Multivariate Analysis of the Effect of Climate Conditions on Gold Production in Ghana*

[1]S. T. Appiah, [1]A. Buabeng and [1]N. K. Dumakor-Dupey
[1]University of Mines and Technology, P. O. Box 237, Tarkwa, Ghana

Abstract

The change in climatic conditions and its catastrophic effect on mining activities has become a source of worry for mining industries and therefore needs due attention. This study examined the effect some climate factors have on gold production in Ghana. First, a direct Multiple Linear Regression was applied on the climate factors with the aim of determining the relative effect of each factor on gold production which exhibited a time series structure. The consequence is that, the estimates of the coefficients and their standard errors will be wrongly estimated if the time series structure of the errors is ignored. In order to eliminate these deficiencies and better understand the effect of these climate factors on gold production, regression with ARIMA errors technique was employed after its appropriateness has been tested. The model was then compared in terms of prediction accuracy which resulted a MAPE of 9.78%. It was concluded that, gold production in Ghana is positively related to Temperature whilst negatively to Rainfall and Precipitate. It was recommended that mine operators in Ghana could base on this analysis to optimise their production planning and scheduling.

Keywords: Gold Production, Climate, Multicollinearity, VIF, Regression Models with ARIMA Errors

1 Introduction

The mining and minerals industries play a critical role in driving Ghana's economic growth. The mining of minerals such as gold in Ghana serves as a pillar to its economy, both in terms of contribution to gross domestic product, tax revenues, and also as an industry that directly employs millions of workers (Anon., 2010). The sector indirectly provides further employment and delivers significant additional economic benefit such as seeking construction services, manufacturing, wholesale and retail trade, as well as technical, scientific and professional services.

Notwithstanding these invaluable benefits to the economy, the changes in climatic conditions present physical threats to the mining and minerals industry because these industries are often located in extreme and challenging geographies (Kumar and Kumar, 1992; Kumar *et al*., 1992; Blischke and Murthy, 2011; Kumar *et al*., 2012; Lanke *et al*., 2016). These threats range from droughts and floods to a greater incidence of vector-borne diseases such as malaria.

Variations in weather conditions also affect the mining and metals industry in a variety of ways as material properties change significantly (Lanke *et al*., 2016). This is manifested as hardened steel and other metals become brittle, Hydraulic oil, the main working component of shovels, shows different characteristics under the influence of cold weather (Lanke *et al*., 2016). Furthermore, variation in temperature also causes stress on equipment, harming its ability to work efficiently (Larsen and Markeset, 2007; Gao *et al*., 2010; Loset, 1995).

The mining industries, although experienced at identifying and managing threats, increased temperatures, changes in precipitation, among other climate conditions may become additional stressors with the potential to intensify existing risks managed by these mining companies.

Hence, having a foreknowledge of the impacts of the changes of some climate components on gold production, mining assets and operations will provide the mining industries an edge in assessing climatic threats and vulnerabilities. It is, therefore, imperative for mine operators to anticipate and plan for the uncertainties of climate change. This is because the issue of climate change will profoundly continue to affect all aspects of mining operations in several ways *i.e*., physical risks to assets and infrastructure arising from climate variations, supply chain risks arising from disruption to transport networks and increased competition for climate-sensitive resources such as water and energy.

Therefore, this work seeks to examine as well as ascertain the relative effect of some climatic factors on Gold production in Ghana using multivariate analysis approach.

2 Resources and Methods Used

2.1 Multiple Linear Regression

The general form of a multiple regression is:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_q x_{q,t} + \varepsilon_t \quad (1)$$

where y_t is the variable to be forecasted and $x_{1,t}, \ldots, x_{q,t}$ are the q predictor variables. The coefficients $\beta_1, \ldots, \beta_q$ measure the effect of each predictor after taking into account the effect of all other predictors in the model. Thus, the coefficients measure the marginal effects of the predictor variables. When forecasting, the general assumption required for the errors $(\varepsilon_1, \ldots, \varepsilon_t)$ are as follows:

(i) $E(\varepsilon_i) = 0$, for all $i = 1, 2, \ldots, t.$ (linearity)

