Computers & Geosciences 103 (2017) 51–58

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

Optimal ordering of realizations for visualization and presentation

George de Barros*, Clayton V. Deutsch

Centre for Computational Geostatistics (CCG), School of Mining and Petroleum Engineering, Department of Civil and Environmental Engineering, University of Alberta, 6-247 Donadeo Innovation Centre for Engineering, 9211-116 Street, Edmonton, Alberta, Canada T6G 1H9

ARTICLE INFO

Keywords:
Uncertainty
Geostatistical realization
Simulated annealing
Image morphing
Animation

ABSTRACT

In geostatistical simulation, a realization represents one possible outcome of the spatial uncertainty model. Tens to hundreds of realizations are generated in order to understand response property variation. There are ways to summarize local uncertainty, but visualizing all realizations is important to understand joint uncertainty between multiple locations. There is no straightforward manner to visualize all realizations at the same time or in sequence. This paper presents a new method to sequentially display multiple geostatistical realizations. The proposed algorithm performs an ordering of the visible portion of the realizations (images), according to the distance between realizations. The concept of distance corresponds to the differences computed cell by cell for every realization pair or to the differences computed from a moving window filtering applied to each realization. To define an optimal sequence of realizations, the shortest path route through the realizations is established by a simulated annealing technique. The gradual transition between realizations is enhanced by an image morphing technique where intermediate images are introduced between the original images. The final result consists of an animation that shows the sequence of realizations and allows better understanding of the uncertainty model.

1. Introduction

Geostatistical realizations must be treated as an ensemble. Although they may be ranked by a response variable such as in-place resources, two adjacent realizations may appear completely different due to areas of high and low values occurring in different locations. This creates a challenge for visualizing the uncertainty in an ensemble of 3D model realizations, which is an integral part of geomodeling applications (Viard et al., 2011). Note that visualizing one realization or a single estimated model does not convey uncertainty.

In this context, advances in computational speed and storage have made it possible to study the development of complex and dynamic systems and represent results accordingly. There is an increase in productivity due to newly developed automated strategies and shared computation. Although the number of realizations of geostatistical simulation is increasing and has been automated, we believe human inspection is still required for quality control and to analyze the results. Tools to help process and analyze an ensemble of realizations in a qualitative and productive manner are required.

The challenge is to handle the information generated by an uncertainty analysis from a geostatistical approach. Geostatistical simulation techniques are relatively well established. The application of geostatistical techniques with geospatial data and the computation of uncertainty in the resulting model are reasonably well understood. In terms of visualization, there are many programs to display and manipulate complex 3D models and their internal properties. Representation of high dimensional uncertainty, however, is still a challenge because there are a restricted number of available visual channels to represent it, such as data position, color, texture and opacity.

According to Potter et al. (2012), this quantification must be simplified in order to display many realizations in an appropriate and effective visual manner that allows for human perception. Lamigueiro (2014) presents the uncertainty by showing realizations side by side (simultaneous) and with superimposed results (condensed). Obermaier and Joy (2014) focus on visualization methods for understanding the similarities, differences and trends among the members of the realization group. Viard et al. (2011) present three methods for uncertainty representation. In the visually separable methods, the uncertainty space is subdivided into a set of smaller ranges, each of which is assigned a different pattern, resulting in a categorization of uncertainty levels (e.g. Cedilnik and Rheingans, 2000; Interrante, 2000; Rhodes et al., 2003; Djurcilov et al., 2001). The visually integral methods fill clarity and blurring objects according to their relevance or uncertainty degree (e.g. MacEachren, 1992; Kosara et al., 2001; Robinson, 2006). The animation methods use time to transmit a sense of uncertainty, displaying maps sequentially ordered, according to a relevant objective function and using some interpolation

* Corresponding author.
E-mail addresses: debarros@ualberta.ca (G. de Barros), cdeutsch@ualberta.ca (C.V. Deutsch).

http://dx.doi.org/10.1016/j.cageo.2017.03.001
Received 8 July 2016; Received in revised form 27 February 2017; Accepted 1 March 2017
Available online 02 March 2017

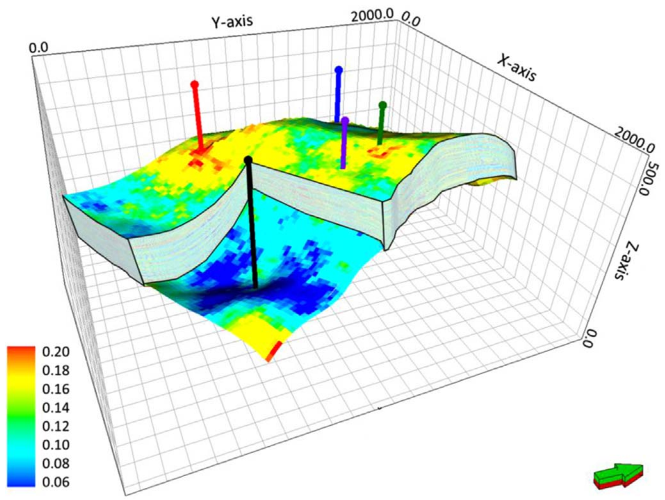

Fig. 1. An arbitrary set of view planes from a 3D geological model. The color in this case represents the porosity of a reservoir layer.

algorithm between the frames (e.g. Fisher, 1993; Srivastava, 1994; Ehlschlaeger, 1997; Davis and Keller, 1997; Dooley and Lavin, 2007).

