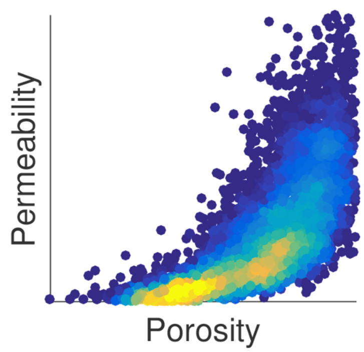
Multivariate Reservoir Property Modeling



March 18, 2016

Prepared by Ryan M. Barnett

Summary

- A workflow that is frequently applied to reservoir modeling uses cosimulation to simulate select reservoir variables, before using the cloud transformation to simulate additional variables. Issues with this conventional workflow lead to poor reproduction of the multivariate properties that exist between reservoir variables. In turn, this reduces the accuracy of reservoir performance prediction.
- These issues motivate the use of a Projection Pursuit Multivariate Transformation (PPMT) workflow, where reservoir properties are decorrelated while simultaneously removing multivariate complexities. Independent Gaussian simulation may then proceed, before the multivariate back-transformation restores the original distributions and potentially complex multivariate relationships.
- This provides a simpler workflow than the conventional reservoir modeling workflow. More importantly, it improves the reproduction of multivariate relationships that are often critical to reservoir performance prediction.
- Data from an oil sands SAGD project is used for comparing the conventional and PPMT workflows. The simpler PPMT workflow is demonstrated to substantially improve the reproduction of multivariate properties and hence, reservoir performance prediction.

1 The Setting

Subsurface modeling is essential to reservoir evaluation and management. It usually follows a hierarchical framework that begins with modeling of surfaces that separate major geologic zones. Modeling of facies then proceeds within each zone before modeling continuous variables within each facies. Geostatistical tools are commonly employed for modeling variables that may include porosity (ϕ), water saturation (*Sw*), oil saturation, (So), permeability (*K*), shale volume (*Vsh*), and others that are required for reservoir performance prediction. Geostatistical methods generate realizations of the reservoir, where the set of realization targets the reproduction of representative statistics such as the distribution and variogram of each variable, as well as the multivariate relationships between the variables (Chiles & Delfiner, 2012). The set of realizations quantify the reservoir uncertainty that exists due to incomplete data, geologic heterogeneity, and parameter uncertainty. Transfer functions such as flow simulation or proxy models may then be applied to the realizations to transfer geological uncertainty to reservoir performance uncertainty. This uncertainty may then be integrated in decision making to mitigate and manage risk.

The conventional workflow for generating realizations of continuous reservoir variables uses cosimulation for select reservoir variables, before using the cloud transformation to simulate additional variables (Dull, 2004; Ma & Gomez, 2011; Moore, 2011; Pyrcz & Deutsch, 2014). More formally, let the well data available for reservoir modeling be denoted by the matrix $\mathbf{Z} : z_{\alpha}, i, \alpha = 1, , n, i = 1, , M$, where n is the number of observations and M is the number of variables. Steps in the conventional workflow are as follows:

1. Normal score m variables ($m < M$) to transform them to Gaussian distributions using the well established quantile transform (Bliss, 1934; Verly, 1983), $Y_i = G^{-1}\left(F_i(Z_i)\right)$, where F_i is the cumulative distribution function (CDF) of the Z_i variable and G is the standard Gaussian CDF. In the case study to follow, the $m = 3$ variables are ϕ, *Sw* and *Vsh*;
2. Perform conditional Gaussian cosimulation of the $Y_1, ..., Y_m$ Gaussian variables, where their covariance matrix, $\mathbf{\Sigma}(\mathbf{h}) : C_{i,j}(\mathbf{h}), i, j = 1, , m$ is assumed to characterize the multivariate distribution according to coregionalization models such as the Markov model (colocated cokriging)(Almeida & Journel, 1994);
3. Normal score back-transform the Gaussian realizations of the m variables according to $Z_i = F_i^{-1}\left(G\left(Y_i\right)\right)$.

4. Sequentially simulate the o variables ($o = Mm$) using the cloud transform (Bashore et al., 1994). In the case study to follow, the $o = 2$ variables are *Kh* and *Kv*. This step decomposes into the following sub-steps:

   i. Construct a conditional bivariate CDF of the variable to be simulated Z_i (e.g., *K*), conditional to another variable Z_j (e.g., ϕ), yielding $F(Z_i|Z_j)$;

   ii. Generate probability field (p-field) (Froidevaux, 1992) realizations. No conditioning or coregionalization is used for the simulation, as only the regionalization model of Z_i (fitted model of $\gamma_{i,i}(\mathbf{h})$) is required. The output is realizations of probability values at the N grid locations $p(\mathbf{u}_\alpha), \alpha = 1, ..., N$;

   iii. Given $F(Z_i|Z_j)$ from Step (i), the simulated or known conditioning value $z_j(\mathbf{u}_\alpha)$ is used to determine the univariate conditional CDF $F(Z_i|z_j(\mathbf{u}_\alpha))$. The associated probability value from Step (ii) $p(\mathbf{u}_\alpha)$ is then used to draw a value $z_i(\mathbf{u}_\alpha)$ from $F(Z_i|z_j(\mathbf{u}_\alpha))$. Repeating this process over the grid, a realization $z_i(\mathbf{u}_\alpha), \alpha = 1, ..., N$ is simulated by using $z_j(\mathbf{u}_\alpha), \alpha = 1, ..., N$ and $p(\mathbf{u}_\alpha), \alpha = 1, ..., N$ to sample from $F(Z_i|Z_j)$.