(ii) $\text{var}(\varepsilon_i) = \sigma^2$, for all $i = 1, 2, \ldots, t.$
(constant variance)

(iii) $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$, for all $i \neq j.$
(uncorrelated error terms)

(iv) $\varepsilon_i \sim N(0,1)$ (error terms be normally distributed)

2.2 Regression Models with ARIMA Errors

When performing regressions using time series variables, it is common for the errors (residuals) to have a time series structure. Normally, this violates the usual assumption of independent errors made in ordinary least squares regression. The consequence is that the estimates of coefficients and their standard errors will be wrong if the time series structure of the errors is ignored. It is possible, though, to adjust estimated regression coefficients and standard errors when the errors have an ARIMA structure (Ankrah *et al*., 2014). Thus, the errors from the regression model are autocorrelated (serially correlated). Due to this violation, the error term ε in regression equation (Equation (1)) is replaced by ω, where ω is assumed to follow an ARIMA structure.

2.2.1 Model Estimation

Suppose that y and x_q in Equation (1) are time series variables, and if ε follows an ARMA (1,1) model, it can be written as:
$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_q x_{q,t} + \varepsilon_t,$$
$$(1 - \phi_1 B)\varepsilon_t = (1 - \phi_1 B)\omega_t \quad (2)$$
where ω_t is a white noise series.

Notice that the model in Equation (2) has two error terms; that is the error from the regression model which we denote by ε_t and the error from the ARMA model which we denote by ω_t. Only the ARMA model errors are assumed to be white noise.

An important consideration in estimating a regression with ARMA errors is that all variables in the model must first be stationary. If the model is estimated while any of these variables are non-stationary, the estimated coefficients can be incorrect. One exception to this is the case where non-stationary variables are co-integrated. Thus, when series are differenced for stationarity, regression model (Equation (2)) with ARIMA (1,1,1) errors becomes:
$$y_n' = \beta_0 + \beta_1 x_{1,t}' + \beta_2 x_{2,t}' + \cdots + \beta_q x_{q,t}' + \varepsilon_t',$$
$$(1 - \phi_1 B)\varepsilon_t' = (1 - \phi_1 B)\omega_t \quad (3)$$

where
$$y_t' = y_t - y_{t-1}, \ x_{i,t}' = x_{i,t-1} - x_{i,t-1}, \ \varepsilon_t' = \varepsilon_t - \varepsilon_{t-1}$$
which is a regression model in differences with ARMA errors. In a more general form, a regression model with ARIMA (p,1, q) is written as:
$$y_n' = \beta_0 + \beta_1 x_{1,t}' + \beta_2 x_{2,t}' + \cdots + \beta_q x_{q,t}' + \varepsilon_t',$$
$$\phi(B)\varepsilon_t' = \theta(B)\omega_t \quad (4)$$

where the autoregressive (AR) and the moving average (MA) characteristic operators are given as:
$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) \text{ and}$$
$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) \text{ respectively.}$$

3 Results and Discussion

To assess the effect of climate on gold production, secondary data on production and some climatic variables (from gold mining areas in Ghana) spanning from 1960-2016 were obtained. The descriptive statistics of variables considered are summarised in Table 1.

3.1 Multiple Regression Analysis on Variables

Using Production (PROD) values as the dependent variable and the climate variables as the predictors, the result of the analysis is shown in Table 2. As observed, the p-value from the F-test (<0.01), shows that the model is statistically significant. Also, two out of four of the parameters are significant, which is an indication of how each variable contributes significantly to the model. The VIF values of all variables are close to 1.00, hence an indication of a nearly complete elimination of multicollinearity.

Table 3 shows the test on the assumptions regarding the regression model in Table 2. As observed in Table 3, all the assumptions of the model were satisfied except that the errors were serially correlated (p-value from Breusch-Godfrey test was <0.05). This is also confirmed by the Box-

Pierce tests (p-value <0.05) in Table 2. This suggests that the residuals of the regression model have a time series structure which could result in wrong estimation of parameters and standard errors. Thus, regression model with ARIMA errors technique would be appropriate (Ankrah *et al.*, 2014).

