Statistical analysis of the outlier events from seismic stress inversion

Y Abolfazlzadeh *Queen's University, Canada*

SL Penhall *Queen's University, Canada*

SD McKinnon *Queen's University, Canada*

Abstract

The seismic stress inversion method has been developed and used in the field of crustal seismology, but due to the technical challenges and lack of a mining-specific application methodology, it has rarely been used for the analysis of mining-induced seismicity. The technique is based on assumptions that the stress is uniform in the volume of study (less satisfied in mining environments and near geological structures), and slip on any fault plane occurs in the direction of maximum resolved shear stress. Currently, there is a variety of seismic stress inversion algorithms available which use multiple seismic events to constrain the fault plane solutions for determining the four parameters of the stress tensor (orientations of principal stresses and their relative magnitude). First motion stress inversion (MOTSI) is a widely used code that uses first motion data in the inversion. For a group of events, MOTSI runs a test on two factors; dS and dM. The parameter dS specifies rejection of the homogeneous stress hypothesis when it exceeds 2.32 (95% confidence level) and the parameter dM defines the similarity between the stress-constrained and unconstrained focal mechanisms (dM = 1, represents no change). Events with large deviations from acceptable levels of dS and dM should not be used in the stress inversion process. These events are expressed as outliers because there is not enough information to infer their relation to the other events. The goal of this paper is to use statistical techniques to explore the characteristics of these outlier events. Outliers are of particular interest in this case because by indicating non-compliance with the homogeneous stress hypothesis, they may be good indicators of zones where there is strong influence of geological structures on the stress field. The seismic events for this study are recorded over a period of ten months during the early stages of mining (to satisfy the uniform stress assumption) and from a specific area (to have a better coverage of focal sphere) of the Nickel Rim South Mine in Sudbury, Canada.

More than 500 seismic events were manually processed and after removing noise and applying different filters, we were able to determine the orientation of the principal stresses through seismic stress inversion from the middle level of the mine. The aggregate of the inversions inferred a north–south, northeast– southwest orientation of the maximum principal stress and near vertical minimum principal stress.

In each inversion processed, several events were tagged as outliers according to the mentioned criteria. Several statistical analyses including mean, standard deviation, variance, t-test, box plot, Kurskal–Walles test, and principal component analysis (PCA) were conducted over 10 seismic source parameters (i.e. source radius, seismic moment, etc.) to determine the relation between outliers and used data.

Keywords: *microseismic monitoring, seismicity, stress mapping, seismic stress inversion, statistical analysis*

1   Background

The following sections provide an overview of the Nickel Rim South Mine layout, mining methods and seismic monitoring system set-up. A summary of the seismic stress inversion technique completed with data collected from the mid-level mining zone is also summarised.

