Cross Validation for Selection of Variogram Model and Kriging Type: Application to IP Data from West Virginia

Bora Oz (boz@gpu.srv.ualberta.ca) and Clayton V. Deutsch (cdeutsch@civil.ualberta.ca)
Department of Civil & Environmental Engineering, University of Alberta

Abstract

A variogram model and kriging type must be chosen prior to the generation of geostatistical models. Cross Validation gives us some ability to assess the impact of our many choices such as the variogram models, kriging type, and search strategies. For estimation, one can plot the estimated values versus true values and look for the scenario with the highest correlation coefficient. Other tools are available to check the results of simulation.

The aim of this paper is to review the methodology to choose parameters such as the candidate variogram model and kriging type. An example study with initial potential data from Barbour County, West Virginia is presented to clarify the application of this methodology.

KEY WORDS: simulation, accuracy, precision, uncertainty, cyclicity

Introduction

In practice, we would like to compare different geostatistical prediction scenarios and choose the one that works best. These scenarios consist of different weighting procedures, search strategies, variogram models and so on. The comparison can be based on the difference between true and estimated values. "Cross Validation" allows us to compare estimated and true values using the information available in our sample data set. The sample values are temporarily discarded from the sample data set; the value is then estimated using the remaining samples. The estimates are then compared to the true values.

Since its intoduction in the geostatistical literature [2, 3], cross validation has been used widely for different purposes. Delfiner [3] used it to search for the generalized covariance function to be used in his estimation procedure among a finite number of candidates. Cross validation is used to find the best model among the competitors. Clark [1] reviews the history of validation and its usefulness in geostatistics. She points out that this type of comparison was used to compare methods of estimation [2, 7] and to justify the use of kriging as an estimation method [8]. Clark demonstrates that use of cross validation for selecting a semivariogram model may be acceptable, but may not be sensitive enough to be very useful.

Notwithstanding some limitations, the results of cross validation provide practitioners an ability to examine and reformulate models to better utilize and conform to the data at hand. Cross validation is primarily suited to exploratory purposes: it does not provide a definitive measure of the goodness of an estimation / simulation method.

We consider cross validation to help determine the variogram model and kriging type. An example application to initial potentials in Barbour County, West Virginia, is used to clarify the use of cross validation in estimation and simulation modes.

Methodology

Determine Scenarios to Test

Variograms are used to describe spatial variability. For a given field or data set, one can consider different variograms to describe the spatial distribution of the variables under consideration; a number of different variogram models may appear to fit the data set equally well. It is for that reason that cross validation technique is used to test the calculated variograms

Values are estimated by kriging with the candidate variogram models. We can use different kriging types such as Simple Kriging or Ordinary Kriging. We can also propose different search parameters. A reasonable set of scenarios consisting of a variogram model, kriging type, and search strategy are proposed for testing with cross validation.

Cross Validation

The cross validation exercise is repeated for search scenario and then the results are compared. Each sample point is left out of the data set in turn and estimated with the remaining data and appropriate scenario parameters.

In estimation, an error is calculated as the estimated value minus the true value. These error values can be analyzed in different ways. The most common ones:

- A scatterplot of the true values versus the estimated values. There should be few outliners and a high correlation coefficient.
- The histogram of errors should be symmetric, centered on a zero mean, with minimum standard deviation.
- The plot of the error versus the estimated value should be centered around the zero-error line, a property called "conditional unbiasedness".

In simulation, we have a distribuition of simulated values instead of a single estimated value; different techniques must be used to evaluate the results. Details can be found in the study of Deutsch [4]. In short, we check that the probabilistic model is both accurate and precise while minimizing the unceratinity.

Application to Initial Potentials in Barbour County

The Data

The study area is located in the Barbour County, West Virginia. Hydrocarbon is being produced from Upper Devonian sandstones and siltstones. The data set, taken from Hohn [6], consists of 674 wells from a 22 km by 22 km area.

A location map of the field is given in Figure 1. The potential values are higher in the middle of the field. There are large variations even between closely-spaced wells.

The histogram of the data is presented in Figure 2. It is clear from the histogram that most of that fall into the range of 30-1030 Mscfd. Values up to 16000 Mscfd are seen. The mean value is 1239.3 Mscfd.

