Kriging in High Dimensional Attribute Space using Principal Component Analysis

Katrine Lange[1], Thomas Mejer Hansen[1], Juan Luis Fernández Martínez[2], Jan Frydendall[1] and Klaus Mosegaard[1]

[1]Technical University of Denmark, Copenhagen, Denmark. Email: katla@imm.dtu.dk
[2]Universidad de Oviedo, Oviedo, Spain

Abstract

Interpolation between measured well log properties is a well-known problem in seismic exploration. Most often 3D seismic data is available from which a large number of seismic attributes can be extracted quantifying various properties of the seismic data. Such attributes have been used to guide interpolation between well log parameters, using for example neural network, linear regression and cokriging. Full cokriging using seismic attributes as secondary data is though not feasible due to the complexity of inferring a full cross-covariance model. A recently proposed method suggests a kriging alternative to full cokriging, where kriging is performed in the attribute space, relying on a distance measure in the attribute space. One limitation is, however, that all attributes are considered uncorrelated, whereas in reality they are not. We propose to use principal component analysis (PCA) to transform the coordinate system in attribute space to a coordinate system given by the principal components. By construction the principal components are uncorrelated, and we can therefore apply kriging in the principal component space. In addition PCA naturally orders the principal components according to their variance contribution, and can therefore be used to select only the most significant principal components for the kriging, allowing easier inference of the covariance model.

Keywords: seismic inversion, seismic attributes, kriging interpolation.

1. INTRODUCTION TO KRIGING IN ATTRIBUTE SPACE

The objective of our work is to produce accurate and reliable estimations of the porosity levels in the subsurface of a reservoir. To do so we exploit knowledge of seismic attributes, such that the estimated porosity at a certain location relies on the subsurface geology rather than just the spatial location.

We propose to use the kriging technique[1] to interpolate between known values of the porosity level. One of the advantages of kriging is that, unlike traditional interpolation methods such as

linear regression and neural networks, it provides not only an estimate of the value itself but also an uncertainty estimate. The interpolation will be based on a distance measure in the attribute space rather than the more traditional distance measure in the physical XYZ space.

Outline of the method:

1) Normal score transformation of the porosity data to meet the assumption of data being samples from a continuous Gaussian distribution. Normalization of the position coordinates in attribute space for computational reasons.
2) PCA transformation of the attributes to ensure that their cross-correlation is zero.
3) Definition of a distance measure in the principal component space and computation of optimal covariance parameters by use of maximum likelihood estimation[4] given the choice of covariance model type.
4) Formulation of the kriging system given the covariance model, and the system is then solved with respect to the kriging weights.
5) Back transformation of the porosity values estimated in normal score space.

Using a subset of the attributes, and/or a subset of the principal components, the dimension of the space in which the interpolation is conducted can be reduced.

2. CASE STUDY: The South Arne Field

Hess Copenhagen provided us with data from the South Arne Field, which is a chalk reservoir situated in the Danish North Sea. The data set contains known values of 8 seismic attributes of nearly 76.000 points in a regular 2D XY-grid. The attributes associated with each point of the grid are the following: spatial coordinates UTM X, UTM Y and UTM Z, the two way travel times to the top and the base of the reservoir, both have been established from the seismic data, the amplitude and the dip at the top has also been extracted as attributes, and finally the acoustic impedance.

There are additionally 213 well site points where both the values of the seismic attributes and also the values of the porosity levels are known. These points are used to validate our results. To do so they are divided into two sets. The first set contains the data used in the interpolation, and the second set acts as a set of blind wells. Porosity levels are estimated at the location of the blind wells using only the data in the first set and the results are then compared to the known true values. The partition of the data set is done in accordance to the method described in[3]. This yields six different subsets of blind data with respectively {106, 53, 21, 178, 99, 144} points of the total 213 observation points as known data and the remaining data will act as blind data.

Figure 1 shows the known values of the porosity levels in the XY domain of the total data set and known value of each of the six data sets that holds blind data. This implies that the data in the second to the seventh subplot are subsets of the data shown in the leftmost subplot.

