International Journal of Coal Geology 152 (2015) 80–93

Contents lists available at ScienceDirect

International Journal of Coal Geology

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

Mapping of coal quality using stochastic simulation and isometric logratio transformation with an application to a Texas lignite

Ricardo A. Olea [a,*], James A. Luppens [b]

[a] U.S. Geological Survey, 12201 Sunrise Valley Drive, Mail Stop 956, Reston, VA 20192, USA
[b] U.S. Geological Survey, W 6th Ave. & Kipling St., Mail Stop 939, Lakewood, CO 80225, USA

ARTICLE INFO

Article history:
Received 31 July 2015
Received in revised form 1 October 2015
Accepted 1 October 2015
Available online 9 October 2015

Keywords:
Two-point geostatistics
Compositional data analysis
Geochemistry
Proximate analysis
Uncertainty

ABSTRACT

Coal is a chemically complex commodity that often contains most of the natural elements in the periodic table. Coal constituents are conventionally grouped into four components (proximate analysis): fixed carbon, ash, inherent moisture, and volatile matter. These four parts, customarily measured as weight losses and expressed as percentages, share all properties and statistical challenges of compositional data. Consequently, adequate modeling should be done in terms of a logratio transformation, a requirement that is commonly overlooked by modelers. The transformation of choice is the isometric logratio transformation because of its geometrical and statistical advantages. The modeling is done through a series of realizations prepared by applying sequential simulation for the purpose of displaying the parts in maps incorporating uncertainty. The approach makes realistic assumptions and the results honor the data and basic considerations, such as percentages between 0 and 100, all four parts adding to 100% at any location in the study area, and a style of spatial fluctuation in the realizations equal to that of the data. The realizations are used to prepare different results, including probability distributions across a deposit, E-type maps displaying average properties, and probability maps summarizing joint fluctuations of several parts. Application of these maps to a lignite bed clearly delineates the deposit boundary, reveals a channel cutting across, and shows that the most favorable coal quality is to the north and deteriorates toward the southeast.

Published by Elsevier B.V.

1. Introduction

Coal is one of the most heterogeneous substances naturally found on earth. Coal may contain up to 76 of the 98 naturally occurring chemical elements, most of them in the form of traces (Schweinfurth, 2009). Carbon, which by definition is required to be a dominant element, is the only one present as a native element. Ranking of coal according to purity varies by country, but a common designation, starting with the most impure form, is a classification into peat, lignite, subbituminous coal, bituminous coal, anthracite and graphite (Kwiecinska and Petersen, 2004; Schweinfurth, 2009).

Methods for coal chemical analysis also abound. Proximate analysis is the least detailed but at the same time one comprising all possible constituents. The four components reported as weight percentages in a proximate analysis are moisture (water), volatile matter, ash and fixed carbon. Ideally, the moisture of interest is the in situ moisture also known as inherent moisture, a component hard to measure exactly primarily because of the disturbance produced during the drilling used to cut cores. Moisture is ordinarily determined by drying pulverized coal at about 100 °F (38 °C). "As received moisture" is the percentage weight loss relative to the original weight. Inherent moisture is obtained by applying empirical corrections to the "as received moisture". Volatile matter refers to the additional components liberated after rising the temperature to about 900 °F (482 °C) in the absence of air. It comprises primarily hydrocarbons, sulfur and carbon dioxide. Ash is the solid residue left after complete combustion of the coal. Finally, fixed coal is not directly measured; it is the difference to 100% of the sum of the other three components (American Association for Testing of Material (ASTM), 2013). In a following paper intended to be published in this same journal, we will map the results of ultimate analysis and calorific value to complete the spatial analysis of coal quality.

Proximate analysis provides the contribution to a total of the four partial components, hence, the four parts fall in the category of compositional data (e.g. Bacon-Shone, 2011). In variable space, drill cores have 4 components with all of the parts having a numerical value mathematically defining a vector. If, like in our case, the interest is in the geographical variation of such components, each one of them is a regionalized variable, which are best modeled using geostatistics (e.g. Caers, 2011; Chilès and Delfiner, 2012). Both compositional data analysis and geostatistics have been around for several decades, but their combined application has been sporadic since first attempted by Pawlowsky (1984). Among the multiple methods offered by geostatistics, the vast majority of the applications to compositional data have used cokriging

* Corresponding author.
E-mail address: rolea@usgs.gov (R.A. Olea).

http://dx.doi.org/10.1016/j.coal.2015.10.003
0166-5162/Published by Elsevier B.V.

and have not been related to proximate analysis: Pawlowsky-Glahn and Olea (2004) and Tolosana-Delgado et al. (2011) dealt with petroleum geology; Tolosana-Delgado and van den Boogaart (2013) and Niang et al. (2014) with soils; Lark et al. (2012) and Park and Jang (2014) with sedimentology; Boezio et al. (2012); Tolosana-Delgado and van den Boogaart (2014) with mineral exploration; and, Pawlowsky-Glahn et al. (2015a) with hydrochemistry.

For our full system with 4 components, the main objectives here are: (a) modeling spatial fluctuations in coal proximate analysis components across a deposit using the newer formulations of isometric logratio transformation in compositional data analysis and stochastic simulation in geostatistics; (b) using the results to map trends and uncertainty separately for each component, and (c) combining results of several coal components in the form of probability maps to delineate favorable and unfavorable areas across the coal deposit.

2. Methodology

Cutting edge modeling of spatial fluctuations of compositional data requires borrowing from two bodies of knowledge: geostatistics and compositional data analysis. We follow the standard practice of preparing two-dimensional pixel maps, which requires interpolation of values at regular arrangement of locations, called grids when referring to the centers (nodes) or tessellations when referring to the Voronoi polygons (cells) defined by the nodes (Aurenhammer, 1991).