Cell declustering was used to establish a more representative mean and sample histogram. Figure 3 shows a cross plot of the declustered mean versus cell size. A Cell-size of 3 km, corresponding to a value 1068 Mscfd, was chosen for declustering. These weights are used in simulation and for the simple kriging mean. The new histogram, with a mean of 1068.0 Mscfd, is presented in Figure 4.

Candidate Variograms / Kriging Type

The experimental omminidirectional normal score semivariogram and the fitted traditional variogram model (Candidate-1) are presented in Figure 5. A relatively high nugget value of around 0.45 is observed, which implies abrupt changes in the variable over small distances. The variogram rises too quickly for small values of lag distance which causes the range of correlation appears at about 3 km.

The experimental variogram appears to oscillate. The general trend of the experimental variogram shows high and low values between successive lag distances, which is called hole effect or cyclicity. Therefore, our second candidate variogram model will be the one that reflects the effect of the cycylicity (Candidate-2). In Figure 6, the fitted cyclic model and the experimental variogram are presented.

Simple and Ordinary Kriging will be tested with both variogram models. We have four different alternatives:

- Simple Kriging with Traditional Variogram Model (Variogram Candidate-1)
- Simple Kriging with Cyclic Variogram Model (Variogram Candidate-2)
- Ordinary Kriging with Traditional Variogram Model
- Ordinary Kriging with Cyclic Variogram Model

Figure 7 shows the scatter plot of estimated values versus true values for the four alternatives. It is clear ordinary kriging with cyclic variogram has the highest correlation. Another interesting conclusion is that cyclic variogram with simple kriging has higher correlation than the traditional variogram with ordinary kriging, which seems to place importance on the cyclic variogram.

The mean of the error distribution is often referred to as the *bias* and a goal for the any estimation method is to produce unbiased estimates. Ideally, one would like the mean to be zero which shows our overestimates and underestimates are in balance. Besides, we prefer to have a more symmetric error distribution.

Figure 8 shows a histogram of the errors for the four alternatives. Although the mean of the four distributions are not equal to zero, when we compare them within the whole range of error, they are small and assemble around zero. The Ordinary Kriging with cyclic variogram model has the mean closest to zero. One clear observation is that all of them have large underestimates reaching to a absolute error value of 14000.

CCDF Model	Variance	Mean^2	MSE
SK- Variogram Candidate-1	3,404,025	4,409	3,408,434
SK- Variogram Candidate-2	2,030,625	388	2,031,013
OK-Variogram Candidate-1	2,030,629	1,043	2,031,672
OK-Variogram Candidate-2	1,965,504	21	1,965,525

Table 1: Summary of the mean squared errors, *MSE*, for the four alternatives

Another feature one want to have is that a small spread of error distribution. The Mean Squared Error, *MSE*, incorporates both the mean (or bias) and the spread of the error distribution and can be calculated by,

$$MSE = E\left\{[z^* - z]^2\right\} = errorvariance + bias^2 \tag{1}$$

The results of Mean Squared Error, *MSE*, for the four alternatives are given in Table 1. It is clear that while the three candidates have approximately same Mean Squared Error, the Simple Kriging with traditional variogram (Variogram Candidate-1) has the largest *MSE*, reaching twice as much as the alternatives.

Another tool to analyze the error data is to check for the "conditional unbiasedness". Ideally, we would like to subdivide our estimates into many different groups and have an unbiased error distribution within each group (i.e. *conditional unbiasedness).* One way of checking for conditional bias is to plot the errors as a function of the estimates values. Ideally, this scatter plot would result in; (1) no correlation, and (2) no increase in estimation variance.

Error versus estimate plots for simple and ordinary kriging with cyclic variogram model are given in Figure 9. It is clear that both of the models result in overall no correlation; however, for both of these models, estimation variances increase with the estimates. The reason for this may be explained by the "proportional effect" or "heteroscedasticity". It is common to find that the data values in some regions are more variable than in others, called *heteroscedasticity.*

The *proportional effect* is a positive correlation between the local means and the local variances. After applying moving window averaging technique, by 2 km by 2 km windows and ignoring any resultant window which contains less than 3 data; the scatter plot of the local means versus local variances is given in Figure 10 for the Ordinary Kriging with cyclic variogram model (Variogram Candidate-2). It is clear that there is strong correlation between the local means and the local variances. This explains the reason why the estimation variance increases with the estimates. The data show (1) an increase in local variance with an increase in local mean, and (2) large variability between the data values.