1.1 Nickel Rim South Mine

The Nickel Rim South Mine is an underground base metal operation located on the east side of the Sudbury Basin, Ontario, Canada. Open stoping is completed on three mining levels, ranging from 1.1 to 1.7 km depth below ground surface. Nickel comprises the hanging wall of the orebody, whereas copper comprises the footwall, totalling approximately 16 million tonnes of ore reserve. Primary, secondary, and tertiary stope sequencing is used, producing 2,000 to 4,000 tonnes per day since 2009. The seismic array has a good aerial coverage around both hanging wall and footwall for all levels. The middle level has better seismic resolution as it contains seven triaxial geophones and 11 uniaxial accelerometers (Jalbout & Simser 2014). Two shafts allow access to three main stoping levels, located at the 1,280 m level for the nickel deposit, 1,480 m level for nickel and copper transition zone and 1,660 m level for copper (Figure 1).
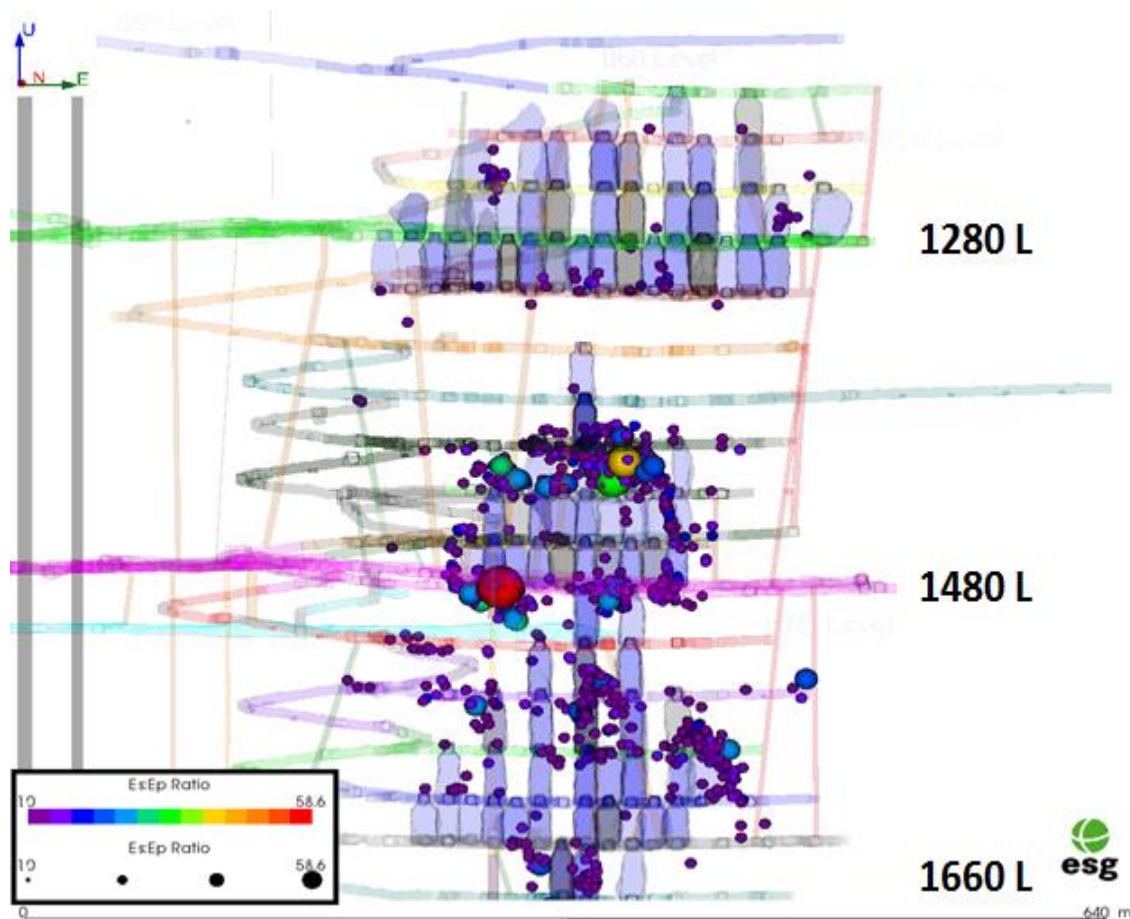

Figure 1   Three main stoping levels (looking north) and seismic events in 2011

1.2 Stress field characterisation using seismic stress inversion

Abolfazlzadeh and McKinnon (2017) used MOTSI (Abers & Gephart 2001) to analyse first motion data in the stress inversion process, and to estimate regional in situ stresses for the Nickel Rim South Mine. The MOTSI code generates a stress tensor by taking in the first motion polarity of seismic events. Clusters of seismic event data collected over a period of seven months during the early mining stages and with double couple component more than 50% were initially analysed to develop and refine the seismic inversion process.

MOTSI uses a statistic, denoted dS, to test for heterogeneity within a population of data. Events within a population are rejected where values of dS exceed 2.32 (95% confidence level) and the homogeneity stress hypothesis fails. A second statistic is employed, denoted dM, which represents the similarity between constrained and unconstrained focal mechanisms (a dM value of one represents no change between mechanisms).

The software then employs two parameters, α and γ, to constrain the volume of events with high confidence. These parameters describe the probability of the first motion pick being correct (Abers & Gephart 2001). The theoretical P-wave amplitude is represented by α. Picks below this value indicate those of decreased reliability. The estimate of the probability of a first motion P-wave pick being incorrect is approximated by the parameter γ. The parameters may be fixed at values of 0.1 and 0.2, respectively, to accommodate noisy data populations within a constrained volume.

For individual clusters of events, MOTSI can be used to estimate the spatial variations in stress conditions. This helps in dividing the seismically-active, three-dimensional volume into more manageable volumes. Stresses within these volumes are approximately constant and of lower stress gradient. Naturally, error is introduced when the size of the volume is increased, as the number of events per cluster is increased, which is observed as a spatial gradient of poor resolution. A decrease in data resolution within a particular volume relaxes the assumption of constant stress, which ultimately affects the results of an inversion.

