JCPT

DISTINGUISHED AUTHORS SERIES

# What in the Reservoir is Geostatistics Good For?

C.V. DEUTSCH
University of Alberta

**Clayton Deutsch** is a professor in the Department of Civil and Environmental Engineering at the University of Alberta. Prior to joining the University of Alberta, Dr. Deutsch was an associate professor (research) in the Department of Petroleum Engineering at Stanford University. Dr. Deutsch has published two books, over 45 peer-reviewed technical papers, and over 40 papers in conference proceedings.

## Abstract

Geostatistics provokes strong reactions. There are champions who believe the application of geostatistics adds value in almost any reservoir modelling situation. There are skeptics who do not think that a geostatistical model will have a meaningful impact on reservoir management decisions. The majority of engineers and geoscientists, however, are seeing an increasing use of geostatistics and are not sure when geostatistics should be used and how the results affect reservoir decisions.

There are three specific cases where geostatistics can provide valuable support for decision-making: 1) calculating maps of uncertainty over large areas to support resource calculations and well placement; 2) reconciling well and seismic data into high resolution reservoir models; and, 3) constructing representative models of heterogeneity to provide input to flow simulation and support reservoir forecasting. These three cases are developed without excessive theoretical detail. Realistic examples are presented without getting lost in the details of a particular reservoir. Limitations and pitfalls are discussed.

## Framework of Geostatistics

*Geostatistics* refers to the theory of regionalized variables and the related techniques that are used to predict variables such as rock properties at unsampled locations. Matheron formalized this theory in the early 1960s[1]. Geostatistics was not developed as a theory in search of practical problems. On the contrary, development was driven by engineers and geologists faced with real problems. They were searching for a consistent set of numerical tools that would help them address real problems such as ore reserve estimation, reservoir performance forecasting, and environmental site characterization. Reasons for seeking such comprehensive technology included: 1) an increasing number of data to deal with; 2) a greater diversity of available data at different scales and levels of precision; 3) a need to address problems with consistent and reproducible methods; 4) a belief that improved numerical models should be possible by exploiting computational and mathematical developments in related scientific disciplines; and, 5) a belief that more responsible decisions would be made with improved numerical models. These reasons explain the continued expansion of the theory and practice of geostatistics. Problems in mining, such as unbiased estimation of recoverable reserves, initially drove the development of geostatistics. Problems in petroleum, such as realistic heterogeneity models for unbiased flow predictions, were dominant from the mid 1980s through the late 1990s. More recently, the problems of realistic geologic modelling and reliable uncertainty quantification are driving development.

The main focus of geostatistics is constructing high-resolution 3D models of categorical variables, such as facies, and continuous variables, such as porosity and permeability. It is necessary to have *hard* truth measurements at some volumetric scale. All other data types including geophysical data are considered *soft* data and must be calibrated to the hard data. It is neither possible nor optimal to construct models at the resolution of the hard data. Models are generated at some intermediate geological modelling scale, and then scaled to an even coarser resolution for resource calculation or flow simulation. A common goal of geostatistics is the creation of detailed numerical 3D geologic models that simultaneously account for a wide range of relevant data of varying degrees of resolution, quality, and certainty. Much of geostatistics relates to data calibration and reconciling data types at different scales.

Geostatistical modelling requires spatial statistical control on the nature of the variability. Often, however, there are often insufficient data to provide reliable statistics. For this reason, data from reservoir analogues are used to help infer spatial statistics that are impossible to calculate from the available data. There are general features of each geological setting that can be transported to other reservoirs of similar geological setting.

The conceptual framework of geostatistics starts from an admission that the true variability of reservoir properties is important, but will never be accessible. Numerical tools are used to create numerical models that mimic the patterns of variability that we believe exist in the reservoir. These models are rarely based on depositional or diagenetic principles, but they are useful to appreciate variability and the consequent uncertainty.

## Case One: Mapping Uncertainty

We are often interested in mapping over a large area with sparse well control, a number of seismic variables, and some conceptual geologic maps. One goal is to predict the best values of reservoir variables, such as storativity ($\phi h$) and transmissibility ($kh$), and risk qualified values, such as the $P_{90}$ and $P_{10}$ estimates. Geostatistics does not help much in a purely exploration setting; good geologic sense and past experience are required. A requirement of geostatistics is enough data to calculate correlations and perform

