

Spatial uncertainty quantification of siliciclastic reservoirs, using the distance kernel method

Mihály Apró*

Department of Geology and Paleontology, University of Szeged, Szeged, Hungary

Received: July 28, 2017; accepted: January 8, 2018

Assessing the uncertainty in reservoir performance is a necessary step during the exploration phase. To examine the uncertainty in flow response, a large set of realizations must be processed. There are several stochastic geostatistical algorithms capable of simulating multiple equiprobable realizations. Although these can show us the possible realities highlighting the spatial uncertainty, their handling is time- and CPU-consuming during the later processes, such as flow simulations. Consequently, only a small number of realizations can be post-processed in industrial practice. The purpose of this work is to develop a method, which will reduce the huge number of realizations in a way that the remaining ones retain the spatial uncertainty of a reservoir's flow behavior, as would be demonstrated by a larger set of realizations. To solve this problem, ranking methods can be applied. Traditional ranking techniques, such as probability selection, are highly dependent on the applied static properties. In this paper, an alternative selection method is parameterized for measuring the pairwise dissimilarity between geostatistical models, with a distance function based on the hydrodynamic properties of the hydrocarbon reservoirs. The effectivity of the method is highly dependent upon the selected criteria. Thus, the distance function refers to the flow responses and allows visualizing the space of uncertainty through multidimensional scaling. A kernel transformation of the MDS data set is required to obtain a feature space where the K-means algorithm can discover non-linear structures in the basic data set. The final step of the method is the selection of the Earth models closest to the cluster centers. This tool allows for the selection of a subset of representative realizations, containing similar properties to the larger set.

Keywords: uncertainty, stochastic realization, clustering, kernel, response

*Corresponding address: Department of Geology and Paleontology, University of Szeged, Egyetem u. 2–6, Szeged H-6722, Hungary

E-mail: apromisi@gmail.com

This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.

Introduction

Stochastic spatial simulation is a widely used method to quantify and display spatial uncertainty. Multiple, alternative, and equiprobable realizations are used to represent the spatial uncertainty of the given simulated variables. This set of realizations is a space sampling of uncertainty of a spatial phenomenon. In most cases, these realizations are not enough to assess the uncertainty; further processing must be applied to represent the process. For instance, in reservoir engineering, the generated geologic models are submitted for flow simulation. They can be petrophysical and/or lithological models. These are suitable to assess reservoir flow performance, to assess the effect of drilling new wells, and to locate their placements (Scheidt and Caers 2007).

Handling of many realizations (100–1,000) is both time- and CPU-consuming. Currently, traditional ranking techniques can be used to select representative models from a set, such as probability-based, or quantile selection (P10, P50, and P90), which is often based on static parameters. In general, a “transfer function” can be applied to post-process the set of Earth models to reduce the huge amount of data, focusing on the flow response of the reservoir, while retaining the information of spatial uncertainty. This method considers the hydrodynamic properties of the realizations and defines the dissimilarity between each model. There are several similar methods, which can cluster the realizations based on flow properties, such as streamline simulations, or Original Oil in Place (OOIP) estimations. These require dynamic assessments, such as fast-flow simulations, while the flow parameters of the presented technique were originated from static properties (Caers 2011).

The aims of the research are to create a “transfer function,” which can make the realizations comparable pairwise, and to quantify and visualize the spatial uncertainty in terms of the flow response of the reservoir. The main goal of this study is to find a method which can reduce the huge number of realizations for hydrocarbon reservoirs in a way that the remaining ones retain the information of spatial uncertainty for the reservoir’s flow behavior as represented by the larger set of realizations. This could be useful for further processes, such as dynamic simulations of the selected models or evaluation of the reservoirs.

Data set

The input data set may contain several hydrocarbon realizations simulated by any kind of sequential simulation algorithm due to the representation of the spatial uncertainty. These may be generated in 3D, but averaging the values for every grid node along the Z axis is necessary (2D) to ensure their easier handling (Fig. 1A). These maps should be inverted for further processes.

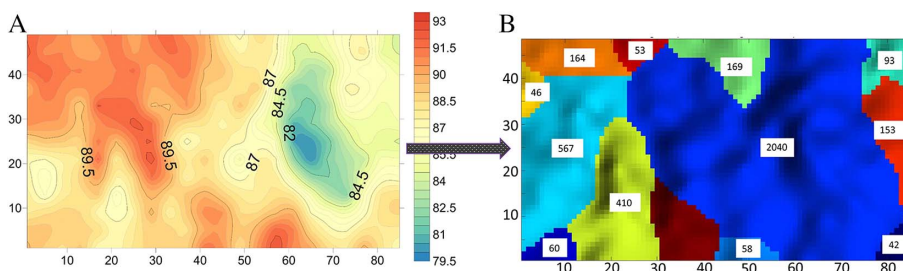


Fig. 1
An average, inverse porosity map (A), and its drainage basins (B)

Methods

Measure of dissimilarity

As was previously mentioned, a “transfer function” is necessary to define the dissimilarity (distance) between each pair of map data points. This approach was presented by Arpat (2005) and Suzuki and Caers (2006). The distance can provide information about how similar two reservoir models are, considering geologic and/or flow-related properties. However, these measurements depend on the pre-determined goals and assumptions of the spatial uncertainty examination. The most common definition of the “transfer function” is based on dynamic approaches, such as streamline simulation or OOIP, etc. (Caers 2011).