2 The Problem

Although this workflow (referred to here as Conventional) effectively reproduces univariate properties of the variables, it often fails to reproduce multivariate properties that exist between them. Cosimulation assumes the variables follow a multivariate Gaussian (multiGaussian) distribution, which is often not the case due to the multivariate complexities that exist in geological data (see Figure 1).

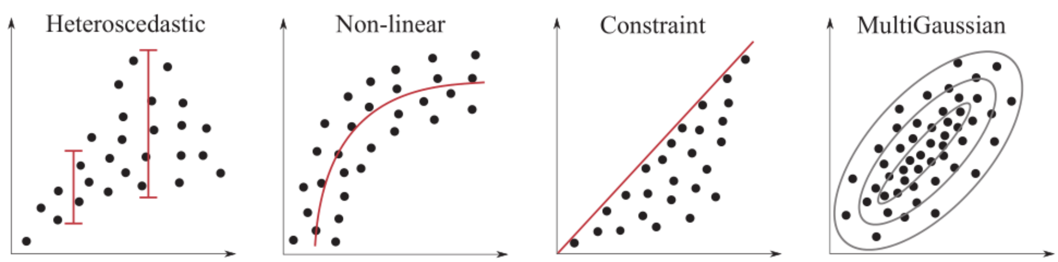


Figure 1: Schematic illustrations of multivariate complexities and a multiGaussian distribution.

When these complexities exist, $\mathbf{\Sigma}(\mathbf{h})$ will not characterize the multivariate distributions of the normal score transformed variables, meaning that realizations will not reproduce the multivariate distributions of the reservoir variables following cosimulation and back-transformation.

Although the cloud transform can reproduce complex features, it only targets the bivariate relationship that exists between the conditioning variable, Z_i, and simulated variable, Z_j, according to the modeled CDF, $F(Z_i|Z_j)$. The assumption is made that the full joint distribution, $F(Z_{1,...,M})$, will be reproduced indirectly, which is often not the case. These two issues cumulatively represent a significant problem since it is critical for geostatistical models to reproduce the multivariate distributions of reservoir variables.

3 The Solution

An alternative to the conventional workflow for modeling continuous reservoir variables will transform the data variables to an uncorrelated multiGaussian distribution where the transformed variables are independent. This facilitates independent simulation of the transformed variables. The associated multiGaussian back-transformation returns the original units, distribution, and multivariate relationships to the simulated realizations.

To perform this multiGaussian transformation, the CCG has recently developed the Projection Pursuit Multivariate Transformation (PPMT) for geostatistical modeling (Barnett et al., 2014). The PPMT is based on modified components of projection pursuit density estimation (Friedman, 1987; Hwang et al., 1994), which transforms data of virtually any multivariate form, M variables, and n observations to an uncorrelated multiGaussian distribution, $\mathbf{X}$. Consider that projecting a multiGaussian distribution onto any arbitrary vector, $\boldsymbol{\theta}$, will yield a univariate Gaussian distribution, $\mathbf{p} = \mathbf{X}\boldsymbol{\theta}$. The premise of projection pursuit is to find the vector, $\boldsymbol{\theta}$, that yields the most non-Gaussian projection, $\mathbf{p}$. The multivariate data is then transformed to make its projection Gaussian. This search and Gaussianize procedure is iterated until the least Gaussian projection approaches the univariate Gaussian model. Although the PPMT involves several steps and iterations, they are all accomplished within a single program is very straight forward to execute.

4 Case Study

The described Conventional and PPMT workflows are demonstrated using data from a heavy oil reservoir that is produced using steam assisted gravity drainage (SAGD). The workflows are evaluated based on their reproduction of properties that are representative of the reservoir and critical to SAGD performance prediction.

Data Background and Inventory

The bituminous (extra-heavy) oil of the Athabasca oil sands has a high viscosity that does not permit it to flow under normal reservoir conditions. To counter this viscosity, SAGD uses a horizontal well to inject steam into the reservoir, which forms a steam chamber that heats the bitumen (Denbina, 1998; Edmunds & Sugget, 1994). Figure 2 presents a schematic illustration of the SAGD process.

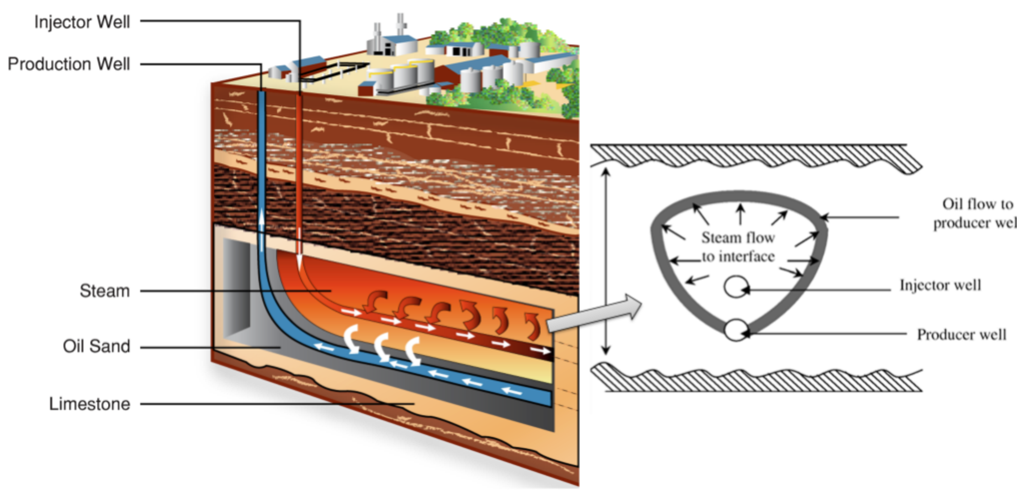


Figure 2: Schematic of the SAGD process ((Hadavand & Deutsch, 2015)).

