



# Optimal Backfilling Materials with High Compressive Strength Based on Multiple Linear Regression

Ehab Gomaa<sup>1</sup> · K. A. El-Nagdy<sup>1</sup>

Received: 14 March 2023 / Accepted: 15 September 2023 / Published online: 3 October 2023  
© The Author(s) 2023

## Abstract

Backfilling material such as tailing (mine wastes) mixing with cement or gypsum has grown throughout the world's underground mines. However, despite their growing popularity, the typical hydraulic and mechanical fill types utilized in many mines still exist. Deep underground mining has increased due to the lack of commercial minerals nearby. Mine wastes were considered the main part of backfilling to prevent environmental pollution, ground subsidence after mine abandonment, and mine collapse during deeper extraction phases. The cemented backfill technique is the principal technique used in underground mines, which include cement with fly ash and/or filter dust, cement with tailing material and fly ash, gypsum with fly ash, and synthetic anhydrite with fly ash and have been reviewed. It has concluded that a backfilling material must be selected based on further goals, available material near the mine site, and economic factors. This paper analyzes different backfill material mixtures to create a technique that will increase safety in underground mining conditions and foresees an appropriate formula that gives high uniaxial compressive strength. The multiple linear regression (MLR) on the collected data from the experimental works to construct the relationship between the uniaxial compressive strength (UCS) of the mixture and the components of the backfilling and the prediction formula for expected compressive strength was obtained. The results revealed that the predicted regression equation was robust and reliable to predict the (UCS) for the new components of the filling (cement (CE), filter dust (FD), water content (WC), and time (T)).

**Keywords** Backfilling materials of mine · Multiple linear regression · Uniaxial compressive strength

## 1 Introduction

Backfilling process in the mines is filling the unused mine voids and openings with waste or mixtures consisting of those wastes with hardness or bearing strength greater than the loads and weights resulting from the upper layers, to achieve safety and a suitable environment for the underground mining process. Backfilling was also used for economic and environmental reasons, to mitigate the impacts of subsidence on the surface and avoid fires and explosions, enhance mine ventilation, strengthen the ground stability of rocks, and improve ground stability. [1, 2].

The most common backfill materials in mines consist of three basic elements. The majority of the backfill is an inert material, typically made up of binding and chemical

material. The most important inert materials used are tailing from the concentration process, sand-gravel, mining tailings, and industrial slag. As for the bonding material must have an adhesion quality, such as cement, slag, gypsum, or fly ash. Chemical compounds are used to improve the permeability of the mixture, slurry flowability, and compacting properties of backfill such as flocculants, accelerators, and retarders. Most of the raw materials used in backfilling technology come from mining-related sectors, including overburden, mining, and mineral processing tailings [3–6].

Carefully designed and effectively operated backfill systems can greatly improve the mining process, achieve mining goals, and improve work performance. By contrast, a poorly designed and executed backfill can greatly impede the operation of a mine and, worst of all, jeopardize safety. Backfill is getting stronger over time, until it reaches its final strength. The physical and mechanical properties of backfill materials and the technique of monitoring and ensuring performance are necessary to control and perform the filling [3]. The physical, chemical, and mechanical properties of

✉ Ehab Gomaa  
ehabgomaa7@yahoo.com

<sup>1</sup> Mining Engineering Dept., Faculty of Petroleum and Mining Engineering, Suez University, Suez, Egypt

different mixtures were studied and their impact on obtaining the best, cheapest, and strongest suitable mixture of backfill materials for the selection and systematic application of backfill in underground mines [7, 8].

Among the most important of these effective factors for advancing backfill technology is the work of backfill materials and the interactions and influence of each component on the other [5, 6]. The study emphasized waste and readily available, inexpensive resources like gypsum and natural and synthetic anhydrite. There are opportunities to lessen the number of surface tailings and their environmental effects through the use of backfill as part of the mining cycle, the advancement of existing procedures, selective mining, the use of novel processing techniques, and the use of backfill materials in beneficial deep geological formations. Currently, cemented backfilling technology is widely used all over the world. Backfill placement, application, and system selection in mines include many disciplines.

