A Conditional Finite-Domain (CFD) Approach to Parameter Uncertainty

Olena Babak and Clayton V. Deutsch

Centre of Computational Geostatistics
Department of Civil & Environmental Engineering
University of Alberta

A new method for assessing uncertainty in the input histogram parameters to the modeling process is introduced. This method, referred to as the Conditional Finite-Domain (CFD) approach, accounts for the two important factors related to the study area: size of domain and conditioning data. It is a stochastic approach based on a multivariate Gaussian distribution. The CFD approach is shown to be convergent in the sense of limiting uncertainty calculation, design independent and parameterization invariant. The performance of the CFD approach is illustrated in a case study focusing on the impact of the number of data and the range of correlation on the limiting uncertainty in the parameters. The Spatial Bootstrap method and Conditional Finite-Domain approach are compared. The importance of the lower and upper possible values of the variable of interest in determining the limiting uncertainty using the CFD approach to uncertainty in the statistic of interest is documented and uncertainty associated with changes in the variogram is discussed.

Introduction

The input parameters to the modeling process are often uncertain. These parameters are the histogram and the variogram or the training images that contain all spatial features relevant to describe the reservoir under consideration. Presently, the only practical technique for quantification of the uncertainty in the histogram is the spatial bootstrap method (Journel and Bitanov, 2004). This approach allows assessment of uncertainty by resampling original data accounting for their spatial correlation/configuration.

One of the apparent limitations of the spatial bootstrap is that it allows only quantification of uncertainty of order one in the histogram. It is not evident how one would proceed to use this approach to address the question of uncertainty in the uncertainty in the histogram and, in general, uncertainty of any order in the histogram as well as the limiting (expected) uncertainty in the histogram.

In this paper we propose a new technique, referred to as the Conditional Finite-Domain (CFD) approach to parameter uncertainty, which allows quantification of the limiting uncertainty. This technique is based on sequential Gaussian simulation (Deutsch, 2002). Thus, it allows direct accounting for the two important study area factors: conditioning data and size of domain. Note that the spatial bootstrap approach accounts for neither of these factors.

The other disadvantage of the spatial bootstrap is that it does not account for all possible data in the area of interest. It is almost always the case that by additional sampling, we can observe in the area of interest some lower and higher values of the variable of interest than those previously sampled. Therefore, not including such possibility in the uncertainty assessment process can lead to underestimation (or overestimation) of uncertainty. To address this issue, we also propose an approach for determination of the possible higher and lower values of the variable of interest. The obtained values can be easily incorporated into the limiting uncertainty assessment strategy.

We start with short explanation and description of the methodology underlying the Conditional Finite-Domain approach to parameter uncertainty. Then, for investigation of the properties of the proposed approach, we conduct a case study based on the small GSLIB (Deutsch and Journel, 1998) dataset cluster.dat. In the section Real Data Example we compare the spatial bootstrap uncertainty in the mean with the one predicted by Conditional Finite-Domain approach for the real data set, Amoco 2D data. Finally, we summarize and discuss the obtained results.

Methodology

It is common to view one full grid of nodes as an observation from the multivariate distribution. For example, if the distribution of the variable of interest is normal, a realization simulated by sequential Gaussian simulation (SGS) on the grid of m by m nodes can be viewed as an observation from m^2-variate normal distribution. We propose to approach this problem differently.

Assume that n observations of the variable of interest come from the distribution that is fully defined by its mean, and covariance matrix, which is completely specified by a known variogram model. Then if we simulate the full grid of node values in the area of interest, we can base uncertainty assessment not on the full grid of simulated values, but rather on the sub samples of it. This is because it follows from the above made assumption that every set of simulated data which have the same configuration as the original data can be considered as an observation from the same underlying distribution as the original data. Therefore, the assessment of uncertainty in the statistic of interest made based on the simulated data with the same configuration as the original data is reasonable and fair.

When applying SGS to obtain the full grid of simulated values, we do not work directly with the original data. We transform the data to standard normal distribution prior to simulation, generate data using a Gaussian distribution (distribution which is fully defined by its mean and variance), and then back transform data to original units. Thus, it becomes clear that even if the data do not belong to the distribution which is fully defined by its mean and covariance matrix, we can base the assessment of uncertainty in the statistic of interest on the simulated data by SGS, if these data have the same configuration as the original data.

Let us now describe the procedure for choosing the simulated data from the N generated realizations of the study domain for subsequent uncertainty assessment.

Based on the full grid of simulated values, one can select any desired number of data combinations, say K, using translation and/or rotation with respect to some centre of the original data, which have the same configuration as the original data and belong to the study domain. The same K simulated data combinations can be found for all other N-1 simulated realizations. Then using all combinations, we assess and quantify the uncertainty in the statistic of interest in the usual way. For example, if the statistic of interest is the mean, the distribution of uncertainty in the mean is given by the NK mean values for the data combination. The uncertainty in the mean is then quantified, say with standard deviation of the uncertainty in the mean distribution.

The above described procedure for finding uncertainty can be applied several times with the aim of finding the limiting uncertainty in the statistic of interest. Specifically, in order to assess the limiting uncertainty, one needs to use reference distributions in the subsequent generation of the *N* realizations. Reference distributions are assembled based on the simulated realizations in the following way. The interval of variability of the variable of interest for each realization is divided into several parts (subintervals) and the number of observations obtained in each part is recorded separately. In subsequent uncertainty assessment, the midpoint of these subintervals and their associated probability, calculated as number of observation divided by total number of observations, are used as a reference distribution. The same procedure is repeated for each simulated realization.

Note that, when we assess the uncertainty of order one in the statistic of interest, we either apply SGS without any reference distributions or perform SGS with reference distributions taken as realizations of the spatial bootstrap approach applied with a reference variogram model of the variable of interest. Alternatively, one can perform SGS with arbitrarily chosen reference distributions provided they describe adequately the study domain. We will show later in the case study that the 'correct' (as explained above) starting reference distributions do not change the limiting uncertainty.

Now we will summarize the proposed approach, to which we will refer to as Conditional Finite-Domain (CFD) approach to parameter uncertainty, in the following algorithm.

1. Apply SGS to create *N* possible realizations of the variable of interest;
2. Calculate and quantify the uncertainty of order 1 in the statistic of interest. Establish the reference distributions to be used in the subsequent assessment of uncertainty in the statistic of interest.

…

2*k*-1. Use the reference distributions obtained in Step 2*k* - 1 to create *N* realizations of SGS using available conditioning data;

2*k*. Calculate and quantify the uncertainty of order *k* in the statistic of interest. Establish new reference distributions;

where $k = 2, \ldots, \infty$.

The CFD approach for the assessment of uncertainty of order *k* in any statistic of interest is convergent. However, one should not expect convergence in the usual sense. The CFD approach relies on the simulations, thus, it is a stochastic algorithm. Therefore, similar to the Markov Chain Monte Carlo (MCMC) approach, we observe a 'burn-in' period, when the uncertainty in the parameter of interest will increase, or decrease to the point where the parameter uncertainty starts to stabilize. 'Stabilization' phase corresponds to the fluctuation of the limiting parameter uncertainty around some constant value, which defines the expected limiting parameter uncertainty.

