Journal of Petroleum Science and Engineering 90–91 (2012) 112–123

Contents lists available at SciVerse ScienceDirect

Journal of Petroleum Science and Engineering

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

A joint stochastic simulation method using the Bernstein copula as a flexible tool for modeling nonlinear dependence structures between petrophysical properties

V. Hernández-Maldonado [a,*], M. Díaz-Viera [a], A. Erdely [b]

[a] Instituto Mexicano del Petróleo, Eje Central Lázaro Cárdenas No. 152, Col. San Bartolo Atepehuacan, México D.F. 07730, Mexico
[b] Universidad Nacional Autonóma de México, FES Acatlán, Mexico

ARTICLE INFO

Article history:
Received 26 June 2011
Accepted 25 April 2012
Available online 12 May 2012

Keywords:
porosity
permeability
dependence
Bernstein copula
nonparametric estimation
geostatistical simulation

ABSTRACT

The statistical dependence between petrophysical properties (porosity, permeability, water saturation, etc.) in heterogeneous formations is usually nonlinear and complex; therefore, traditional statistical techniques based on assumptions of linearity are not appropriate for modeling these dependence relationships. Also, these methods may not reproduce the extreme values and data variability, which may represent impermeable barriers or high permeability zones. A modern way to model the petrophysical dependence structure between random variables is using copulas.

Copula functions have been previously applied to this kind of problems, but it seems to be very restrictive that a single copula family be flexible enough to model the nonlinear dependence structure between petrophysical properties in highly heterogeneous porous media. For this reason, in this work we have resorted to a nonparametric approach, where the Bernstein copula is used to model the empirical petrophysical relationship without imposing any distributional constraint.

The copula based stochastic method proposed here, basically consists on applying the simulated annealing method with a joint probability distribution model estimated by a nonparametric Bernstein copula. This approach has several advantages, among others we can mention that does not require the assumption of normality or other probability distribution, and is not restricted to the case of linear dependence between the variables. The proposed method provides a very flexible tool to model the complex dependence relationships between pairs of petrophysical properties. It is shown a case study where this tool is applied to model the permeability–porosity nonlinear relationship in carbonate double-porosity formations with complex microstructure of pore. It is discussed a comparative study between methods already established and the proposed one.

1. Introduction

Currently, the integrated reservoir modeling is the most accepted way to perform a reservoir study. In a simple manner it consists of modeling, from a geological and petrophysical point of view, the spatial distribution of facies and petrophysical properties (Cosentino, 2001). In integrated reservoir modeling, permeability estimation is normally a common objective. A way to perform this task is establishing dependence models between petrophysical properties such as the porosity–permeability relationship. Since it is difficult to have direct data about permeability, we use the porosity information to estimate it (Landa et al., 1996). In particular, in the context of integrated reservoir modeling, we propose a stochastic method to estimate permeability, at well log scale, using the available information of porosity.

In carbonate double-porosity formations with complex microstructure of pore space the permeability prediction becomes more difficult to estimate because it usually depends on classes of porosity, such as vugular and fracture porosity (secondary porosity). Even more, in such cases permeability is directly related to the connectivity structure of the pore system. This fact makes permeability prediction a challenging task.

By far the most used permeability predictor is the porosity–permeability relationship (Balan et al., 1995). It has long been assumed that most reservoir rocks show a reasonably linear relationship between these parameters in a semi-log scale, which allows for the estimation of permeability when a porosity profile is available. This normally requires a calibration data set that is represented by one or more key wells where comprehensive information is available in terms of core and log data. This calibration data set is used to build the predictor and to test the reliability of the results.

The regression approach tries to predict a conditional average, or expectation of permeability, corresponding to a given set of

* Corresponding author. Tel.: +52 5591756473.
E-mail address: vmhernann@yahoo.com.mx (V. Hernández-Maldonado).

0920-4105/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.petrol.2012.04.018

parameters. A different predictive equation must be established for each new area or new field. The main drawback of traditional regression methods is that the complex variability of data may not be effectively captured just in terms of variance or standard deviation and therefore the predicted permeability profile will be ineffective in reproducing extreme values. This situation is critical because, from a fluid flow point of view, these values may represent either high permeability zones or impermeable barriers.

Neural networks provide an alternative to the traditional techniques of estimating permeability. The method has been introduced in the last years, following the widespread availability of powerful computing resources and has rapidly found a number of applications. In contrast to any other estimation method, neural networks do not make use of a predefined relationship, since the estimation function is built through experience during the training phase (Rogers et al., 1995). In this respect, neural networks are model-free estimators.

An interesting point is that the predicted permeability distribution does not obey usual statistical rules, e.g., the preservation of the mean value. In fact, one of the drawbacks of the predictions made through regressions, i.e., the smoothing effect and the loss of the extreme values, is not a concern in the case of neural networks. In fact, the technique allows the actual variability of the data to be preserved (Cosentino, 2001).

Neural networks have some disadvantages, too. First, the training process has to be done with caution and can be a lengthy process, which requires a good calibration data set. The good results obtained by this technique are reached utilising a comprehensive training data set, which is not always available in real cases. On the other hand, failing in correctly calibrating the network may result in aberrant results. Another point to take into consideration is that the methodology is not yet an '*off the shelf*' application and requires expertise by the geoscientist (Cosentino, 2001).

The idea of constructing numerical models of the reservoir that honor all available data (core measurements, well logs, seismic and geological interpretations, etc.) having sparse knowledge of rock properties leads us to consider the stochastic simulation approach (Deutsch, 1992). This is not a new concept (Haldorsen and Damsleth, 1990; Journel and Alabert, 1990), stochastic models of physical systems are used extensively in many disciplines.

Stochastic simulation is the process of building alternate, equally probable models of the spatial distribution of the unsampled values. The simulation is said to be "conditional" if the resulting realizations honor the hard data values at their locations. The most straightforward algorithm for generating realizations of a multivariate Gaussian field is provided by the Sequential Gaussian Simulation (SGS) and it is extensively used to perform permeability simulations (Holden et al., 1995), another algorithm used to make permeability predictions is the Sequential Indicator Simulation. The sequential Gaussian simulation (SGS) and sequential indicator simulation (SIS) algorithms are state-of-the-art simulation methods that have contributed significantly to reservoir modeling and risk-qualified decision making. Nevertheless, these methods and their enhancements (Journel and Zhu, 1990; Suro-Perez, 1992) are limited to cases when the spatial continuity is characterized by stationary two-point statistics and to data that is defined on the same support[1] (Deutsch, 1992).

Deutsch and Cockerham (1994) used the simulated annealing technique to specify the dependence relationship between petrophysical properties and its spatial structure by minimizing a Multivariate Objective Function (MOF). The terms of the MOF to model the dependence structure are two individual marginals, which describe the empirical distribution of each petrophysical property; a correlation coefficient, which makes a linear description of the relationship, and a conditional distribution. To describe the spatial distribution, a semivariogram model of the permeability is used. The terms of the Multivariate Objective Function describe the dependence structure and spatial distribution of the petrophysical properties, however, the application of this approach do not show a natural representation of them.

A modification of the above methodology was given by Díaz-Viera and Casar-González (2005). Here, it was proposed the use of bivariate t-copula to construct the joint distribution function rather than using the joint distribution function of the sample. In this case, the structure relationship between petrophysical properties was specified by dependence measures such as Kendall's τ and Spearman's ρ (Díaz-Viera and Casar-González, 2005). This methodology was applied to simulate permeability from a porosity profile in a double porosity carbonate systems restricted to one-dimensional case at well-log scale (Díaz-Viera et al., 2006).

One of the main advantages of Díaz-Viera and Casar-González proposal is that it is used a modern statistical tool (the t-copula) capable to catch and reproduce the most important statistical characteristics of each variable (mean, median, variance, standard deviation, etc.), and their joint behavior as well. Hence, a copula function is ideal to model the complex dependence pattern that frequently appears between petrophysical properties. Another advantage is that the MOF of simulated annealing technique is considerably reduced, because each univariate marginal distribution is reproduced automatically, therefore it is no longer necessary to include them into the MOF. Moreover, as a consequence of this, the MOF becomes lighter and computationally faster.

While Díaz-Viera and Casar-González proposal can reproduce many features of the observed data, including their extreme values, the critical problem of this approach is that the t-copula is a parametric type copula, i.e., it is based on a given distribution function, the *student t*. Consequently, expecting a single copula family to be capable to model any kind of bivariate dependency seems to be too restrictive, at least for the petrophysical properties under consideration.

It is worth mentioning that the copula approach has been successfully used to model dependence patterns in other areas of oil industry, for example, for field development decision process (De Melo e Silva Accioly and Yassuo Chiyoshi, 2004); or to model dependence in petroleum decision making (Al-harthy et al., 2005). In this context, copulas have become the new way of modeling dependence structure between variables.

In this work we continue using copulas, extending their application to the nonparametric case. In other words, we do not impose a priori a parametric joint distribution function to characterize the behavior of the sample data, instead, we use a copula function to model the intrinsic distribution of data values through a nonparametric approach, using the Bernstein copula (Sancetta and Satchell, 2004; Sancetta, 2007).

