Stochastic simulations of dependent geological variables in sandstone reservoirs of Neogene age: A case study of the Kloštar Field, Sava Depression

Kristina Novak Zelenika¹ and Tomislav Malvić¹,²

¹Sector for Geology and Reservoir Management, INA-Oil Industry Plc., Šubićeva 29, 10 000 Zagreb, Croatia; (kristina.novakzelenika@ina.hr, tomislav.malovic@ina.hr)
²Department for Geology and Geological Engineering, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, 10 000 Zagreb, Croatia

doi: 104154/gc.2011.15

ABSTRACT

The research presented herein is the first attempt to perform geostatistical simulations on three geological variables, porosity, thickness, and depth to reservoir, in the Croatian part of the Pannonian Basin. The data were collected from a reservoir of Lower Pontian age in the Kloštar Field, located in the western part of the Sava Depression. All three variables were analyzed using sequential Gaussian simulations (SGS). Information regarding present-day depth, thickness, and locations of areas with higher porosity values were used to reconstruct palaeo-depositional environments and the distribution of different lithotypes, ranging from medium-grained, to mostly clean sandstones and to pure, basin marls. Estimates of present-day thickness and depth can help to define areas of gross tectonic displacement and the role of major faults that have been mapped in the field. However, since mapping of the raw data (including porosities) does not allow the reconstruction of palaeo-depositional environments, sequential indicator simulations (SIS) were applied as a secondary analytical tool. For this purpose, several cut-off values for thickness were defined in an effort to distinguish the orientation of depositional channels (main and transitional). This was accomplished by transforming porosities to indicator values (0 and 1) and by applying a non-linear indicator kriging technique such as the “zero” map for obtaining numerous indicator realizations by SIS.

In the SGS and SIS approaches, the simulations encompassed 100 realizations. A representative realization was then selected using purely statistical criteria, i.e., two realizations were almost always chosen in accordance with the order of the calculation. The 1st and 100th realizations were selected for each variable in the SGS and SIS and five indicator kriging maps were chosen for the thicknesses cut-offs.

Keywords: Sequential Gaussian simulations (SGS), Sequential indicator simulations (SIS), indicator kriging (IK), sandstone reservoir, Kloštar Field, Sava Depression

1. INTRODUCTION

Geostatistics is currently one of the standard geological tools applied in the exploration and development of hydrocarbon reservoirs. As a form of applied statistics, it allows the introduction of mathematical exactness to explain the results. In hydrocarbon reservoirs in Croatia, the use of deterministic methods such as kriging, cokriging, and stochastic simulations have led to better reservoir descriptions than derived using previous, simpler mathematical models (e.g., MALVIĆ & DUREKOVIC, 2003; SMOLJANOVIC & MALVIĆ, 2004, 2005; MALVIĆ, 2005; MALVIĆ & VELIĆ, 2010).

Stochastic simulations represent one form of a probabilistic approach aimed at creating a geological model. Here we present a first probabilistic approach to Croatian hydro-
carbon geology in a systematic application to estimate geological risk. Pioneering works on this topic were published in masters theses by FRANK (1992) and NOVINC (1992), with research later continued by HERNITZ et al. (2001) and MALVIĆ (2003).

The probabilistic approach is exclusively based on the Monte Carlo algorithm and has long been used as an estimation method in the petroleum industry. The Monte Carlo approach has been applied to probabilistically calculate hydrocarbon reserves in Molve, the largest Croatian field, and in several traps, e.g., the deeper parts of the Ferdinandovac-Vizvar Field and prospects in Syria. The approach can be very useful in the first stage of exploration, when data from only a few wells are available.

Stochastic simulation tools that include the Monte Carlo algorithm represent a logical upgrade to the probabilistic approach as applied in estimating reservoir variables and hydrocarbon reserves. These are deterministic methods that draw on a variogram model and kriging or cokriging as the “zero” or base realization. Although interpolation can be made using deterministic methods, mostly by inverse distance weighting, or simple or ordinary kriging, similar results can be obtained with stochastic simulations, by performing estimations (rather than interpolations) through the model’s cells. As in kriging, the stochastic method works on minimizing error estimation, but it also results in uncertainty, which accompanies each location where the estimation is performed. Running a simulation is often connected with the later selection of several characteristic realizations. Many realizations can be compared cell by cell and for each location, along with a calculated range of minimum and maximum values. There are also other numerical combinations that can be calculated for a set of realizations obtained by simulation. In general, the larger the data-set, the more accurate the obtained statistics.

