Physically Based Susceptibility Assessment of Rainfall-Induced Shallow Landslides Using a Fuzzy Point Estimate Method

Hyuck-Jin Park *, Jung-Yoon Jang and Jung-Hyun Lee

Department of Geoinformation Engineering, Sejong University, Seoul 06005, Korea; jyjang5904@naver.com (J.-Y.J.); jhlee6086@gmail.com (J.-H.L.)

* Correspondence: hjpark@sejong.ac.kr; Tel.: +82-2-3408-3965

Academic Editors: Chaoying Zhao, Zhong Lu, Norman Kerle and Prasad S. Thenkabail

Received: 3 March 2017; Accepted: 13 May 2017; Published: 16 May 2017

Abstract: The physically based model has been widely used in rainfall-induced shallow landslide susceptibility analysis because of its capacity to reproduce the physical processes governing landslide occurrence and a higher predictive capability. However, one of the difficulties in applying the physically based model is that uncertainties arising from spatial variability, measurement errors, and incomplete information apply to the input parameters and analysis procedure. Uncertainties have been recognized as an important cause of mismatch between predicted and observed distributions of landslide occurrence. Therefore, probabilistic analysis has been used to quantify the uncertainties. However, some uncertainties, because of incomplete information, cannot be managed satisfactorily using a probabilistic approach. Fuzzy set theory is applicable in this case. In this study, in order to handle uncertainty propagation through a physical model, fuzzy set theory, coupled with the vertex method and the point estimate method, was adopted for regional landslide susceptibility assessment. The proposed approach was used to evaluate susceptibility to rainfall-induced shallow landslides for a regional study area, and the analysis results were compared with landslide inventory to evaluate the performance of the proposed approach. The AUC values arising from the landslide susceptibility analyses using the proposed approach and probabilistic analysis were 0.734 and 0.736, respectively. However, when the COV values of the input parameters were reduced, the AUC values of the proposed approach and the probabilistic analysis were reduced to 0.722 and 0.688, respectively. It means that the performance of the fuzzy approach is similar to that of probabilistic analysis but is more robust against variation of input parameters. Thus, at catchment scale, the fuzzy approach can respond appropriately to the uncertainties inherent in physically based landslide susceptibility analysis, and is especially advantageous when the amount of quality data is very limited.

Keywords: landslide; uncertainty; fuzzy number; point estimate method; probability of failure; GIS

1. Introduction

Landslides are the movements of earth materials down a slope on one or more failure surfaces [1], and are very common in mountainous and hilly environments in all parts of the world. Landslides are complex slope instability phenomena that are influenced by many external factors (e.g., heavy rainfall, earthquake, and volcanic eruption) and intrinsic factors (geometry of slope, lithology, geological structure, geotechnical properties). That is, landslides can be triggered by a wide range of geological and meteorological processes. In natural hillslopes, one of the most frequent triggering factors is heavy rainfall, which can directly infiltrate the slope surface or can indirectly provide subsurface water from the bedrock [2]. Rainfall-induced slope failures are normally shallow and of small volume, with a depth of failure of less than 3 m, and occur on steep soil slopes of 30°–50° [3,4]. Deep-seated
landslides after rainfall are also occasionally reported but shallow landslides are more typically induced by intense rainfall [5–10]. Since rainfall-induced landslides repeatedly cause fatalities, damage to property, and economic losses around the world, various researchers have sought to predict landslide occurrence and reduce the damage they cause.

In recent years, landslide susceptibility (or hazard) analysis and risk assessment have become a major subject in landslide studies and significant progress has been made. Rapid advances in Earth observation techniques supported by the application of geospatial technologies, such as image analysis and geographic information systems (GIS), have greatly enhanced landslide hazard analysis capabilities over recent years. Remote sensing data has contributed particularly to advances in two aspects of landslide research; landslide detection and identification, and landslide monitoring. Landslide detection and mapping are important in landslide studies because an accurate landslide inventory, including the spatial distribution and geometry of landslides, is essential information. Remote sensing imagery provides invaluable information facilitating the geological interpretation of a landslide and hence better understanding of landslide geometry and kinematics [11]. Most landslide detection and identification studies have used aerial photographs, associated with optical imagery [12–20]. Optical data generally provide better results because of their high spatial resolution and sensor look angle [21]. High spatial resolution optical satellite imagery has also been used to identify the distribution of causative landslide factors such as topography, geology, land use and vegetation cover [22–28]. In addition, Light Detection and Ranging (LiDAR) has been used in landslide studies since LiDAR (or laser scanning) can provide high-resolution point clouds of the topography [29–35]. In addition, according to Singhroy and Molch [36], high resolution synthetic aperture radar (SAR) images can be used to characterize landslides on the basis of an observed close relationship between SAR measurements and debris distribution. On this basis, SAR techniques have been used in landslide detection and mapping [37–42].

As well as mapping landslides post-event, remote sensing is a key tool for measurement of ground movements [11]. Optical remotely sensed imagery has been used in monitoring landslide movements [43–47]. According to Kääb [47], analytical and digital photogrammetry techniques are suitable for monitoring geometric changes in high mountain areas. In recent decades, SAR interferometry, whether satellite or ground based (InSAR and DInSAR), has been a heavily researched technique for landslide monitoring [48]. SAR interferometry uses the phase measurements of two or more SAR images, acquired at different times, of the same scene. The interferogram represents small changes that can be related to topography or surface deformation. Consequently, there is an extensive scientific literature on the use of SAR interferometric techniques for landslide monitoring [36,49–55]. Moreover, satellites are able to provide accurate estimates of precipitation as demonstrated by the Tropical Rainfall Measuring Mission (TRMM) [56]. Remote rainfall measurements can be used to predict rainfall-induced landslides as part of landslide susceptibility analysis [57–60].

