

STOCHASTIC SIMULATION
And
RESERVOIR MODELING WORKFLOW

C&PE 940, 21 October 2005

Geoff Bohling
Assistant Scientist
Kansas Geological Survey
geoff@kgs.ku.edu
864-2093

Overheads and other resources available at:

<http://people.ku.edu/~gbohling/cpe940>

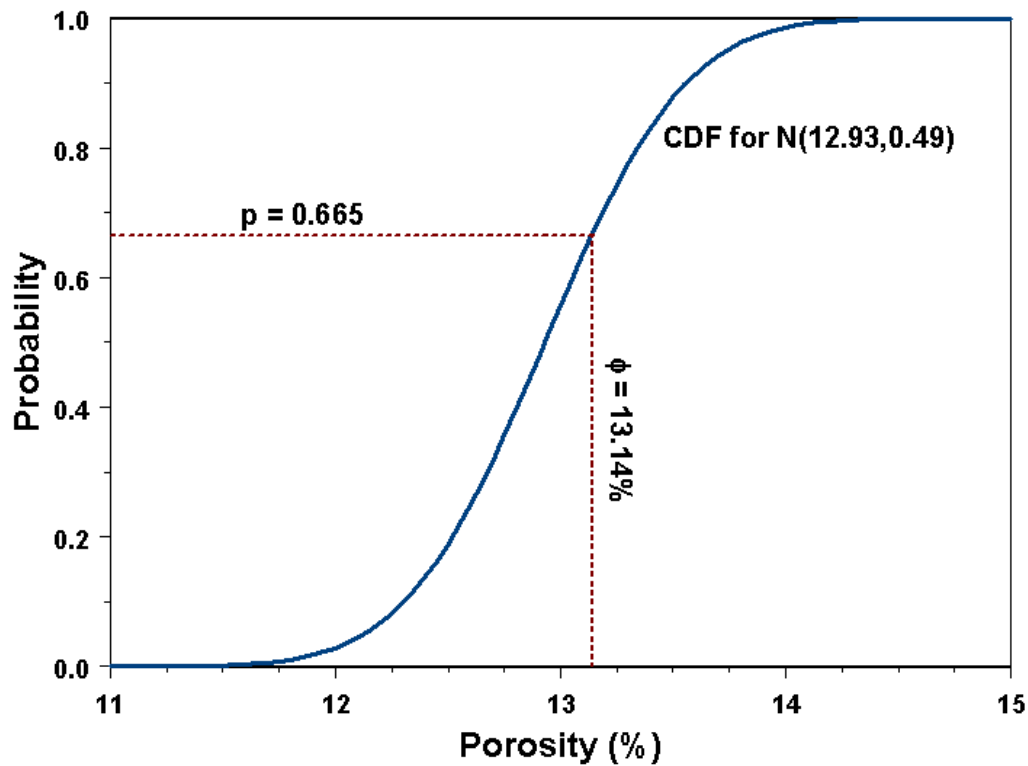
Stochastic Simulation

Stochastic simulation is a means for generating multiple equiprobable realizations of the property in question, rather than simply estimating the mean. Essentially, we are adding back in some noise to undo the smoothing effect of kriging. This possibly gives a better representation of the natural variability of the property in question and gives us a means for quantifying our uncertainty regarding what's really down there.

The two most commonly used forms of simulation for reservoir modeling applications are sequential Gaussian simulation for continuous variables like porosity and sequential indicator simulation for categorical variables like facies.

The basic idea of sequential Gaussian simulation (SGS) is very simple. Recall that kriging gives us an estimate of both the mean and standard deviation of the variable at each grid node, meaning we can represent the variable at each grid node as a random variable following a normal (Gaussian) distribution. Rather than chooses the mean as the estimate at each node, SGS chooses a random deviate from this normal distribution, selected according to a uniform random number representing the probability level.

For our six-point example in the porosity data, ordinary kriging gave a mean estimate of 12.93% with a standard deviation of 0.49%. In this case, if we happened to generate a uniform random number of $p = 0.665$ for this grid node, then the assigned porosity would be 13.14%, the corresponding value of the cumulative normal probability function:



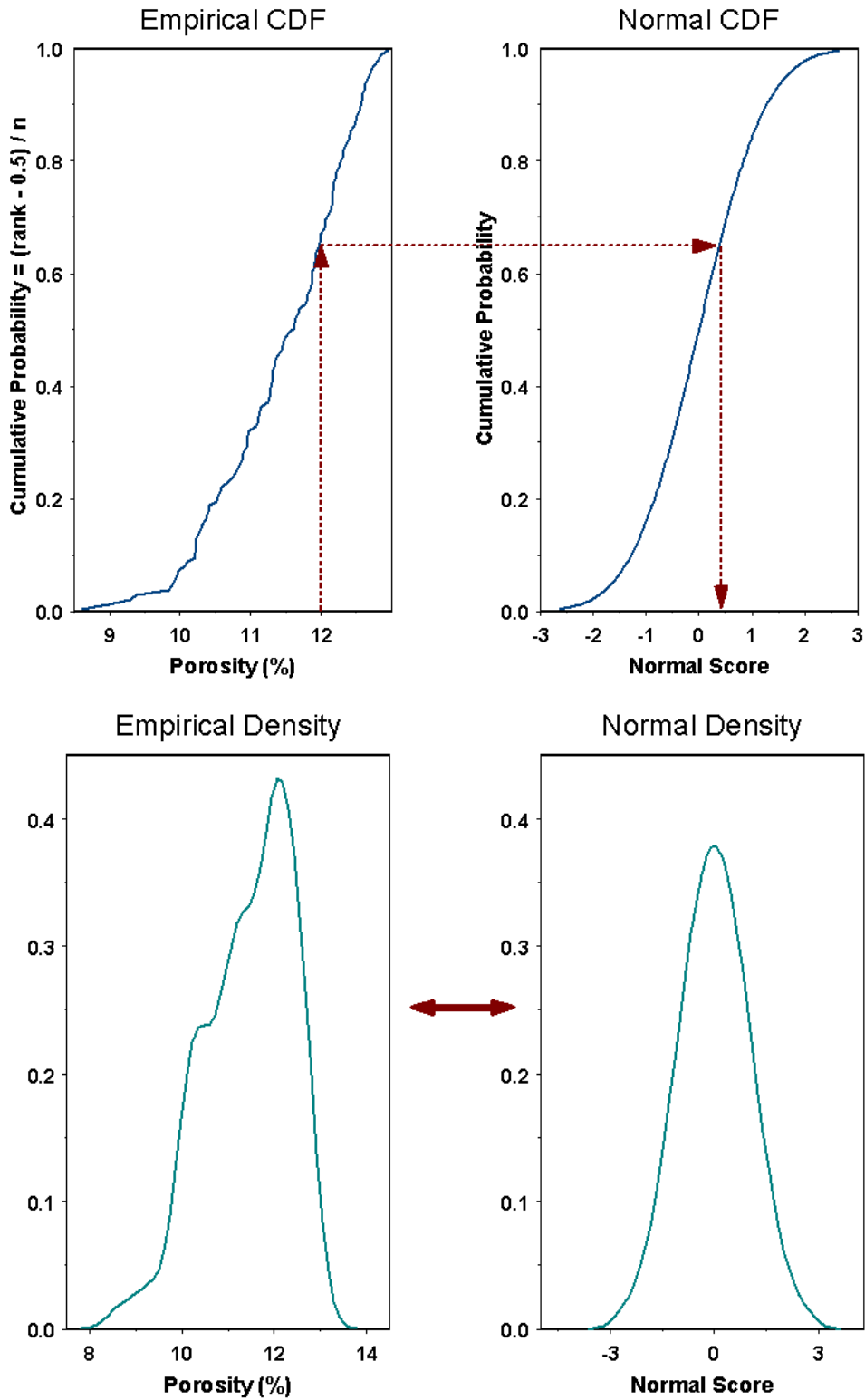
So, the basic steps in the SGS process are:

- Generate a random path through the grid nodes
- Visit the first node along the path and use kriging to estimate a mean and standard deviation for the variable at that node based on surrounding data values
- Select a value at random from the corresponding normal distribution and set the variable value at that node to that number
- Visit each successive node in the random path and repeat the process, including previously simulated nodes as data values in the kriging process

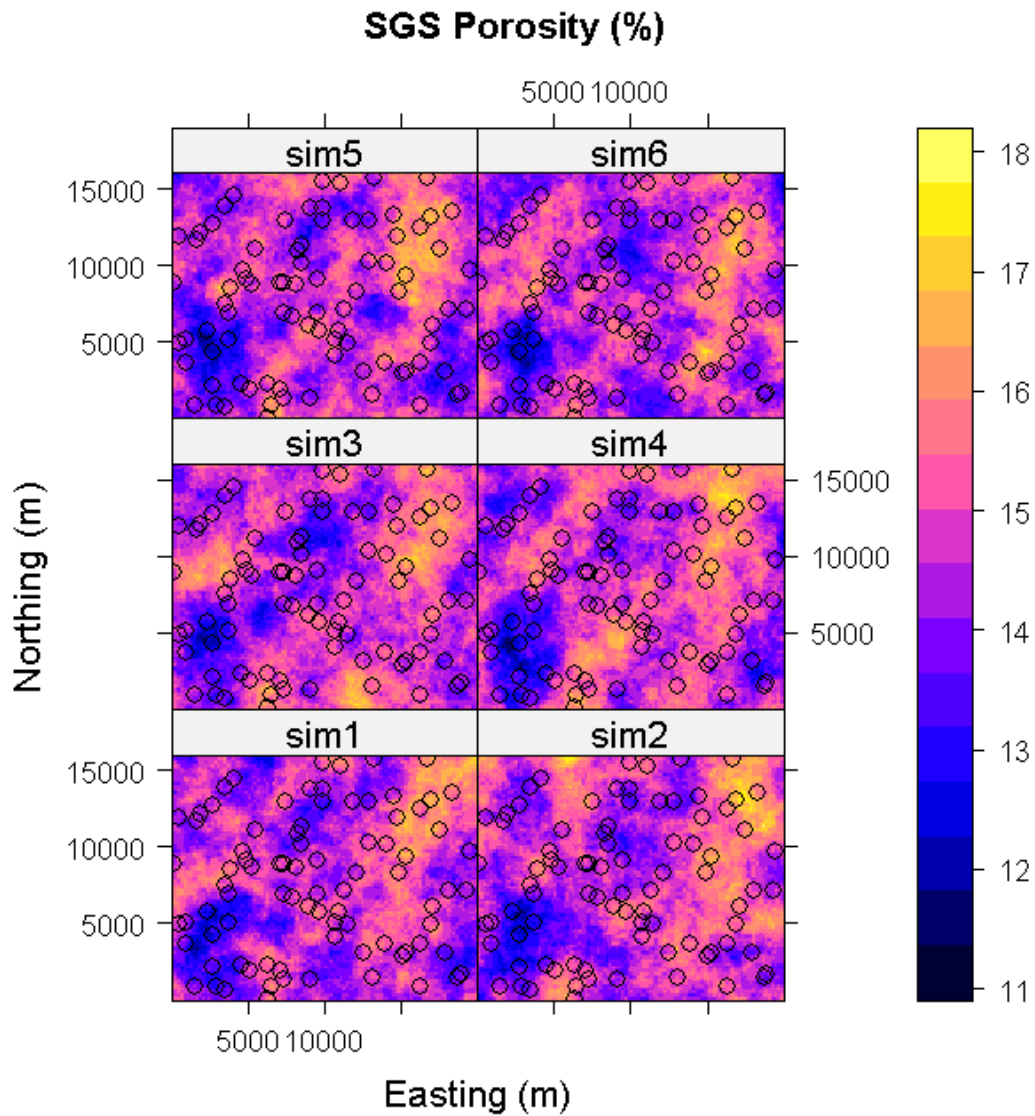
We use a random path to avoid artifacts induced by walking through the grid in a regular fashion. We include previously simulated grid nodes as “data” in order to preserve the proper covariance structure between the simulated values.

Sometimes SGS is implemented in a “multigrid” fashion, first simulating on a coarse grid (a subset of the fine grid – maybe every 10th grid node) and then on the finer grid (maybe with an intermediate step or two) in order to reproduce large-scale semivariogram structures. Without this the “screening” effect of kriging quickly takes over as the simulation progresses and nodes get filled in, so that most nodes are conditioned only on nearby values, so that small-scale structure is reproduced better than large-scale structure.

For SGS, it is important that the data actually follow a Gaussian distribution. If they do not, we can use a normal score transform:



Here are six sequential Gaussian simulations of our porosity data, using the spherical semivariogram model and a 16 nearest-neighbor search:

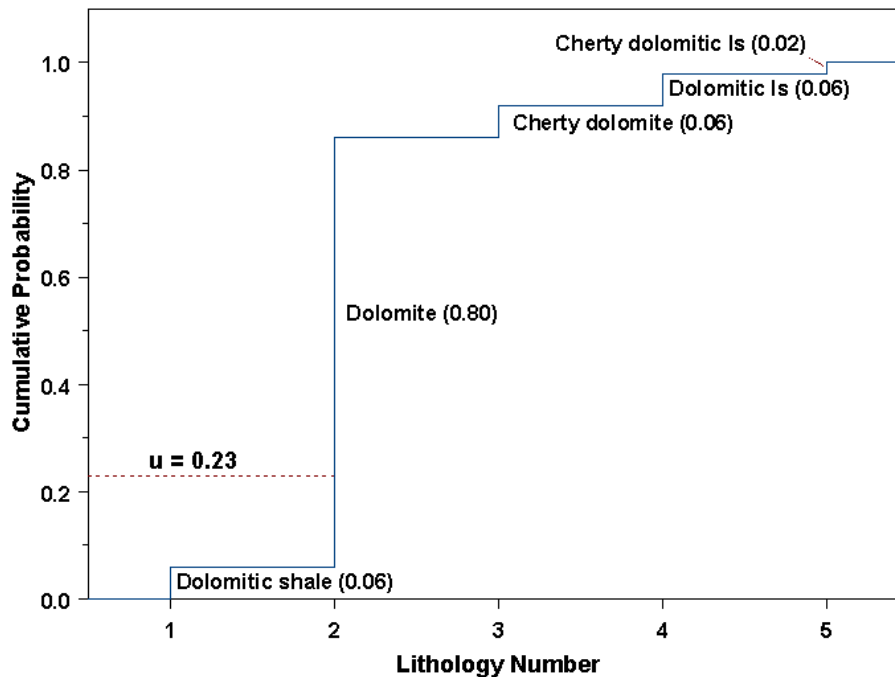


Sequential indicator simulation (SIS) is very similar to sequential Gaussian simulation, except that indicator kriging is used to build up a discrete cumulative density function for the individual categories at each case and the node is assigned a category selected at random from this discrete CDF.

