APPLICATION OF GEOSTATISTICAL SIMULATION MODELS IN THE CHARACTERIZATION OF COMPLEX GEOLOGICAL STRUCTURES

Sideri D.\textsuperscript{1}, Modis K.\textsuperscript{1} and Rozos D.\textsuperscript{1}

\textsuperscript{1} National Technical University of Athens, School of Mining and Metallurgical Engineering, 9 Heroon Polytechniou St, 15780 Athens, Greece, dsideri@metal.ntua.gr, kmodis@mail.ntua.gr, rozos@metal.ntua.gr

Abstract

Geostatistical simulation methods are able to generate numerical models or relations of the spatial distribution of a continuous geologic variable (grade, thickness, density, etc.) or a categorical variable (geological units and lithofacies or rock types). In this work, a review of traditional simulation techniques, as the Sequential Indicator Simulation (SIS), reveals a major pitfall that comes from theoretical difficulties in the development of a valid cross covariance model. On the contrary, a valid indicator cross covariance model is automatically defined in the framework of the Truncated Gaussian Simulation Method (TGS). This method is based on the concept that the categorical variables are obtained by truncating one standard multigaussian random variable at different thresholds. Plurigaussian Simulation Method (PGS) is an extension of the TGS Method but based on the simultaneous truncation of several multigaussian variables. An application of Plurigaussian method to simulate the lithofacies in the alluvial formations of the West Thessaly Basin is finally presented. This method was shown to be effective in reproducing the spatial characteristics of the different lithofacies and their distribution across the studied area.

Key words: Plurigaussian, Lithofacies modelling, Categorical variables.

Περίληψη

Η χρήση μεθόδων προσομοίωσης στη γεωστατιστική μπορεί να οδηγήσει στην ανάπτυξη αριθμητικών μοντέλων χωρικής κατανομής γεωλογικών μεταβλητών (περιεκτικότητα, πάχος, πυκνότητα, κλπ) ή κατηγορικών μεταβλητών (γεωλογικοί σχηματισμοί και λιθολογικές φάσεις ή τύποι πετρωμάτων). Στην παρούσα εργασία, η ανασκόπηση των κλασικών μεθόδων προσομοίωσης, όπως η Sequential Indicator Simulation (SIS), αναδεικνύει ένα σημαντικό μειονέκτημα που προκύπτει από τις θεωρητικές δυσκολίες στην ανάπτυξη ενός έγκυρου μοντέλου συνδιασποράς. Αντίθετα, ένα παρόμοιο μοντέλο μπορεί να οριστεί αυτόματα στο πλαίσιο της Truncated Gaussian Method (TGS). Η μεθόδος αυτή βασίζεται στη δημιουργία κατηγορικών μεταβλητών μέσω της αποκοπής μιας πολλαπλά κανονικής τυχαίας μεταβλητής σε διάφορα ορία. Η Plurigaussian Simulation Method (PGS) αποτελεί επέκταση της προηγούμενης με τη διαφορά στην ταυτόχρονη αποκοπή περισσότερων της μιας τυχαίας μεταβλητών. Στη συνέχεια της εργασίας παρουσιάζεται μία...
1. Introduction

To model the geology and patterns of compressible structures in heterogeneous alluvial deposits and other complex geological formations, conceptual and deterministic models are not in general as flexible and realistic to adequately represent the internal geometry (Falivene et al., 2007a). Although a possible exception is the Inverse Distance Weighting (IDW) interpolation method, geostatistics is preferable (Modis et al., 2008). For this reason, stochastic models such as geostatistical estimation and simulation are most commonly used (Bierkens and Burrough, 1993; Deutsch, 2006; Mariethoz et al., 2009; Staflou et al., 2011) particularly for the characterization of hydrofacies distribution within the alluvial deposits.

Geostatistical estimation is used to interpolate the value of a given attribute at a given location, by minimizing the error and bias of the estimate. It facilitates quantification of the spatial features of geological parameters and enables spatial interpolation (Komnitsas and Modis, 2006).

