Estimation with Non Stationary First and Second Moments

David F. Machuca-Mory and Clayton V. Deutsch

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

The calculation and modelling of local variograms was developed in the previous paper. These local variograms and correlograms, as well as local mean and variance values are built by weighting the sample pairs used in variogram calculation. They prove to better reflect the local spatial behaviour of the variable under study and the results are represented as 2-D or 3-D maps of the model parameters values; there is no longer a stationary set of variogram model parameters. Once the locally varying first and second moments are available at every location, the next step is to use them in estimation. A non stationary Kriging approach and program are developed below. The variogram or correlogram are assumed stationary for all locations within the search of a particular point or block.

Theoretical Background

The well known Ordinary Kriging (OK) system of equations is written as:

$$\begin{cases} \sum_{\beta=1}^{n} \lambda_{\beta}^{(OK)}(\mathrm{u}) C(\mathrm{u}_{\beta} - \mathrm{u}_{\alpha}) + \mu(\mathrm{u}) = C(\mathrm{u} - \mathrm{u}_{\alpha}), \quad \alpha = 1, \ldots, n \\ \sum_{\beta=1}^{n} \lambda_{\beta}^{(OK)}(\mathrm{u}) = 1 \end{cases} \tag{1}$$

Where the $\lambda_{\beta}^{(OK)}(\mathrm{u})$'s are the OK weights and $\mu(\mathrm{u})$ is the Lagrange parameter associated with the constraint $\sum_{\beta=1}^{n} \lambda_{\beta}^{(OK)}(\mathrm{u}) = 1$ (Deutsch and Journel, 1999).

Under a stationary framework a single covariance function $C(\mathrm{u}_{\beta} - \mathrm{u}_{\alpha})$ is valid for all the pairs u_{β} and u_{α}, and u and u_{α} regardless the location of the estimated point. But if fully non stationary is accepted the covariance function can be different for each pair of sample locations, and according to the location of the estimated point were the search ellipsoid is centred. However, in this case the OK system of equations can easily become unsolvable.

A middle way approach, which could be called Quasi-stationarity, is therefore considered. This consists in assuming that for a given point or block to be estimated, its local covariance model is valid for every sample-sample or sample – estimated pair. Thus the OK system remains unchanged:

$$\begin{cases} \sum_{\beta=1}^{n} \lambda_{\beta}^{(OK)}(\mathrm{u}) C_{\mathrm{O}}(\mathrm{u}_{\beta} - \mathrm{u}_{\alpha}) + \mu(\mathrm{u}) = C_{\mathrm{O}}(\mathrm{u}_{\mathrm{O}} - \mathrm{u}_{\alpha}), \quad \alpha = 1, \ldots, n \\ \sum_{\beta=1}^{n} \lambda_{\beta}^{(OK)}(\mathrm{u}) = 1 \end{cases} \tag{2}$$

With the exception that the covariance function $C_{\mathrm{O}}(\mathrm{u}_{\beta} - \mathrm{u}_{\alpha})$ changes according the location u_{O} of the point to be estimated.

This Quasi stationary approach requires counting with a definition of the covariance model at the same resolution of the estimation. This can be achieved by interpolating the variogram parameters obtained from fitting the local weighted variograms or correlograms at each control point.

Program Implementation

The new FORTRAN Kriging program, called `KT3D_lp`, is very similar to its predecessor, with the difference that, if the local variogram parameters are provided in grid files of the same size as the estimation grid, it is able to use the non stationary variograms for local estimation. If these files are not provided, `KT3D_lp` will perform as the old `KT3D` program.

The parameter file of this new kriging program is also similar to the old one, with the difference that extra lines for the variogram parameter file names has been added (see figure 1).

```
                  Parameters for KT3D_lp
                  ******************

START OF PARAMETERS:
../data/cluster.dat              -file with data
0  1  2  0  3  0                 -   columns for DH,X,Y,Z,var,sec var
-1.0e21   1.0e21                 -   trimming limits
0                                -option: 0=grid, 1=cross, 2=jackknife
xvk.dat                          -file with jackknife data
1   2   0    3    0              -   columns for X,Y,Z,vr and sec var
3                                -debugging level: 0,1,2,3
kt3d.dbg                         -file for debugging output
kt3d.out                         -file for kriged output
50   0.5    1.0                  -nx,xmn,xsiz
50   0.5    1.0                  -ny,ymn,ysiz
1    0.5    1.0                  -nz,zmn,zsiz
1    1      1                    -x,y and z block discretization
4    8                           -min, max data for kriging
0                                -max per octant (0-> not used)
20.0  20.0  20.0                 -maximum search radii
 0.0   0.0   0.0                 -angles for search ellipsoid
0     2.302                      -0=SK,1=OK,2=non-st SK,3=exdrift
0 0 0 0 0 0 0 0 0                -drift: x,y,z,xx,yy,zz,xy,xz,zy
0                                -0, variable; 1, estimate trend
extdrift.dat                     -gridded file with drift/mean
4                                -  column number in gridded file
1    0.2                         -nst, nugget effect
1    0.8  0.0   0.0   0.0        -it,cc,ang1,ang2,ang3
         10.0  10.0  10.0        -a_hmax, a_hmin, a_vert
kt3d_c0.out                      -local nugget effect (same grid)
kt3d_c1.out                      -local sill of structure one
kt3d_ah1.out                     -local range hmax
kt3d_ah2.out                     -local range hmin
kt3d_ah3.out                     -local range vertical
kt3d_ang1.out                    -local angle 1
kt3d_ang2.out                    -local angle 2
kt3d_ang3.out                    -local angle 3
```

Figure 1:

At each location `KT3D_lp` reads the corresponding parameters for the local variogram model, if a parameter file is absent, its value from the global variogram is taken instead.

