26-027

Dynamic Failure Predictions Using Machine Learning and Geochemical Data

Heather Lawson
NIOSH-SMRD, Spokane, WA, USA

Tom Meuzelaar
Life Cycle Geo, Longmont, CO, USA

Alice Alex
Life Cycle Geo, Toronto, ON, Canada

Sam J.S. Wright
Life Cycle Geo, Spokane, WA, USA

D. Morgan Warren
Life Cycle Geo, Atlanta, GA, USA

David Hanson
NIOSH-SMRD, Spokane, WA, USA

ABSTRACT

Dynamic coal failure events remain a safety concern in underground coal mining due to their sudden nature and high fatality rate. Previous research utilized Principal Component Analysis on field samples revealing strong empirical correlations between positive failure history and low sulfur content (S) and high volatile matter (VM). Further studies indicated that bumping coals tend to be less mature and less well-cleated than non-failure coals. However, these in situ studies faced the challenge of distinguishing intrinsic coal properties from external geological and stress-related influences. This study augments previous findings through laboratory experimental testing of coal samples that included unconfined compressive strength, chemical composition analysis, and qualitative observations on cleating, combined with compositional data analysis and machine learning techniques, to isolate and understand coal's inherent bursting capacity. These results further confirmed the relationship between coal chemistry and failure classification.

INTRODUCTION

Dynamic failures, or "bumps," are a category of ground failure occurring in underground coal mines that involve the sudden expulsion of coal and rock debris into a working area of an active mine. They are an imperative safety concern for the global coal industry. Dynamic failure events have occurred in nearly every coal mining nation of the world, and there have been an estimated greater than 30,000 fatalities as the result of dynamic failure events globally (Bodziony and Lama, 1996; Kidybinski, 2011). Kidybinski (2011) describes them as "the most dangerous natural phenomenon creating a threat to life and health of miners working in hard coal mines."

In the United States, dynamic failures are significantly less common than in many other coal mining countries. However, when they do occur, they result in a proportionately extremely high rate of accidents and fatality, with more than 60% of cases reported to the Mining Safety and Health Administration (MSHA) resulting in worker injury up to and including death (MSHA, no date). The driving mechanisms behind these events are not yet fully understood. Research aimed at optimizing pillar design and mining practices has resulted in a significant decrease in dynamic failure occurrence over the past three decades. However, events continue to occur, and new research focused on engineering controls has failed to yield continued, measurable results. It is likely that to produce a further reduction in the rates of dynamic failure accident occurrence, novel lines of inquiry must be pursued. Recent research performed by the National Institute of Occupational Safety and Health (NIOSH), Spokane Mining Research Division (SMRD) has sought to address this need through a focus on the complex interaction between mining induced stressors and risks innate to the host rock mass through improved incorporation of geologic variables into hazard assessment.

Background

Specifically, recent studies have focused on the compositional attributes of coal that correlate with a reportable history of dynamic failure event occurrences in the United States (Lawson, Weakley and Miller, 2016; Berry, Warren and Hanson, 2019; Lawson, 2020; Lawson and Hanson, 2021; Hanson and Lawson, 2023). A dynamic failure event is reportable if it results in worker injury, impairs ventilation, impedes passage, disrupts regular mining activity for one hour or more, or results in worker entrapment for a period of 30 minutes or greater, per the Code of Federal Regulations, Part 30 (30 CFR), subsections 50.2 and 50.10. Key findings from these studies indicate that within the United States, coals prone to dynamic failure fall into two categories: 1) Those that are relatively thermally mature, and 2) those that are relatively immature.

Figure 1 (Lawson, 2020) illustrates how coal data from the Pennsylvania State Coal Sample Databank (PSCSD), which has been cross-referenced with MSHA accident reports, plots on a Van Krevelen diagram. This figure shows that there are two distinct clusters of dynamic failure prone coals: 1) close to the axis origin, and 2) another, larger group falling further from the axis in an arc. This suggests that, within the sample set considered, coals with a history of reportable dynamic failure tend to be thermally immature, but that a second, smaller set of samples have experienced dynamic failure at higher maturity, and these cases show evidence of internal gas pressure as a contributing risk factor (Lawson, 2019).

MSHA Accident Narratives associated with these reports suggest that high-maturity coals in this group experienced failure in which there was a significant internal gas pressure component, while less mature coals experienced failure types associated with unfavorable stress concentrations and/or association with geologically hazardous conditions. This is intuitive; Agapito and Goodrich (2000) note a correlation between cleat development and dynamic failure occurrence in western U.S. coals. Coals with poorly developed cleating are more capable of storing large amounts of potential energy, which can then be released kinetically in the form of a dynamic failure event. Conversely, well-developed cleating allows for ready planes of failure such that coal pillars may fail gradually, rather than explosively. As cleat development is a function of catagenic processes during coalification, it is then unsurprising that most U.S. coals that have a history of reportable dynamic failure are thermally immature.

The most consistent compositional attribute of dynamic failure prone coals, however, was a gross sulfur content of less than 1%. In most cases with a history of dynamic failure, compositional sulfur (measured in weight percent by volume, moisture-ash-free basis) was much less than 1%, with an average value of only 0.71%. This correlation, moreover, was consistent among both the high-maturity and low-maturity dynamic failure groups. Hanson and Lawson (2023) performed a Recursive Feature Elimination with Cross Validation (RFECV) to identify those compositional attributes that most closely correlate with a positive history of MSHA-reportable dynamic failure phenomena. A RFECV will assign a rank to each attribute based upon the number of times the algorithm selected it as a good indicator of dynamic failure history. The RFECV was performed using eighty-six coal sample records from the PSCSD that were established to have experience MSHA-reportable dynamic failure events after 1983, and a series of randomly selected eighty-six coal records from the coals that have not experienced reportable dynamic failure, to serve as control dataset. Out of three thousand trials, low sulfur content was assigned Rank 1 in 100% of trials. These results corroborate the findings of Lawson, Weakley and Miller (2016) and Berry, Warren and Hanson (2019) that also suggest a very strong correlation between dynamic failure event occurrence and low-sulfur content (Table 1).

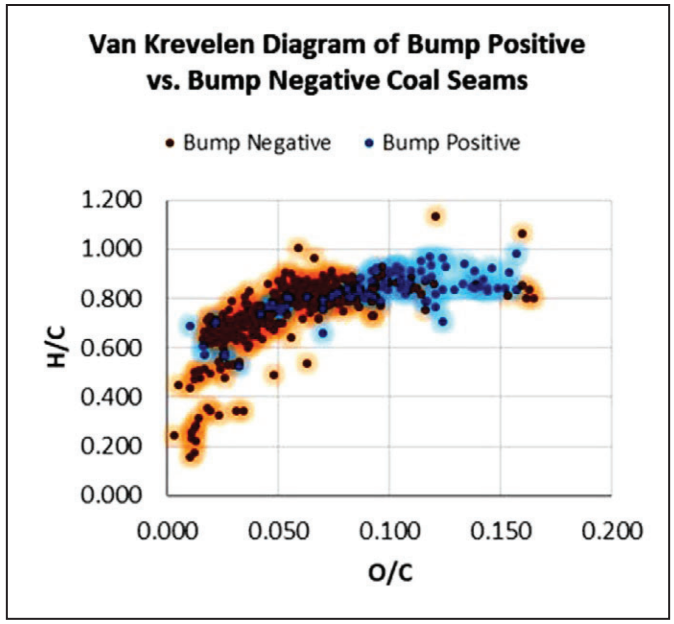


Figure 1. Coal sample records from the PSCSD

The correlation between low-sulfur content and dynamic failure event occurrence raises questions, most notably in terms of how this may contribute to failure mechanistically. Sulfur content has not been cited in any of the preceding literature as a contributor to dynamic failure occurrence, and the mechanism(s) behind this phenomenon remains elusive. One possibility may be that low sulfur