Simulation mode: Accuracy, Precison and Uncertainity

We need to check the goodness of the probabilistic model itself. In other words, we have to assess our model and answer the question of "How good are our probabilities?". Two

CCDF Model	Accuracy	Precision	Uncertainity
SK-Variogram Candidate-1	0.757	0.712	0.56
SK-Variogram Candidate-2	0.823	0.74	0.494
0K-Variogram Candidate-1	0.792	0.729	0.513
0K-Variogram Candidate-2	0.865	0.77	0.432

Table 2: Summary of the results for simulation mode.

concepts, *accuracy* and *precision* help us answer this question .

The accuracy and precision of the previous four alternatives are checked using the program `accplot` from GSLIB [5]. The accuracy plots for the four alternatives are given in Figures 11.The accuracy plots show that until the probability interval of 0.3, all the methods are accurate and precise, but for the other intervals, points are away from the 45 degree line which makes the models inaccurate. That is the variance of the local distributions is too high.

In Table 2, the results of the accuracy, precision and uncertainity are given respectively for the four alternatives. It is again clear that cyclic variogram is more successful than the others. When we look at the uncertainity values, cyclic variogram model with Ordinary Kriging has less uncertainity than the alternatives.

Discussion and Conclusion

Many of the statistics including the mean of the error distributions and the *MSE* values are ambiguous and do not help for selection purposes with the Barbour County data. Nevertheless, the highest correlation between estimate and true identifies ordinary kriging with the cyclic variogram model (Variogram Candidate-2) as the best alternative.

In simulation mode, lone again, ordinary kriging with the cyclic variogram model has the highest accuracy and precision values with the lowest uncertainity.

By looking at the results of estimation and simulation modes, the cyclic variogram with Ordinary Kriging appears as the “best” model among the others. Nevertheless, we should note that ‘best” applies in a restricted sense. A variogram model is best, for instance, only under the choice of discrepancy measure, data number, predictive function, and number if models to be evaluated. If a change exists in any of these, a new ‘best” model may emerge. If one model performs better than another on a particular data set, it does not mean that, in general, that the model will always perform better. The result does not guarantee the success of the final performance becauese our ability to produce good estimates at sample locations may have little relevance to the final estimation study.

The high correlation between the local means and local variances (or proportional effect) entails a correlation between the error variance and the estimate.

Cross validation is used best as an exploratory technique. It can lead the user to reexamine and reformulate models to better utilize and conform to the data at hand.

References

[1] I. Clark. The art of cross validation in geostatistical application. *Proceeding of APCOM, pages 211-220*, 1986.

[2] M. David. The practice of kriging,. *Advance Geostatistics in Mining Industry, pages 31-48*, 1976.

[3] P. Delfiner. Linear estimation of nonstationary spatial phenomena,. *Advance Geostatistics in Mining Industry, pages 49-68*, 1976.

[4] C. V. Deutsch. Direct assessment of local accuracy and precision. In E. Y. Baafi and N. A. Schofield, editors, *Geostatistics Wollongong '96*, volume 1, pages 115–125. Kluwer Academic Publishers, 1996.

[5] C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 1992.

[6] Michael Edward Hohn. *Geostatistics and Petroleum Geology, Second Edition*. West Virginia Geological Survey.

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

[8] H. M. Parker, A. G. Journel, and W. C. Dixon. The use of conditional lognormal probabiliy distribution fo the estimation of open-pit ore reserves in strata-bound uranium deposits-a case study. *Proceeding of APCOM, Society of Mining Engineers, pages 133-148, New York*, 1979.