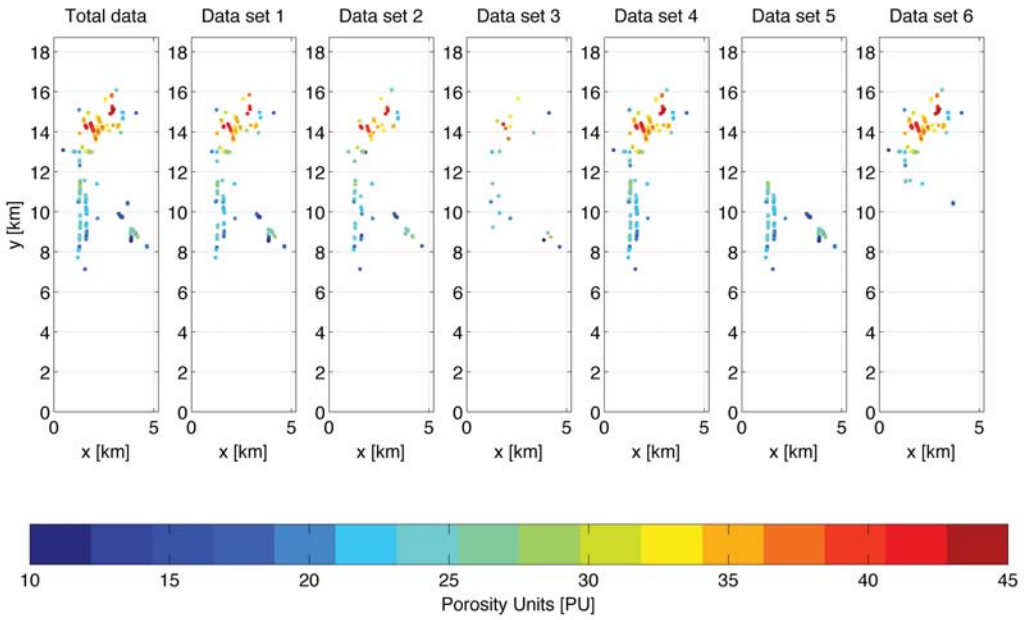

Figure 1: Known values of the porosity levels of the total data set and each of the six data sets containing blind data.

In the following subsection we present and discuss the kriging results. As a measure of performance of our method we will evaluate the root mean squared (RMS) error of the porosity estimates at the blind wells and the uncertainty estimates of the porosity levels estimated for the entire grid for the six different partitions of the known data.

2.1. Parameter choices

Before we evaluate the performance of our method we must discuss the effect of decisions that we have to make during the formulation and solution of the problem. The choices we have are regarding the following aspects:

1) *Partition of the known data to create a blind set.* We have the previously mentioned six different partitions that can also be seen in Figure 1.
2) *Selection of the attributes used for the interpolation.* It is likely that we can reduce the dimension of the transformed attribute space without reducing the quality of our results by only selecting the most relevant of the attributes and simply ignoring the rest.
3) *Number of principal components used.* It is also worthwhile to consider if one can reduce the dimension of the transformed attribute space further by selecting only a subset of the principal components.
4) *Type of covariance model.* We will choose a covariance model for each principal component consisting of two terms. The first term is a nugget effect model and the second term is a model of either of the types spherical, exponential or Gaussian. This term will be of the same type in all directions.
5) *Type of kriging method.* For this extended abstract we will use only kriging with a trend,

which in each direction of the principal components is modeled as a polynomial of degree one.

2.2. Attributes vs. principal components

Figure 2 shows the eight different attributes and Figure 3 shows the resulting principal components. It is interesting to see how the principal components, some more than others, resembles single attributes. This can be an inspiration in choosing the trend type of the kriging method.