With the recent advances in remote sensing and GIS, various landslide susceptibility analysis methods have been proposed in the scientific literature. These can be divided into qualitative (or knowledge-driven) and quantitative approaches, depending on how they treat landslide-inducing factors and models. Generally, qualitative approaches are based entirely on the judgment of experts conducting the susceptibility assessment [61]. Therefore, qualitative approaches are not commonly used in landslide susceptibility assessments over large areas because of the lack of an established physical mechanism by which to predict slope failures [62]. Quantitative landslide susceptibility assessment methods can be divided into two categories: data-driven methods and physically based modeling approaches [63]. In data-driven landslide susceptibility assessment methods, the statistical relationships between the locations of landslides occurred in the past and landslide-inducing factors are evaluated, and then, quantitative predictions are made for landslide free areas with similar conditions. The most significant challenge in the application of data-driven methods is the collection of necessary data regarding landslide distribution and factor maps over large areas [64]. In addition, data-driven methods only consider statistical relationships between landslide occurrences and related
factors based on past observations: they do not ascribe a failure mechanism. Moreover, the statistical models generally ignore the temporal aspects of landslides, and are not able to predict the impact of changes in landslide controlling conditions [64]. The physically based approach, on the other hand, analyzes the mechanical condition of slopes and evaluates their stability using mathematical calculations [5,65–76]. This approach couples distributed hydrological and stability models that can evaluate dynamically the risk of rainfall-induced landslide. The coupled models estimate the spatially distributed safety factor using a physical slope model, such as the infinite slope model, together with the crucial mechanical and hydrological parameters of soils. Recently, physically based models have been widely used because they have a higher predictive capability than other approaches and are the most suitable for quantitatively assessing the influence of individual parameters that contribute to landslide initiation [77]. However, a limitation of using a physically based and spatial distributed model is that it is not always possible to obtain reliable mechanical and hydrological soil parameter data from a broad natural area. An inability to characterize fully the geotechnical and hydrological parameters of slope materials causes inevitable uncertainty and may have a significant impact on model results. Uncertainty in model parameter evaluation has been recognized as an important cause of mismatches between simulated and observed distributions of landslide occurrence [78]. In addition, natural materials, comprising most slopes, have innate variability that is difficult to predict; this variability of geologic material is another major source of uncertainty.

To account for these uncertainties, a probabilistic approach has been investigated and used in slope stability analysis at both site-specific and regional scales [76,79–87]. Probabilistic analysis has been used for representing uncertainty in the failure model and in the material properties. However, the dealing with the uncertainty in the probabilistic analysis procedure is associated with difficulties caused by the lack of information, which is typical in large-area regional landslide studies. According to Haldar and Mahadevan [88] and Giasi and Masi [89], an adequate number of reliable observations are important to implement probabilistic analysis accurately and to estimate uncertainty. However, these data, which are needed for utilization of moment values (e.g., mean and variance) as input parameters and for the evaluation of probability distributions in the probabilistic analysis, are frequently not available, in practice, to a sufficient extent and quality for regional scale analysis. Therefore, probabilistic analysis, in regional landslide studies, is associated with the difficulties because of the lack of information. In addition, some uncertainties connected to measured geotechnical parameters may be nonstochastic [90], being, in conditions of limited knowledge, of cognitive origin [91]. Under such conditions of limited information, it appears reasonable to base estimation on the concept of fuzzy sets [92]. Fuzzy set theory has been found to be effective and suitable for modeling uncertainty in geotechnical parameters when data are insufficient to define a probability distribution fully [93]. Consequently, fuzzy set theory has been used in several site-specific slope stability analyses at hillslope scale [89,92,94–99].

In this study, we propose a fuzzy-set-based analysis method for rainfall-induced shallow landslide susceptibility analysis in a regional area to account for the uncertainty in input parameters. In the landslide susceptibility analysis at catchment scale, the input data used in physically based models are often limited in extent and have imperfect quality. Therefore, the uncertain parameters, used in catchment scale analysis, should be considered as fuzzy numbers. Consequently, in this study, fuzzy set theory was employed to evaluate physically based and spatially distributed landslide susceptibility over a broad area. For that, we developed an algorithm to model the uncertainties with fuzzy numbers and to handle the uncertainty propagation through a physically based model. The proposed algorithm was applied to a practical example to verify its feasibility and validity.
2. Materials and Methods

2.1. Framework for Fuzzy-Set-Based Analysis

2.1.1. Uncertainty Modeling with Fuzzy Numbers

Probabilistic analysis has been a formidable tool for dealing with uncertainty. However, the probabilistic characteristics of a random variable can be described completely if the form of the distribution function and the associated parameters are specified [92]. In addition, probabilistic analysis is carried out on the premise that for an uncertain parameter, the precise mean, precise standard deviation, and an appropriate probability density function can be obtained [100]. However, in many practical cases, the form of the distribution may not be known, and frequently, only the interval between minimum and maximum values can be obtained. Under this condition, if it is judged that not all values in the interval have the same degree of support, an uncertain parameter may be expressed as a fuzzy number [90].

In landslide susceptibility analysis, based on a physically based modeling approach, the strength parameters of soil, such as friction angle and cohesion, are considered as the essential input parameters. However, these parameters cannot be spatially determined over a large area because of limited input data availability. Therefore, uncertainty in soil parameters is best dealt with using fuzzy set theory. Fuzzy set theory has been shown to be effective and suitable for modeling uncertainty in soil parameters when the data are insufficient to define a probability distribution fully [93,97,101–103]. Therefore, a fuzzy number—a subset of a fuzzy set—was used to represent an uncertain soil parameters using fuzzy set theory.

Fuzzy set theory, introduced by Zadeh [104], facilitates analysis of nondiscrete natural processes [91]. While classical set theory defines an object as a member of a set if it has a membership value of 1, or as not a member if it has membership value of 0, the membership of a fuzzy set is expressed on a continuous scale from 1 (full membership) to 0 (full nonmembership) [105]. That is, in fuzzy set theory, a more flexible sense of membership is possible. A fuzzy set is mathematically defined as a set of ordered pairs \([x, \mu_A(x)]\), where an element \(x\) belongs to the set \(A\) to a degree defined by its membership function \(\mu_A(x)\). The membership function ranges from 0 to 1 and is used to characterize the degree of belief that a member \(x\) belongs to \(A\).

