Various Rainfall Forecasting Methods for Estimation of Pit Lake Flooding Duration in Indonesia

Ginting Jalu Kusuma, Abie Badhurahman, Sendy Dwiki, Salmawati, Rudy Sayoga Gautama, IK Dwika Paramananda

Department of Mining Engineering, Institut Teknologi Bandung, Indonesia

Abstract

Indonesia is the largest coal exporting country in the world, with most mines using open-pit mining methods due to the geological conditions of the coal deposits. Most of these open pit mines leave voids at the end of their mining period, due to technical reasons and/or operating at a low stripping ratio (< 5), which results in a lack of backfill material. The voids can fill with water, forming pit lakes that can be utilized for water supply, flood control, recreational opportunities, environmental rehabilitation, and protection. It is imperative to be able to calculate and predict the variability water balance of the pit lake including inputs from runoffs of surrounding catchment areas and groundwater, as well as output in the form of evaporation, runoff to spill point, and groundwater. The water balance of the pit lake adjacent to the stage volume curve of the pit lake is important to predict the flooding duration of the pit lake. Runoffs are a major component of the water balance and are directly governed by rainfall; thus, characterizing rainfall conditions and their variability, and using appropriate forecasting methods are important.

Most of the previous studies are undermining the variability of rainfall conditions in Indonesia and using a limited rainfall dataset. This study aims to evaluate each method of rainfall forecasting method to predict the flooding duration of a pit lake in Indonesia using long rainfall datasets. Daily rainfall datasets of 26 years are used in this study. The first 15-year dataset is used as input for rainfall forecasting methods whilst the last 11-year dataset is used to verify the methods. Different methods show that substantial difference in flooding duration. On the other hand, different methods also show different abilities to portray the uncertainties of the mid-term and long-term hydrometeorological conditions such as climate changes and climate pattern

Selecting the most appropriate method of rainfall forecasting is important to calculate the water balance of a pit lake and its flooding duration, especially in Indonesia. On the other hand, the ability of the model to capture different hydrometeorological conditions is an advantage to include uncertainties in the prediction of flooding duration.

Keywords: Void, pit lake, coal mine, rainfall forecasting, flooding

Introduction

Indonesia is the largest coal exporting country in the world, with most mines using open-pit mining methods due to the geological conditions of the coal deposits. Most of these open pit mines leave void at the end of their mining period, due to technical reasons for mining operations and/or operating at a low stripping ratio (< 5) resulting in a lack of backfilling material. The voids, hydrologically can be flooded, forming pit lakes which can be utilized for water supply, flood control, recreational opportunities, environmental rehabilitation, and protection. It is imperative to be able to calculate and predict the variability water balance of the pit lake including inputs from runoffs of surrounding catchment areas, groundwater as well as output in the form of evaporation, runoff to spill point, and groundwater. The water balance of the pit lake alongside the stage volume curve of the pit lake is important to predict the flooding duration of the pit lake. Runoffs are a major part of water balance and are directly governed by rainfall depth; thus, the characterizing of rainfall conditions and their variability and the selection of forecasting methods are important. This paper aims to explore the applicability of various models and approaches in the forecasting of monthly rainfall depth.

Methods

There are extensive models and approaches in the forecasting of monthly rainfall. In this study, models and approaches used in the forecasting of monthly rainfall are as follows (Latif et al., 2023; Maia & de Carvalho, 2011; Nwokike et al., 2020):

1. Monthly average (overall). The approach is an average using all training data
2. Linear regression. The approach uses a linear trendline of all training data
3. Statistical Approach (seasonality). The training data is divided into groups based on months (January-December). The mean/average value, quartile-1(Q1) and quartile-2 (Q2) of each month are calculated
4. Winter-Holt's and Seasonal Arima Methods. The approach is based on exponential smoothing and seasonality forecasting using training data
5. Time Series Decomposition. The approach is based on calculating the trend and seasonality of the training data to make the forecasting.