The use of estimation methods for obtaining facies reconstruction has not been as widespread in the literature as the use of simulation methods for obtaining facies models (see next paragraph). Johnson and Dreiss (1989) and Ritzi et al. (1995) applied a categorical estimation method (Indicator Kriging: IK) to obtain facies reconstructions in clastic aquifers, in both cases with only two different facies categories, making the results of this method very similar to those obtained with a continuous method. Falivene et al. (2007b) used a categorical method (IK) to reconstruct facies distribution in a fine-grain alluvial fan. Moreover, Falivene et al. (2007a) compared visually and statistically several facies reconstruction methods (Truncated Inverse Distance Weighting: TIDW, Truncated Kriging: TK, Indicator Inverse Distance Weighting: IIDW, and Indicator Kriging, among others) applied to a heterogeneous coal seam, this work was based on the same dataset as here. Other studies have focused on estimation of continuous parameters that can directly be related to facies (mud fraction in Flach et al., 1998; grain-size compositions in Koike et al., 1998; or results of geotechnical cone penetration tests in Lafuerza et al., 2005). This approach is conceptually similar to continuous methods for facies reconstruction (TK and TIDW). In all cases, published facies reconstructions achieved by estimation methods have been derived from extensively sampled sites. The scarcity of facies estimation in the literature (compared to simulation facies models, see below) is due to: a) the lack of digital high-density (with respect to facies heterogeneity) detailed facies descriptions; and b) the smoothing effect of facies reconstructions yielding usually optimistic results compared to the real heterogeneity distribution (i.e. more homogeneous distributions), limiting their predictive use. Moreover, as the density of the dataset with respect to facies heterogeneity decreases, the smoothing effect of facies reconstructions increases.

Geostatistical simulations on the other hand, consist of a set of methods being able to generate numerical models or relations of the spatial distribution of a continuous rock variable (grade, thickness, density, etc.) or a categorical variable (geological units and lithofacies or rock types). The aim is to generate equiprobable images of the real situation showing the characteristics of the phenomena, such as variability of the dataset and probability distribution function (Deutsch and Journel, 1997; Chiles and Delfiner, 1999). Simulated images exhibit the complexity of the dispersion as equally probable scenarios of the internal structure of the alluvial deposits.
The objective of this study is to review and explain the application of geostatistical simulation methods for the characterization of complex geological structures. As an example, the powerful and flexible PGS method (Le Loc’h et al., 1994) is applied to investigate the spatial variability of compressible units in the West Thessaly basin, where extensive land subsidence has been observed the last decades (Kallergis, 1973; Marinos et al., 1995; Rozos and Tzitziras, 2002).

2. Materials and Methods

2.1. Related Concepts from Geostatistics

In geostatistics, the variable under study is considered as a Random Function (RF) in space-time. A RF can be viewed as a collection of correlated random variables, say, \( x_{\text{map}}(p_1, \ldots, p_m, x_k) \), where the symbol \( x_{\text{map}} \) is adopted because the goal is to obtain maps displaying estimates at points \( p_k \) of the unknown values \( \chi_k \) of the natural variable from its observed values \( \chi_1, \ldots, \chi_m \). A realization of the RF at these points is denoted by the vector \( \chi_{\text{map}} = (\chi_1, \ldots, \chi_m) \).

Prob\[ \chi_1 \leq X_1 \leq \chi_1 + d\chi_1, \ldots, \chi_m \leq X_m \leq \chi_m + d\chi_m, \chi_k \leq X_k \leq \chi_k + d\chi_k \] = \( f_G(\chi_{\text{map}}) \) d\chi_{\text{map}}

where the subscript \( G \) denotes the general knowledge base (Christakos, 2000) used to derive the pdf.

A generally incomplete but in many practical applications satisfactory characterization of this RF is provided by a limited set of statistical moments, generally calculated from the experimental data.

In most practical applications, the underlying pdf is considered as Gaussian, and as a result, first and second order moments are sufficient to characterize. Since the first moment (mean) function is not generally known, a special form of the second moment, namely the variogram function (Journel and Huijbregts, 1978), is most widely used.

In the case of categorical attributes (e.g. geological facies) indicator variables are used to define whether a given point lies within a certain facies (Deutsch, 2006). The indicator variable takes the value 1 if the point belongs to the facies in question or 0 otherwise. Hydrofacies in our example (see section 2.3) is a categorical variable. The spatial average of the indicators is the average probability representing the proportion of that type of facies in the study field.

2.2. Geostatistical Simulation Methods

The traditional simulation techniques for several facies, as the Sequential Indicator Simulation (SIS), for indicators are based on the use of a random path through the field to sequentially simulate each node of the simulation grid (Gomez-Hernandez and Srivastava, 1990; Goovaerts, 1994; 1997). For the current node the indicator is firstly cokriged from those points already simulated and the conditioning data. The resulting estimate is assumed to represent an estimate of the probability that the node belongs to the different facies. The major pitfall in applying this technique comes from the cross covariance model required for the cokriging step. Direct and cross indicator covariances must satisfy the positive definiteness conditions for a valid cross covariance model to be defined (Emery, 2004). To avoid this very difficult step, an intrinsic correlation model is often assumed (Journel and Huijbregts, 1978). All direct and cross covariances are assumed to be proportional to one underlying covariance model, which reduces the cokriging to kriging.

