Computers & Geosciences 93 (2016) 96–102

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

Data Envelopment Analysis as a tool for the exploration phase of mining

Tommi Kauppinen

Aalto University School of Electrical Engineering, Department of Electrical Engineering and Automation, PO Box 15500, 00076 Aalto, Finland

ARTICLE INFO

Article history:
Received 10 February 2016
Received in revised form
11 May 2016
Accepted 12 May 2016
Available online 14 May 2016

Keywords:
Data Envelopment Analysis
Raman analysis
Exploration phase of mining
Principal Component Analysis

ABSTRACT

The exploration of mining has often been limited by time-consuming methods of analysis. This paper introduces Data Envelopment Analysis (DEA) as a new tool for the exploration phase of mining. DEA is a non-parametric method for data fusion, and it is used alongside with the on-site Raman analysis. Ten meters of halved rock drillcore from the Kittil mine (Suurikuusikko deposit) were pulverised and homogenised, thus ensuring that each meter had a representative sample. These 10 samples, one for each meter, were subsequently measured with a grid measurement (32×32 measurement each) using the Raman setup. All the data points were analysed using the point-count method. After identifying the frequency at which potentially valuable minerals appear in the samples, this information was analysed using DEA. The study ends by presenting an efficiency score for each meter of drillcore. These efficiency scores enable geologists to judge more rapidly which parts of the drillcore must be logged more carefully. In addition, Principal Component Analysis (PCA) is discussed as an alternative for producing similar results to DEA.

1. Introduction

In recent years, expediting the exploration phase of mining has sparked considerable interest in the mining industry due to the fact that it is highly time-consuming process. One needs to cut through bedrock with a diamond drill, analyse the drillcores obtained, and then decide the feasibility of establishing a mine, as well as the exact mining location. The exploration phase depletes large amounts of both natural and economic resources in the phase of extraction of drillcores and their analysis. However, this is the sole way to acquire evidence of ore as well as estimate the amount available in the rock (Computers & Geosciences, 2011). Moreover, after the mine has been started, the mine planning includes drilling new holes to ascertain where the deposit is located.

Possessing accurate information on the contents of the drillcores is vital for the mining industry (Ma et al., 2010). The ore excavated is completely dependent on the analysis carried out on the drillcores. Time is also of the essence, even though mining occurs over a period of years. By improving the speed of the analysis of drillcore samples, geologists can more quickly provide accurate information to decision-makers.

This paper focuses on affordable but precise methods of analysing drillcores. These methods might speed up the process for the geologist from the phase of acquiring the drillcores from the ground to the next phase of passing on their findings. The phases involved are: drilling, analysing and presenting the findings. This paper concentrates on the analysis part of the exploration. A new method is proposed to help geologists analyse the drillcore faster but still in a reliable manner.

Analysing the drillcore samples is a time-consuming process requiring one to two months to acquire the laboratory results. The methods to circumvent the laboratory analysis using on-site analysis are further described elsewhere (see Kauppinen et al. (2013, 2014a)) but here it might be noted that by using on-site analysis, the time required to produce reliable results is measured in hours, not months. One of the candidates for on-site analysis is the Raman analysis (see e.g. Ishikawa and Gulick (2013)), which is used in this paper to produce data from rock drillcores. In addition, geologists need to analyse the rock when it is extracted from the ground. This analysis is called logging. Logging is hard to quickly and reliably execute, hence the methodology presented in this paper. By the method introduced, geologists can produce results for the decision-makers from a high number of drillholes in a matter of days.

The object of research in this paper is to show that Data Envelopment Analysis (DEA) can be used to help geologists log the drillcores more efficiently (for mineral prospecitivity mapping with DEA, see Hosseini and Abedi (2015)). The case studied is Kittilä mine in Finland, but the general application for drillcores is also discussed. DEA is a non-parametric efficiency measure. It is used here to fusion the spectral data received from Raman

E-mail address: tommi.kauppinen@aalto.fi

http://dx.doi.org/10.1016/j.cageo.2016.05.005
0098-3004/

experiments into a single efficiency score available for the use of the geologists. Furthermore, the mathematics of DEA is discussed and its linkages to other data fusioning methods (see e.g. Abedi et al. (2012) and Chauhan et al. (2016)) are considered. This paper also shows that DEA can be considered to be a reliable tool even though it is by essence non-parametric.

2. Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric method for evaluating Decision Making Units (DMUs). Typically, DEA is used to evaluate branches of a companies or medical facilities. The first contributions for DEA were made by Farrell (1957), who established the foundations for DEA by introducing the comparative efficiency of companies using a linear programming (LP) problem for input–output data set.

Ever since Farrell's discovery, the DEA method has grown in popularity. Charnes et al. (1978) published a paper in the field of operations research. In this and the following paper, Charnes et al. (1979), presented the Charnes–Cooper–Rhodes (CCR) model, which was and still is the first basic model for DEA. The CCR model evaluates the efficiency of the weighted sums of the input–output ratio. In addition, the efficiency scores of DMUs are given between [0,1] in CCR modelling, which is even today the basic premise for many DEA models. After the CCR model, Banker et al. (1984) introduced the Banker–Charnes–Cooper (BCC) model. In BCC modelling, variable returns to scale of DMUs define the efficiency frontier, while the CCR model assumes constant returns to scale.