Before proceeding with the regression with ARIMA errors technique, the stationarity of the variables considered are tested. The Augmented Dickey-Fuller (ADF), as well as the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test were computed to validate this claim. As observed in Table 4, it clearly indicates that all the variables under consideration were nonstationary (p-value of ADF>0.05; p-value of KPSS<0.05) at order zero. However, the variables became stationary (p-value of ADF<0.05; p-value of KPSS>0.05) and integrated of order one after taking the first difference.

Table 1 Variables Considered for the Analysis

Variable	Initials	Unit	Mean	Standard Deviation	Min	Max	Correlation with PROD (Sig.)
Production	PROD	Oz	1118544.05	1072125.76	167115.00	4472378.84	1.0000
Rainfall	RNFL	mm	1164.87	143.04	810.03	1756.00	-0.2222 (0.0322)
Temperature	MAXT	^{0}C	27.29	0.39	26.35	28.35	0.3404 (0.0008)
Precipitate	PPT	mm	1227.72	17.11	1186.50	1263.75	-0.0437 (0.6773)

Table 2 Summary of Analysis

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	VIF
Intercept	-33380000.00	9623000.00	-3.47	<0.01	0.00
RNFL	363.90	836.30	0.44	0.67	1.10
TEMP	1440000.00	422300.00	3.41	<0.01	1.72
PPT	-4182.00	8319.00	-0.50	0.62	1.84

F-Value=5.43, P-Value=<0.01, AIC=1395.10, Box-Pierce=30.11 (<0.01)

Table 3 Test of Assumptions on Model

Assumption	Hypothesis	Statistic (P-Value)
Linearity (Runs)	H_0: $E(\varepsilon_i)=0$ H_1: $E(\varepsilon_i)\neq0$	0.0000
Normality (Shapiro-Wilk)	H_0: Normally distributed H_1: Not normally distributed	0.9814 (0.5992)
Serial Correlation (Breusch-Godfrey)	H_0: True autocorrelation=0 H_1: True autocorrelation $\neq0$	35.3640 (**<0.0001**)
Homoscedasticity (Breusch-Pagan)	H_0: Constant variance H_1: Non-constant variance	5.2506 (0.0724)

Table 4 Summary of Stationarity Test on Variables

Test	Hypothesis	Variables	Differencing (P-Value)	
			Order 0	Order 1
Augmented Dickey-Fuller (ADF)	H_0: Not Stationary H_1: Stationary	PROD	0.93	0.04
		RNFL	0.06	0.01
		TEMP	0.42	0.01
		PPT	0.25	0.01
Kwiatkowski-Phillips-Schmidt-Shin (KPSS)	H_0: Stationary H_1: Not Stationary	PROD	0.01	0.06
		RNFL	0.01	0.10
		TEMP	0.01	0.10
		PPT	0.01	0.10

Model Identification: In selecting the 'best' ARIMA errors for the regression model, nine competing models were fitted and the result is shown in Table 5. The 'best' model is the model with the minimum information criteria. Thus, the 'best' model is ARIMA (1,1,1).

Table 5 Competing ARIMA Models and Information Criteria

Model (p, 1, q)	AIC	AICc	BIC
(0,1,0)	1355.06	1356.42	1366.90
(1,1,0)	1354.53	1356.49	1366.00
(0,1,1)	1355.25	1357.20	1366.72
(1,1,1)	**1353.51**	**1356.18**	**1364.62**
(2,1,0)	1354.40	1357.07	1367.78
(2,1,1)	1354.97	1358.48	1370.27
(0,1,2)	1356.00	1358.67	1369.39
(1,1,2)	1354.58	1358.09	1369.88
(2,1,2)	1355.58	1360.08	1372.79

Estimation of Parameters and Diagnostic Checking: The regression model with ARIMA errors estimates are given in Table 6. Regarding the model diagnostic checking, the p-values from the Box-Pierce test (>0.05) indicates that the residuals of the selected model is now uncorrelated. Also, the Shapiro-Wilk test (p-value >0.05), indicates that the residuals are normally distributed. This is also confirmed by the diagnostic plots in Fig. 1.