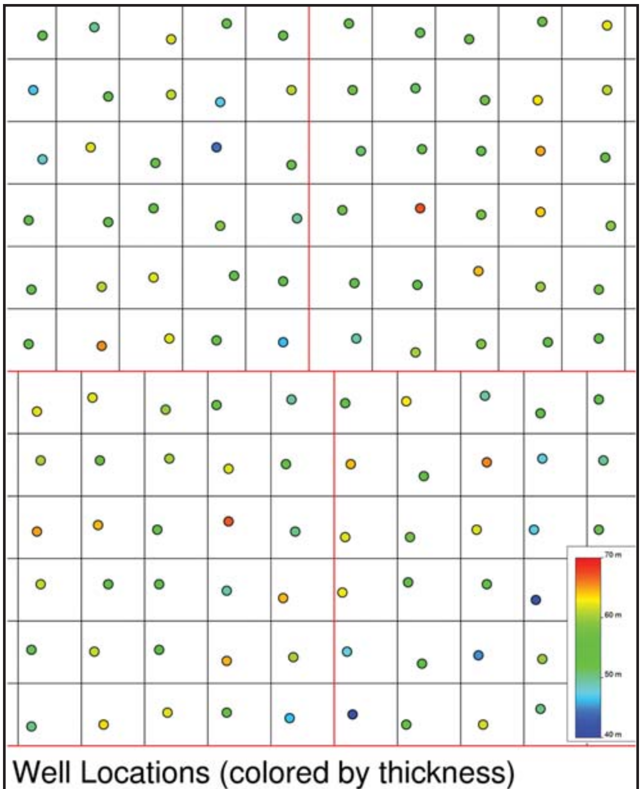


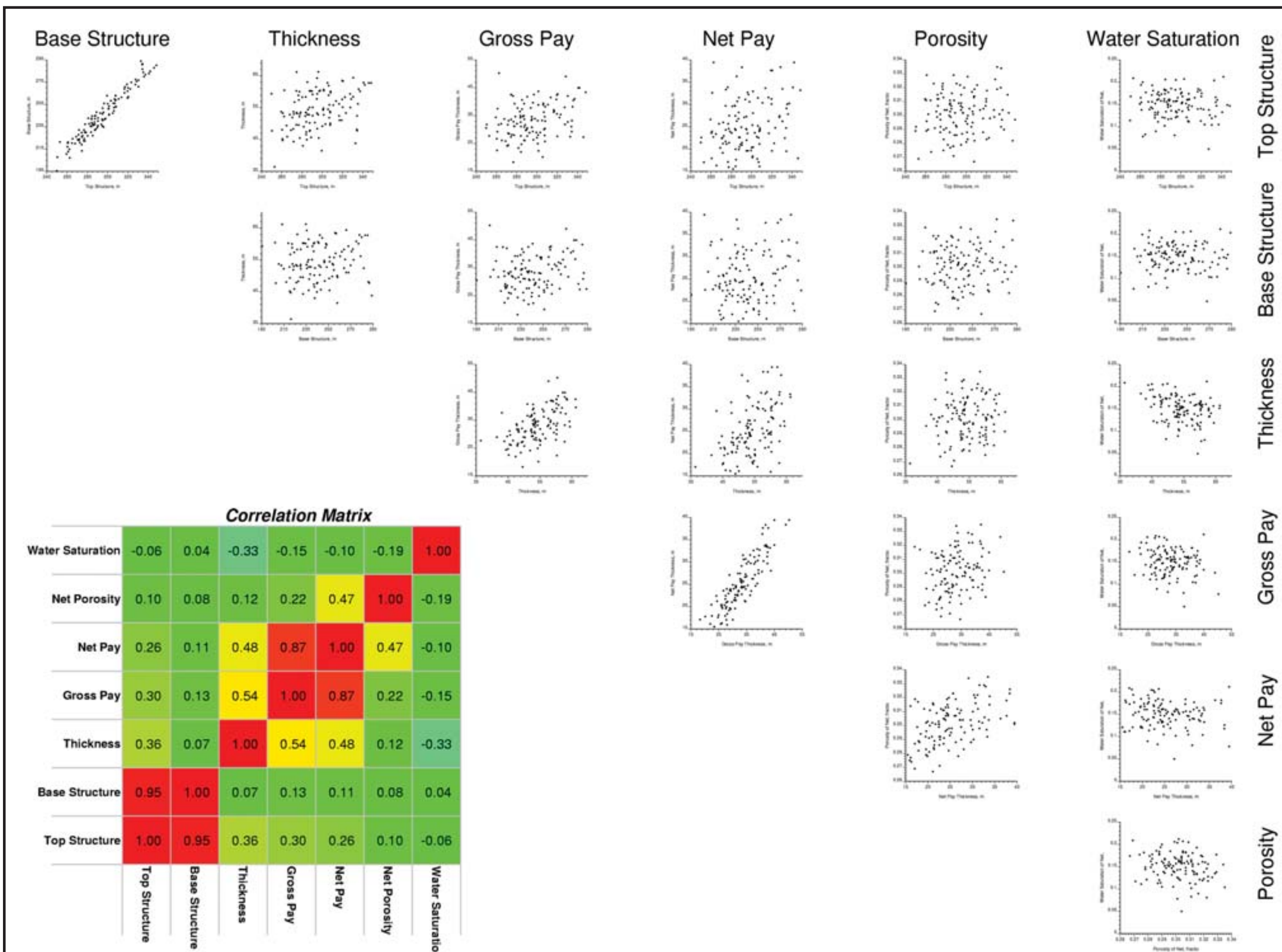

**FIGURE 1: Location map of 117 wells for Case Study One. The well locations are coloured by gross reservoir thickness in metres.**

statistical analysis; in the context of this example, a minimum of 2 – 10 wells are required.

The example presented here is fashioned after a number of real examples. Consider the 117 wells shown on Figure 1. The area is about 16 km by 19 km. The top reservoir structure, the base reservoir structure, and the thickness (which is directly derived from the top and base structure) will be considered as independent secondary variables. There are four variables we must map with uncertainty: gross pay thickness in the reservoir thickness, net pay thickness, net porosity, and water saturation in the net reservoir. The relationships between these variables are shown on Figure 2. The cross plots show one point per well. The cross plots are essentially unreadable; however, we look at them for strange values that do not belong, non-linear trends relationships, and other features. These cross plots do not reveal anything unusual. The variables were transformed to Gaussian units and the matrix of correlation coefficients was calculated (see the lower left table on Figure 2). Mapping of the four variables of interest should respect these correlations.

Mapping also requires measures of spatial correlation. The isotropic variograms for the Gaussian transform of the four variables are shown on Figure 3. The lack of short scale information is typical. Experience is required to infer $\gamma(h)$ for small h. The intercept near $h = 0$ is low for most reservoir variables. ~~As before,~~ the required covariance values $C(h)$ are calculated as $1 - \gamma(h)$.