The DEA method has grown in popularity in academic publications worldwide. Therefore, the scientific community has possibly already accepted or is in the process of accepting the DEA method as a valid tool for decision making. At present, this method includes many different models for different usages. For example, Cooper et al. (2006) describe 20–50 different models, while the exact number is related to the definition of a model. Of course, for the user of the DEA method, this poses a new problem of finding the appropriate model from a selection of models.

In this paper, output-oriented BCC modelling is used to ascertain the efficiency scores for ten (10) consecutive samples of rock drillcore. BCC modelling assumes varying returns to scale, meaning that even smaller quantities of ore can show a high efficiency score. Considering drillcore analysis, this is highly acceptable: the cases where there are multiple minerals in small quantities might also be a valuable area for geologists to study in more detail.

2.1. Introduction to mathematics of data envelopment analysis

Data Envelopment Analysis is a linear programming (LP) based technique for measuring the relative performance of data points (see e.g. Kauppinen and Khajehzadeh (2015)). The usual measure of efficiency, i.e.:

$$\text{efficiency} = \frac{\text{output}}{\text{input}} = \frac{u_1 y_{10} + \cdots + u_s y_{s0}}{v_1 x_{10} + \cdots + v_m x_{m0}} \tag{1}$$

is often inadequate due to existence on multiple inputs and outputs related to the data in question. In Eq. (1), u and v are weights for output and input variables y and x, respectively.

From Eq. (1), a fractional linear optimisation problem can be formulated, by giving limits to the weights u and v and for the efficiency as a whole. Then one has (for decision-making units (DMUs) $o = 1, \ldots, n$)

$$\begin{aligned}
&\max_{v,u} \frac{u_1 y_{1o} + \cdots + u_s y_{so}}{v_1 x_{1o} + \cdots + v_m x_{mo}} \\
&\text{s. t.} \\
&\frac{u_1 y_{1j} + \cdots + u_s y_{sj}}{v_1 x_{1j} + \cdots + v_m x_{mj}} \le 1, j = 1, \ldots, n \\
&u_1, \ldots, u_s \ge 0 \\
&v_1, \ldots, v_m \ge 0
\end{aligned} \tag{2}$$

Eq. (2) is a fractional problem, which can be modified into Linear Programming (LP) form by defining $v_1 x_{1o} + \cdots + v_m x_{mo} = 1$. The result in Eq. (3) is treated as a problem akin to a standard Linear Programming (LP) problem, also known as the input oriented CCR DEA model for DMUs $o = 1, \ldots, n$.

$$\begin{aligned}
&\max_{v,u} u_1 y_{1o} + \cdots + u_s y_{so} \\
&\text{s. t.} \\
&v_1 x_{1o} + \cdots + v_m x_{mo} = 1 \\
&(u_1 y_{1j} + \cdots + u_s y_{sj}) \le (v_1 x_{1j} + \cdots + v_m x_{mj}), j = 1, \ldots, n \\
&u_1, \ldots, u_s \ge 0 \\
&v_1, \ldots, v_m \ge 0
\end{aligned} \tag{3}$$

After developing the CCR model, yet another model can be introduced, known as the BCC model, which is used in this study for the actual computations. The multiplier form of input oriented BCC model can be developed from the Eq. (3) by adding a variable u_o to the cost function. u_o is free in sign and it is substracted from the cost function otherwise equal to the CCR model (see Eq. (4)).

Fig. 1 shows graphically how CCR and BCC efficiency frontiers differ form each other for a 200 data points. The data is random data, including one input and one output variable.

Graphically u_o marks the difference between the efficiency frontiers of CCR and BCC models (see Fig. 1), and therefore u_o varies for different DMUs o. The points on the efficiency frontier always have an efficiency of one. Furthermore, the rest of the points score a better efficiency scores as the distance from the efficiency frontier is somewhat reduced. Therefore, the BCC model gives higher efficiency scores for the data set under study

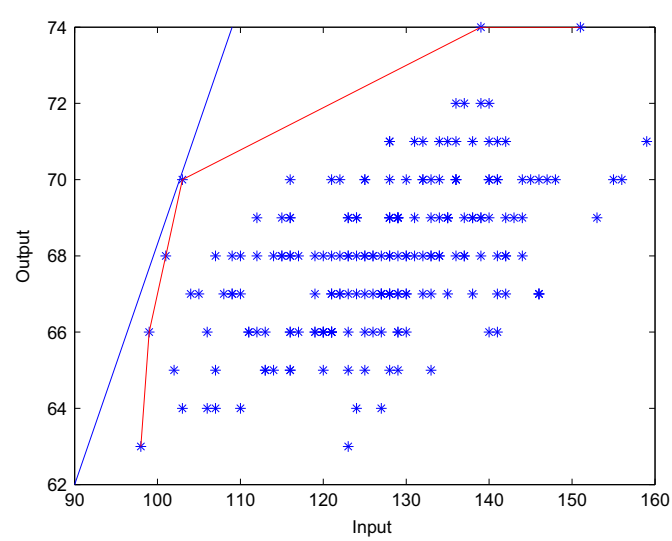

Fig. 1. Efficiency frontiers for both DEA CCR model (blue) and DEA BCC model (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)