If the sample comprises N specimens taken at locations $\mathbf{u}_i$ and each specimen comprises the same D parts, the data are designated as $\mathbf{z}_D(\mathbf{u}_i) = [z_1(\mathbf{u}_i)\ \ z_2(\mathbf{u}_i)\ \ \dots\ \ z_D(\mathbf{u}_i)]$, $i = 1, 2, \dots, N$.

2.1. Geostatistics

The body of knowledge today associated with geostatistics is an extension of formulations originally made to advance the estimation of ore reserves (Krige, 1951; Matheron, 1963). Geostatistics is presently a widely accepted toolbox for the modeling of any attribute having spatial fluctuations. The early methods were different forms of a generalization of least square regression, in which the main objective was the minimization of errors in the inference of values at locations, $\mathbf{u}_o$, not considered in the sampling (e.g., Srivastava, 2013). The collective name for those methods is "kriging", which produces an estimate, $z^*(\mathbf{u}_o)$, as a linear combination of measurements—certainly not compositional—$z(\mathbf{u}_i)$, ordinarily the M closest ones:

$$z^*(\mathbf{u}_o) = \sum_{i=1}^{M} \lambda_i \cdot z(\mathbf{u}_i), \tag{1}$$

where the λ_i's are weights resulting from solving a system of equations minimizing the estimation error, which can take different forms depending on assumptions about the mean. The value for the minimum error variance is also called kriging variance. These weights are a function of the nature of the trend, distances to the estimation location, the distances among measurements, and the style of spatial fluctuation of the attribute. Cokriging is a generalization for the joint estimation of two or more spatially correlated attributes. For examples and the mathematical formulation of kriging or cokriging, the reader may consult Olea (2009). Kriging still remains the best method for minimizing estimation errors, but has been progressively abandoned in favor of stochastic simulation because kriging values are conditionally biased. Although on average a collection of kriged values above and below the average has a mean error of zero, away from the mean, the estimated values have a systematic deviation; when the true values are small, kriging gives estimates that are systematically high and conversely for high true values. This bias has consequences easy to detect and observe. For example, maps prepared using kriging for the interpolation of values away from sampling locations have a smoothing that is more pronounced for sparse samples. This filtering property of kriging is a problem, for instance, in applications related to fluid flow because the smoothing distorts the geometry and extent of flow barriers and flow paths. Another problem is the alteration of the statistics: sample histograms and sample semivariograms are significantly different from those of the estimated values (e.g., Olea, 2009). Finally, kriging does not directly model the distribution for the errors at any given location, distribution that is necessary to model uncertainty. Kriging is restricted to providing two moments only: the mean and the variance of the estimation. The form of the distribution is missing. When interested in modeling uncertainty, kriging requires the analyst to assume a distribution form that is completely defined with just two moments: the mean and the variance. The alternative of common choice is the normal distribution, which may or may not be a close approximation depending on the characteristics of the attribute(s) being modeled. Typically, for example, when the data are lognormally distributed, it is not reasonable to expect that the errors will be normally distributed (Goovaerts, 1997).

These weaknesses prompted the development of alternative methods, of which stochastic simulation has been the most successful approach going back more than 40 years (Journel, 1974). The top objective of stochastic simulation is the reproduction of the global characteristics of the fluctuations, such as the semivariogram and the histogram. The approach can be regarded as an automation of the experience of passing around to several people posted values on a map and requesting them to prepare contour maps manually. Typically, the outcome is to have as many different maps as authors because of individual ideas regarding the vast possibilities about what is going on in between the control points. Stochastic simulation generates any number of different maps called "realizations", all honoring the style of fluctuation suggested by the data. Traditionally, such a style has been summarized in the form of a semivariogram, a two-point statistics. More recently the latest trend is to use "training images" instead, increasing the discrepancies between simulation and kriging (Mariethoz and Caers, 2014). The realizations are said to be "unconditional" when all that is required is to honor the style of fluctuation and "conditional" when additionally it is required to honor values and the histogram of some control points. The down side of stochastic simulation is that, on average, for any realization, the estimation errors are greater than those of kriging. Nevertheless, the E-type map, which is the average of all realizations, is similar to a kriged map, which takes care of the deficiency. Last but not least, stochastic simulation elegantly solves the problem of modeling uncertainty at any node. In general, at each node, each realization contributes one possible value for the attribute at that location. The set of all those values, as many as realizations generated, numerically model the distribution for all values the true value may take.

From the great variety of simulation methods presently available (Caers, 2011), we have selected sequential simulation for not having restrictions in sample size, as well as its ease in honoring data, quick execution using minimal CPU capacity, and being relatively free of the danger in other methods of generating artifacts. A basic idea in geostatistics is that spatial uncertainty can be modeled using random functions of a dimension as large as the number of cells to consider in the model. Sequential simulation takes advantage of a recursive application of the Bayes's theorem so that these complex random functions can be modeled as the product of univariate distributions, one per location (Goovaerts, 1997; Pyrcz and Deutsch, 2014). The term sequential refers to the fact that the values at the grid nodes making a realization are generated one after the other. What is not implied in the name is that the visitation path is different for each realization, which is a fundamental detail needed to produce multiple unique realizations. An important feature in the practical implementation is the fact that, as the generation of a realization progresses, partial results are added to the initial measurements and used as data in further calculations. Another common feature is that kriging is the inference engine for all forms of sequential simulation. Hence, while straight use of kriging is being abandoned, it is alive in sequential simulation in the form of a kriging with feedback.