Application of Geostatistics for Cleat Data Validation in Sungai Lilin, Musi Banyuasin Region, South Sumatra

Stevanus Nalendra Jati¹, Idarwati¹, Ildo M. Putra¹, Fadlan Atmaja Nursiwan¹
¹Geology Study Program, Sriwijaya University.
s.nalendra@unsri.ac.id

Abstract. Cleat on coal seam is a slope stability controller, as well as permeability and porosity for fluidal flows. It also has become an interesting issue on coal mining, nevertheless the finite quantitative data. As a matter of facts, there is actually some available data about cleat as attributes, but this study only objecting on the orientation, cleat spacing, and cleat aperture. Cleat attributes measurement through field observation at two coalfields; Baturona and PMC in the area of Sungai Lilin, Musi Banyuasin Region. Next, those 667 faces and butt cleats were measured from 8 scanlines being analyzed using geostatistic methods. The purpose of geostatistic analysis in this study is to see the relation between variables (cleat attributes) to evaluate data quantitatively until the normal distribution is obtained. The data filter was done as a treatment until valid data is reached through the boxplot model. Early boxplot result that almost entirely had outliers data in each seam. In fact, it is commonly seen some extreme outlier data. This suggests that outlier data filter is treated as a treatment until a normal distribution system is achieved on each cleat attribute. Identification of cleat outliers has been affected by mining activity, or due to weathering processes in the coal seams that have been exposed.

1. Introduction
One of the most interesting things to explore in coal mining is fracture or cleat that formed in an inner part of the coal seam. According to Kuncoro et al. (2007) cleat becomes study object because it relates to the presence of pyrite mineral inside the coal seam, coal quality, CBM exploration and exploitation, mining activities (mining direction, mining location, mining technology application, and slope stability), stacking, up to marketing. Hence, it’s not a surprise that cleat has been known for years in coal mining, and study/paper regarding to cleat has been published long with fractures suspecting since 19th century (Kendall and Briggs, 1993, in Laubach et al., 1998).

Geostatistical application in this study was done with a boxplot model in order to validate cleat data distribution. The boxplot can provide information about the range, mean, median, normality and slope of the data distribution (Sun and Genton, 2011). Boxplot’s advantages that can be optimized are the ability to visualize outlier data from a data (Andrea et al., 2013). The existence of outlier data will interfere in the process of advanced data analysis and should be avoided in many ways, including cleat data.

The study area was conducted at two coal mining companies, one is PT. Baturona Adimulya and the other is PT. PMC (Fig. 1). Both are administratively located in Sungai Lilin, Musi Banyuasin Region, South Sumatra Province. PMC can be reached through Palembang-Jambi crossing Km. 105, and Baturona can be reached through Palembang-Sarolangun cross road, continued to Supat Babat village.
Tectonically, study area is a back arc basin, located in the South Sumatra Basin in the part of Tamiang Highland section north of the North Palembang Sub-basin (Figure 1). Images of the map in this page contains a folded path with two anticlinorium known as Tamiang Anticline, located on the southern part of the sheet cut by two horizontal faults and one normal fault; in the north area known as Anticline Bentayan cut by three strike slip faults. The general direction of the hinge line is NW-SE. Stratigraphic of study area includes three formations, from the oldest to earlier, Namely Airbenakat, Muaraenim, Kasai Formation, as well as swamp and alluvial deposit. All three of the formations constituents are lithology evenly distributed in the fold system of Tamiang Anticlinorium and Bentayan Anticlinorium. The Muaraenim formation is the coal bearing formation, broadly open on anticline flanks. In this formation, vast amount of coal outcrops is found; all of them are part of M4 coal seam-group (Shell, 1978).

The dip on the south limb of the Tamiang Anticline ranges from 8-12° with the NE 110-140°, while the dip on the north limb is generally sloppierat 5-9° with the NE 275-320°. On the south limb of the Bentayan Anticline, the dip is steeper, range from 22-40° with a NE 102-130°. This formation was deposited as a continuation of the regression phase, consisting intercession of sandstone with mudstone and little amount of clay stone, siltstone and coal, aged in Late Miocene-Early Pliocene with a transitional deposit environment.

Figure 1. Position of study area based on tectonic settings (left). Geological maps show the position of both areas, namely Baturona and PMC with different structural control intensities (Gafoer et al., 1995).

The dip on the south limb of the Tamiang Anticline ranges from 8-12° with the NE 110-140°, while the dip on the north limb is generally sloppierat 5-9° with the NE 275-320°. On the south limb of the Bentayan Anticline, the dip is steeper, range from 22-40° with a NE 102-130°. This formation was deposited as a continuation of the regression phase, consisting intercession of sandstone with mudstone and little amount of clay stone, siltstone and coal, aged in Late Miocene-Early Pliocene with a transitional deposit environment.