Within the steam chamber, the heated bitumen has its viscosity lowered to the point where gravity causes it to flow down to the horizontal producer well (which lies below the horizontal injector well). Geostatistical modeling of a SAGD reservoir should include several variables. As with conventional petroleum reservoirs, modeled ϕ, *K* and *Sw* are required for determining the resource and forecasting reservoir production. Beyond those applications, *Sw* is also essential for identifying thief zones that increase steam requirements. Both horizontal (*Kh*) and vertical permeability (*Kv*) are also modeled, as *Kv* is the dominant property in determining whether steam can flow vertically up to heat oil above the injector well, and in turn, whether heated oil can flow vertically down to the producer well.

The public data for this case study is from an active SAGD operation. It consists of 4,998 observations that sample ϕ, *Vsh*, *Sw*, *Kh*, and *Kv*. The observations have been sampled using vertical wells that are drilled from surface to the basement of the reservoir formation. Figure 3 presents the relative spatial configuration of the observations.

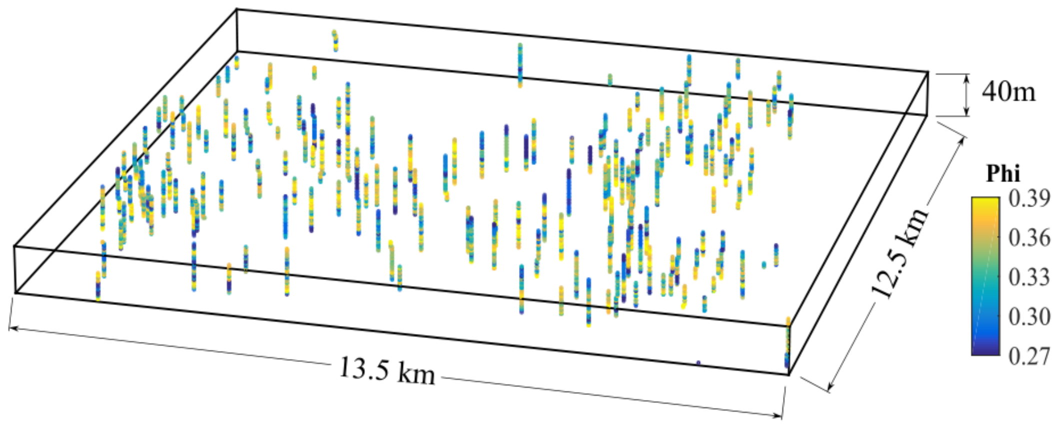


Figure 3: Relative locations of well observations that are colored by ϕ (20:1 vertical exaggeration).

As described, this study is concerned with the modeling of continuous properties, which follows the hierarchical modeling of major geologic domains or strata, and facies or lithofacies. To simplify the evaluation of results, the presented data is drawn from a stratigraphically flattened subset of a larger dataset. It is almost entirely composed of inclined heterolithic stratification (IHS), which is a vertical succession of sand and mud drapes that are deposited by fluvioestuarine point bars (Thomas et al., 1998). The assumption is made that all of the samples lie within a stationary domain that does not require further facies subdivision.

Bivariate properties are displayed in Figure **??** using scatterplots and the correlation coefficient, ρ. The scatterplots are colored according to the bivariate Gaussian kernel density estimate (KDE) that is calculated at each observation. This is referred to as KDE scatterplots, which may aid in observing the bivariate density of the data, and later, in judging whether simulated realizations reproduce those densities. All of the multivariate complexities that were previously schematically represented are present in the oil sands scatterplots, including non-linearity, heteroscedasticity and constraints. Consequently, this multivariate distribution is not expected to be reproduced by geostatistical workflows that fail to remove these complexities prior to the application of Gaussian simulation algorithms.

PPMT Workflow Results

The applied PPMT workflow is summarized as:

1. Apply the PPMT to transform the five variables to be uncorrelated and multiGaussian;
2. Independently simulate the five variables;
3. Apply the PPMT back-transform to return the simulated realizations to original space.

Figure 5 presents KDE scatterplots of the PPMT data (following Step 1), where zero correlation and typical multiGaussian density contours are observed.