A fuzzy number is a special case of a fuzzy set that is normal and convex, which means that the shape of the membership function is single humped and has at least one value whose membership grade is 1. If there is no reason to suggest otherwise, the shape of the membership function may be assumed to be triangular because of its simplicity in formulation and ease of computation [90]. A triangular fuzzy number is characterized by three values: a lower bound, an upper bound, and a modal (or peak) value. The modal value has a membership function of 1, the highest possible set membership for uncertain parameters. As the value of the parameter departs from the modal value, the degree of belief that this value truly represents the soil parameter decreases, and when the value reaches the lower bound (or upper bound), the degree of belief is reduced to zero [98,103]. A triangular fuzzy number is expressed as follows:

\[
\text{TFN} [a, m, b] \tag{1}
\]

where \(a\) is the lower bound, \(m\) is the modal value, and \(b\) is the upper bound (Figure 1).
where \( \mu \) and \( \sigma \) are the mean, standard deviation, and \( \text{COV} \) of \( \phi \) and \( C \), respectively. A higher \( \text{COV} \) (20% for friction angle and 30% for cohesion) was used in setting the parameter [83, 103, 107–110]. The \( \text{COV} \) values reported in previous studies were 10–20% for friction angle and 25–30% for cohesion. Because the number of data in this study was small relative to the size of the study area and it was expected that a higher uncertainty would be involved in the analysis and may be estimated by adopting the published \( \text{COV} \) for a given soil unit, \( k \), ranges from 0.5 to 3, depending on the data availability and the accuracy of the results desired [92]. However, when data are limited, the standard deviation cannot be exactly determined and may be estimated by adopting the published \( \text{COV} \) for a given soil parameter [83, 103, 107–110]. The \( \text{COV} \) values reported in previous studies were 10–20% for friction angle and 25–30% for cohesion. Because the number of data in this study was small relative to the size of the study area and it was expected that a higher uncertainty would be involved in the analysis procedure, higher values of \( \text{COV} \) (20% for friction angle and 30% for cohesion) were used in setting the standard deviation values for the uncertain parameters. In addition, the difference between the upper and lower bounds was considered as 4 \( \sigma \) by taking \( \pm 2\sigma \) from the mean. That is, the fuzzy numbers for friction angle and cohesion are: \[ x_{ub} = \mu_x + k\sigma_x \] \[ x_{lb} = \mu_x - k\sigma_x \] where \( \mu_x \) is the mean value of the uncertain parameter, and \( \sigma_x \) is the standard deviation. The sigma unit, \( k \), ranges from 0.5 to 3, depending on the data availability and the accuracy of the results desired [92]. However, when data are limited, the standard deviation cannot be exactly determined and may be estimated by adopting the published \( \text{COV} \) for a given soil parameter [83, 103, 107–110]. The \( \text{COV} \) values reported in previous studies were 10–20% for friction angle and 25–30% for cohesion. Because the number of data in this study was small relative to the size of the study area and it was expected that a higher uncertainty would be involved in the analysis procedure, higher values of \( \text{COV} \) (20% for friction angle and 30% for cohesion) were used in setting the standard deviation values for the uncertain parameters. In addition, the difference between the upper and lower bounds was considered as 4 \( \sigma \) by taking \( \pm 2\sigma \) from the mean. That is, the fuzzy numbers for friction angle and cohesion are: \[ x_{ub} = TFN\left[ \mu_{\phi} - 2\sigma_{\phi}, \mu_{\phi}, \mu_{\phi} + 2\sigma_{\phi} \right] = TFN\left[ \mu_{\phi} - 2(\mu_{\phi} \times \text{COV}_{\phi}), \mu_{\phi}, \mu_{\phi} + 2(\mu_{\phi} \times \text{COV}_{\phi}) \right] \] \[ x_{lb} = TFN\left[ \mu_{C} - 2\sigma_{C}, \mu_{C}, \mu_{C} + 2\sigma_{C} \right] = TFN\left[ \mu_{C} - 2(\mu_{C} \times \text{COV}_{C}), \mu_{C}, \mu_{C} + 2(\mu_{C} \times \text{COV}_{C}) \right] \] where \( \mu_{\phi}, \sigma_{\phi} \) and \( \text{COV}_{\phi} \) are the mean, standard deviation, and \( \text{COV} \) of the friction angle, and \( \mu_{C}, \sigma_{C} \) and \( \text{COV}_{C} \) are the mean, standard deviation, and \( \text{COV} \) of the cohesion.

2.1.2. Fuzzy Point Estimate Method

After the fuzzy numbers are defined for each uncertain parameter, the vertex method [111] is used to evaluate the performance function (in this case, the physically based slope model) and to analyze the propagation of fuzzy numbers through the slope model. According to the extension principle of
Zadeh [104], algebraic operation on real numbers can be extended to fuzzy numbers for evaluation of the performance function. However, the implementation of the computation is not trivial, as the algebraic solution procedure may be complicated and may not guarantee a result. Therefore, the vertex method, one of the simplest solution procedures, has been used in several studies [89,92,98,99,103].