Selecting a cluster of events is influenced by both the event rate and the time period over which the events occur. A mine that experiences high seismic activity will have a relatively high event count in a short time period. This is an ideal case, in that it represents a sufficient dataset for spatial discretisation, thus allowing a reliable stress inversion. Conversely, if seismic activity is infrequent or inconsistent, seismic events must be collected over a longer period to build a sufficient data population. Collecting less frequent seismic events over a longer time period tends to affect the reliability of a stress inversion, as it is possible that the volume of interest may experience changes in the induced stress field from progressive mining activity.

An ideal cluster of data will initially have moderate seismic activity with minimal stress gradient, followed by subsequent changes in the stress field as mining progresses. When the progression of mining is established, the data for the same volume may be examined over incremental time periods. This allows the determination of the sensitivity of the MOTSI inversion to the relaxation of stresses caused by mining-induced changes in stress field magnitude and orientation.

Data clusters were selected from the early stages of mining, well-contained within the system of uniaxial and triaxial sensors of the monitoring system. The clusters were also selected away from geological structures to avoid potential influence in stress change from mining or fault slip movement (Abolfazlzadeh & McKinnon 2017). More than 500 events over a period of seven months (February to September 2011) were analysed in several clusters using MOTSI.

The results of the stress inversion show that the analysed clusters of data from the initial mining stages are in reasonable agreement with the pre-mining stress estimates obtained from overcoring and borehole breakout observations (Figure 2).
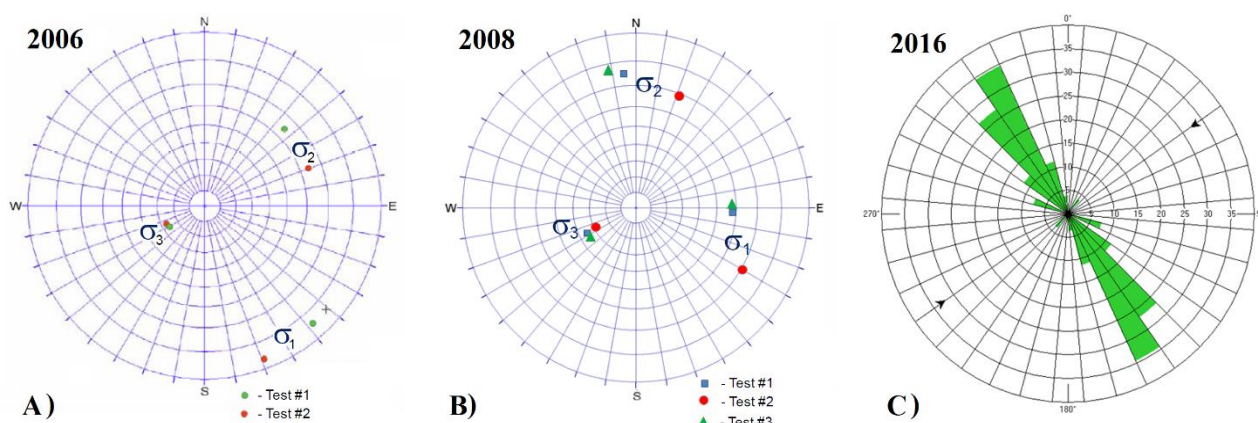

Figure 2 Orientation of major principal stresses obtained based on overcoring data from the test on 1480 level in: (a) 2006; and, (b) 2008 (Turichshev & Brummer 2008). (c) rose diagram for orientations of the borehole breakouts in 2016 (image courtesy of NRS) (Abolfazlzadeh & McKinnon 2017)