The Bernstein copula is a function that fits its joint distribution behavior to the joint empirical distribution, consequently, it does not need to specify the porosity–permeability relationship using dependence measures such as Kendall's τ and Spearman's ρ. As a consequence, the simulated values of permeability are reproduced in a much more natural way and this occurs because we do not impose any dependence relationship, instead, we allow the data to 'speak'. The simulated annealing technique is used to model the spatial distribution of permeability. Unlike previous cases the Multivariate Objective Function has only two terms, one for modeling the dependence structure between petrophysical properties, and the other one for modeling the spatial distribution. In this way the MOF

[1] Data of different supports may only be considered if the averaging is linear, which is not the case for permeability.

is lighter and hence it becomes computationally faster. Another advantage is that this methodology is not limited to petrophysical properties modeling, and can be used in many areas of geostatistics.

In next sections we will make a brief introduction to copulas; next, we will model porosity–permeability relationship using the Bernstein copula; then, we will show the two steps method for spatial permeability modeling, and we will conclude this part with a detailed case study. To finish, we perform a comparative study of three stochastic simulation techniques (our proposal included).

2. A short introduction to copulas

The basic idea of the method proposed here consists of using a stochastic simulation approach for modeling permeability based on the permeability–porosity observed dependence patterns. To explain it, we will give a very short introduction to copulas.

According to Sklar's (1959) Theorem, the underlying *bivariate copula* associated to a bivariate random vector (X,Y) represents a functional link between the joint probability distribution H and the univariate marginal distributions F and G, respectively:

$$H(x,y) = C(F(x),G(y)) \tag{1}$$

for all x,y in the extended real numbers system, where $C: [0,1]^2 \to [0,1]$ is unique whenever X and Y are continuous random variables. Therefore, all the information about the dependence between continuous random variables is contained in their corresponding copula. Several properties may be derived for copulas (Nelsen, 2006), and among them we have an immediate corollary from Sklar's Theorem: X and Y are independent continuous random variables if and only if their underlying copula is $\Pi(u,v) = uv$.

Let $\mathcal{S} := \{(x_1,y_1),\ldots,(x_n,y_n)\}$ be observations of a random vector (X,Y). We may obtain empirical estimates for the marginal distributions of X and Y by means of

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}\{x_k \le x\}; \quad G_n(y) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}\{y_k \le y\} \tag{2}$$

where $\mathbb{I}$ stands for an indicator function which takes a value equal to 1 whenever its argument is true, and 0 otherwise. It is well-known (Billingsley, 1995) that the empirical distribution F_n is a consistent estimator of F, and that, $F_n(t)$ converges almost surely to $F(t)$ as $n \to \infty$, for all t.

Similarly, we have the *empirical copula* (Deheuvels, 1979), a function C_n with domain $\{i/n : i = 0,1,\ldots,n\}^2$ defined as

$$C_n\left(\frac{i}{n},\frac{j}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}\{rank(x_k) \le i, rank(y_k) \le j\} \tag{3}$$

and its convergence to the true copula C has also been proved (Fermanian et al., 2004). The empirical copula is not a copula, since it is only defined on a finite grid, not in the whole unit square $[0,1]^2$, but by Sklar's (1959) Theorem C_n may be extended to a copula.

3. Porosity–permeability data modeling using Bernstein copula

We model porosity as an absolutely continuous random variable X with unknown marginal distribution function F, and permeability as an absolutely continuous random variable Y with unknown marginal distribution function G. Having bivariate observations from the random vector (X,Y), for simulation of continuous random variables, the use of the empirical distribution function estimates (2) is not appropriate since F_n is a step function, and therefore discontinuous, so a smoothing technique is needed. Since our main goal is simulation of porosity–permeability, it will be better to have a smooth estimation of the marginal quantile function $Q(u) = F^{-1}(u) = \inf\{x : F(x) \ge u\}$, $0 \le u \le 1$, which is possible by means of Bernstein polynomials as in (Muñoz-Pérez and Fernández-Palacín, 1987):

$$\tilde{Q}_n(u) = \sum_{k=0}^{n} \frac{1}{2}(x_k + x_{k+1})\binom{n}{k} u^k (1-u)^{n-k} \tag{4}$$

and the analogous case for marginal G in terms of values y_k. For a smooth estimation of the underlying copula we make use of the Bernstein copula (Sancetta and Satchell, 2004; Sancetta, 2007):

$$\tilde{C}_n(u,v) = \sum_{i=1}^{n}\sum_{j=1}^{n} C_n\left(\frac{i}{n},\frac{j}{n}\right)\binom{n}{i} u^i (1-u)^{n-i}\binom{n}{j} v^j (1-v)^{n-j} \tag{5}$$

for every (u,v) in the unit square $[0,1]^2$, and where C_n is as defined in (3).

3.1. Nonparametric simulation algorithm

In order to simulate replications from the random vector (X,Y) with the dependence structure estimated from the observed data $\mathcal{S} := \{(x_1,y_1),\ldots,(x_n,y_n)\}$, according to a result in Nelsen (2006) 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

$$C_u(v) = \frac{\partial \tilde{C}(u,v)}{\partial u} \tag{6}$$

and $\tilde{C}$ is obtained by (5)
(3) The desired pair is $(x,y) = (\tilde{Q}_n(u),\tilde{R}_n(v))$, where $\tilde{Q}_n$ and $\tilde{R}_n$, are the estimated and smoothed quantile functions of X and Y, respectively, according to (4).

For details about this section see Erdely and Díaz-Viera (2009).

4. Permeability modeling by nonconditional stochastic simulation method using the Bernstein copula

Basically the methodology can be described as a two-stage algorithm. In the first one, the empirical porosity–permeability bivariate distribution is reproduced by a nonparametric copula modeling, using the Bernstein copula. The aim is to reproduce the joint dependence pattern between these two petrophysical properties. In the second one, a geostatistical simulation of permeability is performed using simulated annealing technique, whose objective function is the variogram model (Deutsch and Journel, 1998).

A description of each stage of the algorithm is described below:

(1) Petrophysical properties dependence modeling using Bernstein copula.
(2) Stochastic spatial simulation of permeability profile, using porosity as secondary variable. The simulated annealing algorithm involves the following:
   (a) Generation of an initial spatial configuration of permeability, in stochastic simulation this configuration is known as '*seed*'. It means that a simulation must initially have a set of permeability values to begin with. It is not necessary these values fulfill the variogram model, they will be provided by the Bernstein copula.

(b) Definition of the MOF (Multivariate Objective Function): For this case we only need to define the variogram model, hence, we have an univariate objective function.
(c) Measuring the energy of the *seed*: In a simplistic way, this step gives us an idea of how far this configuration is from the objective function.
(d) Calculating the simulated annealing starting parameters (initial temperature (t_0) and cooling schedule) based on Dreo's procedure (Dréo et al., 2006).
(e) Performing the simulation: It finalizes when the objective function error is reached, or an accumulation of three stages without change occurs, or when the maximum number of attempted perturbations is completed.

In the next section, we will show the application of this proposal to a case study. First, a short description of the data set will be done; then, each step of the algorithm above described will be applied; finally, results will be discussed.

4.1. A case study

We model the relationship between permeability and porosity of the double porosity carbonate formations of a South Florida Aquifer in the western Hillsboro Basin of Palm Beach County, Florida (Fig. 1). The characterization of this aquifer for the borehole and field scales is given by Parra et al. (2001) and Parra and Hackert (2002) and a hydrogeological situation is described by Bennett et al. (2002). The interpretation of the borehole data and determination of the matrix and secondary porosity and secondary-pore types (shapes of spheroids approximating secondary pores) were presented by Kazatchenko et al. (2006), where to determine the pore microstructure of aquifer carbonate formations the authors applied the petrophysical inversion technique that consists in minimizing a cost function that includes the sum of weighted square differences between the experimentally measured and theoretically calculated logs (Kazatchenko et al., 2004).

We used the results of inversion obtained by Kazatchenko et al. (2006) for carbonate formations of South Florida Aquifer that includes the following petrophysical characteristics: matrix porosity, secondary vuggy and crack porosities. It should be noted that the secondary porosity system of this formation has complex microstructure and corresponds to a model with two types of pore shapes: cracks (flattered ellipsoids) with the overall porosity of 2% and vugs (close to sphere) with the porosity variations in the range of 10–30%. Such a secondary porosity model can be interpreted as the interconnected by microfractures and channels vuggy formation.

Details of the statistical properties analysis of this results are largely explained by (Díaz-Viera et al., 2006). Based on this study we determine that the highest observed dependence between the petrophysical properties corresponds to vuggy porosity (PHIV) and permeability (K).

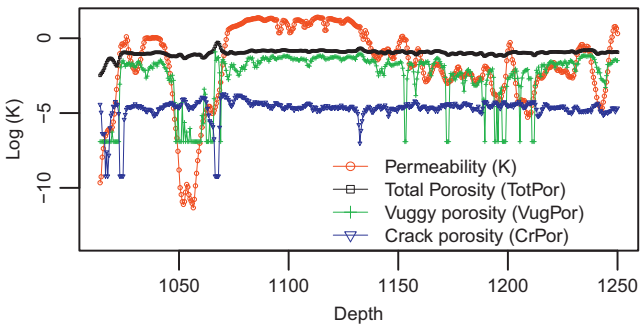

Fig. 1. Crack porosity (PHICR), vuggy porosity (PHIV), total porosity (PHITOT) and permeability (K).

