Mining of Mineral Deposits
Volume 18 (2024), Issue 4, 26-33

https://doi.org/10.33271/mining18.04.026

Enhancing slope failure forecasting model by implementing Archimedean copula to model the error-term

Barlian Dwinagara [1*], Vega Vergiagara [1], Ömer F. Uğurlu [2], Singgih Saptono [1], Rania Salsabila [3], Aldin Ardian [1]

[1] *Universitas Pembangunan Nasional "Veteran" Yogyakarta, Yogyakarta, Indonesia*
[2] *Istanbul University-Cerrahpasa, Istanbul, Turkey*
[3] *PT. GroundProbe Indonesia, Balikpapan, Indonesia*

**Corresponding author: e-mail barliandn@upnyk.ac.id*

Abstract

Purpose. In the open-pit mining, monitoring slope failure is a crucial activity. Many mining projects used a manual approach to measure the slope displacement, but an advanced technology (e.g., slope stability radar) offers more precise data collection. Despite these technological strides, accurately predicting slope failure remains essential to prevent costly incidents and interrupted production. Simple linear regression is commonly and widely used with time on the *x*-axis and inverse-velocity on *y*-axis, is a prevalent method. However, it often fails to forecast slope failure accurately, as the events tend to occur sooner than predicted. This misprediction might be attributed to the oversight of error-term within the model.

Methods. The error-term distinct structure was analyzed to improve the accuracy of the existing linear regression model. To address this, copula models were employed, as they effectively capture the complex dependence patterns among random variables, providing a robust method for incorporating the error-term analysis into the model.

Findings. The proposed approach showed its powerful technique to predict slope failure more accurately than the existing simple linear regression model. According to the slope failure datasets, the linear regression model was $y = 50475.270 – 1.123\,x$. Furthermore, the error-term was modeled through Reversed Gumbel-Hougaard (GH-RR) copula model under parameter $\theta = 1.12$. As a result, the prediction missed by 86.4 seconds, compared to 277.284 seconds when using the existing approach (i.e., linear regression without copula-based error-term modeling).

Originality. This approach emphasizes the analysis and incorporation of the error-term in the model, which is often overlooked in the simple linear regression method commonly used.

Practical implications. Implementing the error-term copula-based model could significantly improve the accuracy of slope failure predictions, thereby preventing costly incidents and ensuring uninterrupted production in mining projects.

Keywords: *slope failure, inverse-velocity, copula, error-term modeling, linear regression*

1. Introduction

Slopes failure prediction is a crucial aspect in the mining industry. Displacements of land surface in natural and man-made slopes are considered as the source of potential environmental risks, including the man-made ones, which often result in emergencies or disaster [1]. Landslides occur when masses of rocks, soil material, or muddy flow move down a slope, caused by disturbances in the natural stability of a slope [2]. The prediction plays a crucial role to ensure the safety and operation stability of open-pit mines [3], [4]. Over the years, various methods have been developed and employed for this purpose. Inverse-velocity method proposed by Fukuzono [5] is commonly used in slope failure prediction. He utilized a relationship between inverse of the slope displacement velocity, i.e., inverse-velocity, (*y*-axis), and time (*x*-axis), where the failure is expected when the inverse-velocity reaches zero $\left(\frac{1}{v} = 0\right)$.

Fukuzono's method has been validated by numerous applications, demonstrating its reliability and effectiveness in predicting slope failures in the open-pit mining industry. The method is valuable due to not only its reliability and effectiveness [6], but also its simplicity, clear algorithm, and relatively easy to understand. Furthermore, the use of continuous monitoring and including updated data in the analysis are highly recommended on using this method to ensure the effectiveness [7].

However, while the inverse-velocity method by Fukuzono [3] looks promising, it relies on certain assumptions and may not fully account for the uncertainty, which is often observed in real-world data. The model assumes slope failure follows a linear trend and assumes the error-term is ignored in the prediction analysis. This oversight can lead to inaccuracies. Thus, Rose and Hungr [8], as well as Carlà, Intrieri [7], recommend conducting regular time-window updates in the model to maintain the model accuracy.

Received: 5 August 2024. Accepted: 4 November 2024. Available online: 30 December 2024

Mining of Mineral Deposits. ISSN 2415-3443 (Online) | ISSN 2415-3435 (Print)

Despite these efforts, the continuous monitoring and time-window updates, failures may still occur sooner than anticipated. This inaccuracy might be due to the non-linear trend or the underestimated error-term presence [9]. On the other hand, the error-term, somehow, is one of the uncertainties in the regression model that is barely incorporated into the analysis, whether it is a linear or non-linear model.

Given that the uncertainty cannot be eliminated and accurately predicted, several papers proposed approaches to model the uncertainties in the slope failure prediction. Based on the literatures, the proposed approaches can be classified into three categories namely:

– incorporating probability concept such as Monte-Carlo simulation (MCS) [10]-[12] and bootstrapping procedure [13], [14];

– modeling the variation through copulas [15], [16] and stochastic approaches [17], [18];

– the combination of both.

The utilization of copulas in the slope failure prediction is quite new and limited. Zhou, Jing [15] incorporated a copula model in the Generalized F discrepancy minimization technology to understand the correlation between several geotechnical parameters. Xu and Zhou [16] compared copulas with regular risk assessment technique (MCS).

Copulas are considered powerful statistical analysis due to their ability to capture the dependence pattern between multiple random variables. Unlike traditional methods that often assume normally distributed error-terms, copulas allow for modeling different types of probability distributions [19]. This flexibility makes them useful for capturing complex relationships and dependences in a dataset. In other words, by using copulas, one can accurately describe the joint distribution of variables and handle heteroscedasticity, which is the condition of having non-constant variability in a dataset. Thus, implementing copula methods can provide a more rational statistical model where unexplained variances are taken into account.

