Comparison of Cosimulation Techniques

Oy Leuangthong (oy@ualberta.ca)
Department of Civil & Environmental Engineering, University of Alberta

Abstract

This paper compares the results of conventional cosimulation algorithms applied to a real petroleum data set. The techniques employed are a direct consequence of adopting some of the models of coregionalization discussed in a companion paper. Four practical methods are compared: (1) cosimulation with full cokriging, (2) cosimulation with collocated cokriging, (3) cosimulation with stepwise conditionally transformed variables, and (4) indicator simulation with the Markov-Bayes model. Multiple realizations are generated for each technique and then processed through a simple flow simulation. Each method is compared based on the resulting flow performance, as well as reproduction of the multivariate distribution.

Introduction

In spite of the fact that real data are often multivariate in nature, there are only a few cosimulation algorithms that are applied in practice. A balance is struck between honouring the available information and the simplicity/practicality of the modelling process. The most complex step in geostatistical modelling is the choice and subsequent fitting of a coregionalization model. The type of simulation that follows is a direct consequence of this decision.

Full cosimulation results from the adoption of a linear model of coregionalization (LMC), while the simpler (and more common) collocated cosimulation is a consequence of adopting the Markov hypothesis. Commercial and public domain software implement the conventional simulation algorithms, so the actual simulation step of a project is simple. This note compares the simulation results of a petroleum data set using different cosimulation techniques. Performance of the reservoir after flow simulation shows the impact of each technique on the response variable. As well, reproduction of the multivariate distribution will also be examined.

The Data

The data set is the well used West Texas "Amoco" data provided for development and testing of geostatistical algorithms. There are three variables of interest: porosity, permeability and seismic data. The seismic data is considered as soft data and is available in 2-D (see Figure 1), which will be used as secondary data in the cosimulation of porosity and permeability. Data sampling is fairly regular (Figure 1), with large areal extents and relatively minor vertical depth. This, combined with the 2-D seismic data, results in the decision to construct 2-D simulations.

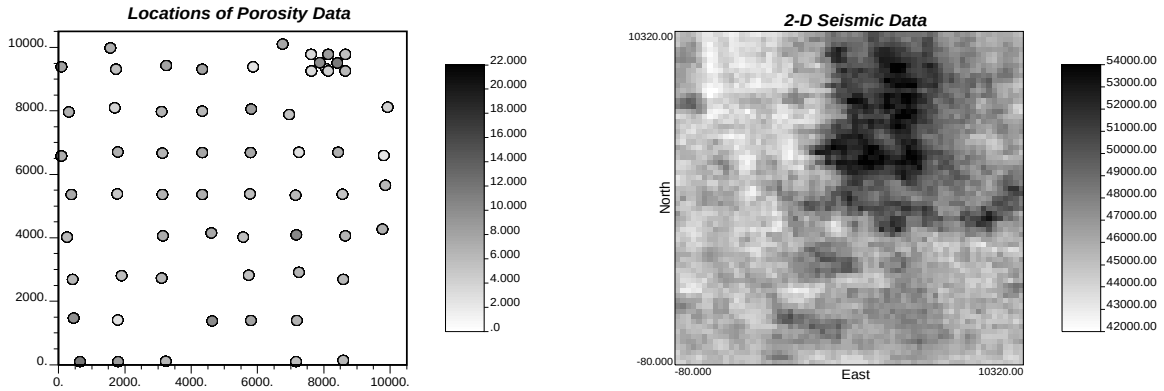

Figure 1: Location map of available hard data (left). Only porosity is shown (collocated permeability data is available). Map of available seismic data for use as soft data (right).

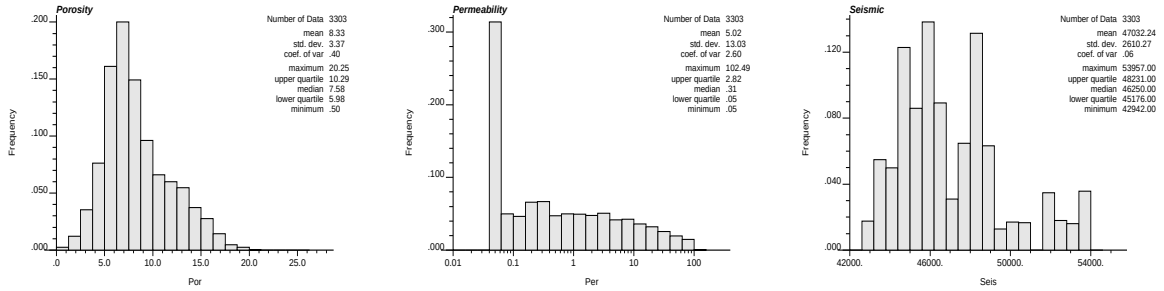

Figure 2: Histogram of porosity (left), permeability (centre) and collocated seismic data (right).

Porosity and permeability are hard data and represent the two variables that will be simulated. The distributions for each variable are shown in Figure 2, and the bivariate relations between all three variables are presented in Figure 3. It is clear that a functional relationship with porosity was used to obtain permeability data.

In the following sections, multiple cosimulation algorithms will be applied to this data set. Since the coregionalization models used in each technique employed may differ, the variogram model(s) used in simulation will be presented in that particular section.