Table 6 Summary of Regression Model with ARIMA Errors

Variable	Parameter Estimate
Intercept	62482.95
RNFL	-12.73
TEMP	25610.92
PPT	-839.26
AR1	0.84
MA1	-0.63

AIC=1353.51, AICc=1356.18, BIC=1364.62
Box-Pierce=0.18 (0.67), Shapiro-Wilk= 0.96 (0.11)

As observed in Fig. 1 (a), the standardised residual plot displays a normally distributed residuals as the points displays zero trace of trend, no outliers, and in general, no changing variance across time. Fig. 1 (b), the ACF of residuals plot, shows that only lag one (1) out of the seventeen (17) lags of the series residuals exceeds the significant bounds. This is negligible, since the probability of a spike being significant by chance is about one (1) in seventeen (17). This simply gives an indication of non-significant autocorrelation, since it is expected that at most one (1) out of seventeen (17) sample autocorrelations to exceed the 95% significance

bounds. From Fig. 1 (c), the p-value for Ljung-box statistic plot; shows that all p-values (the accumulated residual autocorrelation from lag 1 up to the last lag) are positioned above the 0.05 line which is an indication of an adequate model. Collectively, these tests suggest that the model fits very well.
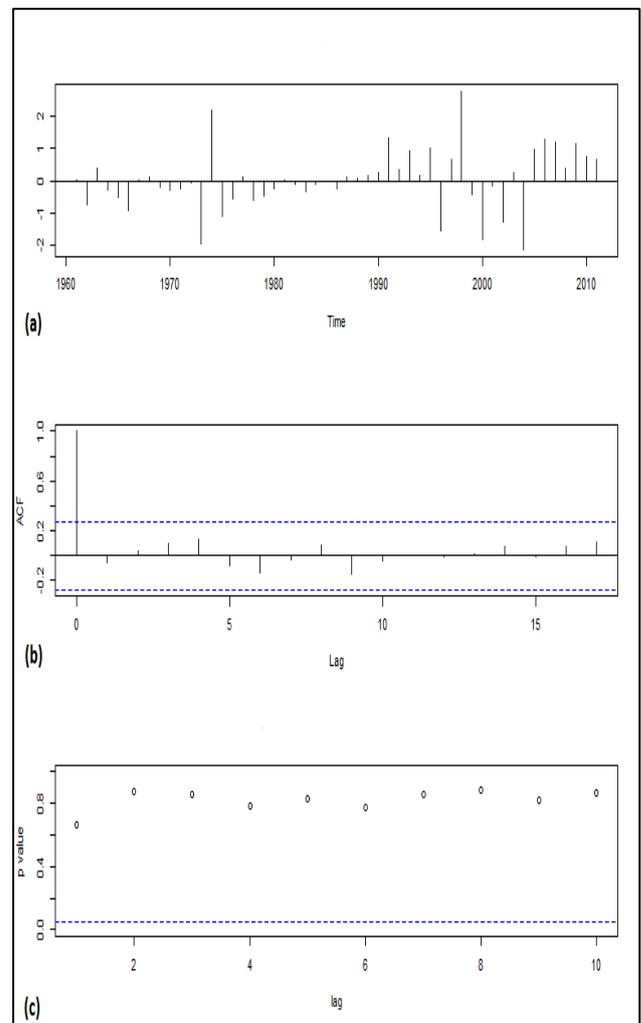

Fig. 1 Diagnostic of Regression Model with ARIMA (1,1,1) Errors

Thus, the regression model with ARIMA (1,1,1) errors is expresses as Equation (5)

$$PROD_t^{'} = 62482.95 - 12.73RNFL_t^{'} \\ + 25610.92TEMP_t^{'} - 839.26PPT_t^{'} + \varepsilon_t^{'} \quad (5)$$

where $\varepsilon_t^{'} = 0.84\varepsilon_{t-1} + \omega_t + 0.63\omega_{t-1}$

A critical observation of the coefficients of the developed model (Equation (5)) indicates that gold production in Ghana is positively related to Temperature (TEMP) whilst negatively to Rainfall (RNFL) and Precipitate (PPT).

