Geostatistical Prediction of Reservoir Petrophysical Properties by Copula Based Dependence Models between Seismic Attributes and Petrophysical Properties*

Martín A. Díaz Viera¹, Arturo Erdely², Raúl del Valle García¹, Tatiana Kerdan¹, and Francisco Mendoza-Torres³

Search and Discovery Article #70246 (2017)**
Posted March 27, 2017

*Adapted from oral presentation given at AAPG 2016 International Convention and Exhibition, Cancun, Mexico, September 6-9, 2016
**Datapages © 2017 Serial rights given by author. For all other rights contact author directly.

¹Instituto Mexicano del Petróleo (IMP), Mexico City, Mexico (mdiazv@imp.mx)
²FES Acatlán (UNAM), Mexico City, Mexico
³Posgrado de Ciencias de la Tierra (UNAM), Mexico City, Mexico

Abstract

In the context of geological and petrophysical reservoir modeling proper prediction of petrophysical property spatial distributions is a crucial task because the correct estimation of reserves and the development and optimal exploitation of the reservoir heavily depend on it. Only in recent years, it has imposed on the oil industry a comprehensive and multidisciplinary approach that combines all available information sources such as core data, geological models, seismic surveys and well logs, by the application of geostatistical models in a systematic way. One of the most common ways to combine seismic data with well logs is to establish correlations between seismic attributes and petrophysical properties. These models are quite restrictive because in most cases they assume that variables follow a Gaussian distribution and a strong linear dependence exists between them. Moreover, also the classical multivariate geostatistical models as Cokriging and Sequential Gaussian simulation method (Parra and Emery, 2013) also consider these assumptions. A non-parametric (distribution-free) method is proposed, which does not assume linear dependence, but rather seeks to represent, reproduce and exploit the underlying dependency between attributes and petrophysical properties: a Bernstein copula dependence model that was successfully applied for petrophysical simulation at well log scale. The methodology consists of two steps: firstly, a dependence model between seismic attributes and petrophysical properties at well log scale is established and then this model is used to estimate (median regression approach) or to simulate (stochastic approach using simulated annealing) petrophysical properties to seismic scale. The application of the methodology is illustrated in a case study where the results are compared with sequential Gaussian method.
References Cited


Geostatistical Prediction of Reservoir Petrophysical Properties by CopulaBased Dependence Models Between Seismic Attributes and Petrophysical Properties

Martín A. Díaz Viera\textsuperscript{1}, Arturo Erdely\textsuperscript{2}, Raúl del-Valle-García\textsuperscript{1}, Tatiana Kerdan\textsuperscript{1}, Francisco Mendoza-Torres\textsuperscript{3}

\textsuperscript{1} Instituto Mexicano del Petróleo (IMP)
Eje Central Norte Lázaro Cárdenas 152, CP 07730 México, D.F.
e-mail: mdiazv@imp.mx
\textsuperscript{2} FES Acatlán (UNAM), \textsuperscript{3} Posgrado de Ciencias de la Tierra (UNAM)

September 8, 2016
Outline

- Introduction
- Methodology
- Case study
- Final remarks and future work
- Bibliography
Modeling the **spatial distribution** of petrophysical properties in reservoir characterization is a crucial and difficult task due to the lack of enough data and hence the degree of uncertainty associated with it.

For this reason, a **stochastic simulation** approach for the spatial distribution of petrophysical properties has been adopted.

Seismic attributes have been extensively used as secondary variables in static reservoir modeling for petrophysical property prediction but usually assuming **linear dependence** and Gaussian distribution (Parra & Emery, 2013).
Quite recently, *copulas* have become popular for being a flexible means of representing dependency relationships in the financial sector and applications are already emerging in the field of geostatistics (Bardossy & Li, 2008) (Kazianka & Pilz, 2010).