In this case, a geographical information system (GIS) tool called TopoToolbox was used to examine the 100 maps as topography maps and to derive hydrodynamic attributes from each one. There is a command package in the toolbox, which can calculate the drainage basins of each porosity/topography map (Fig. 1B; Schwanghart and Kuhn 2010). It is important to note that the areas that have higher porosity attributes will behave as basins, whereas the other areas will behave as mountains.

For the distance definition, the number of drainage basins, the area, and the heterogeneity of the largest basin (Fig. 2A) was considered. The largest basin always occurred at nearly the same area of the maps, representing the hydrocarbon reservoir (Fig. 1B), which is why it was chosen as a distinguishing property. The heterogeneity is represented by the median and the standard deviation of the cells associated with different velocity of the largest drainage basin (Fig. 2B; Schwanghart and Kuhn 2010).

In addition to the 2D properties described above, the volume of porosity higher than 15% was also measured in each 3D realization. Measuring the pairwise dissimilarities of the realizations, based on these properties, resulted in a dissimilarity distance matrix.

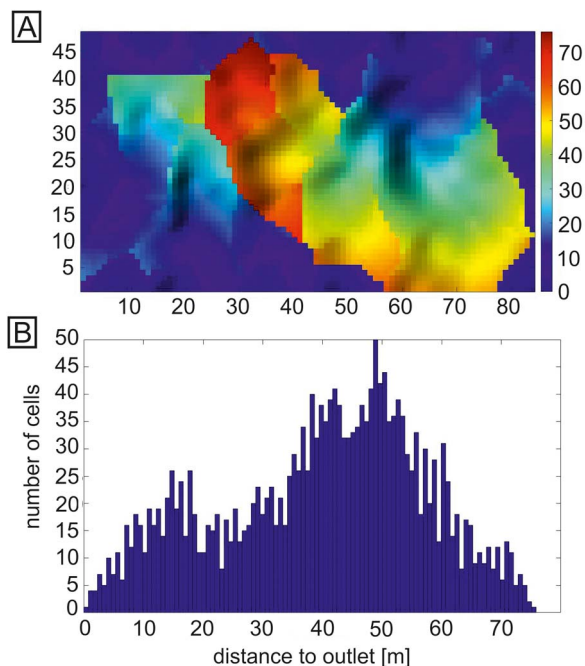


Fig. 2

The velocity map of the largest drainage basin (A), its histogram of number of cells (Y axis), associated with different velocity (X axis) (B)

Multidimensional scaling (MDS)

MDS is a technique applied to transform the dissimilarity matrix into a configuration of points (realizations) in n-dimensional Euclidean space. The distances between the points essentially correspond to the dissimilarities of the maps. There are two kinds of algorithms designed for examining a single dissimilarity matrix: classical and non-metric MDS. With these methods, the spatial uncertainty can be visualized. Supposing that the distance between the points is well-correlated to the flow behavior, points that are close to each other have similar flow attributes (Scheidt and Caers 2007). In this case, classical MDS was employed, because it assumes that the matrix displays metric properties, such as distance measured from the map; thus, the distances retain the intervals and ratios between the points as much as possible (Scheidt and Caers 2007).

Clustering of the realizations

Cluster analysis can discover the inner organizations of a data set by searching for structures within the point cloud. Thus, the data are divided into a number of groups,

containing a similar number of objects. The number of clusters are specified by density-based hierarchical clustering called OPTICS and dendrograms (Ankerst et al. 1999).

After choosing the proper number of clusters, K-means algorithms were applied to assigned points by minimizing the expected squared distance between points and the cluster centers (Scheidt and Caers 2007). Generally, the algorithm works properly, but it is sensitive to the distribution of points, because the initialization of randomly selected cluster centers is determined by the arrangement of points at the beginning of clustering process. A data set with less linearity can result in an erroneous initialization, which may group the object in a wrong way (Fig. 3).

The function

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right), \quad \sigma > 0,$$

where σ is the flexibility of the function, x and y are the coordinates of the points, was applied to transform the points into feature space, and arrange them more linearly (Caers 2011). The process is shown in Fig. 3. The meaningful K-means clustering is ensured in this space and inner structures may be discovered within the data set after a back-transformation (Scheidt and Caers 2007).