Cosimulation Algorithms

All of the techniques that will be compared share one common feature: all employ a sequential simulation approach. The basic steps in sequential simulation, excluding any prior transforms, are:

1. Define a random path visiting each location in the domain.
2. At each location:

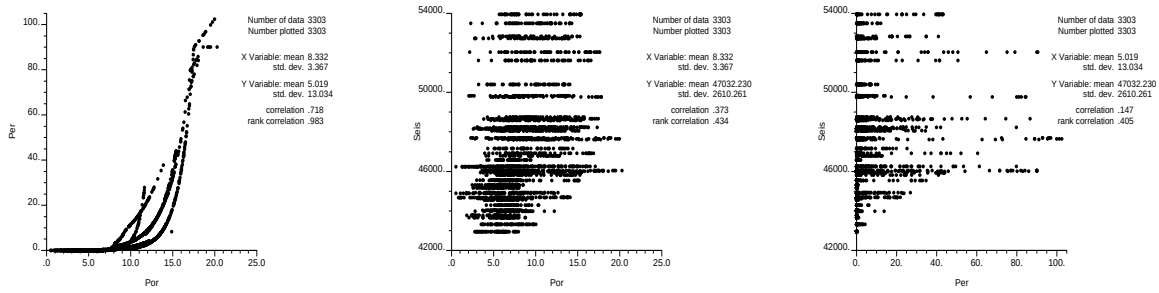

Figure 3: Cross plot of porosity-permeability (left), porosity-seismic (centre) and permeability-seismic data (right).

(a) Search for nearby primary and secondary data, and previously simulated nodes.

(b) Perform (co)kriging to determine the kriged mean and variance of the conditional cumulative distribution function (ccdf).

(c) Draw from the ccdf, defined by the kriged parameters, using Monte Carlo simulation.

(d) Proceed to next location, and repeat until all locations are visited.

Both the conventional full and collocated cosimulation follow the above procedure almost exactly (after normal score transformation) - the main difference between the two lies in the number of secondary data used in the co-kriging step. These two approaches follow directly from adopting either an LMC or a Markov Model, respectively.

Recent development of the stepwise conditional transformation to improve multivariate Gaussian simulation presents yet another cosimulation approach. A complex model of coregionalization is an implicit result of applying this multivariate transform; however, the resulting independent Gaussian variables only require independent simulation, making the approach simple to apply.

Alternatively, a sequential indicator approach may be used to define the *shape* of the ccdf (rather than assuming a simple parametric distribution). The Markov-Bayes coregionalization model permits the consideration of soft data, such as seismic impedance.

These four cosimulation algorithms are the main approaches that will be compared. A brief summary of the specific coregionalization models can be found in the companion paper [7]. Details of the variogram modelling are provided in the next sections. For all techniques, only a 2-D simulation at mid-depth is performed . One hundred realizations are generated for each variable. In each simulation, the entire data set is used for 3-D variogram modelling, but only the samples at the midpoint in each well are used to condition the simulation. These conditioning data are *not* assigned to grid nodes.

Cosimulation using Full Cokriging

Since simulation will be performed in normal or Gaussian space, a normal score transform is independently applied to each variable. The resulting crossplots are shown in Figure 4, and clearly show non-Gaussian features.

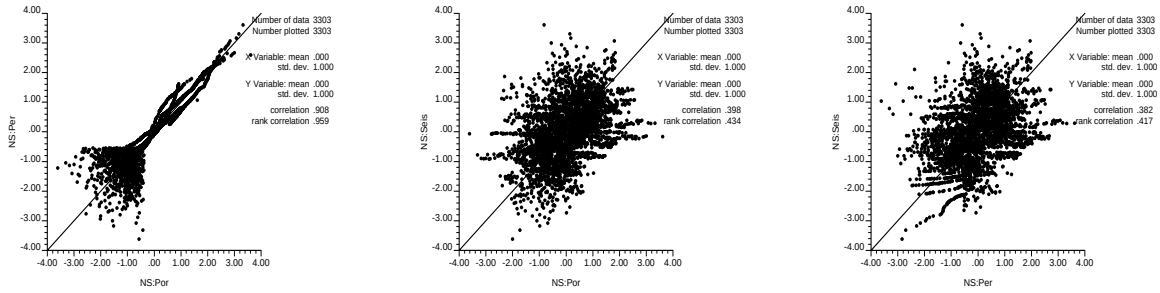

Figure 4: Cross plot of normal scores for porosity-permeability (left), porosity-seismic (centre) and permeability-seismic data (right).

Variable-Variable	Nugget	**cc Structure 1**	**cc Structure 2**	**cc Structure 3**
Type		Exponential	Spherical	Spherical
Range Parameters		$a_v = 35$ $a_{hmax} = 1.0$ $a_{hmin} = 1.0$	$a_v = 35$ $a_{hmax} = 16000$ $a_{hmin} = 5000$	$a_v = 35$ $a_{hmax} = \infty$ $a_{hmin} = 5000$
$\phi - \phi$	0.10	0.55	0.20	0.15
$K - K$	0.20	0.40	0.29	0.11
$AI - AI$	0.00	0.30	0.55	0.15
$\phi - K$	0.10	0.45	0.24	0.12
$\phi - AI$	0.00	0.00	0.25	0.15
$K - AI$	0.00	0.00	0.26	0.12

Table 1: Table of variance contribution (cc) of each nested structure in the six variograms constituting an LMC model for full cokriging.

Challenges in modelling an LMC for more than two variables make this approach cumbersome to employ. Nevertheless, an LMC model is required to define the spatial relationships between the hard porosity and permeability data to the soft seismic data. The six experimental and model variograms are shown in Figure 5, and are tabulated in Table 1.

Cosimulation using Collocated Cokriging

This alternative is attractive simply due to inference of fewer variograms. Note that only Markov Model I will be applied (henceforth simply referred to as the Markov Model). This model only requires that the normal score variogram for the primary data be modelled, and the collocated secondary data be used in cokriging.

Again Gaussian simulation will be performed, so a normal score transform is applied to porosity and permeability. The correlation coefficients in the cross plots shown in Figure 4 are required for this simulation. The required normal scores variograms are the same as the direct variograms listed in Table 1.