The compiled results of the seismic stress inversion over the mid-level mining zone implies an overall north–south, northeast–southwest major principal stress, σ_1, orientation. The orientation of the minor principal stress, σ_3, had an east–west trend with a plunge near vertical (Figure 3). Effects of localised mining-induced stresses were not removed from the results, and thus do not reflect the pre-mining state of stress required for the specification of numerical stress analysis boundary conditions.
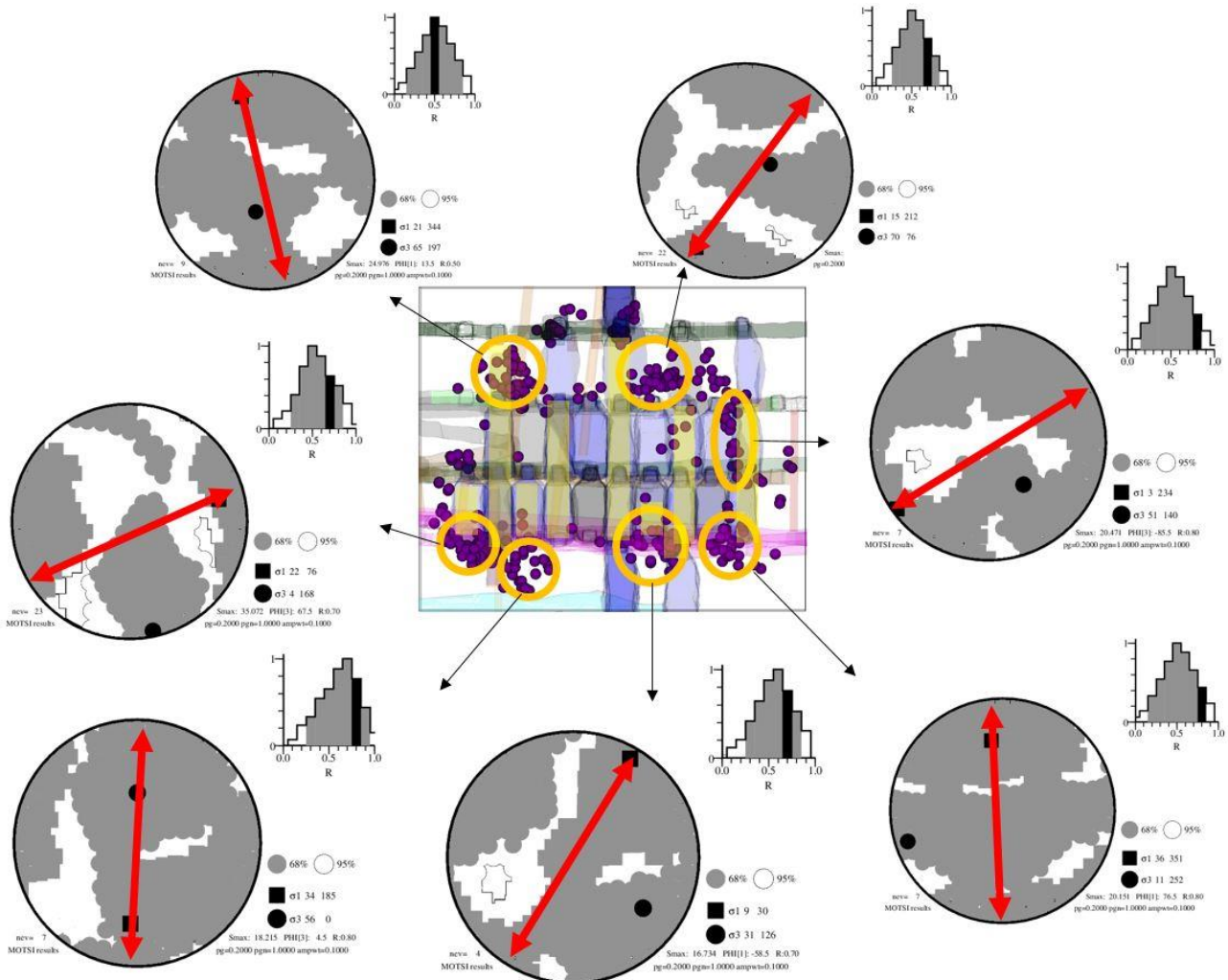

Figure 3   Location of the seismic events (purple circles), selected clusters (orange ellipsoids), final layout of the sublevel open stoping (blue rectangles), the extracted stopes (yellow rectangles), MOTSI inversion results (red arrows show σ_1 orientation) (Abolfazlzadeh & McKinnon 2017)

Manual selection of the data points significantly improved the identification and rejection of data points, resulting from the noisy acoustic environment of the mine activity, and the lack of automated event recording, detecting true first arrival polarities. Altering the size of the volume, the length of the time period, as well as the constant stress assumption largely influences the inversion result as larger volumes show high values of dS.

The seismic events that exhibited large deviations from the acceptable levels of dS and dM were not used in the stress inversion process, and thus were excluded from the analysis. Statistical techniques were employed to assess any additional characteristics of the seismic events that would contribute to the generation of outlier data. The methodology and results of the statistical analysis are described in the following sections.

2   Methodology

For various types of statistical analyses, it is necessary to have a sufficient number of samples. The time interval used for the stress inversions by Abolfazlzadeh and McKinnon (2017) was therefore extended by three months, resulting in an expanded time interval from February to mid-December 2011. Additional events were predominantly located in the upper right region of the middle level (Figure 4) and the stress inversion result was identical to the previous months. The outlier seismic events over the total period of ten months were examined through statistical analysis, to explore any potential trends in the additional source parameter data, the results of which are described in Section 3.