`KT3D_lp` performs quasi stationary Ordinary Kriging and Simple Kriging, in the latter case, the grid file containing the local mean must also be provided. At its present version `KT3D_lp` accepts only local variogram models with a single structure, latter versions will allow to use more than one structure in order to better model zonal anisotropies and other variogram features.

Synthetic Data Example

The synthetic data used for this example has been described in the paper "Non Stationary Variograms based on Continuously Varying Weights". The training image containing two zones of different anisotropy orientation was sampled at 2 x 2, 4 x 4 and 4 x 4 pixels quasi random grids (see figure 2). For each one of these grids the local weighted variograms were calculated using a 10 x 10 pixel control points grid. These variograms were fitted automatically and the resultant parameters interpolated in a 1 x 1 pixel grid using Ordinary Kriging with a Gaussian variogram model and a nugget effect of 10%.The gridded variogram parameters corresponding to each sample grid are presented in figures 3 to 5. It can be noticed in these figures that the local major angle of anisotropy and most of the other parameters are correctly depicted with reasonably accuracy by the local weighted variograms, even for the coarsest grid size.

Figure 6 presents the resultant traditional and quasi non-stationary OK maps for different sampling grid sizes. Traditional OK was performed using a single variogram for both zones, assuming that the boundary between them and the true variogram for each zone are unknown. If the sampling grid is dense enough, the difference between both types of kriging is not appreciable, since data availability outweighs the information provided by the local variogram, but for coarser sampling grids the maps show a clearer difference. However, if the sampling schema is very sparse, the parameters of local weighted variograms become noisier, which can hinder their advantage over the traditional variograms in estimation.

The scatterplots of true versus estimated values (figures 7 to 9) show a higher correlation for the quasi non-sationary OK over the traditional OK, particularly for the 4x4 pixels sampling grid. It is also worthy to notice that, in these graphs the OK performed with locally varying variograms present less bias, as well as a higher variance of estimates.

Real Data Example

The two data sets used for this example are the clustered sampling and the semi regular sampling at a 10m x 10m grid, both taken from the Walker Lake data set (see figure 10). The local varying variogram parameters for both were defined in the previous paper above mentioned. The interpolated correlogram model parameters are shown in figures 11 and 13 for the clustered and gridded sample sets, respectively. For the clustered data set the local correlogram model parameters look more variable and difficult to correlate with the sample values. In the case of the semi regular grid (figure 13), the local correlogram model parameters show features that can be traced in the data, such as the predominantly N-S orientation of grade continuity at the west, the E-W orientation of grade continuity at the east and the long continuity range for the high grade zone.

When used for ordinary kriging, the non stationary correlograms of the clustered data do not improve considerably the estimates, the maps of the traditional and quasi non-stationary OK look similar, overall (see figure 12). And the correlation between estimated vs. true values is indeed lower for the latter (figure 15). In the case of the semi regular sampling grid, the correlation estimated vs. true values is practically similar for both types of OK(see figure 16), however the map corresponding to the quasi non-stationary OK estimates (figure 14, right) show a better local definition concordant with the local orientations of anisotropy and a increased connectivity of very high and very low values.

Discussion and Conclusions

Estimation with non stationary first and second moments calculates improved estimates. The improvement is most noticeable when there are sufficient data to calculate locally varying parameters, but not so much data that the estimates are insensitive to the local variations. Estimation in presence of very sparse data is always problematic. Estimation in presence of very dense data is robust. Locally, the estimates obtained using the non stationary moments reproduce better the local features of spatial continuity of the data. Additionally, they present higher, more realistic, connectivity. When data are clustered, the local weighted variograms may be overly influenced by the spatial features of the over sampled areas, particularly at short lag distances. In these cases the local variogram parameters, the angle of the major anisotropy axis specially, can be modified by the user on the basis of the geological knowledge of the area of study. Information provided by the structural geology, sedimentology and ore genesis of the deposit can serve for such purpose.

References

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

Isaaks, E. H. and Srivastava R. M., 1989. *An introduction to Applied Geostatistics*. Oxford University Press, New York.

Ren, W. and Deutsch, C.V., 2006, *Bayesian Updating with local Varying Correlation*. Paper 104, Eight Annual Report of the Centre for Computational Geostatistics, September 2006.