Table 1. Relative abundance of rank 1 feature assignment out of 3,000 trials using recursive feature elimination with cross validation (RFECV) analysis

Feature	%
Sulfur	100
Oxygen	91.3
Volatile Matter	72.6
Ash	62.6

content serves as a geochemical proxy for known dynamic failure risk factors. Specifically, low sulfur content has been linked with inland depositional settings which may, in theory, have higher frequencies and greater thicknesses of paleochannels and strong sandstone members in the mine roof. These have been associated with bridging effects and unanticipated distributions of stress in the mine workings (Rice, 1935; Holland and Thomas, 1954; Iannachione and Zelanko, 1995; Agapito and Goodrich, 2000; Whyatt, 2008; Whyatt and Varley, 2010; Mark and Gauna, 2016; Lawson et al., 2017). Successful mine design in sedimentary deposits lies not in preventing failure of the rock mass, but rather in strategically controlling it, such that mining-induced stresses can be shifted and anticipated to maintain maximum, albeit temporary, stability. This requires what is referred to in mining vernacular as "good caving". A "good cave" will be tight behind the longwall shields or pillar line. Moreover, proximity of the stiff unit to the coal mine roof has been linked to increased risk, such that increasing distance corresponds to decreased risk as the bridging effect produced by poor caving is more effectively dispersed in the overlying rock mass (Lawson et al., 2017).

Ultimately, however, failure results from a combination of factors, although the precise balance of risk associated with any particular failure may vary from mine to mine, and even between locations within a single mine. Vardar et al. (2018) state that dynamic failure events occur when the bearing capacity of coal is abruptly exceeded in a stress setting of sufficient magnitude to generate violent failure as the result of some triggering event. In the absence of laboratory studies to augment observations derived from in situ failure, the influence of the mining environment cannot be extricated from failure mechanics innate to the coal itself. This then points to the need for laboratory studies in which dynamic failure is simulated under controlled conditions in an array of coal samples with known, diverse compositions. Toward this end, a series of laboratory tests were conducted whose results were then analyzed using a machine learning (ML) approach to determine which compositional and physical attributes of the coal were most closely associated with a sample's tendency to fail, either gradually or explosively, under controlled conditions.

LABORATORY EXPERIMENTAL PROCEDURES

An experiment was designed to simulate dynamic failure in coal samples with diverse compositions under laboratory-controlled conditions at the SMRD. Samples were acquired from multiple sources, and subject to different sample handling procedures following initial collection which did not uniformly include wrapping to prevent degassing. Because of this, methane desorption during testing could not be included in tested variables. Cleat spacing was mapped and descriptive information such as the presence of secondary mineralization was collected prior to sample testing. Dynamic load was simulated under consistent conditions. All load tests conducted in this work were performed in uniaxial compression using an 890 kN (200,000 lbf) stiff frame, servo-controlled test machine (Tinius Olsen 1000SL). The sampling rate was 1000 samples/sec and the accuracy of the machine is ± 0.1% (1.8 N/0.4 lbf) of the indicated load from 0.2% (1.78 kN/400 lbf) to 100% (890 kN/200,000 lbf) of the frame capacity. All samples were preloaded to 890 N (200 lbf) and tests were performed at a constant deformation rate of 12.7 μm/sec (0.0005 in/sec).

A total of 36 tests were completed. Due to challenges in preparing consistent cores from available coal samples, four cylindrical diameters were prepared. These were 3.8 cm (average height 4.9 cm), 4.1 cm (average height 6.1 cm), 5.4 cm (average height 9.4 cm), and 7.5 cm (average height 9.2 cm). In addition, five cubic samples were prepared. These samples ranged from 5.3 cm to 7.9 cm on a side. To avoid error in measuring unconfined compressive strength (UCS) due to stress concentrations within test samples, ASTM standards require sample surfaces be plane to within 0.005 cm (0.002 inch) and deviation of sample vertical axis is less than 0.50. All samples tested were prepared to this standard. Figure 2 shows a characteristic cylindrical sample prior to testing. During testing both load and axial deformation were monitored and recorded, allowing stress-strain curves to be constructed.

Quantifying the degree of bursting behavior was one of the challenges encountered during the testing procedure. An attempt was made to quantify the energy of failure using several simultaneous lines of investigation. A major objective of this series of tests was to examine a wide variety of properties simultaneously and determine if there were correlations between these properties and the failure mode of the sample that provided a characteristic measure of coal

Figure 2. Photograph of test apparatus prior to a uniaxial compression test. The clock shown provided a time reference for optical and infrared (IR) videos. The grid in the background consisted of 2.54 cm (one inch) squares that allowed computation of debris velocity

instability under loading. The secant modulus of elasticity was computed using platen displacement, sample height, and load at 50% peak strength (Obert, 1973). The sample temporal load history and platen displacement were collected and recorded by the internal Tinnius Olsen software. The UCS was defined as the maximum load attained during a test. Numerous works have examined the manifestation of elastic strain energy during testing as a measure of burst propensity. These have included Kidybinski (1981), Li et al. (2017), and Akdag et al. (2021). To examine this parameter and correlate it with other measured quantities, both the dissipated elastic energy and effective elastic strain energy release rate (EESERR) as defined by Lu et al. (2019) were computed.

In addition to the load and deformation, high speed optical videos were made with a Photron Mini camera using a frame rate of 250 frames/sec. These videos were used to find the velocity of particles ejected during tests. Akdag et al. (2021) and Akdag et al. (2018) state that the potential intensity of a burst can be quantitatively related to the ejection velocity of debris. Using the 2.54 cm (1 inch) grid present at the rear of the test frame as a distance scale (Figure 2), an estimate of velocity was computed. In addition to optical images, infrared images were collected with a FLIR 300 camera at 15 frames/second (fps). This camera is normally used in optical gas imaging (OGI) and that was the original intention here. However, no gas could be identified and so the videos were used to image infrared variations during loading and failure. This camera is designed to image absorption from hydrocarbons (such as methane) and operates at 3.2 to 3.4 micron wavelength. This is in contrast to wider band cameras used in other works that report a sensitivity between 7 and 14 microns.

Table 2. Elemental Variables collected from simulated dynamic failure test debris and associated ASTM standards

ASTM Standard	Variable	Unit of Measure
D 3302, 3173	Total moisture	wt. %
D 3174	Ash	wt. %
D 3175	Volatile matter	wt. %
D 4239	Total sulfur	wt. %
D 2492M	Pyritic sulfur	wt. %
D 2492M	Sulfate sulfur	wt. %
D 2492M	Organic sulfur	wt. %
D 5865	Gross calorific value	Btu/lb
D 5865	Net calorific value	wt. %
D 5373	Carbon	wt. %
D 5373	Hydrogen	wt. %
D 5373	Nitrogen	wt. %
Calculation	Oxygen	wt. %

All samples were tested to failure, defined by either the sample disintegrating or plastic post-peak behavior. At the conclusion of each test, debris was carefully collected and saved for later analysis. This included both a measure of the size fraction of the particles and a geochemical analysis. Sample debris was collected into clean containers for later testing and the test apparatus platens were cleaned using methanol between tests. Sample debris was then tested for composition through proximate and ultimate analysis, and petrographic microscopic analysis. Petrographic analysis was completed using 500-point maceral counts under oil immersion. Data collected and associated standards regarding proximate and ultimate analysis are listed in Table 2.