On the other hand a valid indicator cross covariance model is automatically defined in the framework of the Truncated Gaussian Method (Galli et al., 1994). In this method (see end of this section), the lithofacies are not simulated directly: a stationary gaussian random function is simulated first, and then, it is transformed into the lithofacies variable by truncation (Armstrong et al., 2011). It is based on the notion that the indicators are obtained by truncating one standard
multigaussian random variable at different thresholds that are calculated from the known means of
the various indicators.

Plurigaussian Simulation Method (Le Loc’h et al., 1994; Armstrong et al., 2011) is an extension of
the Truncated Gaussian Method but based on the simultaneous truncation of several multigaussian
variables thus allowing the simulation of different facies displaying different spatial anisotropies.
The Plurigaussian technique provides a valid indicator cross covariance model because the
indicator covariances are directly deduced from the covariance model of the Gaussian variables.
Assigning a standard covariance model (for example spherical, exponential or Gaussian) to Y1 and
Y2 ensures that the indicator covariances are valid covariance functions. The two Gaussian
variables will be assumed to have stationary covariances.

2.3. Geology and Dataset

Thessaly basin is a lowland area in Central Greece, of about 4,520 km², mainly drained by Pinios
River. The basin is subdivided by a group of hills in two sub-basins, the Western and the Eastern
one. They are two main individual hydro-geological units, developing high potential aquifers.

The alluvial deposits of the Western one, which is the examined area of this study, constitute a
system of unconfined shallow aquifers, extending in the upper layers, and successive confined
- artesian aquifers developing in the dipper permeable layers (Marinos et al., 1995; 1997). This
system besides the percolated surface water is also supplied by water through the lateral
infiltration from the karstic aquifers of the alpine carbonate formations, outcropping in the margins
of the basin. In general, the richest aquifers are developed in the western sub-basin of Thessaly
plain, due to their rich supply both from the big infiltrating part of the surface runoff and the lateral
infiltration. The overexploitation of these aquifers during the last decades led to the manifestation
of land subsidence phenomena (Marinos et al., 1995; 1997), with extended damages in certain sites
(Rozos et al., 2010).

According to geological studies of Western Thessaly basin (Mariolakos et al., 2001; Rozos and
Tzitziras, 2002), Mesozoic Alpine formations outcrop in the margins of the study area, while, post
alpine deposits are presented in the lowland of the basin. The Mesozoic Alpine formations consist
of Schist-chert formation, Ophioliths, Limestones and Flysch sediments. They belong to the
Pelagonian (Subpelagonian) geotectonic zone and they constitute the bedrock of post Alpine
deposits of the West Thessaly basin.

The post-alpine deposits consist of: (a) Fluvial and lacustrine Neogene or molasse deposits,
including various phases such as clayey silts, clayey sands with scattered gravels and pebbles of
various origin and semi coherent conglomerates, (b) Terrestrial Pleistocene deposits consisting of
clayey silts, sandy silts & clays and sands, with grits and gravels and (c) Alluvial deposits
consisting of clays to sandy clays of variant thickness.

Throughout the Western Thessaly basin, several hundreds of sampling boreholes were drilled, in
order to provide data for various geotechnical and hydro geological studies. Based on the above
data, but also on some extra laboratory tests in samples collected from various soil horizons from
the study area, all geotechnical parameters were analyzed and interpreted. Following to this action,
a dataset was created, including soil characterization of distinguished horizons and their range of
main physical characteristics and mechanical properties.

This dataset includes lithological information of 8029 samples taken from 1039 boreholes (Figure
1).
After the examination of the well logs, 8 geologic types were identified and coded as following: Peat (PT), Clay gray (CG), Sandy clay gray (SCG), Clayey silt brown (CSB), Sandy clayey silt brown (SCSB), Sand (SA), Sand with gravels (SWG) and Basement formations (BASE). These lithotypes were further simplified for the modelling purposes into five lithofacies. The different lithofacies identified in the region are described in terms of categorical variables, or indicators, and represent the following categories: fine materials, fine materials with organic matter, sand, gravels and basement formations (Figure 2). The differentiation between the first two classes is based on the variation in compressibility, a characteristic that mainly depends on the presence of the organic matter in the soil.

3. Results and Discussion

3.1. Application of the Plurigaussian Method in the West Thessaly Basin

According to the available information (SOGREAH SA, 1974; Marinos et al., 1995; 1997) the two alluvial aquifer systems existing in the West Thessaly basin, can be approximately defined by three boundary surfaces, as shown on Figure 3. A thin impermeable layer separates the two aquifer systems, while the lower surface is the limit between the alluvial deposits and the bedrock.
The upper aquifer lies between elevations 70 m and 170 m, while the bottom surface of the lower aquifer extends up to -450 m. The total working area is discretized into 180 x 140 x 71 (1,789,200) blocks of 500 x 500 x 10 m. The coordinates of the center of the lower left block are x = 290,000 m, y = 4,330,000 m, z = -450 m. Discretization of the working domains and further application of Plurigaussian simulations were performed using ISATIS® software from Geovariances.