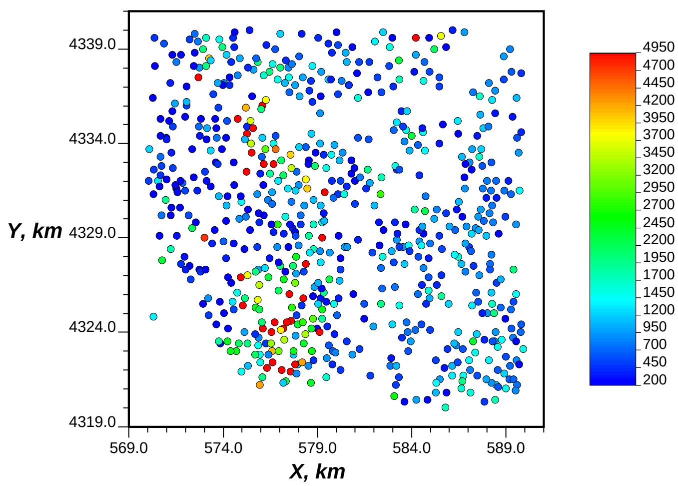

Figure 1: Location map of the initial potentials.

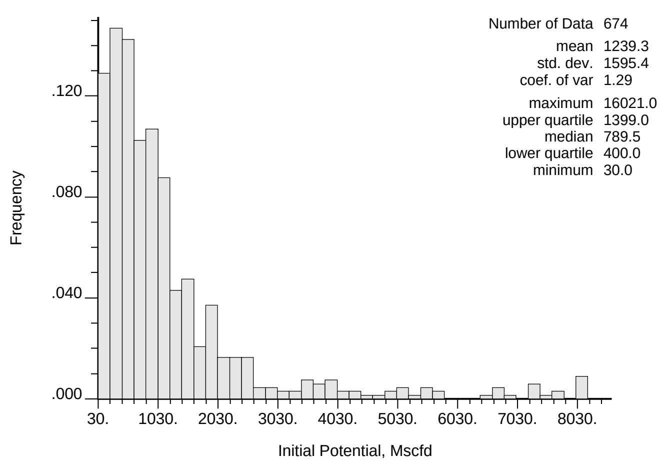

Figure 2: Histogram of the initial potentials.

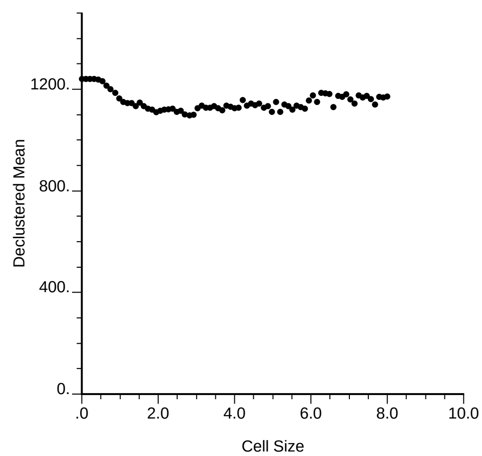

Figure 3: Declustered mean versus cell size

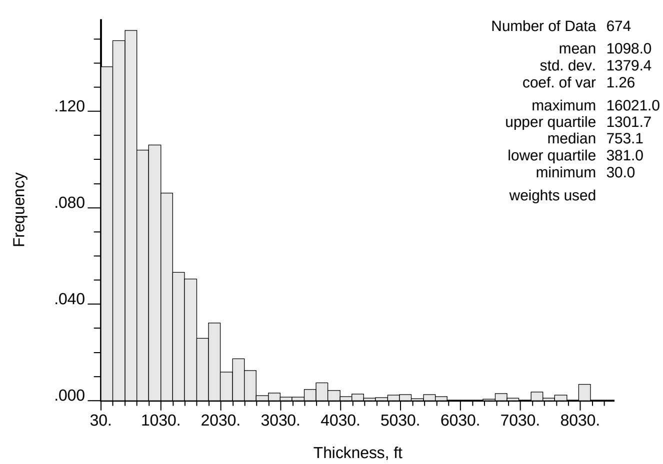

Figure 4: Histogram of the initial potentials (weight used).