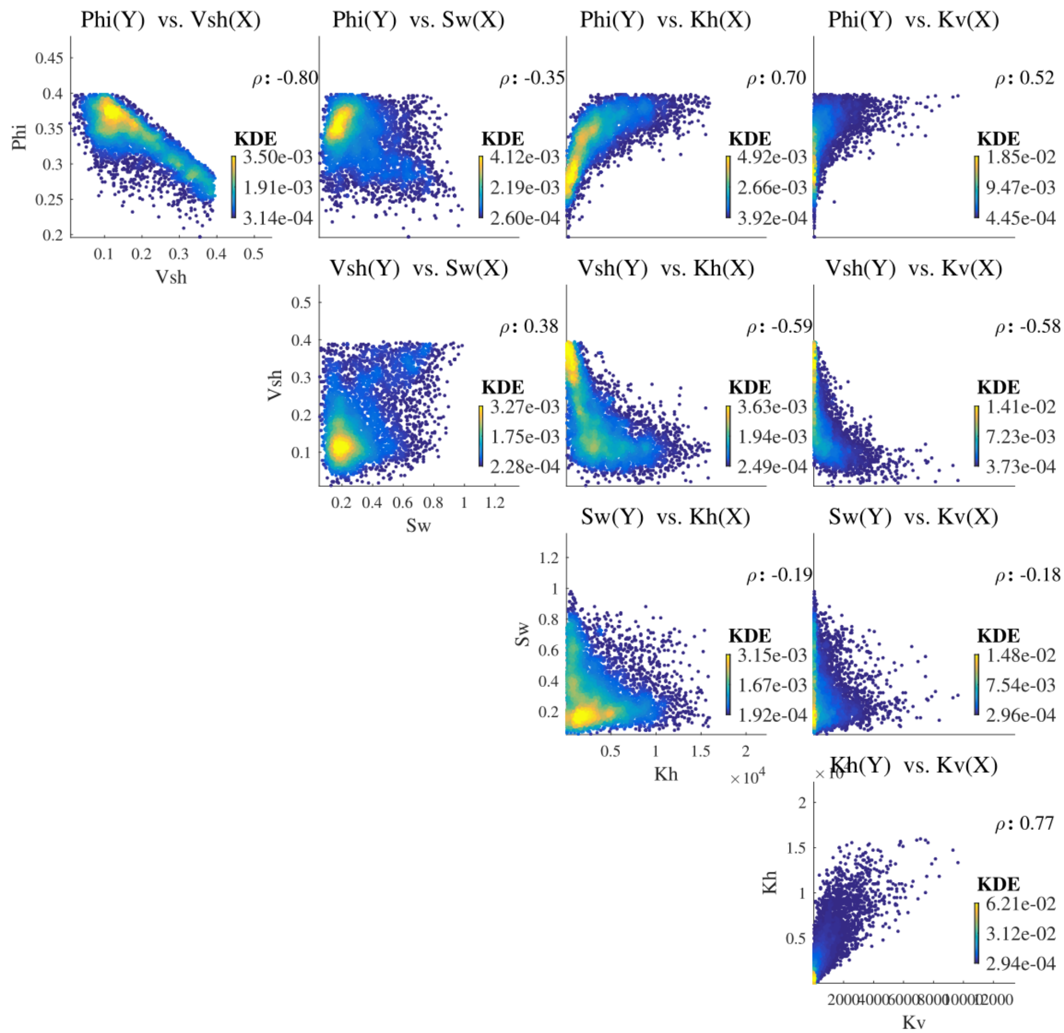


Figure 4: KDE scatterplots and correlation of the original data.

Sequential Gaussian simulation (SGSIM) (Deutsch & Journel, 1998) is used for generating 100 realizations of the reservoir in Step 2. The simulation grid is composed of 135 x 125 x 40 grid nodes in the x, y and z directions respectively. Each node is separated by 100 metres in the horizontal direction and 1 metre in the vertical direction. Following simulation, the realizations are back-transformed in Step 3 to original units. Slices of one back-transformed realization appear in Figure 6, where they are colored by select variables for visual reference.

Reproduction of bivariate properties is inspected in Figure 7 where KDE scatterplots of the data are compared with that of a realization. Excellent reproduction is seen based on visual comparison of the densities and the correlation statistic. The displayed root mean squared error (RMSE) in each bivariate plot is calculated as the square root of the average difference between the data KDE and realization KDE. This KDE RMSE quantifies the reproduction of each bivariate density, which is used to compare the PPMT and Conventional workflow in the next section.