Stepwise Conditional Transform

From Figure 4, the trivariate distribution of the normal scores is clearly non-Gaussian. Applying sequential Gaussian simulation using the normal scores would violate an inher-

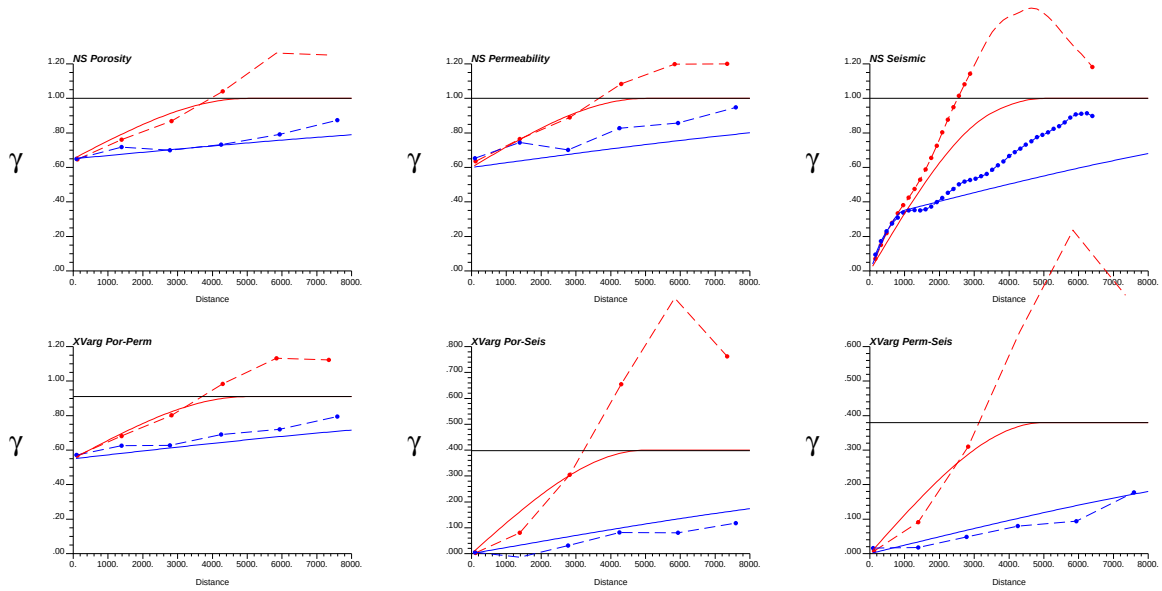

Figure 5: Direct variograms for porosity (top left), permeability (top centre) and seismic data (top right). Cross variograms between porosity-permeability (bottom left), porosity-seismic (bottom centre) and permeability-seismic (bottom right). Dashed lines indicate experimentally calculated variograms, solid lines are the fitted models. Dark lines correspond to the direction of minimum horizontal continuity, while lighter lines correspond to direction of maximum horizontal continuity.

ent multi-Gaussian assumption in the technique. For cases such as these, the stepwise conditional transform can be applied.

Since seismic data is available at all locations in the 2-D grid, it is selected as the primary variable. Porosity is chosen as the secondary variable and permeability is the last variable to be transformed (conditional to both seismic and porosity data). With over 3000 data available, ten probability classes are chosen, resulting in a minimum of 30 data to define each conditional distribution.

The multivariate distribution of the stepwise conditional (SC) scores are shown in Figure 6. A banding effect is apparent. This is partially attributed to the fact that seismic data is 2-D, and so the seismic data that is collocated with the well data is a constant at each well. Secondly, the strong functional relationship between porosity and permeability (in Figure 3) may have been transferred to the transformed Gaussian space.

Only two variograms are required to be modelled - SC porosity and SC permeability (since seismic is available everywhere, there is no need to simulate this variable). The cross variograms for $\mathbf{h} > 0$ between the transformed variables were checked, with a result showing a maximum correlation of 0.07. Since this correlation is quite low, independence of these variables is assumed. The direct variograms for SC porosity and SC permeability are shown in Figure 7. Compared to the variograms for the normal scores (Figure 5), the SC scores exhibit a higher nugget effect in both directions (and note that a horizontally isotropic

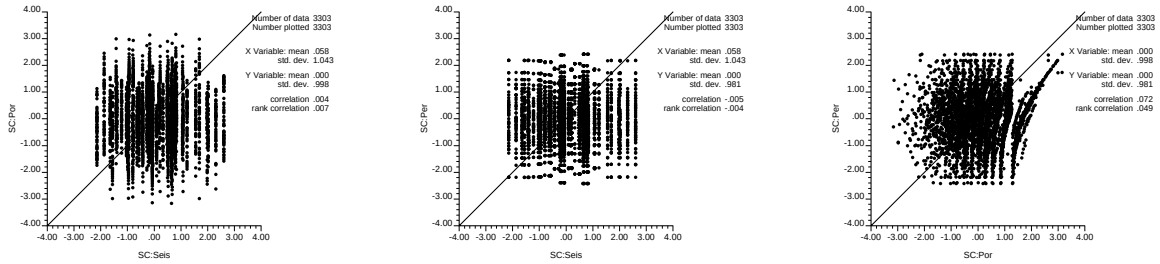

Figure 6: Cross plot of stepwise conditionally transformed scores for porosity-permeability (left), porosity-seismic (centre) and permeability-seismic data (right).

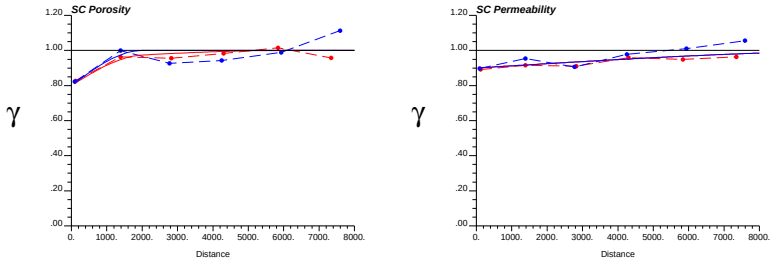

Figure 7: Direct variograms for stepwise conditional scores of porosity (left) and permeability (right). Dashed lines indicate experimentally calculated variograms, solid lines are the fitted models. Dark lines correspond to the direction of minimum horizontal continuity, while lighter lines correspond to direction of maximum horizontal continuity.

variogram could be modelled here).