The outlier seismic events were examined relative to the events from the same time period that were included in the stress inversions in two steps. The first step included all events – both used and outlier – over the entire time period, and the second step assessed the events over two distinguished event clusters (B and G in Figure 4) to consider the effect of mining location.
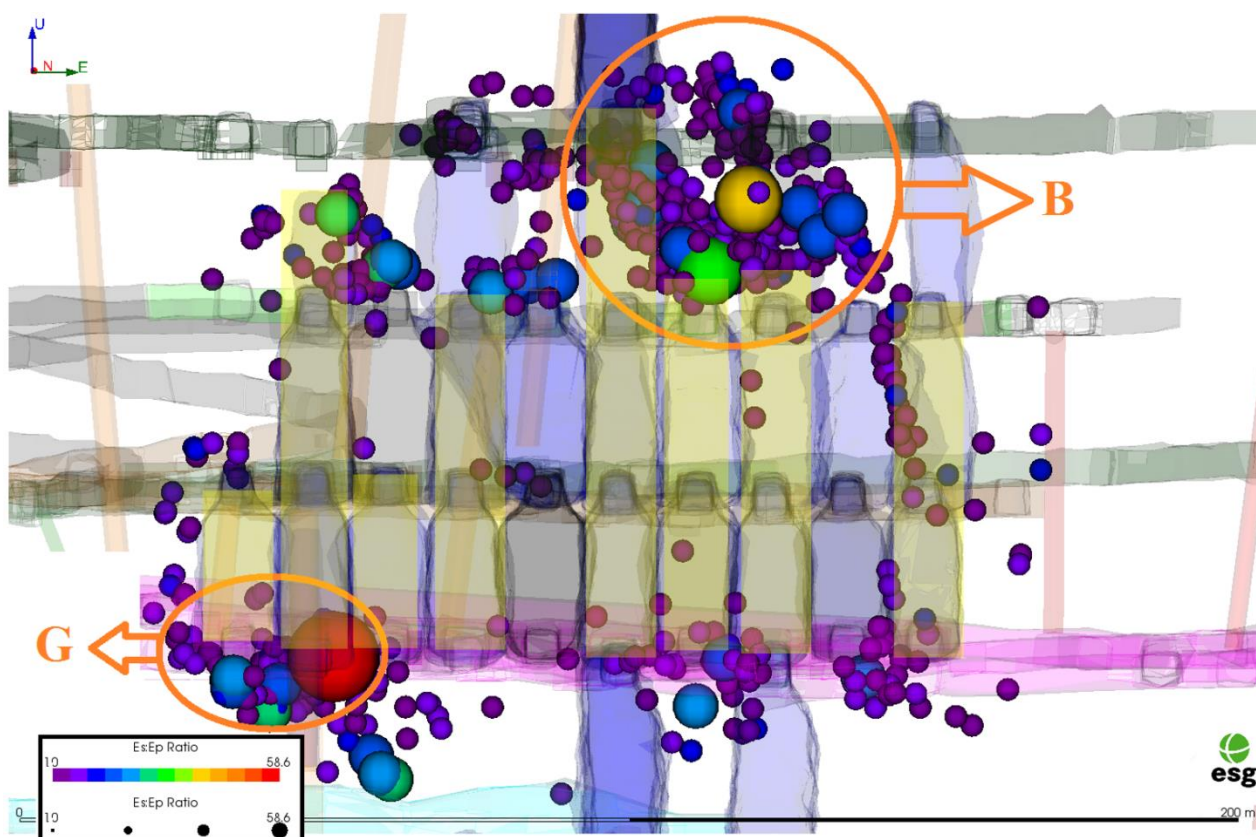

Figure 4   Middle level seismic events (purple circles), clusters assessed (B and G), final layout of the sublevel open stoping (blue rectangles), and the extracted stopes (yellow rectangles), looking north

The source parameters examined for each event include seismic moment, seismic energy, E_s/E_p, source radius, asperity radius, static stress drop, apparent stress, dynamic stress drop, peak particle velocity (PPV), and peak particle acceleration (PPA). The locations of outlier and used events within each previously identified cluster were also closely examined. The Statistical Package for the Social Science (SPSS) software suite was employed in the statistical analysis. Descriptive statistics, including the mean, minimum, maximum, standard deviation and variance were determined for source parameters of each seismic event, to quantitatively describe the population of used and outlier data. Box plots were produced as a visual depiction for the 25th percentile (first quartile), median, and 75th percentile (third quartile), for each source parameter. Box plots are useful for comparing multiple distributions of variables (in this case, source parameters), as well as identifying any outliers within a population of data. The box plots were arranged to compare used events against outlier events, for each source parameter.