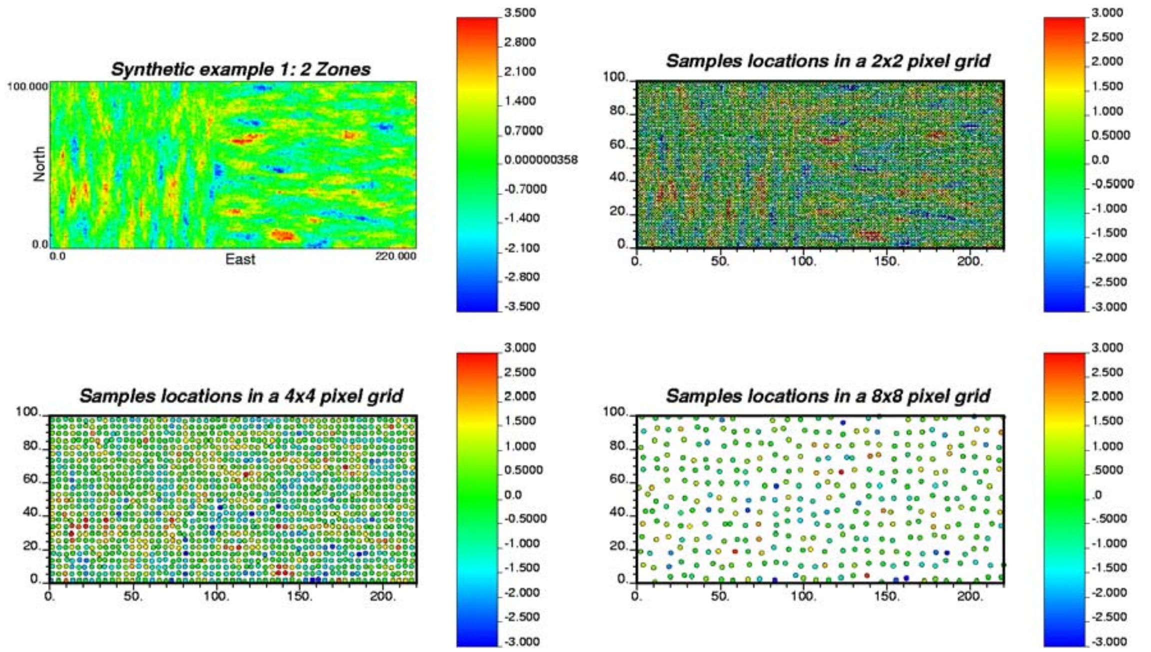

Figure 2: Simulated imaged with two zones of different anisotropy and its corresponding sampling in different grid sizes.