4.1.1. Porosity–permeability data modeling: generating the initial configuration

Simulated Annealing (SA) has proved its effectiveness in various fields such as the design of electronic circuits, image processing, collection of the household garbage, or the organization of the data-processing network (Dréo et al., 2006). The simulated annealing method is a global optimization algorithm which stochastically finds one of the best solutions. To start, the method needs an initial configuration, in other words, needs a possible solution which is not necessarily the best, this configuration is known as *seed*. At this point the method randomly selects a single value of the initial configuration and proposes a new one, this process is known as elemental perturbation. Then, it asks if this new configuration is closer to a good solution, if it is true, it keeps the new configuration and randomly selects another value to do the same; if it is false, it accepts or rejects the new configuration by a probability, determined by the Boltzmann distribution. The simulated annealing method repeats this task many times in order to reach a good solutions.

In our case, to perform a Simulated Annealig Realization (SAR) it is necessary to propose an initial configuration (or '*seed*') of the primary variable, is this case the permeability. The Bernstein copula is used to generate it through the nonparametric simulation algorithm (presented before). Even more, this algorithm is used to generate each new permeability value when the SAR is performed.

In Fig. 2 it is shown a scatter-plot distribution and histograms of porosity–permeability real data (sample size of $n=380$) taken from Kazatchenko et al. (2006). Using the nonparametric simulation algorithm, we generate a single porosity–permeability data set of the same size of observations (Fig. 3); this nonparametric simulation will be considered as the initial configuration, or seed. It should be noticed that the Bernstein copula reproduces very well the marginal distribution and the joint distribution of porosity and permeability. In fact, this is the reason why we do

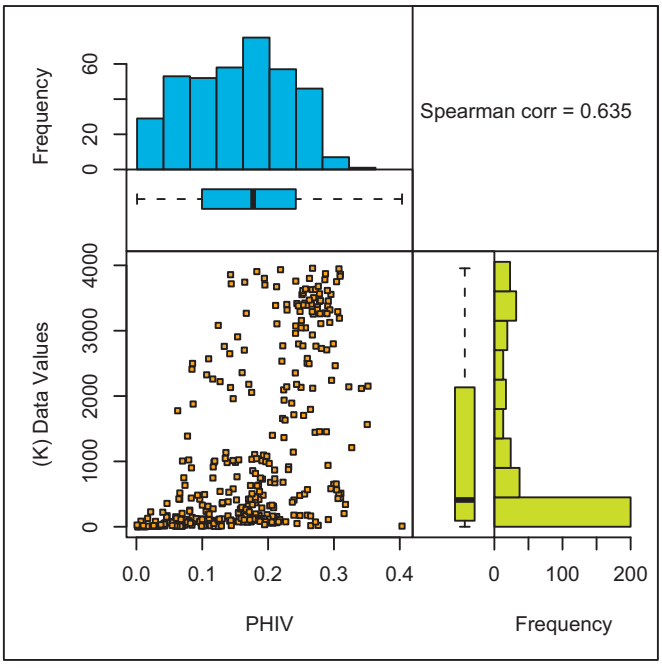

Fig. 2. Scatter-plot distribution and histograms of porosity and permeability of real data values.

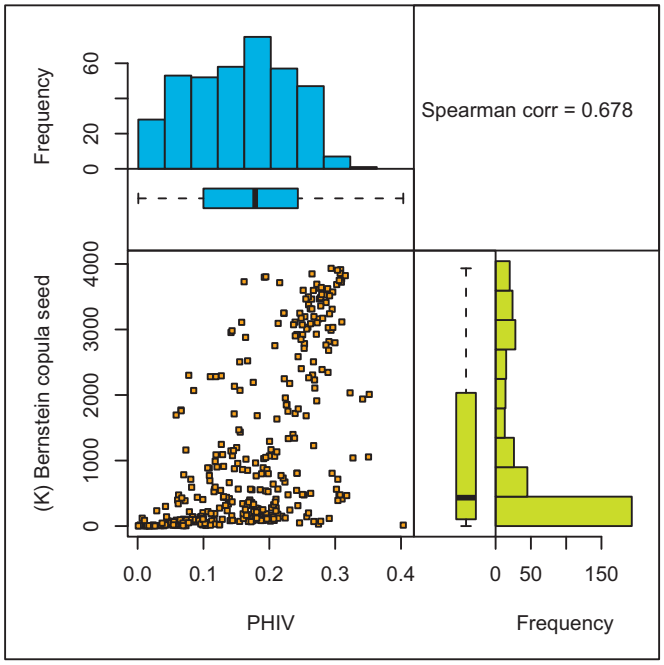

Fig. 3. Scatter-plot distribution and histograms of porosity and permeability of initial configuration (*seed*).

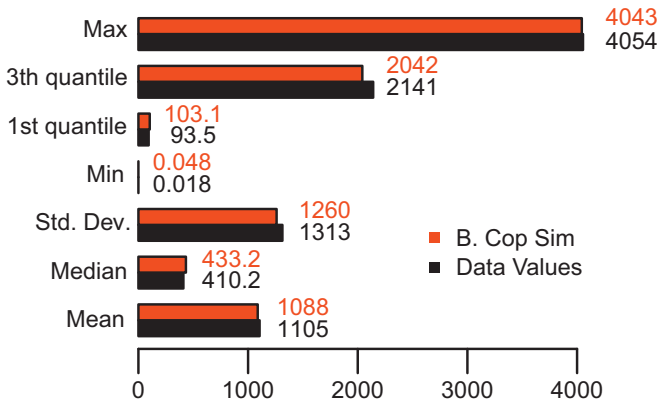

Fig. 4. Comparative plot of some statistics to real data and simulated *K* values using the Bernstein copula.

not need any dependence measures or marginal distributions as Díaz-Viera and Casar-González (2005) and Deutsch and Cockerham (1994) methods, respectively, do.

In addition, the Bernstein copula can reproduce the variability and the extreme values of the original data. Both configurations (data values and seed) show very similar statistics of permeability (Fig. 4), scatter-plot distribution and marginal distributions of each petrophysical property.

Finally in Fig. 5 it is shown the spatial distribution of permeability (red) versus a single nonconditional Bernstein copula simulation, using PHIV as secondary variable (blue). The dependence structure between porosity and permeability is very well reproduced, but the spatial structure has a lot of variability.

4.1.2. *Defining the objective function*

The simulated annealing technique runs many, often millions of perturbations in order to achieve an acceptable realization, and that implies a high computational effort. Therefore, it is not recommended

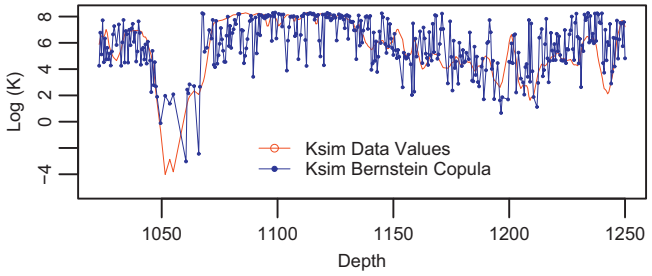

Fig. 5. Spatial distribution of the real data permeability and a single nonconditional Bernstein copula simulation, using PHIV as secondary variable. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

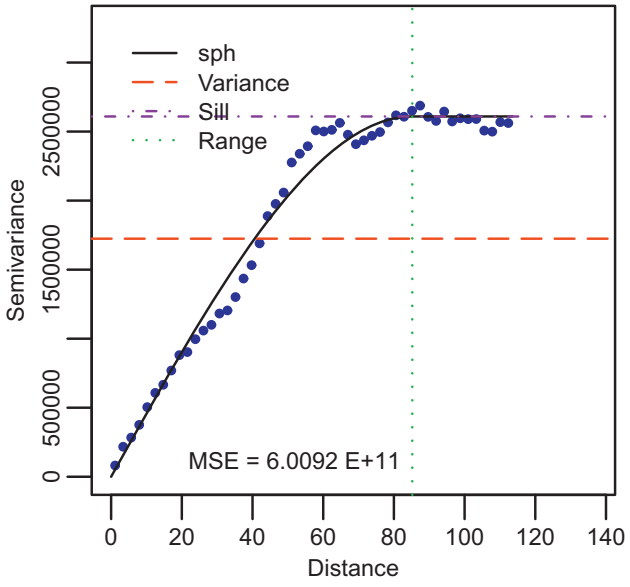

Fig. 6. Estimated variogram for *K* (real data values), and its fitted model.

having too many components into the objective function, and also, each one of them should be reasonably easy to compute (Deutsch and Journel, 1998). Since Bernstein copula allows us to model the dependence pattern between permeability and porosity, we can propose an univariate objective function to satisfy the spatial distribution, then, the objective function is the variogram model (Deutsch and Journel, 1998).

$$FO = \sum_i \frac{[\gamma^*(h_i) - \gamma(h_i)]^2}{\gamma(h_i)^2} \tag{7}$$

where

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \tag{8}$$

4.1.3. *Measuring the energy of the initial configuration*

According to the objective function, we use Eq. (7) to obtain the energy of the seed, $E_i = 159.63$. This energy represents or reflects how far we are from a good solution. In Figs. 6 and 7 we plot the variogram model of each configuration, data values and seed, respectively. In the initial configuration we have a great nugget effect; it means that we have high small-scale variability. RGEOESTAD program was used to analyze these results (Díaz-Viera et al., 2010).