Paste, hydraulic, and dump backfill technologies are thoroughly examined, as are combined backfilling methods [9]. There is a good probability that one of the several material combinations studied can be employed as an underground mine backfill that is both technically and economically feasible. The test work may be necessary to verify the compatibility of the materials on hand for a certain system [6]. The main deterrent to backfilling in certain underground mines is the cost. Regardless of the mineral deposit's geometry, depth, and mining technique.

Thanks to technology, backfill can be buried using dry, slurry, or paste materials. Another way to improve backfill placement, remove some challenges with underground tailings disposal, and lower backfill placement costs is to update or modify mining procedures [2]. When relevant to successful backfill operations, design and planning considerations, material evaluation, and economic and other backfilling techniques are described. The components

could be extracted from mine waste, tailing, or local resources. The highest possible usage of resources from mine and industrial waste should be given priority when choosing the materials [10–13].

The backfill mixture works to maintain the stability of the mine, and this is generally done by increasing the proportion of binders to achieve a higher fill strength, which may lead to an increase in the total costs. Therefore, the required strength can usually be achieved by changing the proportion of materials in the mixture used, so testing is necessary to determine the optimal mixture of materials available. To complete this study, samples of backfill materials were taken from different sources (Germany gypsum mine, Thailand potash mine, and China coal mine).

There are 7 different mixtures used in this work (synthetic anhydrite and fly ash (SA + FA); natural anhydrite and fly ash (NA + FA); cement and fly ash (C + FA); cement and filter dust (C + FD); cement, fly ash, and filter dust (C + FA + FD); gypsum and fly ash (G + FA); tailing, cement, and fly ash (T + C + FA)). The result of the uniaxial compressive strength of the last one (T + C + FA), as shown in Table 1. However, all the results are illustrated on the flowing drive

[https://drive.google.com/drive/folders/1lwDAcDKrisrIB8LptUC2Ajxb-Fu0BAAtY?usp=share\\_link](https://drive.google.com/drive/folders/1lwDAcDKrisrIB8LptUC2Ajxb-Fu0BAAtY?usp=share_link)

Any feasibility analysis must show that underground backfill is more expensive than storing it on the surface [10]. Operating and capital expenses are included in the price of backfill placement. Backfilling technology's economic and technical viability is directly impacted by two important factors: transportation and geotechnical concerns [11, 12]. It enables engineers to maximize the benefits of the backfill design while lowering costs. This study will investigate backfill methods and materials suitable for underground mining. Consequently, the selection of backfill materials is the focus of this study [13, 14].

**Table 1** Uniaxial compressive strength (UCS) of the mixture of tailing, cement, and fly ash

No	Cement %	Tailing %	Fly ash %	7 days UCS (MPa)	28 days	90 days
T1	5	47.5	47.5	1.25	2.4	3.5
T2	5	50	45	1.1	2.3	3
T3	5	45	50	1.3	2.3	3.6
T4	10	45	45	1.8	4	5.1
T5	10	50	40	1.8	3.8	5
T6	10	40	50	2.1	4.2	5.3
T7	15	42.5	42.5	2.9	5.9	7.2
T8	15	50	35	2.5	5.7	6.9
T9	15	35	50	3.2	5.6	7.4
T10	20	40	40	4.1	8	9.1
T11	20	50	30	3.9	8	8.7
T12	20	30	50	4.2	8.2	9.4

The study aims to enhance the statistical analysis of a selection of backfill materials for underground mines. Based on the engineering qualities and compaction characteristics of the available materials, backfill materials should be chosen. Using (MLR) on the data gathered from experimental works of the literature review to study the relationship between the (UCS) of the mixture and the components of the filling, backfill materials should be chosen based on compaction characteristics of the materials available [1, 2]. This study assesses backfill's technical, long-term safety, environmental, and economic aspects [15].

## 2 Methodology

This study examines backfill methods and materials appropriate for the systematic selection and use of backfill in underground mines. Several backfill materials and mixes underwent physical, chemical, and mechanical testing in the lab. Manoon Masniyom's Ph.D. thesis, published in Germany's Technical University Bergakademie, contained the information gathered during the experimental study [1, 2].