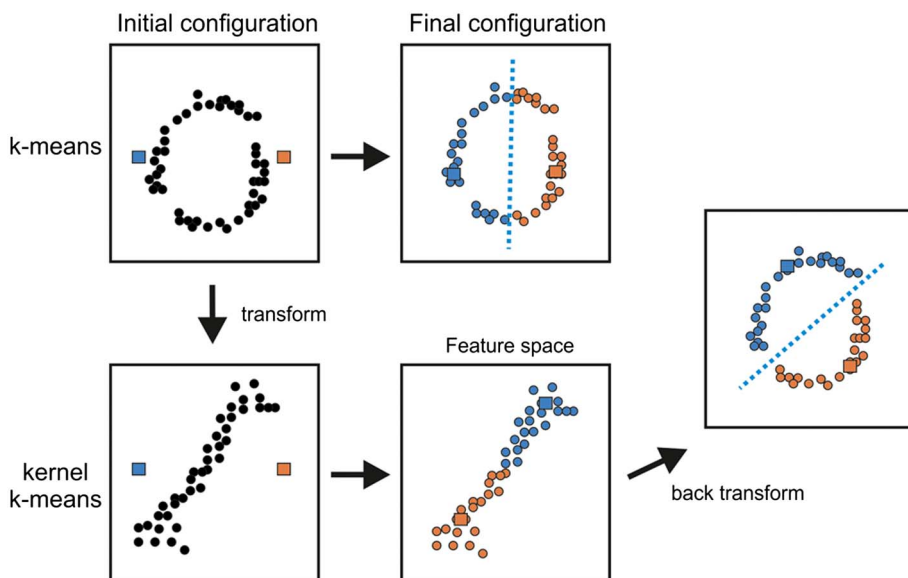


Fig. 3
Difference between K-means and kernel K-means clustering

Selection of the representative realizations and evaluation of the clusters

Finally, the selection of the representative Earth models closest to the cluster centers is the last phase of the method of uncertainty visualization and quantification (Caers 2011). The assessment of cluster membership is the key point to assess the quality of the clustering. Evaluation of the flow response of the selected Earth models and history matching can validate the method in the most appropriate way.

Geologic environment and input data set

The method was applied to a Lower Pannonian turbiditic sand body in Hungary. The input data for the Sequential Gaussian Simulation were porosity logs from 22 boreholes, resulting in 100 equiprobable realizations, with application of an omnidirectional, exponential variogram model, and $84 \times 48 \times 48$ grid resolution. After the dimensionality reduction ($3D \rightarrow 2D$), applying the GIS tools of Matlab, the dissimilarity distance matrix was computed, and the MDS coordinates were plotted (Fig. 4).

The geologic environment or the depositional system was not specified, because there were no core descriptions available. Based on previous studies, the sand body is part of a turbiditic system.

Results and discussion

In most cases, three realizations (P10, P50, and P90) are chosen for dynamic simulations; however, in this case, the density-based hierarchical clustering (OPTICS, dendrogram) resulted in four clusters.

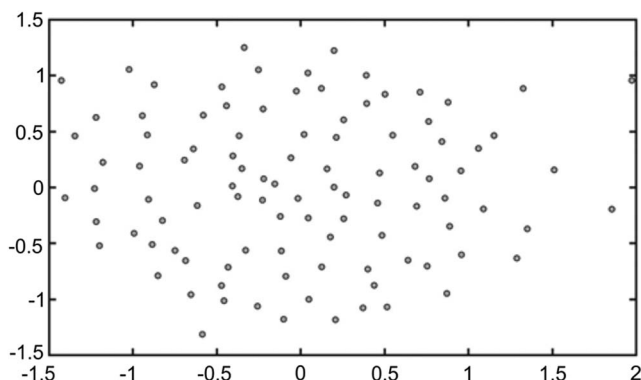


Fig. 4
Result of multidimensional scaling (X axis: components of dimension 1, Y axis: components of dimension 2)

Since the data structure of the MDS plot had less linearity, the radial basis function (RBF) was applied with $\sigma = 25$, which assigned the flexibility of the function (Fig. 5).