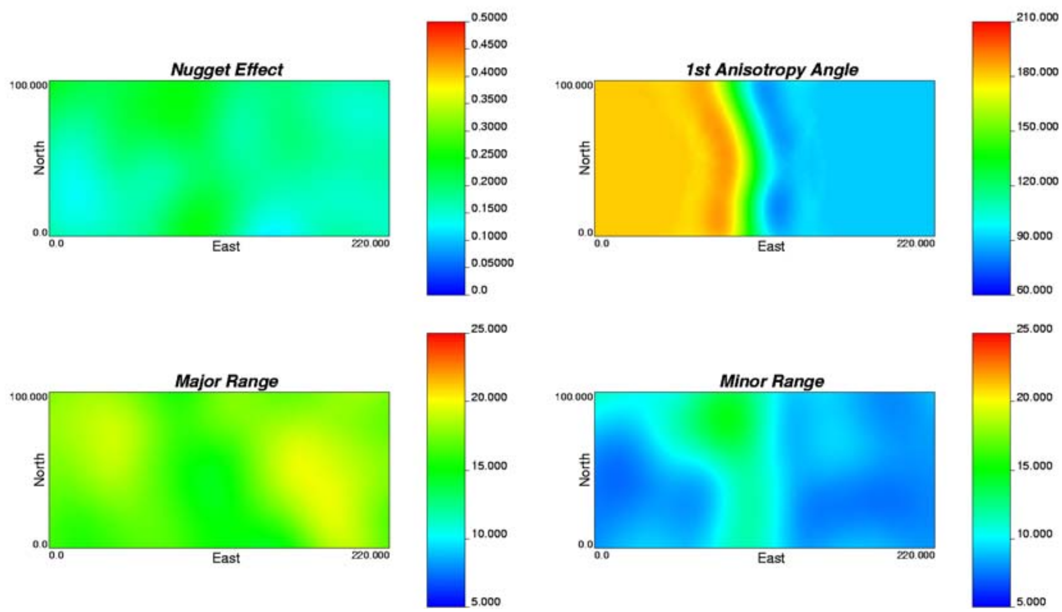

Figure 3: Gridded variogram parameters for 2x2 pixel grid

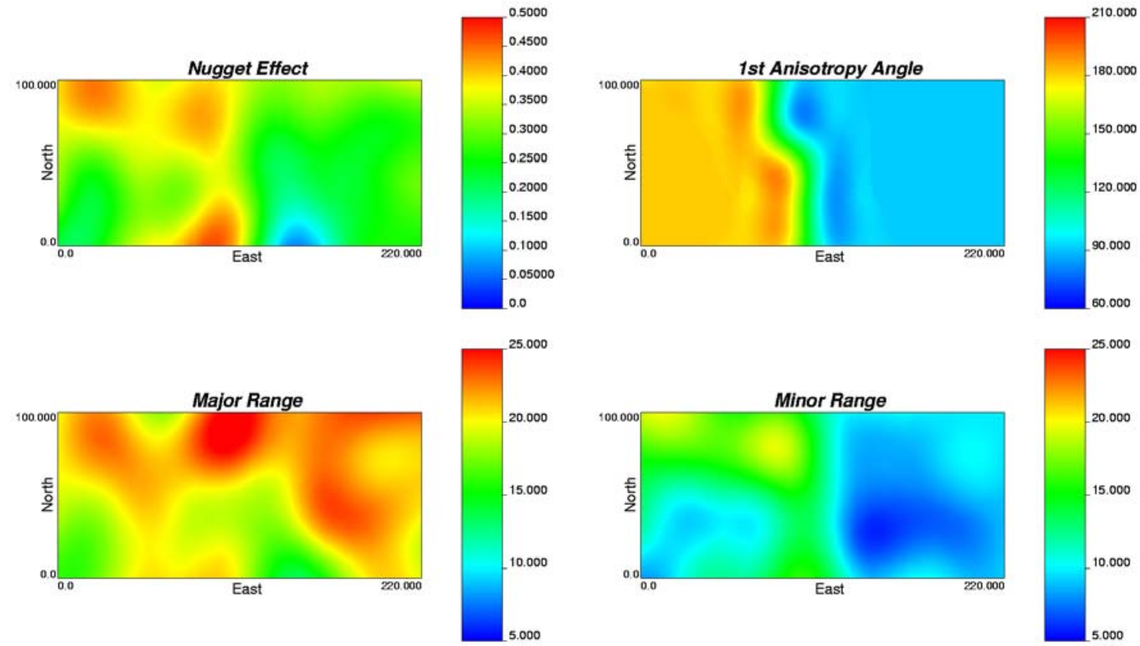

Figure 4: Gridded variogram parameters for 4x4 pixel grid

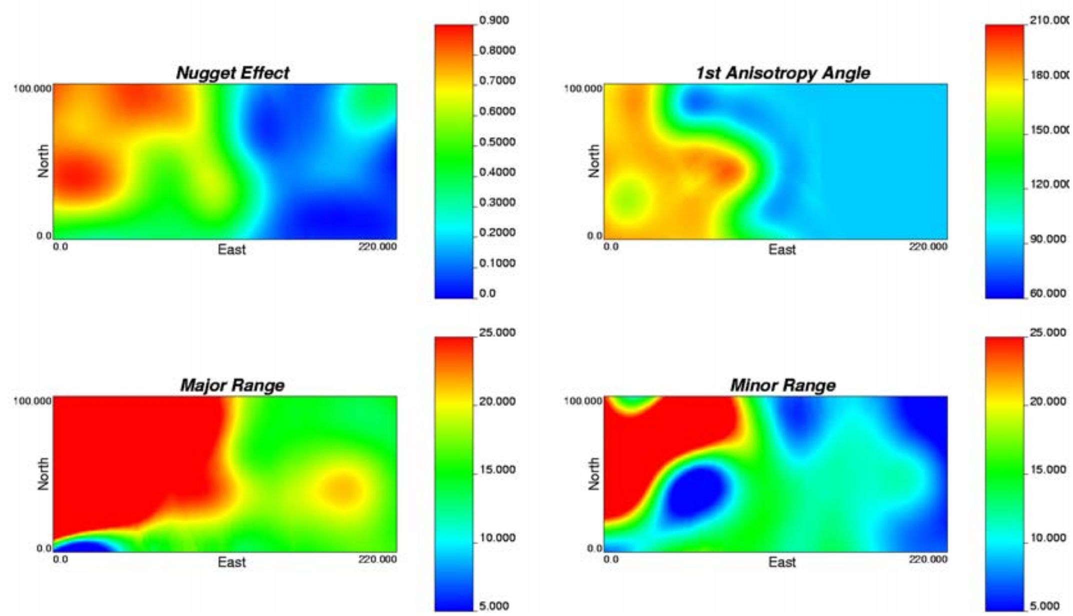

Figure 5: Gridded variogram parameters for 8 x 8 grid

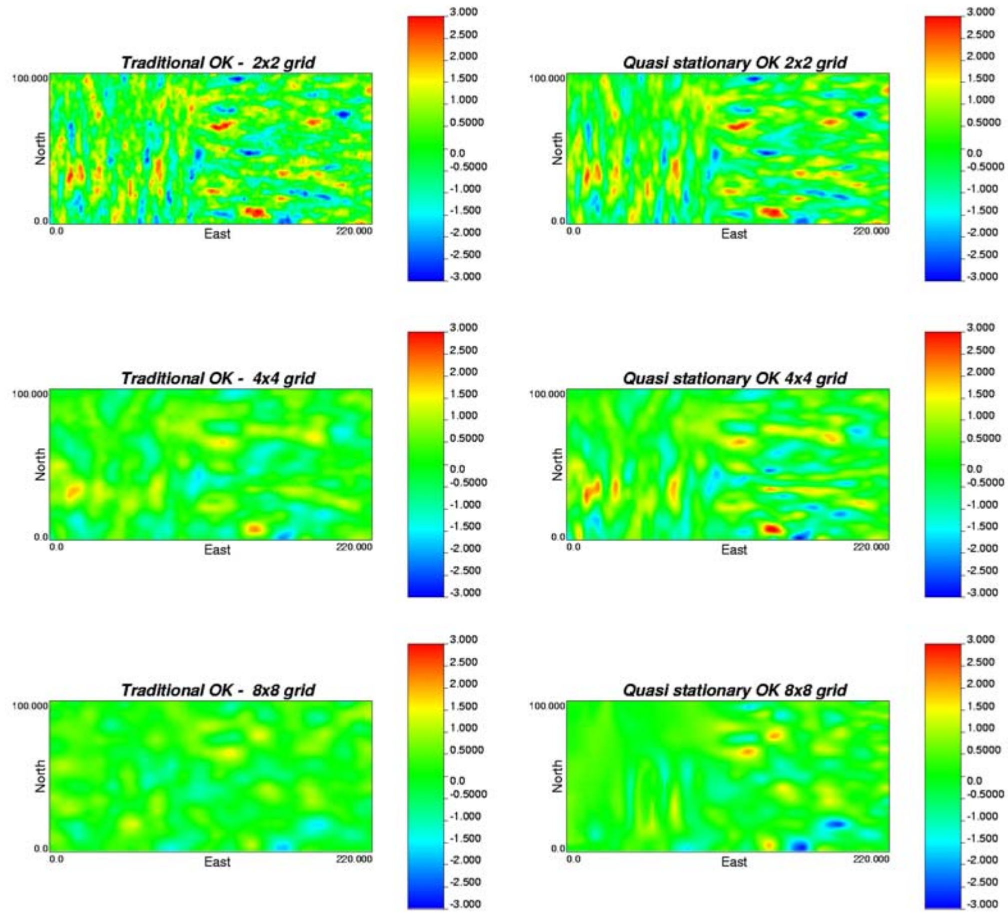

Figure 6: Estimation maps for traditional OK (Left) and Quasi non-stationary OK with data sampled at different grid sizes.