Materials produced as byproducts and other easily accessible, affordable resources, such as fly ash, filter dust, and tailings, were considered when cement, gypsum, or synthetic anhydrite was used as a cement ingredient. There is a good probability that one of the several material combinations studied can be employed as an underground mine backfill that is both technically and economically feasible. Yet, the test work may be necessary to verify the compatibility of the readily available materials for a certain system.

Analyses typically concentrate on lowering approximation error to ensure accurate forecasts. Due to their inherent overfitting qualities, pattern detection models like (MLR) may not be the best match for predicting evolutionary trends. Still, they can provide adequate results when used to predict potential outcomes. In contrast, regression models are more frequently used to forecast trending behaviors; as a result, they might be more useful when examining the development of a given set of data, symbols, or series. A simple regression model, however, has an unsatisfactory prediction error rate and can be quite wrong. This paper shows a mathematical method to predict the suitable backfill material that could be chosen using a linear regression technique [15].

A statistical method known as MLR uses many explanatory variables to forecast the result of a response variable. It is the purpose of MLR to model the connection among the explanatory and response (independent and dependent) variables [16]. Several explanatory factors are included in a multiple regression model, which is based on various assumptions (such as a linear connection between dependent and independent variables). The independent variables do not have a lot of correlation among themselves. The  $y_i$

observations are chosen randomly and independently from the population; residuals should have a mean and variance normally distributed with a mean of 0 and variance  $\sigma$ .

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 X_{i2} + \dots \dots \dots + \beta_p X_{ip} + \varepsilon \quad (1)$$

where  $i = n$  observations,  $y_i$  = dependent variable,  $X_i$  = explanatory variables,  $\beta_0$  =  $y$ -intercept (constant term),  $\beta$  = slope coefficients for each explanatory variable,  $\varepsilon$  = the model's error term (also known as the residuals).

A statistical metric known as the coefficient of determination ( $R$ -squared) is used to determine how much variation in the independent factors can be explained in terms of the result. Even though the variables may not be connected to the outcome variable,  $R^2$  always rises when more predictors are included in the MLR model.  $R^2$  can only have a value between 0 and 1, with 0 indicating that no independent variable can predict the outcome and 1 indicating that it is possible to anticipate the event without error from the independent variables [17]. The assumption behind multiple regressions is that the dependent and independent variables have a linear relationship. Moreover, no discernible association between independent variables is a presumption.

Several multilinear regression equations were developed to predict the relationship between UCS and one or more combinations of the different backfilling materials. One of the trickiest methods is the capacity to create precise linear models to discover patterns or correlations between variables. In mining science, extrapolative applications are frequently employed to forecast values [16].

A common method for forecasting the future values of a revolutionary object is data analysis. Linear patterns are typically investigated first when predicting business expansion, corporation profit, stock values, and other patterns. A technique that accurately forecasts trend behaviors can be a very useful tool, whether used in fields like econometrics, biology, mathematics, or business intelligence. When the average daily prices of their stock values are given, the writers of this article analyze the trend of a stock market that was arbitrarily chosen [15]. Finding consistent and predictable trends in the evolution of a thing is one of the most popular techniques for accurately predicting its behavior.

## 3 Results and Discussion

Multiple linear regression (MLR) is applied to the collected data from the experimental works. The MLR is used to construct the relationship between the compressive strength and the filling components, filling water content, and the time the compressive strength is measured. The components of each proposed filling are different according to its material and the concentrations of these components. Six filling compounds

**Table 2** Statistical coefficients of the applied data on the first filling experimental results

Regression statistics	
Multiple <i>R</i>	0.956886869
<i>R</i> -square	0.91563248
Adjusted <i>R</i> -square	0.910570429
Standard error	0.833827055
Observations	54

are proposed, the first consists of cement and filter dust, the cement, fly ash, and filter dust constitute the second compound, and gypsum replaces the cement and filter dust in the first compound with fly ash to constitute the third compound.