Very briefly, an indicator representation for a categorical variable such as facies would be formulated as

$$i(\mathbf{u}_a; k) = \begin{cases} 1 & \text{if facies } k \text{ is present at } \mathbf{u}_a \\ 0 & \text{otherwise} \end{cases}$$

where you would have one indicator variable for each of the K different facies. We can then use kriging (based on indicator semivariograms) to produce a set of facies membership probabilities at each grid point, build up a CDF from the probabilities, and select a facies at random from the CDF:



For a continuous variable such as permeability, indicator variables are built by comparing data values to a set of thresholds, z_k :

$$i(\mathbf{u}_a; k) = \begin{cases} 1 & \text{if } z(\mathbf{u}_a) \leq z_k \\ 0 & \text{otherwise} \end{cases}$$

We might define thresholds, for example, at the 10th, 25th, 50th, 75th, and 90th percentiles of the data distribution, for example. In this case, kriging the indicator values for the k^{th} threshold, z_k , gives estimates of $P[Z(\mathbf{u}) \leq z_k]$ at each estimation point. Since this already is a cumulative probability, we don't need to go through the process of summing to get a CDF, although we will need to correct any violations of the expected order relationships, $P[Z(\mathbf{u}) \leq z_k] \leq P[Z(\mathbf{u}) \leq z_{k+1}]$, that happen to occur.

In this case, SIS assigns each node to a corresponding range (e.g., upper 10%) by random selection from the CDF and the resulting indicator vector gets added to the conditioning data for the remaining nodes.

Typical Reservoir Modeling Workflow

Basically, work from large-scale structure to small-scale structure, and generally from more deterministic methods to more stochastic methods:

- Establish large-scale geologic structure, for example, by deterministic interpolation of formation tops; this creates a set of distinct zones
- Within each zone, use SIS or some other discrete simulation technique (such as object-based simulation) to generate realizations of the facies distribution – the primary control on the porosity & permeability distributions
- Within each facies, use SGS (or similar) to generate porosity distribution and then simulate permeability distribution conditional to porosity distribution, assuming there is some relationship between the two

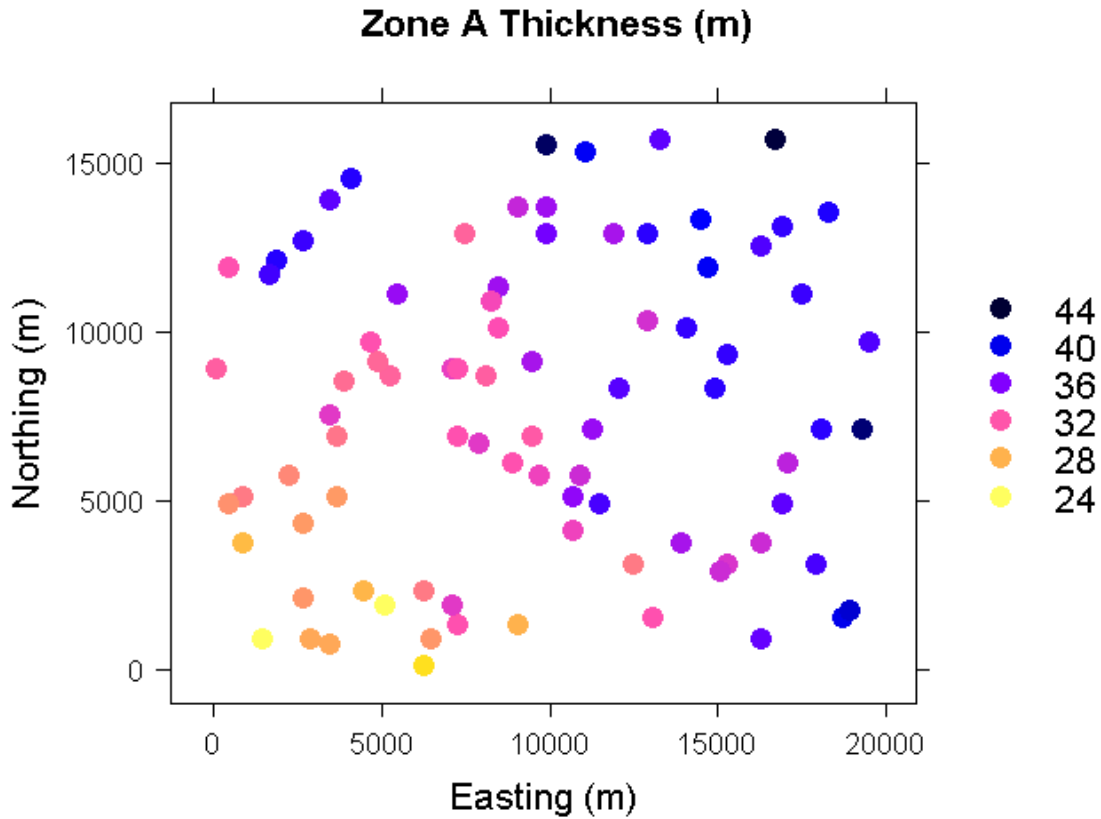
Porosity and facies simulations could be conditioned to other secondary data, such as seismic. Methods also exist for conditioning to well test and production data, but these are fairly elaborate and *probably* not in very common use as yet. More typical (maybe) to run flow simulations after the fact and rank realizations by comparison to historical production & well tests.

For more on reservoir modeling workflow and assessment of uncertainty from multiple realizations, see

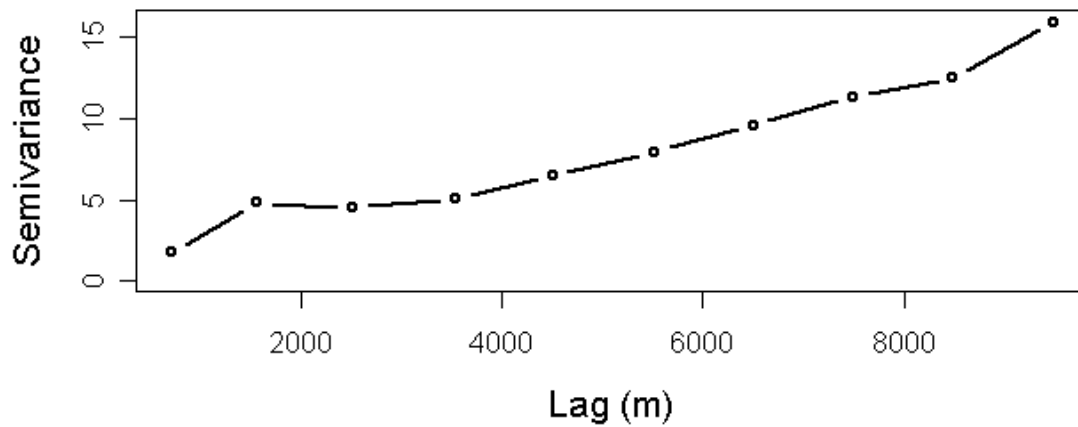
C. V. Deutsch, 2002, *Geostatistical Reservoir Modeling*, Oxford University Press, 376 pp.