A geostatistical simulation method, based on *Bernstein copula* approach as a tool to represent the underlying dependence structure between petrophysical properties and seismic attributes, is proposed.
Introduction

- The procedure basically consists of applying the *simulated annealing* method with a joint probability distribution model estimated by a *Bernstein copula* in a completely non-parametric fashion (Hernández-Maldonado, Díaz-Viera, & Erdely, 2012).
- The method has the advantages of not requiring linear dependence or a specific type of distribution.
- The application of the methodology is illustrated in a case study where the results are compared with sequential Gaussian co-simulation method.
Methodology

• The main goal of this work
• show the application of a *Bernstein copula-based spatial stochastic co-simulation (BCSCS)* method for petrophysical property predictions using seismic attributes as secondary variables and its
• compare with the classical *Sequential Gaussian co-simulation (SGCS)* method
Sequential Gaussian co-simulation (SGCS)

• Usually this method is applied with a linear model of coregionalization (Chiles & Delfiner, 1999) which is mostly unnatural, forced, very complicated and difficult to establish.

• The method assumes the existence of very strong linear dependence between primary and secondary variables, which is its main assumption and at the same time its main drawback.

• Here we choose to use an alternative variant, the Markov Model, given in (Chiles & Delfiner, 1999, p. 305) and implemented in SGeMS (Remy, Boucher, & Wu, 2009).
The procedure consists of two stages:

1. A *dependence model*, using a *Bernstein copula*, is established from which a number of sample values are generated.

2. A *stochastic spatial simulation* is performed using a *simulated annealing* method with a variogram model and a bivariate distribution functions as objective functions *(Deutsch & Cockerham, 1994)*, *(Deutsch & Journel, 1998)*.
• Sklar (1959) proved that there exists a function $C_{XY}: [0,1]^2 \rightarrow [0,1]$ such that

$$H_{XY}(x, y) = C_{XY}(F_X(x), G_Y(y))$$

• $C_{XY}$ is called \textit{copula function} associated to $(X,Y)$ and contains all the information about the dependence relationship between $X$ and $Y$, independently from their marginal probabilistic behavior.
Copula-based dependence modeling

- Some copula function properties:
  
  - \( C(u, 0) = 0 = C(0, v) \)
  - \( C(u, 1) = u, \quad C(1, v) = v \)
  - \( C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \) if \( u_1 \leq u_2, v_1 \leq v_2 \)
  - \( C \) is uniformly continuous on its domain \([0,1]^2\).
Copula-based dependence modeling

• The word *copula* is a Latin noun that means “A link, tie, bond”
• Copula function is a flexible tool for building *joint probability distributions*
• Univariate models for the random variables of interest and the copula function may be chosen separately
• There are parametric, semi-parametric and non-parametric approaches.
• Particularly appropriate for non-linear dependencies
Copula-based dependence modeling

- When $F_X$ and $G_Y$ are known and $H_{XY}$ is unknown, if $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ is an observed random sample from $(X, Y)$, the set $\{(u_k, v_k) = (F_X (x_k), G_Y (y_k)) : k = 1, \ldots, n\}$ would be an observed random sample from $(U, V)$ with the same underlying copula $C$ as $(X, Y)$, and since $C = F_{UV}$ we may use the $(u_k, v_k)$ values (called copula observations) to estimate $C$ as a joint empirical distribution:

$$\hat{C}(u, v) = \frac{1}{n} \sum_{k=1}^{n} 1\{u_k \leq u, v_k \leq v\}$$

- Strictly speaking, the estimation $\hat{C}$ is not a copula since it is discontinuous and copulas are always continuous.
Copula-based dependence modeling

- $F_X$ and $G_Y$ are estimated by univariate empirical distribution functions:

$$
\hat{F}_X(x) = \frac{1}{n} \sum_{k=1}^{n} 1\{x_k \leq x\} \\
\hat{G}_Y(y) = \frac{1}{n} \sum_{k=1}^{n} 1\{y_k \leq y\}
$$