The fourth compound is a mixture of synthetic anhydrite and fly ash; the fifth is the third compound, but the cement concentration is increased up to 30% of the full filling compound. The six backfilling materials are a mixture of cement, fly ash, and filter dust with different concentrations. The objective of using MLR is to select the best compound with a linear relationship between the filling compressive strength and the compound's components. Therefore, the correlation coefficient of the relation is higher than the other to construct the prediction equation relating the response (filling compressive strength) and the input variables, which is the components of each compound.

The MLR is then applied to the results of the measurements, and Table 2 illustrates the correlation coefficients (multiple *R*) and the determination coefficient (*R*-squared) of each compound. The multiple *R* coefficient indicates the strongest or weakest relationship between two or more variables, whereas +1 or −1 refers to a strong relationship between the variables. The +1 indicates the strong direct proportion between the variables, but the −1 refers to the strong inverse proportion between the variables. On the other hand, if multiple *R* is zero, then there is no relation between the variables. The determination coefficient (*R*-squared) measures the variation of one variable when the other variable is changed and refers to how strong the linear relationship between the variables. The adjusted *R*-squared coefficient compares the explanatory power of the regression model containing several predictors. When adding a new term, the adjusted *R*-squared improves the prediction more than the expectation, but if the new term improves the regression model with less than the expectation, the

**Table 4** MLR regression equation of the results of the first filling compound

	Coefficients	Standard error	<i>t</i> Stat	<i>P</i> -value
Intercept	−0.18334788	1.27390005	−0.14393	0.886137101
CE	0.268812698	0.013288186	20.22945	9.94898E−26
WC	−1.35	2.779423517	−0.48571	0.629292983
<i>T</i>	0.037163206	0.003220298	11.5403	1.04425E−15

adjusted *R*-squared decreases. Another interpretation of the adjusted *R*-squared is that it compares the goodness of fit of the regression models containing different independent variables and increases when the new term improves the model fit more than that expected by chance. The standard error indicates the accuracy of the sample mean compared with the population's mean. The smaller the standard error refers to the sample mean is closest to the population mean and the standard error for a good model is very small. The number of observations indicates the number of samples used to construct the regression model.

### 3.1 First Filling Compound

The first filling compound was installed from cement and filter dust. The cement percentage varies from 0 to 25 % of the filter dust. The water content ranges from 40 to 50 % of the compound's total volume. The uniaxial compressive strength (UCS) was measured for different concentrations of the components. The MLR was applied to the 54 data samples and the regression statistics were developed. Regression statistics measure the strength of the relationship between two or more variables. These statistics include the coefficient of determination (*R*-squared), the *F*-statistic, the *t*-statistic, and the *P*-value. The coefficient of determination measures how well a model fits a given data set. In contrast, the *F*-statistic and *t*-statistic measure how significant the model is in explaining the variation in the data. The *P*-value is used to assess the statistical significance of a certain model.

Table 2 shows several statistical values that affect how well a model may fit a particular data collection. By dividing the covariance of two variables by the product of standard deviations, the multiple *R*, which describes the strength of the link between the two variables, comes out at 0.956, indicating a stronger linear relationship. The *R*-squared number

**Table 3** ANOVA statistics and source of variance of the results of the first filling compound

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Significance <i>F</i>
Regression	3	377.2835554	125.7612	180.8817105	7.83E−27
Residual	50	34.7633779	0.695268		
Total	53	412.0469333			

**Table 5** Statistical coefficients of the applied data on the second filling experimental results

Regression statistics	
Multiple <i>R</i>	0.91231627
<i>R</i> -square	0.832320976
Adjusted <i>R</i> -square	0.829142227
Standard error	1.061710235
Observations	216

describes how closely the data resemble the fitted regression line. For multiple regression, it is sometimes referred to as the coefficient of determination or the coefficient of multiple determination.

$$R - \text{squared} = \text{explained variation} / \text{total variation} \quad (2)$$

It runs from 0 to 100%, with 0% signifying that the model does not explain any of the response data variability around the mean. On the other hand, 100% means that the model explains all of the variability in the response data around the mean. According to Table 2, the *R*-squared value of 91% indicates that the response data are highly variable around the mean. The model matches the observed data fairly well, as indicated by the adjusted *R*-squared of 0.91. The adjusted *R*-squared rise only when the additional term enhances the model more than would be predicted by chance. The standard deviation of a sample is multiplied by the square root of the sample size to determine the standard error of the mean (SEM). Here, the number 0.833 denotes the standard deviation from the population mean.