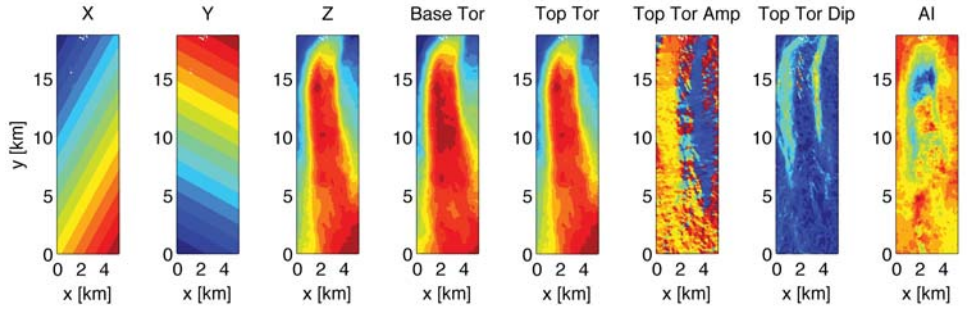

Figure 2: Each of the eight attributes plotted in the physical domain. Notice that the plots have been rotated slightly in order to minimize white space between the subfigures. The attributes do not share a common color scale, but red simply indicate high relative values and blue indicates low relative values.

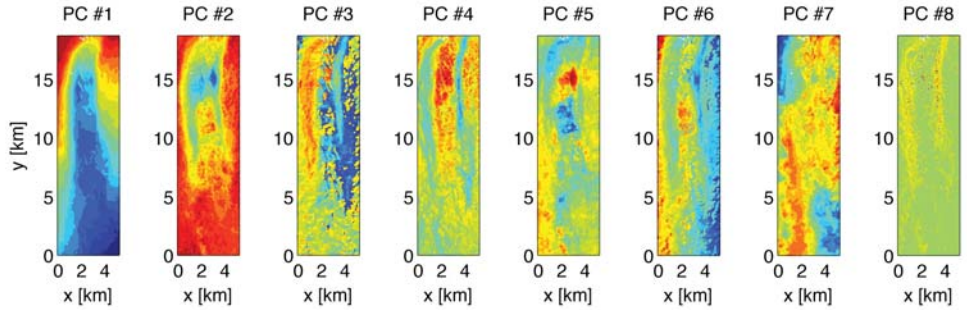

Figure 3: Each of the eight principal components computed based on the attributes shown in Figure 2. As in Figure 2 the principal components do not share a common color scale, but red and blue indicate respectively high and low relative values.

Figure 4 shows the transformation matrix used to transform the attributes of Figure 2 into the principal components seen in Figure 3. To compute the *i*th principal component the attributes has been weighted by the entries of the *i*th column of the transformation matrix. Notice for instance that to compute the third principal component, attribute six has been given a very high positive weight compared to the other attributes. Hence the structure in the third principal component resembles the structure in the top amplitude attribute.

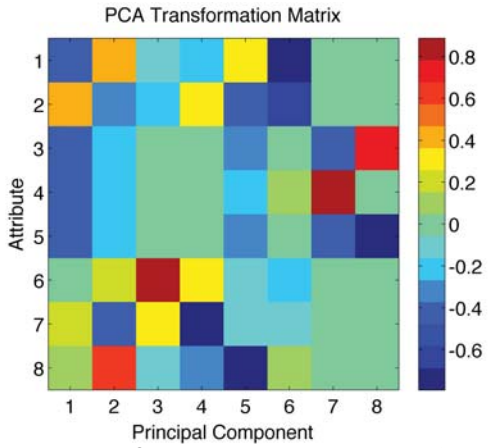

Figure 4: Transformation matrix used to transform the attributes in Figure 2 into the principal components shown in Figure 3.

2.3. Interpolation using a subset of the attributes

We choose the second term of the covariance model to be a Gaussian model. We now compute the RMS error for six different partitions of the data varying the subset of the attributes used in the interpolation. In all cases the full set of principal components is used. The results can be seen in Table 1. For each of the six data sets containing blind data this table shows the RMS error at the blind wells when using different sets of attributes and the complete set of principal components.