The vertex method is based on the \( \alpha \)-cut concept of fuzzy numbers and involves an interval analysis. The \( \alpha \)-cut of a fuzzy set \( A \) is the crisp set comprised of all the elements \( x \) of a universe of discourse \( X \) for which the membership function of \( A \) is greater than or equal to \( \alpha \). The basic idea of the \( \alpha \)-cut concept is the discretization of a fuzzy number into a group of \( \alpha \)-cut intervals. For each uncertain parameter, an \( \alpha \)-cut of its fuzzy number will give an interval having two points; i.e., the upper and lower bound values for a particular \( \alpha \)-cut. At each \( \alpha \)-cut level, the intervals for each input variable are obtained, and combinations of vertices can be formed. Therefore, if we have a performance function with \( N \) fuzzy input variables whose membership functions are assigned, the number of vertex combinations will be \( 2^N \). Then, at each \( \alpha \)-cut level, by replacing fuzzy numbers in the performance function with the vertex combinations, the results of the performance function are obtained by fuzzy computation. In this study, the point estimate method (PEM) proposed by Rosenblueth [112] and Harr [108] was used to compute the performance function using fuzzy input parameters. Conceptually, the PEM uses two point estimates \((x_+, x_-)\) of each uncertain parameter and evaluates the performance function with \( 2^N \) combinations of point estimates. Therefore, because the vertex method can evaluate two vertices, which are the upper and lower bounds \((x_+^\alpha, x_-^\alpha)\) of an \( \alpha \)-cut interval (Figure 1), PEM can use the upper and lower bounds of an \( \alpha \)-cut interval as two point estimates and thus evaluate the performance function. Consequently, the PEM coupled with the vertex method can obtain outputs for the performance function, which are factors of safety (FS) according to the physically based model. Because two uncertain parameters, friction angle and cohesion, were considered as fuzzy numbers in this study, two point estimates for friction angle \((\phi_+^\alpha, \phi_-^\alpha)\) and cohesion \((c_+^\alpha, c_-^\alpha)\) were evaluated at each \( \alpha \)-cut level, and therefore, four vertex combinations were obtained at each \( \alpha \)-cut level. Subsequently, four outputs of the performance function \((FS_{+\alpha}^\alpha, FS_{-\alpha}^\alpha, FS_{+\alpha}^{\alpha\alpha}, FS_{-\alpha}^{\alpha\alpha})\) were obtained using four vertex combinations. That is:

\[
FS_{+\alpha}^\alpha = FS(\phi_+^\alpha, c_+^\alpha), \quad FS_{-\alpha}^\alpha = FS(\phi_+^\alpha, c_-^\alpha), \quad FS_{+\alpha}^{\alpha\alpha} = FS(\phi_-^\alpha, c_+^\alpha), \quad FS_{-\alpha}^{\alpha\alpha} = FS(\phi_-^\alpha, c_-^\alpha)
\]

(4)

Then, the mean and standard deviation of the safety factor (FS), which is the output of the performance function, can be obtained as follows:

\[
E[FS] = \frac{\sum_{i=1}^{M} (p_{++} \times (FS_{+\alpha}^\alpha) + p_{+-} \times (FS_{+\alpha}^{\alpha\alpha}) + p_{-+} \times (FS_{-\alpha}^\alpha) + p_{-\alpha} \times (FS_{-\alpha}^{\alpha\alpha}))}{M}
\]

\[
E[FS^2] = \frac{\sum_{i=1}^{M} (p_{++} \times (FS_{+\alpha}^\alpha)^2 + p_{+-} \times (FS_{+\alpha}^{\alpha\alpha})^2 + p_{-+} \times (FS_{-\alpha}^\alpha)^2 + p_{-\alpha} \times (FS_{-\alpha}^{\alpha\alpha})^2)}{M}
\]

\[
\sigma_{FS} = \sqrt{E[FS^2] - (E[FS])^2}
\]

\[
p_{++} = p_{-\alpha} = \frac{1+\rho}{2}
\]

\[
p_{+-} = p_{-\alpha} = \frac{1-\rho}{2}
\]

(5)

where \( \rho \) is the coefficient of correlation between the uncertain parameters, and \( M \) is the number of \( \alpha \)-cut levels. In this study, nine different \( \alpha \)-cut levels \((0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)\) were used. As explained previously, in this approach, all \( 2^N \) vertex combinations at each \( \alpha \)-cut level are used in the evaluation of the performance function. However, in previous vertex approaches [89,92,98,99,103], the outputs of performance functions were determined by taking only the minimum and maximum values obtained from \( 2^N \) vertex combinations at each \( \alpha \)-cut level. This study therefore differs, in this way, from previous fuzzy-number-based approaches.

From the two moments (mean and standard deviation) of the performance function, the probability of failure can be obtained. In order to evaluate the probability of failure from PEM, the reliability index, \( \beta \), should be evaluated. The reliability index refers to the distance of the mean factor
of safety from its critical value (i.e., $FS = 1$) in units of standard deviation. Therefore, the reliability index is evaluated with the following equation:

$$\beta = \frac{E[FS] - 1}{\sigma_{FS}} \quad (6)$$

In order to obtain the probability of failure from the reliability index, the factor of safety is assumed to be normally distributed, and the reliability index is, therefore, a standard normal variate [107]. The standardized normal distribution, with zero mean and unit standard deviation, is widely tabulated, and usually, the tabulation expresses the integral, $\Phi$, of the standardized normal distribution between $-\infty$ and a positive value. Therefore, the probability of failure ($P_f$) is evaluated as follows:

$$P_f = 1 - \Phi(\beta) = \Phi(-\beta) \quad (7)$$

2.2. Physically Based Slope Model

In physically based analysis, a physical slope model is used to determine the slope stability by analyzing the force applied to the slope. The infinite slope model, a useful simplification for rainfall-induced shallow landslide [9,66,83,113], was used in this study. This model is appropriate for Korean landslides because the sliding surfaces are commonly located at shallow depth [83,114–116]. In addition, this study adopted a transient rainfall infiltration model, which considers spatiotemporal changes in pore water pressure caused by rainfall infiltration, coupled with the infinite slope model, to analyze landslide susceptibility. The infiltration model estimates transient pore pressure changes by considering changes in rain infiltration over time. In this study, we adopted the hydrogeological model implemented in the transient rainfall infiltration and grid-based regional slope stability (TRIGRS) [117], which is designed for modeling the potential occurrence of shallow landslides by incorporating transient pressure responses to rainfall and downward infiltration [118].