Dealing With Trend: Zone A Thickness Data

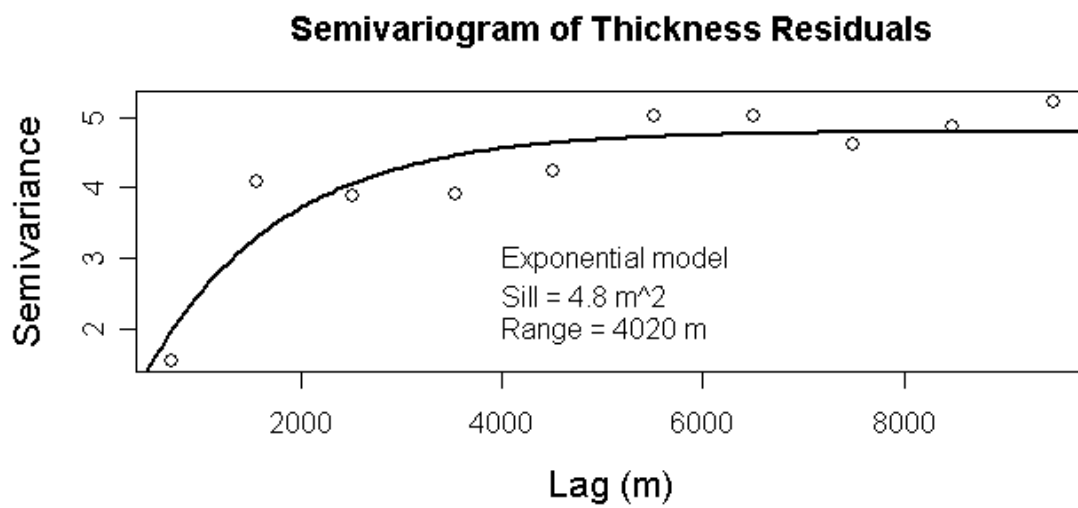
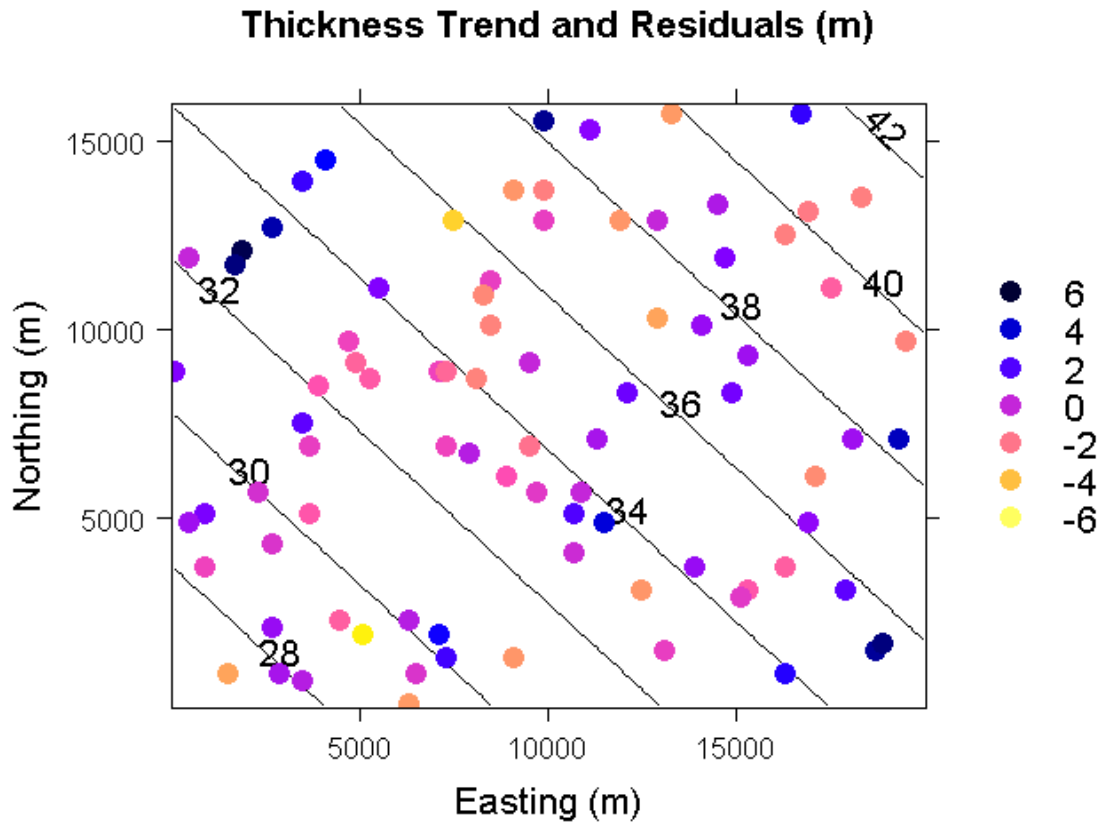
Here are the Zone A thickness values measured in the same 85 wells as our example porosity data:



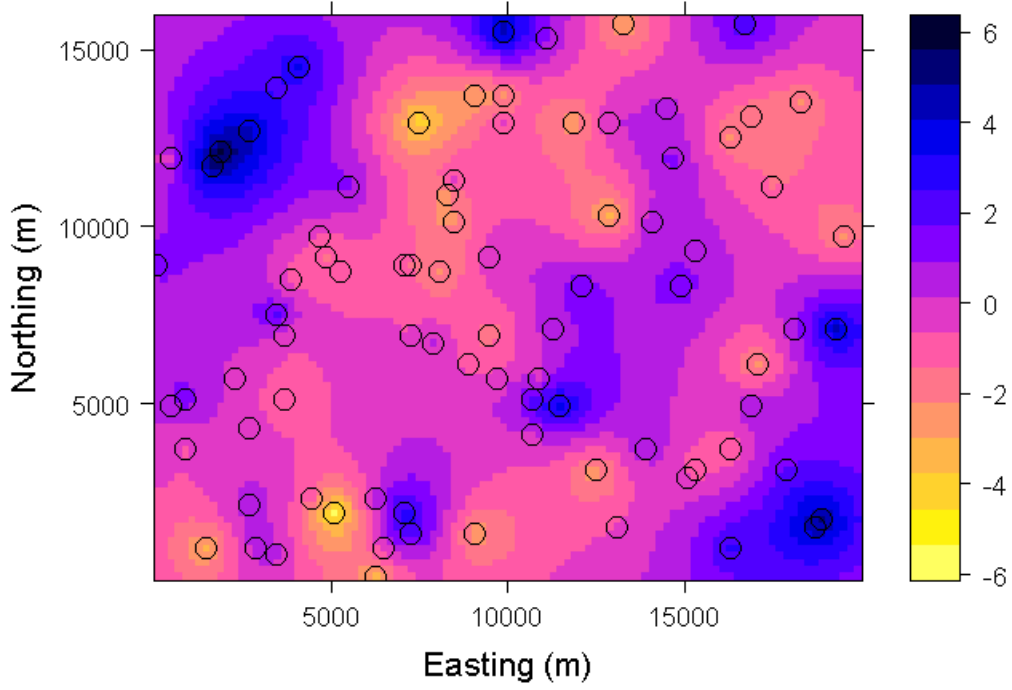
The SW-NE trend affects the omnidirectional semivariogram:



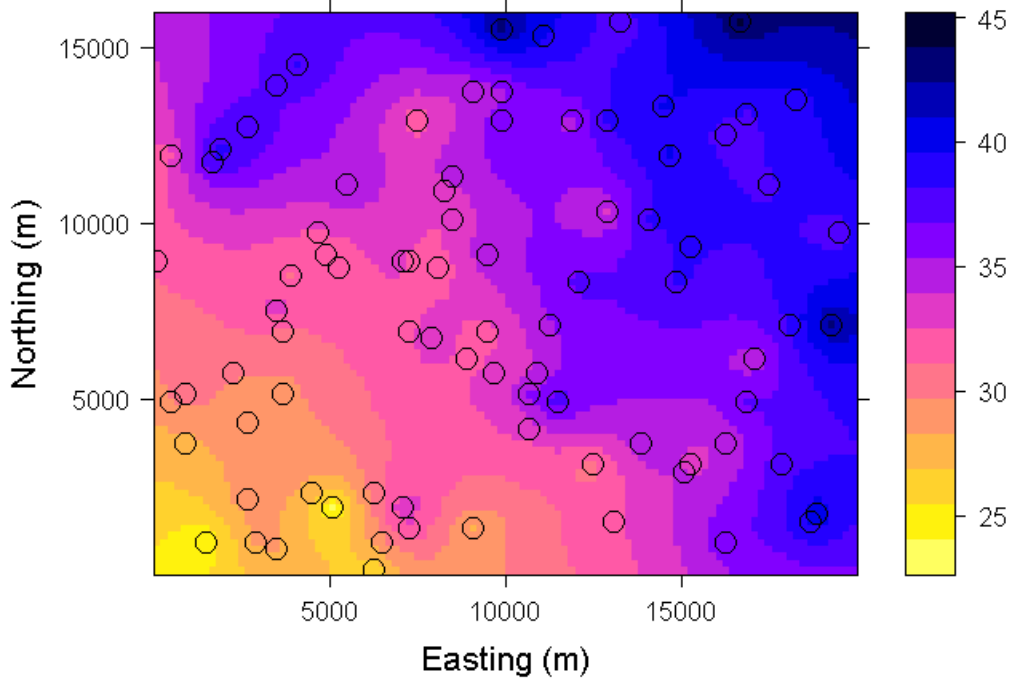
Approach 1: Fit a global linear trend and kriging the residuals



Simple-Kriged Thickness Residuals (m)

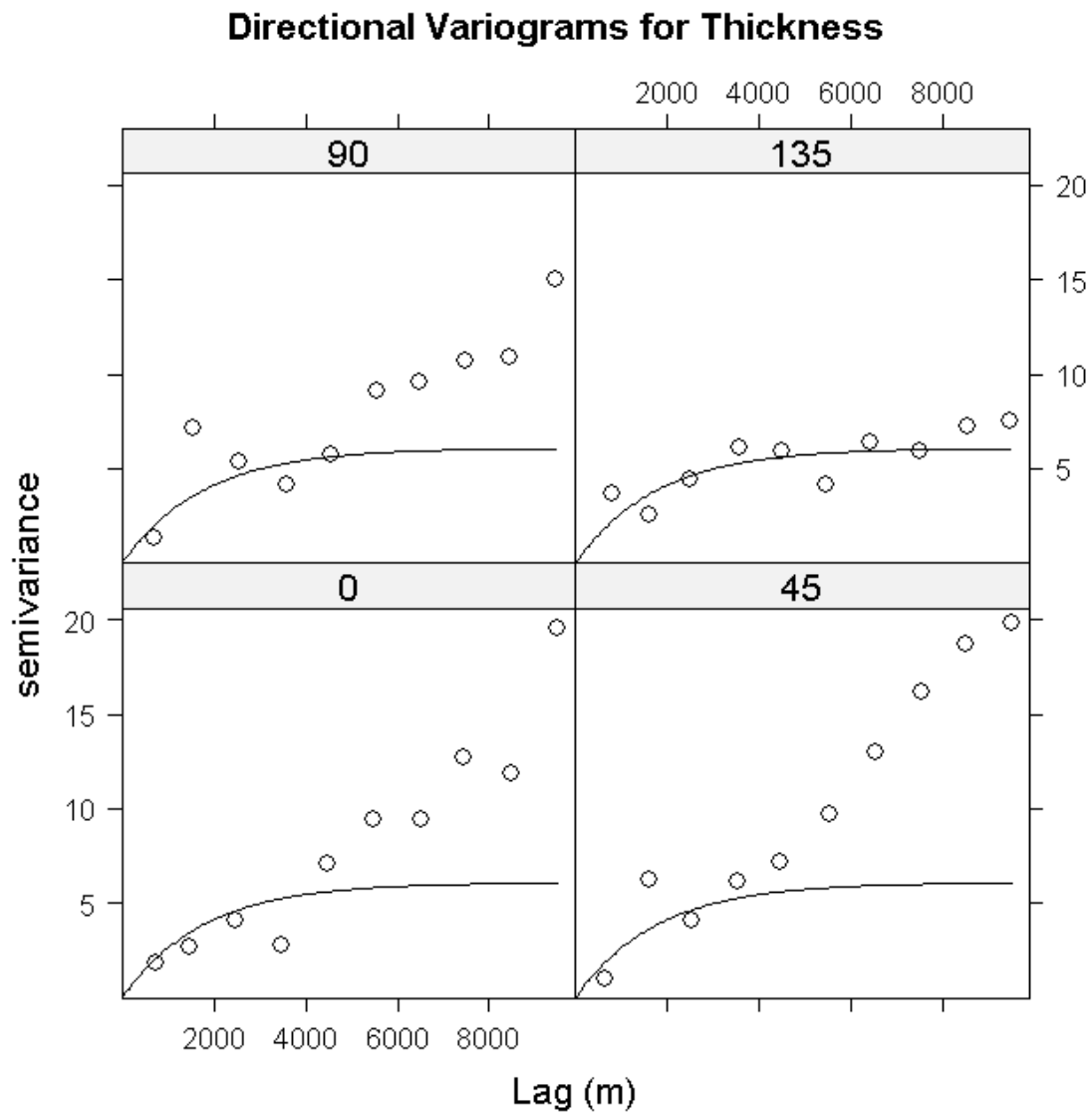


Thickness Trend + Kriged Resid. (m)



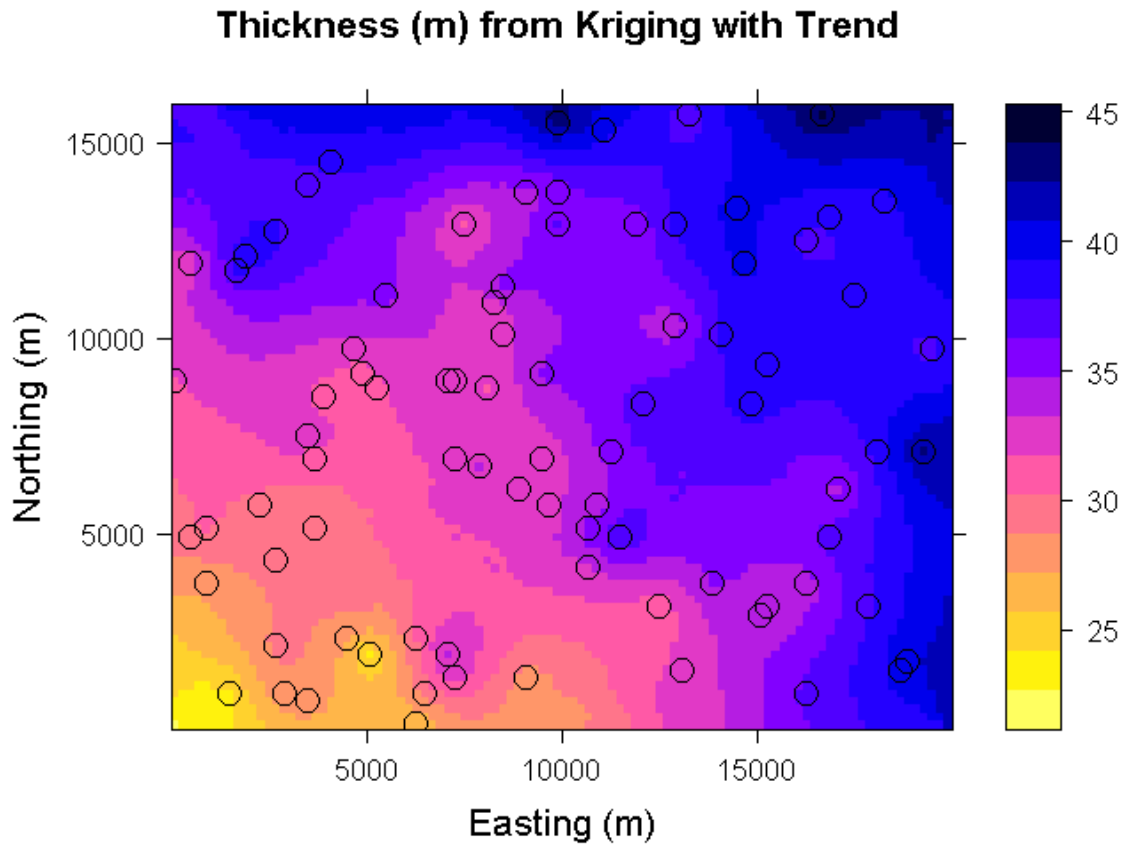
Approach 2: Fit variogram in trend-free direction and use kriging with first-order trend in X & Y