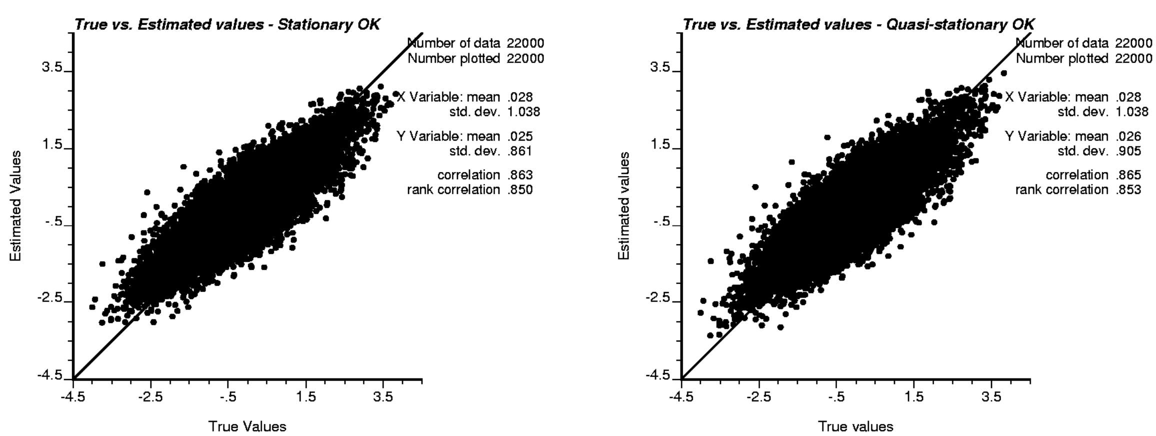

Figure 7: Scatterplots of true vs. estimated values for the 2x2 pixel grid. Left: Traditional OK, right: Quasi stationary OK.

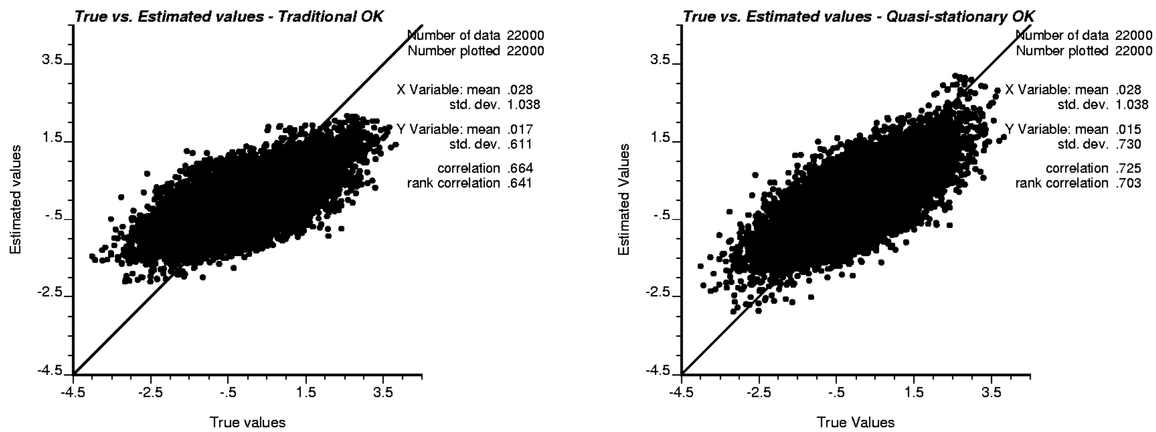

Figure 8: Scatterplots of true vs. estimated values for the 4x4 pixel grid. Left: Traditional OK, right: Quasi stationary OK.

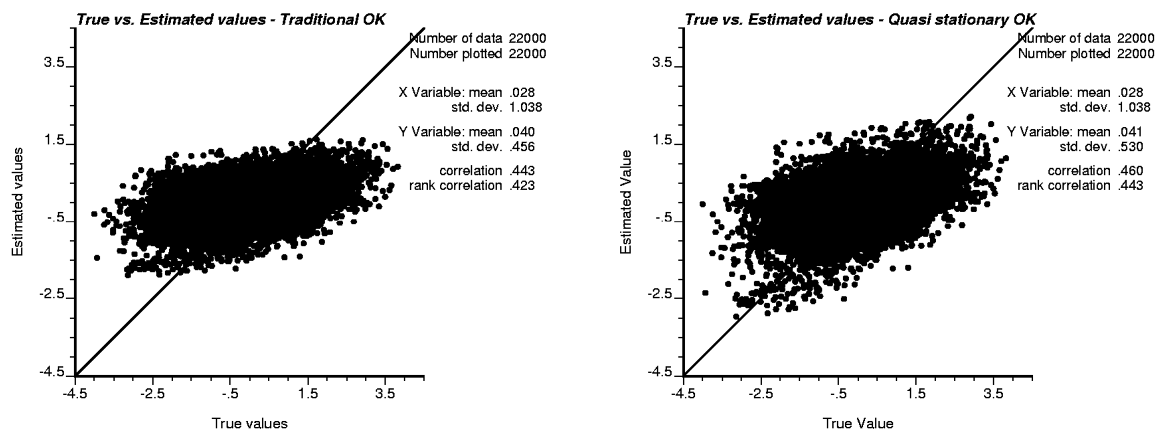

Figure 9: Scatterplots of true vs. estimated values for the 8x8 pixel grid. Left: Traditional OK, right: Quasi stationary OK.