The infiltration models in TRIGRS are based on Iverson’s solution [119] of the Richards’ equation for an infinitely deep impervious basal boundary and the surface condition of constant flux for a specified time, and zero flux thereafter [120]. Iverson [119] provides a theoretical framework for understanding how hydrological processes influence the location, timing, and rate of landslide occurrence, based on a solution to the boundary problem posed by Richards’ equation. This solution evaluates the effects of transient rainfall by modeling the pore water pressure as both a steady component and a transient component. Baum [117] generalized Iverson’s original infiltration model for cases of variable rainfall intensity and duration to develop the TRIGRS software package.

2.3. Study Area and Database Construction

The Inje area, which is located in the middle of Gangwon Province, Republic of Korea, was selected as the study area to assess landslide susceptibility using the proposed analysis method. On 15–16 July 2006, as a result of Typhoon Ewiniar and heavy rainstorm, this area experienced heavy rainfall, 332.5 mm, and approximately 800 landslides were reported. The event caused the death of 17 people, with 12 unaccounted for [121], and most residential settlements were swept away by landslides.

The study area has an approximate area of 31.6 km$^2$ and is located between longitudes 128°11′44.81″ and 128°18′8.99″ and latitudes 38°3′3.93″ and 38°15′58.55″. The altitude ranges from 215 m to 1220 m, with an average altitude of 660 m. Within mountainous terrain, the Deoksancheon stream valley runs from southeast to northwest through the center of the area. The many mountainous parts of the study area have steep slopes (>30°) and are susceptible to slope failure [122]. Geologically, the area is composed mainly of Mesozoic Inje granite and partly Precambrian biotite gneiss with limited deposits of alluvium at the lower end of the stream (Figure 2). The soils in this area are weathered granite soils, exhibiting a wide range of conditions depending on the degree of weathering. When rainfall infiltrates the ground, the granite soils become partially or fully saturated. This increases
the fragility of the soil, resulting in a significant loss of shear strength. Rainfall-induced landslides in this area of Korea are usually shallow failures in such weathered granite slopes. They are mainly triggered by rainfall infiltration, which induces decreased matric suction and shear strength [116].

In order to make a reliable map that predicts landslide occurrence in a certain area, it is crucial to have insight into the spatial distribution or frequency of landslides. Therefore, a landslide inventory is one of the most important factors in landslide susceptibility analysis. This is also because the accuracy of the landslide prediction models can be evaluated using a landslide inventory map. A landslide inventory map can be constructed by collecting historical information of landslide events or by consulting satellite imagery and aerial photographs coupled with field surveys. In this study, landslide locations were mapped using visual interpretation of aerial photographs and then confirmed by field surveys. Visual interpretation is a widely used method and results in landslide inventories of high resolution [77,123]. After the landslide event in 2006, digital aerial imagery from the UltraCam-X sensor, with a ground resolution of 0.5 m, was obtained from Samah Aerial Survey Co. Ltd., Goyang, Korea, in October 2010. These were compared with analog aerial photos taken, using the RC-20 camera, prior to the event in August 2005. Aerial photographs were projected into the International Terrestrial Reference Frame (ITRF) 2000 datum and Transverse Mercator (TM) coordinate system to match the other thematic maps. Based on the comparison of aerial imagery, landslide occurrence locations were
digitized and constructed as an initial point layer inventory map. The reliability of the landslide inventory map was then enhanced by a field survey of some of the landslide locations. Through this process, a total of 877 landslides were located, and a final GIS-based landslide inventory map was constructed for further study (Figure 2).

Geomorphological characteristics, such as slope angles and elevation, were extracted from 1:5000 digital topographic maps, provided by the National Geographic Information Institute of Korea. Contour vectors, which included altitude information, were extracted, and a triangulated irregular network (TIN) was generated from this layer and converted into a grid raster layer as a digital elevation model (DEM) with a 10 m resolution. The DEM was processed to create thematic maps of geomorphological characteristics, such as slope angle and elevation (Figure 3a,b). Furthermore, the soil thickness in the study area was acquired from 1:25,000 scale digital soil maps produced by the National Institute of Agricultural Science.

In order to evaluate the landslide susceptibility using a physically based model, the geotechnical input parameters such as cohesion and friction angle for the soils should be obtained. In this study, the requisite input parameters were obtained from laboratory tests of soil samples collected from the study area. The soil samples were collected from the landslide occurrence locations in each geological unit. For each sampling location, six to nine soil samples were obtained for the laboratory tests, and direct shear tests were performed to obtain shear strength parameters for each soil type. The test results were connected to the sample locations in the point map layer and used to construct

![Figure 3. The distributions of (a) slope angle and (b) elevation.](image-url)
thematic maps of the input parameters (Figure 4a,b). The mean values of cohesion and friction angle were 7.6 kN/m² and 31.5°. However, as mentioned in previous works [62,124–127], cohesion and friction angle of slope materials are considered to be the major sources of uncertainty because of spatial variability and limited sampling. Thus, cohesion and friction angle were considered as uncertain variables. That is, the cohesion and friction angle were treated as fuzzy numbers in the proposed analysis model. In addition, as in many analyses, the uncertain variables were assumed to be independent [72,76,87,109,126,128,129].

The unit weights (Figure 4c) of soil were also obtained from laboratory tests, and the mean value in this area was 15.4 kN/m³. In addition, 20 constant-head permeability tests were carried out in the laboratory to obtain the hydraulic conductivity (Figure 4d). The mean value in this area was 0.0046 m/h. The soil thickness was evaluated from the depth to bedrock and used as soil depth in the infinite slope model. Applying the Z-model [130], soil thickness (h) was estimated, according to altitude, by Equation (8).

\[ h_i = h_{\text{max}} - \left( \frac{z_i - z_{\text{min}}}{z_{\text{max}} - z_{\text{min}}} \right) (h_{\text{max}} - h_{\text{min}}) \]  

(8)

Here, \( h_{\text{max}} \) and \( h_{\text{min}} \) refer to the maximum and minimum soil thickness values, respectively, whereas \( z_{\text{max}} \) and \( z_{\text{min}} \) are the maximum and minimum ground elevations, respectively. In this way, using the thematic map of elevation, soil thickness in the area was also thematically mapped (Figure 4e).