Table 1: RMS error in porosity units of the kriging results when using different subsets of the attributes and all of the principal components.

Data set	X, Y, Z	Impedance	X, Y, Z and impedance	X, Y, Z, amplitude and impedance	All attributes
#1	2.96	4.04	2.91	2.98	3.05
#2	4.22	4.25	3.58	3.82	3.94
#3	5.11	4.80	4.14	4.41	4.51
#4	6.65	4.64	5.52	5.61	5.79
#5	10.2	9.15	9.37	8.91	9.29
#6	13.6	5.82	5.20	5.33	5.90

From the results in Table 1 we see that while the choice of blind data has great influence on the level of the RMS error we experience approximately the same behavior of the RMS error when we change the subset of the attributes used in the interpolation. The second column shows the result when we use only the spatial coordinates while the third column holds the results for when we only use the impedance, which we expect is strongly correlated with the porosity.

For the majority of the data sets the lowest RMS error is achieved by using the spatial coordinates together with the impedance. Adding the amplitude at the top as a fifth attribute seems, in most cases, to increase the error slightly. The same applies for adding the remaining attributes, i.e. for the errors in the last column.

According to the table, using data set #5 results in a relative high RMS error compared to using the other data sets. Figure 1 can explain this. Here we see that the known values of the porosity levels for data set #5 all are relatively low values (all dots are blue or green). This data set is therefore not a representative subset of porosity levels for the entire grid and that complicates the interpolation.

2.4. Reducing the dimension of the principal component space

We now turn to reducing the dimension of the transformed attribute space by reducing the number of principal components included. The principal components are chosen according to their variance contributions.

First we consider the test case from the fourth column of Table 1, i.e. we include four attributes namely the spatial coordinates and the impedance. Table 2 shows the RMS errors when using one to four principal components respectively. Four attributes yields in total four principal components. The table shows a trend, which is that the more principal components we include the lower is the error. Data set #2 is the exception for which three and four principal components yield approximately the same error.

Table 2: RMS error in porosity units of the kriging results when using a subset of the principal components. We have used the attributes X, Y, Z and impedance that result in the last column of Table 2 being a duplicate of the fourth column of Table 1.

Data set	1 out of 4	2 out of 4	3 out of 4	All 4 principal components
#1	4.67	3.56	3.02	2.91
#2	5.58	4.11	3.55	3.58
#3	5.55	4.47	4.28	4.14
#4	5.74	5.46	6.02	5.52
#5	10.3	10.6	9.27	9.37
#6	7.87	6.73	5.86	5.20

In order to derive the general trend we consider all attributes and compute the RMS error when selecting one to all of the eight principal components. These errors can be seen in Figure 5. Notice that data set #5 has been omitted from the figure as its errors are of significantly greater magnitude due to the previously mentioned difficulties regarding the partition of the data.

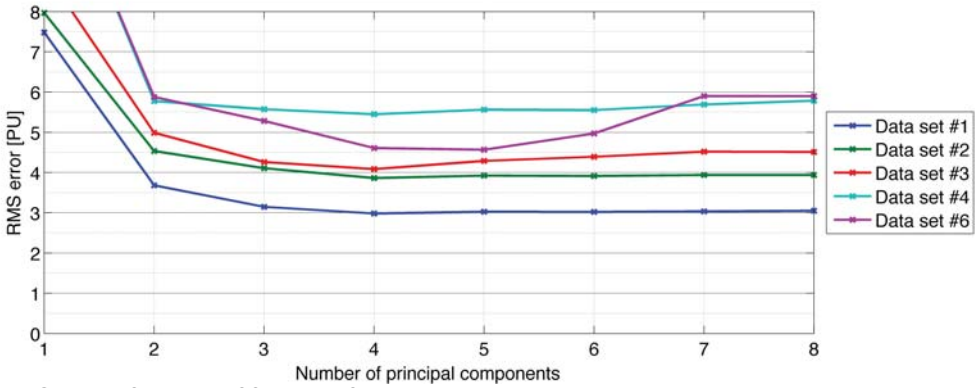

Figure 5: RMS errors for each of five out of the six data sets holding blind data. The interpolation has been done using between one and all principal components. All attributes are used. Data set #5 has been omitted as the errors are of significantly greater magnitude

According to Figure 5 the minimum RMS error is not always achieved by simply selecting all of the principal components. On the contrary, the figure indicates that for data set with a large number of attributes it is some times favorable to only include a subset of the principal components, for instance for data set #3 and data set #6.