As we have already discussed in the Introduction section, it is very important to incorporate the range of all possible values of the variable of interest in the limiting uncertainty assessment. This is especially true if the number of data on the variable of interest is limited. Incorporation of the range of all possible values of the variable of interest can be achieved by the lower and upper tail options in SGS. To determine the tail values we can be use expert judgment or some other information. If, unfortunately, we do not have any information on the lower and upper level on possible values of the variable of interest, then we propose to find them approximately using the following simple procedure.

1. Fit a nonlinear regression to the cdf $F(z)$ of the variable of interest. Check if the fit is satisfactory. If yes, then proceed to 2);
2. Determine the lower and upper tail values as the intersection of the fitted nonlinear regression curve with horizontal lines $F(z)=0$ and $F(z)=1$, respectively.

We will show later in the case study that the lower and upper tail values have a major effect on the limiting uncertainty value as well as on the mean and other parameters of the limiting uncertainty distribution.

Case Study Setup

To give an illustration of the CFD approach, the case study was conducted using the GSLIB dataset from the file cluster.dat. This dataset consists of about 100 observations that are sampled on a random stratified grid and 40 observations that are clustered in high valued areas. We discarded the clustered data. The distribution of data is approximately lognormal with a mean of 2.5 and a standard deviation of 5.0. The data variogram is isotropic spherical with range 10 and contribution of 0.7,

$$\gamma(h) = 0.3 + 0.7Sph\left(\frac{h}{10}\right)$$

The 2-D study area is of the size 50 by 50 meters.

A mean of the global distribution was considered to be the statistic of interest. The uncertainty in the mean was measured by the standard deviation. The standard deviation is used because it is in the units of the data and linearly related to the width of a probability interval (approximately).

In our analysis we vary the number of data and the variogram range. The subsets of size 5, 10 and 20 values are considered. The minimum and maximum data values of 0.16 and 5.05, respectively, are chosen to coincide, in these three subsets. Moreover, the subsets are inclusive, that is, subset of 20 data contains the subset of 10 data and the subset of 10 data contains the subset of 5 data entirely. The data locations and data distributions are shown in Figure 1 for each of three chosen subsamples. The four cases are considered for the range of correlation: 5, 10, 20 and 50 meters.

The uncertainty in the input parameters will be affected by the amount of local data and the spatial correlation. Of course, one would expect that uncertainty is the least for the largest dataset and should increase as the size of dataset decreases; however, the change in the uncertainty as the range of correlation increases is not so predictable. Further, one might think that the uncertainty should increase with increase in the range as it increases in the spatial bootstrap approach. On the other hand, we can think that the uncertainty should decrease because the data values become more correlated to all other data in the area of interest, thus, in estimation and simulation more information is used, and the results are more reliable.

The main interest of this case study is investigate two important questions related to the influence of the number of data and the range of correlation on the limiting uncertainty in the CFD approach to parameter uncertainty. Moreover, we will also conduct the comparison study of the spatial bootstrap and the CFD approach to parameter uncertainty and investigate such important aspects of the CFD algorithm as:

- the influence of parameterization, that is, the scale on which the data is simulated on the limiting uncertainty. This aspect of the study is related to consistency of the proposed method. A good method should provide the same result for the limiting uncertainty independent of the scale on which result is simulated. Such a method is considered consistent. If the method is inconsistent, we would not know which scale is the best and results in a decision on which scale should be chosen to give the limiting uncertainty;
- the effect of the input reference distributions on the limiting uncertainty. A good approach to parameter uncertainty should be independently designed; that is, if the reference distributions equally describe the study domain adequately, they should yield the same result for the limiting uncertainty.

The other problem which we will consider is the effect of the lower and upper tail values on the limiting uncertainty. Of course, the upper and lower possible values of the variable of interest significantly impact the limiting uncertainty. The magnitude of this impact depends on the difference between the minimum and maximum observed values and the tails as well as on the number of data.

The data in CFD approach are simulated on a grid of 100x100 blocks of the size 0.5 by 0.5 meters. For each parameter uncertainty assessment step, 100 SGS realizations are used. The input reference distributions to the CFD algorithm are chosen to be the same as the data, that is, no reference distributions are used. The combinations for determining the subsequent reference distributions are obtained by rotation on angle $\alpha \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ of the original data locations with respect to the randomly selected centre $o = [o_x, o_y]$, where $o_x, o_y \sim U(5,45)$. The data combinations are constrained to be within the area of 50 x 50 meters. In total, for each uncertainty calculation 100 successful data rotations are considered. The limiting uncertainty in the statistic of interest is calculated as the average of the uncertainty values in the ‘stabilization’ phase. For comparison purposes, the tail options in SGS are chosen first the same as minimum and maximum data values. Then they are chosen to be equal to 0.1 (lower tail) and 10.27 (upper tail); these values are chosen from the same initial data set ‘cluster.dat’.

Case Study Results

Sensitivity of the limiting uncertainty in the mean to the change in the number of data is shown in Figure 2. This figure clearly shows the ‘burn-in’ period and ‘stabilization’ phase. The effect of the variogram range on the limiting uncertainty in the mean is shown in Figure 3. Note that the scales in the plots of Figures 2 and 3 are quite different. Looking at Figures 2 and 3, one can note slight stochastic fluctuations in the limiting uncertainty. To decrease the magnitude of fluctuations it is necessary to use a larger number of realizations and/or finer simulation grid.

Table 1 and Figure 4 summarize teh results shown in Figures 2 and 3. Table 1 as well as Figure 5 shows the corresponding results for uncertainty in the mean obtained based on 100 realizations of the spatial bootstrap approach.

Note that each experimental uncertainty point shown in Figures 2 and 3 was obtained following the steps of the CFD algorithm. Schematic representation of the calculations performed in one step is shown in Figure 6.

Figure 7 illustrates that with the change of parameterization (simulation grid), the limiting uncertainty predicted by the CFD approach remains virtually unchanged. Figure 8 shows the invariance of the limiting uncertainty with respect to the 'correct' input reference distributions. As 'correct' input reference distributions are considered spatial bootstrap realizations obtained based on the same variogram model as in the subsequent assessment of the limiting uncertainty and reference distributions are taken to be the same as the data at hand. Figure 9 illustrates the impact of 'incorrect' reference distribution. In the case shown in Figure 9, an 'incorrectly' chosen reference distribution increased the magnitude of the limiting uncertainty in the mean.

And, finally, Table 2 summarizes the effect of the tails (upper and lower possible values on the variable of interest) on the limiting uncertainty in the mean for the subsets of size 5, 10 and 20 data considered in this case study.

Real Data Example

The following example is based on Chu and Xu (1995). A 10500 by 10500 ft reservoir layer is considered. There are 62 wells in the study area. The location map of the well data is shown in Figure 10. The variable of interest is porosity. The naïve and declustered histograms of porosity are shown in Figure 11. The two horizontal variograms of porosity are shown in Figure 12. The variograms are in North-South direction (direction of major continuity) and East-West (direction of minor continuity).

The main interest of this study is assessing uncertainty in the mean of the porosity distribution as well as investigation of the variogram uncertainty. Despite the fact that the CFD approach to parameter uncertainty assumes that the variogram is known, it is often the case, like in the present example, that the true variogram model is not available. Thus, it would be interesting to assess how variograms corresponding to the mean uncertainty in the stabilization phase are different from the input variogram model.