The lithofacies proportions at each domain must be modeled as a function of the point considered in the field. This modeling procedure is guided mainly by the experimental data. Vertical Proportion Curves (VPCs), first proposed by Matheron et al. (1987), are a simple tool for quantifying the evolution in the amount of each facies or lithotype present as a function of depth. They are computed along lines vertical to the chosen reference level. The results are presented as a graph showing the proportion of each facies at each level.

The global VPCs in the case of Western Thessaly basin are shown in Figure 4 for the two working domains. The differentiation between these hydrogeological units is apparent. A higher proportion of clays and silts with organic matter (black color) are observable in the second VPC. This is an indication of increased compressibility of the lower aquifer related to the upper one. Also, sand with gravels proportions (blue color) are grater in the lower aquifer and this suggests a more permeable structure.
In our case (Figure 5) we used two Gaussian functions G1 and G2. G1 represents the transition between greyish clays & silts with organic matters and sand with gravels, sand with gravels and sand, sand and basement formations. G2 represents the transition between brownish clays and silts and all the other lithofacies. This rule reflects the fact that the clay & silt brown facies is in contact with all other facies as seen from the histograms on the right column of Figure 5, which are calculated from the borehole data. Experimentation with alternative lithotype rules showed that the above rule leads to improved variogram fitting, as can be seen in the next paragraphs of this section.

Figure 5 - Lithotype rule with facies transition proportions on the right column. The correspondence between colors and lithofacies is shown in Figure 2.

The next step of the Plurigaussian technique is the inference of the variogram models for the underlying multi-Gaussian functions. Direct adjustment to the experimental variograms is not possible since the only available experimental variograms are the variograms of the indicator functions describing the lithofacies (one per lithofacies, plus all the bivariate combinations), while the two variograms needed for the model are the variograms of the underlying and continuous multi-Gaussian functions. The variogram inference was based on an inverse procedure in which the ranges of the variograms of the multi-Gaussian fields were adjusted iteratively through an inverse procedure, until an acceptable match was obtained between the experimental and the computed variogram (Figure 6).

Figure 6 - Horizontal indicator simple and cross variograms of the lower aquifer.
3.2. Simulation Results and Discussion

The resulting model of the Plurigaussian simulation is shown in Figure 7. A spatial differentiation concerning facies proportions is clear in this model, as for example the gravel percentage in the NW part, which is in agreement with observations concerning water recharge variations, as verified in relative studies (Kallergis, 1973; Bathrellos, 2005). The amount of grayish clays & silts with organic matters is clearly grater in the lower aquifer of the Plurigaussian model compared to the upper one, being in accordance to the proportion curves (Figure 4). Also, the existence of basement formations in the lower edges of the model is justified by the definition of the model boundaries. In general, the distribution of spatial characteristics of the basin is acceptably represented by the Plurigaussian model (Figure 7).

![3D Plurigaussian model of the West Thessaly aquifer systems.](image)

Figure 7 - 3D Plurigaussian model of the West Thessaly aquifer systems.

4. Conclusions

A theoretical drawback in applying the traditional simulation techniques for several facies comes from the cross covariance model required for the cokriging step. Direct and cross indicator covariances must satisfy the positive definiteness conditions for a valid cross covariance model to be defined. On the contrary, the Plurigaussian technique provides a valid indicator cross covariance model because the indicator covariances are directly deduced from the covariance model of the Gaussian variables.

The main characteristic of the Plurigaussian technique is that it allows incorporating a simple geological concept in the stochastic simulations. This is an important feature as in most applications, a detailed geological model is difficult to establish. The geological rule is not derived only from a statistical analysis of the borehole data but also from a geological analysis of all the available information.

In the case of western Thessaly basin, the Plurigaussian technique was shown to be effective in reproducing the spatial characteristics of the different lithofacies and their distribution across the studied area, as determined by the proportions of the borehole samples and other sources of empirical knowledge concerning water recharge variations.
5. Acknowledgements
The authors acknowledge Edafos S.A., Ferrovial Agroman S.A. and HellasLab for the supply of all necessary borehole data for the analysis.

6. References
Falivene O., Cabrera L. and Sáez A. 2007. Large to intermediate- scale aquifer heterogeneity in fine-grain dominated alluvial fans (Cenozoic As Pontes basin, NW Spain): Insight based on 3D reconstruction, Hydrogeology Journal, 15: 861–876 (b)