Indicator Simulation using Markov-Bayes Model

Unlike Gaussian techniques, the indicator approach is a non-parametric method. The conditional distribution is defined by the kriged estimates for a series of thresholds. This requires that variograms are modelled for each threshold. Technically, the cross-variogram between thresholds should also be modelled; however, this quickly becomes cumbersome and often impractical as the number of thresholds increases (e.g. for 5 thresholds, 10 variograms are required to be simultaneously modelled as per the LMC).

For this data, porosity is modelled using nine thresholds (corresponding to the 10 deciles). The indicator variograms for each threshold are shown in Figure 8. They show similar structures as the Gaussian variograms, except with slightly higher nugget effects.

The original permeability data consists of a large number of 0.05 values, comprising almost 30 per cent. For this reason, permeability is model by defining seven thresholds: 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, and 0.90. The corresponding indicator variograms are provided in Figure 9. Similar to the Gaussian variograms, the indicator variograms for both porosity and permeability show high nugget structures along with zonal anisotropy.

To simulate porosity, the gridded seismic data is calibrated to obtain the scaling factors

Threshold	$B(\mathbf{u})$	
	Porosity-Seismic	Permeability-Porosity
0.10	0.0510	–
0.20	0.0916	–
0.30	0.1329	0.8814
0.40	0.1433	0.8647
0.50	0.1645	0.8858
0.60	0.1899	0.9038
0.70	0.1534	0.7704
0.80	0.0999	0.6902
0.90	0.0490	0.6198

Table 2: Table of Calibration Factors between the hard local data (left in column title) and available soft data (right in column title).

($B(z)$) to define the cross-covariance between the hard and soft data, as well as the covariance of the soft data. For the simulation of permeability, the porosity data is considered as soft data, and the corresponding calibration factors must also be calculated. Table 2 lists the calibration factors for the indicator simulation of both porosity and permeability. From Table 2, low calibration factors for simulating porosity indicate that the soft seismic data is not as informative then if these factors were high, as in the case of simulating permeability using soft porosity data.

Remarks

Alternative multivariate simulation approaches exist, including principal component analysis (PCA) and direct sequential cosimulation. PCA could be used in multivariate simulation, with a primary objective of reducing the dimension of the required cosimulation. For this particular data set, only two variables are simulated. Depending on the model of coregionalization we wish to adopt, there may be relatively little to gain from employing this technique (that is, unless the LMC is chosen, there is little reduced effort in variogram modelling but a rather inconvenient additional transformation and back-transformation that must be applied).

Direct sequential cosimulation is another alternative. The reason for not pursuing this avenue is simple: the theoretical and numerical development is currently unavailable. A discussion of the anticipated challenges and some ideas being pursued to advance research in this area are given by Leuangthong and Deutsch [8].

Comparison of Cosimulation Algorithms

The first realization of porosity and permeability are arbitrarily chosen for the purpose of a simple visual comparison. Figure 10 shows the realizations of porosity and permeability, resulting from the four approaches employed. Overall, all four techniques show low values in the west and south-east quadrant, with a high region in the north-east quadrant of the field.