To compare the errors seen in Figure 5 to the errors of Table 2, Table 3 shows the errors for each of the six data sets when the number of principal components selected is four. This means the dimension of the transformed attribute space in both cases is four.

Table 3: RMS errors when using all attributes and four out of the eight principal components. The values in this table can also be seen in Figure 2. These errors are comparable to the last column of Table 2, as the dimension in both cases has been reduced to four.

Data set	4 principal components
#1	2.98
#2	3.86
#3	4.09
#4	5.45
#5	8.45
#6	4.61

Comparing the values of Table 3 to the values in the last column of Table 2 one can conclude that in most of the cases selecting four relevant attributes and using all four principal components yield smaller RMS errors than using all of the eleven attributes but only selecting the four most significant principal components.

2.5. Type of covariance model

Again we turn to the problem of interpolating using the four attributes X, Y, Z and the impedance and all four principal components. Up until now all covariance models has had a Gaussian model as their second term, but Table 4 and Figure 6 show respectively the RMS errors and the porosity estimate varying the type of covariance model.

Table 4: RMS errors when using four attributes – X, Y, Z and the impedance – and all principal components. The second term of the covariance matrix has been modeled as respectively a spherical, an exponential and a Gaussian model. The kriging results can be seen in Figure 6.

Data set	Spherical	Exponential	Gaussian
#1	3.98	3.99	2.91
#2	4.84	5.01	3.58
#3	6.71	6.74	4.14
#4	8.75	7.30	5.52
#5	11.7	12.5	9.37
#6	12.6	8.97	5.20

The table reveals that a Gaussian model yields by far the smallest RMS errors, and the spherical and exponential type of covariance model results in errors of approximately the same magnitude. Which of the latter two models that in this specific test case is the most favorable depends on the data set.

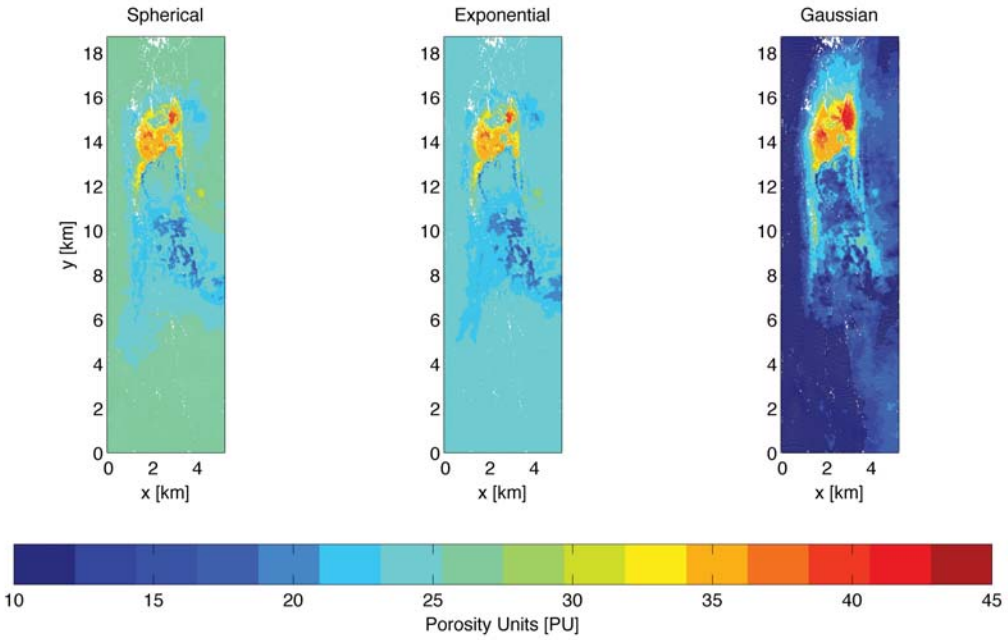

Figure 6: Estimated porosity levels for the same test case but where the second term of the covariance model has been modeled as respectively a spherical, an exponential and a Gaussian covariance model. We have used data set #1. The results have the RMS errors seen in Table 4, second row.