ANOVA stands for analysis of variance is shown in Table 3. To compare the means of two or more groups, one uses statistics. It assesses whether the means of two or more groups differ significantly. ANOVA can be used to compare the means of different populations, such as comparing the mean scores of students in different classes or the mean salaries of employees in different departments. ANOVA can also be used to compare the means of different treatments, such as comparing the effectiveness of different drugs on a particular disease. The parameters of ANOVA statistics are a degree of freedom (*df*), a variance of the sum of squares (*SS*), variance mean square (*MS*), *F*-test, and significance *F*. The degree of freed of

regression which is 3 refers to the independent variables in the regression minus one. The degree of free within treatments or groups (residual) is 50. The variance sum of squares (*SS*) assesses the variability between different groups in an experiment. It is calculated by summing the squared differences between each group’s mean and the overall mean. It can be used to compare the means of several groups and to assess whether there is a statistically significant difference between two or more groups.

The calculation of the *SS* can be determined as follows:

$$SST = \sum Y_{ij} - \frac{(Y_{..})^2}{N} \quad (3)$$

where *SST* refers to the total sum of squares, *Y<sub>ij</sub>* refers to the observation *j* in sample *i*, where *i* varies from 1 to several samples *m* and *j* varies from 1 to several variables *n*. *N* refers to the product of *m* × *n*.

The term *Y<sub>..</sub>* refers to the sum of observations in all samples where

$$Y_{..} = \sum_{i=1}^m \sum_{j=1}^n Y_{ij} \quad (4)$$

The total sum of squares between groups (*SS<sub>t</sub>*) can be calculated as follows:

$$SS_t = \frac{\sum Y_i^2}{m} - \frac{(Y_{..})^2}{N} \quad (5)$$

The sum of squares of error (*SS<sub>e</sub>*) can be computed from the following equation:

$$SS_e = SST - SS_t \quad (6)$$

From the data samples, *SS<sub>t</sub>* is 377.28, *SS<sub>e</sub>* is equal 34.76, and finally, the *SST* is 412.04.

ANOVA MS stands for analysis of variance mean square. A chi-square test is a statistical method for comparing the means of two or more groups. Determine whether there is a statistically significant difference between the group means using this method. ANOVA MS can be used to compare the means of different treatments, products, or services to determine which has the greatest impact on a given outcome.

The variance means square (*MS*) can be computed as follows:

**Table 6** ANOVA statistics and source of variance of the results of the first filling compound

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Significance <i>F</i>
Regression	4	1180.610293	295.1525733	261.83914	1.35197E–80
Residual	211	237.8452396	1.127228623		
Total	215	1418.455533			

**Table 7** MLR regression equation of the results of the first filling compound

	Coefficients	Standard error	<i>t</i> Stat	<i>P</i> -value
Intercept	0.742569511	1.044110873	0.711197949	0.477747044
Cement	0.314609259	0.014448046	21.7752112	7.14611E–56
Filter dust	–0.018166667	0.012922725	–1.405792202	0.161256653
Water content	–2.652777778	1.769517059	–.499153549	0.135328473
Time	0.041801665	0.002050199	20.38907538	8.87854E–52

**Table 8** The statistics regression output of the five-regression model for the studied compound

	Compound 1	Compound 2	Compound 3	Compound 4	Compound 5	Compound 6
Multiple <i>R</i>	0.96	0.91	0.78	0.73	0.92	0.91
<i>R</i> -square	0.92	0.83	0.61	0.53	0.85	0.83
Adjusted <i>R</i> <sup>2</sup>	0.91	0.83	0.58	0.52	0.84	0.83
Standard error	0.83	1.06	2.34	2.56	0.98	1.58
Observations	54	216	54	90	108	162

$$MSt = \frac{SS_t}{df_{regression}} \quad (7)$$

$$MSe = \frac{SSe}{df_{residual}} \quad (8)$$

A statistical test, the ANOVA *F*-test, assesses the differences between two or more groups. It assesses whether the means of multiple separate groups differ significantly. The ANOVA *F*-test is based on *F*-distribution, a probability distribution that compares the variability between groups to the variability within each group. The test statistic for ANOVA *F*-test is calculated by dividing the variance between groups by the variance within each group. If the ratio of these two variances is statistically significant, then it can be concluded that there is a significant difference between the means of two or more groups.