There are a variety of geostatistical techniques to simultaneously account for multiple variables. A practical and robust Bayesian Updating technique will be illustrated here. The method is theoretically equivalent to the common collocated cokriging implementation[2] with a Bayesian interpretation[3]. As with virtually all geostatistical techniques, the data variables are transformed to be Gaussian, all




| | Top Structure | Base Structure | Thickness | Gross Pay | Net Pay | Net Porosity | Water Saturation |
|---|---|---|---|---|---|---|---|
| Water Saturation | -0.06 | 0.04 | -0.33 | -0.15 | -0.10 | -0.19 | 1.00 |
| Net Porosity | 0.10 | 0.08 | 0.12 | 0.22 | 0.47 | 1.00 | -0.19 |
| Net Pay | 0.26 | 0.11 | 0.48 | 0.87 | 1.00 | 0.47 | -0.10 |
| Gross Pay | 0.30 | 0.13 | 0.54 | 1.00 | 0.87 | 0.22 | -0.15 |
| Thickness | 0.36 | 0.07 | 1.00 | 0.54 | 0.48 | 0.12 | -0.33 |
| Base Structure | 0.95 | 1.00 | 0.07 | 0.13 | 0.11 | 0.08 | 0.04 |
| Top Structure | 1.00 | 0.95 | 0.36 | 0.30 | 0.26 | 0.10 | -0.06 |

**FIGURE 2: Bivariate summary of multiple variables. The cross plots are shown in the upper right and the matrix of correlation coefficients is shown in the lower left.**

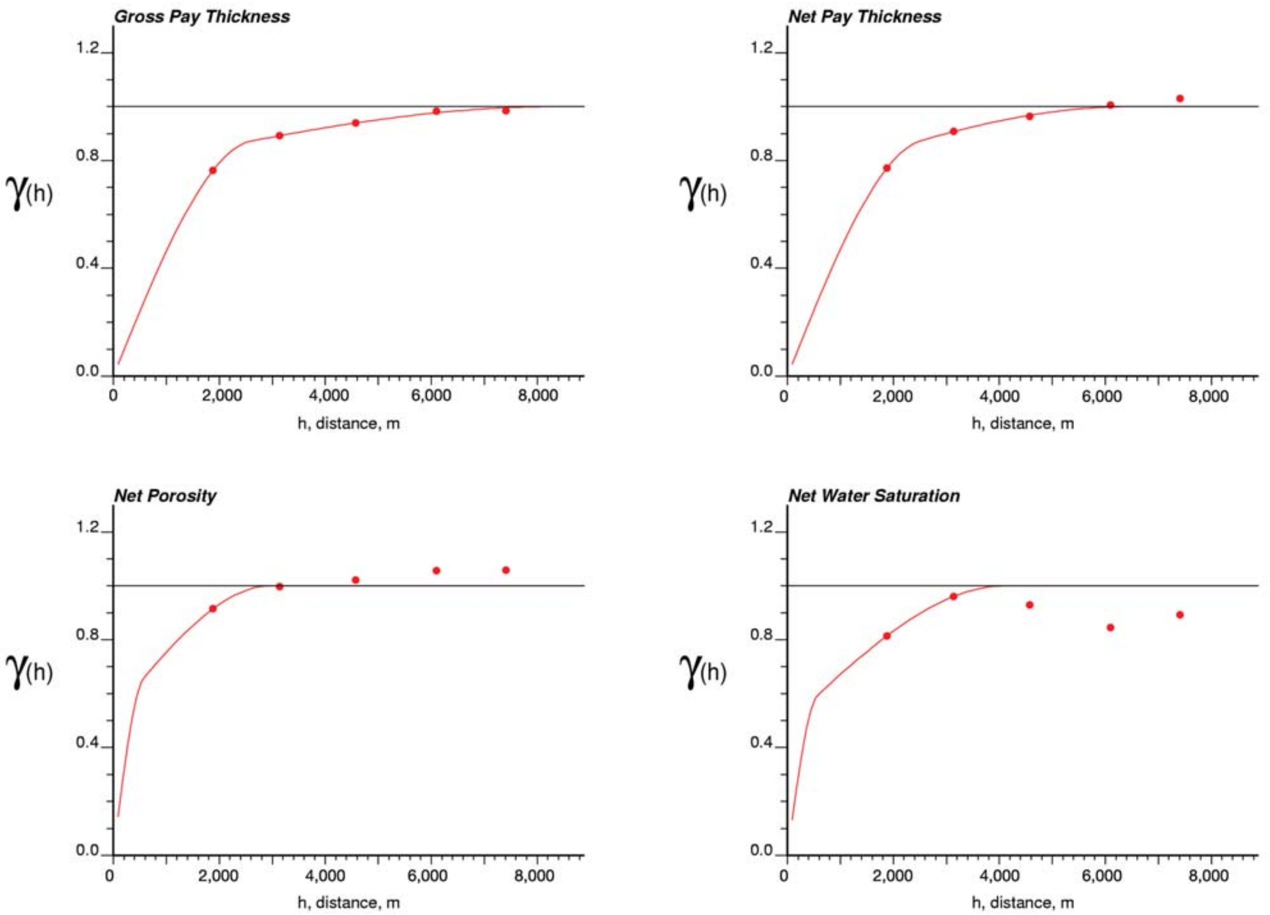



**FIGURE 3: Variograms of the four variables under consideration. The dots are the calculated points and the solid line is the fitted curve. These variograms are isotropic.**

calculations are performed in Gaussian units, and the results are back transformed.