MATA-LIMA (2008) used stochastic simulations (direct sequential simulations and co-simulations) to predict history matching, i.e. permeability distribution. According to the author, sequential simulations of continuous variable in general easily process variable with Gaussian distribution or indicator transformation of such variable. But, a direct sequential simulation approach can also reproduce models of covariance, i.e. simulated values can follow a simple kriging estimation, and their variance can match simple kriging variance.

Furthermore, example of fluvial sandstone reservoir simulations is described in KEOGH et al. (2007) on the North Sea hydrocarbon reservoir. The authors recommended stochastic algorithms in 3D geological modelling in fluvial reservoirs, where it is possible, more or less successfully, to show a series of depositional environments characteristic of different lithofacies in fluvial environments such as channel

Figure 1: Kloštar Field, located in the western part of the Sava Depression.
fill, crevasse splays etc. Sometimes is even possible to show internal features at the metre scale. Stochastic realizations are able to cover multiple scenarios of possible channel distributions, transitions facies or even sedimentation rates.

Moreover, ROBERTSON et al. (2006) also presented algorithms of direct sequential simulations, but which reproduce histograms. The authors also showed experimental semi-variograms calculated on a variable (permeability) that is not transformed in normal distribution. Results showed that applied algorithms led to given semivariograms very similar to the “regular” (normal distributed) model, but with a lower variogram sill. The final maps reproduced the spatial variability very well.

The BlockSIS software is shown in DEUTSCH (2006). Although the WinGslib package was used here to obtain the SIS maps, it is good to know that there are other options including BlockSIS, which allows the use of soft (secondary) data from geological interpretations or geophysical measurements. There are nine different techniques (e.g. simple kriging, ordinary kriging, collocated cokriging, block (co)kriging etc.) available in the program. DEUTSCH (2006) described an example tested with all nine techniques, and concluded that it is very difficult to decide on the most appropriate approach, although block kriging produced very good results. In cases with clustered data and secondary variables derived from geophysics, some other algorithm could be preferred.

Here, a stochastic analysis was performed on a dataset collected in the Kloštar Field (Figure 1). This area was previously analyzed by BALIĆ et al. (2008) and CVETKOVIĆ et al. (2008), who used geomathematical methods to demonstrate that the most appropriate method for porosity analysis is the kriging interpolation. These authors also established that a geostatistical approach is the most appropriate method for mapping geological variables in all clastic reservoirs of Neogene age in the Croatian part of the Pannonian Basin. MALVIĆ & BUREKOVIĆ (2003), for example, showed that, as an interpolation technique, ordinary kriging is better than inverse distance weighting, based on cross-validation, even in the case of a modest input dataset (about 15 points). Also, secondary seismic variables enable the use of collocated cokriging for the same size data set. Moreover, BALIĆ et al. (2008) tested ordinary kriging in a sandstone reservoir of the Kloštar Field and showed that it was a better interpolation approach than either the inverse distance weighting, moving average, or nearest neighbourhood method.

For the Kloštar Field, the stochastic approach was applied as an analytical tool to estimate three geological variables, porosity, thickness, and reservoir depth, using data collected in the same sandstone reservoir, referred to as “T” (a description of this reservoir is given in BALIĆ et al., 2008). This dataset is considered to be representative not only of present-day structural relationships (depth and thickness), but also of the depositional palaeo-environment as recorded in terms of reservoir porosity distribution and thickness. Moreover, by carefully selecting mapped data using an indicator approach and the appropriate cut-offs, “hidden” or subtle trends in the distribution of these two variables can be revealed and, consequently, the margins of the distribution system determined. Proper selection of cut-offs is especially important in the absence of a clear relationship between core lithology and average log-porosity values, as was the case in this study, in which only trends were used to select the limit values between marl, marly-sandstone transitions, and sandstone. All other cut-offs were selected using (a) the criteria of a minimum of five cut-offs (NOVAK ZELENIKA et al., 2010) and (b) a larger number of cut-off classes around the median values in the interval between minimum and maximum porosity (or thickness).