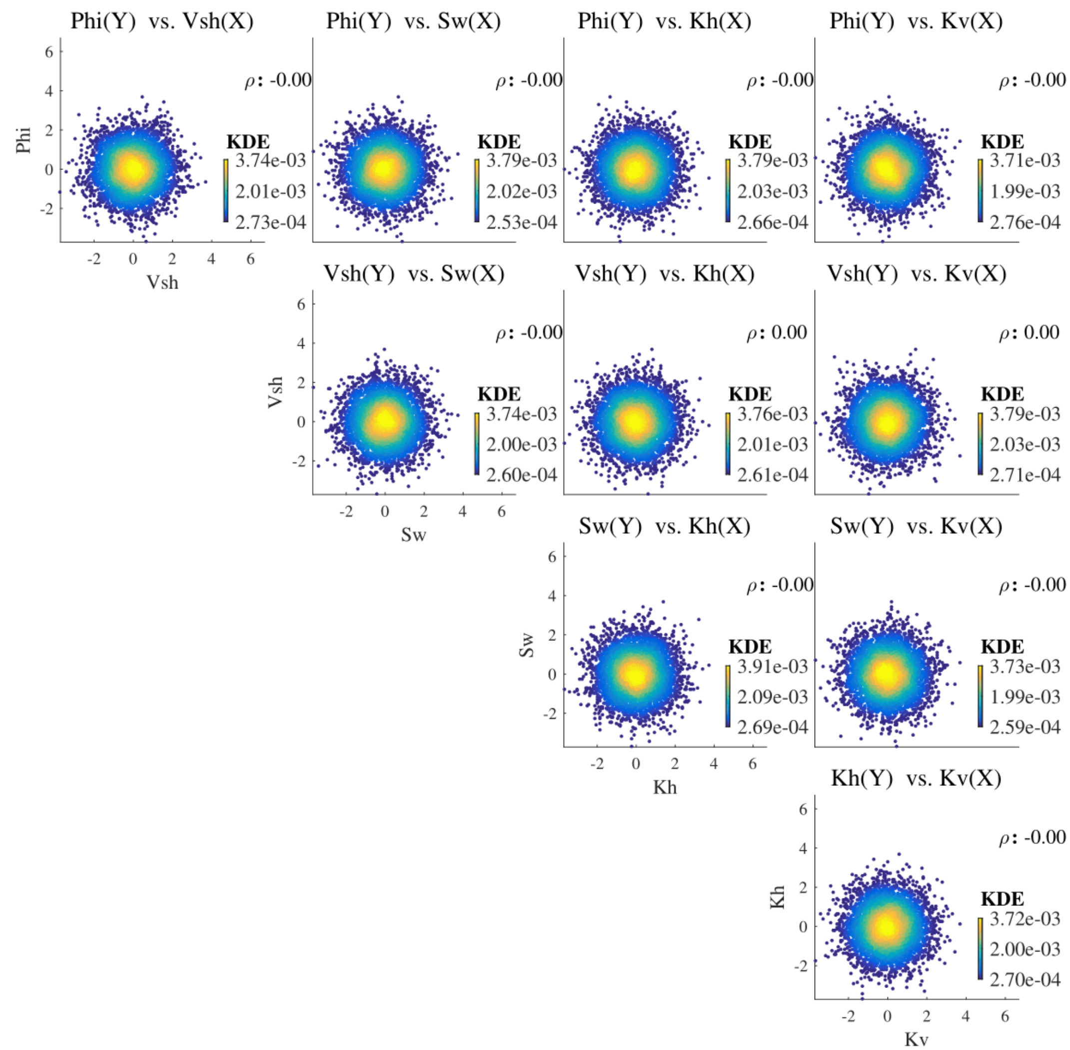


Figure 5: KDE scatterplots and correlation of the PPMT variables.

Conventional Workflow

The Conventional workflow is used to provide a relative benchmark for the PPMT workflow results. Aside from described differences of the two workflows, other modeling parameters are held constant to allow for a fair comparison. This includes the use of the same variogram modeling approach, grid definition and number of realizations. The Conventional workflow is summarized as:

1. Normal score transform ϕ, *Sw* and *Vsh*;
2. Perform conditional colocated cosimulation of ϕ, *Sw* and *Vsh*;
3. Return the ϕ, *Sw* and *Vsh* realizations to their original units using the normal score back-transformation;
4. Generate p-fields of *Kh* and *Kv* using independent and unconditional simulation;
5. Model the conditional bivariate CDFs F(*Kh* | ϕ) and F(*Kv* | *Kh*) using the discretized approach;
6. Perform the cloud transformation to generate realizations of *Kh*. The *Kh* p-fields from Step 4 are used to sample the F(*Kh* | ϕ) CDF from Step 5 conditional to the simulated ϕ from Step 3;

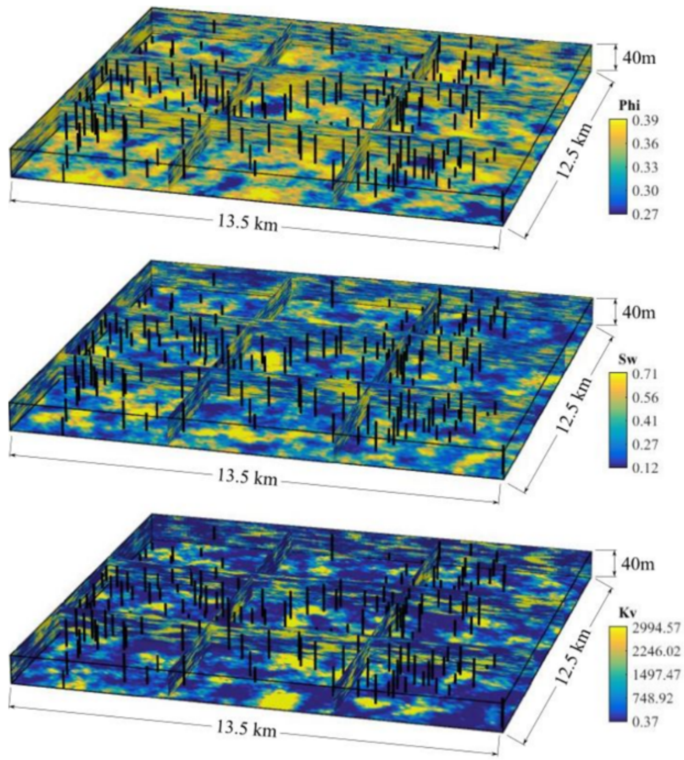


Figure 6: Slices of select variables for a realization that is generated by the PPMT workflow (20:1 vertical exaggeration).

7. Perform the cloud transformation to generate realizations of *Kv*. The *Kv* p-fields from Step 4 are used to sample the F(*Kv* | *Kh*) CDF from Step 5 conditional to the simulated *Kh* from Step 6.

Note that the Conventional workflow requires more steps than the PPMT workflow from the previous section. Even if the two workflows yielded similar results, the PPMT workflow would lend value to this modeling scenario since it requires fewer steps and associated effort. Observe from the original data scatterplots that the variables chosen for cosimulation, ϕ, *Sw*, and *Vsh* have relatively non-complex multivariate relationships between each other according to their KDE scatterplots.

Although these scatterplots do not follow ideal multiGaussian model contours, they are far less complex than the non-linear relationships that are observed in scatterplots that include *Kh* and *Kv*. This is the primary motivation for using sequential cloud transformations for the simulation of *Kh* and *Kv*. Despite the relative multiGaussian nature of the cosimulation variables, however, Figure 8 illustrates that any deviations from the multiGaussian model in original space will generally manifest themselves in normal score space. In this figure, KDE scatterplots of the normal score variables (output from Step 1) are compared to that of the simulated Gaussian realizations (output from Step 2). Observe that multivariate complexity remains between the normal score transform data that is not captured by the displayed correlation statistic. The colocated cosimulation only considers the correlation, leading to the displayed scatterplots that follow the typical multiGaussian contours. This creates obvious discrepancies between the density of the data and that of the realizations in Gaussian units, which will lead to similar issues in original space.