A research to enhance the inverse-velocity theory by incorporating a probability concept was conducted by Manconi and Giordan [20]. The proposed method was enriching the regression model by applying bootstrap procedure to add confidence interval in the model. Bootstrapping, a resampling technique, allows for estimation of the sampling distribution of a static by repeatedly sampling with replacement data. By applying bootstrapping, they were able to construct confidence interval for their model parameters. In addition, the proposed approach by Manconi and Giordan [20] demonstrated a significant advancement in the application of inverse-velocity method. In short, the proposed technique shows a convenience, intuitive, and robustness.

Another regression enhancement was also conducted by Li, Xu [21], who explored the integration of a linear ensemble-based extreme learning machine and copulas within an autoregressive moving average (ARMA) model. The copulas used in the model were applied to capture the variation in the random variables of the displacement, reservoir water level, and the precipitation. As a result, Gumbel-Hougaard showed the best performance based on the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) number. These criteria are widely used for model selection, with lower values indicating a better fit to the data. The superior performance of the Gumbel-Hougaard highlighted its ability to capture the tail dependences and the asymmetrical nature of the relationship among variables.

On the other hand, the use of a linear ensemble-based extreme learning was due to its ability on fast-learning speed and high generalization performance. The model was suitable for complex trends and non-linearity of the datasets. By combining copulas and linear ensemble-based extreme learning, Li, Xu [21] successfully developed a robust and efficient approach to model the stochastic nature of the slope failure, especially in terms of capturing the dependence patterns of the random variables.

Currently, copulas are increasingly recognized as a vital tool in risk assessment and statistical modeling due to their ability to capture the dependence structure between random variables, regardless of the complexity of their relationship [22]. This feature makes copulas valuable in cases where traditional correlation measures (e.g., Pearson and Spearman correlation) are insufficient. Al-Harthy, Begg [23] observed the efficacy of copula in modeling dependences that are not easily modeled by linear correlations.

In another intensive application, Singh, Ardian [24] utilized copulas to model heteroscedasticity within a fractionally integrated moving average autoregressive model of a generalized autoregressive time series model with conditional heteroscedasticity (ARFIMA-GARCH). The ARFIMA itself is suitable for capturing long memory properties in time series data, while the GARCH component accounts for volatility over time. By integrating copulas into ARFIMA-GARCH model, Singh, Ardian [24] aimed to better represent the stochastic dependences and co-movements among random variables, which are critical for model accuracy in the forecasting and risk analysis.

In general, uncertainty in the regression model can be modeled by two techniques:

– minimizing the error-term based on historical data through statistical techniques (e.g., ARMA, ARFIMA, and regression);

– capturing variation through simulation techniques (e.g., Monte-Carlo, bootstrapping, and stochastic processes).

The statistical technique typically involves developing models that seek to minimize the discrepancy between observed values and the predicted values. This is often achieved by fitting a model to the historical data and optimizing the parameters to reduce the error-term, where ordinary least square (OLS) algorithm is commonly used. In contrast, the simulation techniques encompass methods that aim to replicate the underlying stochastic nature of generating the data. The simulation is useful for quantifying the uncertainty in complex systems where analytical solutions may be intractable. This paper, thus, focuses on improving the regression model (i.e., statistical approach) by modeling the error-term variation through copula (i.e., simulation approach).

Eventually, slope failure prediction is a critical component in geotechnical engineering, directly impacting the safety and reliability of the operation, especially in the mining industry. Traditional methods for slope failure prediction often rely on oversimplification assumptions, especially about the dependence relationship among random variables, which limit the accuracy and the reliability of the model. To address these limitations, this research, thus, proposes an innovative approach. The proposed approach employs error-term simulation modeling through a copula method to enhance the existing methodologies for slope failure prediction. By employing copulas, the complex relationship and dependences inherent in geotechnical data, leading to more precise and reliable risk assessment, can be better captured. In addition, the proposed technique can potentially benefit decision makers in slope failure risk profiling.

2. Methods

2.1. Simple linear regression

The common simple linear regression, where the inverse-velocity and the time are the *y*-axis and *x*-axis, respectively, was still used to forecast the slope failure in this research. The linear regression was then enhanced by modeling the error-term through a best fitted copula, regardless of its error-term variation (i.e., homoscedasticity or heteroscedasticity). The linear regression itself is expressed by Equation (1) where y is the dependent variable (inverse-velocity), a is the intercept, β is the slope, x is the independent variable (time), and ε is the error-term:

$$y = a + \beta x + \varepsilon\,; \tag{1}$$

$$\beta = \frac{Cov(x, y)}{Var(x)}\,; \tag{2}$$

$$a = \bar{y} - (\beta)(\bar{x})\,. \tag{3}$$

The β and α can be calculated using the ordinary least square (OLS) algorithm, which follows Equations (2) and (3), respectively. The OLS algorithm ensures that the estimates of β and α result in the lowest possible value of the summation of squared error (i.e., minimizing error). However, this method typically assumes that the error-term (ε) is random and cannot be predicted and commonly ignored in the model. These assumptions can be problematic, especially in the presence of heteroscedasticity.

2.2. Error-term variation modeling through copula-based modeling

One important assumption in the regression model is the homoscedasticity, while the error-term is purely random and exhibits perfectly constant variance across all levels of the independent variables. Despite its importance, achieving that ideal condition in practical application is rare. This violation of the homoscedasticity assumption can result in inefficient estimators and biased standard errors, which furthermore can compromise the reliability of the model and its confidence intervals [25]. The presence of heteroscedasticity might come from several sources such as outliers, omission of relevant variables [26], measurement errors [27], and so on. These factors result in non-constant variance of the error-term, complicating the modeling process. In addressing heteroscedasticity, copula-based models offer a robust solution. Copulas enable capturing dependences within the dataset. By modeling the joint distribution of variables separately from their marginal distributions, copulas can effectively account for the varying relationship and dependences that contribute to heteroscedasticity. In addition, if the model is homoscedasticity, copulas can still be used as effectively as in heteroscedastic cases.