In Table 1, we present the variogram models of each configuration.

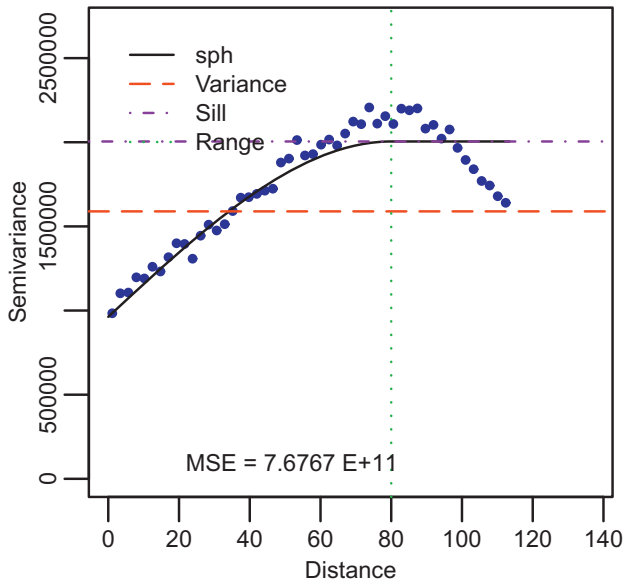

Fig. 7. Estimated variogram for *K* (initial configuration or seed configuration) and its fitted model.

Table 1
Variograms models of the data values and initial configuration (*seed*).

Configuration	Nugget	Sill	Range
Data values	0.00	261,1621.52	85.07
Seed	962,645.81	2,004,671.28	80.00

4.1.4. Calculating the initial temperature and the most suitable annealing schedule

To calculate the annealing schedule the procedure proposed by Dréo et al. (2006) is followed. The initial temperature is calculated based on the following equation:

$$T_0 = \left(-\frac{\Delta E}{\log(\tau_0)}\right) \tag{9}$$

where $\Delta E = E_{perturb} - E_{initial}$ is the mean of the energy differences between 100 perturbed configurations and the *seed*. The perturbation mechanism consists in choosing a given value of porosity and generating a new permeability value using the Bernstein copula algorithm (see Section 3.1).

Using Eq. (9) and considering an acceptance rate $\tau_0 = 0.5$ (Dréo et al., 2006), the initial temperature is

$T_0 = 1.14$

Annealing schedule:

(1) $12*n$ accepted perturbations $= 12*380 = 4500$
(2) $100*n$ attempted perturbations $= 100*380 = 38\,000$

where n is the data number.

The decreasing temperature values are calculated by geometric law (10):

$$T_{k+1} = 0.8*T_k \tag{10}$$

4.1.5. Results

We have all necessary elements to perform simulated annealing realizations, i.e., initial configuration, objective function, initial temperature, cooling program schedule and a new permeability value generator (the Bernstein copula). Now it is the moment to execute realizations, we choose to make a single nonconditional simulation and a median of 10 nonconditional simulations, both will be compared to the original data values. It must be clarified that a nonconditional simulation means that it is a simulation where any permeability value is fixed.

The single nonconditional simulation shows a very good agreement in terms of histograms, variogram and scatter-plot reproduction (Fig. 8). Also its variogram and its fitted model show very good agreement with respect to real data (Fig. 9).

As we could anticipate, the dependence structure between these two petrophysical properties is well represented, the extreme values, and in general, all of the most important statistics are quite well reproduced (Fig. 10).

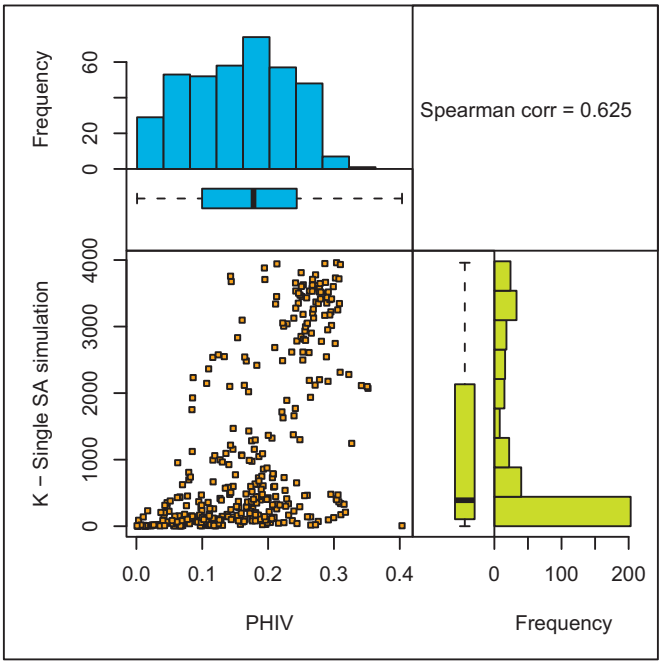

Fig. 8. Scatter-plot distribution and histograms of PHIV and *K* simulated values (a single simulation).

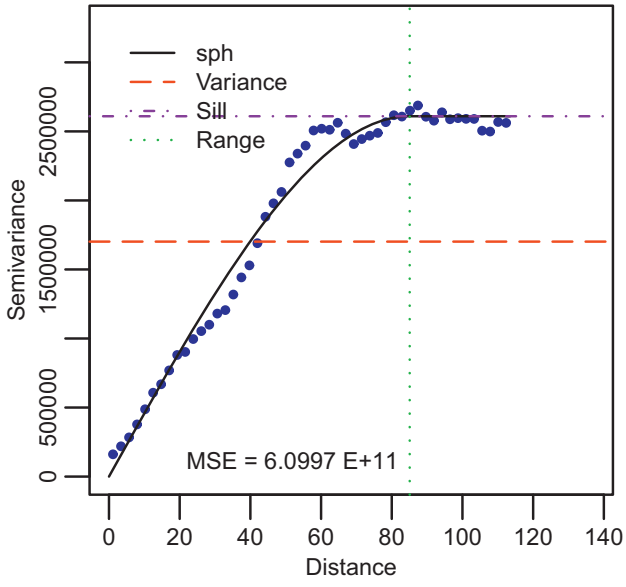

Fig. 9. Estimated variogram for *K* (a single simulation), and its fitted model.

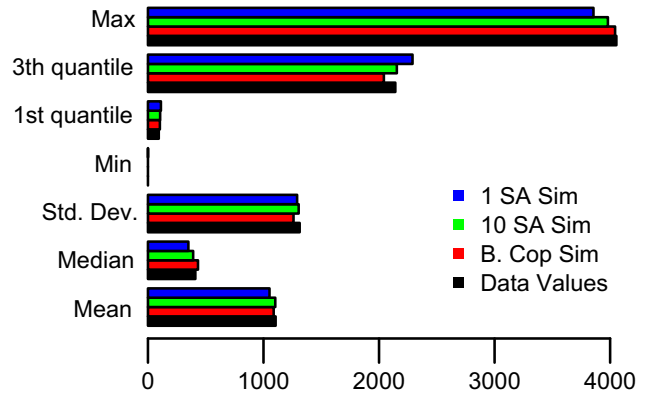

Fig. 10. Statistical comparison of to the original data, seed, a single SA simulation, and the median of 10 SA simulations.

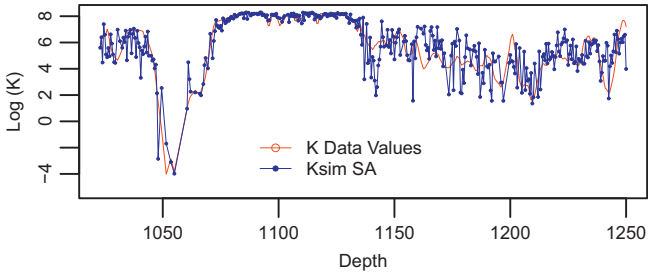

Fig. 11. Spatial distribution of permeability in a single nonconditional SA simulation of *K*, with bivariate Bernstein copula, using PHIV as secondary variable.

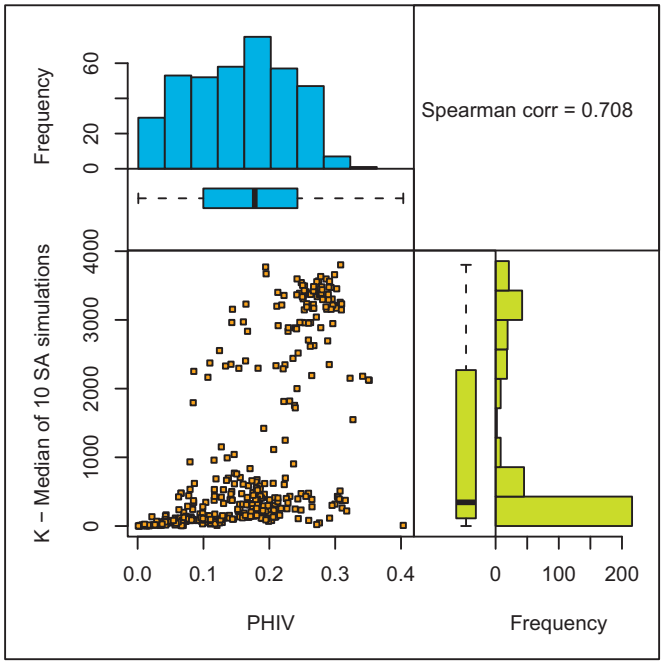

Fig. 12. Scatter-plot distribution and histograms of PHIV and *K* simulated values (median of 10 simulations).