Skipping ahead to Step (5), the conditional CDFs F(*Kh* | ϕ) and F(*Kv* | *Kh*) that are input to cloud simulation are displayed in Figure 9. The large number of data permits the generation of smooth

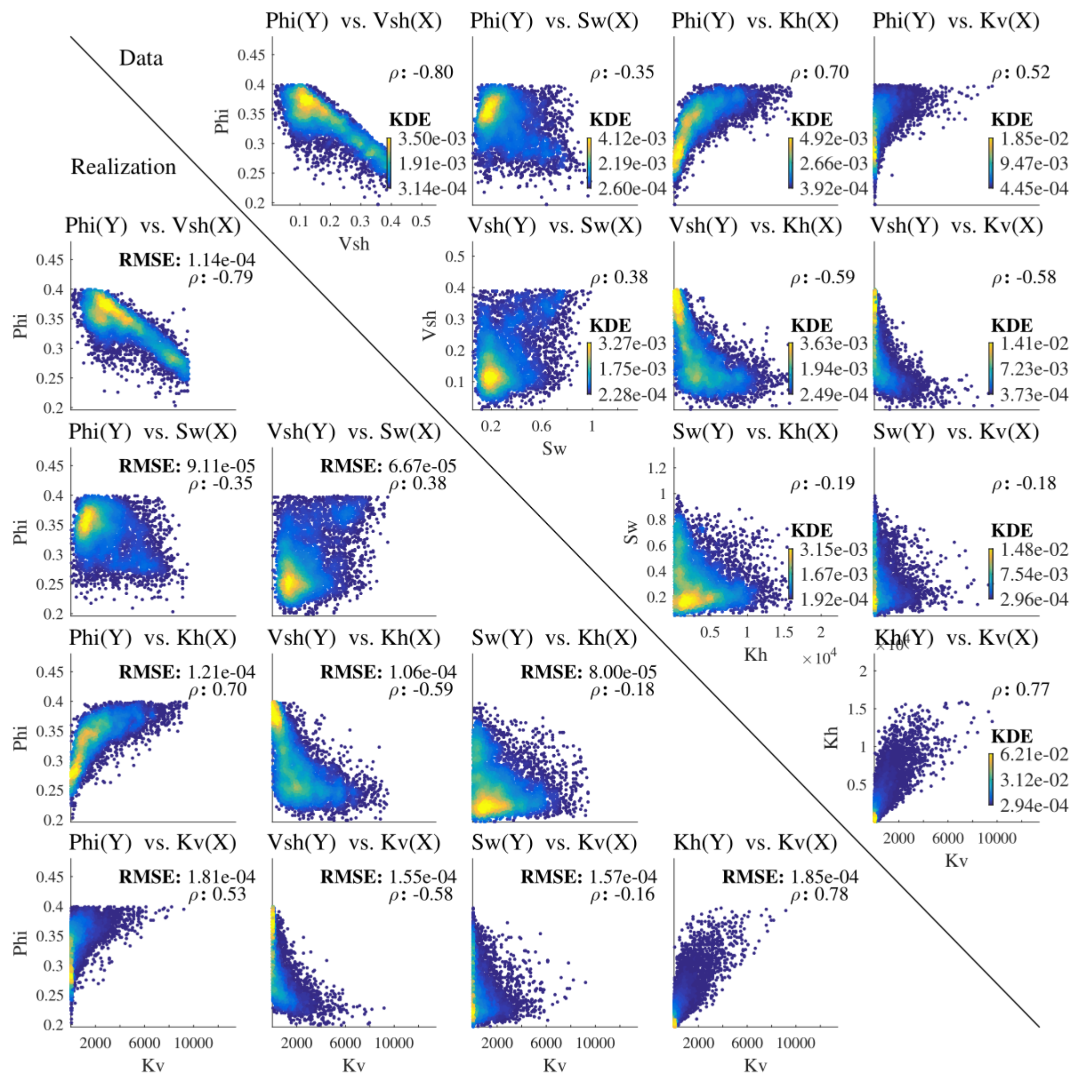


Figure 7: KDE scatterplots and correlation of one PPMT workflow realization (lower triangle), which are compared to that of the original data (upper triangle). The KDE coloring of the realization plots are scaled according to the presented color scale in the associated data plot.

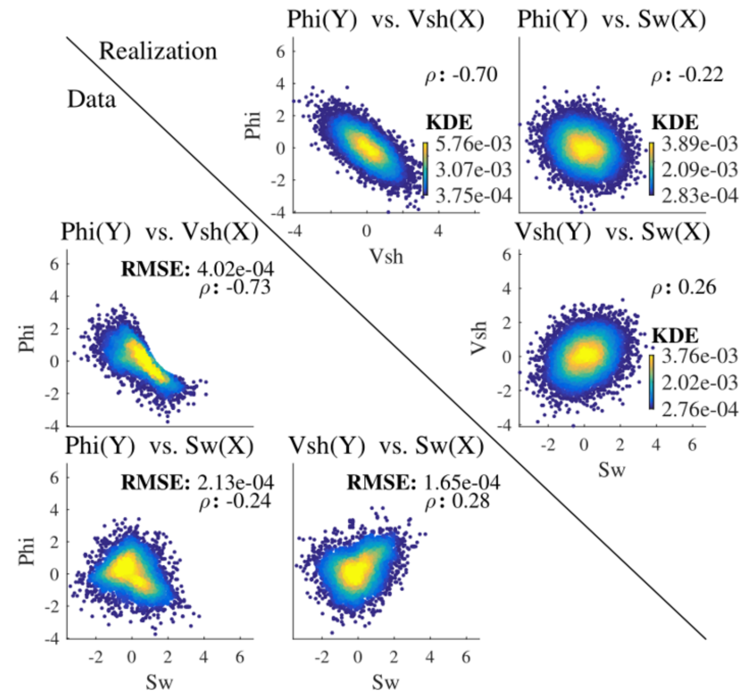


Figure 8: KDE scatterplots and correlation of one cosimulation realization (upper triangle), which are compared to that of the original data (lower triangle). The KDE coloring of the data plots are scaled according to the presented color scale in the associated realization plot.

and well informed conditional CDFs, which should allow for an effective application of the cloud transformation.