Figure 4 shows maps of the $P_{90}$ low values and the $P_{10}$ high values for each variable. The $P_{90}$ low value map is used to identify areas that are surely high; where the $P_{90}$ low value is high, then the variable is surely high. The $P_{10}$ high value map is used to identify areas that are surely low; where the $P_{10}$ high value is low, then the variable is surely low. P50 maps are not significantly different from maps of local mean values. These maps can be used for planning well locations.

Geostatistical simulation techniques could be applied to assess joint uncertainty. Simulation must be conducted jointly in space and jointly between multiple variables. There are a number of cosimulation techniques for this purpose. Assessing uncertainty over a large area with multiple variables is an important goal in petroleum geostatistics. Constructing detailed 3D models that reproduce all of the available data is another important goal.

## Case Two: Data Integration

A reservoir model is required for resource assessment and to help with devising a reservoir management strategy. There is almost always seismic data and a limited number of well data. These data must be integrated under a sound conceptual geologic model. A small example leading to optimization-based well placement is developed.

Figure 5 shows a map of a seismic attribute and two wells: one well is in a higher quality area and the other is in a poorer quality area. There are a number of algorithms to simulate the variability of facies and petrophysical properties. Illustrations of facies modelling techniques are deferred to the next example. The porosity and permeability for this example are simulated using sequential Gaussian simulation (SGS). The SGS algorithm is an industry standard algorithm that assumes a multivariate Gaussian distribution between all variables (seismic, porosity, and permeability) and all locations, after each variable is transformed to have a Gaussian histogram. A unique feature of the multivariate Gaussian distribution is the sole requirement for covariances or variogram values between all variables and locations. The variogram for the seismic attribute ~~(recall Figure 5)~~ is not a problem since it is available (virtually) everywhere. The variograms of porosity and permeability are more of a problem.

The vertical variogram of porosity is well established by just two well data; however, it is not possible to calculate a horizontal variogram from the well data. The horizontal variogram is calculated from the seismic data. The normalized variograms are shown in Figure 6. There is a strong assumption that the seismic attribute is providing a good measure of continuity for porosity. This seems reasonable in this case given the close correspondence between the well average porosity and the collocated seismic value. Analogue data would be used if this was deemed unacceptable (see Chapter 4 of *Geostatistical Reservoir Modelling*). Details of permeability modelling will not be shown here because of space constraints. A conventional SGS cosimulation approach was used.

An additional parameter is required to use the seismic—the correlation between the seismic attribute (at a relatively coarse vertical resolution) and the porosity at a small scale. A correlation coefficient of 0.5 was chosen for the modelling. This led to a correlation of 0.8 between the vertically averaged porosity and the seismic attribute, which is considered reasonable.

A total of 50 realizations were created in flattened stratigraphic coordinates. The spatial variability away from the well locations looks natural relative to the well locations. A variety of uncertainty maps could be created. The realizations could be ranked by increasing oil in place and selected quantiles, say the $P_{10}$, $P_{50}$, and $P_{90}$ used in flow simulation. These applications are common

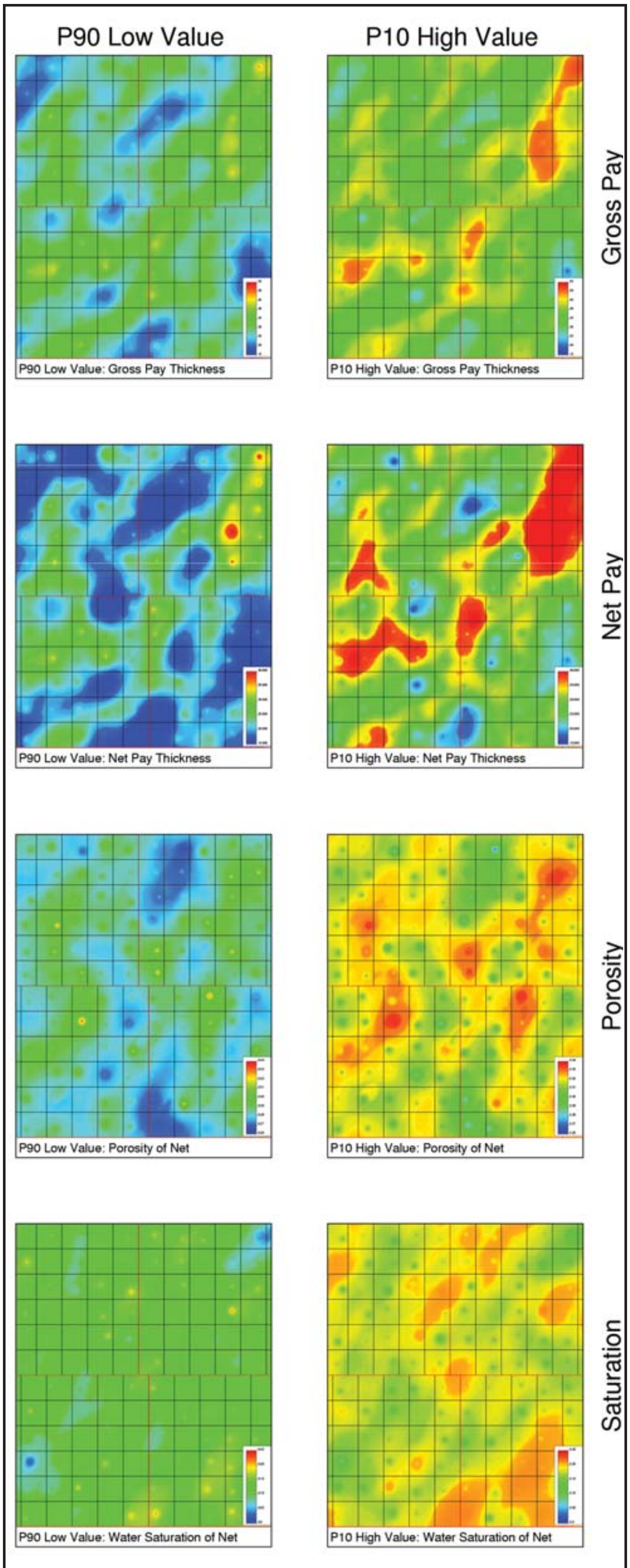



**FIGURE 4: P90 low and P10 high values for each variable being predicted. The maps are in the units of the original data variables.**

practice in petroleum geostatistics. A slightly different application will be demonstrated here.