The error of the model or approach forecasting of monthly rainfall is evaluated using RSME/ root mean squared error, which is calculated as follows:

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{p,i}-P_{a,i}\right)^2}$$

where $P_{p,i}$ is the predicted monthly rainfall and $P_{a,i}$ is the actual monthly rainfall. The error value is expected to be minimum for a good forecasting model.

A hypothetical void of a coal mine is used with a specific stage-volume curve and the catchment area. The void started to be filled, at the same time as the start of the forecasting time. The majority of water input is from runoff, thus the runoff volume of water into the void to form the pit lake is roughly estimated using the following equation

$$\textbf{Water Input at n-th month, } Vn = \sum_{i=1}^{n}(\mathrm{C} \times P_{p,i} \times \mathrm{A})$$

Where C is the runoff coefficient (0.7), $P_{p,I}$ is the predicted monthly rainfall and A is the catchment area of the mine void. The pit flooding volume curve thus is developed to show different increments of volume water inside the void using various methods and approaches. The actual monthly rainfall depths are also used to estimate the actual flooding duration. Pit lake flooding duration is the time in the n-th month that $Vn \geq$ maximum volume of void.

Data

Monthly rainfall datasets of 26 years, ranging from January 1993 to December 2018, are used in this study. The first 15-year dataset (from January 1993 to December 2007) is used as input for rainfall forecasting methods whilst the last 11-year dataset (from January 2008 to December 2018) is used to verify the methods. The time series of monthly rainfall datasets and boxplots of monthly rainfall are shown in Fig. 1 and Fig. 2. The RSME is evaluated for verification dataset from January 2008 to December 2018

A hypothetical void of a coal mine is used with a specific stage-volume curve and the catchment area as shown in Fig. 3 and Fig. 4. The void is characterized as a coal mine void with an estimated catchment area of 440 ha. The longest dimension of the void is 1800 m, whilst the shortest dimension of the void is 730 m. The void is estimated to be filled with a maximum of 48×10^6 m^3 of water.

Results and Discussion

Models and approaches for the prediction of monthly rainfall

The predicted result of monthly rainfall for each model and approach vs. the actual monthly rainfall depth from January 2008 to December 2018 is shown in Fig. 5 to Fig. 8.

The approaches using monthly averages (overall) and linear regression show no seasonality and difference for each month and year of prediction. This method simplifies the expected value of precipitation but does not capture the seasonality or wet and dry

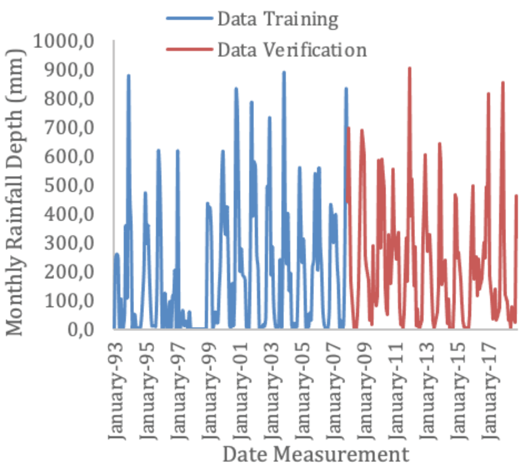

Figure 1 Time series of monthly rainfall datasets (Data Training and Verification)