Based on Fig. 1, spatially, Baturona is located in the fault zone, which is a part of Anticline Keluangpatterns. Whereas the PMC area is not in the geological structure zone so it is relatively more stable. This will affect the development of the cleat system in the coal seam from each location.

2. Data and Method
This study is descriptive-observative, so the data source is based on the results of field observations, then supported by geostatistical analysis. The methods used are data acquisition, data analysis, and synthesis. Data acquisition or cleat data retrieval applies the scan line method by measuring the orientation of the cleat’s field (face and butt), which intersects the observation lines stretched along the coal seam. The function of the scan line in this study is to minimize duplication of cleat data and validate between face and butt cleats data collected from the field before further analysis (Watkins et al., 2015). The scan line implementation in this study is by selecting the observation location as long as 5 m, the scan line length is determined from the average thickness of the coal layer to be observed. Scan line length is the same at all observation locations to reduce measurement bias (Apriyani et al.,
2014; Sapiie et al, 2014; Sapiie and Rifiyanto, 2017). Scan line direction is also important because it is related to the appearance of faces and butt cleats. So that the scan line will be ideal if the direction of the azimuth is diagonal or ± 45° from coal seam strike.

The object of observation is the measurement of cleat attributes, which includes the aperture and spacing by using a geological compass, meter tape, ruler, and calipers (Table 1). The results of these observation objects become variables in the analysis or processing of data using geostatistical methods. Thus, geostatistics is the object of this study.

The advantage of geostatistical science is its ability to characterize the application of spatial structures with consistent probabilistic models. Geostatistical application in this study is to see the relation between variables (cleat attributes) to evaluate data quantitatively until the normal distribution is obtained. The data filter was done as a treatment until valid data is reached through the boxplot model.

Boxplot is a graphic technique developed by Tukey and is often used for exploratory data analysis (Buttarazi et al, 2018). Boxplot is relatively easier to made manually, but in this study, the help from StatPlus software: mac LE from AnalystSoft, Inc. is used. Boxplots are also known as boxes and whisker diagrams, which visually show the data center, distribution (Marmolejo-Ramos and Tian, 2011). The basic elements of the graphic’s form (Fig.2), are: Box contains 50% data, the top edge of the box is called Q3 (75% data) and the bottom edge of the box is called Q1 (25% data), Transverse line in the box is the median (Q2). If Q2 is not in the middle box, it means that data distribution is not symmetrical (skew), Last point of the vertical line is the maximum and minimum value, unless there is an outlier in the data. Vertical line length is 1.5 interquartile range (IQR = Q3-Q1). Box’s space shows the size of the datadiversity. and Point outside the line is called outlier.

**Figure 2.** Actual observations are represented by a rug on the left axis in each plot: 1) smallest observations, 2) lower limit of whiskers, 3) Q1, 4) median, 5) Q3, 6) upper limit of whiskers, and 7) mid and extreme outliers (any if).

The face cleat in the coal which is aligned in the direction of the alignment can be compared to the occurrence of tension and open tension pulling. Face cleats are relatively smooth, planar, continuous and elongated, present at regular intervals, while the surface cleats are rough and end on a face cleat (Sapiie and Rifiyanto, 2017). Su et al. (2001), concluded that face cleats can be compared with planar systematic rigidity, while butt cleats with cross joints. Face cleats are aligned in the direction of regional stress, while the cleat butt is parallel to the face cleat (Fig. 3).

**Figure 3.** Schematic model of cleats pattern on folds.
The observation of cleat attributes in the eight measurement locations was revealed very well. Face cleat orientation on each scanline path shows a pattern parallel to the main stress or perpendicular to the fold axis of the Tamiang anticline. On the basis of straight, parallel, regular distance cleats and having a parallel alignment orientation, it can be concluded that face cleats throughout the scanline have developed as systematic cleats.

