaluation of high-resolution hydrologic components based on TOPLATS land surface model. *Atmosphere* **2012**, *22*, 357–365. [[CrossRef](#)]
46. Smith, M.B.; Seo, D.J.; Koren, V.I.; Reed, S.M.; Zhang, Z.; Duan, Q.; Moreda, F.; Cong, S. The distributed model intercomparison project (DMIP): Motivation and experimental design. *J. Hydrol.* **2004**, *298*, 4–26. [[CrossRef](#)]
47. Beven, K.; Quinn, P.; Romanowicz, R.; Rreer, J.; Fisher, J.; Lamb, R. *TOPMODEL and GRIDDATB, A Users Guide to the Distribution Versions (94.03)*; CRES Technical Report TR110/94; Lancaster University: Lancaster, UK, 1994.
48. Bormann, H. Impact of spatial data resolution on simulated catchment water balances and model performance of the multi-scale TOPLATS model. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 165–179. [[CrossRef](#)]
49. Monteith, J.L. Evaporation and environment. In *The State and Movement of Water in Living Organism*; Foggy, G.T., Ed.; Cambridge University Press: Cambridge, UK, 1965; pp. 205–234.
50. Milly, P.C.D. An event-based simulation model of moisture and energy fluxes at a bare soil surface. *Water Resour. Res.* **1986**, *22*, 1680–1692. [[CrossRef](#)]
51. Gardner, W.R. Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Sci.* **1985**, *85*, 228–239. [[CrossRef](#)]

52. Brooks, R.J.; Corey, A.T. *Hydraulic Properties of Porous Media*; Hydrology Paper 3; Colorado State University: Fort Collins, CO, USA, 1964.
53. Sivapalan, M.; Beven, K.; Wood, E.F. On hydrologic similarity. 2. A scaled model of storm runoff production. *Water Resour. Res.* **1987**, *23*, 2266–2278. [[CrossRef](#)]
54. Beven, K.J.; Lamb, R.; Quinn, P.F.; Romanowicz, R.; Freer, J. TOPMODEL. In *Computer Models of Watershed Hydrology*; Singh, V.P., Ed.; Water Resources Publications: Colorado, CO, USA, 1995; pp. 627–668.
55. Rawls, W.J.; Brakensiek, D.L. Prediction of soil water properties for hydrological modeling. In *Watershed Management in the Eighties*; Jones, E., Ward, T.J., Eds.; ASCE: Denver, CO, USA, 1985; pp. 293–299.
56. Famiglietti, J.S.; Wood, E.F. Multiscale modelling of spatially variable water and energy balance processes. *Water Resour. Res.* **1994**, *30*, 3061–3078. [[CrossRef](#)]
57. McDonnell, J.J. Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response. *Hydrol. Process.* **2003**, *17*, 1869–1875. [[CrossRef](#)]
58. Smith, G.E. *Development of a Flash Flood Potential Index Using Physiographic Data Sets within a Geographic Information System*; University of Utah: Salt Lake City, UT, USA, 2010.
59. Davis, R.S. The Flash Flood (FF) Index: Estimating Flash Flood Severity. In *Proceedings of the Symposium on Managing the Extremes, Floods and Droughts, Environmental and Water Resources*; Institute of American Society of Civil Engineers: Roanoke, VA, USA, 2002.
60. Clark, R.A.; Gourley, J.J.; Flamig, Z.L.; Hong, Y.; Clark, E. CONUS-wide evaluation of National Weather Service flash flood guidance products. *Weather* **2014**, *29*, 377–392. [[CrossRef](#)]
61. Schroeder, A.J.; Gourley, J.J.; Hardy, J.; Henderson, J.J.; Parhi, P.; Rahmani, V.; Reed, K.A.; Schumacher, R.S.; Smith, B.K.; Taraldsen, M.J. The development of a flash flood severity index. *J. Hydrol.* **2016**, *541*, 523–532. [[CrossRef](#)]
62. Sung, T.J.; Si, K.J. *Methodology*; Hakjisa: Seoul, Korea, 2014.
63. Warner, R.M. *Applied Statistics: From Bivariate through Multivariate Techniques*; Sage Publications: Los Angeles, CA, USA, 2008.
64. Chatterjee, S.; Hadi, A.S. *Regression Analysis by Example*; John Wiley & Sons: Hoboken, NJ, USA, 2015.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).