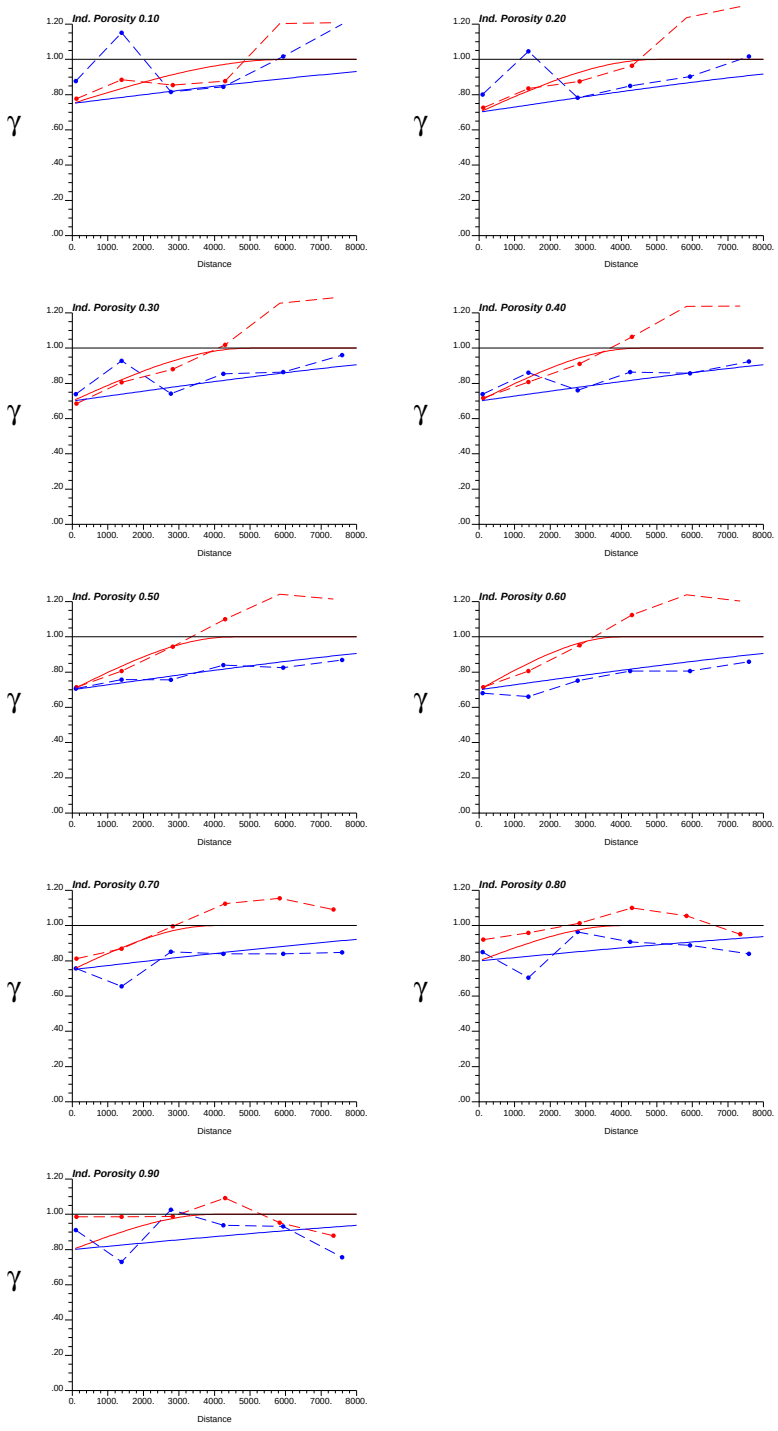

Figure 8: Indicator variograms for porosity for each of the nine thresholds. Dashed lines indicate experimentally calculated variograms, solid lines are the fitted models. Dark lines correspond to the direction of minimum horizontal continuity, while lighter lines correspond to direction of maximum horizontal continuity.

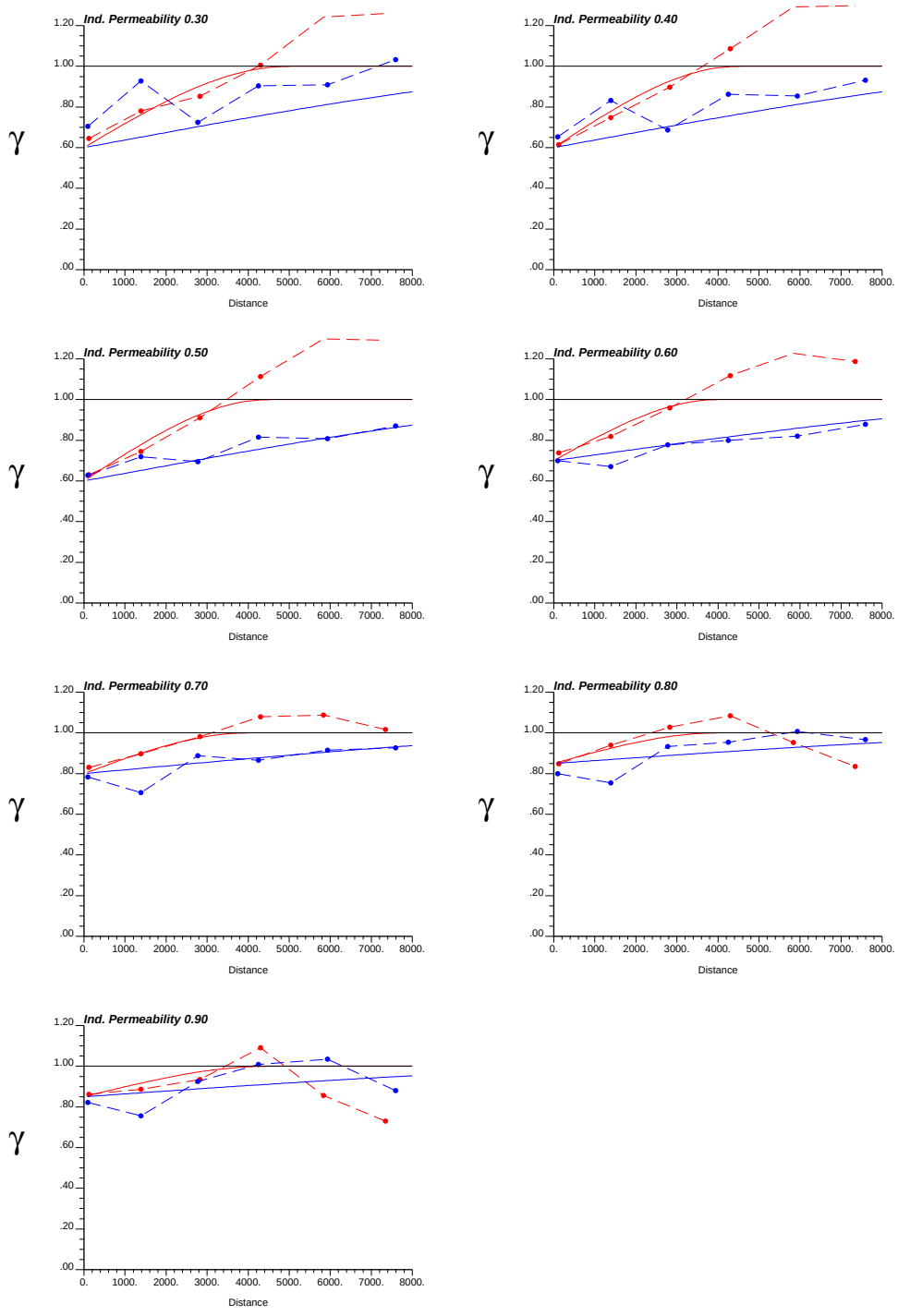

Figure 9: Indicator variograms for permeability for each of the seven thresholds. Dashed lines indicate experimentally calculated variograms, solid lines are the fitted models. Dark lines correspond to the direction of minimum horizontal continuity, while lighter lines correspond to direction of maximum horizontal continuity.

These features correspond to those exhibited in the soft data (see Figure 1). Application of indicator simulation with soft data yields the least variable realizations. Gaussian-based simulations all show higher variability - a result of the nugget structure in the Gaussian variograms.

Aside from visually, it is difficult to compare models generated by each approach without knowledge of the true reservoir. Instead, two comparisons are made: (1) a simple transfer function illustrates the effect on flow performance of the models from each cosimulation method, and (2) an examination of the resulting bivariate distributions shows how well the methods are able to reproduce higher order statistics (extending beyond the traditional histogram and variogram reproduction).