### Table 1. Characteristics of measurement cleats at PMC and Baturona.

| Cleat   | Face Cleat | Butt Cleat |
|---------|------------|------------|
|         | Orientasi (N.°/E) | Spacing (cm) | Aperture (mm) | Frekuensi | Orientasi (N.°/E) | Spacing (cm) | Aperture (mm) | Frekuensi |
|         | Data | Mean | Data | Mean | Data | Mean | Data | Mean | Data | Mean | Data | Mean | Data | Mean | Data | Mean | Data | Mean |
| PMC     | Barat NE-SW | 229 | 0.5-20 | 2.7 | 1-21 | 1 | 156 | 284-355 | 325 | 1-9 | 3 | 1-7 | 2 | 19 |
|          | Timur NE-SW | 227 | 1-20 | 3 | 1-17 | 2 | 104 | 306-336 | 324 | 1-11 | 4 | 1-39 | 3 | 39 |
| Seam 1  | Barat 22-90 | 61 | 1-18 | 5 | 1-30 | 3 | 129 | 82-167 | 147 | 0.5-12 | 4 | 1-8 | 2 | 59 |
|          | Timur 201-248 | 229 | 1.8-10 | 5.5 | 2-24 | 7 | 31 | 106-155 | 124 | 1-9 | 3 | 1-7 | 2 | 21 |
| Baturona | Barat 222-258 | 238 | 0.9-20 | 4.9 | 0.1-40 | 0.2 | 78 | 265-358 | 330 | 0.19-58 | 7.5 | 0.1-9 | 0.3 | 41 |
|          | Timur 210-285 | 242 | 0.3-85 | 5 | 0.1-4 | 2 | 54 | 104-178 | 136 | 0.4-43 | 10 | 0.1-5 | 2 | 35 |
| Seam 1  | Barat 44-86 | 67 | 0.5-45 | 3.25 | 0.1-2 | 0.1 | 74 | 100-172 | 142 | 0.1-11 | 63 | 10-0.12 | 2 | 47 |
|          | Timur 40-80 | 62 | 0.25-49 | 6 | 0.1-60 | 0.4 | 41 | 92-165 | 135 | 1-53 | 12.5 | 1-3 | 1 | 23 |

3. Result and Discussion

The initial step in validating data is to study the characteristics of the data. For this reason, this study had to know the concentration and dissemination of data from mean and median, extreme values and outliers, and several other measurements. There are several techniques for describing the characteristics and distribution of these data. However, this study will discuss boxplots. The boxplot visualization applied in this study is the multiple boxplot, which is realized with the aim of comparing the face and butt cleats in each seam of the two Baturona and PMC locations.

Tukey boxplots were applied to visualize cleat spacing in Baturona and PMC that almost entirely had data outliers in each seam (Fig. 4). In fact, it is commonly seen some extreme outlier data and cleat spacing tends to appear on face cleat patterns, such as Seam 1B and 1T in Baturona, then at PMC on Seam 1T and 2B. Outliers are data that deviate from other data sets (Mate and Arroyo, 2015). Whereas according to Li et al. (2016), extreme outliers are the data resulting from observations that do not follow most of the patterns and are located far from the data center that may have a large influence on the regression coefficient. Cleat measurements in this study do not rule out the possibility of obtaining a cleat value that is different from other cleat values. This may be caused by errors during the measurement or the presence of different types of cleats, namely induced cleats due to mining activities such as heavy vibration or the excavator bucket.

Based on Fig. 4, although the results of the cleat spacing showed outlier data, there was an equal distribution of data that appeared on the Seam 2T face cleat in Baturona and the Seam 2B and 2T butt cleat data at the PMC. The data in these locations do not indicate an outlier; this indicates that the field conditions at these locations have a more regular cleat pattern (Fig. 5).

**Figure 4.** Spacing cleat has not validated yet.

The aperture cleat measurement result that embodied in the boxplot shows that the data distribution still has a lot of deviation marked by the emergence of data outliers and even extreme outliers (Fig. 6). Extreme outliers from the aperture data are very abundant in Baturona. The
uniformity of data in Baturona is reflected in box with an average aperture of 1-9 mm, but the aperture is 39 mm in the seam 1T cleat. No less interesting aperture cleat data at PMC is on the seam 1B face cleat, because the data density is concentrated up to 5mm, but there is a high outliers gap at 15mm and 40mm. According to Rousseeuw et al. (1999) and Qarmalah et al. (2016), in space statistics, data outliers must be viewed against the position and distribution of other data so that they will be evaluated for the causes of the data outliers. Thus, due to extreme outliers with large gaps on the Seam 1B and 2T PMC face cleat apertures, evaluation of the cause needs to be through the cleat genetic factor approach.

Figure 5. Regular appearance of butt cleat spacing on Seam 2T PMC, so that the reflection of measurement data is relatively uniform without outliers.

Figure 6. Aperture cleat has not validated yet.

As noted in the previous discussion of cleat spacing, obtaining representative cleat data on the surface is not easy. This also applies to the aperture data, considering that the aperture cleat on the surface is hard to believe because it has experienced various symptoms due to geological processes and the effects of the mining process. The existence of data outliers in the spacing and aperture of Baturona and PMC is a problem in data validity, so a treatment that can identify outliers problems is required, one of which is by setting aside the outliers from the data group then analyzing the data without outliers (Mate and Arroyo, 2015).