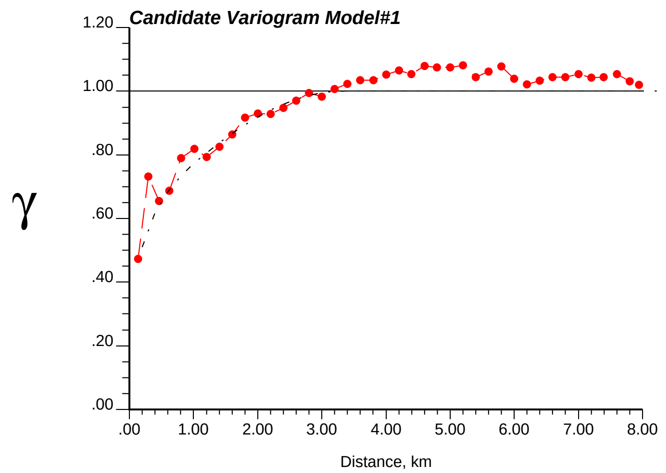

Figure 5: Experimental normal score semivariogram and the first candidate model (omminidirectional)

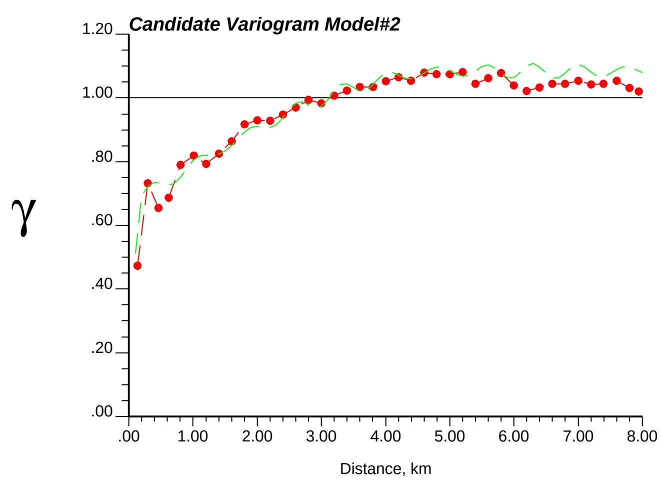

Figure 6: Experimental normal score semivariogram and the second candidate variogram model (cyclicity)