Flow simulation

A simple flow simulation program was used to test the dynamic performance of the realizations from each simulation method. The program `flowsim` by Deutsch [1] provides a quick and simple algorithm that permits all realizations to be processed. The algorithm involves defining the dimensions of both the input grid and the desired grid (the latter must be some integer, such that the input grid is divisible). Directional effective permeabilities are calculated by setting no flow boundary conditions on parallel sides. For instance, to calculate kx_{eff}, the north and south edges of the grid are set as no flow boundaries, flow equations are solved to determine the effective permeability in the X direction.

Performance Measure. Using this flow simulation program, all 100 realizations from each method were processed. Since only an overall comparison is desired, the output grid is set to $1 \times 1 \times 1$. This will provide an overall kx_{eff} and ky_{eff} for each realization. The methods are then compared based on the $p10$, $p50$ and $p90$ values for kx_{eff} and ky_{eff} (see Table 3).

Method	kx_{eff}			ky_{eff}		
	p_{10}	p_{50}	p_{90}	p_{10}	p_{50}	p_{90}
Full CK	0.191	0.223	0.270	0.350	0.458	0.569
Collocated CK	0.180	0.196	0.218	0.486	0.608	0.817
Stepwise	0.237	0.267	0.304	0.271	0.334	0.381
Markov-Bayes	0.169	0.175	0.181	0.857	1.000	1.257

Table 3: Summary of p_{10}, p_{50} and p_{90} effective permeability in X and Y directions from each of the cosimulation methods employed.

There is significant spread in both the kx_{eff} and ky_{eff} values. For the kx_{eff} values, the p_{10} and p_{90} between all the techniques overlap. Indicator simulation with the Markov-Bayes model yields the smallest range, indicating a very low variance in kx_{eff} values among all realizations. The ky_{eff} values show an even larger spread from one method to another. Gaussian simulation with the stepwise conditional transform produces the smallest ky_{eff} values as well as the smallest spread between the p_{10} and p_{90} values. Conversely, indicator simulation results in the highest ky_{eff} values.

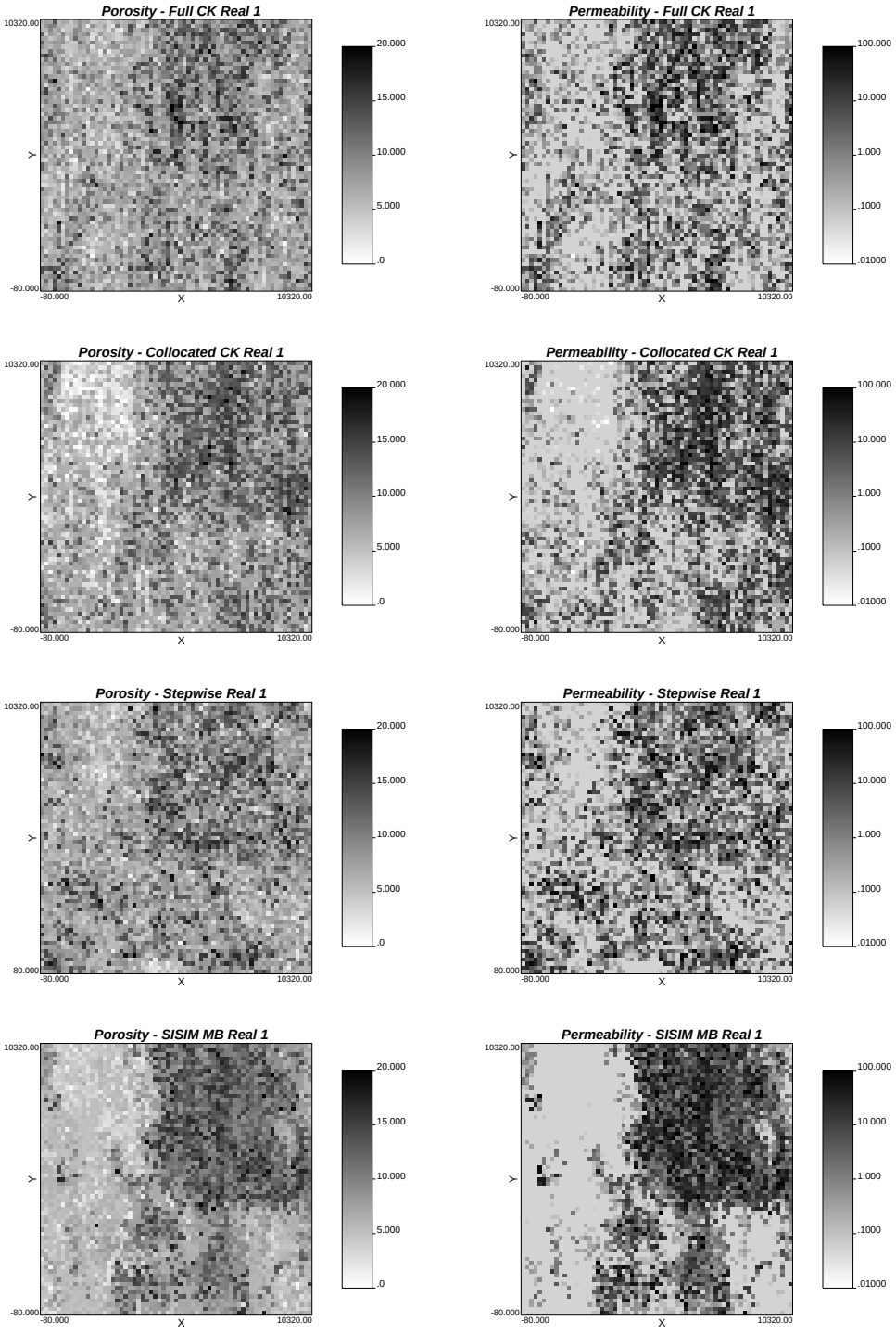

Figure 10: First realization of porosity (left) and permeability (right) for each simulation approach: full cokriging (top row), collocated cokriging (second row), indicator with Markov-Bayes (third row) and stepwise conditional (bottom row).