The t-test was used as it evaluates the statistical significance between two sets of data, through a hypothesis test, and is determined by the magnitude of the difference between the mean values of the two datasets. Outlier events were tested relative to the used events, within each cluster of seismic events.

The final statistical analysis employed was the PCA. PCA is a technique that reduces a relatively large dataset to a set of components or factors. This reduction allows the interpretation of a large population of data in a smaller and possibly more meaningful way. Seismic events, both outliers and those used in the stress inversion, were analysed by data cluster.

3   Results

The results of descriptive analysis showed very small differences in minimum, maximum, mean, standard deviation and variance for all parameters associated with the outlier and used events in both clusters, implying that the used and outlier events are very similar (Figure 5).

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Log of Seismic Moment_U	167	6.94	9.35	7.5905	.50890	.259
Log of Energy_U	167	.26	3.80	1.4989	.74794	.559
Es/Ep_U	167	10.01	33.90	13.6197	3.45559	11.941
SourceRo_U	167	1.83	7.84	3.2986	1.07335	1.152
AspRadius_U	167	.32	3.06	.7362	.39610	.157
Log of Static SD_U	167	5.01	6.48	5.7235	.32090	.103
Log of AppStress_U	167	3.71	5.13	4.4500	.30313	.092
Log of DyStressSD_U	167	5.79	7.18	6.5345	.31083	.097
PeakVelPar_U	167	.00	.28	.0363	.04839	.002
Log of PeakAccPar_U	167	4.48	5.76	5.2694	.28003	.078
Valid N (listwise)	167					

(a)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Log of Seismic Moment_O	90	6.95	9.64	7.6701	.57051	.325
Log of Energy_O	90	.31	4.52	1.5608	.80919	.655
Es/Ep_O	90	10.11	58.59	14.5468	6.16521	38.010
SourceRo_O	90	1.82	7.72	3.6064	1.29996	1.690
AspRadius_O	90	.35	5.61	.8090	.73244	.536
Log of StaticSD_O	90	5.07	6.65	5.7012	.30903	.096
Log of AppStress_O	90	3.77	5.42	4.4322	.29773	.089
Log of DyStressD_O	90	5.72	7.28	6.5470	.31015	.096
PeakVelPar_O	90	.00	.55	.0416	.07825	.006
Log of PeakAccPar_O	90	4.39	5.78	5.2623	.27312	.075
Valid N (listwise)	90					

(b)

Figure 5   Descriptive analysis of used events for 10 parameters (a); descriptive analysis of outliers for 10 parameters (b)

The results of the box plot analysis of all events (Table 1) shows that the outliers have higher first quartiles in seismic moment, energy, E_s/E_p, and source radius. And used events have a higher 3rd quartile in energy, asperity radios, and apparent stress. Among all parameters, it seems used events have a higher asperity radius, as they have higher quartiles than outliers. However, the source radius in the outliers is showing higher values in all quartiles. Peak particle acceleration also exhibited a higher median value in all groups in used events. Higher 3rd quartile source radius and median in seismic moment in all groups belong to outliers.

Table 1   Representation of box plot results of different groups with their highest quartile and their parameters

Parameter	Used			Outlier		
	1st quartile	Median	3rd quartile	1st quartile	Median	3rd quartile
Log of seismic moment				All, B, G	All, B, G	
Log of energy			All, B, G	All, B, G		
E_s/E_p				All, B, G		
Source radius				All, B, G	All, B, G	All, B, G
Log of asperity radius	All, B, G	All, B, G	All, B, G			
Log of apparent stress			All, B, G			
Log of PPA		All, B, G				