The “T” reservoir is made up of medium-grained sandstone in its central part, giving way laterally to sandy marlstone at the margins (Figure 2). It is the largest sandstone reservoir in a monotonous series of alternating marls and sandstones. It is also referred to as the first sandstone series (informal lithostratigraphic name). The sandstone is of Lower Pontian age (Figure 2) and was deposited in a deep, brackish lake (up to 200 m depth), characterized by a mostly calm environment (deposition of marl) but interrupted periodically by turbidity currents (deposition of sandstones). This depositional system is described in VRBANAC et al. (2010).

Here we examined the ability of stochastic analysis to provide more detailed information on the distribution of reservoir variables than that obtained by kriging, especially in the presence of several irregularly distributed lithofacies that are laterally discontinuous (sandstone, marly sandstone, sandy marl and marl), with borders that are not strictly defined. It was postulated that it is more appropriate to define a “fluid” border using stochastic methods, allowing it to change position in different realizations. This approach is especially valid for understanding reservoir porosity and thickness. The input dataset consisted of 19 data points for thickness, depth, and porosity (Figure 3). These points were selected from approximately 100 well points based on criteria obtained from the latest computer analysis of well logs, including average porosity values (previously not calculated), precise, stratigraphically determined reservoir tops and bottoms, and thickness (mostly from e-log markers).

2. THE BASICS OF STOCHASTIC SIMULATIONS

Stochastic simulations comprise a group of geostatistical tools based on an algorithm that differs from interpolation methods. Below we provide a short review of sequential Gaussian simulations (SGS), which is the stochastic tool most frequently used in the analysis of hydrocarbon reservoirs. Our review is based on the theory published in e.g. DUBRULE (1998), KELKAR & PEREZ (2002), and MALVIĆ (2008). Additional background information can be found in other geostatistical textbooks or references.

2.1. Processing of input data and “zero” realization

1. Existing measurements (point data from wells) are kept as constant values in each realization. These values are considered “hard data” and the type of simulation is referred to as a conditional simulation. This is in contrast to unconditional simulations, in which “hard” points are also estimated.
and referenced to a cell in which each point value is derived from the surrounding area, including the uncertainty of position).

2. A variogram model is constructed together with input for the kriging interpolation; this deterministic kriging map is called a “zero” realization.

3. In this type of deterministic solution the following values are always known:
   a) Mean value, variance ($\mu$, $\sigma^2$)
   b) Kriging variance ($\sigma^2_k$)
   c) The interval allowed for simulated values (determined from the relationship between the mean and the variance)
d) If the allowed interval encompasses three standard deviations (±3σ) around a mean (µ) each cell, then 99% of all possible solutions are included.

4. When all previous values are known and the type of simulation is defined, then the values of all model “blank” cells can be estimated using the SGS method.

2.2. Simulation

After normal transformation, the data are most often defined by the properties of a normal distribution, \( N(\mu, \sigma) \), and an estimation is performed on empty cells. Random choice of location is used to estimate one cell (the first introduction of randomness, i.e., stochasticity). The value of the selected cell is estimated mostly by kriging from other points in a spatial model. It is important to note that these points can be hard data or previously simulated points. After a cell is estimated deterministically, a new value is “surrounded” by the interval ±3σ (using the variance of the “zero” solution). Using a random choice (the second introduction of randomness), any value from this interval can be accepted as simulated for the cell. The procedure is repeated until the estimates of all cells have been made.

2.3. Calculation of a set of realizations

The purpose of running a simulation is to obtain an abundant set of realizations, since it is not possible to select one realization as the best, regardless of the criteria applied. It is possible, however, to select “the most appropriate” realization using certain criteria (mostly statistical quartiles or minimum and maximum sums of simulated cells). The large number of realizations in a simulation allows the inclusion of almost all uncertainties, i.e., each cell is sampled with enough values to calculate a reliable, probability density function (PDF). It is of course unnecessary to calculate several hundreds of thousands of realizations as this would be overly time-consuming for simulation and later post-processing. An acceptable number in a simulation is 100 realizations, which is large enough to be representative of the stochastic characteristics of a particular dataset. Among these realizations, several can be outlined as useful for modelling purposes; that is, if realizations represent volumes, the letter “P” (probability) followed by a number expresses the percentage of how many of the others have a smaller total volume. For example, P10 means that 10% of all realizations have a smaller volume.