However, all approaches used to quantify bursting behavior failed to yield a consistent and repeatable method for assessing the degree to which any given sample failed dynamically. Instead, it was observed that total failure of a sample was achieved through multiple discrete bursting events, each with its own measurable load drop and resultant strain, and the cumulative effect of these did not capture the quality of explosive failure in a way usable for the purposes of this study. This highlights the difficulty in measuring burst proneness of coals. Consequently, test results were binned into one of two qualitative designations: "Crushed" or "Burst". "Crushed" samples were those in which the bulk of the sample failed non-explosively, although failure may have been preceded by high-velocity ejections of small sample particles. "Burst" samples were those in which the entire sample failed energetically, through one or more bursting events involving the high-speed ejection of material. These categories of failure were visually distinct, with no overlap or instances of ambiguity. The body of data generated by multiple lines of synchronous investigation remain available for future work.

DATA ANALYSIS

Data Preparation and Organization

Specific data preparation and organization were conducted to ensure the coal geochemical dataset was suitable for unsupervised multivariate statistical analysis and supervised ML. The coal geochemical dataset comprised the following compositional data: oxygen (O), carbon (C), nitrogen (N), hydrogen (H), sulfate sulfur (SS), pyritic sulfur (PS), and organic sulfur (OS), with total moisture, ash content, vitrinite reflectance, and cleat spacing. The dataset was labeled with target variable classifications of 'Burst' or 'Crushed' indicating how the coal failed during the UCS experiments, providing the basis for predictive analysis.

No missing data or detection limit issues were identified in the geochemical data, so imputation or detection limit handling was unnecessary. However, certain calculations were performed to complete sub-compositions, such as determining the pyritic sulfur component of the total sulfur content. All data analysis, preparation, and modeling in this study were conducted using the Python (Van Rossum and Drake 2009) programming language.

Data Preprocessing

Before applying statistical and ML models, several preprocessing steps were undertaken to ensure data integrity and consistency across analyses for the predictor and target variables: a center-log ratio transformation (CLR) to address statistical issues inherent in compositional datasets and scaling to address any large differences in magnitude between features. These preprocessing steps are discussed in the following sections.

Compositional Data Analysis

Traditional statistical methods cannot be directly applied to compositional data, such as bulk rock chemistry, due to the inherent dependencies between the variables because compositional data represents proportions of a whole (Aitchison 1986). These dependencies create a closed dataset structure, where all points lie within a constrained simplex. If left uncorrected, these interdependencies can lead to spurious correlations and invalid statistical inferences. To address this issue, compositional data were transformed into real space using the CLR transformation (Aitchison 1986 and Egozcue et al. 2011). The CLR transformation ensures that the data is "opened" into an unconstrained space, allowing the application of standard multivariate statistical techniques such as principal component analysis (PCA). This processing transformation is necessary for properly analyzing and interpreting the relationships between the geochemical variables.

Scaling

Scaling is a critical step in preparing the data for analysis, especially for ML models that can be sensitive to variations in magnitude across different features (Bishop 2006; Han et al. 2011). Two distinct scaling methods were applied: (1) For descriptive analytics, scaling was performed on the entire dataset to normalize feature values for exploratory analysis, and (2) For predictive analytics, scaling was incorporated within the model pipeline. The scaler was fit on the training data and then applied to the test data during cross-validation (CV), ensuring that scaling did not lead to information leakage or overfitting (Stone 1974; Hastie 2001).

Machine Learning Techniques

ML techniques were applied in this study for two different objectives: 1) unsupervised ML used for dimensionality reduction and feature extraction, and 2) supervised ML used for predicting coal behavior as either 'Burst' or 'Crushed'. ML was used because it provides a suite of tools ideal for elucidating complex relationships between variables, assisting with feature engineering to transform or create new variables which capture more value, and to generalize a relationship based on the chosen variables to make predictions.

Principal Component Analysis

PCA was used to explore relationships within the dataset by reducing its dimensionality while preserving statistical variance (Pearson 1901; Greenacre 2022). PCA generates principal components (PCs), which are linear transformations of the original variables that maximize variance and are uncorrelated, aiding in feature selection and interpretation. Each PC is a linear combination of the original variables, with eigenvectors representing directions of maximum variance. Factor loadings, indicating the correlation between variables and PCs, reveal each variable's contribution to the overall variance structure.

Biplots visualize the relationships between two PCs (typically just the first three to four PCs are visualized as that represents a large proportion of dataset variance), displaying transformed sample scores as points and factor loadings as vectors, illustrating variable contributions and dataset variance (Gabriel 1971; Greenacre 2010). The variance explained by each component is indicated on the biplot axes to help interpret the amount of information captured by the graphed components. Additional information not directly used in the PCA, such as the categorical variable representing 'Burst' or 'Crushed', was overlain onto biplots. This allows for the interpretation of the relationship between these variables and the underlying high-dimensional geochemical data. The overlain data helps to visually assess how target variables and other meta data interact with the PCs and overall variance structure.

Supervised Learning: Predictive Modeling

Multiple supervised ML methods were applied in this study to predict coal seam burst risk, which are described in the sections that follow.

Decision Trees

A non-parametric decision tree (DT) algorithm, chosen for its flexibility in handling non-linear relationships without requiring assumptions about data distribution, was used as a first approach to predicting burst proneness. DTs partition the dataset iteratively, aiming to minimize misclassification using the Gini index to measure subset homogeneity (Breiman et al., 2017). For instance, partitions may be defined by a threshold concentration of an analyte (e.g., SS > 5 wt.%), generating decision rules for classifying bursts. Various iterations of the decision tree model were tested, incorporating different input data streams such as:

- Geochemical data alone
- Geochemical data combined with categorical variables (e.g., cleating, vitrinite reflectance)
- Categorical data alone

This workflow allowed the impact of each variable and variable type to be evaluated and determine which variable combination best captured patterns related to coal seam bursts, a precursor to demonstrating predictability.

Despite their advantages, DTs can overfit the training data, especially when training on a dataset with noisy (e.g., highly variable) data or numerous features, leading to poor generalization. Small changes in input data can also result in significant variations in the tree structure. Although DTs are often considered interpretable, their complexity can make it difficult to generalize decision rules, and their output must be interpreted carefully, particularly with wide (high-dimensional) datasets.

Logistic Regression

For descriptive analytics, a parametric approach was utilized to further explore the relationships between geochemical features and the target classification. Due to the relatively small dataset and the complexity of the underlying relationships, logistic regression was selected as a suitable method for descriptive analysis. Logistic regression is highly interpretable and robust for small datasets, making it ideal for understanding the key potential predictor variables of the target variable (Cox 1958; Pregibon 1981; Menard 2002; Hosmer & Lemeshow 2000; Harrell 2015).

The logistic regression model maps the log-odds of an event (i.e., coal seam burst) as a linear combination of the predictor variables. The model applies a sigmoidal "logistic" function to convert the linear output into probabilities, bounded between 0 and 1. The model's coefficients indicate the magnitude and direction of the effect of each predictor on the target variable, while statistical measures such as p-values assess the significance of these relationships (Cox 1958). The logistic regression model was trained using the Newton-Raphson optimization method, which iteratively minimizes the loss function to achieve the best fit for the data.

RESULTS

Unsupervised Learning

Principal Component Analysis

PCA results revealed approximately 75% of the total variance in the system can be explained by the first two principal components (PC1 and PC2), representing 52.7% and 22.3% of the explained variance, respectively, as shown in Figure 3. PC3 contributed an additional 12.3% of the variance, bringing the total variance explained by the first three components to nearly 87%. The fact that 87% of the variance is explained by the first three components is significant because it indicates most of the complex, multidimensional relationships within the dataset can be captured and interpreted through a highly reduced set of variables. This reduction in dimensionality enables clearer visualization of the key correlated groups of parameters and their relationships to burst proneness while retaining most of the information needed for subsequent analysis. It lends credibility to relationships observed on the biplot, facilitating optimal construction of supervised learning models.