Since the true reservoir, and hence its effective permeabilities, are unknown, a comparison of the flow performance of the models generated by each method is not a sufficient measure to compare the performance of the techniques. Instead, we can examine the ability of each method to reproduce the original data distributions.

Reproduction of Multivariate Distribution

Recall that the conditioning data was only a small subset of the entire data set (only midwell samples were used). We know that simulation reproduces the histogram, and use of simple kriging in simulation will reproduce the variogram. Histograms and variograms were verified, and confirmed these expectations.

Reproduction of the bivariate distribution of porosity and permeability were also examined. These are shown in Figure 11 for realizations 1 and 100. The Gaussian techniques show good reproduction of the bivariate distribution, with the stepwise approach reproducing the shape most accurately. The indicator method does reproduce the overall shape of the bivariate distribution; however, it produces simulated values with higher variance than those generated by the Gaussian techniques.

Further, a comparison of the bivariate distribution in Gaussian space is compared for the three Gaussian techniques. This bears further investigation due to the non-Gaussian features of the normal scores cross plot in Figure 4, which corresponds to the transformed data used in the conventional full and collocated cosimulation. Figure 12 shows the simulated porosity and permeability values prior to back transformation for the three Gaussian methods. Clearly the full cokriging approach assumed a bivariate Gaussian distribution which is inconsistent with the input bivariate distribution. Both the collocated cokriging and stepwise method reproduced the input Gaussian bivariate distribution.

Discussion

Among the conventional Gaussian cosimulation methods employed, full cokriging was expected to yield better realizations than collocated cokriging since it uses more secondary information. Further, most implementations of collocated cokriging uses the Markov Model I (MMI) assumption, which is considered to be a poor assumption when the secondary data has a much larger support than the local hard data. For this data, this is precisely the case. Utilization of the Markov Model II (MMII) assumption may yield better results.

Since the multiGaussian assumptions are violated in the applications with the classical Gaussian simulations, the indicator approach was expected to be better since it is a nonparametric technique. However, calibration values of seismic for porosity simulation are very low suggesting that the soft data are not as informative for simulation of the hard data. The high calibration factors of porosity for the simulation of permeability capitalizes the information provided by soft data (porosity in this instance).

Geostatistical simulation is proven to reproduce both the histogram and the variogram of the model variable; however, there is no assurance of reproducing the multivariate distribution. Gaussian simulation using the stepwise conditional transform performed better than the other three methods in the sense that it reproduced the bivariate distribution while satisfying the multiGaussian assumptions inherent in the simulation.