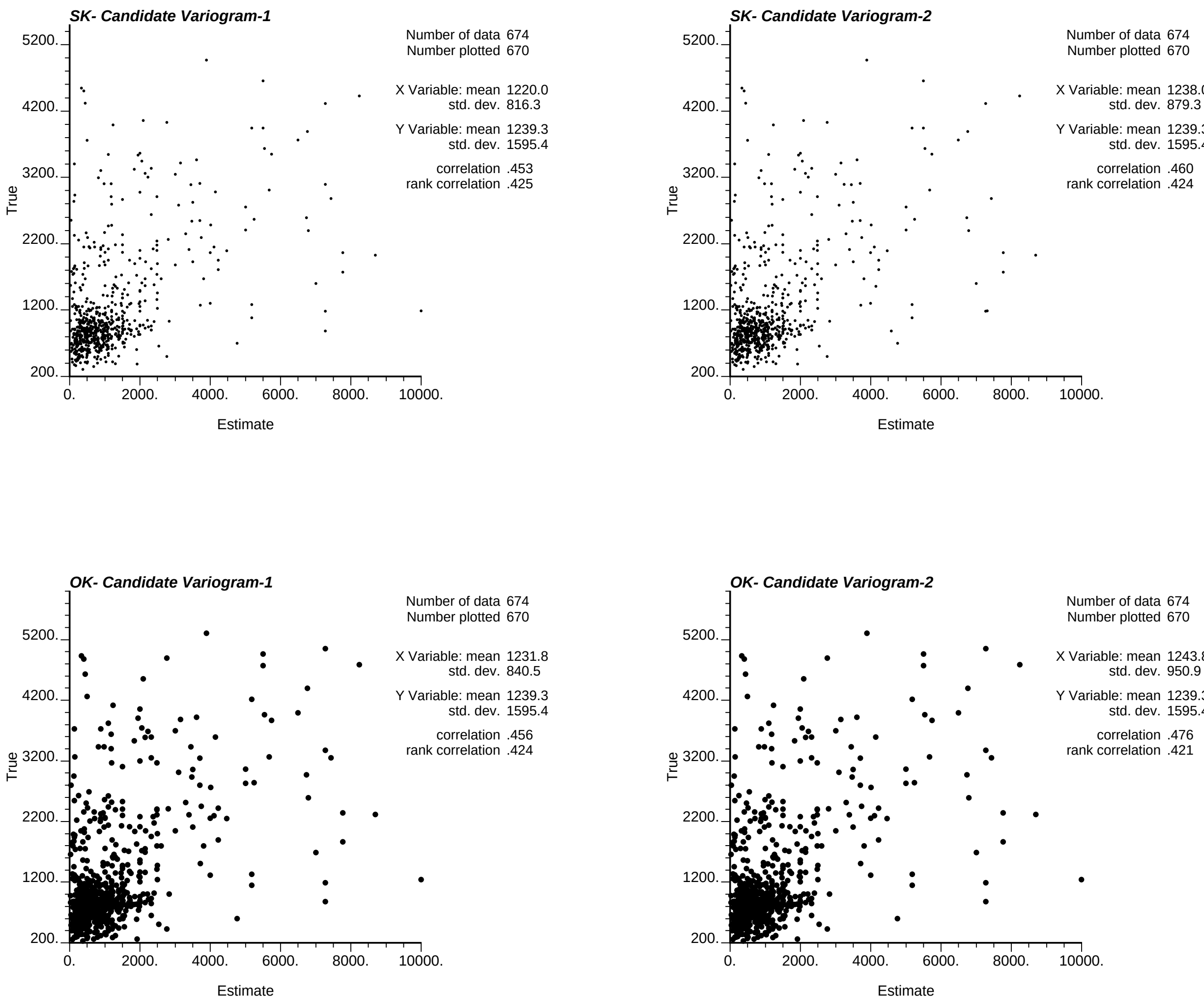

Figure 7: Cross Validation results for the four alternative scenarios.