The data in the CFD approach are simulated on the grid of 105x105 blocks of size 100 by 100 ft. For each step of parameter uncertainty assessment, 100 SGS realizations are used. The combinations for determination of the subsequent reference distributions are obtained by rotation on angle $\alpha \sim U\left(-\frac{\pi}{30}, \frac{\pi}{30}\right)$ of the original data locations with respect to randomly selected centre $o = [o_x, o_y]$, where $o_x, o_y \sim U(5100{,}5300)$. Data combinations are constrained to be within the area of 10500 x 10500 ft. In total, for each uncertainty calculation 25 successful data rotations are considered. The lower and upper tail options in the SGS are set to 4 and 11.4, respectively. The limiting uncertainty in the porosity mean is then calculated as the average of the uncertainty values in the 'stabilization' phase.

The uncertainty in the mean of porosity as well as the mean of the uncertainty in the mean distribution as functions of the uncertainty order are shown in Figure 13. The limiting uncertainty in the mean obtained based on the CFD approach is shown in Figure 14. Note that this figure also shows the limiting uncertainty in the mean obtained based on 1000 realizations of the spatial bootstrap approach. While results of the two methods for the mean of the uncertainty distribution are virtually the same, the standard deviations of the uncertainty distribution, measures of uncertainty, are very different. The limiting uncertainty predicted by CFD approach is much smaller than the one predicted by the spatial bootstrap due to the fact that the 62 data show significant spatial correlation with each other.

Uncertainty in the variogram is shown in Figure 15. Clearly, variograms calculated based on the reference distributions are significantly different from the input variogram to SGS in the direction of major continuity. Reference distributions are even more continuous in this direction than Amoco data. For the direction of minor continuity the variograms calculated based on the reference distributions are very close to the input variogram to SGS.

Conclusions

In this paper a new method for assessing uncertainty in the statistic of interest is proposed. This method called the Conditional Finite Domain (CFD) approach to parameter uncertainty directly accounts for the conditioning data and size of the domain. To our knowledge, this method is the first approach that allows direct incorporation of these two important factors. The CFD approach is a stochastic approach based on a multivariate Gaussian model to determine the limiting (expected) uncertainty in the statistic of interest. The proposed method is convergent, design independent and invariant under parameterization.

The effect of varying the number of data and the range of correlation on the limiting uncertainty in the CFD approach was investigated through the case study. In general, the results were what we expected. That is, we observed that the uncertainty decreases quickly as the number of data increases. This result is also consistent with the spatial bootstrap predictions. With respect to the change in range of correlation, we observed that as the range of correlation increases, the uncertainty in the statistic of interest decreases. This is due to the fact that the conditioning data are more correlated with each other and more correlated to the locations being simulated. Note that the spatial bootstrap approach predicts the reverse the impact, that is, the uncertainty decreases if the range of correlation decreases. In general, we can conclude that for large relative (dimensionless) ranges of correlation, the uncertainty in the statistic of interest calculated by the CFD approach would be smaller than that predicted by the spatial bootstrap method. Note that this conclusion is also shown to be true in the real data example. On the other hand, we can also conclude that for small relative (dimensionless) ranges of correlation, the uncertainty in the statistic of interest calculated by the CFD approach would be generally higher than that predicted by the spatial bootstrap method. The higher uncertainty in the statistic predicted by the CFD approach is reasonable since geological data is never random. Note that for a pure nugget effect variogram, the uncertainty predicted by the CFD approach is significantly higher than that predicted by the spatial bootstrap method.

The other interesting question is variogram uncertainty. We never really know the true variogram model. Thus, investigation and incorporation of the variogram uncertainty in the geostatistical modeling is of great importance. In this paper, variogram uncertainty was investigated for the real geological data example. It was shown that the variogram of the reference distribution can be very different from the input variogram to sequential Gaussian simulation. Thus, when we are interested in the assessing the full uncertainty in the statistic of interest, we recommend that variogram uncertainty also be incorporated in the limiting uncertainty assessment. This can be done simply by applying SGS each time not only with a different reference distribution, but also with a different input variogram corresponding to that reference distribution.

Furthermore, in the case study, the impact of the lower and upper tail values on the limiting uncertainty value obtained using the CFD approach was investigated. It was shown that the limiting uncertainty can be unreliable if we do not consider the full range of possible data values on the variable of interest. A method for the calculating the lower and upper possible value of the variable of interest was also proposed in this paper.

Future areas for more detailed examination include: consider different statistics of interest for Conditional Finite Domain approach; study the impact of parameter uncertainty on the global uncertainty, for example, the impact of the parameter uncertainty on the global uncertainty in the recoverable resources; and investigate more closely the issue of variogram uncertainty.

References

Deutsch, C.V., 2002, *Geostatistical Reservoir Modeling*. Oxford University Press, New York.

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

Journel, A.G. and Bitanov, A., 2004, *Uncertainty in N/G ratio in early reservoir development*, Journal of Petroleum Science and Engineering, 44 (1-2), pp.115-130.