Pyritic sulfur and sulfate sulfur factor loadings (Figure 3) are positively aligned with PC1 and exhibit high variance and are inversely related to other variables such as total moisture, hydrogen, carbon, nitrogen, and oxygen. Both biplots suggest that sulfur content in the coal has an opposing relationship to other primary coal chemical constituents. Additionally, sulfate and pyritic sulfur appeared to be collinear, while organic sulfur was obliquely aligned with PC2, further suggesting that sulfur speciation in coal is a potentially important differentiator of coal types with respect to mechanical behavior.

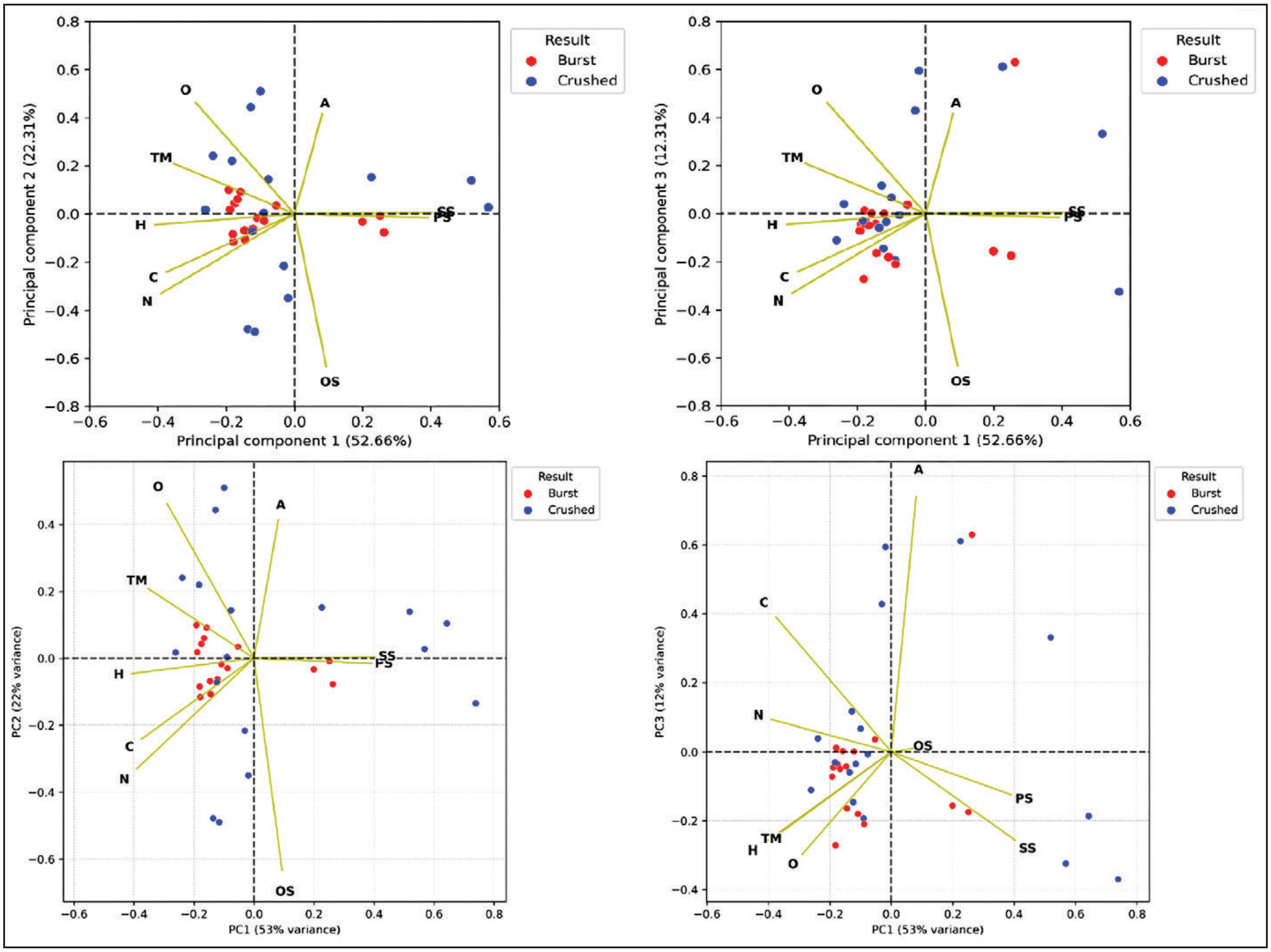


Figure 3. PCA Biplot of Coal Composition. PC1 and PC2 (left) and PC1 and PC3 (right) for all variables. TM. total moisture, O. oxygen, A. ash, SS. sulfate sulfur, PS. pyritic sulfur, OS. organic sulfur, N. nitrogen, C. carbon, and H. hydrogen

Variables such as total moisture, oxygen, hydrogen, carbon, and nitrogen clustered together away from sulfur components, indicating a close association between these variables. Ash, on the other hand, was uncorrelated with the other variables, suggesting it behaves independently within the compositional dataset.

Samples classified as 'burst' exhibited clear clustering within the PCA space, separating well along PC1 and PC2 and forming a tight grouping near the origin. Conversely, crushed samples spread outward along both PC1 and PC2, indicating more heterogeneous variation in chemistry. Comparing PC1 and PC3 shows a similar, though less distinct, pattern, suggesting that the first two principal components capture most of the separation between burst and crushed samples. Overall, the PCA indicates that the burst–crushed separation is well-expressed in the data, suggesting that these variables are suitable for descriptive modeling.

Cleating Analysis Using Decision Trees

The degree of cleat development of coal seams was analyzed using decision tree models to explore its relationship to crush and burst events (Figures 4 and 5). DT models effectively captured the relationships between cleating and the likelihood of bursting, confirming that the quality of cleats plays a role in these outcomes. Poor cleating was strongly tied to bursting events, with 6 of the 14 bursting events occurring in "Very poorly cleated" samples and another 2 of the 14 being "Poorly cleated". Despite these relatively strong relationships, cleating was considered a poor metric for determining bursting because the qualitative nature of