As discussed above, the calculation of each new realization means that the order of simulated cells is set completely randomly. The consequence is very interesting, as different realizations for the same cell can be estimated using a different number for the “hard” data inside the variogram ellipsoid.

2.4. Advantages and disadvantages of SGS

The main advantage of SGS is the possibility of estimating values in all cells of a model through a set of realizations. Furthermore, the input data and errors of simulated values are characterized by a normal distribution. The second goal of SGS is to take advantage of new, abundant data to construct a histogram of simulated variables. The resulting histogram is much more reliable than a histogram derived from pure input data. For example, a model based on 50×50 cells and 100 realizations gives 250,000 output values, with an input set made up of only 10–20 measurements.

There are, however, disadvantages, e.g., simulated values can be significantly different in neighbouring cells. This may not be a problem in cases with large cells, which may better describe the subsurface than an interpolation. Due to the fact that simulation does not have one representative realization, it is necessary to use several criteria from a probability curve to select those that are more characteristically representative.

3. RESULTS OF SEQUENTIAL GAUSSIAN SIMULATIONS (POROSITY, THICKNESS, DEPTH)

All three selected variables analyzed by the SGS method and the respective results are presented in Figures 4–6. For each variable (porosity, thickness, and depth), two realizations were selected using purely statistical criteria, i.e., the 1\textsuperscript{st} and 100\textsuperscript{th} (first and last) calculated realizations.

There is a visible trend of higher porosity in a north-south direction (“stripes” of denser red colour). The respective values are located in the western part of the Kloštar Field, which is near the deeper part of the Sava Depression. The higher porosity in the deeper parts of the palaeo-environment corresponds to an area in which more porous sandstone would have been deposited by high-energy turbidity currents.

The stochastic maps of thickness shown in Figure 5 depict the 1\textsuperscript{st} and 100\textsuperscript{th} realizations. It is interesting, but not
surprising, that the trends are the same as those observed for porosity (Figure 4). This suggests that the most permeable sandstones were also deposited in the deeper palaeo-environmental setting; consequently, they are of maximum thicknesses.

The values for greater reservoir depth occur where the thicknesses are also greater, suggesting that structural inversion was not completed. It is well-documented (e.g., MALVIĆ, 2003; ROYDEN, 1988) that, in the Pannonian Basin, transpression during the Pliocene-Quaternary inverted many structures, thus accounting for the numerous anticlines and hydrocarbon traps. In the study area, this process occurred only partially, mostly due to the location of the Mt. Moslavča Gora, which is a subsurface, uplifted, pre-Neogene palaeo-relief located on the eastern side (depicted by the blue zones in Figure 6).

The SGS results portray a very clear correlation among the trends observed on the maps for each of the three variables (porosity, thickness, and depth). Also, the variation between particular realizations is apparent in all three pairs of 1st and 100th realizations. It is important to emphasize that realization nos. 1 and 100 do not match P1 and P100, because it was impossible to perform depth volume calculations, thus preventing this criterion to be generally applied. Accordingly, we performed another, purely statistical criterion selection.

4. RESULTS OF SEQUENTIAL INDICATOR SIMULATIONS AND INDICATOR KRIGING (THICKNESS)

Sequential indicator simulation is an alternative estimation method that uses original as well as indicator data for variogram development and mapping. A comparison of Figures 5 and 7 shows thickness values mapped by the SGS and SIS methods, respectively. It is clear that the SIS estimation provides more uniform and lower cell values; consequently, the differences between realizations are not as large as in SGS.
Figure 6: The 1st (left) and 100th (right) SGS realizations for depth (scale 600–1100 m).

Figure 7: The 1st (left) and 100th (right) SIS realizations for thickness (scale 0–30 m).

Figure 8: SIS probability maps (0–1) for thickness cut-offs of 5 m (left) and 9 m (right).
Given that the original data range is the same (0–30 m), then why does the SIS map have more final cell values closely estimated around the mean? The answer is that the variance of an indicator variable is $F'(z_k) [1.0 - F'(z_k)]$, where $F'(z_k)$ is the cumulative distribution function (CDF) of the continuous random variable $z$, defined as $F'(z_k) = \text{Pr} \{ z_k \leq z \}$. Also, five cut-offs were used to define the indicator classes. The greater the number of cut-offs used, the larger the reduction in within-class noise (DEUTSCH & JOURNEL, 1998). The intention of using SIS was to obtain indicator maps based on stochastic realizations for values higher than that of the selected cut-off (the post-process option in WinGslib™ for 100 realizations previously obtained with SIS was employed.) Note that the SIS probability maps represent probabilities that a value will be greater than the selected cut-off (Figures 8–10).