Here are the directional variograms (directions are azimuths from north) computed with a directional tolerance of 22.5°:



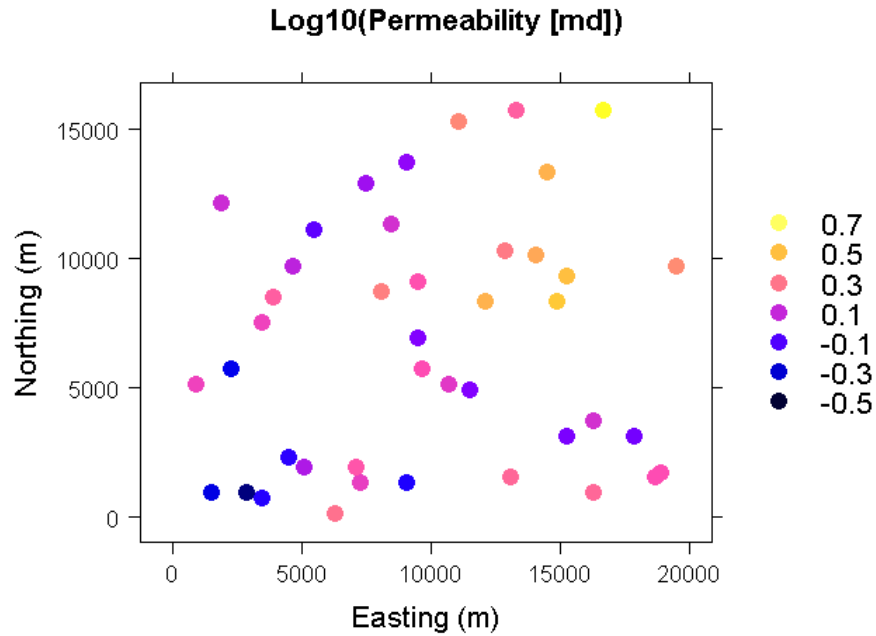
The model shown is that fitted to the semivariogram for N 135° E, which seems to be reasonably trend-free. The model is exponential with a sill of 6.1 m² and a range of 5292 m.

We then kriged using a first-order trend model in X & Y and the presumably trend-free semivariogram model for N 135° E:

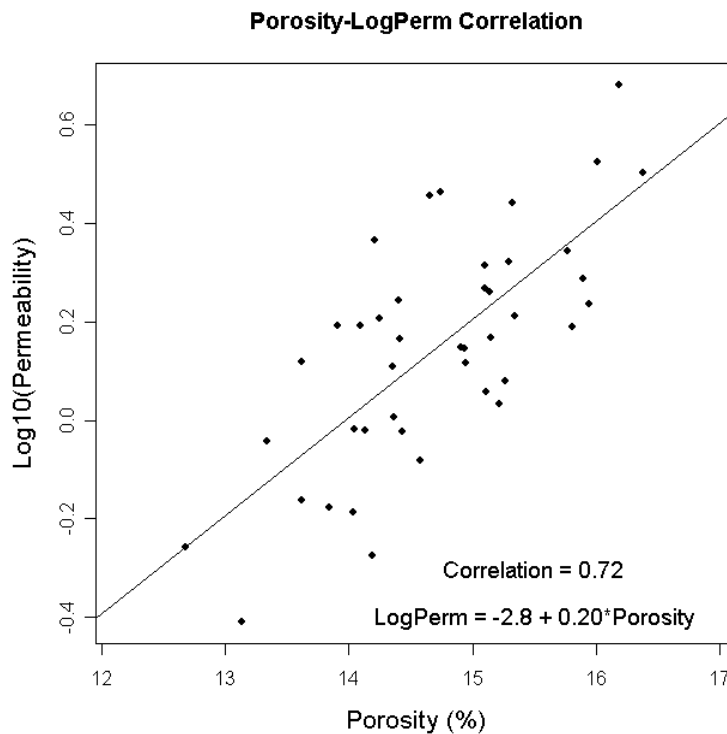


Accounting for Porosity-Permeability Correlation

Permeability data are available from 42 of the 85 Zone A wells:

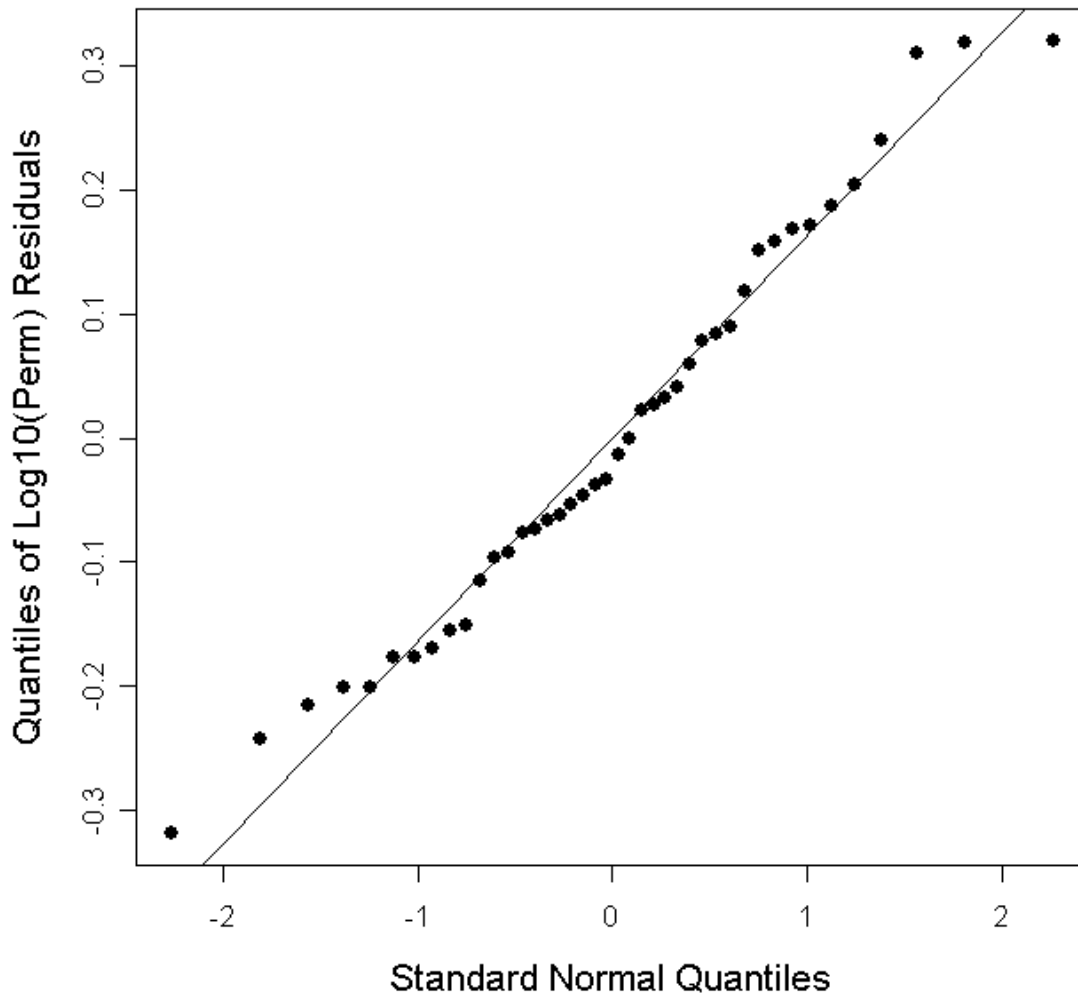


There is a fairly strong correlation between the log(perms) and porosity at the 42 wells:

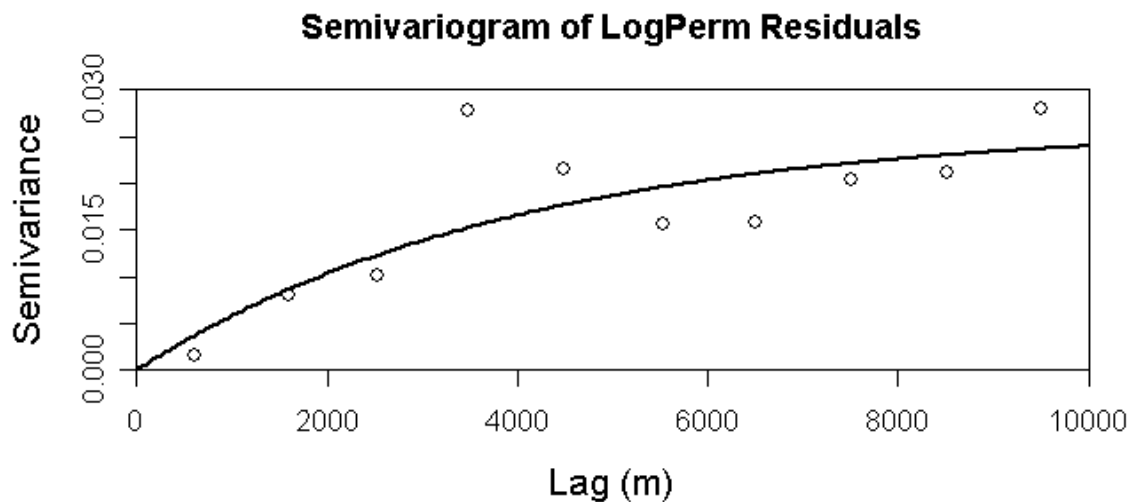
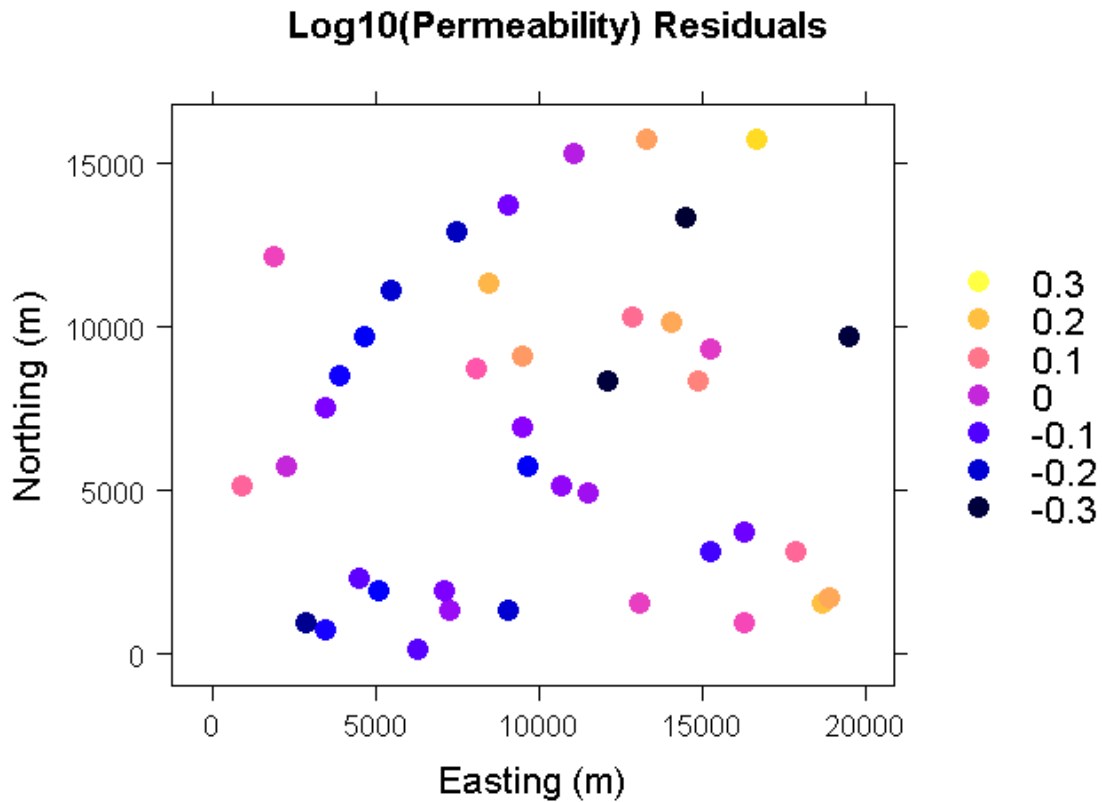


We want to account for the observed permeability-porosity relationship for two reasons: 1) to preserve this relationship in our modeled permeability and porosity distributions, and 2) to take advantage of the more abundant porosity data in our estimation of the permeability. To do this, we will work with the LogPerm residuals (actual LogPerms minus those predicted from the porosity) and add the kriged or simulated residuals back into a mean LogPerm grid predicted from the porosity grid.