From this interpretation, the value of gold production in Ghana is bound to increase by approximately 25610 Oz if there is a unit increase in Temperature (TEMP). This could be attributed to the fact that stable and ambient climate/weather conditions favour continuous mine equipment operations resulting improved mine production; smooth loading and hauling operations. Generally, atmospheric temperature conditions often influence the operational status of mining fleet (shovel-truck system), that is to say that a fall in temperature below operational limit would invariably affect equipment components including working component, tyres (TKPH), hydraulic oil, *etc*. According to Lanke *et al*. (2016), at -10°C material properties of mechanical components begins to response to temperature differently; steel hardens and other metals becomes brittle. This induces stresses in the equipment and reduces its operational efficiency.

Also, the negative coefficients of Rainfall (RNFL) and Precipitate (PPT) suggests that a unit increase in these factors would decrease Ghana's gold production by approximately 12 and 839 Oz respectively. Of the two, Precipitate is more impactful; as precipitation increases, production is substantially reduced. Rainfall is less impactful; even so, with increases in rainfall, production is reduced.

This could be as a result of the effects of adverse mining conditions such as high levels of rainfall rendering haul roads unmotorable and flooding of pits floor. During heavy rainfalls which are usually accompanied by severe thunderstrikes and lighting, mine operations including loading and hauling are halted for safety reasons. This also affects the working ability of operators and causes reduction in human-machine interface, alertness, productivity and quality of work. Further, presence of water in pits impedes movement of haulage trucks and decreases production since additional cost is incur in dewatering the mining area. This corroborates the findings of Lanke *et al*. (2016), that rainfall and snowfall adversely influence mine production volumes.

Finally, the positive constant (62482) implies that aside these climate factors considered in this study, there are other variables contributing to gold production in Ghana. This is manifested in the usage of heavy machineries, technology, supervision, operator efficiency, among others.

Table 7 shows the observed and predicted values of the model with respective errors. As observed, the model produced an MAPE of 9.78%. Fig. 2 shows the plot of observed as well as the predicted production values for five years.

Table 7 Observed and Predicted Values with Respective Errors

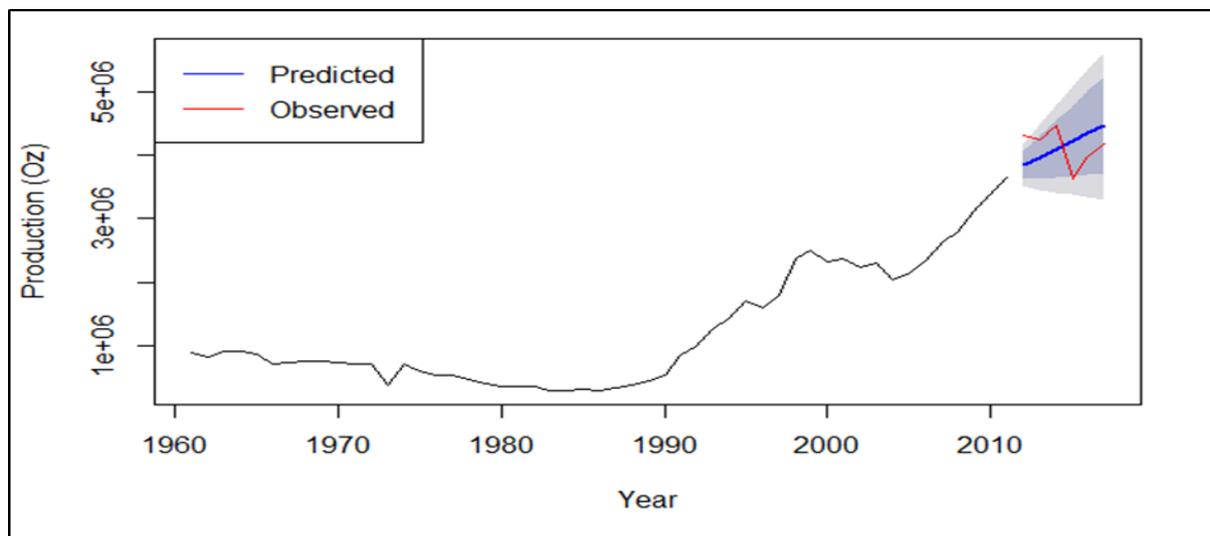| Year	Observed	Predicted	Lo 80	Hi 80	Error
2012	4313190	3849351	3632688	4066013	-463839
2013	4249902	3971136	3631776	4310497	-278766
2014	4472379	4105784	3653887	4557681	-366595
2015	3623742	4234972	3675707	4794237	611230
2016	3984051	4356693	3694018	5019368	372642
2017	4175000	4464440	3701943	5226938	289440
MAPE (%)					**9.78**