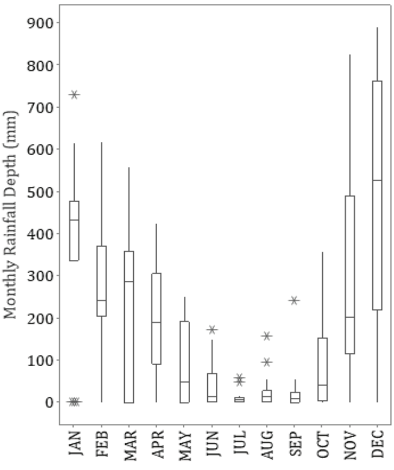

Figure 2 Boxplot of Monthly Rainfall

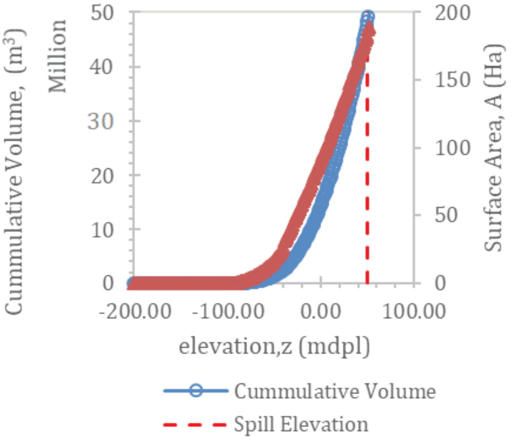

Figure 3 Staging curve (elevation, surface area and volume) of void

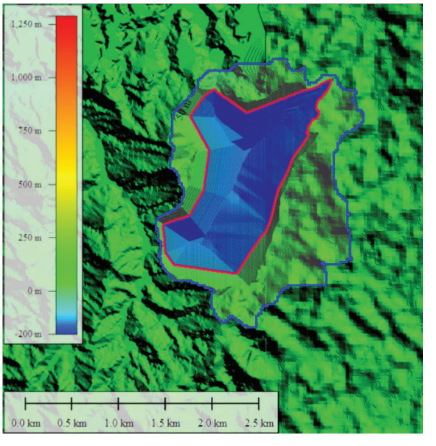

Figure 4 Catchment and Elevation map of void

seasons. The linear regression approach shows the trend of increment in the rainfall depth due to increasing yearly rainfall depths.

The approaches using a statistical approach based on seasonality show the ability to capture the interchanging wet and dry seasons, yet the prediction is bound and limited to the statistical parameter selected and does not accommodate the yearly trend.

Models using Winter-Holt and seasonal ARIMA approaches show the ability to capture the seasonal value of rainfall and yearly trends, but the accommodated yearly trend tends to be exaggerated over time. On the other hand, time series decomposition additive methods show a less steep yearly-trend.

RSME for and approaches for the prediction of monthly rainfall

The error resulting from the difference between predicted and actual monthly rainfall is summarized in Fig. 9. Linear approaches, including monthly average using all data (overall), and linear regression produced the lowest RSME (257.09 and 162.45 mm, respectively). The next to the lowest RSME values were produced by time

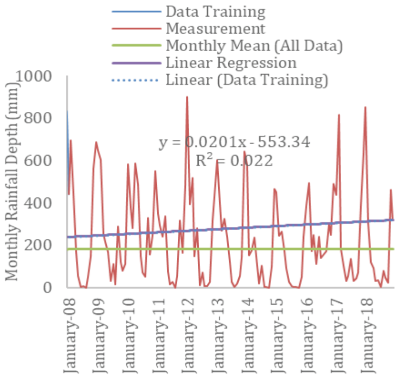

Figure 5 Actual values vs prediction of monthly rainfall using overall monthly average and linear regression for January 2008 – December 2019

Figure 6 Actual values vs prediction of monthly rainfall using statistical approach (mean seasonality, Q1-seasonality, Q2-seasonality) for January 2008 – December 2019

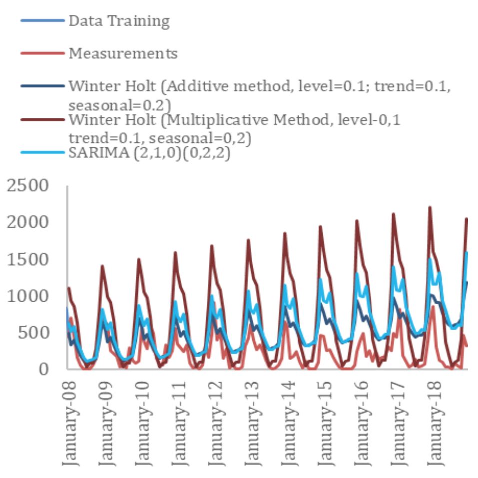

Figure 7 Actual values vs prediction of monthly rainfall using the Winter-Holt and seasonal ARIMA approaches for January 2008 – December 2019