In the t-test, the null hypothesis is that the variances of the two groups are approximately equal (similar distribution of competency scores). The level of significance was set to 0.05. The results of the t-test over all groups showed that the source radius and the asperity radius have values less than 0.05, indicating the variances of the used and outliers are significantly different or the used and outlier groups are not similar (Figure 6). But the results of the t-test on B and G groups showed that none of the parameters (in both used and outlier datasets) had a value less than 0.05. This implies that the variances of the used and outliers are not significantly different or the used and outlier groups are similar. A verification using the Kruskal–Wallis test also resulted in the same outcomes as t-test for groups B and G.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Log of Seismic Moment_U_O	Equal variances assumed	.739	.391	-1.147	255	.253	-.07965	.06946	-.21645	.05714
	Equal variances not assumed			-1.108	165.387	.269	-.07965	.07188	-.22158	.06227
Log of Energy_U_O	Equal variances assumed	.014	.906	-.615	255	.539	-.06191	.10067	-.26017	.13634
	Equal variances not assumed			-.601	170.449	.549	-.06191	.10308	-.26539	.14156
Es/Ep_U_O	Equal variances assumed	3.326	.069	-1.546	255	.123	-.92708	.59980	-2.10827	.25411
	Equal variances not assumed			-1.319	119.846	.190	-.92708	.70273	-2.31846	.46430
SourceRo_U_O	Equal variances assumed	4.890	.028	-2.034	255	.043	-.30788	.15136	-.60595	-.00981
	Equal variances not assumed			-1.921	155.181	.057	-.30788	.16024	-.62441	.00864
AspRadius_U_O	Equal variances assumed	4.360	.038	-1.035	255	.302	-.07277	.07034	-.21130	.06575
	Equal variances not assumed			-.876	117.698	.383	-.07277	.08307	-.23727	.09173
Log of Static SD_U_O	Equal variances assumed	.376	.541	.538	255	.591	.02231	.04143	-.05928	.10389
	Equal variances not assumed			.545	188.381	.587	.02231	.04096	-.05849	.10311
Log of AppStress_U_O	Equal variances assumed	.291	.590	.450	255	.653	.01774	.03939	-.05984	.09532
	Equal variances not assumed			.453	185.218	.651	.01774	.03918	-.05956	.09504
Log of DyStressSD_U_O	Equal variances assumed	.052	.821	-.309	255	.757	-.01257	.04061	-.09255	.06741
	Equal variances not assumed			-.310	182.724	.757	-.01257	.04059	-.09265	.06751
PeakVelPar_U_O	Equal variances assumed	1.665	.198	-.666	255	.506	-.00527	.00791	-.02085	.01031
	Equal variances not assumed			-.582	126.580	.562	-.00527	.00906	-.02319	.01266
Log of PeakAccPar_U_O	Equal variances assumed	.180	.672	.196	255	.844	.00713	.03631	-.06437	.07863
	Equal variances not assumed			.198	186.342	.843	.00713	.03603	-.06396	.07822

Figure 6   T-test results of used versus outliers for all events

The PCA results showed up to three components in used and two components in outliers. For a more detailed assessment in the difference between used and outliers, the first and second components of used and outliers in each group were compared in sequence, with the intention of finding which parameter had less correlation in each component. The PCA results for all events are shown in Figure 7, where component one in the used events is very similar to the component one of outliers, with a high correlation of parameters, except E_s/E_p. Resemblance of components can also be observed in the second component, where there is a high correlation of E_s/E_p. This pattern was identical in groups B and G.
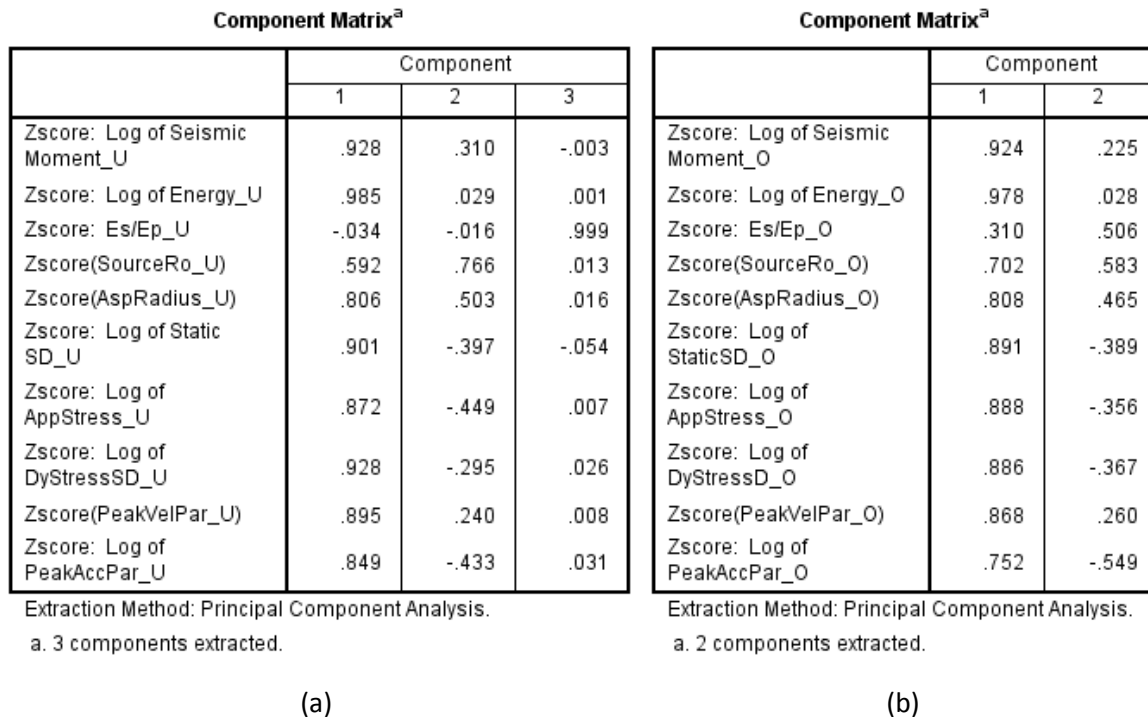
Component Matrix^a