Figure 4. Cont.
Figure 4. The distributions of (a) friction angle, (b) cohesion, (c) unit weight, (d) hydraulic conductivity, and (e) soil thickness.

Precipitation input, such as rainfall intensity, for the physically based model is one of the most important parameters to be obtained. Rainfall intensity values were obtained from the Inje automatic weather system (AWS), which is the closest and most representative rain gauge in the study area. The rainfall intensity for the study area was obtained from the hourly rainfall records during the rainstorm on 15–16 July 2006. On July 15, the total rainfall from the time when the rain started (3:00 a.m.) to the time when the landslides occurred (11:00 a.m.) was 153.5 mm, and the average rainfall intensity was 19.19 mm/h [131].

2.3.1. The Analysis Procedure

In this study, the regional slope stability analysis, based on PEM coupled with the fuzzy input parameters, was conducted using the following procedure. First, uncertainties of the geomechanical parameters, such as cohesion and friction angle, were quantified as triangular fuzzy numbers.
Then, each of the input fuzzy numbers was discretized into a set of $\alpha$-cut intervals. In this study, nine $\alpha$-cut levels were considered—i.e., 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9—in order to represent the possible variability in the input parameters. At each selected $\alpha_i$-cut level, the upper bound ($x_{\alpha_i}^+$) and lower bound ($x_{\alpha_i}^-$) of the intervals for each input fuzzy number were obtained, and these were considered as point estimates in the PEM procedure. The combinations of point estimates (i.e., $(\phi_{\alpha_i}^+, c_{\alpha_i}^+)$, $(\phi_{\alpha_i}^+, c_{\alpha_i}^-)$, $(\phi_{\alpha_i}^-, c_{\alpha_i}^+)$, $(\phi_{\alpha_i}^-, c_{\alpha_i}^-)$) were formed at each $\alpha_i$-cut level, and the combinations were used as input values to solve the physical slope model and to obtain FS values at each $\alpha_i$-cut level. Because there were $2^N$ combinations, there would be $2^N$ FS values for each $\alpha_i$-cut level. Then using PEM in Equation (5), the expected value and standard deviation of FS were obtained, and subsequently, the reliability index, $\beta$, and the probability of failure were evaluated. Using a GIS platform, this procedure was applied for spatially distributed slope stability analysis across the regional area. In order to apply the landslide susceptibility analysis in the GIS environment, all input data were first distributed in a grid-based framework with information for each cell assigned (see Figures 3 and 4). For each layer used in the model, a thematic map was generated using a $10\times10$ m pixel/cell size and raster format. The raster format is widely used in spatial analysis and modeling because of its simple data structure and easy integration with other data.

In addition, because the study area was composed of 316,479 pixels and the fuzzy PEM had to be implemented in all the pixels, the calculation of FS and probability of failure required much time and effort. Therefore, to carry out the analysis procedure efficiently, we developed a fuzzy PEM module using Mathworks’ MATLAB platform. In order to evaluate FS and probability of failure, using a grid-based module developed in the MATLAB environment, the grid-format input data were converted to ASCII-format data files. Subsequently, the fuzzy PEM analysis for each pixel was carried out using the developed modules. The results of the analysis of the modules were then reconstructed into ASCII format and transferred to grid-format files using GIS software. The analysis results were presented as maps using the GIS software.

2.3.2. Evaluation of Model Performance

In landslide susceptibility analysis, model validation is essential for understanding model performance. Validation refers to comparison of the model predictions with real-world data to assess the model’s accuracy and predictive power [132]. A receiver operating characteristics (ROC) graph is a commonly used technique for visualizing agreement between the susceptibility analysis and the inventoried spatial distribution of landslides. ROC graphs are developed from a confusion matrix that distinguishes between two classes of events (true and modeled) [133,134] (Figure 5). Two metrics obtained from confusion matrix are used to plot the ROC graph, which is a two-dimensional graph in which the true-positive rate (TPR) is plotted on the y-axis, and the false-positive rate (FPR) is plotted on the x-axis. The TPR is the ratio of the number of true positives to the total number of positives, and it quantifies the accuracy of the predicted results. The FPR is the ratio of the number of false positives to the total number of negatives, and it quantifies the error rate. Model results that plot toward the upper left of the ROC graph are generally considered to be superior [66]. The origin of the graph (0, 0) represents a model result with maximum underestimate. In contrast, the upper right of the graph (1, 1) represents the maximum overestimate, which means that the entire area is predicted to be unstable. A perfect prediction would be located at the upper left (0, 1). The closer the results of the analysis are to the upper left corner of the ROC graph (0, 1), the better the performance of the model is. Because an ROC curve is a two-dimensional depiction of individual model performance, a single derived scalar value is needed to compare performance between models. A common method is to calculate the area under the curve (AUC). The larger the AUC value, the better the performance of the model [66,133].
In this study, a physically based model was used to evaluate rainfall-induced shallow landslide susceptibility across a region. A fuzzy PEM was adopted for this landslide susceptibility analysis to accommodate uncertainties in the input parameters of the physical model. A transient infiltration model can estimate transient pore pressure changes and evaluate slope stability arising from changes in rain infiltration over time. The physically based model was therefore combined with a transient infiltration hydrogeological model to produce landslide susceptibility maps. In the study area, landslide susceptibility was determined 8 h after the beginning of the rain to match the time interval between the time when the rain started (3:00 a.m.) and the time when the landslides occurred (11:00 a.m.). The results of this landslide analysis were compared with the inventoried locations of actual landslides to verify the accuracy of the proposed approach. In slope stability analysis based on a physically based model, correct determination of the level of safety for a slope is the objective. In deterministic analysis, a value of \( FS = 1.0 \) is considered to be the criterion that the safety of the slope can be evaluated. Therefore, \( FS \geq 1.0 \) indicates that the slope is safe, and a value of \( FS < 1.0 \) implies an unstable slope. However, there is no single criterion for stability in probabilistic analysis. Because many uncertainties influence the outcomes of the analysis, we need to estimate the probability that the slope will be unstable in order to determine the level of slope safety [135]. Although the criteria used differ among researchers, many studies have chosen a failure probability of 10% as a threshold to indicate the stability of natural slopes [72,83,135–138]. Thus, areas with a probability of failure greater than 10% were considered to be susceptible to landslides in this study, and this figure was used to evaluate the confusion matrix and determine the AUC.