- Now the set of pairs $\{(u_k, v_k) = (\hat{F}_X(x_k), \hat{G}_Y(y_k)) : k = 1, \ldots, n\}$ is referred to as \textit{copula pseudo-observations}. 
Copula-based dependence modeling

- $F_X$ and $G_Y$ are estimated by univariate empirical distribution functions:

\[
\hat{F}_X(x) = \frac{1}{n} \sum_{k=1}^{n} 1\{x_k \leq x\}
\]

\[
\hat{G}_Y(y) = \frac{1}{n} \sum_{k=1}^{n} 1\{y_k \leq y\}
\]
Copula-based dependence modeling

- $F_X$ and $G_Y$ are estimated by univariate empirical distribution functions (Pérez & Fernández-Palacín (1987)):

\[
\tilde{Q}_n(u) = \sum_{k=0}^{n} \frac{1}{2} (x_k + x_{k+1}) \binom{n}{k} u^k (1 - u)^{n-k}
\]
The **empirical copula** is defined as the following function $C_n: I^2_n \rightarrow [0,1]$, where $I_n = \{ \frac{i}{n} : i = 0, \ldots, n \}$, given by:

$$C_n \left( \frac{i}{n}, \frac{j}{n} \right) = \frac{1}{n} \sum_{k=1}^{n} 1\{rank(x_k) \leq i, rank(y_k) \leq j\}$$

where $C_n$ is not a copula but it is an estimation of the underlying copula $C$ on the grid $I^2_n$ that may be extended to a copula on $[0,1]^2$ by means of, for example, a polynomial approximation.
Copula-based dependence modeling

- The *empirical copula* is defined as $C_n : I_n^2 \rightarrow [0,1]$, where $I_n = \{ \frac{i}{n} : i = 0, \ldots, n \}$, given by:

$$C_n \left( \frac{i}{n}, \frac{j}{n} \right) = \frac{1}{n} \sum_{k=1}^{n} 1_{\{ \text{rank}(x_k) \leq i, \text{rank}(y_k) \leq j \}}$$
Copula-based dependence modeling

As proposed in (Sancetta & Satchell, 2004), using Bernstein polynomials leads to what is known as a *Bernstein copula* non-parametric estimation $\tilde{C}: [0,1]^2 \to [0,1]$ given by:

$$
\tilde{C}(u, v) = \sum_{i=0}^{n} \sum_{j=0}^{n} C_n \begin{pmatrix} i \over n \end{pmatrix} \begin{pmatrix} j \over n \end{pmatrix} \binom{n}{i} u^i (1-u)^{n-i} \binom{n}{j} v^j (1-v)^{n-j}
$$
Copula-based dependence modeling

- As summarized in (Erdely & Diaz-Viera, 2010) in order to simulate replications from the random vector \((X, Y)\) with the dependence structure inferred from the observed data \(\{(x_1, y_1), \ldots, (x_n, y_n)\}\) we have the following algorithm:

1. Generate two independent and continuous Uniform \((0,1)\) random variates \(u\) and \(t\).
2. Set \(v = c_u^{-1}(t)\) where \(c_u(v) = \frac{\partial \tilde{C}(u,v)}{\partial u}\).
3. The desired pair is \((x, y) = (\hat{Q}_n(u), \hat{R}_n(v))\) where \(\hat{Q}_n\) and \(\hat{R}_n\) are empirical quantile functions for \(X\) and \(Y\), respectively.
The general workflow

- The general workflow is as follows:

1. Univariate data analysis,
2. Bivariate dependence analysis,
3. Variography analysis,
4. Simulations.
Case study

• Data used in the case study are from a deep water reservoir in the Gulf of Mexico.
• The data consist of a total porosity well-log from a well and seismic attribute (P-impedance) obtained in a vertical (inline) section.
• The well-log has a sample interval of 0.1 m.
• The section has a length of 412.5 m and covers an interval of 336.4 m in depth and was chosen so that the well was located in the middle of it.
Case study