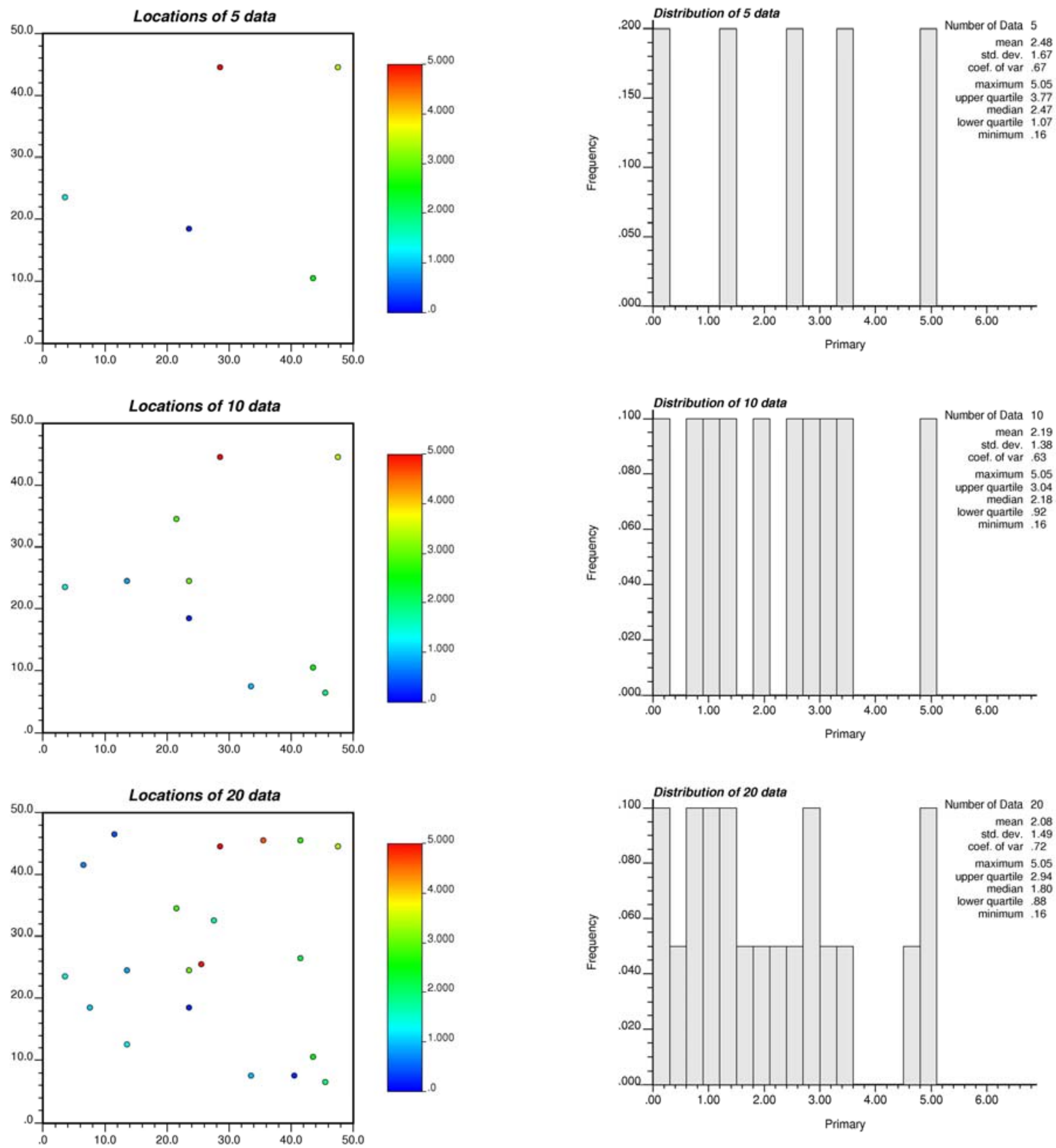
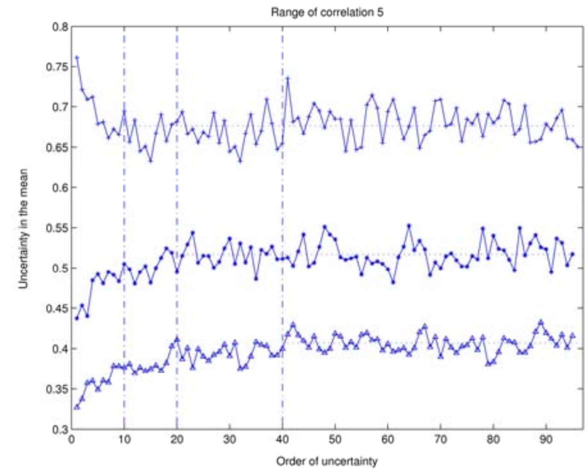
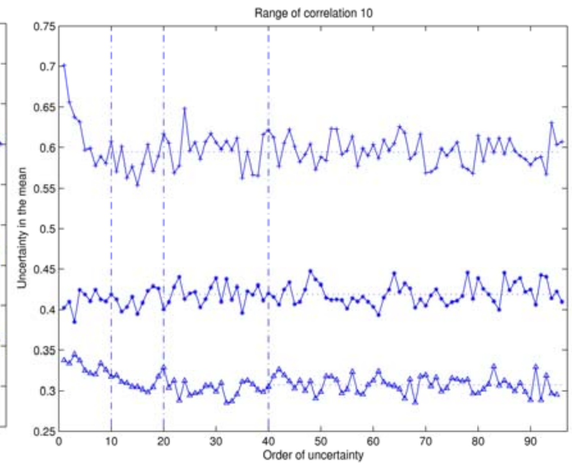


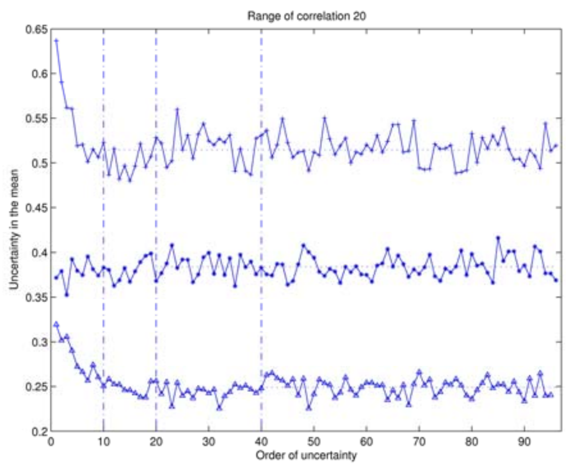
Figure 1: Location maps (left column) and distributions (right column) of the subsamples of size 5, 10 and 20 of the file data set ‘cluster.dat’.



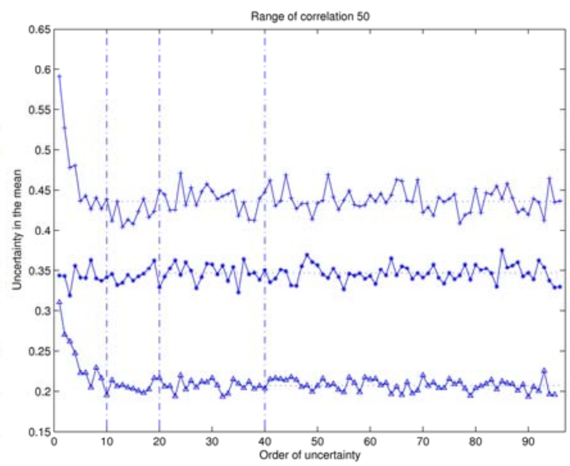
a)



b)



c)



d)

Figure 2: Effect of the number of data on the limiting uncertainty in the mean for the range of correlation a) 5; b) 10; c) 20 and d) 50. Pluses, asterisks and triangles denote the uncertainty in the mean for 5, 10 and 20 data, respectively. Horizontal doted lines denote the respective limiting uncertainties in the mean.