2.3. Copulas model and the parameter estimation

Copula is mathematically defined as a joint distribution function model of the uniformly distributed [0, 1] random variables that has a unique dependence pattern based on its correlation described by its parameter [28], [29]. Mathematically, let H be a joint distribution function for F and G, the $F(x)$ and $G(y)$ are the margin distribution of F and G, then copula C is defined as $H(x, y) = C(F(x), G(y))$. Regardless of the marginal distribution, copulas are classified by class and family based on its generator and parameter [29]. However, in this research, the family of Clayton, Gumbel-Hougaard (GH), and Frank copulas within the class of Archimedean copula were explored due to its simplicity and common patterns.

The simplicity is due to its relatively simple model construction and parameter estimation (θ) through Kendall's tau (τ) correlation model. The τ and the θ for Clayton, GH, and Frank can be calculated by Equations (4), (5), (6), and (7), respectively [30]-[32]:

$$\tau = \frac{2(n_c - n_d)}{n(n-1)}\,; \tag{4}$$

$$\theta = \frac{2\tau}{1-\tau}\,; \tag{5}$$

$$\theta = \frac{1}{1-\tau}\,; \tag{6}$$

$$\frac{D_1(\theta)-1}{\theta} = \frac{1-\tau}{4}\,, \tag{7}$$

where:

$D_1(\theta)$ – the Debye first-order function that follows Eq. (8):

$$D_k(\theta) = \frac{1}{\theta}\int_0^{\theta} \frac{t}{\exp(t)-1}\,dt\,. \tag{8}$$

Subsequently, the copulas model can be constructed based on its families, the Clayton, GH, and the Frank that follow Equations (9), (10), and (11), respectively:

$$C = \left[\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right]^{-\frac{1}{\theta}}\,; \tag{9}$$

$$C = \exp\left(-\left[(-\ln u)^{\theta} + (-\ln v)^{\theta}\right]^{\frac{1}{\theta}}\right); \tag{10}$$

$$C = -\frac{1}{\theta}\ln\left(1 + \frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}\right). \tag{11}$$

Another Archimedean dominance application is due to the pattern produced. The patterns cover three common scatterplot patterns:

– the tight pattern at the beginning and more disperse afterwards;

– disperse at the beginning and tighter afterwards;

– relatively uniform pattern for both at the beginning and at the end. In addition, the modification of each of those families is also simply done by reflecting the model [33]. Figures 1, 2, and 3 show the pattern of each family of Archimedean copula and its reflection, respectively.

The Clayton copula pattern is tight at the beginning and disperses afterwards. Nevertheless, the Clayton copula can be reflected by three more reflection axis:

– reflecting to its *y*-axis;

– reflecting to its *x*-axis;

– reflecting to its both axis, the *x*-axis and *y*-axis [33].

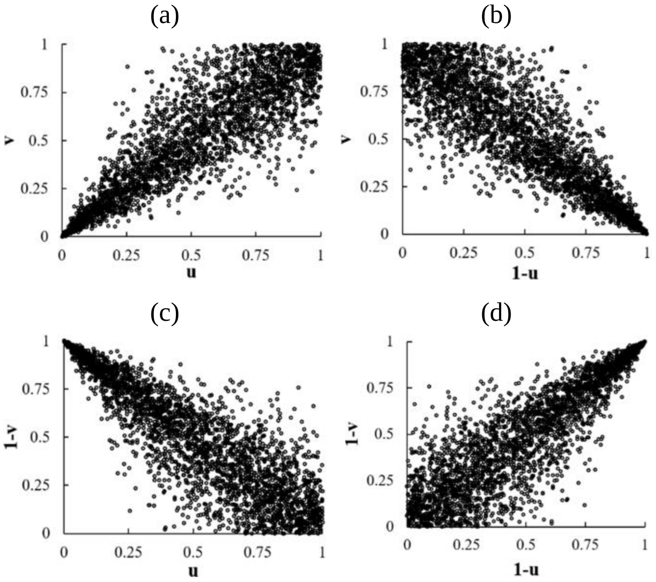

Figure 1.* *Clayton copula dependence pattern: (a) Clayton (C); (b) Clayton RY (C-RY); (c) Clayton RX (C-RX); (d) reversed Clayton (C-RR)

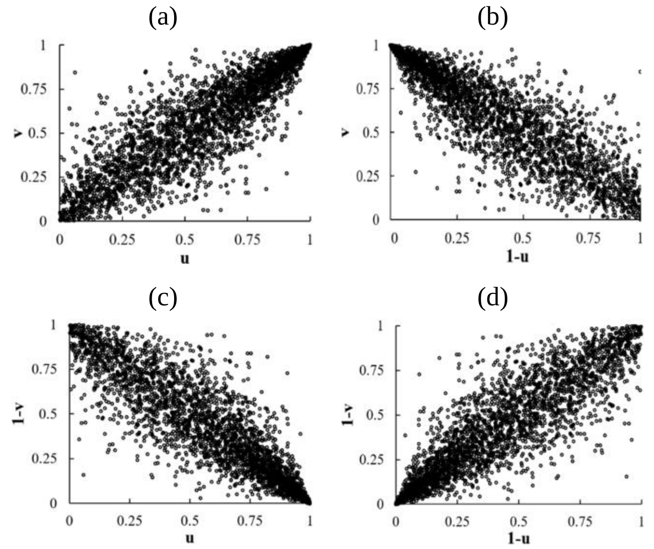

Figure 2. Gumbel-Hougaard (GH) copula dependence pattern: (a) Gumbel-Hougaard (GH); (b) Gumbel-Hougaard RY (GH-RY); (c) Gumbel-Hougaard RX (GH-RX); (d) reversed Gumbel-Hougaard (GH-RR)

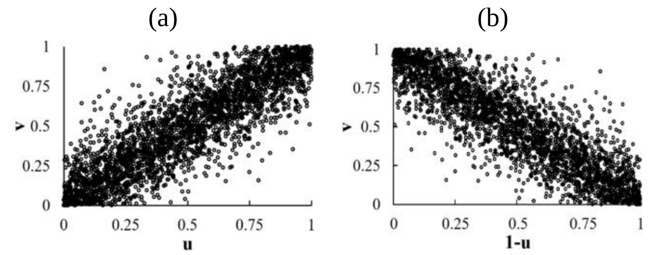

Figure 3. Frank copula dependence pattern: (a) Frank (F); (b) Frank RX (F-RX)

This reflection concept obviously expands the possibility of copula application. The GH copula reflection concept is similar to Clayton reflection concept. It can be reflected to its *x*-axis, *y*-axis, and both. However, instead of being dispersed at the tail, GH tail becomes tight. The dispersion happens in the middle (body).