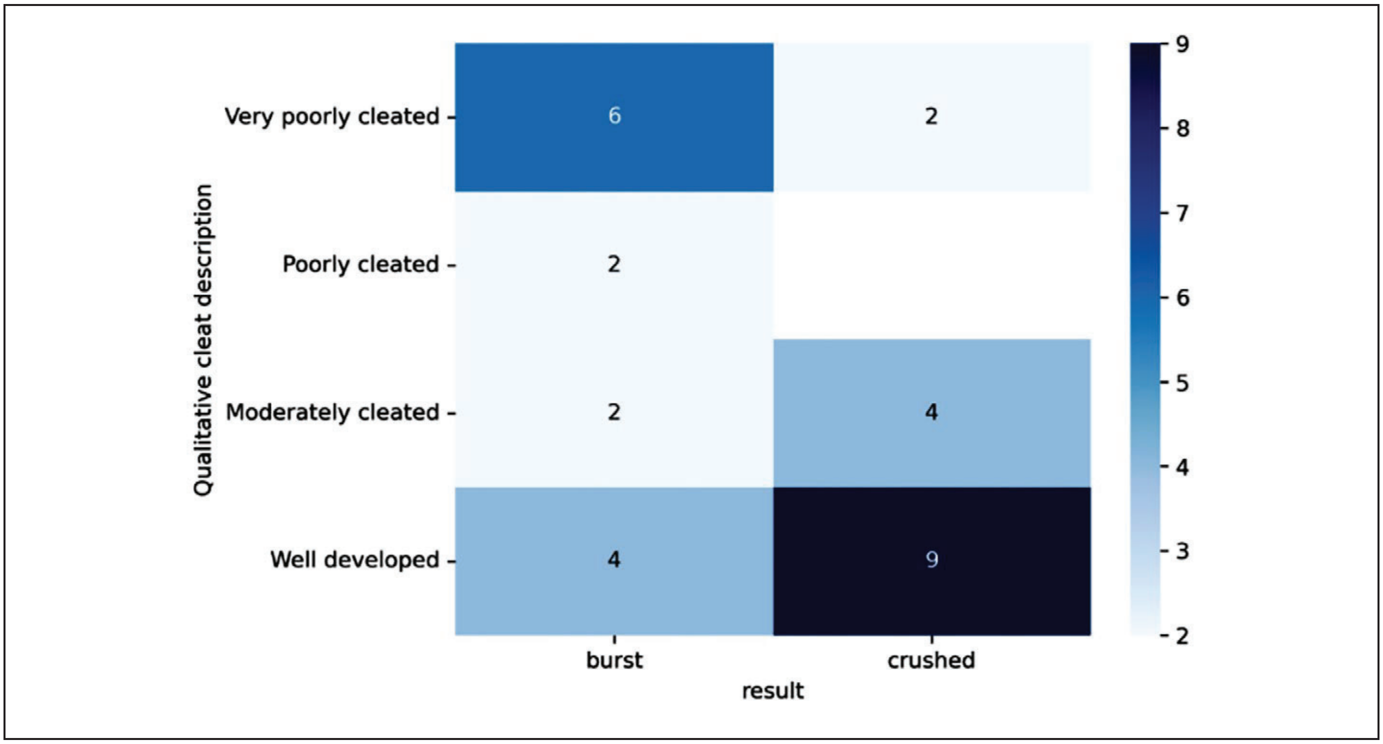


Figure 4. Correlation analysis between the cleating description and the target variable

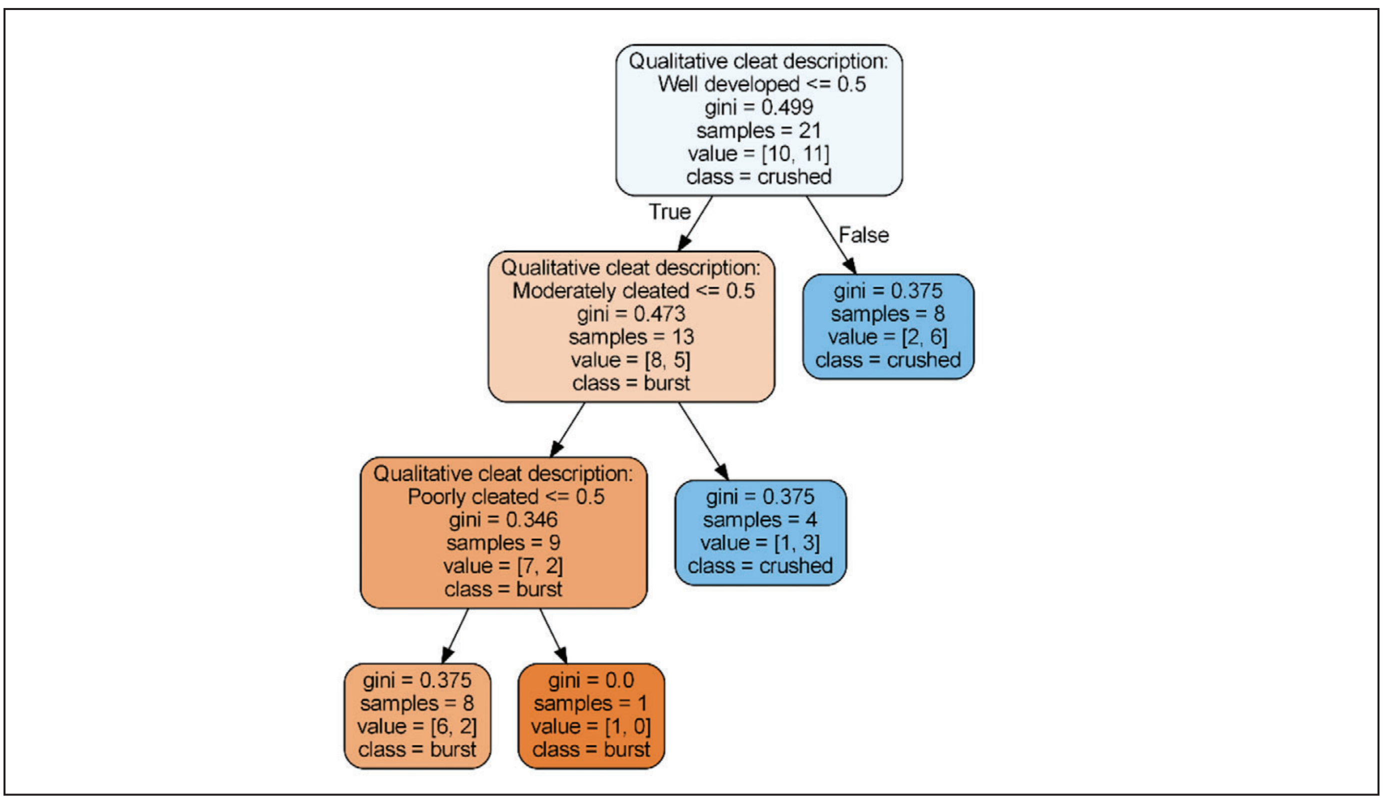


Figure 5. Decision tree diagram classifying samples into crushed or burst classes depending on cleating

the visual observations that lack the quantitative definition needed for its inclusion in rigorous statistics.

Supervised Learning

Compositional Tree Models

Tree-based models were also constructed to explore the relationships between the compositional data and coal seam failure classification (i.e., 'Burst' or 'Crushed'). The complex clustering relationships seen in Figure 3 are ideally predicted using a non-parametric tree model. DTs were employed to predict 'Burst' or 'Crushed' behavior from coal chemistry and generally performed well on the entire dataset. However, given the inherent complexity of the geochemical relationships and the relatively limited size of the dataset, the models struggled to generalize across multiple data splits. Specifically, during CV procedures such as k-fold and train-test splits, the models exhibited high variance and were unable to consistently pass validation test metrics. This suggests that while tree-based methods have predictive value and hold potential for predicting burst and crush events, the current dataset lacks the volume and diversity needed to build a robust, reliable model capable of performing well across unseen data.

As a result, efforts were further focused on more interpretable models and techniques, such as logistic regression and descriptive statistics. Instead of prioritizing predictive modeling, the full dataset was leveraged to explore relationships between the input and target variables. This approach extracted meaningful insights without the limitations imposed by model validation requirements (e.g., CV).

Logistic Regression

A logistic regression model was developed to assess the influence of coal composition on the likelihood of coal seam failure being classified as 'Burst' or 'Crushed'. An initial model with each analyte included was built and produced good performance metrics, with a high receiver operating characteristic-area under the curve (ROC-AUC) and pseudo-R^2. However, none of the coefficients were statistically significant at 5%, with p-values >0.05 (Table 3). To identify coefficients of significance, a stepwise regression approach, informed by domain expertise and exploratory data analysis, was employed to select a subset of the analytes for inclusion in the model. The final model included total moisture, sulfate sulfur, organic sulfur, and carbon as input variables. All coefficients were statistically significant at the 5% level (Table 4), indicating their strong influence on the bursting versus crushing classification. Additionally, model predictions are acceptable as shown in the ROC-AUC curve and the confusion matrix (Figures 6 and 7, respectively)

In the logistic regression model, a positive coefficient for a given feature indicated an increased likelihood

Table 3. Logistic regression results for model with all quantitative variables

Variable	Coefficient	Standard Error	P-Value
Total Moisture	1.328	1.052	0.207
Ash	0.203	0.893	0.820
Pyritic Sulfur	–0.125	0.882	0.888
Sulfate Sulfur	–1.298	0.908	0.153
Organic Sulfur	1.836	0.980	0.061
Carbon	–1.563	2.235	0.484
Hydrogen	–0.188	1.345	0.889
Nitrogen	0.165	2.723	0.952
Oxygen	0.405	1.097	0.711
Intercept	–0.151	0.438	0.731