The placement and timing of production and injection wells is a significant decision in reservoir development planning. Well placement is difficult with a single deterministic model of reservoir structure and petrophysical properties. The duration of plateau production must be maximized, water handling must be minimized, recovery should be maximized, and key economic indicators should also be optimized. These response variables depend on complex nonlinear ways on the placement/timing of the wells, their operating conditions, and the subsurface reservoir description. Many of the interactions are resolved by a combination of sound engineering judgment and flow simulation. The problems associated with well placement become more complex in the presence of multiple geostatistical realizations. The number of flow simulation runs becomes intractable and it becomes impossible to visualize all possibilities. There is a need for numerical measures to assist in the optimization of well locations to minimize risk and maximize reservoir performance.

The conventional approach is to consider a few reasonable well plans and perform some limited perturbations/optimization. This resulting solution is likely to be close to the optimal because experienced reservoir engineers can quickly rule out many bad well configurations. Nevertheless, there is still room for improving the well plan by considering an optimization scheme. The consequences of a minor improvement in the well plan are significant; an incremental improvement in recovery translates to a large monetary value. It is impossible to run flow simulation for many well configurations and geological models. Thus, the following idea is proposed:

1. Run some flow simulations with a number of different well configurations and different geological models. The optimization will be better with more flow simulations; 20 different models would provide a starting point. Summary flow response variables, such as discounted value of the production, are calculated from each flow simulation;
2. Propose static reservoir quality measures that capture the local goodness of the reservoir, for example, connected pore volume discounted by distance and permeability;
3. Calculate the static measures on the models used for flow simulation in Step One. Calibrate the static reservoir measures to the flow response variables using classical multivariate statistical tools. The result is a static reservoir proxy for flow simulation that can be optimized very quickly;
4. Optimize a set of well locations that maximize the calibrated measure of static reservoir goodness over all realizations accounting for their probability. This optimization can be repeated considering changes to the number of wells, initial configuration, and weighting of different factors; and,
5. Validate the results by performing flow simulation on the optimal configuration and reasonable alternatives proposed by the reservoir engineer before (and after) the optimization.

As with all numerical short cuts, there are limitations including: 1) it is difficult to capture the effect of timing and the incremental knowledge gained during the drilling program; 2) the physics of flow are not accounted for directly in the optimization; 3) the location of injectors and the pore volume replaced by water injection are not accounted for; and, 4) the specifics of well completion are not accounted for. Notwithstanding the long list of limitations, there is significant value in the optimization. Attention is focused on what makes a well plan good and important features of the reservoir are better understood. Application to a number of reservoir studies indicates 1 – 5% improvements in hydrocarbon recovery.

Returning to the example, a number of flow simulations were performed with different cases. Three different well patterns were chosen: square, five-spot, and random. Three realizations were chosen: $P_{10}$, $P_{50}$, and $P_{90}$. Three different numbers of vertical wells were chosen: 10, 25, and 35. These decisions were made arbitrarily for the purpose of this example. In practice, the choices are made by the reservoir engineers with reservoir-specific considerations. The field oil production rates were discounted by 10% per year and a cumulative discounted oil production was calculated. A single number, summarizing the flow simulation results, characterizes each realization. Twenty-seven runs were used for calibration (based on three geostatistical realizations). ~~The $\omega_d$ and $\omega_k$ parameters were optimized to be approximately 1.0 and 0.7, respectively.~~ Figure 7 shows the results. The calibrated static measure very closely predicts the actual flow response. The results are robust in the sense that minor changes in the calibration parameters do not significantly change this excellent correlation.