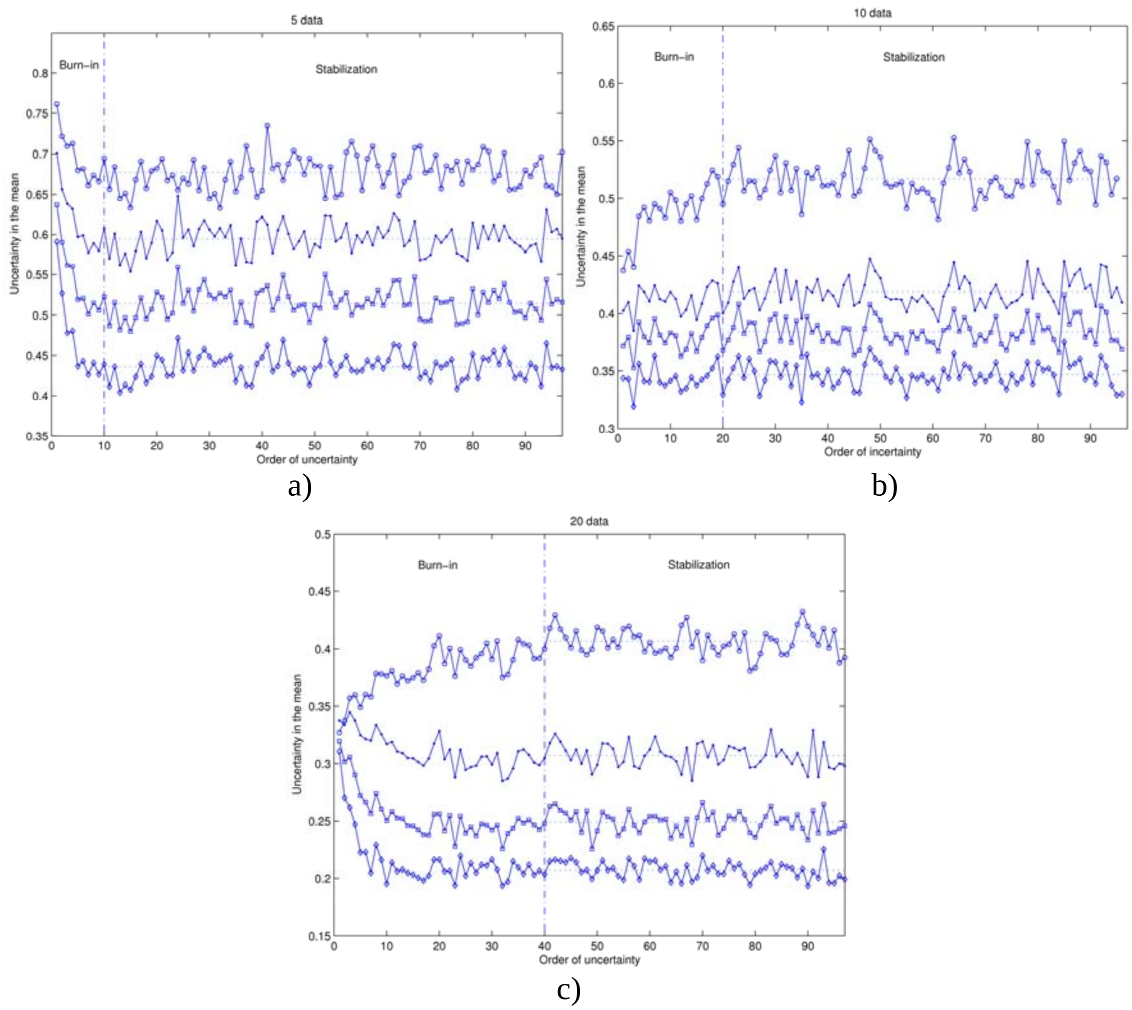
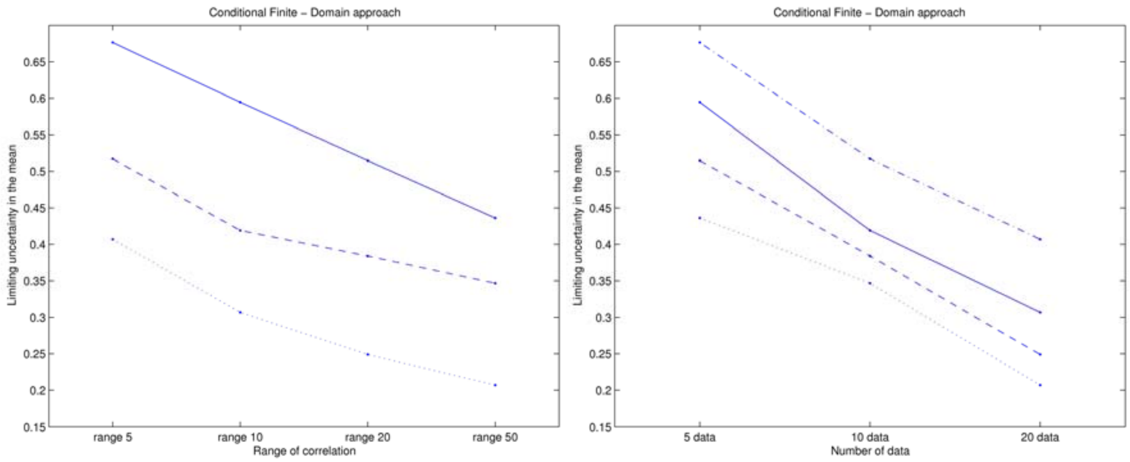


Figure 3: Sensitivity of the limiting uncertainty in the mean to the change in the range of correlation for: a) 5 data; b) 10 data; c) 20 data. Circles, dots, squares and diamonds denote the uncertainty in the mean for the ranges of correlation 5, 10, 20 and 50, respectively. Horizontal doted lines denote the respective limiting uncertainties in the mean.

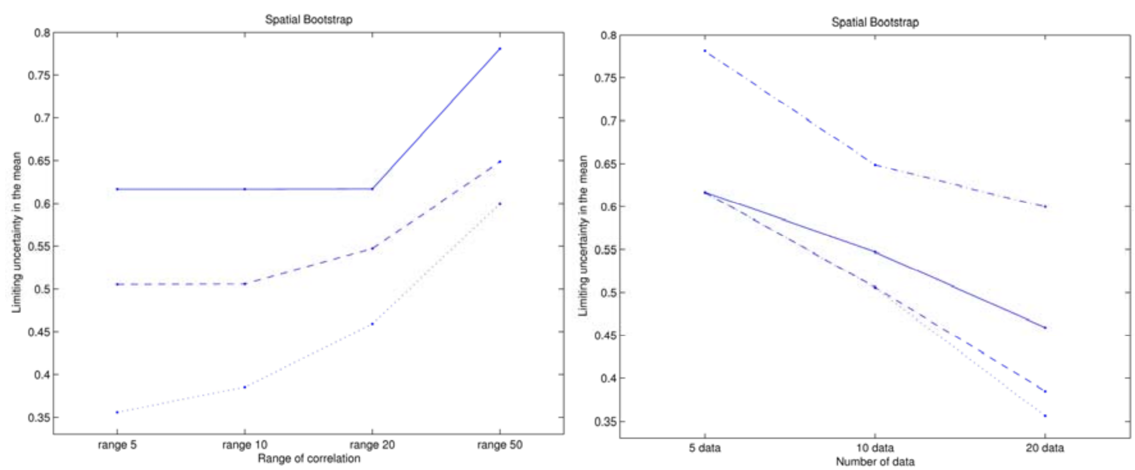
Table 1: Comparison of the Spatial Bootstrap and the Conditional Finite-Domain approach to the limiting uncertainty in the mean for 5, 10 and 20 data and ranges of correlation 5, 10, 20 and 50.

Number of data	**Mean (& std) of data**	**Method**	**Range 5**		**Range 10**		**Range 20**		**Range 50**	
			Mean of uncert. distrib.	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**
5 data	2.4820	SB	2.4618	0.6166	2.4618	0.6166	2.4618	0.6169	2.4623	0.7809
	1.6706	CFD	2.4773	0.6765	2.4887	0.5945	2.4944	0.5147	2.4684	0.4360
10 data	2.1940	SB	2.1697	0.5054	2.1687	0.5059	2.1738	0.5471	2.1696	0.6487
	1.3802	CFD	2.2306	0.5170	2.1700	0.4189	2.1582	0.3840	2.1437	0.3467
20 data	2.0780	SB	2.0660	0.3559	2.0712	0.3850	2.0710	0.4592	2.0614	0.5997
	2.2100	CFD	2.1643	0.4068	2.0438	0.3069	1.9318	0.2490	1.7050	0.2071



a) b)

Figure 4: a) Limiting uncertainty in the mean as a function of the range of correlation for the CFD approach. Solid, dashed and dotted lines correspond to the limiting uncertainty in the mean for 5, 10 and 20 data, respectively; b) Limiting uncertainty in the mean as a function of the number of data for the CFD approach. Dotted, dashed, solid and dash dot lines correspond to the limiting uncertainty in the mean for the range of correlation 50, 20, 10 and 5, respectively.



a) b)

Figure 5: a) Limiting uncertainty in the mean as a function of the range of correlation for the SB approach. Solid, dashed and dotted lines correspond to the limiting uncertainty in the mean for 5, 10 and 20 data, respectively; b) Limiting uncertainty in the mean as a function of the number of data for the SB approach. Dotted, dashed, solid and dash dot lines correspond to the limiting uncertainty in the mean for the range of correlation 5, 10, 20 and 50, respectively.