Figure 6 shows the estimated porosity levels for data set #1. It is noticed that the Gaussian type of covariance model allows us better to capture the variety of the subsurface especially for the part of the XY domain in which we have no known data.

2.6. Uncertainty estimates

Figure 7 and Figure 8 shows uncertainty estimates of the kriging results of the test case with data set #1. Again we have used the four attributes X, Y, Z and the impedance and the full set

of principal components. The second term of the covariance model is a Gaussian model, which means the kriging results are identical to those that can be seen in Figure 6 to the right. The visual difference is due to a change in the color bar.

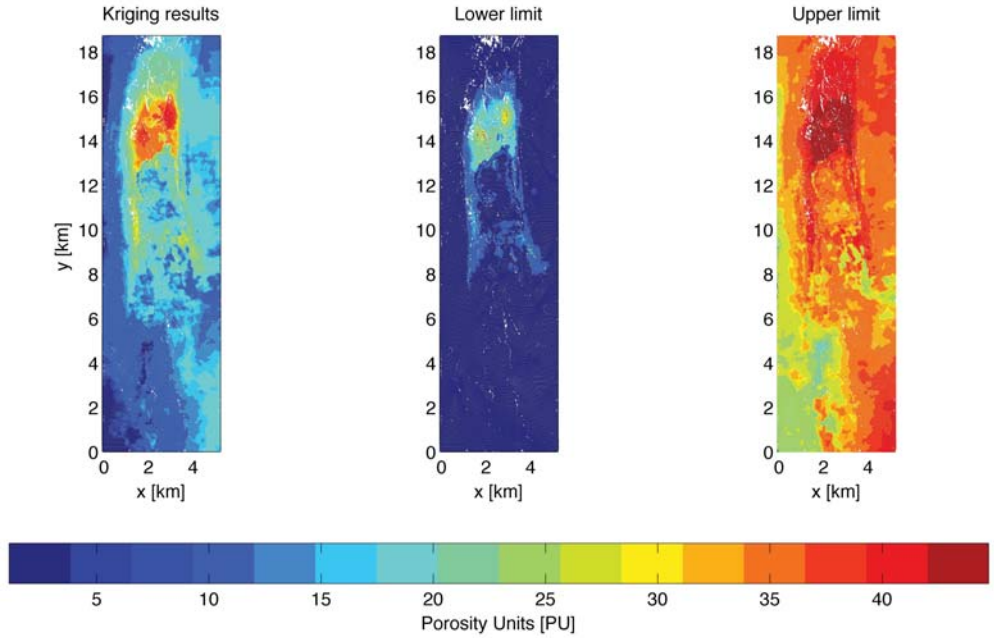

Figure 7: Estimated porosity levels as well as lower and upper limit of the 95% confidence interval for the same test case as shown in the right most plot of Figure 6. Notice the color bar has been altered slightly.