Both, Clayton and GH, have tight head and tail. Frank copula, on the other hand, has relatively dispersed head and tail. Therefore, the Frank copula can only be reflected on its *x*-axis. Frank copula reflection on the *y*-axis or on both axis results in an identical pattern with RX and regular copula, respectively.

The mathematical model for reflected Archimedean copulas is constructed based on Equations (12), (13), and (14) for the reflection of a regular copula pattern on its *x*-axis, *y*-axis, and both axes, respectively [34]. Note that the Frank copula has a symmetrical pattern, thus it may only be reflected on its *x*-axis. Thus, focusing on just three families of Archimedean copula and its modification, ten different common error-term patterns can be modeled:

$$\begin{cases} u \to u \\ v \to 1-v \end{cases}; \tag{12}$$

$$\begin{cases} u \to 1-u \\ v \to v \end{cases}; \tag{13}$$

$$\begin{cases} u \to 1-u \\ v \to v-v \end{cases}. \tag{14}$$

The application of copulas requires transforming the *x* and *y* dataset into Uniform distribution range [0, 1] for *u* and *v*, respectively.

2.4. Copula simulation through MCS algorithm

Construction of the copula dependence model for *x* and *y* using the previous steps is insufficient to show the risk profile. A copula simulation through MCS algorithm is then proposed. The copula MCS was discussed by Mari and Kotz [35], Al-Harthy, Begg [23], and Ardian and Kumral [36]. The Clayton and Frank copulas have the same algorithm, as shown below. Let *u* and *w* are random variables [0, 1]. Reformulate Archimedean copula function as $v = \varphi^{-1}\left[\varphi(w) - \varphi(u)\right]$. Simulate Clayton and Frank by Equations (15) and (16), respectively:

$$v_{Clayton} = \left[(w)^{-\frac{\theta}{1+\theta}} (u)^{-\theta} - u^{-\theta} + 1 \right]^{-\frac{1}{\theta}}; \tag{15}$$

$$v_{Frank} = -\frac{1}{\theta} \ln \left[1 - \frac{\left(1 - \exp(-\theta)\right)}{\left(1 + \exp(-\theta u)\left(w^{-1} - 1\right)\right)} \right]. \tag{16}$$

However, the GH copula has no closed form. The GH simulation is based on the following algorithm. Let the variable *s* ranges randomly [0, 1] and variable *q* – consecutively [0, 1]. Simulate GH by finding *u* and *v* by Equations (17) and (18).

$$u = \exp\left[s^{1/\theta} \ln(q) \right]; \tag{17}$$

$$v = \exp\left[\left(1 - s^{1/\theta}\right) \ln(q) \right]. \tag{18}$$

Using algorithms, find u and v for three copulas. Eventually, a back transformation must be done to return the copula simulation into its initial distribution, as well as to its initial unit.

2.5. Model selection

In this research, Akaike information criterion (AIC) was used to select the best model [37]. The AIC was used to select the best copula and distribution fitting [38]. The model with the lowest AIC value is considered as the best model. The AIC follows Equation (19):

$$AIC = \left(\frac{2n}{n-k-1}\right)k - 2\ln\left[L_{\max}\right], \tag{19}$$

where:

n – the sample size;

k – the number of parameters;

ln [L_{max}] – the maximized value of log-likelihood.

Although the lowest AIC number is considered as the best model, the second or other lowest number might be used due to its intuitive consideration [33].

2.6. The proposed technique algorithm

In order to show the steps of the proposed approach, an algorithm of the proposed approach in this paper is explained subsequently. First of all, when conducting the proposed approach, a simple linear regression was done based on the given dataset. The dataset was the velocity and the time to failure. Note that, the velocity is the y-axis, but it is inversed, whereas the time is the x-axis. Next step was to build the linear regression model, and the model can be seen in Equation (1). When constructing a linear regression model, the slope and intercept must be estimated in advance. The slope and intercept estimation through the OLS estimator follow Equations (2) and (3), respectively.

Second, perform the error-term modeling. The error-term modeling was done through copula. The error-term dataset was gathered from calculating the difference between observed value (i.e., historical data) and the predicted value which was obtained from the previous step (i.e., the simple linear regression). Next, the rank correlation of the error-terms was estimated by Equation (4). After that, the copula parameters were estimated by Equations 5), (6), and (7), respectively.

Third, the error-term was fitted into one of the Archimedean copulas by conducting the AIC method in Equation (9) before copula simulation could be done. After selecting the best copula, the copula simulation was done by following Equations (9), (10), or (11), for Clayton, GH, and Frank, respectively. The simulation used MCS algorithm to reproduce the error-term. Clayton and Frank simulation follows Equations (15) and (16), respectively. The GH follows Equations (17) and (18).

Finally, the reproduced error-terms were incorporated into the linear regression model in Equation (1). In other words, the scatterplot of inversed-velocity vs time was also reproduced. Assuming that the inversed velocity is less than or equal to 0 on x-axis $\left(\frac{1}{v} \le 0\right)$ as the slope failure time, thus all the reproductions $\left(\frac{1}{v} \le 0\right)$ were thus fitted into probability distribution to show the slope failure risk profile.