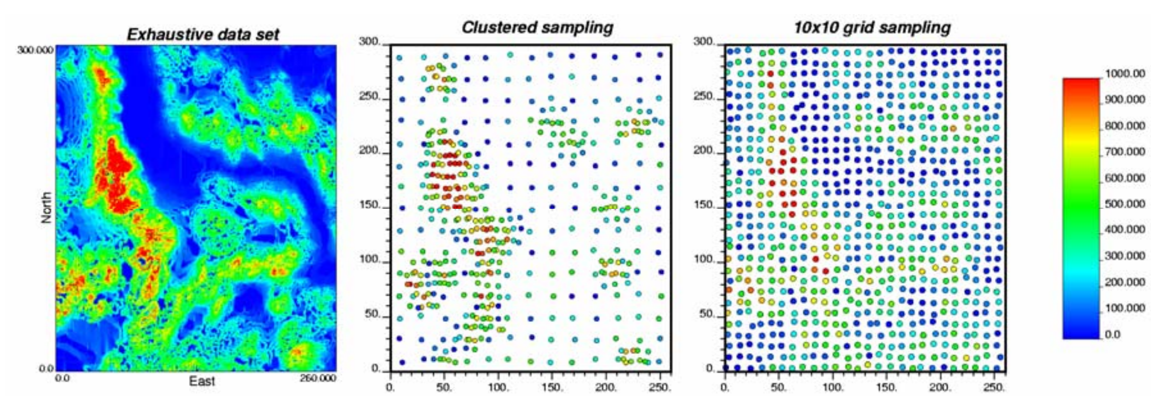

Figure 10: Walker Lake data sets.

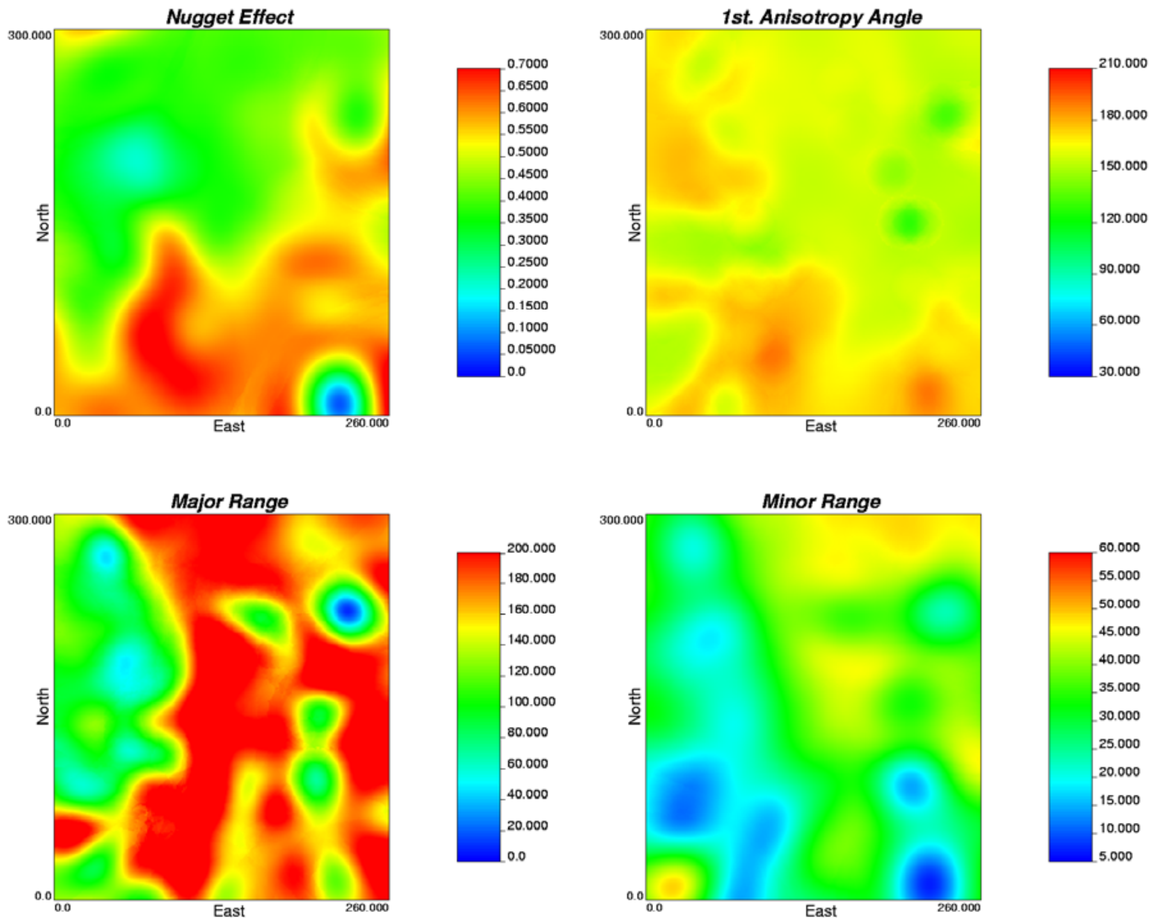

Figure 11: Local Correlogram model parameters for the clustered data set.