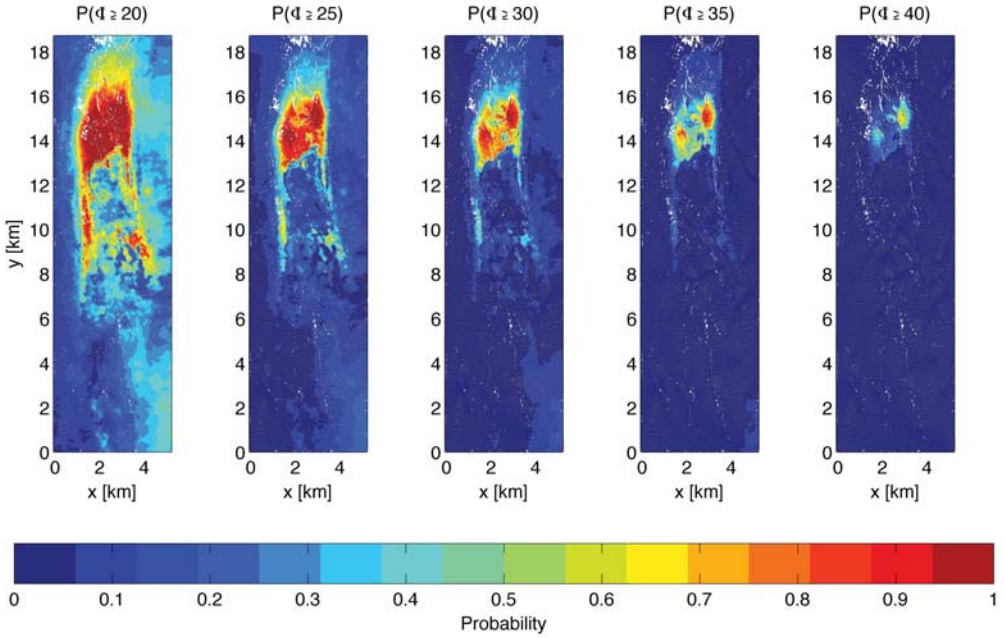

Figure 7: Probability of the estimated porosity levels being higher than certain values. Again we have chosen the test case with four attributes and the complete set of principal components.

2.7. Running times

To give an idea of the performance of the algorithm when it comes to running times we have timed the algorithm solving the test case of data set #1, all attributes and all principal components. Solving the test case we have run MATLAB® R2010a on a MacBook Pro with a 2.66 GHz Intel Core 2 Duo processor and 4 GB of RAM.

The time the algorithm spends has been classified on the following three main tasks:

1)	Initialization, scaling and transformation	0.160 seconds
2)	Computation of optimal covariance parameters	1.80 seconds
3)	Kriging in more than 76,000 grid points	47.0 seconds

It should be noted that the optimization uses a well-qualified initial guess, which of course lowers the number of iterations needed. The kriging itself is in any case the most time consuming tasks, but considering the number of grid points 47 seconds definitely is acceptable.

It should also be noted, that the running time of the kriging task of course also depends on the size of the kriging system. The dimensions of the kriging system translate to the number of known data points. Data set #1 has approximately 100 known data points.

3. ACKNOWLEDGEMENTS

The present work was sponsored by the Danish Council for Independent Research - Technology and Production Sciences (FTP grant no. 274-09-0332) and DONG Energy.

4. REFERENCES

[1] Goovaerts, P. (1997). Geostatistics for Natural Resources Evaluation: New York, Oxford University Press.

[2] Hansen, T. M., Mosegaard, K., Pedersen-Tatalovic, R., Uldall, A., Jacobsen, N. L. (2008). Attribute-guided well-log interpolation applied to low-frequency impedance estimation. Geophysics , 73 (6), 83-95.

[3] Hansen, T. M., Mosegaard, K., Schiøtt, C. (2010). Kriging interpolation in seismic attribute space applied to the South Arne Field, North Sea. Personal communication, accepted for publication in Geophysics.

[4] Pardo-Igúzquiza, E. (1998). Maximum Likelihood Estimation of Spatial Covariance Parameters. Mathematical Geology , 30 (1), 95-108.

[5] Pedersen-Tatalovic, R., Uldall, A., Jacobsen, N. L., Hansen, T. M., Mosegaard, K. (2008, May). Event-based low-frequency impedance modeling using well logs and seismic attributes. The Leading Edge , 592-603.

[6] Shlens, J. (2009, April 22). A Tutorial on Principal Component Analysis. Retrieved Marts 2010, from http://www.snl.salk.edu/~shlens/pca.pdf