The spatial distribution of the permeability follows the same general behavior of the original one (Fig. 11). However, there is still a small-scale variability which does not exist in the original data values; but certainly, it is smaller than the initial configuration (Fig. 5).

The median of 10 nonconditional simulated annealing realizations was calculated in order to reduce small-scale variability (Fig. 14). Histograms and scatter-plot reproduction also show a very good agreement with respect to real data (Fig. 12), as well as the variogram and its fitted model (Fig. 13).

Table 2 is a summary of the variogram models. Observe that these values are very similar between each data set, except, obviously, the seed values.

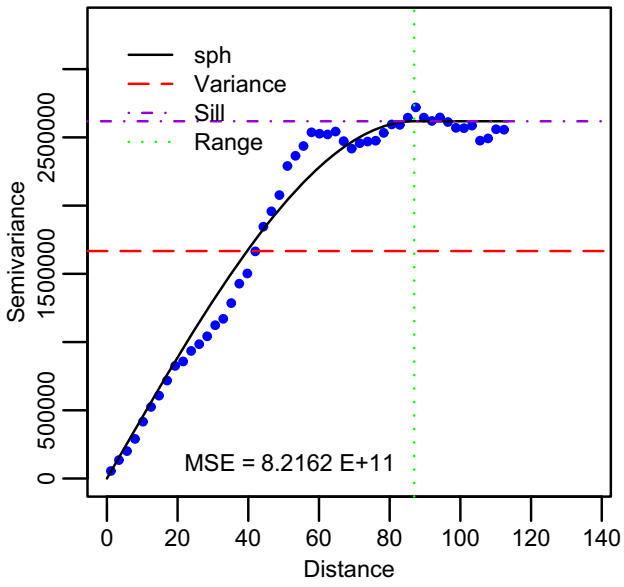

Fig. 13. Estimated variogram for *K* (a median of 10 simulations), and its fitted model.

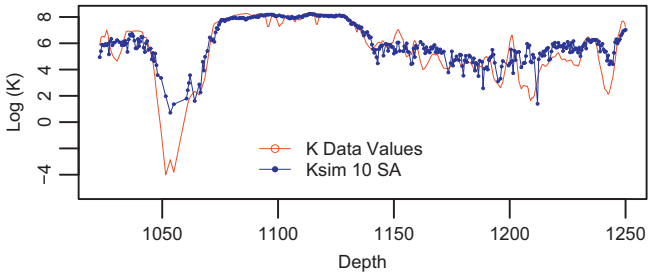

Fig. 14. Spatial distribution of permeability in a median of 10 nonconditional SA simulations of *K*, with bivarite Bernstein copula, using PHIV as secondary variable.

Table 2
Variograms models of the data set, seed, single nonconditional SA simulation and the median of 10 SA nonconditional simulations.

Configuration	Nugget	Sill	Range
Data values	0.00	2.611×10^6	85.07
Seed	962,645.81	2.004×10^6	80.00
K SA 1	0.00	2.612×10^6	84.99
K SA median	0.00	2.620×10^6	86.84

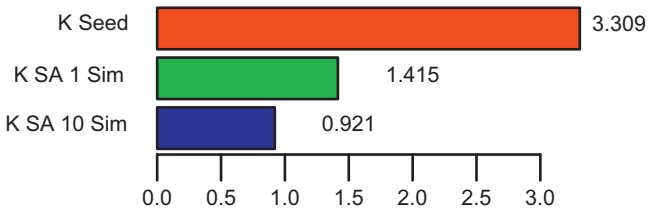

Fig. 15. Mean Square Error (MSE) comparison between the initial configuration (seed), a single SA simulation, and the median of 10 SA simulations.

Finally, Fig. 15 shows a comparison chart in terms of mean square error (MSE). The data set values that has the greatest MSE is the seed (3.309). The simulated annealing realizations using bivariate Bernstein copula show very low MSE values (1.415 for single simulation and 0.921 for median of 10 simulations).

5. Comparison of three different spatial stochastic methods

We have proposed the use of Bernstein copula as an alternative to traditional methods based on linear assumptions like regressions or semi-logarithmic transformations. The reason is because these methods eventually end up giving results that are far away from the real data distribution. Also we said that a parametric copula approach can represent very well the relationship between porosity and permeability, but due to their complex dependence structure, the data values may not to be naturally estimated, since it is hard to fit the dependence structure between these two petrophysical properties with a single model.

The purpose of this section is to perform a comparative and quantitative study between three different methods to simulate the spatial distribution of permeability (K versus Depth), using as a covariate the porosity profile (PHIV). It will be shown that this proposal is a feasible and consistent tool to model the spatial structure of petrophysical properties using their dependence relationship.

5.1. Stochastic simulation methods

The following three methods are used to perform the permeability simulations:

(1) SASIM (Deutsch and Journel, 1998): SASIM is a program that belongs to GS library (GSLIB) that uses a stochastic simulation method based on simulated annealing technique. It has a MOF consisting in: two individual histograms, a correlation coefficient, a conditional distribution and a semivariogram model (Eqs. (11)–(14)).
(2) t-copula (Díaz-Viera and Casar-González, 2005): stochastic simulation method that uses simulated annealing technique and t-copulas approach to model the dependence structure between petrophysical properties. It has a MOF consisting in a correlation coefficient, a semivariogram model, and conditional distribution (Eqs. (12)–(14)).
(3) The proposed method (Bivariate Bernstein Copula): stochastic simulation method that uses simulated annealing technique and *the Bernstein copula* to model the dependence structure between petrophysical properties. Here it is not assumed any probability distribution function. The objective function consists only of the semivariogram model (Eq. (12)).

Combinations of the following objective functions are used in the above methods.

A histogram or univariate distribution is one statistical measure that stochastic realizations should honor. The cumulative distribution $F^*(z)$ of the simulated realization should match the prespecified cumulative distribution $F(z)$ for some number of z values [chosen equally to discretize the reference cumulative distribution F(z)]:

$$O_1 = \sum_z [F^*(z) - F(z)]^2 \tag{11}$$

A semivariogram captures the two-point spatial variability in the realization. The semivariogram $\gamma^*(h)$ of the simulated realization should match the prespecified semivariogram $\gamma(h)$. The objective function is written:

$$O_2 = \sum_h \left[\frac{\gamma^*(h) - \gamma(h)]^2}{\gamma(h)^2} \right] \tag{12}$$

A correlation coefficient between the primary variable being simulated and a secondary variable (available at the same resolution) captures any linear correlation. The component objective function is

$$O_3 = [\rho^* - \rho]^2 \tag{13}$$

Conditional distribution between the primary variable being simulated and a secondary variable captures much more than a linear correlation coefficient. The objective function is

$$O_4 = \sum_{i=0}^{n_s} \sum_{j=0}^{n_p} [f_i^*(j) - f_i(j)]^2 \tag{14}$$

where n_s and n_p are the number of secondary and primary classes, respectively; the notation $f_i(j)$ is used for the conditional distribution of the primary variable ($j = 1, \ldots, n_p$) given that the collocated secondary variable is in class i. Note that $\sum_{j=1}^{n_p} f_i(j) = 1, \forall i$.

5.2. Describing comparison method

We compare realizations of three stochastic geostatistical methods in order to determine the accuracy of this proposal. As hard reference for comparison we have the variogram model and the Mean Squared Error (MSE). Also a graphical or qualitative comparison is done.

Finally we make a percentage comparison for each MSE value. We take as the highest error the largest MSE value, and then, with respect to this value we calculate the corresponding proportion for the other two methods.

We performed the simulations in the following order:

(1) A single nonconditional simulation and a median of 10 nonconditional simulations of permeability.
(2) A single conditional simulation and a median of 10 conditional simulations of permeability (both conditioned to 10%).
(3) A single conditional simulation and a median of 10 conditional simulations of permeability (both conditioned to 50%).
(4) A single conditional simulation and a median of 10 conditional simulations of permeability (both conditioned to 90%).

We perform the median simulation in order to decrease the small-scale variability; we strictly followed the order above described and we will discuss the results.

5.3. Results and discussion

In a single nonconditional simulation Fig. 16 shows the spatial distribution of permeability, using PHIV as secondary variable. Simulation methodologies (from top to bottom order): SASIM-GSLIb (Deutsch and Journel, 1998), t-copula (Díaz-Viera and Casar-González, 2005), bivariate Bernstein copula.

A complete table of the MSE differences between each methodology, in terms of percentage, is presented (Table 3).

It should be noticed that the methodology with the greatest MSE is SASIM method with a value of 7.76, followed by the t-copula approach with 5.55 (which represents the 71.5% of SASIM-GSLIB methodology); finally, the Bernstein copula with a value of 1.85 (which represents the 23.8% of SASIM-GSLIB methodology).

In median nonconditional simulations, Fig. 17 shows the comparison of the spatial distribution of the permeability, using PHIV as a secondary variable.