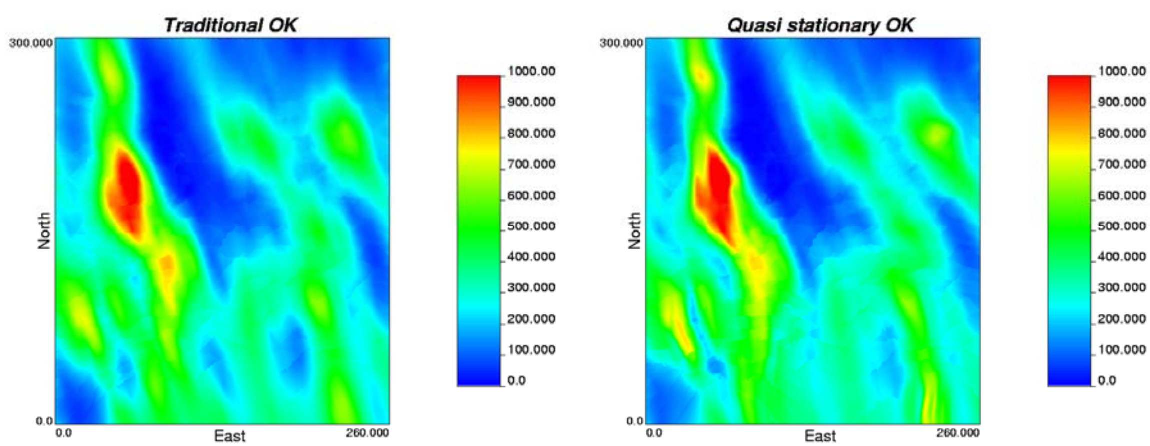

Figure 12: Kriging maps for the clustered data set. Left: Traditional OK, right: quasi stationary OK.

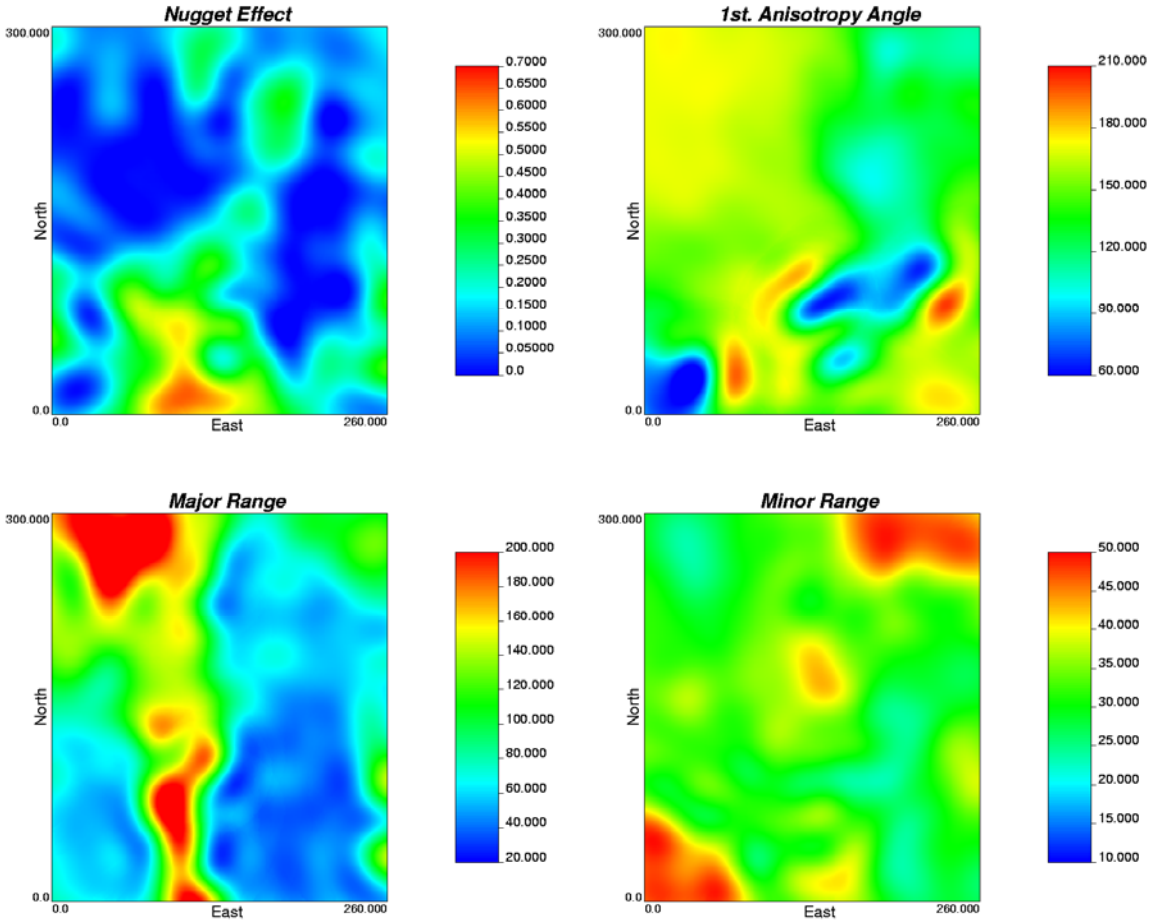

Figure 13: Local Correlogram model parameters for the 10m x 10m sampling grid.