KDE scatterplots of the realizations (following completion of the Conventional workflow) are compared with that of the data in Figure 10. Based on visual inspection, KDE RMSE, and correlation error, it is readily apparent that the Conventional workflow has yielded inferior results, relative to the PPMT workflow. Observe that obvious discrepancies exist between the bivariate densities of the data and the equivalent densities of the Conventional workflow realization, whereas the PPMT realization KDE scatterplots are barely distinguishable from that of the data. The simple PPMT workflow implicitly targets the reproduction of the full multivariate relationship. By comparison, the Conventional workflow uses many sequential steps to target the covariance between select variables (ϕ, *Vsh* and *Sw*), and the complex relationships between select bivariate pairs (*Kh*-ϕ and *Kv*-*Kh*). As a result, the PPMT approach is demonstrated to generate realizations of the reservoir that more effectively reproduce multivariate properties that are important to SAGD performance prediction.

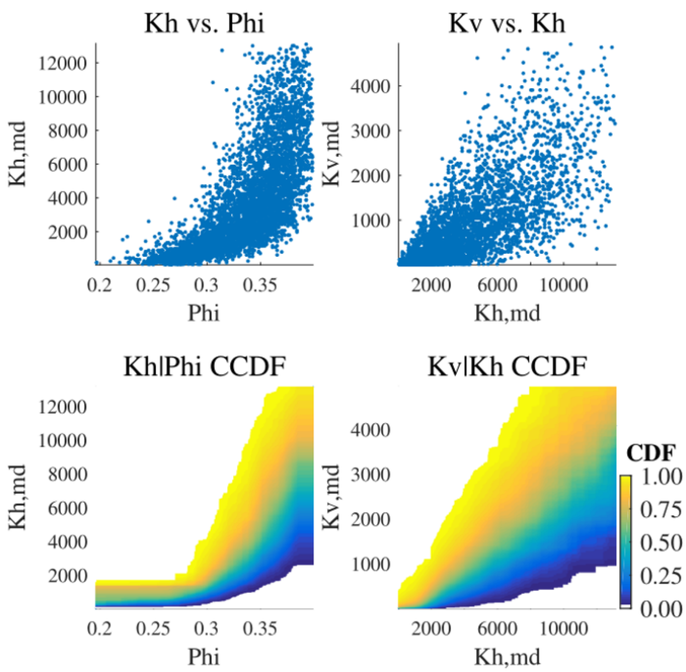


Figure 9: Scatter plot between *Kh* and ϕ (top left), with the associated conditional CDF F(*Kh* | ϕ) that is used for simulating *Kh* (bottom left). Similarly, scatter plot between *Kv* and *Kh* (top right), with the associated conditional CDF F(*Kv* | *Kh*) that is used for simulating *Kv* (bottom right).