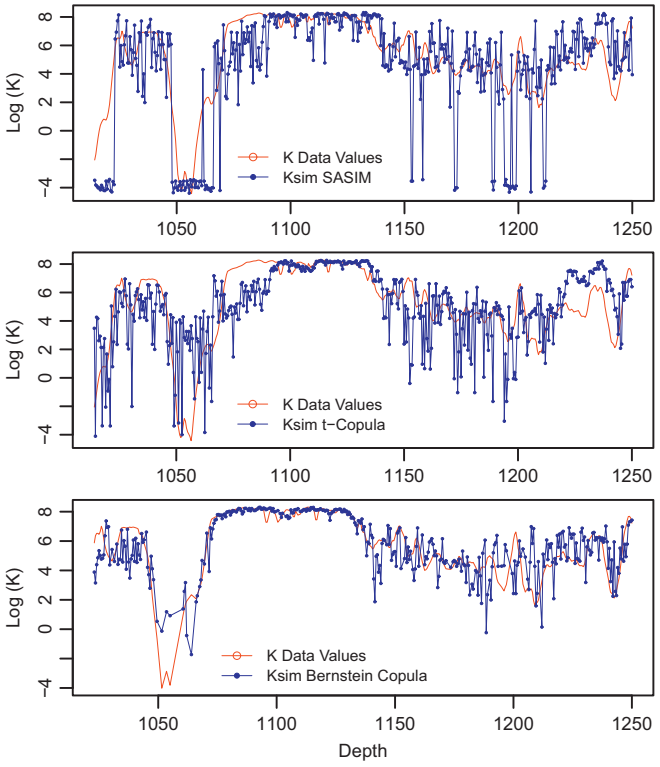

Fig. 16. Spatial distribution of permeability in a single nonconditional simulation. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

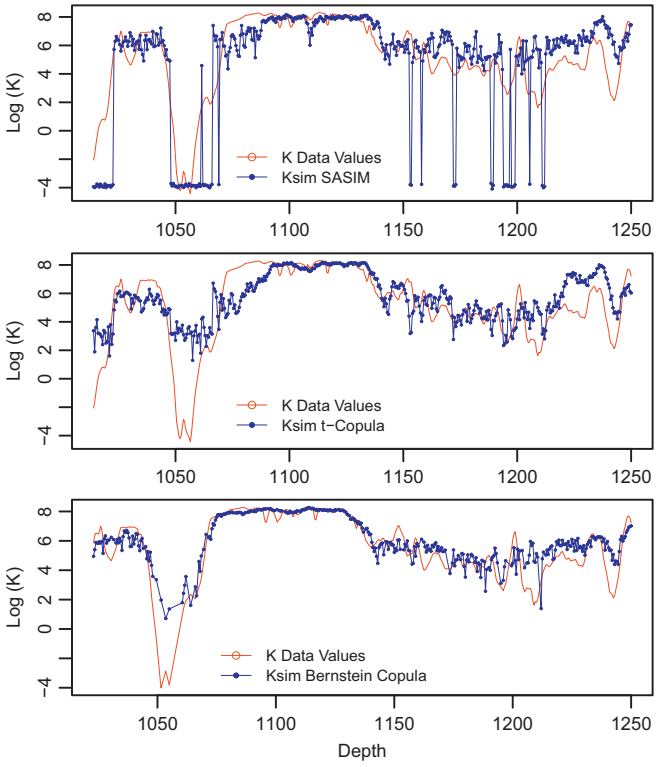

Fig. 17. Spatial distribution of permeability in a median of 10 nonconditional simulations. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

Table 3
Complete MSE table. For a single simulation.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	7.76	100	–
t-Copula	5.55	71.5	100
Bernstein copula	1.85	23.8	33.3

Table 4
Complete MSE table. For a median of 10 nonconditional simulations.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	7.00	100	–
t-Copula	3.82	54.6	100
Bernstein copula	1.17	16.7	30.6

A complete table of the MSE differences of the median of 10 nonconditional simulations between each methodology, in terms of percentage, is presented (Table 4).

In all cases the errors decreased and the low-scale variability also decreased. Once again the methodology with the greatest MSE is SASIM with a value of 7.00, followed by *t*-copula with 3.82 (which represents the 54% of SASIM-GSLIB methodology); finally, the Bernstein copula appears with a value of 1.17 (which represents the 16% respect SASIM-GSLIB methodology).

Between 1 and 10 simulations, each methodology has a MSE reduction, 9% for SASIM (from 7.76 to 7.00); 31% for the *t*-copulas (from 5.55 to 3.82); 36% for Bernstein copulas (from 1.85 to 1.17).

In a single conditional simulation, conditioned to 10%, Fig. 18 shows the spatial distribution of permeability. In this case we observe the same tendency as in the above results, i.e., the method with the greatest MSE is SASIM method and the method with the lowest error is our proposal. The MSE differences between each method are presented in Tables 5 and 6, for single and median conditional simulations, respectively.

Fig. 19 shows the comparison of the spatial distribution of the permeability of the median of 10 conditional simulations.

In simulations performed for conditioned values in higher scales (50% and 90%), we notice the tendency to have better results, but always in all cases the method presenting better results was our proposal. Fig. 20 for single realization and Fig. 21 for median of 10 realizations, for 50% of conditioning; and Figs. 22 and 23 for 90% of conditioning, respectively.

Tables 7 and 8 show the MSE results for 50% of conditioning values, single and median realizations, respectively. In Tables 9 and 10 we present the MSE results for 90% of conditioning values.

As we can see our proposal has shown the most accurate results of permeability values, for nonconditional and conditional realizations. For example, in 10% of conditioning, the methodology with the greatest MSE is SASIM-GSLIB (MSE=5.91), followed by the *t*-copula (MSE=2.90), which represents the 49% of SASIM-GSLIB method; finally, Bernstein copula (MSE=1.02) representing the 17%. The same difference (for our proposal) is viewed in 50% of conditioning (17%), but in 90% of conditioning the difference reaches to 59%. This last case is very interesting because, graphically, it is easy to appreciate how far a simulated value is from the original data. Bernstein copula generates values very close to the original data set.

Another interesting point is that the accuracy between conditioned simulations of the same method, for example, for our proposal the MSE values are 1.85, 1.42, 0.88 and 0.40, for nonconditional and conditional simulations (to 10%, 50% and 90%, respectively). The differences are not too big; this leads us to think that, at least for conditioned simulations, they are not so different. It does not happen for SASIM-GSLIB method because the differences between each conditioning simulation are very large (7.76, 6.66, 3.22 and 0.92). Hence, the results for this method will strongly depend on the quantity of the conditioning data values we have, and this is a crucial point.

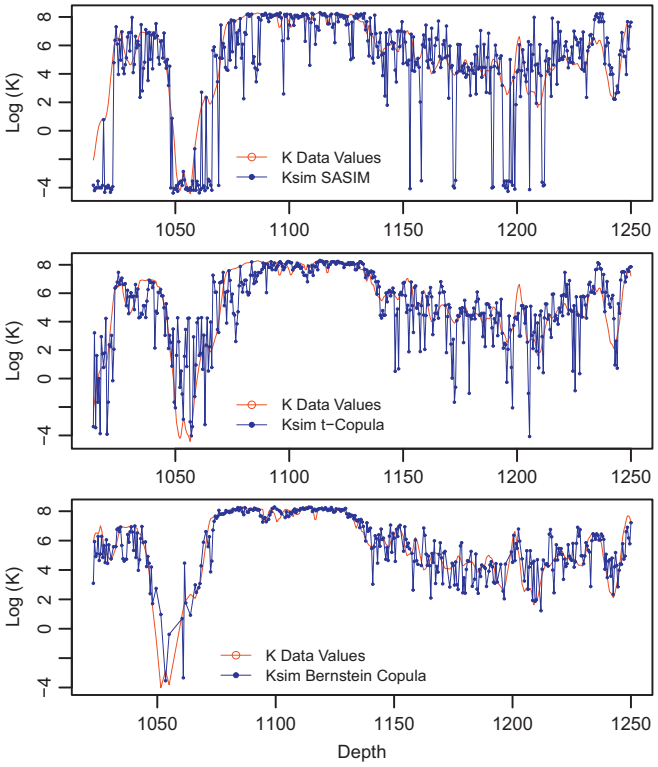

Fig. 18. Spatial distribution of permeability in a single 10% conditional simulation. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

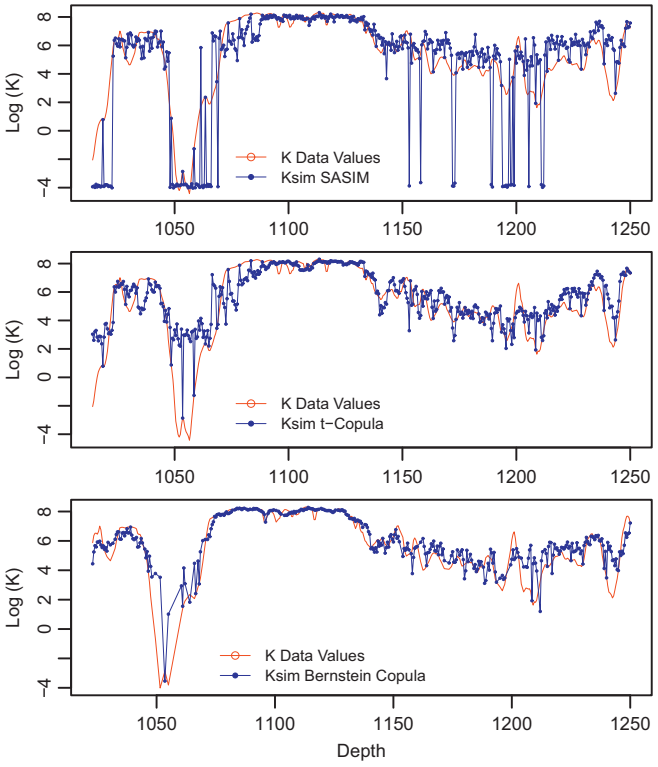

Fig. 19. Spatial distribution in a median of 10, 10% conditional simulations. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

Table 5
Complete MSE table. For a single 10% conditional simulation.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	6.66	100	–
t-Copula	3.58	53.8	100
Bernstein copula	1.42	21.3	39.7

Table 6
Complete MSE table. For a median of 10% conditional simulations.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	5.91	100	–
t-Copula	2.90	49.1	100
Bernstein copula	1.02	17.3	35.2

The above results suggest that the Bernstein copula has a promising potential in geostatistical modeling, due to its accurate results.