Note: each variable is the CLR–transformed version

Table 4. Logistic regression results for model with statistically significant variables

Variable	Coefficient	Standard Error	P-Value
Total Moisture	2.4001	0.967	0.013
Sulfate Sulfur	–2.4766	1.041	0.017
Organic Sulfur	3.0487	1.321	0.021
Carbon	–3.0565	1.17	0.009
Intercept	–0.2935	0.507	0.562

Note: each variable is the CLR-transformed version

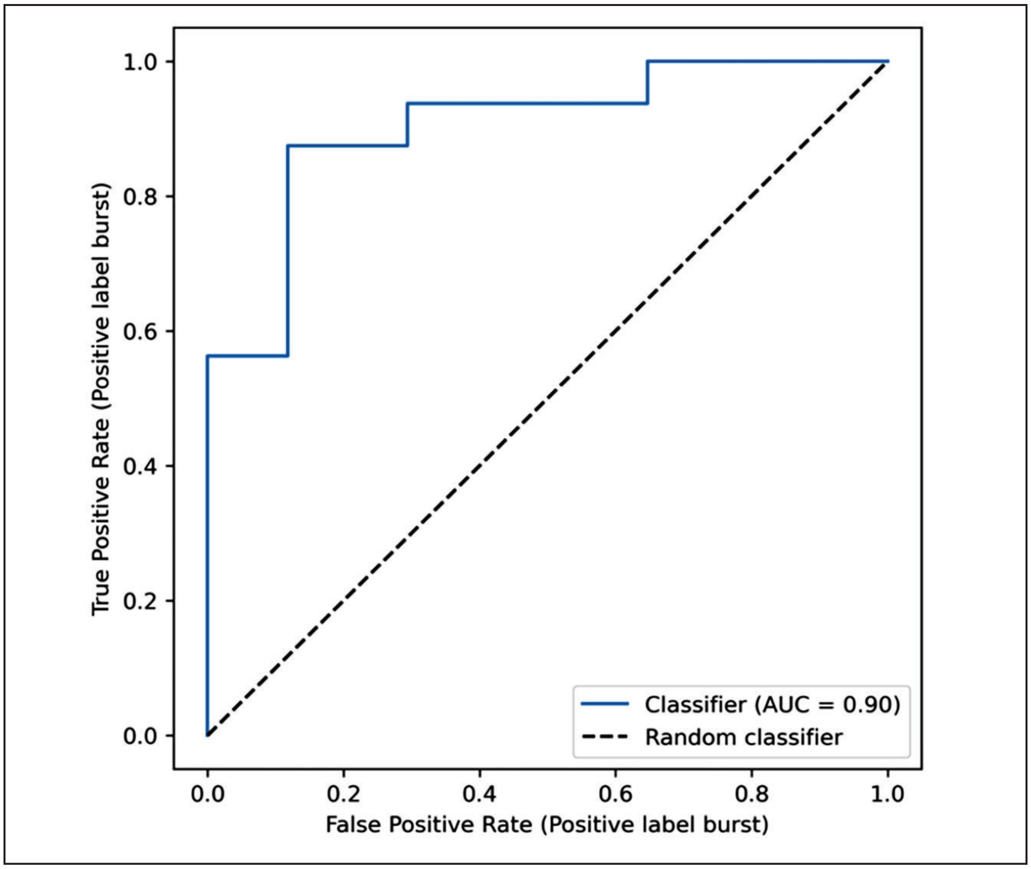


Figure 6. ROC0AUC curve for logistic regression model with chemistry data as model input

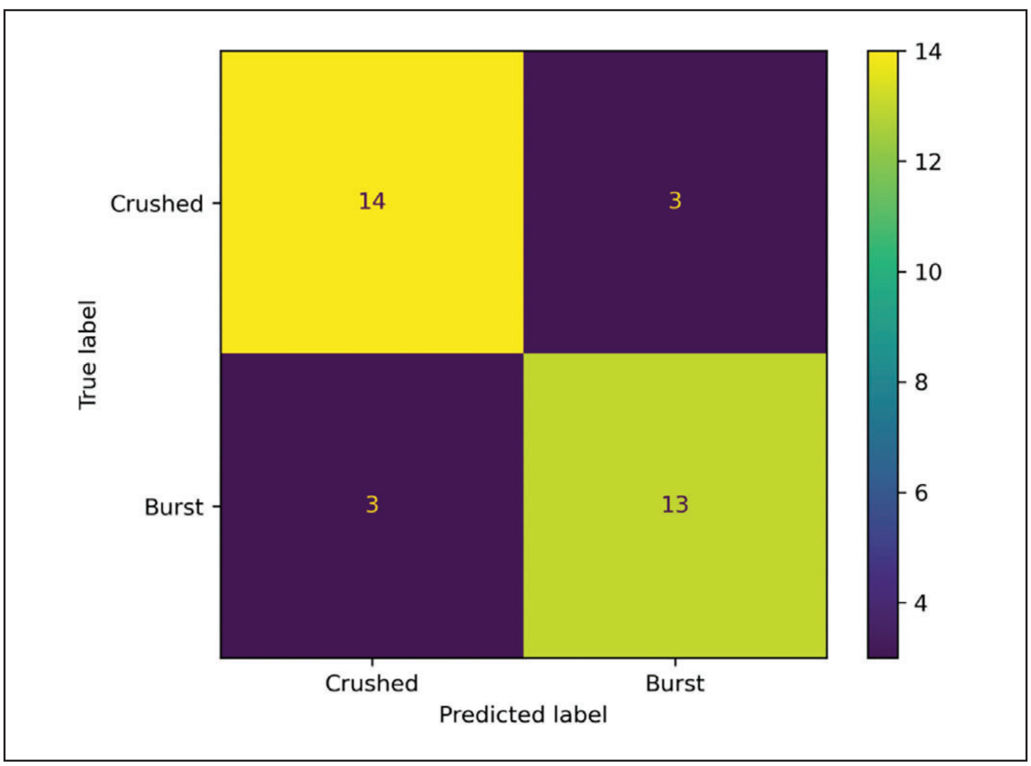


Figure 7. Confusion matrix for logistic regression model with chemistry data as model input

of bursting. Total moisture and organic sulfur were both positively associated with 'Burst', while other variables were more closely tied to 'Crush' classifications. Despite the complexity of the relationships, the logistic regression model performed well in distinguishing the two outcomes.

Overall, the logistic regression model confirmed a statistically significant relationship between coal chemistry and the 'Burst'/'Crush' classification, providing valuable insights into the role of chemical composition in coal seam stability.

DISCUSSION AND CONCLUSIONS

PCA analysis indicates that there is clear clustering of 'Burst' samples within the PCA space, suggesting that samples innately prone to 'Burst' behavior share similar compositional features as defined by the variables in the retained dataset. In short, this validates that 'Burst' behavior is influenced to some degree by compositional controls, regardless of other risk factors associated with the mining environment.

Analysis of cleating using an unsupervised learning DT approach indicated that coal samples that exhibited poor or very poor cleat development were more prone to 'Burst' behavior than coal samples with more well-developed cleat. This echoes observation made by Agapito et al., (2000) that poorly developed cleat is associated with dynamic failure events in coals in the United States. However, in terms of practical application, this observation may be somewhat limited as widescale mapping of cleat in coal mines may be impractical and subject to localized variability.

Logistic regression results further support this finding, through the strong anticorrelation of carbon with 'Burst' behavior. Put more simply, coals with higher compositional percentages of carbon are less likely to exhibit 'Burst' behavior. Compositional proportions of carbon increase with increased thermal maturity as hydrogen and oxygen are expelled as liquid and gaseous hydrocarbons. This corroborates the findings of previous studies drawing correlations between bursting behavior and low vitrinite reflectance values. Further, Laubach et al. (1998) and others have demonstrated repeated correlation of carbon content with degree of cleat development, such that, in the absence of annealing of cleat apertures through secondary mineralization or alteration of the coal fabric, it can be considered a reasonable quantitative proxy for cleat spacing.