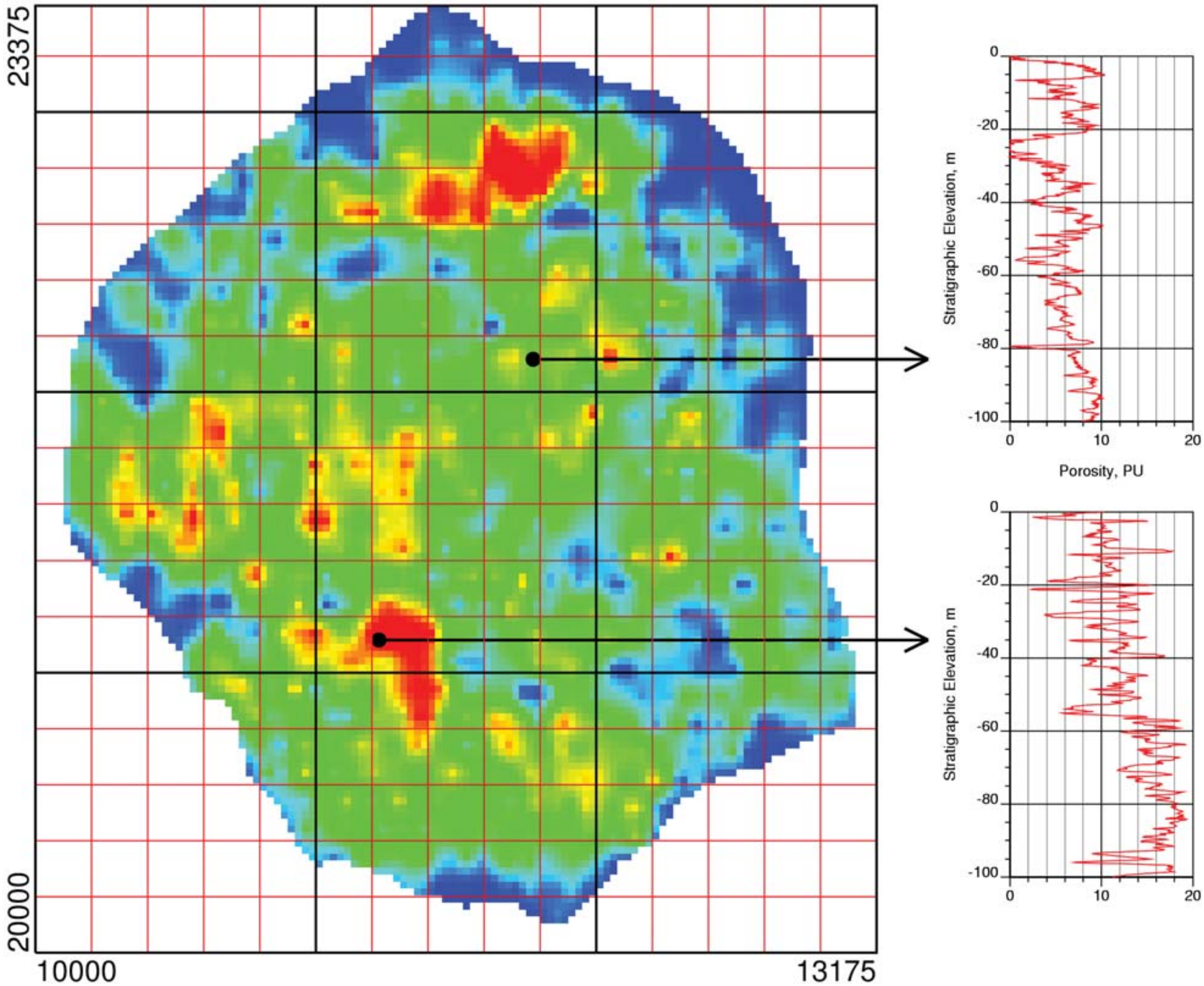



**FIGURE 5: Seismic data attribute and two wells for Case Study Two. The bright colors on the seismic map are high porosity and the dark colours are lower.**

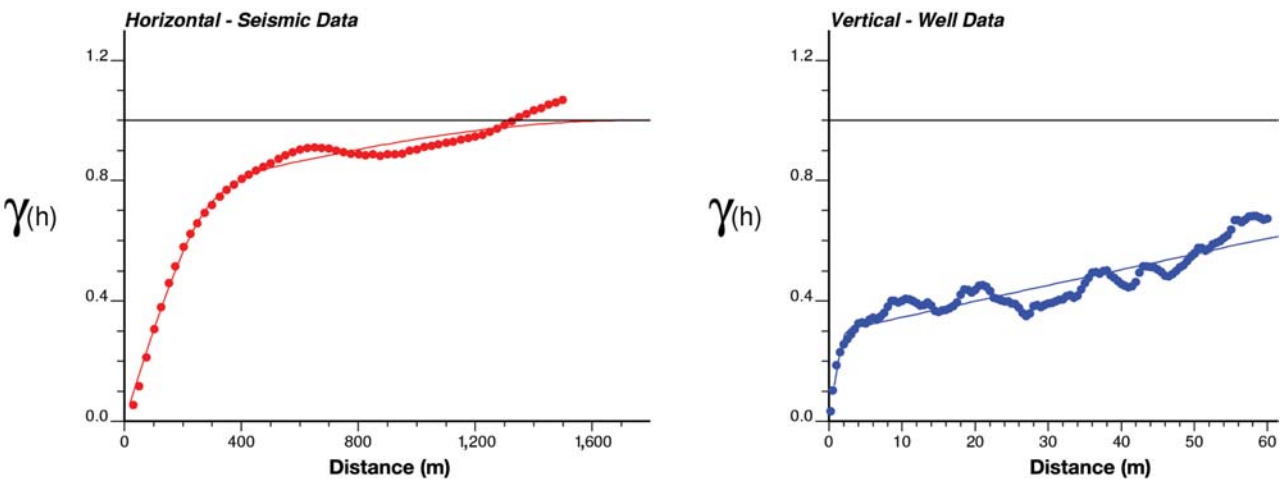



**FIGURE 6: Horizontal and vertical variograms from the seismic and well data, respectively. The fitted variogram model to both simultaneously.**

Optimization was then undertaken with 10, 20, 30, and 40 wells. The results are shown on Figure 8. The underlying colour scale map is one of hydrocarbon thickness—red is high and blue is low. This measure is closely related to the static quality defined above. An initial choice of well locations is iteratively perturbed to maximize the goodness of the well plan. The objective function is to maximize the goodness of the well plan, in terms of the static quality measure, in expected value over all realizations. Each realization is considered equally probable; however, a subjective weight could be assigned to favour specific realizations. The optimized well locations can be tweaked to account for complex considerations not easily coded in an optimization algorithm. Flow simulation with the optimized locations is required to verify that optimizing the static quality measure also improves the actual production.