Vertical (inline) section with P-impedance as a result of seismic inversion. The color scale represents impedance values. In the middle of the section two logs are plotted along a well: in yellow P-impedance and in green total porosity.
Univariate data analysis

-log scale

1m scale

seismic scale
Bivariate dependence analysis

log scale

1m scale

seismic scale

1m scale simulation

Spearman

Pearson

well-log scale

one-meter scale

seismic scale

1m_BCS

-0.589

-0.711

-0.477

-0.657

-0.361

-0.529

-0.576

-0.703
Variography analysis

**PhiT-upscaled**

**Ip-upscaled**

**Ip-seismic**
## Simulations

<table>
<thead>
<tr>
<th>Features</th>
<th>SGCS</th>
<th>BCSCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid</td>
<td>33x60x1</td>
<td></td>
</tr>
<tr>
<td>Variogram model of primary variable</td>
<td>spherical, nugget= 0.0002, structure contribution=0.0016, ranges: max.=160, med.=50, min.=1, angles: x=90, y=0, z=0</td>
<td></td>
</tr>
<tr>
<td>Dependence Model</td>
<td>Corr. coefficient -0.657</td>
<td>Bernstein copula model</td>
</tr>
<tr>
<td>Software</td>
<td>SGEMS</td>
<td>SASIM(GSLIB)</td>
</tr>
</tbody>
</table>
Simulations

Depth (m)

X-direction (m)

SGCS

BCSCS

PhIT_SGC (dec)

PhIT_sasim40000a_swap100u (dec)
Simulation Comparison

<table>
<thead>
<tr>
<th>Stat.</th>
<th>SGCS</th>
<th>BCSCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1980</td>
<td>1980</td>
</tr>
<tr>
<td>Min.</td>
<td>0.1106</td>
<td>0.121</td>
</tr>
<tr>
<td>1st Q.</td>
<td>0.2373</td>
<td>0.2375</td>
</tr>
<tr>
<td>Med.</td>
<td>0.2737</td>
<td>0.2681</td>
</tr>
<tr>
<td>Mean</td>
<td>0.2665</td>
<td>0.2662</td>
</tr>
<tr>
<td>3rd Q.</td>
<td>0.297</td>
<td>0.2969</td>
</tr>
<tr>
<td>Max</td>
<td>0.3406</td>
<td>0.3903</td>
</tr>
<tr>
<td>Var.</td>
<td>0.0019</td>
<td>0.0017</td>
</tr>
</tbody>
</table>
Simulation Comparison

**X-direction**

<table>
<thead>
<tr>
<th>Model</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>spherical</td>
<td>0.0002</td>
<td>0.0018</td>
<td>160.0000</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

**Depth**

<table>
<thead>
<tr>
<th>Model</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>spherical</td>
<td>0.0002</td>
<td>0.0018</td>
<td>50.0000</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
## Comparison

<table>
<thead>
<tr>
<th>Features</th>
<th>SGCS</th>
<th>BCSCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Parametric</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Transformation</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Computational cost</td>
<td>Lower</td>
<td>Higher</td>
</tr>
</tbody>
</table>
Final remarks

• A Bernstein copula-based spatial stochastic co-simulation (BCSCS) method presented in this paper possess several advantages over the classical sequential Gaussian co-simulation (SGCS), among others:
  • Does not require of a strong linear dependence
  • Captures and reproduces the existing dependence
  • Is non-parametric (does not need a specific distribution)
  • Reproduces the variability and the extreme values.
  • Does not need to make back transformations
Future work

• Used a linear combination of attributes (principal component and factorial analysis).
• A multivariate copula with three or more variables.
• 3D extension, but it depends on the computing power available.
• A simpler and efficient alternative, the median regression approach.
Acknowledgments

• The present work was supported by the IMP project D.61037 “Interpretación Sísmica Cuantitativa Guiada por Litofacies para la Caracterización de Yacimientos”.
Thank you for your attention