3. Results

3.1. Simple linear regression modeling

The slope failure dataset (the time and velocity) was obtained from coal mine in Indonesia and provided by GroundProbe. GroundProbe is a real-time geo-hazard measuring and monitoring company. The dataset was modeled by simple linear regression in Equation (1). The scatterplot and the corresponding linear regression model of the dataset can be seen in Figure 4. The scatterplot illustrates the distribution of data points, highlighting the trend, and variability in the measurements. The imposed linear regression on the scatterplot represents the best-fit line that minimizes the sum of squared errors.

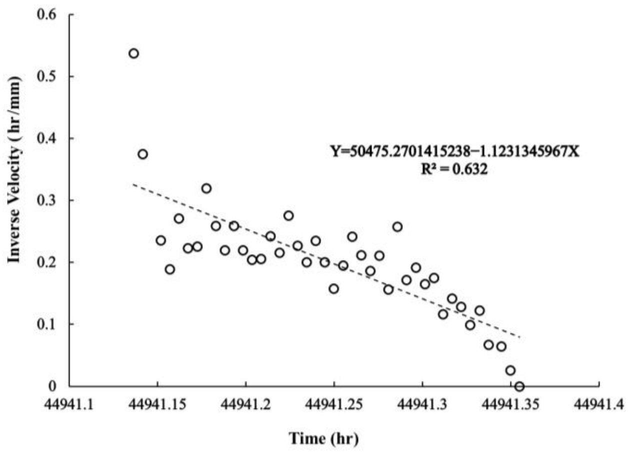

Figure 4. The scatterplot and the linear regression of the dataset

Based on Figure 4, there were 42 datasets, collected over a time interval starting from $t = 44941.136$ hours until the point of failure at $t = 44941.355$ hours. It means that there was duration of 0.219 hours or 13.113 minutes or 786.791 seconds from the initial observation to the occurrence of the actual slope failure. The linear regression model applied to this dataset was defined by equation $y = 50475.270 – 1.123\,x$. According to the model, the failure was predicted to happen in $t = 44941.432$. Comparing this predicted time with the actual failure time, a discrepancy of 277.284 seconds (4.621 minutes) was observed.

This analysis underscores the efficacy of linear regression in approximating the time of slope failure, but the discrepancy should be kept in mind. The 4.621 minutes difference between the predicted and actual failure times highlighted the uncertainty in predictive modeling of geotechnical phenomena. Nevertheless, the application of a linear regression model still provides valuable insights to the decision makers.

3.2. Error-term modeling

The error-terms were derived from the 42 observed data points and their corresponding predicted values – representing the residuals or the difference between observed and predicted values – were subjected to further statistical examination. To model the dependence pattern of these residuals (i.e., error-terms), Archimedean copulas were employed. The selection of the most appropriate copula model was done through the AIC method using Equation (19). Among the evaluated copulas, the Reversed GH (GH-RR) was chosen. The top three AIC values for each copula are summarized in Table 1.

Table 1. The AIC value for each Archimedean copula

Name	AIC
Reversed Gumbel-Hougaard (GH-RR)	2.35
Clayton (C)	2.48
Frank (F)	3.48

In the process of simulating the GH-RR, it was essential to estimate the copula parameter θ using Equation (6). The estimation of θ required the prior calculation of Kendall's tau

(τ), which was derived from the error-terms and using Equation (4). Once the τ was estimated, it could be substituted into Equation (6) for copula parameter estimation process ($\theta = 1.12$). The next step involved using Equations (17) and (18) to reproduce (or simulate) u and v, respectively. The 10000 reproduction of the error-terms was conducted.

To evaluate the performance of the GH-RR copula simulation, a comparison between actual error-terms and the reproduction through GH-RR is presented in Figure 5. Such a visual representation provides an intuitive understanding of how well the copula model replicates the observed error-term dependence pattern.

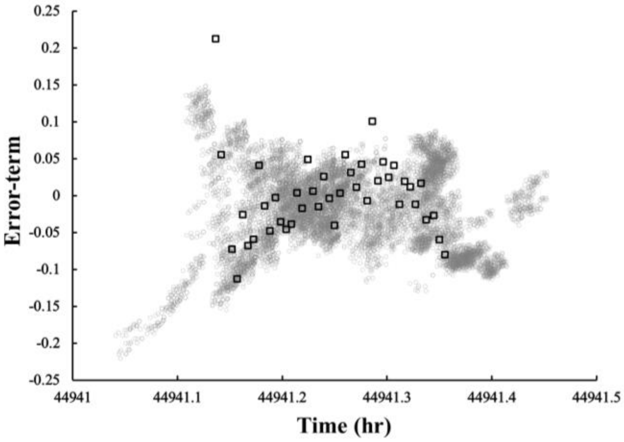

Figure 5. The scatterplot of the error-term

Based on Figure 5, the scatterplot illustrates the distribution of the actual error-terms (represented by square dots) alongside the GH-RR simulated error-terms (represented by grey dots). Upon examination, it was evident that the simulated error-terms effectively covered the actual error-terms, indicating a high degree of a reliability of the GH-RR copula model in capturing the error-terms in the slope failure linear regression model.

This congruence between the simulated and actual error-terms underscores the robustness of the GH-RR copula model in modeling the error-term behavior. The ability to reproduce the observed error-terms through simulation provided confidence in the application of this model to real-world data. Such reliability was crucial, as it validated the GH-RR copula potential to accurately reflect the stochastic nature of error-terms, which were often neglected in traditional linear regression analysis.

To further assess the utility of the GH-RR simulation, these simulated error-terms were incorporated into the existing linear regression model. This step aimed to evaluate the model predictive accuracy concerning slope failure, particularly when the error-terms were accounted for in the analysis. In contrast, traditionally, the error-terms were either simplified or even ignored, which potentially led to an incomplete understanding of the model predictive capacity. Moreover, the incorporation of the simulated error-terms into the regression model allowed for a more comprehensive evaluation of its performance. The model provided a more nuanced and accurate prediction of slope failure, enhancing its practical applicability and management context (i.e., decision making process). The GH-RR simulated error-term incorporation can be seen in Figure 6.

