Engineering Geology 232 (2018) 104–113

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

A spatially focused clustering methodology for mining seismicity

Kyle Woodward*, Johan Wesseloo, Yves Potvin

Australian Centre for Geomechanics, The University of Western Australia, 35 Stirling Highway (M600), Crawley, WA 6009, Australia

A B S T R A C T

Mining seismicity is routinely observed to cluster in space and time due to the spatially distinct rock mass failure processes associated with the temporally dependent process of mining. Assessment of clustered seismicity is important to develop an understanding of and to quantify seismic hazard that is associated with mining.

This article presents a density-based clustering method that is applicable to the assessment of 3D spatial distributions of short-term seismicity. The methodology presented in this article is developed from existing approaches that address the general limitations of density-based clustering algorithms.

Synthetically generated seismicity allows for the assessment of the methodology with respect to external and internal performance measures. The clustering of a dataset with known attributes allows for confidence to be developed in the capability of the clustering method. Additionally, this internal performance evaluation can represent the relative accuracy of outcomes without prior information concerning dataset attributes.

The clustering method is applied to two case studies of mining seismicity. These cases illustrate the general applicability of the clustering method along with the value of evaluating internal performance measures when optimising the selection of parameters and understanding the sensitivity of clustering outcomes to these choices.

1. Introduction

The general spatial characteristics of mining seismicity are controlled by the factors that influence the rock mass failure process. It is this rock mass failure that manifests as a source of seismicity which generates a number of events over a range of magnitudes (Hudyma and Potvin, 2010). Assessing the spatial distribution of seismicity forms the basis for understanding and quantifying seismic hazard associated with mining (Wesseloo, 2014). Multiple sources of seismicity may exist in close proximity to mining excavations and contribute to an overall spatial distribution of seismic events (Hudyma et al., 2003). Fig. 1 shows a hypothetical open stoping operation that experiences seismicity generated by a contrast in rock mass properties, local rock mass failure near stopes, crushing of pillars, stress increase in pillars, and slip on geological features.

Sources of seismicity can exhibit a strong temporal change due to mining activity and as such, the timing of events may not be neglected when spatially clustering seismicity. Spatial clustering can only be performed on subsets of data for which the spatial component did not change. The temporal definition of subsets of data depends on the scope of analysis. This article will focus on short-term subsets of events that are generated by time dependent sources of seismicity (e.g. on the order of days). While further discussion is outside the scope of this article, the definition of short-term subsets can be found as part of a broader methodology to identify and delineate time dependent seismicity in Woodward (2015).

Assessing the spatial distribution of seismicity typically requires the application of a clustering procedure in order to isolate events associated with each individual source of seismicity. The most fundamental aspect of clustering procedures is that elements which share similar characteristics are grouped together (Jain et al., 1999). This condition does not necessarily mean that elements closest in space cluster together, but instead focuses on identifying the underlying structures present within a dataset. The most fundamental question when clustering seismicity is establishing what shared event characteristics should form the basis for clustering. Generally, analysis of mining seismicity aims to delineate spatial clusters of events of variable size, shape, and density.

Studies typically assess the spatial distribution of seismicity, although, most studies will consider additional event characteristics. Examples of these studies include:

- Hudyma (2008) proposed a two-pass spatial clustering methodology;
- Malek and Leslie (2006) presented a method that adopts a normalised entropy metric to represent the degree of spatial clustering;
- Frohlich and Davis (1990) presented Single Link Clustering (SLC) which uses a simple metric to define a space-time distance between

* Corresponding author.
E-mail address: kyle.woodward@research.uwa.edu.au (K. Woodward).

https://doi.org/10.1016/j.enggeo.2017.11.015
Received 23 May 2017; Received in revised form 20 November 2017; Accepted 20 November 2017
Available online 21 November 2017
0013-7952/