Logistic regression results also indicate anticorrelation between sulfate sulfur and burst proneness and positive correlation between organic sulfur and induced bursting behavior. No strong relationship was determined for sulfide sulfur.

This has several implications:

- The first is the importance of sulfur speciation when considering sulfur content in the context of dynamic failure. This is also supported by PCA, which indicated that pyritic and sulfate sulfur were colinear in the PCA space, while organic sulfur was obliquely oriented. However, the reasons behind these apparent relationships merit further investigation. Chou (1997, 2012) and Calkins (1994) suggest that as coal becomes more thermally mature and its carbon content increases, the thiophene fraction of organic sulfur also increases, such that during diagenesis, organosulfur compounds undergo aromatization processes consistent with overall coal maturation processes. In other words, sulfur bonds tend to be relatively weak compared to other bonds typically found in organic molecules and may serve as catalysts for increasing aromatization in high sulfur coals. It may be possible then that the apparent correlation between organic sulfur and bursting behavior is a consequence of an inverse proportion of organic sulfur to carbon, such as organic sulfur decreases, overall carbon content increases.
- It may also be possible, however, that the apparent correlative and anti-correlative relationships between both sulfate and organic sulfur, respectively, are an artefact of the analytic approach in the absence of other, more compositionally significant variables. The amount of non-pyritic sulfur content in coal is generally negligible (Calkins, 1994; Chou, 2012), and this is true for this dataset as well. This cannot be said of the anticorrelation established between carbon and induced bursting, as carbon is the primary compositional constituent of coal.

Ash content does not appear to have a correlative relationship with coal behavior outside of the mining environment. The implication is that it is unrelated to a significant extent to the behavior of the coal itself and is more likely to in fact be a proxy for the depositional setting, as speculated by Lawson and Hanson (2021). This suggests a need for further research regarding the importance of stratigraphic risk in facilitating in mine dynamic failure events, as geochemical data suggest that deposition in sulfur-depleted environments with little water influx may be a near requisite condition for event occurrence in United States coal seams, regardless of failure mechanism. This further suggests a potential weighting of factors such that depositional setting is much more influential on dynamic failure risk in the United States than previously assumed, and that risk is also influenced to a lesser degree by degree of cleat development, for which percent weight of carbon may be a reasonable quantitative proxy. Further research into this area may produce a set of compositional attributes that can be used to predict environments at high risk for bursting behavior, allowing for proactive mitigative action to be taken before an accident occurs.

It is important to reiterate that these observations do not supersede the requirements for dynamic failure occurrence such as high or anomalously concentrated stresses and a triggering event. Nor do these findings imply that coals in depositional settings other than these will not experience dynamic failure. As validated by Babcock and Bickel (1984), most coals can be induced to bursting behavior under sufficient stress. Additionally, Mark (2018) clearly demonstrates that coal can experience dynamic failure in unexpected settings and that complacency in assessing risk is never merited. Rather, findings may suggest that coals deposited in sulfur-depleted settings may experience a lowered threshold for stress required to produce a dynamic failure event, and that this threshold may be lowered further by poor cleat development, which may in theory allow coal to retain greater energy prior to failure.

REFERENCES

Agapito, J. F. T. and Goodrich, R. R. (2000). Five stress factors conducive to bumps in Utah, USA, coal mines. In: Proceedings of the 19th International Conference on Ground Control in Mining. S. S. Peng, ed., Morgantown, WV: West Virginia University, p. 93– 100.

Aitchison, J. (1986). The Statistical Analysis of Compositional Data. Chapman and Hall. London.

Akdag, S., Karakus, M., Taheri, A., Nguyen, G., & Manchao, H. (2018). Effects of thermal damage on strain burst mechanism for brittle rocks under true-triaxial loading conditions. Rock Mechanics and Rock Engineering, 51(6), 1657–1682.

Akdag, S., Karakus, M., Nguyen, G. D., Taheri, A., & Bruning, T. (2021). Evaluation of the propensity of strain burst in brittle granite based on post-peak energy analysis. Underground Space, 6(1), p. 1–11.

Babcock, C. O., & Bickel, D. L. (1984). Constraint-the missing variable in the coal burst problem. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-84). ARMA.

Berry, C., Warren, S., and Hanson, D. (2019). Investigating the Correlation between coal geochemistry and coal bumps. In: Proceedings of the 38th International Conference on Ground Control in Mining. Morgantown, WV:171–177.

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.

Bodziony J, Lama RD (1996) Sudden outbursts of gas and coal in underground coal mines, Outbursting Scoping Study Final Report, ACARP Australia, p. 453–478.

Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. Chapman and Hall/CRC., p. 368.

Calkins, W. (1994). The chemical forms of sulfur in coal: a review. Fuel, Volume 73, Issue 4, p. 475–484.

Chou, C. (1997). Geologic factors affecting the abundance, distribution, and speciation of sulfur in coals. In: Yang, Q. (Ed.), Geology of Fossil Fuels, Proceedings of the 30th International Geological Congress: Part B, VSP, Utrecht, The Netherlands, Volume 18, p. 47–57.

Chou, C. (2012). Sulfur in coals: A review of geochemistry and origins. International Journal of Coal Geology, Volume 100, p. 1–13.

Cox, D. R. (1958). The regression analysis of binary sequences. Journal of the Royal Statistical Society Series B: Statistical Methodology, 20(2), p. 215–232.

Egozcue, J. J., & Pawlowsky-Glahn, V. (2011). Basic concepts and procedures. Compositional data analysis: Theory and applications, p. 12–28. DOI: 10.1002/9781119976462.ch2.

Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. Biometrika, 58(3), p. 453–467.

Greenacre, M. J. (2010). Biplots in practice. Fundacion BBVA.

Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. Nature Reviews Methods Primers, 2(1), p. 100. DOI: 10.1038/s43586-022-00184-w.

Han, J.., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques. 3rd ed., Morgan Kaufmann Publishers Inc.

Hanson, D. R., & Lawson, H. E. (2023). Using machine learning to evaluate coal geochemical data with respect to dynamic failures. Minerals, 13(6), p. 808.

Harrell, F.E. (2015). Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.

Hastie, T., Tibshirani, R. & Friedman, J.H. (2001). The Elements of Statistical Learning. Springer New York.

Holland C.T. & Thomas E. (1954).Coal mine bumps: some aspects of occurrence, cause and control. U.S. Bureau of Mines Bulletin 535, p. 36.

Hosmer, D.W., & Lemeshow, S. (2000). Applied logistic regression. New York: Wiley.

Iannacchione, A. T. and Zelanko, J. C. (1995). Occurrence and remediation of coal mine bumps: A historical review. Proceedings: Mechanics and Mitigation of Violent Failure in Coal and Hard- Rock Mines. Bureau of Mines SP 01–95: p. 27–67.

Kidybiński, A. (1981). Bursting liability indices of coal. In International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts (Vol. 18, No. 4, pp. 295–304). Pergamon.

Kidybiński, A. (2011). The effect of porosity and the strength of coal on the dynamics of coal and methane outburst-the BMP modelling. Archives of Mining Sciences, 56(3), p. 415–426.

Lawson, H., Weakley, A., & Miller, A. (2016). Dynamic failure in coal seams: Implications of coal composition for bump susceptibility. International journal of mining science and technology, 26(1), p. 3–8.

Lawson, H. E., Tesarik, D., Larson, M. K., & Abraham, H. (2017). Effects of overburden characteristics on dynamic failure in underground coal mining. International Journal of Mining Science and Technology, 27(1), 121–129.