Figure 6 shows the spatial distribution map for the probability of failure calculated using the fuzzy PEM and the physically based model. In all, 747 out of total 877 locations at which actual landslides occurred were predicted to be unstable. That is, 85.2% of actual landslides (or the mapped landslides in the inventory map) were evaluated as unstable, meaning that the TPR, which is the ratio of the number of correctly predicted landslide grids (true positives) to the total number of landslide occurrence grids (positives), was 0.852. In addition, 38.3% of nonlandslide grids were predicted as unstable, meaning that the FPR was 0.383 (Table 1). Thus, the evaluated AUC from the ROC graph in Figure 7 was 0.734. Because an AUC greater than 0.7 suggests considerable accuracy [77], the proposed approach had a good predictive performance.

To compare the results of fuzzy PEM with the results of a probabilistic approach, a Monte Carlo simulation was performed as the probabilistic analysis (Figure 8). Monte Carlo simulation is considered to be the most complete probabilistic analysis method because all random variables and the probability of failure that results from the reliability analysis are represented by their probability density functions through repeated calculations [83]. Therefore, this study utilized Monte Carlo simulations to obtain the probability of failure by considering the strength parameters (cohesion and effective friction angle) as random variables. The random input parameters used in Monte Carlo simulation were considered as normally distributed variable, as in much previous research [72,83,97,127,129,139]. The same values of mean and standard deviation as used in the fuzzy analysis were applied to the random
input parameters. TPR and FPR values were evaluated as 0.885 and 0.413, respectively (Table 1). Then the AUC was evaluated as 0.736, which is similar to the performance of the fuzzy PEM (0.734). This means that both the fuzzy PEM and the probabilistic analysis show good predictive performance. In addition, fuzzy PEM demonstrates accuracy similar to Monte Carlo simulation, when using only limited information on the input parameters, without the tedious repeated calculations required in Monte Carlo simulation.

![Figure 6](image1.png)

**Figure 6.** Map showing slope failure probability predicted using fuzzy PEM.

![Figure 7](image2.png)

**Figure 7.** ROC graph comparing the analysis results. A: fuzzy PEM, B: probabilistic analysis (Monte Carlo simulation), C: deterministic analysis, D: fuzzy PEM with reduced COV, E: probabilistic analysis with reduced COV, F: fuzzy PEM with ±1 σ, G: fuzzy PEM with ±3 σ.

![Figure 8](image3.png)

**Figure 8.** Map showing slope failure probability predicted using Monte Carlo simulation.

Furthermore, a deterministic analysis of the factor of safety was conducted for comparison with the fuzzy approach and the probabilistic analysis. In the deterministic analysis, the same infinite-slope
model with transient infiltration model was used, and the mean values of the random variables (cohesion and friction angle) were used as uniform values across the study area. Figure 9 shows the spatial distribution of the calculated factor of safety. Of the inventoried landslides, 47.1% (413 landslides out of total 877) were evaluated as unstable, and 16.3% of nonlandslide grids were predicted to be unstable. Therefore, TPR and FPR values were 0.471 and 0.163, respectively, and the AUC was 0.654 (Figure 7). This value is below 0.7 and considerably smaller than the AUCs of the fuzzy PEM (0.734) and the probabilistic analysis (0.736). Therefore, the fuzzy PEM and Monte Carlo simulation showed superior performance than the deterministic analysis. This result shows that uncertainties in the input parameters affect the analysis results, and therefore, appropriate consideration of uncertainties in input parameters is import in application of a physically based modeling approach.

![Figure 9. Map showing the factor of safety predicted using the deterministic analysis.](image)

However, one of the limitations of the probabilistic analysis is that the results of the analysis are often affected by the statistical characterization of the input parameters. In order to check the effect of varying input parameter statistics, in the fuzzy approach and the probabilistic approach, the COV values for input parameters were changed. The landslide susceptibility analyses were then carried out using the changed COV values. Specifically, the COVs of the internal friction angle and cohesion were reduced to 10% and 20% (from 20% to 30%), respectively. Consequently, the standard deviations of friction angle and cohesion were reduced from 6.3 to 3.15 and from 2.3 to 1.5, respectively. The landslide susceptibility analyses using the proposed approach and Monte Carlo simulation were carried out with reduced COV values as input parameters (Figure 10). As listed in Table 1, the TPR and FPR values for the proposed fuzzy PEM approach using 10% COV of friction angle and 20% COV of cohesion were 0.772 and 0.328, respectively. Then AUC was evaluated as 0.722, which is somewhat less than the AUC (0.734) evaluated using 20% COV of friction angle and 30% COV of cohesion. However, the TPR and FPR values for Monte Carlo simulation using 10% COV of friction angle and 20% COV of cohesion were 0.594 and 0.219, respectively, and AUC was evaluated as 0.688. The AUC value was reduced from 0.736 (the AUC evaluated using 20% COV of friction angle and 30% COV of cohesion) to 0.688 as the COVs of uncertain variables were reduced. Therefore, the variation of AUC in the fuzzy PEM is smaller than the variation of Monte Carlo simulation. Consequently, the fuzzy PEM seems to be more robust against the variation of input parameters. The term ‘robustness’ characterizes the change in the accuracy of the model performance because of errors in the input parameters [140]. In addition, the fuzzy numbers were changed from mean ±2 σ to mean ±1 σ and mean ±3 σ with 20% COV of friction angle and 30% COV of cohesion to check the robustness of the fuzzy PEM model performance (Figure 11). The results showed that TPR and FPR values using the fuzzy numbers of mean ±1 σ were evaluated as 0.708 and 0.281, respectively, and the evaluated AUC was 0.714. If the fuzzy numbers are changed to mean ±3 σ, the TPR and FPR values were evaluated as 0.930 and 0.509, respectively and the AUC was 0.711. Therefore, fuzzy PEM model performance was more robust than probabilistic analysis.
Figure 10. Maps showing slope failure probability predicted using (a) fuzzy PEM with reduced COV and (b) Monte Carlo simulation with reduced COV.