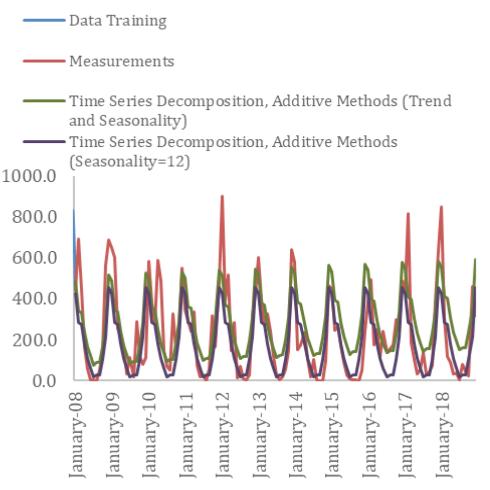

Figure 8 Actual values vs prediction of monthly rainfall using time series decomposition additive methods for January 2008 – December 2019

series decomposition methods followed by statistical approaches. The Winter-Holt and seasonal ARIMA methods produced varying RSME values, which indicates the models are more volatile than the others. The simple models or approaches produced the lowest RSME but could not explain the seasonal nature of the monthly rainfall depths.

Pit Flooding Volume Curve and Estimated Pit Lake Flooding Duration

The pit flooding volume curve shows the changing volume of the water inside the void based on above mentioned assumptions. The curves for each method and approach are shown in Fig. 10. The pit flooding duration for each method and approach are summarised in Table 1

The estimated pit lake flooding duration

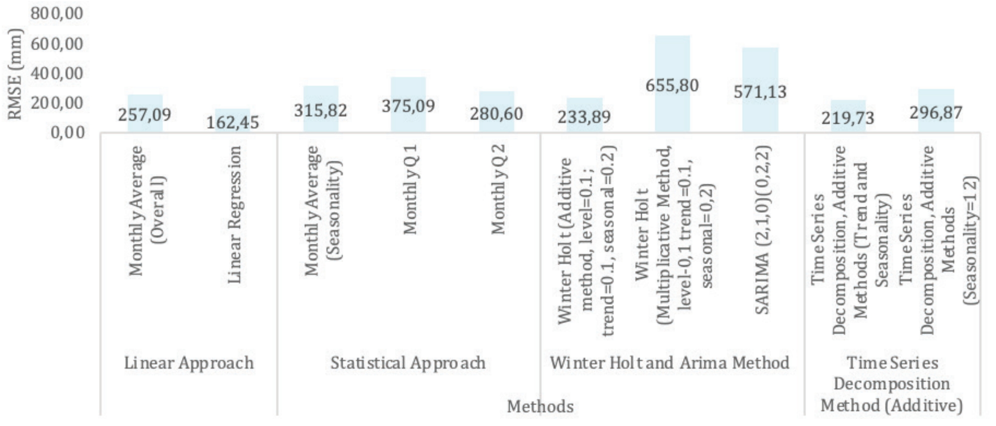

Figure 9 *Root Mean Square Error (RMSE) value of various methods in predicting monthly rainfall depths.*