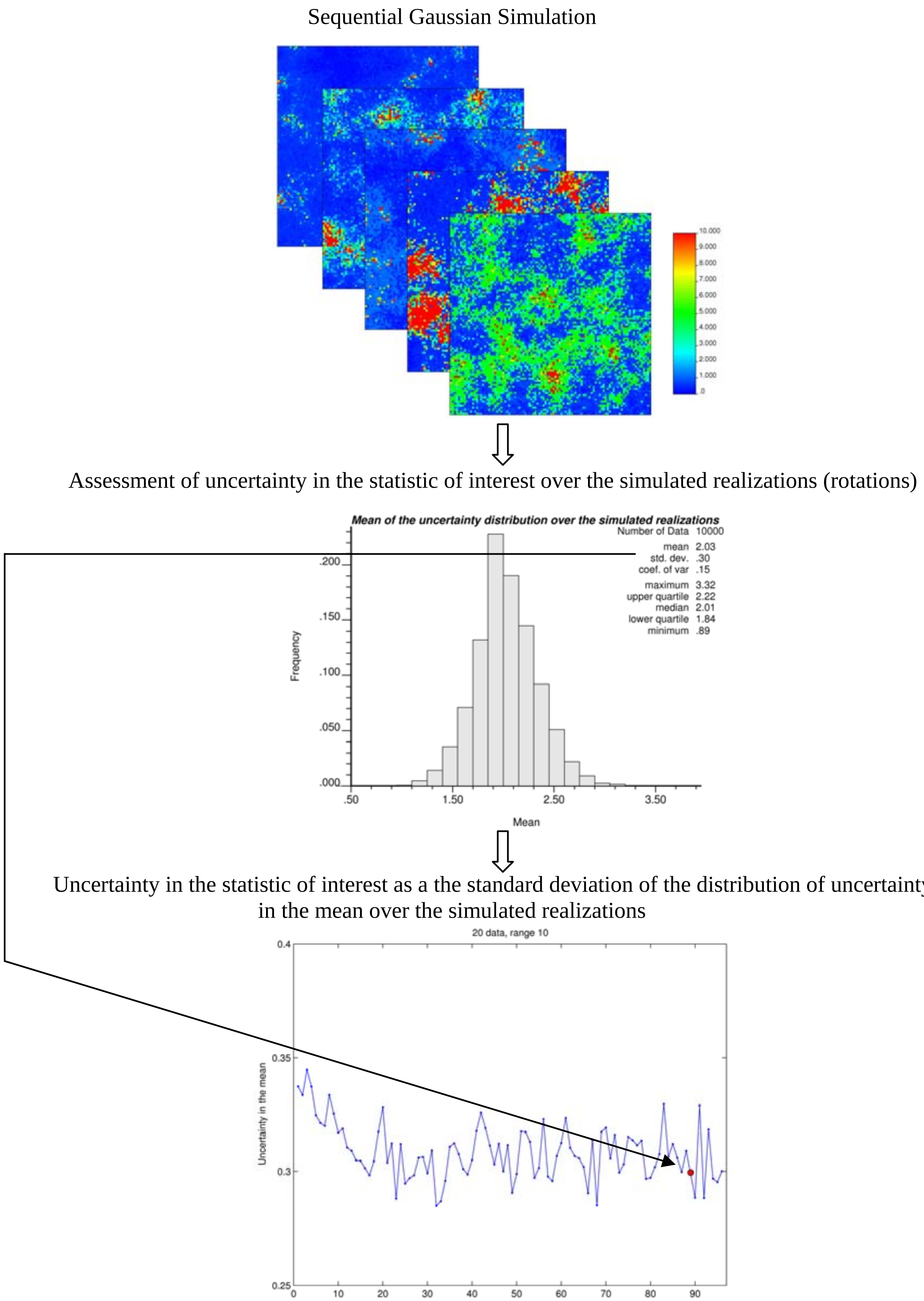
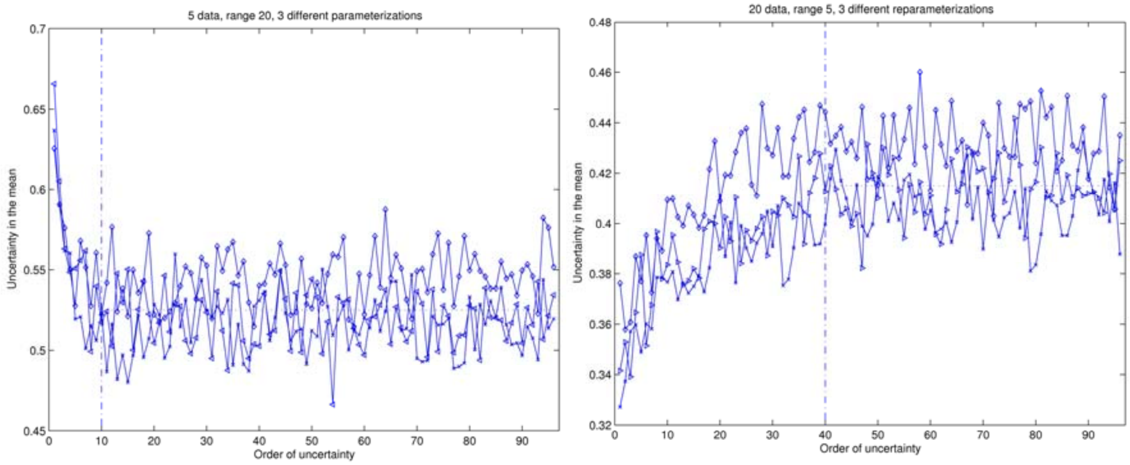
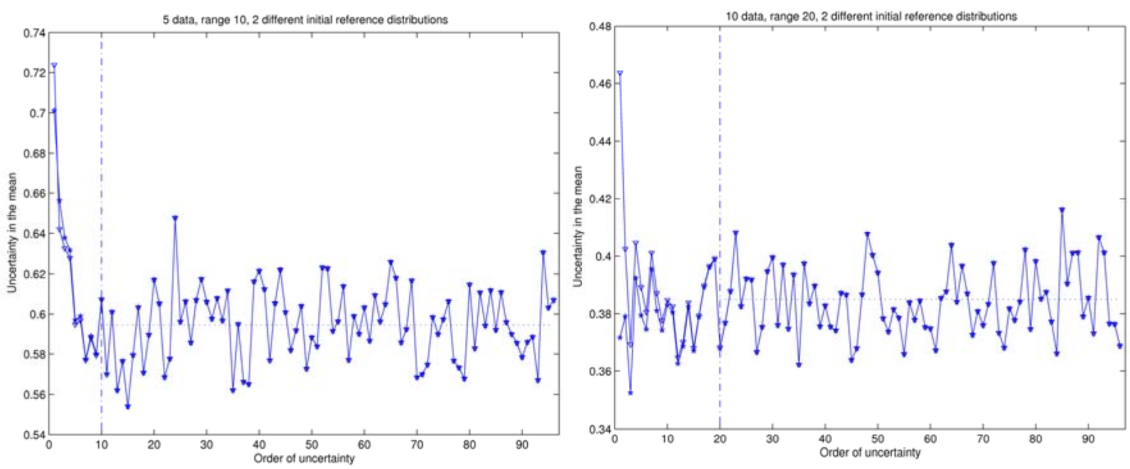


Figure 6: Schematic representation of the calculations performed in one step of the CFD algorithm.