Geostatistical analysis in this study is not merely looking for general patterns of cleats but sorting cleat data that is far beyond the general pattern. It turns out that the results of the identification through the box plot model showed that there were outliers present in each data group, in this case, on the face and butt cleats of each seam in Baturona and PMC. So the treatment removes the mid data and extreme outliers. Firmly stated that descriptive statistics will use cleat data within the range of IQR. Then, modeling the boxplot is done again to achieve valid cleat data.

Figure 7. A (Left): Validated spacing cleat. (B) (Right) Validated aperture cleat.

The results of the outliers analysis are shown in Fig. 7 (A) and Fig. 7 (B), it appears that all the extreme outliers cleats has disappeared. Even though on the spacing and aperture cleat there is still...
mild outliers in three data groups, but mild outliers are not an error data because there is the possibility of storing certain information. So, if a data treatment is about to be done again, it is an act that is very wrong, the data may be part of superior seeds. This is relevant to Andrea et al. (2014) who previously has already done detection of outliers with boxplots in Monte Carlo. If it is associated with generic data, then the probability that occurs is in mid outliers thecleats has the same genes but there are pre depositional fractures (Fig. 9).

Another benefit of the application of geostatistics in cleat data is the availability of information from the boxplot model regarding the distribution of data that has been validated. Interpretation, predominantly validated cleats have an asymmetrical distribution (skew). This can be observed from the position of the box against the whisker line that is not the same length, or can also look at the median line that is not in the middle of the box.

Next, the difference in the diversity of data is reflected in the height of the box (Hubert and Vandervieren, 2007). Validated cleat diversity was found in cleat spacing at PMC with IQR 2.35-13 cm and aperture in Baturona with IQR 1-2 mm. Very limited data distribution is found at the Seam 2B aperture face cleat, which is 0.1 mm, this can be seen from the box shape that is only in line.

4. Conclusion
As a matter of fact, to measure cleat attributes and obtain representative data on the surface is not easy. Considering the cleat data on the surface is hard to believe, especially at the mine site, which does not necessarily reflect the influence of the geological process. Generally in the study area, cleats have been affected by heavy equipment collisions during excavation, opening of overburden during stripping, or due to weathering in the exposed coal seams. As a result of all this, the results of the cleat attribute measurement need to be statistically validated to obtain the pattern and uniformity of the data.

References
[1] Andrea, K., G.L. Shevlyakov, and P.O. Smirnov, 2013, conference paper, download from Research Gate.
[2] Apriyani, N., Suharmono, M. Momen, S. Djaelani, A. Sodli, A. Satria, and A.S. Murtani, 2014, AAPG Int. Conf.
[3] Buttarazzi, D., Pandolfo, G., and G.C. Porzio, 2018, Int. Biometric Soc., 1-10.
[4] Gafoer, S., G. Burhan, and J. Purnomo, 1995, Geological map Lahat sheet, Geological Survey Center.
[5] Hubert, M. and E. Vandervieren, 2007, Comp. Stat. data analysis, preprint submitted to Elsevier.
[6] Kuncoro, B., S. Notosiswoyo, K. Anggayana, C.I., and Abdulla, 2007, Journal of JTM 14, 145-165.
[7] Li, A. M. Feng, Y. Li, and Z. Liu, 2016, Procedia Comp. Sci. 91, 245-251.
[8] Laubach, S.E., R.A. Marret, J.E. Olson, and A.R. Scott, 1998, Int. J. Coal Geol., 35, 175-207.
[9] Marmolejo-Ramos, F., and T.S. Tian, 2010, Int. J. Psychology Research, 3, 37-45.
[10] Mate, C. and J. Arroyo, 2015, ResearchGate articles, 1551-1556.
[11] Qormalah, N.M, J. Einbeck, and F.P.A. Coolen, 2016, Stat Papers, 1-16.
[12] Rousseeuw, P.J., I. Ruts, and J.W. Tukey, 1999, The American Stat., 53, 382-387.
[13] Sapiie, B., A. Rifiyanto, and L.A. Perdana, 2014, AAPG Int. Conf.
[14] Sapiie, B., A. Rifiyanto, 2017, J. Eng. Technol. Sci. 49, 322-339.
[15] Shell Mijnbouw, N.V., 1976, Geological study of the Bukit Asamcoal mines, 18 pp (unpublished)
[16] Su, X., Y. Feng, J. Chen, and J. Pan, 2001, Int. J. Coal Geol., 47, 51-62.
[17] Sun, Y. and M.G. Genton, 2011, J. Computational and Graphical Statistics, 20, 316-334.
[18] Watkins, H., C.E. Bond, D. Healy, and R.W.H. Butler, 2015, J. Structural Geology, 72, 67-82.