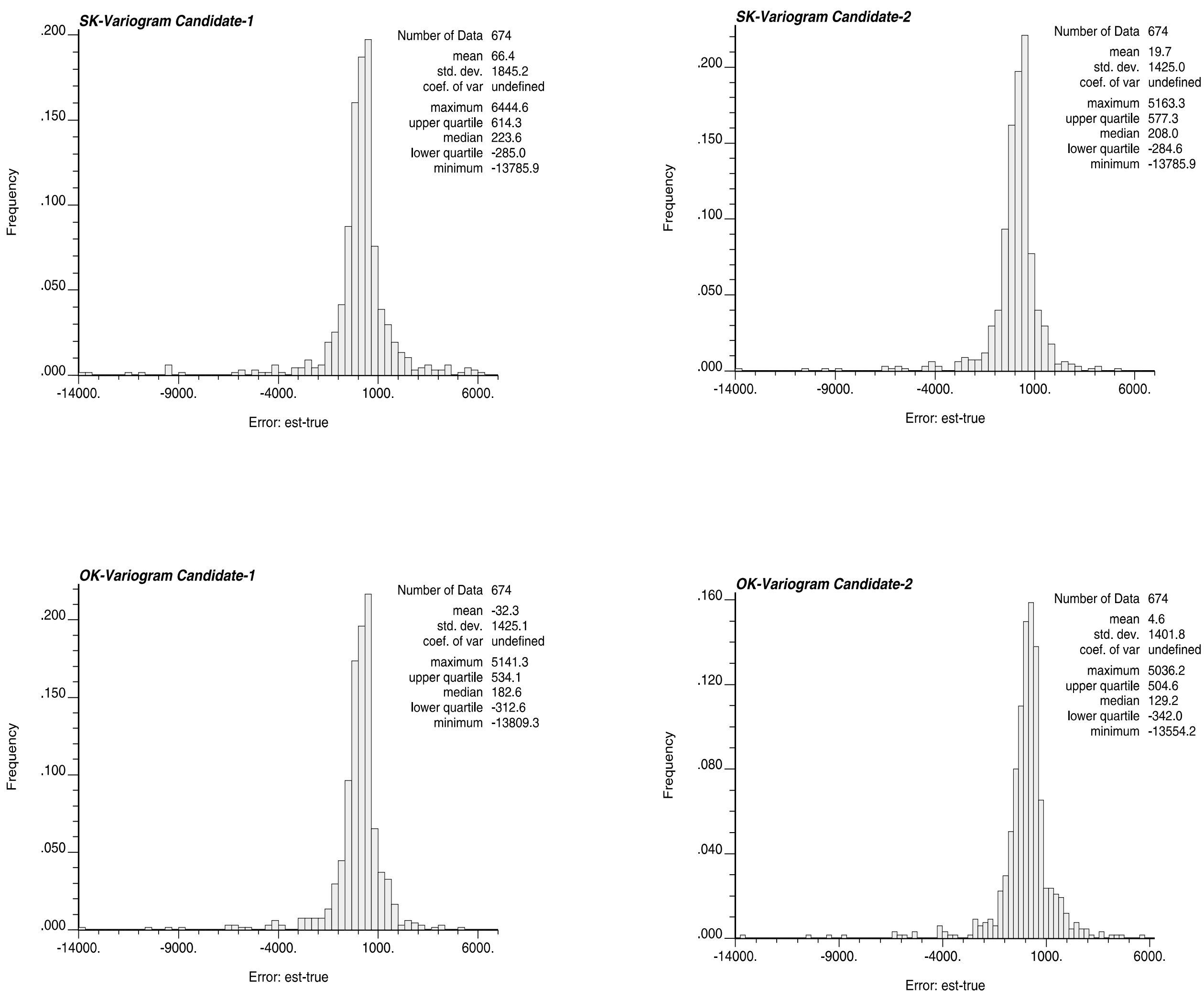

Figure 8: Histogram of the errors for the four alternatives.

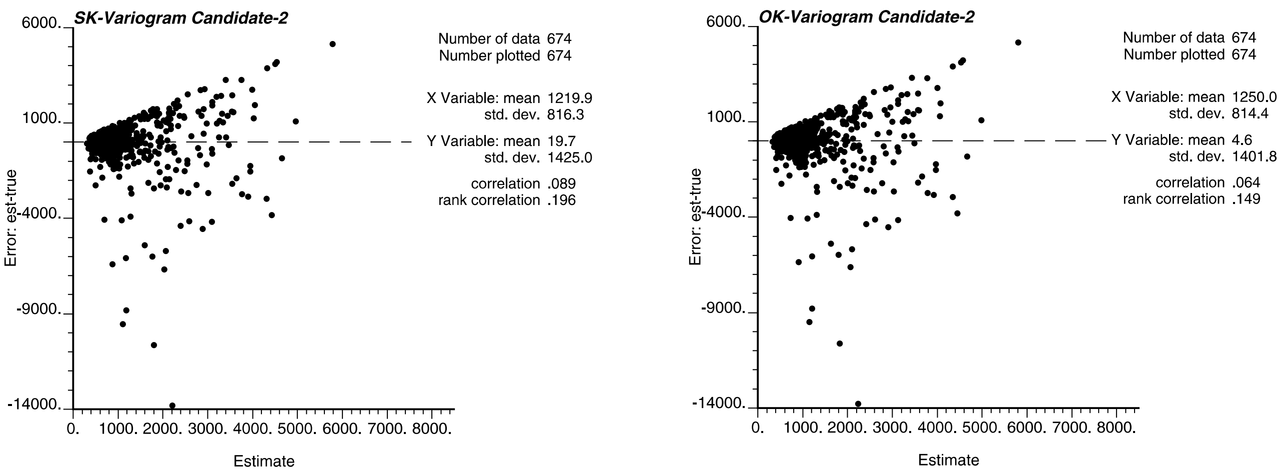

Figure 9: Error versus estimate plots for simple and ordinary kriging with cyclic variogram model.

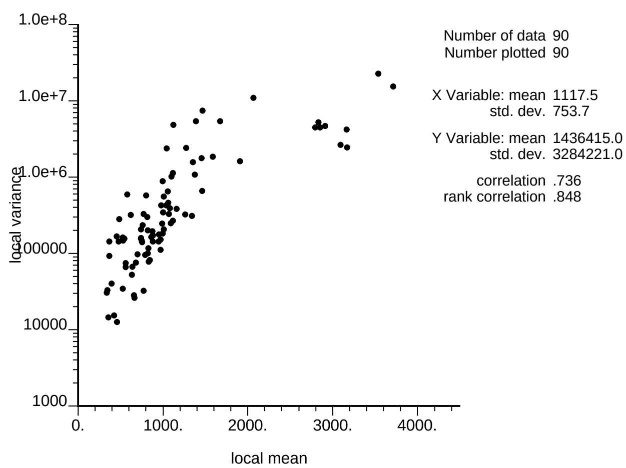

Figure 10: Local mean versus local variance after moving average for ordinary kriging with cyclic variogram model.