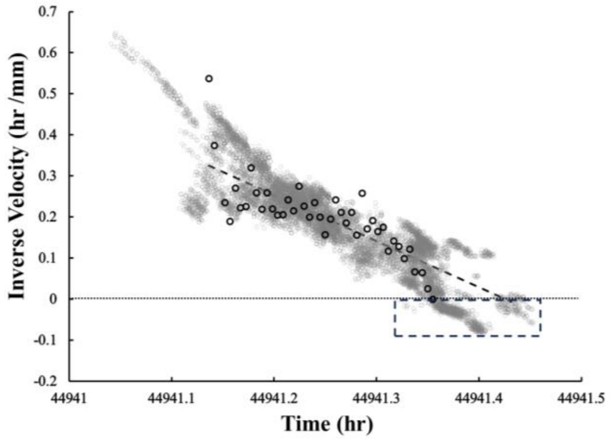

Figure 6. The incorporation of the GH-RR simulated error-term

Figure 6 presents the linear regression of the slope failure model (represented by black linear line) together with its actual dataset (represented by black circles) and the GH-RR simulated error-term (represented by grey dots). More than that, the rectangular dashed lines area highlights the error-terms that exceeded a critical threshold (i.e., $1/v = 0$). The grey dots within the rectangular dashed lines area are the variation of the slope failure prediction under error-term copula-based modeling.

Following the identification of predicted slope failure occurrences and using error-term copula-based modeling in Figure 6, the next step was fitting those predictions into the probability distribution function to display the possibility of the slope failure occurrence. Utilizing the probability distribution function made the model more realistic due to its probability framework, instead of a deterministic one. The AIC method was also used in the process of fitting the distribution to select the best probability distribution function. The AIC values for those distributions are presented in Table 2.

Table 2. The AIC value of the error-term within the rectangular dashed lines

Name	AIC
Ext-Value-Max	-4185.78
Gamma	-4099.99
Normal	-4049.54

Table 2 shows the AIC value of the three probability distribution functions. The AIC values of the Ext-Value-Max, Gamma, and Normal distribution were -4185.78; -4099.99; and -4049.54, respectively. The Ext-Value-Max shows the lowest and significantly differs from the other distributions. Therefore, the Ext-Value-Max distribution was chosen.

3.3. Slope failure risk profiling

Determining the slope failure at one specific time is too risky due to ignoring the uncertainty that might be involved in the slope failure occurrences. Using the probability concept – where a range of slope failure time was provided – made this proposed approach more rational. The Ext-Value-Max in this case was selected and fitted together with the actual time of slope failure and the predicted time as can be seen in Figure 7. The probability distribution function of the Ext-Value-Max distribution (represented by red line) was fitted to the simulated error-term with the parameters a and b, 44941.375 and 0.021, respectively.

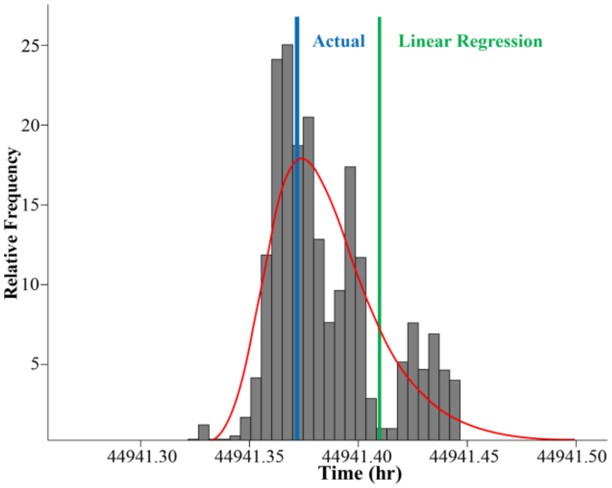

Figure 7. The slope failure profile

The blue line shows the actual time of slope failure, which occurred at $t = 44941.355$. The green line represents the slope failure prediction by linear regression, which estimated the failure time at $t = 44941.432$. In addition, according to Ext-Value-Max distribution, the minimal, maximal, and the mode were 44941.322, 44941.453, and 44941.379, respectively.

In other words, it has been shown that the proposed model predicted the slope failure, which occurred between $t = 44941.322$ (the minimal) to $t = 44941.453$ (the maximal), with the high probability to happen at $t = 44941.379$. Assuming that the mode is the recommended one, the proposed model (i.e., copula-based error-term modeling) was wrong by 0.024 hours (or 86.4 seconds), whereas the conventional linear regression exhibited a higher misprediction of 0.077 hours (277.284 seconds).

The comparison between the actual time, the proposed approach, and the conventional model proved that the predictive capability offered by the proposed approach was not only more powerful than the conventional one, but also might show the slope failure occurrence possibility through a specified continuous probability distribution function. Thus, instead of using a simple linear regression alone, incorporating the error-term copula-based modeling enhanced the existing model. The model benefits from a more comprehensive framework that accounts for stochastic variability in real-world cases.

4. Conclusions

In geotechnical engineering, a simple linear regression stands as a prevalent and extensively employed tool for forecasting slope failures, particularly within the mining industry. The method is generally reliable, but often predicts slope failures occurring later than their actual timing. Such discrepancies underline the oversight of the error-term, which has a unique pattern, but is commonly ignored in the analysis.

Incorporating error-term copula-based modeling shows a significant advancement over relying solely on conventional linear regression models. Copula models offer a more nuanced approach by capturing the complex interdependencies within the error-term structure. This capability allows the error-term copula-based modeling approach to provide more accurate predictions of slope failure timings compared to the simplistic assumptions of linear regression. In addition, the probabilistic framework in the proposed approach makes it more realistic.