In this work, the classic K-means algorithm was applied in feature space to determine the previously defined four subsets. As can be observed in Fig. 6, among

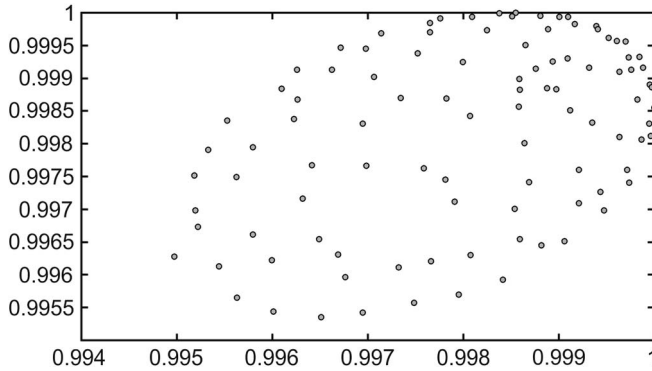


Fig. 5
The data set after the RBF transformation, in feature space, (X axis: components of dimension 1, Y axis: components of dimension 2)

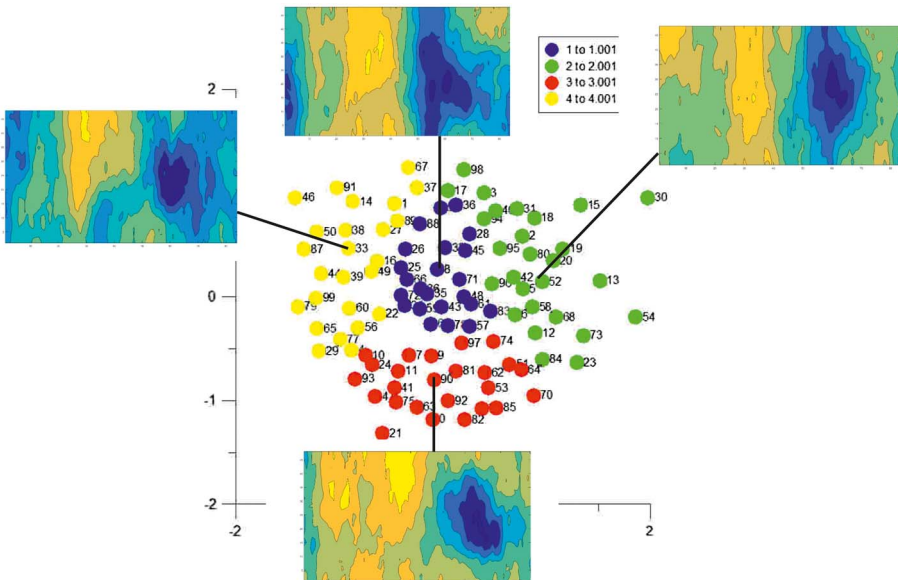


Fig. 6
Result of clustering with the selected realizations

the four clusters identified, the back-transformation one took a central position in the coordinate system and the other three clusters surrounded the central one. Realizations, being closest to the corresponding cluster centers, undoubtedly represent real groups. The distances between any two group centers are proportional to the dissimilarity between these groups (Fig. 6). The selected realizations can be post-processed during further flow-related assessments.

The resultant four groups may provide four possible realities about the spatial appearance and architecture of the hydrocarbon reservoir, based on the selected hydrodynamic features referring to the spatial uncertainty of the reservoir model. The number of the clusters can be user-defined, which may be limited by the available time and CPU capacity available for further processing.

Conclusions

The presented realization-grouping method does not require dynamic simulations, or OOIP estimations. It relies on flow-related and volume properties, and on the relationship between the distance measurement and flow behavior of the realizations. The selected four models represent the four clusters, and they also provide information about the spatial uncertainty of the reservoir, focusing on its performance. The application of this new process showed promising results, but it requires further improvements. Other distance measurements should be tested in more complex cases involving more distinguishing properties; for instance, evaluation of the streamlines can make the distance measurement more accurate and more reliable.

Acknowledgements

The author would like to thank MOL Plc. for providing the well-log data set and reservoir identification report, as well as the Department of Geology and Paleontology (University of Szeged) for the technical background and for professional help. The author would also like to thank Paul Haryott (Rose & Associates) for his professional review.

References

- Ankerst, M., M.M. Breunig, H.-P. Kriegel, J. Sander 1999: OPTICS ordering points to identify the clustering structure. – Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data, Philadelphia, pp. 49–60.
- Arpat, B.G. 2005: Similarity measures. – In: Arpat, B.G. (Ed): Sequential Simulation with Patterns. PhD dissertation, Stanford University, Stanford, pp. 62–82.
- Caers, J. 2011: Modeling response uncertainty. – In: Caers, J. (Ed): Modeling uncertainty in the earth sciences. Stanford University, Stanford, pp. 156–186.

- Scheidt, C., J. Caers 2007: A workflow for spatial uncertainty quantification using distances and kernels. – Stanford Center for Reservoir Forecasting Annual Meeting Report 20, Stanford University, Stanford, 34 p.
- Schwanghart, W., N.J. Kuhn 2010: TopoToolbox: A set of Matlab functions for topographic analysis. – *Environmental Modelling and Software*, 25, pp. 770–781.
- Suzuki, S., J. Caers 2006: History matching with an uncertain geological scenario. – SPE Annual Technical Conference and Exhibition, San Antonio, TX, pp. 24–27.