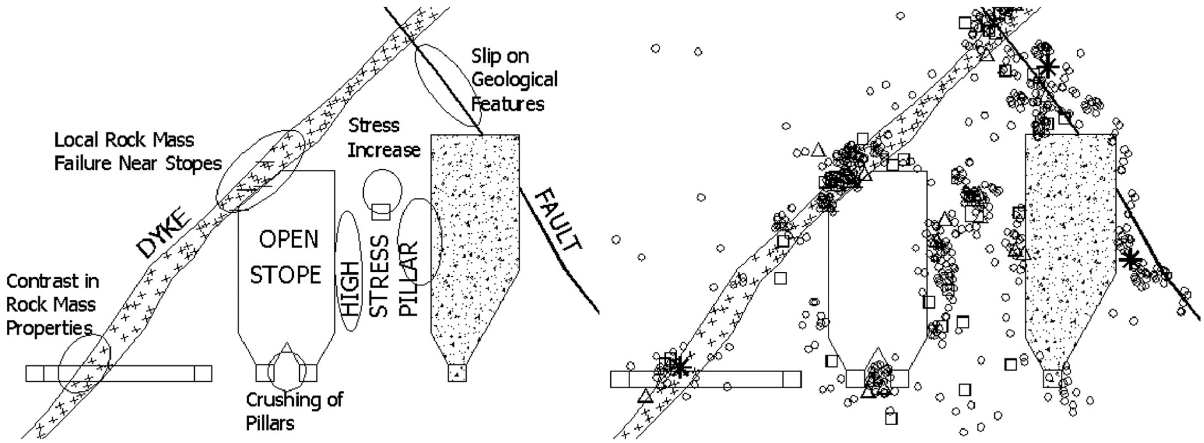

Fig. 1. Hypothetic open stope mining environment showing sources of seismicity (left) and seismicity typically associated with these sources (right) (Hudyma et al., 2003).

seismic events;
- Falmagne (2001) focused on approximating rock mass damage from fracture coalescence by assessing the distance between two events and an effective seismic source radius;
- Cho et al. (2010) implemented the Thirumalai-Mountain metric based on spatial and temporal event occurrence; and
- Rebuli and Kohler (2014) considered density-based clustering in conjunction with a parameter standardisation procedure. This clustering approach considered the location of events with similar temporal, E_S/E_P ratio, and energy index characteristics.

A limitation of methods that consider multiple parameters, is that the cause of strong clustering can become ambiguous as it is not clear which characteristics dominate the measure of similarity. Furthermore, arbitrary scaling factors need to be introduced to allow for the clustering of different characteristics of different units. These assessments also assume that events will share similar characteristics in location, timing and/or source parameter to define a strong measure of clustering. It is preferable to consider event location for a generalised measure of clustering due to these limitations and assumptions.

The ideal characteristics of clustering methods that spatially delineate events are not unique to mining seismicity and are applied to a large number of practical and research applications. As a result, there is a significant amount of literature related to parametric and non-parametric clustering techniques (Jain et al., 1999). Density-based clustering methods are well suited to assess the spatial distribution of seismicity but suffer from several shortcomings, i.e. selecting appropriate parameters, high outcome sensitivity to parameters, and a poor performance for datasets with varying element densities.

Although the underlying concepts of density clustering are generally applicable to the clustering of seismicity, the prevalence of specialised approaches indicates the need for algorithms to be tailored to address general and problem-specific limitations. To achieve consistent application to mining seismicity, density-based methods need to address the clustering of datasets with varying densities and the sensitivity of outcomes to clustering parameters.

This article presents a density-based clustering methodology applicable to mining seismicity and focuses on the spatial assessment of short-term seismicity. A density-based clustering methodology is adapted to address the relevant shortcomings associated with alternative methods. Two modifications are introduced which decrease the sensitivity of outcomes to parameter selection and improves the clustering of datasets with varying event densities.

This article presents a performance evaluation of the proposed clustering methodology. This evaluation is considered with respect to external and internal measures that have been implemented from alternative fields of research. The performance of the methodology is quantified and evaluated by applying the algorithm to synthetic and real seismic data. Furthermore, this article applies the clustering methodology and performance evaluation to synthetic and real examples of mining seismicity.

2. Density-based clustering of mining seismicity

Clustering approaches vary significantly based on the study in question (Jain et al., 1999; Xu and Wunsch, 2005). Clustering is considered in two broad categories: parametric and non-parametric. Parametric approaches produce clusters by the optimisation of a function that describes the likelihood of elements belonging to a set of assumed clusters. These approaches generally require underlying assumptions of the structure of the data and are not suited to the clustering requirements of seismic cluster delineation. In contrast, non-parametric approaches do not require assumptions concerning data structure and will group elements based on similarity (agglomerative) or disassociate elements based on differences (divisive).

Density-based clustering is a non-parametric method suited to mining seismicity for the following reasons (Ester et al., 1996; Kriegel et al., 2011):

- Distinct class identification: Seismic events should be allocated to one unique cluster;
- Minimal requirements of existing dataset knowledge: Assumptions concerning spatial distributions are not likely to be generally applicable due to the variation in sources of seismicity and mining environments;
- Discovery of clusters with arbitrary shapes: Allows for various shaped clusters to be clustered as spatial distributions are controlled by sources of seismicity, e.g. planar faults, spherical stress changes, or cylindrical pillars; and
- Discovery of high-density clusters within low-density areas: Required for the identification of high-density clusters superimposed with sparse clusters.

2.1. DBSCAN

A simple density-based method is Density-Based Spatial Clustering of Applications with Noise (DBSCAN) proposed by Ester et al. (1996) and provides the general framework for density-based approaches. The DBSCAN approach classifies elements as a core, boundary, or noise element by considering the number of neighbouring elements (N_e) with respect to a user-specified minimum (N_{MIN}) within a search distance (D_S). DBSCAN creates clusters from adjacent core elements and their neighbours. Core events are recursively considered and merged if one or more core elements are shared (Ester et al., 1996). The element