Finally, we compare the original spatial data for log-permeability versus a spatial median regression based on porosity (Erdely and Díaz-Viera, 2009), Fig. 24. The mean squared error of this regression is 1.83, the value for a single nonconditional simulation is 1.85, and hence, we can see that a quantile median regression, just using the Bernstein copula can compete with simulated annealing results at least for this case.

6. Conclusions

The proposed method provides a very flexible tool to model the complex dependence relationships between pairs of petrophysical properties such as porosity and permeability. It can model bivariate dependencies in a much more efficient and accurate way. Hence, it is an alternative to traditional methods like linear regression, since it does not need the assumption of linear dependence between variables.

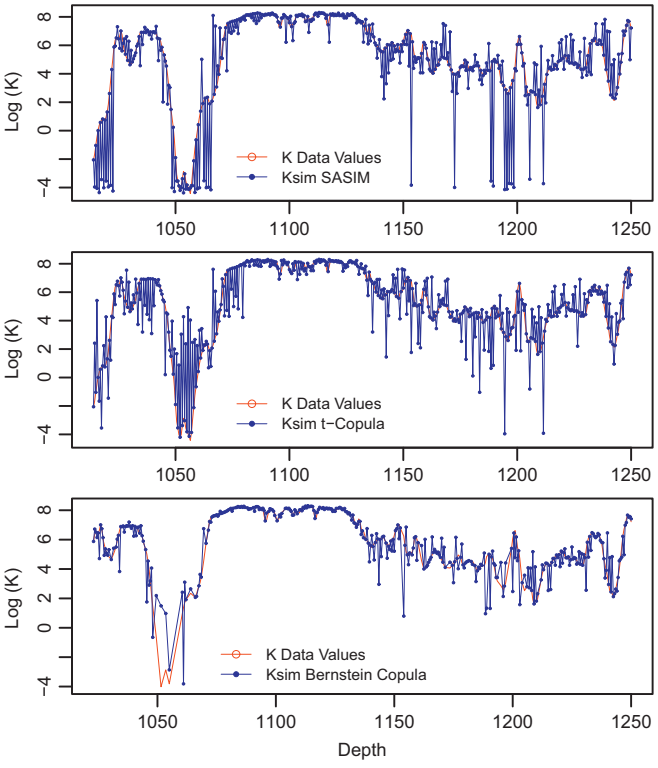

Fig. 20. Spatial distribution in a single 50% conditional simulation. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

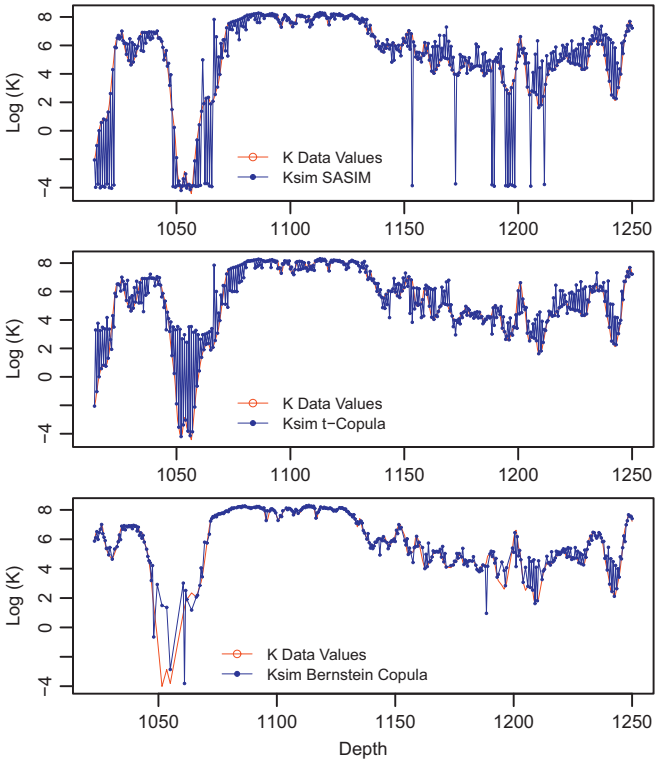

Fig. 21. Spatial distribution in a median of 10, 50% conditional simulations. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

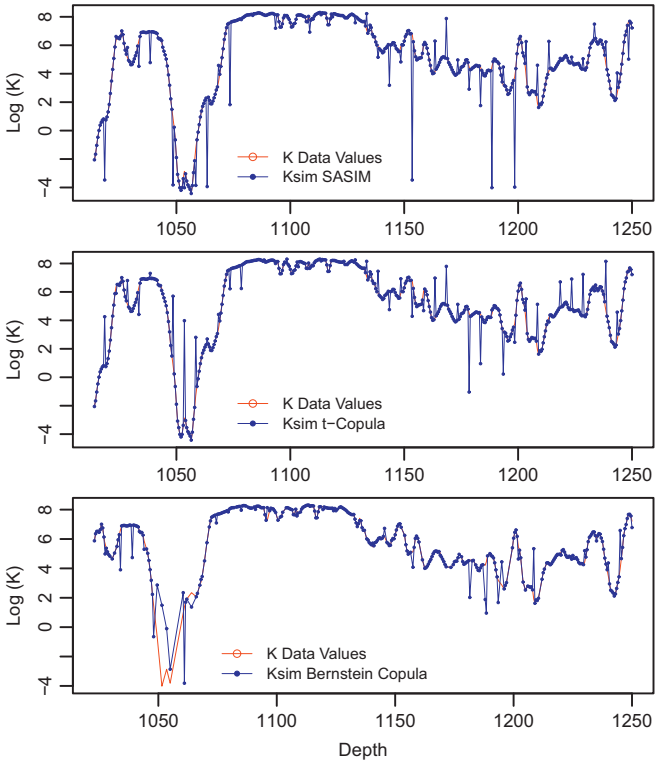

Fig. 22. Spatial distribution in a single 90% conditional simulation. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

The proposed method has three main advantages: First, an easy way to simulate bivariate data with the dependence structure and marginal behavior suggested by already observed data;

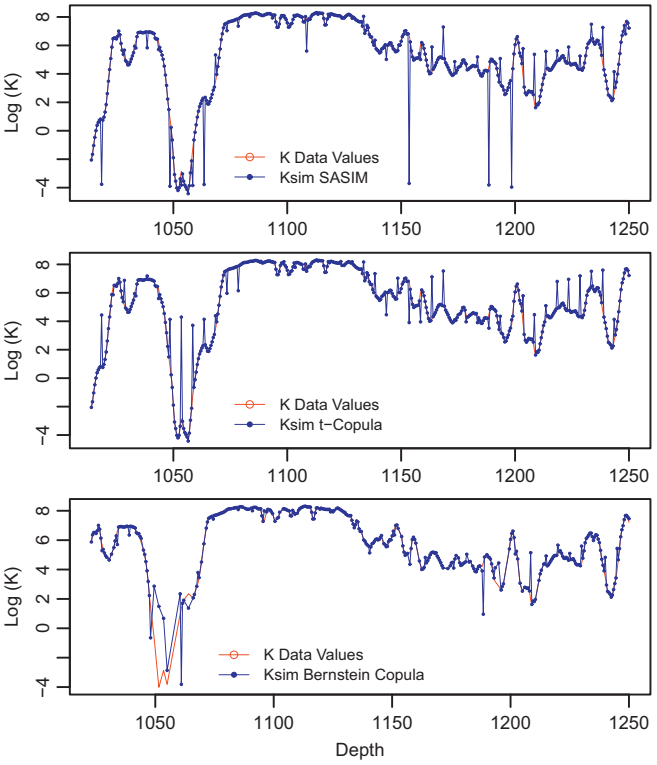

Fig. 23. Spatial distribution in a median of 10, 90% conditional simulations. Simulation methodologies (from top to bottom order): SASIM, *t*-copula and bivariate Bernstein copula.