The value of the *F*-test or *F* can be determined as follows:

$$F - test = \frac{MSt}{MSe} \quad (9)$$

Therefore, the value of the *F*-test is 180.88. Now, the tabulated *F*-test is calculated for the specified significance level (0.05) with the degree of freedom ( $df_{regression}$ ,  $df_{residual}$ ), which is (3, 50) in our case, so the tabulated *F*-test is 2.8. Therefore, the calculated *F*-test is greater than the tabulated *F*-test. Hence, the null hypothesis ( $H_0$ ) is rejected, which says that the means are equal and accept the substitute hypothesis ( $H_1$ ) referring to significant differences between the data groups.

The regression develops an equation that links the input variables (ratio between the cement to filter dust, the water content, and the time needed for the sample to run the

measurement) and the output response uniaxial compressive strength (UCS). The parameters of the equation can be extracted from Table 4 as follows:

$$UCS = -0.1833 + 0.2688 \times CE - 1.35 \times WC + 0.037 \times T \quad (10)$$

The *P*-values in Table 4 indicated that the CE variable and *T* variable have values less than a significant value of 0.05, so the null hypothesis can be rejected, and the substituted hypothesis will be considered, referring that it is significantly different between the data samples. Equation 10 can be approximated to

$$UCS = -0.1833 + 0.2688 \times CE + 0.037 \times T \quad (11)$$

### 3.2 Second Filling Compound

The cement, fly ash, and filter dust constitute the second compound. A total number of 216 data samples were collected from the experimental works. The cement ranges from 5 to 20 %, the fly ash ranges from 30 to 40 % and the rest is for the filter dust. The water content is 0.4 to 0.5, and the time varies from 7, 28, and 90 days. Table 5 illustrates several statistical values of the MLR that were applied to

**Table 9** Statistical parameters of the fifth compound using MLR

	Coefficients	Standard error	<i>t</i> Stat	<i>P</i> -value
Intercept	1.44	1.36	1.056	0.293
CE	0.32	0.019	16.249	3.49005E–30
FD	–0.02	0.017	–1.340	0.183
WC	–3.19	2.31	–1.380	0.171
T	0.04	0.0027	14.723	4.3363E–27

**Table 10** Statistical parameters of the fifth compound using MLR considering the effective parameters based on Table 8

	Coefficients	Standard error	t Stat	P-value
Intercept	-1.126	0.258628	-4.353	3.13E-05
CE	0.318	0.017009	18.6866	3.58E-35
T	0.039	0.002699	14.60612	4.85E-27

the data samples of the experimental run results. Table 6 shows the ANOVA statistics parameters, and the calculated *F*-test is greater than the tabulated *F*-test (2.45 at *df*<sub>residual</sub> equal to 120); hence, the null hypothesis will be rejected and accepted with the substituted hypothesis.

The regression model results in Table 7 indicate that the first compound, which consists of cement and filter dust is the strongest regression model. The multiple *R* is 0.96, which refers to the strong directly proportional between the input variables (cement concentration, water content, and the span time at which the compressive strength is measured). The *R*-squared for the first regression model is 0.92, which indicates the strongest relationship between the compressive strength and the input variables and refers to the fact that any change in one of the input variables significantly influences the output response (compressive strength). The fitness of the regression model is so good because the adjusted *R*-squared is 0.91, where the improvement that occurred is more than that expected by chance. The results also show that the standard error of the first regression model is less

than that of the other model (0.83). The results in Table 8 indicate that the fifth compound's regression model is also significant, and the statistics parameters may decrease with the number of samples increase.