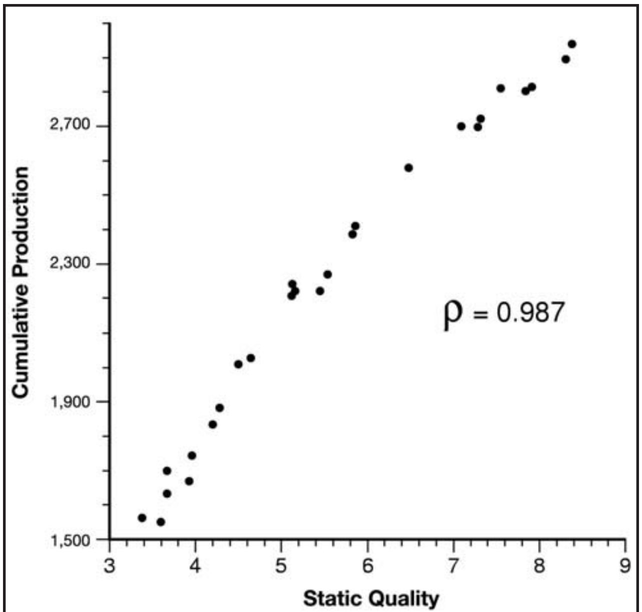



**FIGURE 7: Calibration between calibrated/optimized static reservoir quality and discounted cumulative oil production.**

The optimized well locations may not perform significantly better than those chosen by a thoughtful engineer; however, the advantages are the ability to optimize over a set of alternative geostatistical realizations with a repeatable quantifiable measure of goodness. There will always be multiple realizations when a development plan is being established; the uncertainty will diminish as additional wells are drilled, but it will not go away. Optimizing over multiple realizations and achieving well locations that are robust over the entire space of uncertainty is desirable. There may be a need to customize the optimization to adapt to reservoir-specific factors such as depletion/displacement mechanism.

This example highlights two points: 1) constructing reservoir models that integrate multiple types of data; and, 2) using models of uncertainty for decision-making.

Assessing uncertainty over large areas and constructing detailed 3D models that reproduce all of the available data are important goals in petroleum geostatistics. Yet another goal is the assessment of the impact of heterogeneity on flow performance. Running flow simulation on realistic high resolution geostatistical models is a worthy goal.

## Case Three: Models of Heterogeneity

A reservoir model is required input to flow simulation for forecasting and recovery calculations. Although flow simulation is performed in a minority of reservoirs, ~~virtually all~~ major developments rely heavily on flow simulation for planning and forecasting. There is almost always seismic data, a limited number of well data, and some flow testing. Geostatistics does not work well if there are a large number of wells with extensive production history; the available techniques do not reliably account for extensive historical production data in the geostatistical reservoir models. There is a critical time, however, after expensive exploration and before significant production when a good geostatistical reservoir model combined with flow simulation can answer important questions and facilitate decision-making.

The large-scale structure of a reservoir including volumetrics and compartmentalization has the greatest effect on reservoir performance. The connectivity of high permeability pathways and low permeability barriers, however, can have a large effect on dynamic performance. The flow behaviour of simplistic layercake models and an interpolated (kriged) model are not even in the range of uncertainty from geostatistical simulations. The effect of heterogeneity on flow predictions is variable.

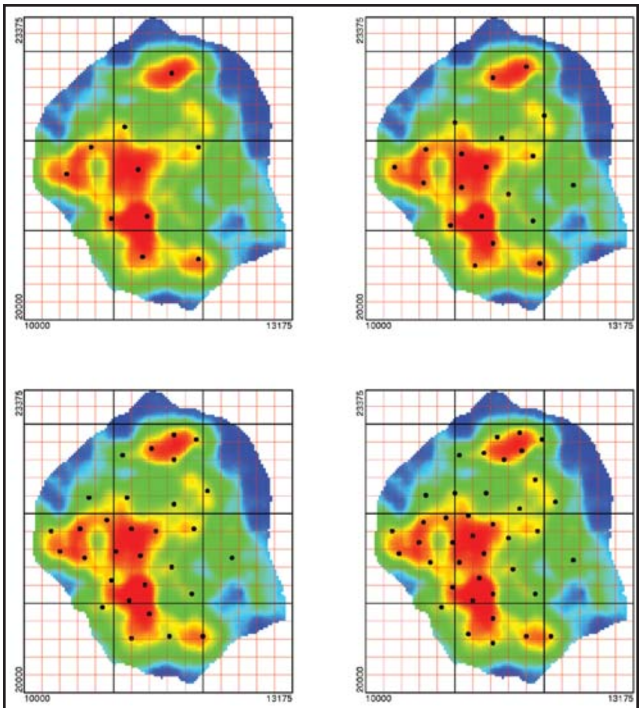



**FIGURE 8: Optimized well locations for 10, 20, 30, and 40 wells. The underlying colour scale map is the expected oil column thickness (average over all 50 realizations).**

There are times when heterogeneity has a dominant effect on flow performance and there are times when the heterogeneity averages out and the model behaves in a *heterogeneously homogeneous* manner. Running the flow simulator with a number of candidate heterogeneous models will establish the importance for a particular reservoir. There are a number of geostatistical tools aimed at creating heterogeneous models. Most geostatistical practitioners agree that the major heterogeneity is captured in facies models. Facies are modelled as categorical variables by a variety of techniques based on cell, object, process-mimicking and multiple point statistics. Continuous properties like porosity and permeability are assigned once facies realizations have been constructed.