Figure 11. Maps showing slope failure probability predicted using fuzzy PEM with (a) fuzzy numbers of mean ±1\(\sigma\) and (b) fuzzy numbers of mean ±3\(\sigma\).
Table 1. Comparison of the model performance.

<table>
<thead>
<tr>
<th>Model</th>
<th>True Positive Rate (TPR)</th>
<th>False Positive Rate (FPR)</th>
<th>Area Under the Curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy PEM</td>
<td>0.852</td>
<td>0.383</td>
<td>0.734</td>
</tr>
<tr>
<td>Probabilistic analysis</td>
<td>0.885</td>
<td>0.413</td>
<td>0.736</td>
</tr>
<tr>
<td>Deterministic analysis with the transient model</td>
<td>0.471</td>
<td>0.163</td>
<td>0.654</td>
</tr>
<tr>
<td>Fuzzy PEM with reduced COV</td>
<td>0.772</td>
<td>0.328</td>
<td>0.722</td>
</tr>
<tr>
<td>Probabilistic analysis with reduced COV</td>
<td>0.594</td>
<td>0.219</td>
<td>0.688</td>
</tr>
<tr>
<td>Fuzzy PEM with ±1σ</td>
<td>0.708</td>
<td>0.281</td>
<td>0.714</td>
</tr>
<tr>
<td>Fuzzy PEM with ±3σ</td>
<td>0.930</td>
<td>0.509</td>
<td>0.711</td>
</tr>
</tbody>
</table>

4. Conclusions

The occurrence of landslide is governed by complex interrelationships between factors, of which some cannot be determined in detail and others can be determined only with a large degree of uncertainty. Therefore, recognition and assessment of uncertainties is of paramount importance in landslide susceptibility analysis based on a physical model. Probabilistic analysis has provided a quantitative way to account for the uncertainties in input parameters. However, the number of reliable observations needed to estimate parameter uncertainty for probabilistic analysis is more than can be reasonably assessed in many practical conditions, in which the amount of data is frequently limited, and the distribution type of the uncertain variable may not be known. This situation makes the application of the probabilistic approach difficult. Furthermore, some uncertainties relating to measured geotechnical parameters may, in fact, be nonstochastic but rather cognitive, arising from incomplete knowledge. Under such conditions of limited information, it appears to be reasonable to adopt fuzzy set theory because a fuzzy-set-based approach only requires mean, minimum, and maximum values of the uncertain parameters. Therefore, in this study, fuzzy set theory was adopted for a regional landslide susceptibility analysis based on a physically based model. The input parameters were considered as fuzzy numbers, and the vertex method and point estimate method were used to handle uncertainty propagation through the physical model. The AUC values arising from the landslide susceptibility analyses using the proposed approach and probabilistic analysis were 0.734 and 0.736, respectively. However, when the COV values of the input parameters were reduced, the AUC values of the proposed approach and the probabilistic analysis were reduced to 0.722 and 0.688, respectively. In other words, the fuzzy approach performed similarly to the probabilistic analysis on the same data but was more robust against variations of input parameters. Moreover, when the fuzzy numbers were changed from mean ±2σ to mean ±1σ and mean ±3σ, the proposed approach showed more robust performance than the probabilistic analysis. Therefore, the fuzzy approach can be an alternative to probabilistic analysis, especially when the amount of quality data is very limited. In addition, previous fuzzy approach studies have been used only for site-specific slope stability analysis, not for regional slope stability analysis as demonstrated here. Therefore, this study is an important step forward, showing that a fuzzy approach can successfully be applied to physically based landslide susceptibility analysis at catchment scale, even if the available information about input parameters is very limited.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) grant funded by Korea government (MSIP) (No. NRF-2016R1A2B4008963).

Author Contributions: Hyuck-Jin Park conceived and designed the experiments, and wrote the paper; Jung-Yoon Jang and Jung-Hyun Lee performed the experiments and analyzed the data.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.
References


46. Hervas, J.; Barredo, J.I.; Rosin, P.L.; Pasuto, A.; Mantovani, F.; Siliano, S. Monitoring landslides from optical remotely sensed imagery: The case history of Tessina landslide, Italy. *Geomorphology* 2003, 54, 63–75. [CrossRef]


52. Prati, C.; Ferretti, A.; Perissin, D. Recent advances on surface ground deformation measurement by means of repeated space-borne SAR observations. *J. Geodyn.* 2010, 49, 161–170. [CrossRef]


61. Van Westen, C.J.; Seijmonsbergen, A.C.; Mantovani, F. Comparing landslide hazard maps. *Nat. Hazards* 1999, 20, 137–158. [CrossRef]


88. Shou, K.J.; Chen, Y.L.; Liu, H. Hazard analysis of Li-Shan landslide in Taiwan. *Geomorphology* 2009, 103, 143–153. [CrossRef]


126. Shou, K.J.; Chen, Y.L. Spatial risk analysis of Li-shan landslide in Taiwan. Eng. Geol. 2005, 80, 199–213. [CrossRef]


130. Saulnier, G.M.; Beven, K.J.; Obled, C. Including spatially variable effective soil depths in TOPMODEL. J. Hydrol. 1997, 202, 158–172. [CrossRef]