### 3.3 The Comparison Results Among the Statistical Results of the Five Filling Compound

Therefore, based on the fifth compound, a prediction model is constructed that links the compressive strength and the input variables of the filling (components concentration, water content, and the elapsed time at which the compressive strength can be measured). However, its statistical parameters are less than the first compound, but the number of observations is more than that for the first compound. In addition, compound 5's components are three components, but only two components for compound 1.

The form of the prediction equation that is constructed to relate the response and input variables should be in the following form:

$$\text{response} = c_0 + c_1 \cdot x_1 + c_2 \cdot x_2 + \dots + c_n \cdot x_n \tag{12}$$

where the response refers to the estimating variable (output); *x*<sub>1</sub>, *x*<sub>2</sub>, ..., and *x*<sub>*n*</sub> refer to the input variables that influence the response; *n* is the number of the variables; and *c*<sub>0</sub> is the intercept coefficient when each *x* is zero and *c*<sub>1</sub>, *c*<sub>2</sub>, ..., and *c*<sub>*n*</sub> refer to the coefficient of each term.

**Table 11** The percentage errors between the actual and predicted UCS based on Eq. (13)

	Cement (CE)	Filter dust (FD)	Water content (WC)	Time (T)	UCS	Predicted UCS	Errors %
1	15	42.5	0.4	90	7.2	7.348649056	2.0645702
2	15	50	0.4	90	6.9	7.179037945	4.0440282
3	15	35	0.4	90	7.4	7.518260167	1.5981104
4	20	40	0.4	90	9.1	8.937871279	1.7816343
5	20	50	0.4	90	8.7	8.71172313	0.1347486
6	20	30	0.4	90	9.4	9.164019427	2.5104316
7	5	47.5	0.45	90	3.9	4.010760167	2.8400043
8	10	45	0.45	90	5.5	5.59998239	1.8178616
9	10	40	0.45	90	5.6	5.713056464	2.0188654
10	15	42.5	0.45	90	7.5	7.189204612	4.1439385
11	15	50	0.45	90	7.1	7.019593501	1.1324859
12	15	35	0.45	90	7.7	7.358815723	4.4309646
13	20	40	0.45	90	9.4	8.778426834	6.6124805
14	20	50	0.45	90	9.1	8.552278686	6.0189155
15	15	42.5	0.5	90	7	7.029760167	0.4251452
16	15	50	0.5	90	6.5	6.860149056	5.5407547
17	15	35	0.5	90	7.2	7.199371279	0.0087322
18	20	40	0.5	90	8.8	8.61898239	2.0570183
19	20	50	0.5	90	8.5	8.392834242	1.2607736
20	20	30	0.5	90	9	8.845130538	1.7207718

The prediction equation that links the compressive strength and the components of the filling (cement (CE), filter dust (FD), water content (WC), and time (T)) can be built based on the results and statistical parameters in Table 2. The results of the MLR to construct the prediction equation are illustrated by Eq. (13):

$$\text{UCS} = 1.44 + 0.32 \times \text{CE} - 0.02 \times \text{FD} - 3.19 \times \text{WC} + 0.04 \times T \quad (13)$$

The effective or significant terms can be evaluated based on the statistical parameters, especially the *P*-value shown in Table 8. The *P*-value is a statistical parameter that indicates each term's significance by rejecting or confirming the null hypothesis. When the *P*-value is less than the significant limit (0.05), then the null hypothesis can be rejected, and confirm the alternative hypothesis, else the null hypothesis is confirmed. The *P*-value in Table 9 refers to the effective terms of Eq. (13) CE and *T* because their *P*-values are less than 0.05, and the non-significant terms are FD and WC because their *P*-values are greater than 0.05 (0.183 and 0.171). Therefore, MLR can be carried out on the CE, *T*, and UCS. Hence, Eq. (13) can be approximated to Eq. (14) based on the new MLR results as in Table 10.

$$\text{UCS} = -1.126 + 0.318 \times \text{CE} + 0.039 \times T \quad (14)$$

Therefore, Eq. (14) can predict the compressive strength (UCS) when the data of CE and *T* are known. The percentage error between the actual and predicted UCS is shown in Table 11 for 20 randomly selected cement and flay ash values with the filter dust and water content. The results indicated the ability of the model to predict the UCS well where the error percentage did not exceed 6 %.