A comparison of the thicknesses on the probability maps obtained with SIS (Figures 8–10) and indicator kriging (IK) (Figures 11 and 12) shows that the SIS maps present probabilities that the cell values are larger than the cut-off, whereas in the IK maps the cell values are less than the cut-off. The contour shapes and interpretation are the same, but the SIS maps are “noisier,” with smaller transitions between values (note that the IK map for a 21% cut-off is not interpolated because heterogeneity would be observable only on the scale $p=0.990–1.000$, which is meaningless).

Stochastic results through realization allow the calculation of histograms from a more abundant dataset (simulated and hard data), thus offering better insight into variable distribution and statistics. Such histograms are shown in Figure 13 for porosity and depth, as calculated by SGS, and in Figure 14, for thickness calculated by SGS and SIS. As expected, the variable depth is not characterized by a normal distribution because the field’s structure thins along Moslavačka Gora, located in the northeast. This implies that the structure occurs in combination with inclined anticlines over a slight monocline dipping towards the southwest. Generally, this is apparent on the maps of reservoir bodies (porosity maps) and structures (depth maps) that are mostly oriented north-south.

5. CONCLUSIONS

A stochastic approach is especially useful when the amount of input data is moderate (18–23 points in this analysis), and the variogram models include large uncertainties. When variograms have a large nugget and there is a need for all models to be equal, as in the indicator approach, the stochastic approach will partially remove the “bull’s-eye” effect. This would emphasize interpolation and result in a noisy transition between cell values and hard data.

The two histograms in Figure 14 were calculated from different simulated sets: SGS (left) and SIS (right). As evidenced by the parallel histograms for the thickness of the analyzed sandstone reservoirs, the simulated values calculated by SIS are clearly more highly concentrated around the mean value. This also explains why the thickness realiza-
Figure 11: IK probability maps (0–1) for thickness cut-offs of 5 m (left) and 9 m (right).

Figure 12: IK probability maps (0–1) for thickness cut-offs of 13 (left) and 17 m (right).

Figure 13: Histogram of porosity (left) and depth (right) values as calculated by SGS.
tions in Figure 7 are significantly more uniform than those in Figure 5 and confirms the statement that “artificial” histograms, which include original and simulated values, need to be calculated from SGS results.

A depositional channel is evident in the probability maps obtained by IK and SIS, but only for certain cut-off values. In the case of the Lower Pontian sandstones in the “T” reservoir of Kloštar Field, this cut-off value corresponds to a thickness of 13 m (Figure 12). NOVAK ZELENIKA et al. (2010) reported similar findings when the porosity of the same reservoir was mapped by IK. In those maps, a depositional channel was also seen, but only in the probability map with a 19% porosity cut-off. Therefore, for mapping by IK-based methods, an adequate number of cut-off values is particularly important.

It is also possible to obtain SIS maps with values instead of probabilities. The final cell values of those maps (Figure 7) are estimated more closely around the mean and do not indicate the presence of a depositional channel.

No depositional channel was observed on maps obtained by SGS; instead, a monocline with several local tops is seen as in Figures 4 and 6. Higher porosity values occurred in the deepest part of the reservoir as well as in the section of greatest thickness, suggesting that tectonic, i.e., transpression, events starting at the end of the Pontian did not result in full inversion in the area of Kloštar Field.

A general conclusion that can be drawn from this study is that SGS can be used in the mapping of structural variables. This is in contrast to IK and SIS approaches, which instead described the morphologies of depositional environments in the Kloštar Field.

ACKNOWLEDMENT

This work represents part of a multidisciplinary geostatistical research project carried out in 2009 and 2010 within the larger project entitled “Stratigraphical and geomathematical research of petroleum geological systems in Croatia” (project no. 195-1951293-0237), financed by the Ministry of Science, Education and Sports of the Republic of Croatia.

REFERENCES


Manuscript received November 10, 2010
Revised manuscript accepted March 05, 2011
Available online June 09, 2011