Cell-based techniques are based on statistical controls such as indicator variograms. A sequential simulation approach is commonly followed where the cells are assigned a facies in a sequential approach considering the well data, seismic data, and previously assigned cells. Cell-based methods capture the coarse features, but they do not always appear realistic. Object-based techniques proceed by filling a model with facies objects. The shape, size, orientation, and relationships between objects are chosen to appear realistic and match the available data.

It is becoming increasingly common to model facies with numerical techniques that have been adapted to mimic complex features of the depositional and diagenetic processes. These techniques are often referred to as process-mimicking techniques. They are hybrids of object-based modelling, cell-based modelling, and depositional processes(4).

Geostatistics makes it relatively easy to generate multiple realizations. Most of the effort goes into setting up a reasonable workflow and establishing the required parameters. It is not possible, however, to process a large number of realizations through flow simulation. The computer requirements are prohibitive. It is common to rank the realizations from low to high based on a measure of static quality. The static quality could be the same as that used for well optimization. More complex ranking measures could be used; however, the calibration shown on Figure 6 is hard to improve upon. It is common to take the 10th, 50th, and 90th percentile realizations and process them through flow simulation.

Heterogeneity is modelled by a set of realizations generated by a variety of statistical techniques. The realizations are often ranked to limit the number of realizations that have to be used in flow simulation. Geostatistical techniques generate models at the small scale of the well data; they are scaled up to more accurately reflect the heterogeneity at the flow simulation scale. Assessing the effect of heterogeneity on recovery and flow performance is an important application of geostatistics.

## When to Avoid Geostatistics

There are some inescapable realities of reservoir modelling: 1) data cleaning, formatting, and checking take much longer than anyone believes possible; 2) bad data lead to a bad model regardless of the techniques used; 3) all numerical models depend on an appreciation of the features being modelled; understanding the geology is critical; and, 4) reasonable models are only obtained if they are built for a particular purpose; there is no universal model or modelling approach. There are some additional sources for concern with geostatistical techniques.

There must be ground truth hard data for model construction. An essential feature of geostatistics is the calibration of extensive soft data with hard data, which results in an assessment of uncertainty at unsampled locations. Geostatistics should not be used when there are no hard data measurements of the variables being modelled.

Another essential feature of geostatistics is the exploitation of multivariate and spatial structure in measurements. At times, there is too little data or too widely spaced data to observe any structure. Perhaps this is the time when geostatistics is most needed to provide a quantitative measure of uncertainty; however, the results will be entirely model driven. Geostatistics should be avoided when there is too little data to make a meaningful decision of statistical populations and probabilistic predictions.

Geostatistical tools are inherently statistical—no physics or process information is embedded in the prediction. It is a mistake to apply geostatistics directly to variables such as pressure, flow rate, or reservoir production particularly if the wells interact with each other. Geostatistics should be used to construct the input static models. Flow simulation, or some suitable process model, should be used for the dynamic predictions.

Geostatistics should not be considered when there is too little time and expertise to effectively apply the techniques and validate the resulting models. Geostatistical techniques are time consuming and finicky to apply. Slick demos by software vendors have not helped; they misrepresent the amount of time it really takes to construct a verified and useful set of geostatistical realizations. It takes time to establish a reservoir-specific workflow, choose reasonable modelling parameters, undertake reasonable sensitivity studies, verify the results, and apply the models to the problem at hand.

# Conclusions

Geostatistical reservoir models are useful to transfer uncertainty in geological parameters through process evaluation to output uncertainty. Basic tools of the Monte Carlo simulation paradigm are adapted to spatially correlated variables. They are often applied hierarchically in an attempt to capture geologic structures and reproduce all of the available data.

What in the reservoir is geostatistics good for? Three examples were given: firstly, the prediction of uncertainty in reservoir properties at unsampled locations using a large number of soft geophysical, geological, and engineering data at many locations combined with a limited number of hard data at a few well locations; secondly, the integration of seismic and well data into plausible geological scenarios for resource assessment and well placement; and, thirdly, the construction of realistic heterogeneity models for recovery predictions. The examples were fabricated to mimic the features of real reservoirs. The data and programs for all synthetic examples are available from the author.

## REFERENCES


1. MATHERON, G., The Theory of Regionalized Variables and its Applications; *Les cahiers du CMM. Fasc. No. 5, Ed. Ecole Nationale Superieure des Mines de Paris, Paris, France, 1971.*
2. XU, W., TRAN, T.T., SRIVASTAVA, R.M., and JOURNEL, A.G., Integrating Seismic Data in Reservoir Modelling: The Collocated Cokriging Alternative; *paper SPE 24742, 1992.*
3. DOYEN, P.M., DEN BOER, L.D., and PILLET, W.R., Seismic Porosity Mapping in the Ekofisk Field Using a New Form of Collocated Cokriging; *paper SPE 36498, 1996.*
4. PYRCZ, M.J., Integration of Geologic Information Into Geostatistical Models; *Ph.D. Thesis, University of Alberta, Edmonton, AB, 2004.*
5. DEUTSCH, C.V. and JOURNEL, A.G., GSLIB—Geostatistical Software Library and Users Guide; *Oxford University press, 2nd Edition, 1998.*
6. DEUTSCH, C.V., Geostatistical Reservoir Modelling; *Oxford University Press, New York, NY, 2002.*
7. GOOVAERTS, P., Geostatistics for Natural Resources Evaluation; *Oxford University Press, New York, NY, 1997.*



**Provenance**—Original unsolicited manuscript, **What in the Reservoir is Geostatistics Good For?** (2005-10-24DAS). Abstract submitted for review November 24, 2005; editorial comments sent to the author(s) January 23, 2006; revised manuscript received November 24, 2005; paper approved for pre-press November 24, 2005; final approval February 15, 2006.