The LogPerm residuals at the wells seem reasonably normally distributed:

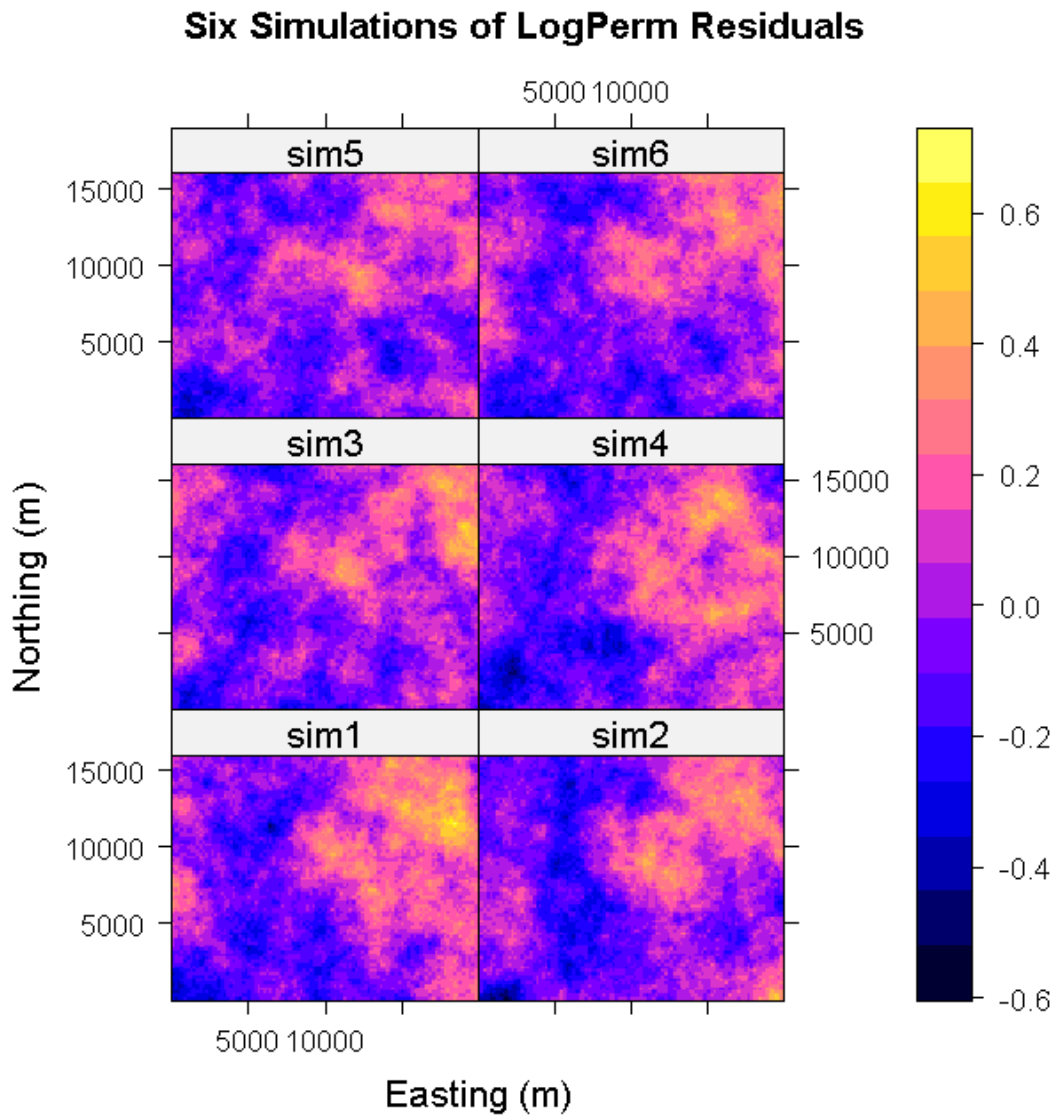


and may or may not show a bit of trend:

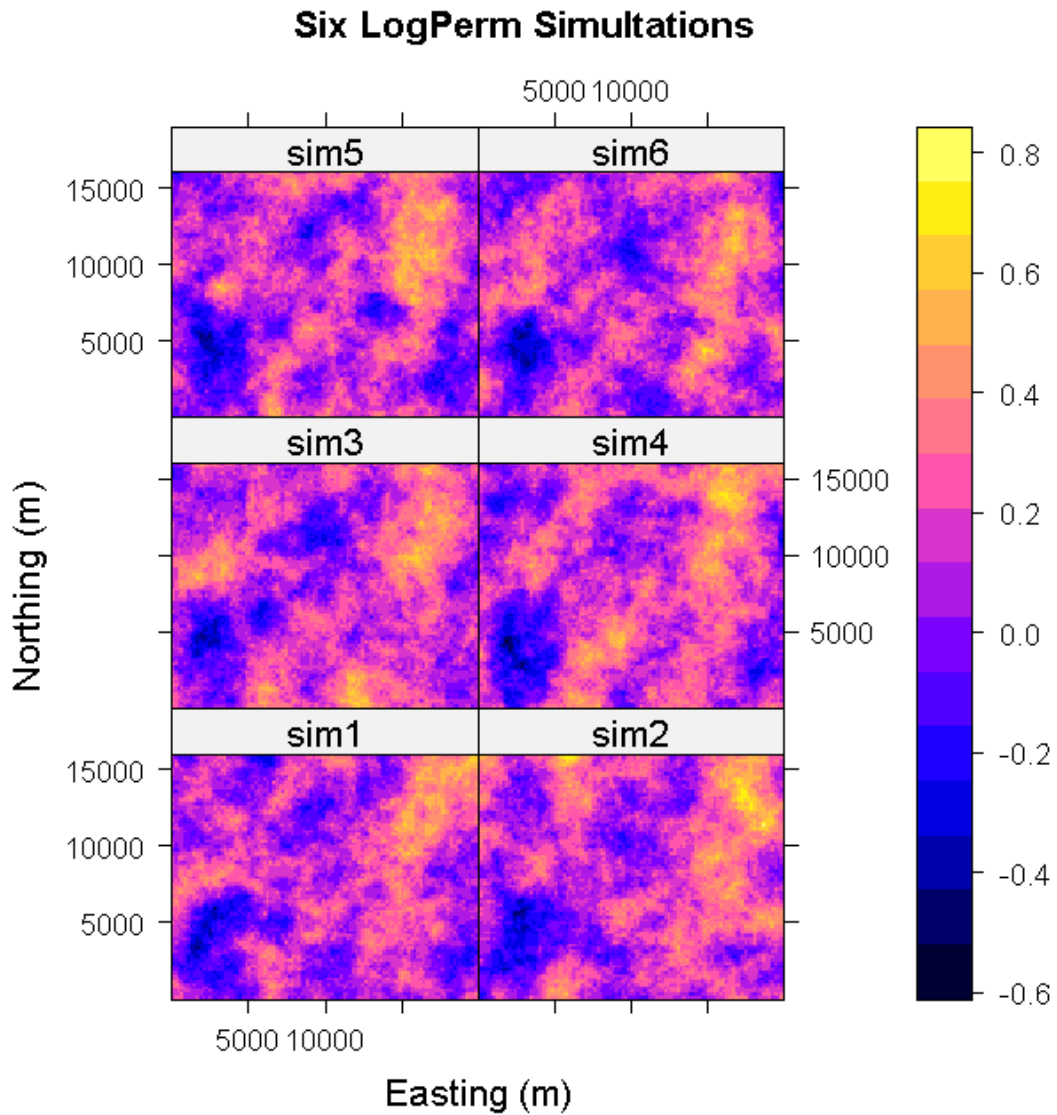


The semivariogram model is exponential with a sill of 0.026 and a range of 11800 m. The fitted sill matches the global variance quite closely, so I am going to treat the residuals as trend-free.

Here are six sequential Gaussian simulations of the LogPerm residuals based on simple kriging with a 16 nearest-neighbor search:



And here are those six simulations of the LogPerm itself based on adding the simulated residuals to the LogPerm values predicted on the basis of the six porosity simulations we developed before, using the regression equation developed on the well data:



Here are the crossplots of simulated LogPerm versus simulated porosity, along with the regression line developed from the well data. The simulated values show a somewhat higher correlation overall (0.79) than seen in the original well data (0.72), probably due to the small positive correlation (0.19) between the simulated LogPerm residuals and simulated porosities.