Author contributions

Conceptualization: BD, VV; Data curation: SS, RS; Formal analysis: ÖU, AA; Funding acquisition: AA; Investigation: VV, RS; Methodology: BD, ÖU; Project administration: VV; Software: VV, AA; Supervision: BD, SS; Validation: BD, SS; Visualization: VV, RS; Writing – original draft: VV, AA; Writing – review & editing: BD, AA. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported financially by internal research grant of Universitas Pembangunan Nasional "Veteran" Yogyakarta (LPPM – UPN "Veteran" Yogyakarta) and PT Studio Mineral Batubara.

Conflicts of interests

The authors declare no conflict of interest.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

References

[1] Kovrov, O., Kolesnyk, V., & Buchavyi, Y. (2020). Development of the landslide risk classification for natural and man-made slopes based on soil watering and deformation extent. *Mining of Mineral Deposits*, *14*(4), 105-112. https://doi.org/10.33271/mining14.04.105

[2] Boubazine, L., Boumazbeur, A., Hadji, R., & Fares, K. (2022). Slope failure characterization: A joint multi-geophysical and geotechnical analysis, case study of Babor Mountains range, NE Algeria. *Mining of Mineral Deposits*, *16*(4), 65-70. https://doi.org/10.33271/mining16.04.065

[3] Kolapo, P., Oniyide, G.O., Said, K.O., Lawal, A.I., Onifade, M., & Munemo, P. (2022). An overview of slope failure in mining operations. *Mining*, *2*(2), 350-384. https://doi.org/10.3390/mining2020019

[4] Yu, S., Ke, Y., Deng, H., Tian, G., & Deng, J. (2021). Experimental investigation of porous and mechanical characteristics of single-crack rock-like material under freeze-thaw weathering. *Minerals*, *11*(12), 1318. https://doi.org/10.3390/min11121318

[5] Fukuzono, T. (1985). A method to predict the time of slope failure caused by rainfall using the inverse number of velocity of surface displacement. *Landslides*, *22*(2), 8-13. https://doi.org/10.3313/jls1964.22.2_8

[6] Cabrejo, A. (2021). Failures in open-pit mines and uncertainty when predicting collapses. *GroundProbe*, 1-9.

[7] Carlà, T., Intrieri, E., Di Traglia, F., Nolesini, T., Gigli, G., & Casagli, N. (2017). Guidelines on the use of inverse velocity method as a tool for setting alarm thresholds and forecasting landslides and structure collapses. *Landslides*, *14*, 517-534. https://doi.org/10.1007/s10346-016-0731-5

[8] Rose, N.D., & Hungr, O. (2007). Forecasting potential rock slope failure in open pit mines using the inverse-velocity method. *International Journal of Rock Mechanics and Mining Sciences*, *44*(2), 308-320. https://doi.org/10.1016/j.ijrmms.2006.07.014

[9] Zhang, J., Yao, H. Z., Wang, Z. P., Xue, Y. D., & Zhang, L. L. (2023). On prediction of slope failure time with the inverse velocity method. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *17*(1), 114-126. https://doi.org/10.1080/17499518.2022.2132263

[10] Wu, Y., & Lan, H. (2019). Landslide analyst – A landslide propagation model considering block size heterogeneity. *Landslides*, *16*(6), 1107-1120. https://doi.org/10.1007/s10346-019-01154-2

[11] Li, D.Q., Zhang, F.P., Cao, Z.J., Zhou, W., Phoon, K.K., & Zhou, C.B. (2015). Efficient reliability updating of slope stability by reweighting failure samples generated by Monte Carlo simulation. *Computers and Geotechnics*, *69*, 588-600. https://doi.org/10.1016/j.compgeo.2015.06.017

[12] Li, A.J., Cassidy, M.J., Wang, Y., Merifield, R.S., & Lyamin, A.V. (2012). Parametric Monte Carlo studies of rock slopes based on the Hoek-Brown failure criterion. *Computers and Geotechnics*, *45*, 11-18. https://doi.org/10.1016/j.compgeo.2012.05.010

[13] Li, D.Q., Tang, X.S., & Phoon, K.K. (2015). Bootstrap method for characterizing the effect of uncertainty in shear strength parameters on

slope reliability. *Reliability Engineering & System Safety*, *140*, 99-106. https://doi.org/10.1016/j.ress.2015.03.034

[14] Efron, B. (1992). Bootstrap methods: Another look at the jackknife. *Breakthroughs in Statistics: Methodology and Distribution*, 569-593. https://doi.org/10.1007/978-1-4612-4380-9_41

[15] Zhou, Y., Jing, M., Pang, R., Xu, B., & Yao, H. (2022, September). A novel method for the dynamic reliability analysis of slopes considering dependent random parameters via the direct probability integral method. *Structures*, *43*, 1732-1749. https://doi.org/10.1016/j.istruc.2022.07.074

[16] Xu, Z.X., & Zhou, X.P. (2018). Three-dimensional reliability analysis of seismic slopes using the copula-based sampling method. *Engineering Geology*, *242*, 81-91. https://doi.org/10.1016/j.enggeo.2018.05.020

[17] Das, I., Kumar, G., Stein, A., Bagchi, A., & Dadhwal, V.K. (2011). Stochastic landslide vulnerability modeling in space and time in a part of the northern Himalayas, India. *Environmental Monitoring and Assessment*, *178*, 25-37. https://doi.org/10.1007/s10661-010-1668-0

[18] Napoli, M.L., Barbero, M., Ravera, E., & Scavia, C. (2018). A stochastic approach to slope stability analysis in bimrocks. *International Journal of Rock Mechanics and Mining Sciences*, *101*, 41-49. https://doi.org/10.1016/j.ijrmms.2017.11.009

[19] Ahooyi, T.M., Arbogast, J.E., & Soroush, M. (2015). An efficient copula-based method of identifying regression models of non-monotonic relationships in processing plants. *Chemical Engineering Science*, *136*, 106-114. https://doi.org/10.1016/j.ces.2015.03.044