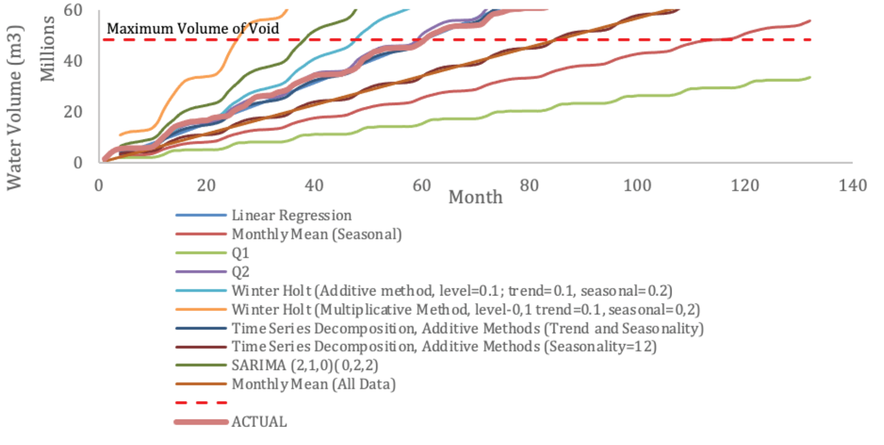

Figure 10 *Pit Flooding Volume Curve*

Table 1 *Estimated pit lake flooding duration*

Methods/Approach	Estimated Time to Flood (months)
Actual	61
Monthly Mean (All Data)	86
Linear Regression	61
Monthly Mean (Seasonal)	116
Statistical Approach Seasonal Q1	>132
Statistical Approach Seasonal Q2	60
Winter-Holt (Additive method, level=0.1; trend=0.1, seasonal=0.2)	48
Winter-Holt (Multiplicative Method, level-0,1 trend=0.1, seasonal=0.2)	26
Seasonal ARIMA (2,1,0)(0,2,2)	39
Time Series Decomposition, Additive Methods (Trend and Seasonality)	61
Time Series Decomposition, Additive Methods (Seasonality=12)	85

using actual rainfall is 61 months. Methods using linear regression, statistical approach (Q2) and time series decomposition, additive methods (trend and seasonality) show the most accurate prediction of flooding (60 – 61 months). Monthly means (seasonal) and monthly means using all data resulted in higher estimated flooding duration due to several low monthly rainfall depths in the training datasets. The same reason affected the predicted time to flood using the statistical approach of quartile-1 (>132 months). The Winter-Holt and seasonal ARIMA methods predicted shorter estimated time to flood due to overestimating the value of the monthly rainfall in the later prediction months.

Conclusions

The monthly precipitation series recorded in the stations of the study area shows seasonal behaviour, reflecting the typical unimodal rain regime from the basin.

The simple models or approaches produced the lowest RSME but could not explain the seasonal nature of the monthly rainfall depths. The next to the lowest RSME values were produced by the time series decomposition methods followed by statistical approaches. The Winter-Holt and seasonal Arima methods produced varying RSME values, which indicates those models are more volatile than the others.

The estimated pit lake flooding duration using actual rainfall was 61 months. Methods using linear regression, statistical approach (Q2) and time series decomposition, additive methods (trend and seasonality) predicted flooding most accurately. Based on the ability to explain the seasonality of the monthly rainfall data, yearly trend, and predictive accuracy, time series decomposition using additive methods (trend and seasonality) were the best option to predict monthly rainfall depth and to calculate the flooding duration.

Different methods show substantial differences in flooding duration. In addition, the methods differ in their abilities to portray uncertainties in mid-term and long-term hydrometeorological conditions, such as climate changes and climate pattern.

Acknowledgements

The authors thank all parties who helped during data acquisition and analysis.

References

Latif, S. D., Alyaa Binti Hazrin, N., Hoon Koo, C., Lin Ng, J., Chaplot, B., Feng Huang, Y., El-Shafie, A., & Najah Ahmed, A. (2023). Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. Alexandria Engineering Journal, 82, 16–25. https://doi.org/10.1016/j.aej.2023.09.060

Nwokike, C. C., Offorha, B. C., Obubu, M., Ugoala, C. B., & Ukomah, H. I. (2020). Comparing SANN and SARIMA for forecasting the frequency of monthly rainfall in Umuahia. Scientific African, 10, e00621. https://doi.org/10.1016/j.sciaf.2020.e00621

Maia, A. L. S., & de Carvalho, F. D. A. (2011). Holt's exponential smoothing and neural network models for forecasting interval-valued time series. International Journal of Forecasting, 27(3), 740–759. https://doi.org/10.1016/j.ijforecast.2010.02.012