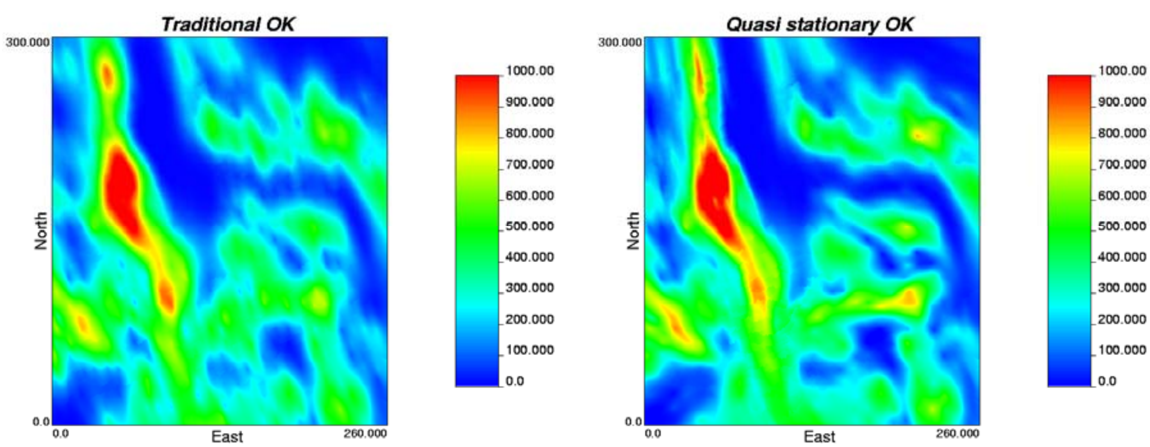

Figure 14: Kriging maps for the 10m x 10m sampling grid. Left: Traditional OK, right: quasi non-stationary OK.

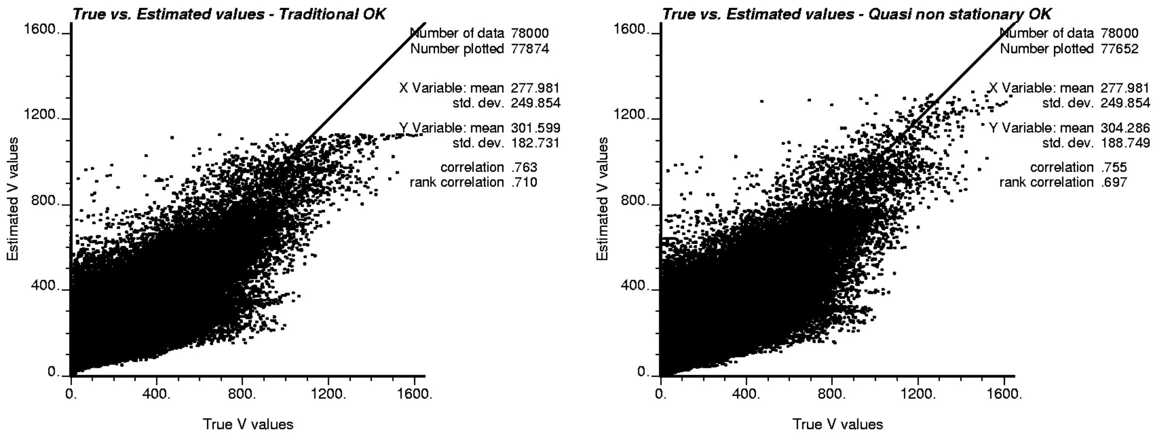

Figure 15: True vs. estimated values scatterplots for the clustered data set. Left: Traditional OK, right: Quasi stationary OK.

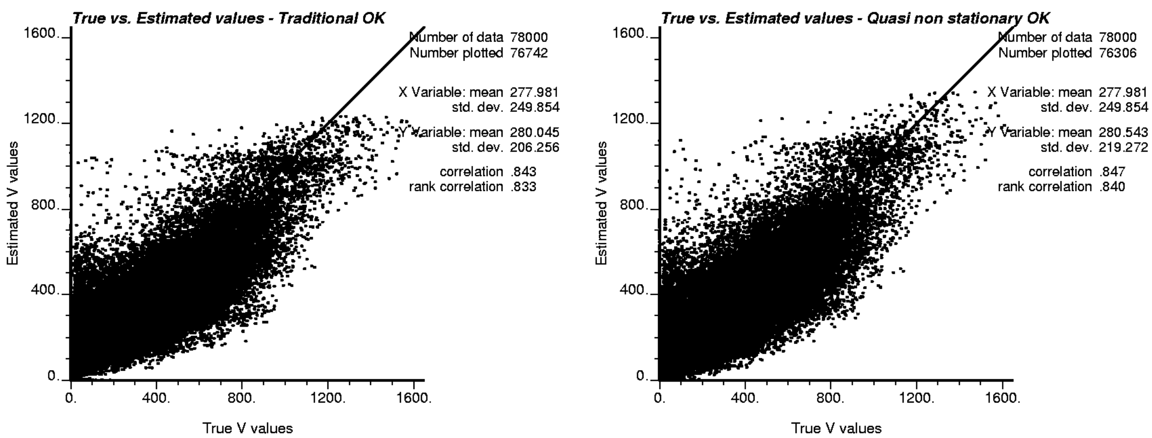

Figure 16: True vs. estimated values scatterplots for the 10m x 10m sampling grid. Left: Traditional OK, right: Quasi stationary OK.