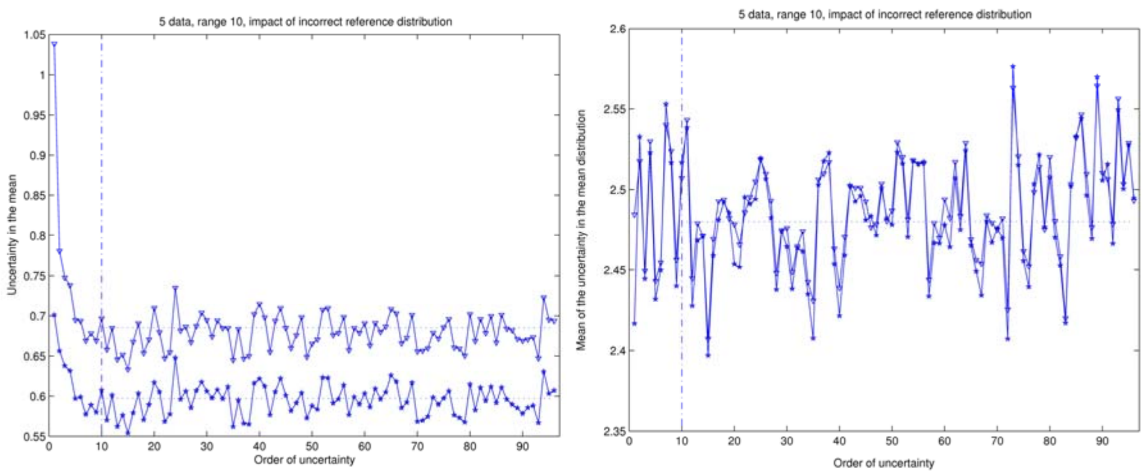
a) b)

Figure 7: Influence of the parameterization for: a) 5 data, range of correlation 20; and b) 20 data, range of correlation of 5. Triangles, crosses and diamonds denote the uncertainty in the mean obtained using simulation on the scale 0.25 by 0.25 blocks, on scale of 0.5 by 0.5 blocks, and on the scale 1 by 1 blocks, respectively.



a) b)

Figure 8: Influence of the 'correct' input reference distributions on the limiting uncertainty for: a) 5 data, range of correlation of 10; b) 10 data, range of correlation of 10. The different input reference distributions were taken as the realizations of the Spatial Bootstrap with true variogram model and the same as the data at hand.



a) b)

Figure 9: Impact of the 'incorrect' reference distribution: a) Standard deviation of the uncertainty in the mean distribution; and b) Mean of the uncertainty in the mean distribution. Stars and triangles denote the results for the uncertainty distribution obtained based on correct input reference distribution (reference distribution taken the same as data at hand) and incorrect reference distribution (reference distribution taken as the realization from the Spatial Bootstrap algorithm with spherical variogram model with range of correlation not 10, but 20) to the Conditional Finite Domain approach, respectively.

Table 2: Influence of the tails on the limiting uncertainty in the mean for 5, 10 and 20 data and ranges of correlation 5, 10, 20 and 50.

Number of data	**Mean (& std) of data**	**Method**	**Range 5**		**Range 10**		**Range 20**		**Range 50**	
			Mean of uncert. distrib.	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**
5 data	2.4820 1.6706	SB	2.4618	0.6166	2.4618	0.6166	2.4618	0.6169	2.4623	0.7809
		CFD (no tails)	2.4773	0.6765	2.4887	0.5945	2.4944	0.5147	2.4684	0.4360
		CFD (with tails)	2.9852	1.1342	2.8752	0.9493	2.7947	0.7706	2.6833	0.5876
10 data	2.1940 1.3802	SB	2.1697	0.5054	2.1687	0.5059	2.1738	0.5471	2.1696	0.6487
		CFD (no tails)	2.2306	0.5170	2.1700	0.4189	2.1582	0.3840	2.1437	0.3467
		CFD (with tails)	2.5593	0.7909	2.2822	0.5246	2.2251	0.4355	2.1929	0.3680
20 data	2.0780 2.2100	SB	2.0660	0.3559	2.0712	0.3850	2.0710	0.4592	2.0614	0.5997
		CFD (no tails)	2.1643	0.4068	2.0438	0.3069	1.9318	0.2490	1.7050	0.2071
		CFD (with tails)	2.8192	0.7585	2.2386	0.4100	2.0889	0.2752	2.0158	0.2433

Table 3: Comparison of the Spatial Bootstrap and the Conditional Finite-Domain approach to the limiting uncertainty in the mean for 5, 10 and 20 data and pure nugget variogram.

	Method	**5 data**		**10 data**		**20 data**	
		Mean of uncert. distrib.	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**	**Mean of uncert. distrib.**	**Std of uncert. distrib.**
Nugget variogram	SB	2.4618	0.6197	2.1499	0.4897	2.0663	0.3561
	CFD (no tails)	2.5030	0.8101	2.3478	0.6566	2.3388	0.5306

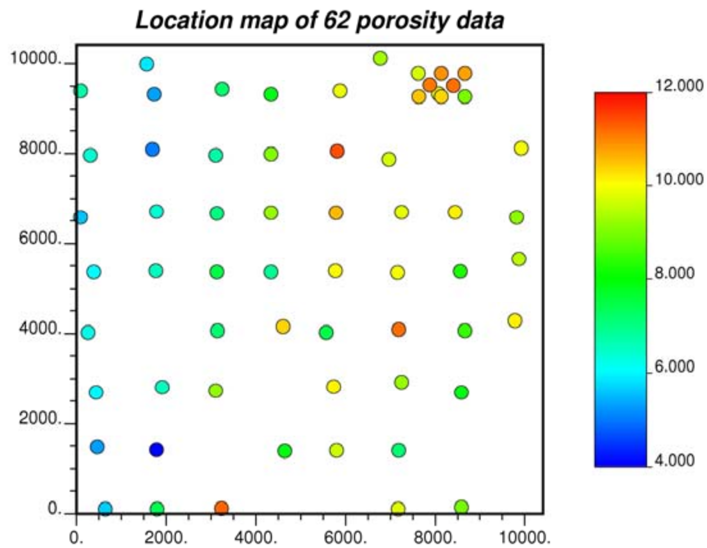
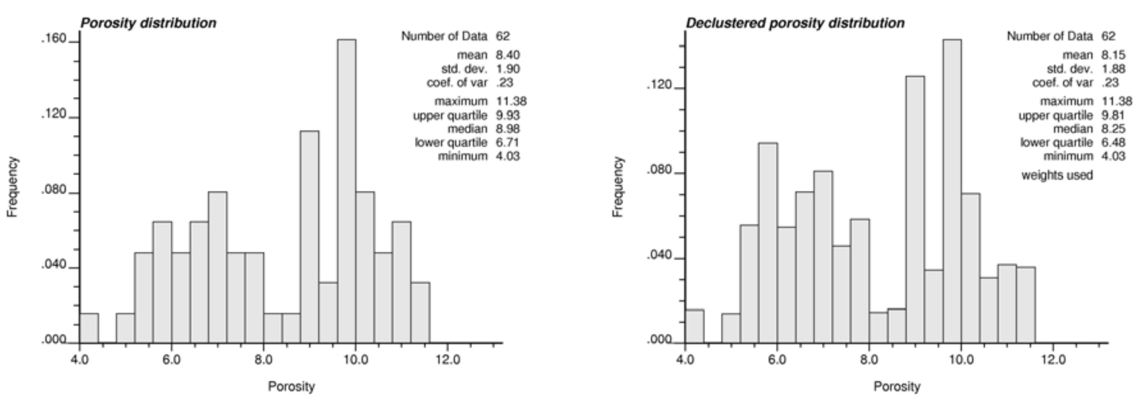


Figure 10: Location map of 62 porosity data of the data set AmocoData2-D.dat.



a) b)

Figure 11: a) The simple histogram (undeclustered) of porosity; b) Representative (declustered) histograms of porosity.