[20] Manconi, A., & Giordan, D. (2016). Landslide failure forecast in near-real-time. *Geomatics, Natural Hazards and Risk*, *7*(2), 639-648. https://doi.org/10.1080/19475705.2014.942388

[21] Li, H., Xu, Q., He, Y., & Deng, J. (2018). Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models. *Landslides*, *15*, 2047-2059. https://doi.org/10.1007/s10346-018-1020-2

[22] Yang, L., Frees, E. W., & Zhang, Z. (2020). Nonparametric estimation of copula regression models with discrete outcomes. *Journal of the American Statistical Association*, *115*(530), 707-720. https://doi.org/10.1080/01621459.2018.1546586

[23] Al-Harthy, M., Begg, S., & Bratvold, R.B. (2007). Copulas: A new technique to model dependence in petroleum decision making. *Journal of Petroleum Science and Engineering*, *57*(1-2), 195-208. https://doi.org/10.1016/j.petrol.2005.10.015

[24] Singh, J., Ardian, A., & Kumral, M. (2021). Gold-copper mining investment evaluation through multivariate copula-innovated simulations. *Mining, Metallurgy & Exploration*, *38*(3), 1421-1433. https://doi.org/10.1007/s42461-021-00424-9

[25] Poole, M.A., & O'Farrell, P.N. (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 145-158. https://doi.org/10.2307/621706

[26] Zhou, Q.M., Song, P.X.K., & Thompson, M.E. (2015). Profiling heteroscedasticity in linear regression models. *Canadian Journal of Statistics*, *43*(3), 358-377. https://doi.org/10.1002/cjs.11252

[27] Tracy, K., Waititu, H., & Sewe, S. (2022). A Copula-based approach for modelling the dependence between inflation and exchange rate in Kenya. *Asian Journal of Probability and Statistics*, *18*(4), 59-84. https://doi.org/10.9734/ajpas/2022/v18i430457

[28] Embrechts, P., Lindskog, F., & McNeil, A. (2001). *Modelling dependence with copulas, rapport technique*. Zurich, Switzerland: Swiss Federal Institutes of Technology.

[29] Nelsen, R.B. (2006). *An introduction to copulas*. New York, United States: Springer, 269 p.

[30] Genest, C., & Favre, A. C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, *12*(4), 347-368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

[31] Clemen, R.T., & Reilly, T. (1999). Correlations and copulas for decision and risk analysis. *Management Science*, *45*(2), 208-224. https://doi.org/10.1287/mnsc.45.2.208

[32] Frees, E.W., & Valdez, E.A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, *2*(1), 1-25. https://doi.org/10.1080/10920277.1998.10595667

[33] Vose, D. (2008). *Risk analysis: A quantitative guide*. West Sussex, United Kingdom: John Wiley & Sons, 752 p.

[34] Ardian, A. (2021). *Investigation into correlations, dependencies, and interactions between uncertain variables in risk analysis of mining projects*. Montreal, Canada: McGill University, 178 p.

[35] Mari, D.D., & Kotz, S. (2001). *Correlation and dependence*. London, United Kingdom: Imperial College Press, 236 p. https://doi.org/10.1142/p226

[36] Ardian, A., & Kumral, M. (2021). Enhancing mine risk assessment through more accurate reproduction of correlations and interactions between uncertain variables. *Mineral Economics*, *34*, 411-425. https://doi.org/10.1007/s13563-020-00238-z

[37] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. https://doi.org/10.1109/TAC.1974.1100705

[38] Torikian, H., & Kumral, M. (2014). Analyzing reproduction of correlations in Monte Carlo simulations: Application to mine project valuation. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *8*(4), 235-249. https://doi.org/10.1080/17499518.2014.966116

Удосконалення моделі прогнозування руйнування схилу шляхом застосування копули Архімеда для моделювання похибки

Б. Двінагара, В. Вергіагара, О.Ф. Угурлу, С. Саптоно, Р. Салсабіла, А. Ардіан

Мета. Удосконалення моделі прогнозування руйнування схилів при видобутку корисних копалин відкритим способом шляхом застосування копули Архімеда для моделювання членів похибки, що є важливим для запобігання витратним інцидентам і перервам у гірничому виробництві.

Методика. Для підвищення точності існуючої лінійної регресійної моделі було проаналізовано чітку структуру членів похибки. Для вирішення цієї проблеми використані копула-моделі, оскільки вони ефективно відображають складні закономірності залежностей між випадковими величинами, забезпечуючи надійний метод включення аналізу членів похибки в модель.

Результати. Запропонований новий підхід показав надійну методологію прогнозування руйнування схилів більш точно, ніж існуюча проста лінійна регресійна модель. Отримано лінійну регресійну модель згідно з масивом даних про руйнування схилів, яка має вигляд $y = 50475.270 - 1.123\,x$. Змодельовано члена похибки за допомогою оберненої копули Гумбеля-Хоугаарда (GH-RR) з параметром $\theta = 1.12$. Встановлено, що прогноз був зроблений із запізненням на 86.4 секунди, порівняно з 277.284 секунд при використанні існуючого традиційного підходу (тобто лінійної регресії без моделювання членів похибки на основі копули). Визначено, що копульні моделі забезпечують точніший підхід, дозволяючи враховувати складні взаємозалежності у структурі помилки, а, ймовірнісна основа пропонованого підходу робить його більш реалістичним.

Наукова новизна. Розроблений новий науковий підхід, в якому акцентується увага на аналізі та включенні члена похибки в модель, що часто не береться до уваги в методі простої лінійної регресії, і, зазвичай, використовується.

Практична значимість. Впровадження моделі з використанням членів похибки на основі копули може значно підвищити точність прогнозування обвалів на схилах, тим самим запобігаючи витратним інцидентам і забезпечуючи безперебійне виробництво в гірничодобувних проєктах.

Ключові слова: *обвал схилу, зворотна швидкість, копула, моделювання із використанням похибки, лінійна регресія*

Publisher's note

All claims expressed in this manuscript are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.