	Component		
	1	2	3
Zscore: Log of Seismic Moment_U	.928	.310	-.003
Zscore: Log of Energy_U	.985	.029	.001
Zscore: Es/Ep_U	-.034	-.016	.999
Zscore(SourceRo_U)	.592	.766	.013
Zscore(AspRadius_U)	.806	.503	.016
Zscore: Log of Static SD_U	.901	-.397	-.054
Zscore: Log of AppStress_U	.872	-.449	.007
Zscore: Log of DyStressSD_U	.928	-.295	.026
Zscore(PeakVelPar_U)	.895	.240	.008
Zscore: Log of PeakAccPar_U	.849	-.433	.031

Extraction Method: Principal Component Analysis.
a. 3 components extracted.

(a)

Component Matrix^a

	Component	
	1	2
Zscore: Log of Seismic Moment_O	.924	.225
Zscore: Log of Energy_O	.978	.028
Zscore: Es/Ep_O	.310	.506
Zscore(SourceRo_O)	.702	.583
Zscore(AspRadius_O)	.808	.465
Zscore: Log of StaticSD_O	.891	-.389
Zscore: Log of AppStress_O	.888	-.356
Zscore: Log of DyStressD_O	.886	-.367
Zscore(PeakVelPar_O)	.868	.260
Zscore: Log of PeakAccPar_O	.752	-.549

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

(b)

Figure 7   Comparison of the PCA results of used (a); and outliers (b) for all events

4   Conclusions

The MOTSI stress inversion software identifies events suitable for inclusion in the inversion process (used events) and those that are not (outliers). Several statistical analyses have been carried out on the source parameters of used and outlier events to determine if there are additional criteria that could be used to identify events unsuitable for the inversion process. The collective results show that the two groups of events are fundamentally similar. There exists some evidence demonstrating that the outliers of all events from the middle level of the Nickel Rim South Mine have higher seismic moment, energy, source radius, and E_s/E_p in their first quartiles, while used events have a higher 3rd quartile in energy, asperity radius, and apparent stress. It seems the source radius and the asperity radius are the parameters, which can be used to differentiate between the used and outliers as they have higher values in all quartiles in the used versus outlier events for all groups and also the t-test showed that the variances of these parameters are significantly different in the used and outlier events. The observed trend does not consistently extend to the two event clusters above and below the mid-level mining area. This implies that the results are influenced by the cluster location.

Plotting the location of outlier events revealed that there is no specific location in the mine where these outlier events originate. Similar to the used event data, the outliers are scattered chaotically throughout the data cluster (typical nature of seismic events). The dS and dM parameters within the MOTSI software remain as the most effective determinants of outlier events, which must be removed from the stress inversion process.

Acknowledgement

References

Abers, GA & Gephart, JW 2001, 'Direct inversion of earthquake first motions for both the stress tensor and focal mechanisms and application to southern California', *Journal of Geophysical Research*, vol. 106, no. 26, pp. 523–526.

Abolfazlzadeh, Y & McKinnon, SD 2017, 'Stress field characterisation in Nickel Rim South Mine using seismic stress inversion', in J Wesseloo (ed.), *Proceedings of the Eighth International Conference on Deep and High Stress Mining*, Australian Centre for Geomechanics, Perth, pp. 247–257.

Jalbout, A & Simser, B 2014, 'Rock mechanics tools for mining in high stress ground conditions at Nickel Rim South Mine', in M Hudyma & Y Potvin (eds), *Proceedings of the Seventh International Seminar on Deep and High Stress Mining*, Australian Centre for Geomechanics, Perth, pp. 189–208.

Turichshev, A & Brummer, RK 2008, *In-Situ Rock Stress Measurements at the 1480 Level of the Nickel Rim South Project*, Itasca Consulting Canada Inc., Xstrata Nickel, Nickel Rim South Project.