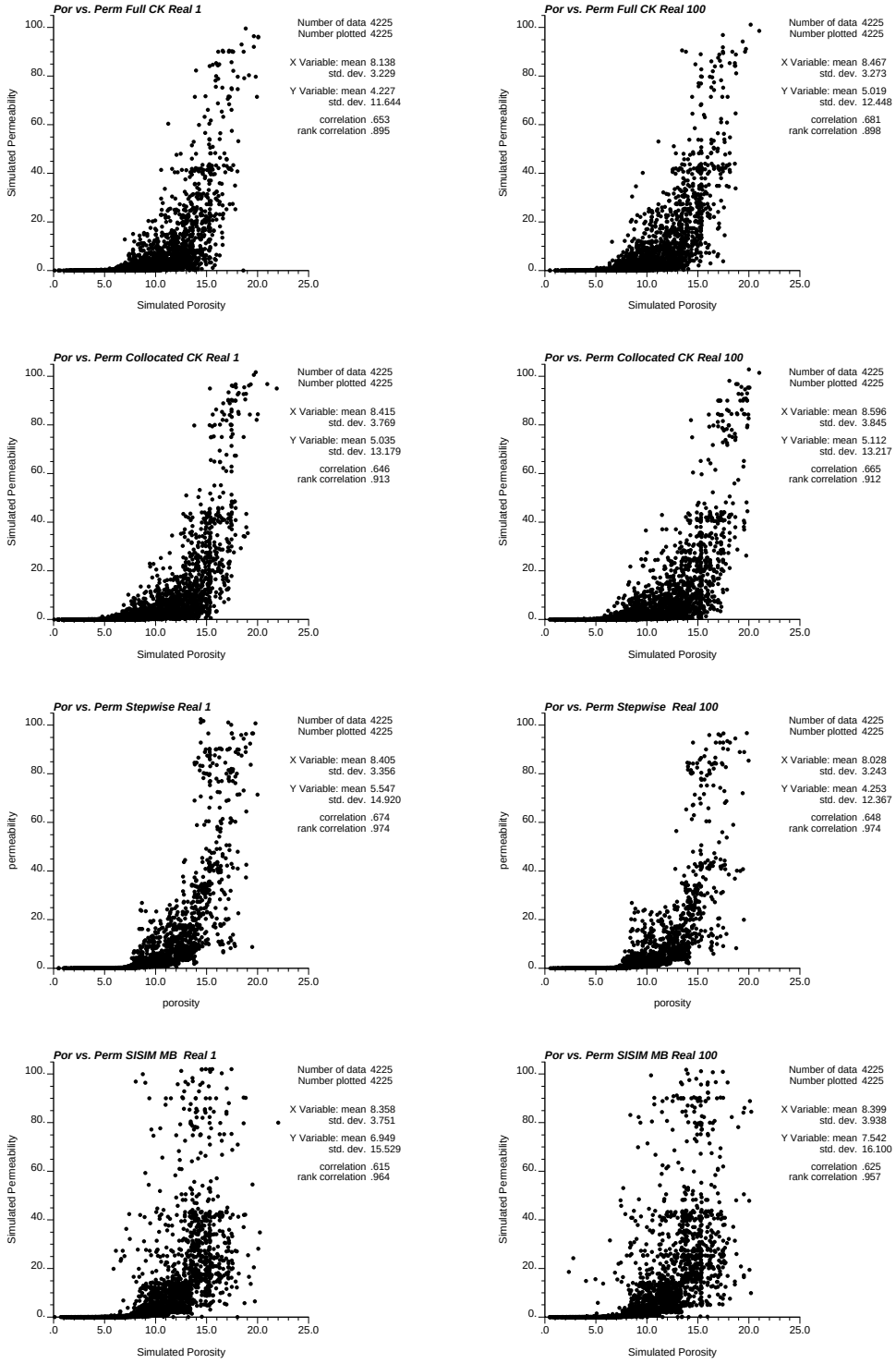

Figure 11: Cross plot of porosity (left) and permeability (right) for first realization from each simulation approach: full cokriging (top row), collocated cokriging (second row), stepwise conditional (third row) and indicator with Markov-Bayes (bottom row).

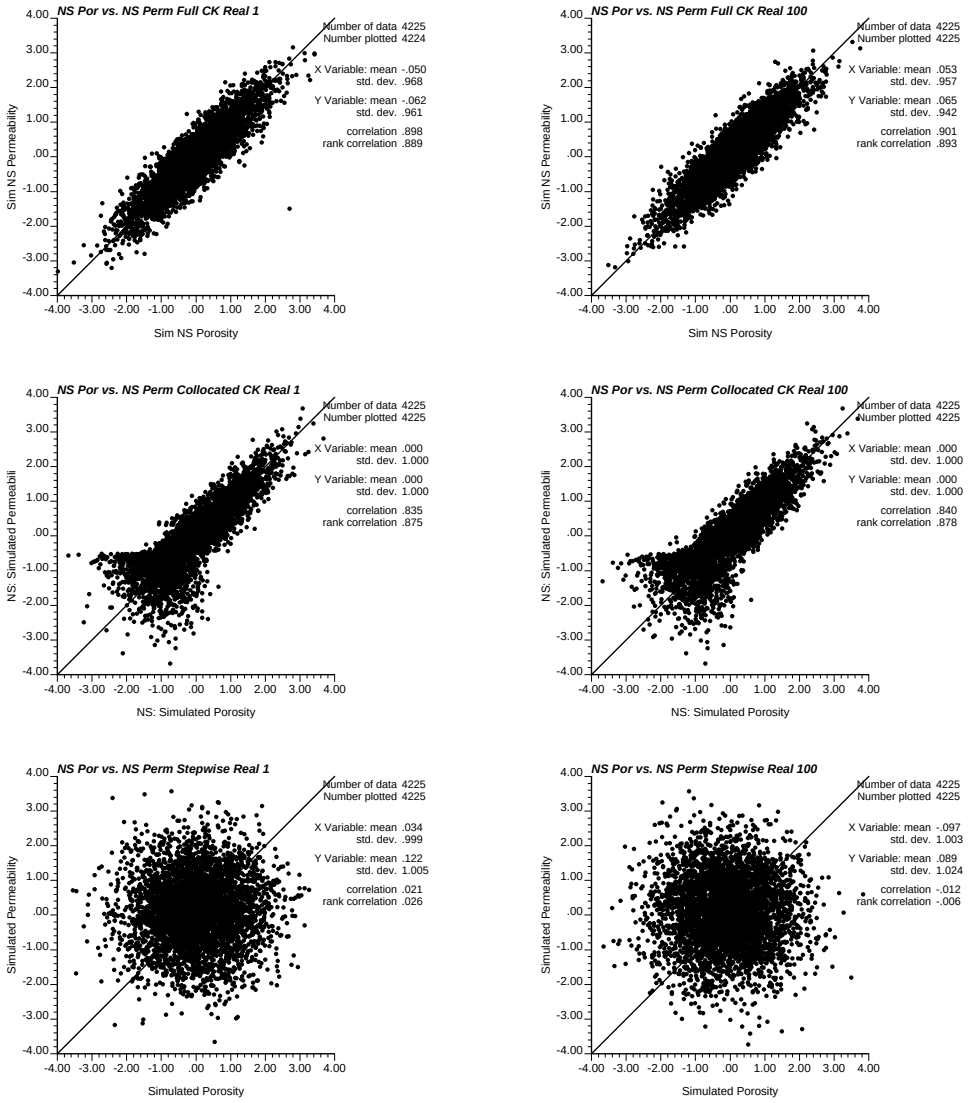

Figure 12: Cross plot of normal scores of porosity (left) and permeability (right) for first realization from each simulation approach (prior to back transformation): full cokriging (top row), collocated cokriging (middle row), and stepwise conditional (bottom row).

In practice, the choice of which cosimulation approach to apply will be affected by the amount and type of available data, the number of variables of interest, and the ease of implementation of the technique. This choice will further impact the response variable, and consequently may have significant effects on future decisions pertaining to reservoir development.

References

[1] C. V. Deutsch. Calculating effective absolute permeability in sandstone/shale sequences. *SPE Formation Evaluation*, pages 343–348, September 1989.

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

[3] P. Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 1997.

[4] A. Journel. Markov models for cross-covariances. *Mathematical Geology*, 31:955–964, 1999.

[5] A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978.

[6] A. G. Journel and H. Zhu. Integrating soft seismic data: Markov-Bayes updating, an alternative to cokriging and traditional regression. In *Report 3, Stanford Center for Reservoir Forecasting*, Stanford, CA, May 1990.

[7] O. Leuangthong. Short note on models of coregionalization. Technical report, Centre for Computational Geostatistics, University of Alberta, Edmonton, AB, March 2002.

[8] O. Leuangthong and C. Deutsch. Modeling of multivariate multiscale data. Technical report, Centre for Computational Geostatistics, University of Alberta, Edmonton, AB, March 2002.

[9] L. Shmaryan and A. Journel. Two markov models and their applications. *Mathematical Geology*, 31:965–988, 1999.