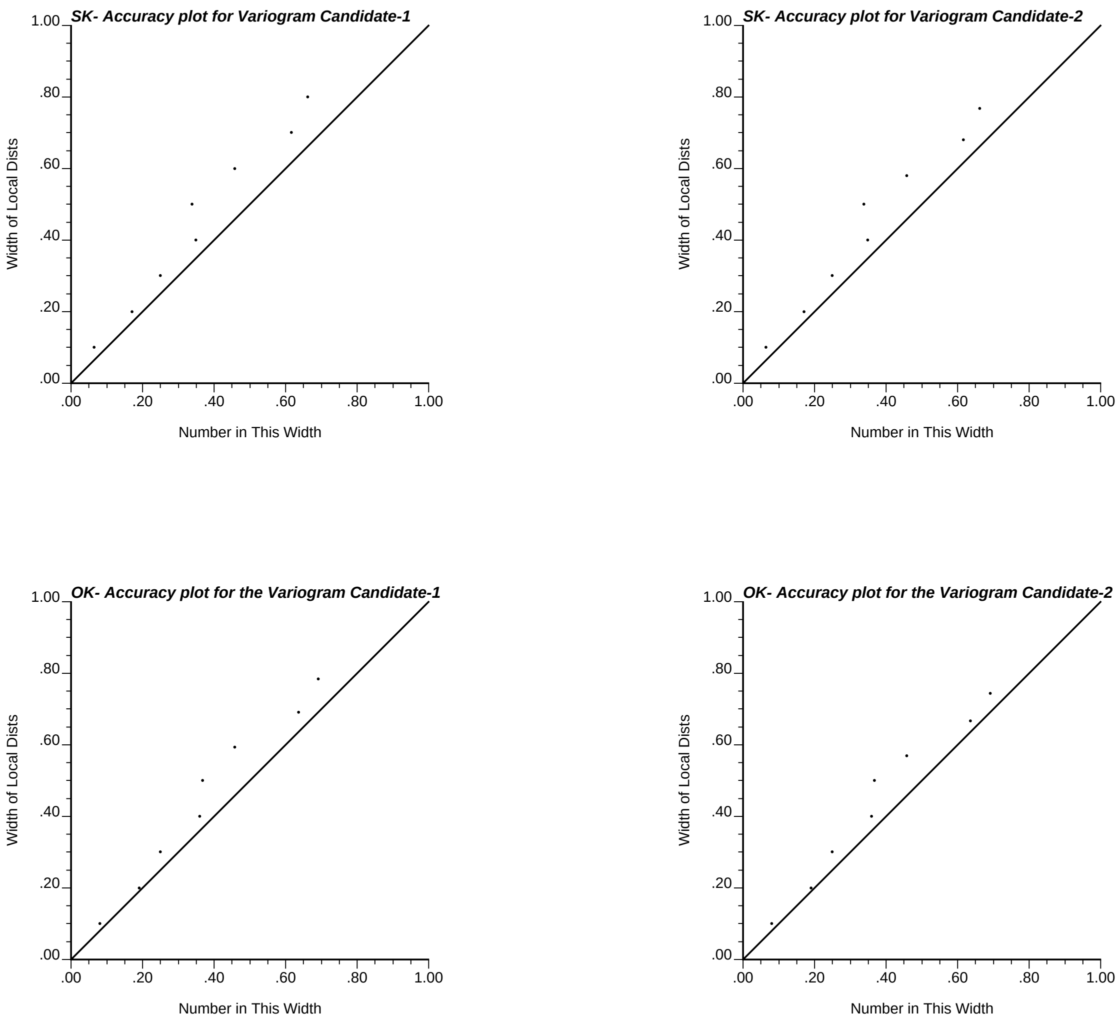

Figure 11: Accuracy plots for the four alternatives.