Fig. 2 Plot of Observed and Predicted Values

4 Conclusion

This study was conducted to examine and ascertain the relative effect of some climate factors on gold production in Ghana. The factors were examined using regression with ARIMA errors technique through which a suitable model was developed. The model adequately described and predicted gold production in Ghana. In terms of prediction accuracy, the model produced an MAPE of 9.78%. The conclusion drawn from the analysis indicated that, gold production in Ghana is positively related to Temperature whilst negatively related to Rainfall and Precipitate. It is recommended that mine operators can base on this analysis to optimise their Production planning and scheduling.

References

Ankrah, S., Peiris, B. L. and Thattil, R. O. (2015), "Weighted Modelling and Forecasting of Cocoa Production in Ghana: A Multivariate Approach", *Tropical Agricultural Research*, Vol. 26, No. 3, pp. 569–578.

Anon. (2010), "Draft National Mining Policy of Ghana", 34 pp. *https://asgmresearch. weebly.com/uploads/3/0/1/6/30160743/mining_ policy_of_ghana.pdf.* Accessed: April 10, 2018.

Blischke, W. R. and Murthy, D. P. (2011), *Reliability: Modelling, Prediction, and Optimization*, Vol. 767, John Wiley & Sons, 821 pp.

Gao, X., Barabady, J. and Markeset, T. (2010), "An Approach for Prediction of Petroleum Production Facility Performance Considering Arctic Influence Factors", *Reliability Engineering and System Safety*, Vol. 95, No. 8, pp. 837–846.

Kumar, D., Klefsjo, B. and Kumar, U. (1992), "Reliability Analysis of Power Transmission Cables of Electric Mine Loaders using the Proportional Hazards Model", *Reliability Engineering and System Safety*, Vol. 37, No. 3, pp. 217–222.

Kumar, D. and Kumar, U. (1992), "Proportional Hazard Model: a useful Tool for the Analysis of a Mining System", *Proceedings of the 2nd APCOM Symposium*, Tucson, AZ, USA, pp. 6–9.

Kumar, R., Barabady, J., Markeset, T. and Kumar, U. (2012), "Improving Maintainability in Extreme Cold Climatic Conditions", *International Journal of Performability Engineering*, Vol. 8, No. 5, pp. 563-572.

Lanke, A., Ghodrati, B. and Hoseinie, H. (2016), "Uncertainty Analysis of Production in Open Pit Mines: Effect of environmental conditions", *Mining Sciences*, Vol. 23, pp. 147-160.

Larsen, A. and Markeset, T. (2007), "Mapping of operations, maintenance and support design factors in arctic environments", *Proceedings of the European Safety and Reliability Conference*, pp. 25–27.

Authors

S. T. Appiah is a Senior Lecturer working at the University of Mines and Technology, Tarkwa, Ghana. He holds the following degrees in Mathematics: BSc and MSc from the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, and PhD from the University of Mines and Technology (UMaT), Tarkwa, Ghana. He is a member of the Ghana Mathematics Society and he is interested in modeling with Differential Equations and Statistics.

A. Buabeng holds BSc and MPhil in Mathematics from the University of Mines and Technology (UMaT), Tarkwa, Ghana. He is currently a demonstrator, a researcher in the Mathematical Sciences Department of the University of Mines and Technology (UMaT), Tarkwa, Ghana. He is a member of the Ghana Mathematics Society. Research interests are in time series analysis and statistical methods for quality improvement. In the time series area, the focus is on building models through dimension reduction methods, multivariate prediction and process control with latent variables, designed experiments.

N. K. Dumakor-Dupey holds BSc and MPhil in Mining Engineering from University of Mines and Technology. He is currently a Teaching and Research Assistant in the Mining Engineering Department of the University of Mines and Technology, Tarkwa, Ghana. Areas of his research interest are modelling and simulation of mining systems, underground mining systems, and sustainable mining.