Phadke et al. (2012) also mention the animation technique to support a set of realizations. The pairwise sequential animation method orders n members of an ensemble (a collection of related datasets), combines subsets of realizations, and presents them as an animated visualization using hue and texture. Höllt et al. (2014) introduce the ensemble visualization approach for spatially distributed data, defined as a collection of n values of a single variable in m dimensions. In our case, this approach is used for an ensemble of realizations generated for a geological model. The ensemble visualization approach appears in several works (e.g. Kao et al., 2001, 2005, Luo et al., 2003, Love et al., 2005, Obermaier and Joy, 2014). This approach can be considered a specific method from the general uncertainty visualization methods, applied when uncertainties are not well represented in quantitative terms.

This paper presents a new method to sequentially display multiple geostatistical realizations, combining both animation methods and the ensemble visualization approach. The emphasis is on the visualization aspect of post processing and not on the details of model construction. There are many references available with model construction details. The assumption here is that an ensemble of realizations is available in a common numerical format. There are likely tens to hundreds of realizations. The realizations likely contain millions of locations; however, not all locations will be visible on a particular display. Some locations will be transparent to see deeper into the model and some locations will be blocked behind locations that are being displayed.

A sequence of realizations is defined by their similarities/dissimilarities. Numerical or quantitative differences between two realizations define a distance for each pair of images. Realizations that are close together are displayed one after another. The results are presented as a dynamic animation, which shows the defined sequence of realizations.

Although some results on a 2D grid are presented, the main idea of this work is to apply the proposed approach to any visualization of 3D geological models (Fig. 1). A user can choose an arbitrary view plane(s) and visualize all realizations available for this plane. Fig. 1 shows an arbitrary view with some horizontal views visible and some cross sections visible. The order of realizations would not change with zooming or panning, but would change when the position of the slices is changed. The visualization ordering would be automatically recalculated when the view plane is changed. Software would cycle the realizations in the optimized sequence at a specified speed until the view is changed. Thus, no one realization is chosen. The realizations do not change at the data locations. The greatest changes would occur away from the data.

The rest of this paper is organized as follows. The second section reviews different techniques used in this research: distance of realizations, simulated annealing and image morphing. The third section describes the new methodology. The fourth section shows and discusses some examples using the proposed technique. Finally, conclusions are presented and future work is outlined.

2. Methodology

2.1. Distance between realizations – continuous variables

Even with computational advances over the years, exploring hundreds of geostatistical realizations simultaneously is not well understood. The spatial pattern of any single realization is not of particular interest. Instead, there is more information in understanding the features that can vary between realizations and the features that are consistent across realizations.

Summary models of uncertainty are useful. The local variance, probability of net reservoir or difference between realizations are useful summary statistics. Visualization of such models would quickly reveal areas that are more uncertain and areas that are less uncertain. However, these visualizations would not show the heterogeneity and joint uncertainty between multiple locations that impacts reservoir performance. Summary models inevitably change smoothly away from well control. Focusing on such models may lead the professional to believe the reservoir properties change smoothly.

Another approach is to define a summary response function to index the geostatistical images. This response could be used to order or classify the realizations. Several different similarity metrics are used for this purpose in the literature. In the environmental sciences, for example, similarity metrics include empirical orthogonal functions (Koch et al., 2015), connectivity analysis (Koch et al., 2016), fractions skill score (Roberts and lean, 2008), feature based analysis (Wolff et al., 2014) and spatial prediction comparison test (Gilleland, 2013).

In petroleum applications, the realizations could be ordered according to static reservoir properties such as porosity, facies proportions or hydrocarbon volume. In terms of dynamic properties, realizations could be ordered by connectivity or flow response. A scalar response would order the realizations, but the spatial similarity between adjacent realizations is not guaranteed; realizations with similar average response could be quite different. A natural ordering of the realizations requires calculating the difference between them. The greater the difference, the greater the distance between any two realizations. Formally, a distance is a function D with nonnegative real values that presents a symmetric property. Considering two different points A and B, the distance calculated from A to B is equal to the distance calculated from B to A, i.e. $D_{A,B} = D_{B,A}$.

A spatial distance function is proposed to compare the entire set of images coming from an underlying pool of realizations. A Euclidean-based distance provides a simple formulation and a clear interpretation. Let I and I' be two realizations that are composed of N visible grid cells. The distance $D_{I,I'}$ is the sum of differences between realizations, which is calculated cell-by-cell as follows:

$$D_{l,l'} = \sqrt{\sum_{m=1}^{N} \{Z_{l,m} - Z_{l',m}\}^2}, \tag{1}$$

where m denotes the grid cells that are visible, N denotes the number of visible grid cells and Z_m denotes the cell values of the two realizations.

The cells could be weighted by how much of them is visible. For example, cells on a cross section may not be as visible as those in plan view (refer back to Fig. 1). To avoid undue influence of extreme values, these distances could be calculated after applying a moving window filtering. This modifies a cell value by taking the average of a fixed