Lawson, H. (2019). Exploration of petrographic, elemental, and material properties of dynamic failure-prone coals. In: Proceeding of the 38th International Conference on Ground Control in Mining. Denver, CO: Society for Mining, Metallurgy and Exploration (SME).

Lawson, H. (2020). Exploration of petrographic, elemental and material properties of dynamic failure prone coals. International Journal of Mining Science and Technology, 30(1), p. 69–75.

Lawson, H. & Hanson, D. (2021). The influence of depositional environment on dynamic failure potential in U.S. coal seams. Preprints of the 2021 SME Annual Meeting and Expo, Preprint 21–024. Society for Mining and Metallurgical Engineering (SME), Denver, CO.

Laubach, S. E., Marrett, R. A., Olson, J. E., & Scott, A. R. (1998). Characteristics and origins of coal cleat: a review. International journal of coal geology, 35(1–4), p. 175–207.

Li, D., Sun, Z., Xie, T., Li, X., & Ranjith, P. G. (2017). Energy evolution characteristics of hard rock during triaxial failure with different loading and unloading paths. Engineering Geology, 228, 270–281.

Lu, Z., Ju, W., Gao, F., Feng, Y., Sun, Z., Wang, H., & Yi, K. (2019). A new bursting liability evaluation index for coal–the effective elastic strain energy release rate. Energies, 12(19), p. 3734.

Mark, C., & Gauna, M. (2016). Evaluating the risk of coal bursts in underground coal mines. International journal of mining science and technology, 26(1), p. 47–52.

Mark, C. (2018). Coal bursts that occur during development: A rock mechanics enigma. *International Journal of Mining Science and Technology*, *28*(1), 35–42.

Menard, S. (2002). Applied logistic regression analysis. SAGE publications.

Mine Safety and Health Administration Data Retrieval System (MSHA MRDS) (no date). https://arlweb.msha.gov/drs/drshome.htm.

Obert, L. (1973). Rock Mechanics, Chapter 6–2, SME Mining Engineering Handbook, eds. Arthur B. Cummins, Ian A. Given, .p 6–13 to 6–52.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), p. 559–572. DOI: 10.1080/14786440109462720.

Pregibon, D. (1981). Logistic Regression Diagnostics. The Annals of Statistics, 9(4), p. 705–724.

Rice, G. (1935). Bumps in coal mines: Theories of causes and suggested means of prevention or minimizing effects. In: A.I.M.E. Transactions. New York, NY: American Institute of Mining, Metallurgical, and Petroleum Engineers.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. Journal of the Royal Statistical Society. Series B (Methodological). Vol. 36, No. 2, p. 111–147.

Van Rossum, G. & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace. https://docs.python.org/3/reference/.

Vardar, O., Zhang, C., Canbulet, I., and Hebblewhite, B. (2018). Numerical modelling of strength and energy release characteristics of pillar-scale coal mass. Journal of Rock Mechanics and Geotechnical Engineering, 11(5), p. 935–943.

Whyatt, J. (2008). Dynamic failure in deep coal: Recent trends and a path forward. In: Proceedings of the 27th International Conference on Ground Control in Mining. S. S. Peng et al., eds. Morgantown, WV: West Virginia University, p. 37–45.

Whyatt, J. and Varley, F. (2010). Regional bumps: Case studies from the 1958 bump symposium. In: Transactions of the Society of Mining, Metallurgy and Exploration. Denver, CO: The Society of Mining, Metallurgy and Exploration, p. 101–105.

Society for Mining, Metallurgy & Exploration

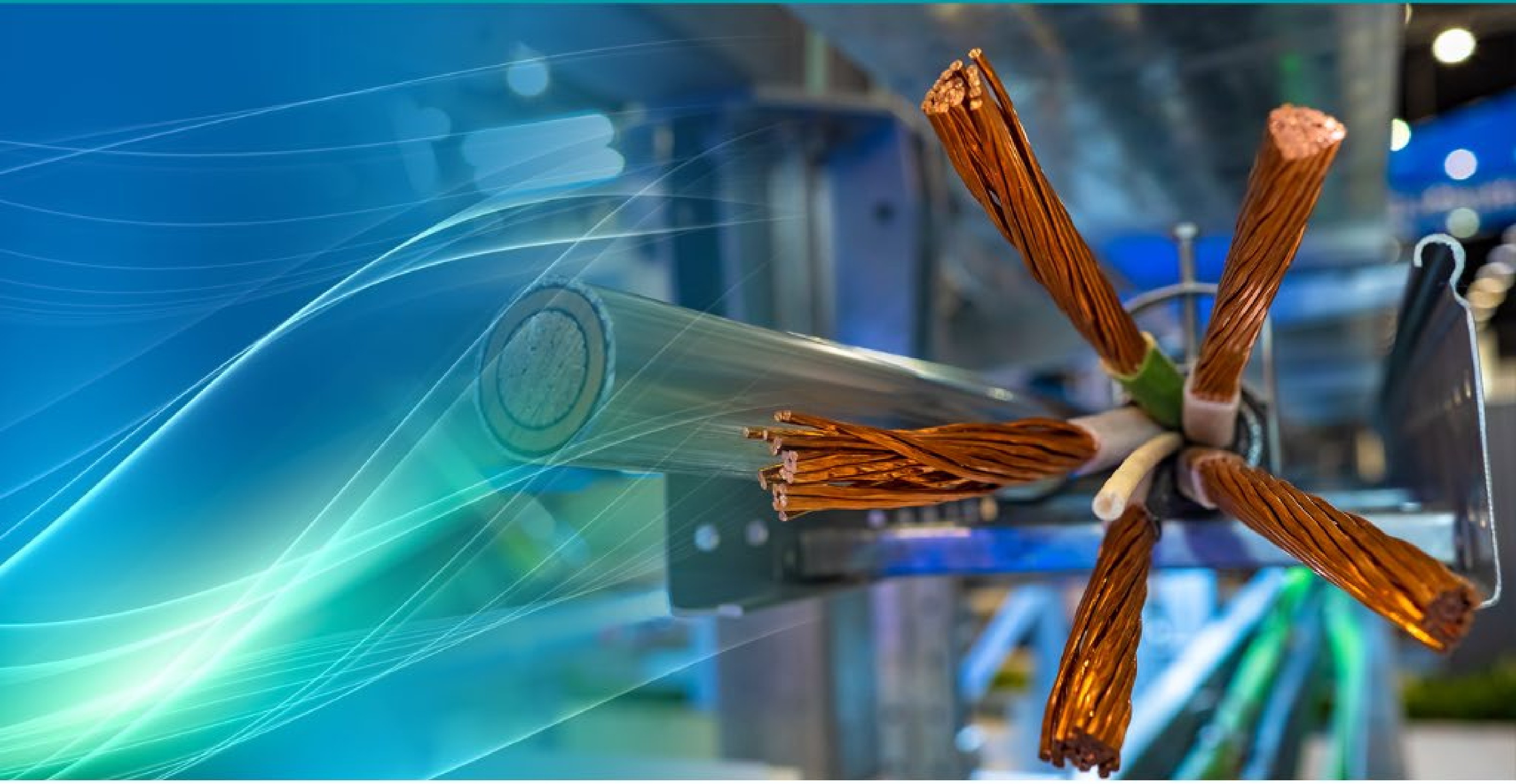
MINEXCHANGE

2026 SME ANNUAL CONFERENCE & EXPO

FEBRUARY 22-25, 2026

SALT LAKE CITY, UTAH

2026 Preprints



Society for Mining, Metallurgy & Exploration (SME)
12999 E. Adam Aircraft Circle
Englewood, Colorado, USA 80112
(303) 948-4200 / (800) 763-3132
www.smenet.org

SME advances the worldwide mining and underground construction community through information exchange and professional development.