## 4 Conclusions

This study used five backfilling compounds to prevent mine collapse in later and deeper extraction phases, ground subsidence after mine abandonment, and environmental degradation. MLR was applied to the results of experimental runs performed on the uniaxial compressive strength of the backfilling materials to support the mine. Six backfilling materials were selected to choose the best one based on the results of the experimental runs, and the regression applied to that data. The MLR results indicated that the five backfilling material compound is the best where statistics parameters decreased with the number of samples increased. Therefore, a prediction model can be installed based on Eq. (13) to predict the UCS value at new component data, such as concentration, water content, and the elapsed time at which the

compressive strength can be measured. May the fifth compound is not the best from the economic concept where the cement quantity is higher than in the first compound, but the prediction results are more accurate. The percentage error did not exceed 6% between the predicted and actual UCS, referring to the constructed equation's robustness.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

## Declarations

**Consent for Publication** It is an original work and has not published or sent for publication elsewhere.

**Conflict of Interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Masniyom M, Drebenstedt C (2009) Systematic selection and application of backfill in underground mines. Doctoral thesis (Technical University Bergakademie Freiberg)
2. Lamos AW, Clark IH (1989) The influence of material composition and sample geometry on the strength of cemented backfill. In: Proceedings of the Innovations in Mining Backfill Technology. Rotterdam, The Netherlands, pp 89–94
3. Potvin Y, Thomas EG, Fourie AB (2005) Handbook on mine fill. Perth, Australia, ACG
4. Wang M, Liu L, Chen L (2018) Cold load and storage functional backfill for cooling deep mine. Adv Civ Eng
5. Zhang J, Deng H, Taheri A, Deng J, Ke B (2018) Effects of superplasticizer on the hydration, consistency, and strength development of cemented paste backfill. Minerals 8:381
6. Belem T, Benzaazoua M (2008) Design and application of underground mine paste backfill technology. Geotech Geol Eng 26(2):147–174
7. Bell FG, Donnelly LJ, Genske DD, Ojeda J (2005) Unusual cases of mining subsidence from Great Britain. Int J Environ Geol 47:620–631
8. Lu H, Qi C, Chen Q (2018) A new procedure for recycling waste tailings as cemented paste backfill to underground stopes and open pits. J Clean Prod 188:601–612

9. Zhang B, Xin J, Liu L (2018) An experimental study on the microstructures of cemented paste backfill during its developing process. *Adv Civ Eng* 2018
10. Cao S, Yilmaz E, Song W (2018) Evaluation of viscosity, strength and microstructural properties of cemented tailings backfill. *Minerals* 2018(8):352
11. Liu L, Fang ZY, Qi CC, Zhang B, Guo LJ, Song K (2018) Experimental investigation on the relationship between pore characteristics and unconfined compressive strength of cemented paste backfill. *Constr Build Mater* 179:254–264
12. Zheng J, Li L, Li Y (2019) Total and effective stresses in backfilled stopes during the fill placement on a pervious base for barricade design. *Minerals* 9:38
13. Yang PY, Li L, Aubertin M (2017) A new solution to assess the required strength of mine backfill with a vertical exposure. *Int J Geomech* 17:04017084
14. Yin Y, Zhao T, Zhang Y, Tan Y, Qiu Y, Taheri A, Jing Y (2019) An innovative method for placement of gangue backfilling material in steep underground coal mines. *Minerals* 9:107
15. Douglas CM, Elizabeth AP, Geoffrey GV (2012) *Introduction to liner regression analysis* 5<sup>th</sup> Edition Copyright ©. Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada
16. Song X, Lu H (2017) Multilinear regression for embedded feature selection with application to fMRI analysis. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, February 4–9. California, USA, San Francisco, pp 2562–2568
17. Speed H (2010) *Statistical models: theory and practice*, revised edition by David A. Freedman. *Int Stat Rev* 78(3):457–458

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.