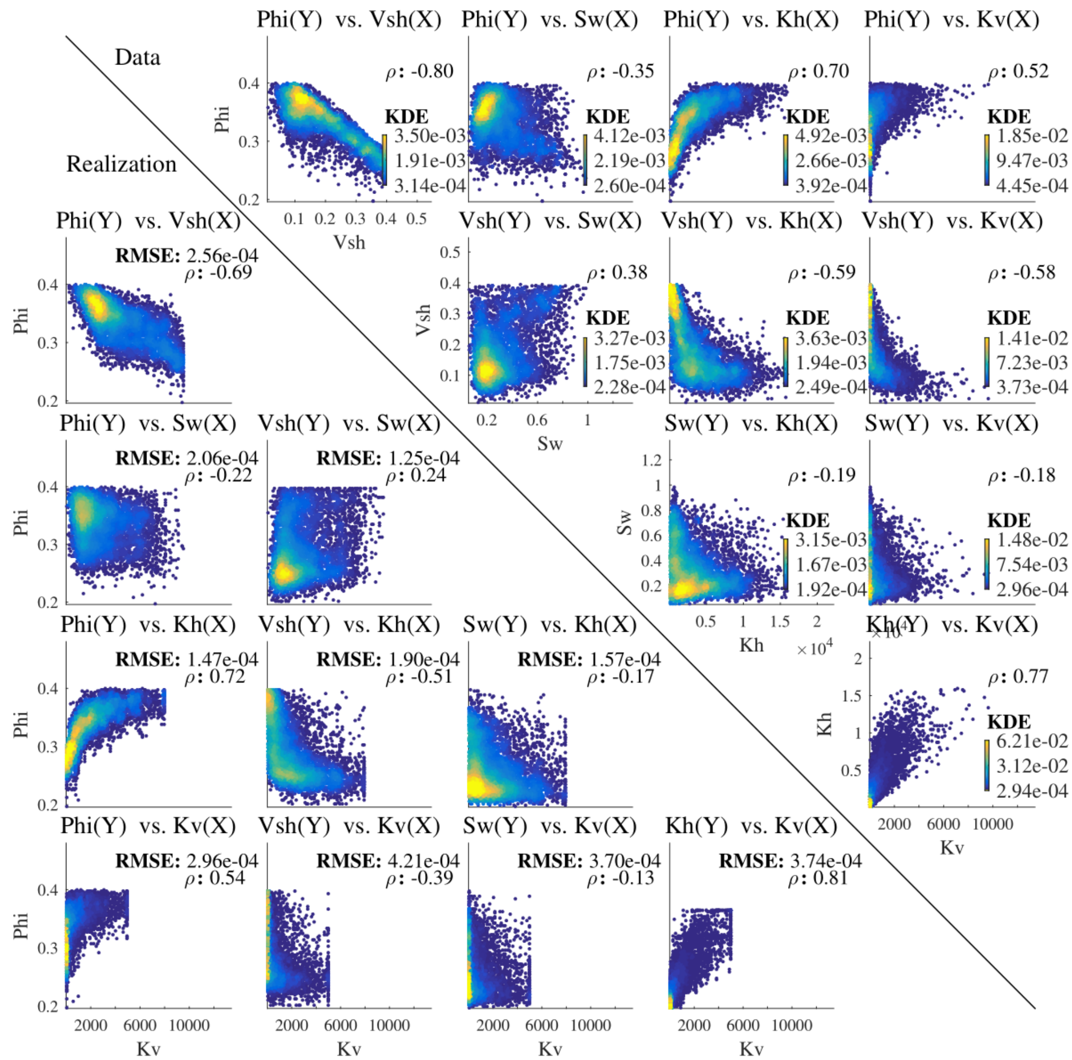


Figure 10: KDE scatterplots and correlation of one Conventional workflow realization (lower triangle), which are compared to that of the original data (upper triangle). The KDE coloring of the realization plots are scaled according to the presented color scale in the associated data plot.

References

Almeida, A. S., & Journel, A. G. (1994). Joint simulation of multiple-variables with a Markov-type coregionalization model. *Mathematical Geology*, *26*, 565--588.

Barnett, R. M., Manchuk, J. G., & Deutsch, C. V. (2014). Projection pursuit multivariate transformation. *Mathematical Geosciences*, *46*, 337--359.

Bashore, W., Araktingi, U., & Levy, M. (1994). Importance of a geological framework and seismic data integration for reservoir modeling and subsequent fluid-flow predictions. In J. M. Yarus & R. L. Chambers (Eds.), *AAPG computer applications in geology 3* (pp. 159--176).

Bliss, C. (1934). The method of probits. *Science*, *79*, 39--39.

Chiles, J. P., & Delfiner, P. (2012). *Modeling spatial uncertainty, 2nd edn.* (p. 734). Wiley, New York.

Denbina, E. (1998). SAGD comes of aGE! *Journal of Canadian Petroleum Technology*, *37*, 9--12.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB: A geostatistical software library and user's guide, 2nd edn.* New York: Oxford University Press.

Dull, D. (2004). Integrating sequence stratigraphy and multiple three-dimensional geostatistical realizations in constructing a model of the second eocene reservoir, wafra field, partitioned neutral zone, kuwait and saudi arabia. In G. M. Grammer, P. M. Harris, & G. P. Eberli (Eds.), *Integration of outcrop and modern analogs in reservoir modeling* (Vol. AAPG Memoir 80, pp. 309--336).

Edmunds, N., & Sugget, J. (1994). Design of a commercial sAGD heavy oil project. In J. M. Yarus & R. L. Chambers (Eds.), *International heavy oil symposium*. Calgary, Alberta.

Friedman, J. H. (1987). Exploratory projection pursuit. *Journal of the American Statistical Association*, *82*, 249--266.

Froidevaux, R. (1992). Probability field simulation. *Proc., Geostatistics, Troia, Kluwer*, *1*, 73--84.

Hadavand, M., & Deutsch, C. (2015). Improving sAGD reservoir modeling by geostatistical enforcement of anomalies in fluid flow, paper 203. In *CCG annual report 17*. University of Alberta, Edmonton.

Hwang, J., Lay, S., & Lippman, A. (1994). Nonparametric multivariate density estimation: A comparative study. *IEEE Transactions on Signal Processing*, *42*, 2795--2810.

Ma, Y., & Gomez, E. (2011). F. In *Frequentist meets spatialist: A marriage made in reservoir characterization and modeling* (p. 12p.). Denver, Colorado.

Moore, M., W.R. (2011). Uncertainty analysis in well-log and petrophysical interpretations. In P. La Pointe (Ed.), *Uncertainty analysis and reservoir modeling* (Vol. AAPG Memoir 96, pp. 17--28).

Pyrcz, M. J., & Deutsch, C. V. (2014). *Geostatistical reservoir modeling, 2nd edn.* (p. 448). Oxford University Press.

Thomas, R., Smith, D., & Wood, J. et a. (1998). Inclined heterolithic stratification-terminology, description, interpretation and significance. *Sedimentary Geology*, *53*, 123--179.

Verly, G. (1983). The multiGaussian approach and its applications to the estimation of local reserves. *Mathematical Geology*, *15*, 249--286.

Who We Are

The CCG was launched by Professor Clayton V. Deutsch with the vision of becoming a leader in the education of geostatisticians and the delivery of geostatistical tools for modeling heterogeneity and uncertainty. The main objective of the CCG is to support the mutual needs of industry and academia in research and education. The benefits to industry include the opportunity to influence geostatistical research and education, interaction with students as potential employees, early access to publications and access to faculty members for discussions and presentations. The CCG provides a mechanism for industry to contribute to and sustain geostatistical research and teaching, which is of long term interest to many companies.

Contact Us

For more information regarding the demonstrated Solution or to discuss another problem that your project presents, please contact Professor Clayton V. Deutsch at: <cdeutsch@ualberta.ca>
Or drop by our offices at:

Centre for Computational Geostatistics
6-247 Donadeo Innovation Centre For Engineering
9211-116 Street, University of Alberta
Edmonton, Alberta, Canada T6G 1H9

Citation

Barnett, Ryan M. (2016) CCG Solutions - Multivariate Reservoir Property Modeling. Centre for Computational Geostatistics, University of Alberta. Retrieved from http://www.ccgalberta.com/solutions/