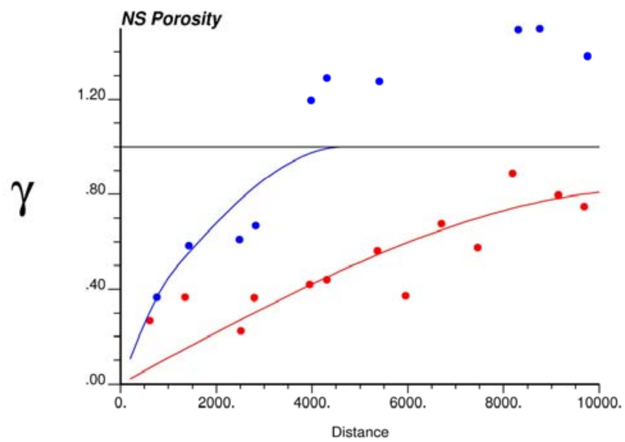


Figure 12: Two horizontal variograms of porosity. The variogram in North-South direction (direction of major continuity) is shown in light color, the variogram in East-West (direction of minor continuity) is shown in dark color.

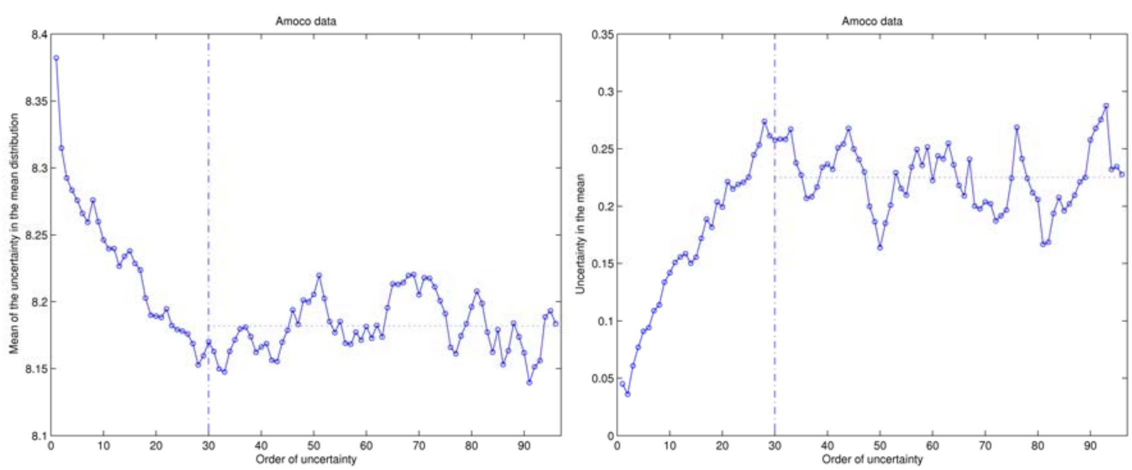
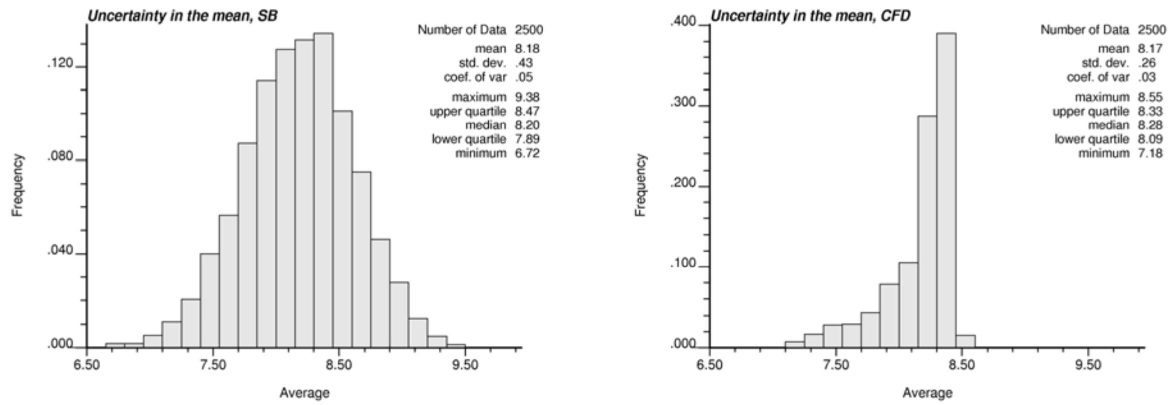
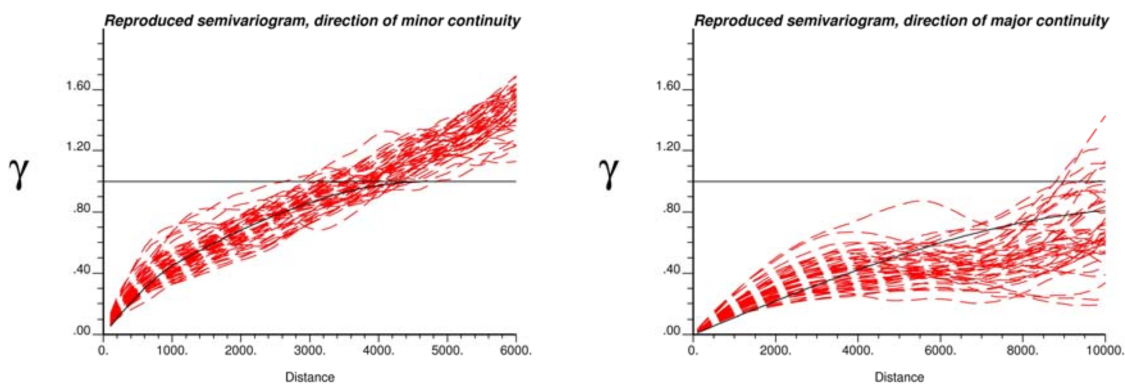


Figure 13: a) Mean of the uncertainty in the mean of porosity distribution as a function of the uncertainty order; b) Uncertainty in the mean of porosity as a function of the uncertainty order.



a) b)

Figure 14: Limiting uncertainty in the mean obtained based on: a) Spatial Bootstrap approach and b) Conditional Finite-Domain approach.



a) b)

Figure 15: Uncertainty in the variogram in the direction of a) minor, and b) major continuity when calculating the limiting uncertainty in the mean for the Amoco data.