Table 7
Complete MSE table. For a single 50% conditional simulation.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	3.22	100	–
t-Copula	2.56	79.5	100
Bernstein copula	0.88	34.4	34.4

Table 8
Complete MSE table. For a median of 50% conditional simulations.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	3.04	100	–
t-Copula	1.65	54.3	100
Bernstein copula	0.53	17.4	32.1

Table 9
Complete MSE table. For a single 90% conditional simulation.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	0.92	100	–
t-Copula	0.51	55.4	100
Bernstein copula	0.40	43.5	78.4

second, a straightforward way to perform nonparametric quantile regression; and third, an easy way to implement a nonparametric copula into a stochastic geostatistical simulation.

In contrast to the parametric approach, the nonparametric one allows us to model nonlinear relationships between petrophysical properties without assuming any distribution function as the

Table 10
Complete table of error differences between each method in terms of percentage. For a median of 90% conditional simulations.

Method	MSE	Versus SASIM (%)	Versus *t*-copula (%)
SASIM	0.81	100	–
t-Copula	0.40	49.4	100
Bernstein copula	0.33	40.7	82.5

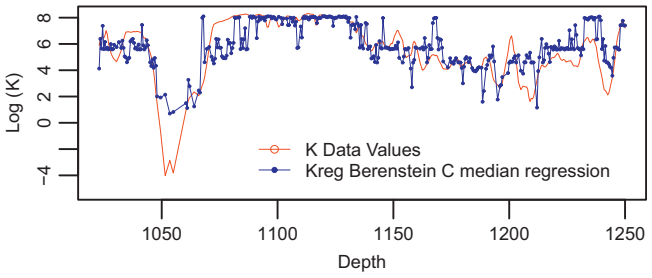

Fig. 24. Spatial Bernstein copula median regression of permeability.

t-copula does, because the Bernstein copula is based on the empirical distribution function, and consequently, it may reproduce the data variability and the extreme values in a more natural way.

Since all information about the dependence structure is contained in the underlying copula, the histogram of permeability and porosity are automatically reproduced. Hence, the objective function is reduced and consequently its computational cost is lowered, we can take this advantage to introduce other components to the objective function in order to get a better solution.

Another advantage about using the Bernstein copula is that there is no need to make logarithmic transformations of permeability, i.e., we do not have to make back transformations that can potentially bias the results. In fact, copulas are invariant under strictly increasing transformations of the variables.

In the case study, in contrast with the other two methods, SASIM of GSLIB and *t*-copula, the Bernstein copula has a mean squared error reduction of about 83% (from 7.00 to 1.17) in nonconditional simulations, consequently, it has more accurate results. It is necessary to say that there must be implemented efficient computational algorithms in order to speed up the computing time, because the Bernstein copula increases the computational effort in higher dimensions.

In this study case we noted that if we just performed a nonparametric quantile regression we obtain results that can compete (spatially speaking) with simulated annealing results (1.83 Bernstein copula median regression versus 1.85 in a single SA simulation).

The use of nonparametric copulas opens a promising line of research to model complex dependence structures between petrophysical properties and their intrinsic spatial dependence. In the geostatistical simulation framework, for joint simulations we propose to use the simulated annealing method, but we can use another optimization method which may give us more accurate results.

Acknowledgments

The present work was partially supported by the two following projects: IN110311 from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) at Universidad Nacional Autónoma de México, and by 116606 (Y.00102) from Fondo Sectorial CONACYT-SENER-Hidrocarburos.

References

Al-harthy, M., Begg, S., Bratvold, R.B., 2005. Copulas: a new technique to model dependence in petroleum decision making. J. Pet. Sci. Eng. 57, 195–208.

Balan, B., Mohaghegh, S., Ameri, S., 1995. State-of-The-Art in Permeability Determination From Well Log Data: Part 1. A Comparative Study, Model Development, SPE 30978.

Bennett, W.M., Linton, P.F., Rectenwald, E.E., 2002. Hydrologic Investigation of the Floridian Aquifer System, Western Hillsboro Basin, Palm Beach County, Florida, Technical Publication WS-8, South Florida Water Management District.

Billingsley, P., 1995. Probability and Measure, 3rd ed. Wiley, New York.

Cosentino, L., 2001. Integrated Reservoir Studies, 2nd ed. TECHNIP, Paris, France.

De Melo e Silva Accioly, R., Yassuo Chiyoshi, F., 2004. Modeling dependence with copulas: a useful tool for field development decision process. J. Pet. Sci. Eng. 44, 83–91.

Deheuvels, P., 1979. La fonction de dépendance empirique et ses porpriétés. Un test nonparamétrique d'independance. Acad. R. Belg. Bull. Cl. Sci. (5) 65 (6), 274–292.

Deutsch, C.V., 1992. Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data. Ph.D. Thesis, Stanford University, Stanford, CA.

Deutsch, C.V., Cockerham, P.W., 1994. Geostatistical Modeling of Permeability with Annealing Cosimulation (ACS), SPE 28413.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide, 2nd ed. Oxford University Press, New York.

Díaz-Viera, M., Casar-González, R., 2005. Stochastic simulation of complex dependency patterns of petrophysical properties using *t*-copulas. In: Proceedings of IAMG'05: GIS and Spatial Analysis, vol. 2, pp. 749–755.

Díaz-Viera, M., Anguiano-Rojas, P., Mousatov, A., Kazatchenko, E., Markov, M., 2006. Stochastic modeling of permeability in double porosity carbonates applying a Monte-Carlo simulation method with *t*-copula. In: SPWLA 47th Annual Logging Symposium.

Díaz-Viera, M., Hernández-Maldonado, V., Mendez-Venegas, J., 2010. RGEOESTAD: Un programa de código abierto para aplicaciones geoestadísticas basado en R-Project, México ⟨http://mmc2.geofisica.unam.mx/gmee/paquetes.html⟩.

Dréo, J., Pétrowski, A., Siarry, P., Taillard, T., 2006. Metaheuristics for Hard Optimization, 1st ed. Springer-Verlag, Berlin, Heidelberg, Germany.

Erdely, A., Díaz-Viera, M., 2009. Nonparametric and semiparametric bivariate modeling of petrophysical porosity–permeability well log data. In: Jaworski, P., et al. (Eds.), Copula Theory and its Applications, Lecture Notes in Statistics, vol. 198. Springer, pp. 267–278.

Fermanian, J.D., Radulović, D., Wegcamp, M., 2004. Weak convergence of empirical copula processes. Bernoulli 10 (5), 847–860.

Haldorsen, H., Damsleth, E., 1990. Stochastic modeling. J. Pet. Technol. April, pp. 404–412.

Holden, L., Madsen, R., Skorstad, A., Jakobsen, K.A., Tjolsen, C.B., Vik, S., 1995. Use of well test data in stochastic reservoir modelling. In: SPE Annual Technical Conference and Exhibition held in Dallas, USA, 22–25 October 1995.

Journel, A., Alabert, F., 1990. New method for reservoir mapping. J. Pet. Technol. February, pp. 212–218.

Journel, A., Zhu, H., 1990. Integrating soft seismic data: Markov–Bayes updating, an alternative to cokriging and traditional regression. In: Report 3, Stanford Center for Reservoir Forecasting, Stanford, CA.

Kazatchenko, E., Markov, M., Mousatov, A., 2004. Joint inversion of acoustic and resistivity data for carbonate microstructure evaluation. Petrophysics 45 (2), 130–140.

Kazatchenko, E., Markov, M., Mousatov, A., Parra, J., 2006. Carbonate microstructure determination by inversion of acoustic and electrical data: application to a South Florida aquifer. J. Appl. Geophys. 59, 1–15.

Landa, J.L., Kamal, M.M., Jenkins, C.D., Horne, R.N., 1996. Reservoir Characterization Constrained to Well Test Data: A Field Example. In: SPE Annual Technical Conference and Exhibition held in Denver, Colorado, USA, 6–9 October 1996.

Nelsen, R.B., 2006. An Introduction to Copulas, 2nd ed. Springer, New York.

Muñoz-Pérez, J., Fernández-Palacín, A., 1987. Estimating the quantile function by Bernstein polynomials. Comput. Stat. Data Anal. 5, 391–397.

Parra, J.O., Hackert, C.L., Collier, H.A., Bennett, M., 2001. A Methodology to Integrate Magnetic Resonance and Acoustic Measurements for Reservoir Characterization. Report DOE/BC/ 15203-3, National Petroleum Technology Office, Department of Energy, Tulsa, Oklahoma.

Parra, J.O., Hackert, C.L., 2002. Permeability and porosity images based on cross-well reflection seismic measurements of a vuggy carbonate aquifer at the Hillsboro site, South Florida, In: 72nd Annual Meeting of the Society of Exploration Geophysicists, Paper VCD P1.2.

Rogers, S.J., Chen, H.C., Kopaska-Merkel, D.C., Fang, J.H., 1995. Predicting permeability from porosity using artificial neural network. AAPG Bull. December.

Sancetta, A., Satchell, S., 2004. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. Econom. Theory 20, 535–562.

Sancetta, A., 2007. Nonparametric estimation of distributions with given marginals via Bernstein–Kantorovic polynomials: L_1 and pointwise convergence theory. J. Multivar. Anal. 98 (7), 1376–1390.

Sklar, A., 1959. Fonctions de répartition á *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris. 8, 229–231.

Suro-Perez, V., 1992. Indicator Kriging and Simulation Based on Principal Component Analysis. Ph.D. Thesis, Stanford University, Stanford, CA.