Geostatistics in ILWIS

J. Hendrikse
ILWIS Development
IT Department
ITC

e-mail:
hendrikse@itc.nl

KEYWORDS: geostatistics, variogram, G.I.S, remote sensing, vector, raster, georeference, interpolation, prediction, estimation, regionalized variable, kriging, anisotropy, trend, co-kriging, indicator, error variance, semi-variogram model, cross-variogram, goodness-of-fit, co-regionalization, spatial statistics.

1. ABSTRACT

The newest version of the ITC GIS and Remote Sensing software **ILWIS** offers an integrated set of modules to perform geostatistics.

Its functionality for image processing, geographic analysis and digital mapping combined with extensive geostatistical tools make it a modern product for education and research in the field of G.I.S.

It incorporates spatial correlation tools like semi-and cross variogram modelling.

Anisotropy and spatial trend in the data can be studied. Unlimited numbers and types of attributes can be linked to both vector and raster maps. Many options of statistical interpolation (Kriging) and error estimation are implemented and described in this article. Especially, the flexible integration of different data formats is emphasized.

2. INTRODUCTION

Geographic data have a geometry that is often structured either in regular grid patterns ('raster') or as randomly distributed sample points, lines or polygons ('vector'). Statistical analysis on 'raster' data can involve image processing techniques that make frequently use of concepts like histogram equalization, clustering and filtering methods, multi-band image classification, auto- and cross-correlation.

In the '**vector**' case the incoming data can have a rather irregular geometric position. The geometric information contains point coordinates in some local or global (e.g. latlon) system.

The semantic information about the points, lines or polygons can be stored in any (user-defined) range of values, This is called "**domain**" in ILWIS. (Koolhoven and Wind 1996)

In case of statistical models the domain must have a numerical content. This means that the vectors should have at least one attribute with numerical values as domain.

The statistical study of data available in this point vector format (in particular as a result of measurement samples of any kind) is called **geostatistics**.

It makes use of the study of spatial continuity of so-called regionalized variables.(Burrough and McDonnell, 1998)

Geostatistics assumes that one can set up a **stochastic model** to predict measurement outcomes at unvisited locations based on measurements at visited ones.

The stochastic modelling implies for instance the test on normality or log-normality, followed by a search for **covariance or variogram models**. The combination of point maps with associated tabular data is therefore essential.

International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B4. Amsterdam 2000.

365

To each measurement location one can link various properties by means of an attribute table. These tables accept quasi-unlimited amounts of records and fields and allow many table- and join-operations.

Moreover ILWIS offers **variogram surfaces** and curve fitting techniques to investigate the stochastic properties of the data.

After finding the model that matches the observations best, the prediction of values at unvisited locations is done by gridding. Gridding is an interpolation technique that transforms a point map into a raster map covering the same geographic area.

The **geostatistical interpolation** follows one of the methods derived from the work of Krige or his successors (Krige 1951, Matheron 1965).

The '**Kriging**' method is statistical in the sense that it supposes knowledge of the correlation between point measurements at different locations and that it produces an estimate of the variance (and standard deviation) in the error of the prediction.

This error variance is minimal compared to error variances of other statistical interpolation methods. The estimation is unbiased.

IN BRIEF , this article wants to point out three aspects of the geostatistics in ILWIS:
Modelling: spatial correlation and variogram analysis on point map data,
Kriging prediction: interpolation by gridding using the models found and statistical interpretation using so called error maps.
Integration of geostatistics with other G.I.S. tools to allow teachers and researchers to place their study case in a realistic context and a correct georeference.

3. SPATIAL CORRELATION AND EMPIRICAL SEMIVARIOGRAM

Geostatistics is the science of modelling and estimating spatial variability. So it studies data that fluctuate in space (and time) and their predictability at unvisited locations.

A useful characteristic that expresses spatial correlation, or spatial continuity, is the semivariogram. Assume $Z(x_j)$ is a random function of the location vector xj In case of intrinsic stationarity the semivariogram is defined using the variance of $Z(\mathbf{x_j}) - Z(\mathbf{x_j}+\mathbf{h})$: (Deutsch and Journel 1992)

$$\gamma(\mathbf{h}) = E (Z(\mathbf{x_j}) - Z(\mathbf{x_j}+\mathbf{h}))^2 / 2 \quad \text{which is a function of the separation vector } \mathbf{h} \text{ only.}$$

The **empirical** semivariogram being an estimator of $\gamma(\mathbf{h})$ is calculated from the separation vectors between all point pairs, subtracting attribute values $z(\mathbf{x_j})$, $z(\mathbf{x_k})$ measured in pairs of distinct points.
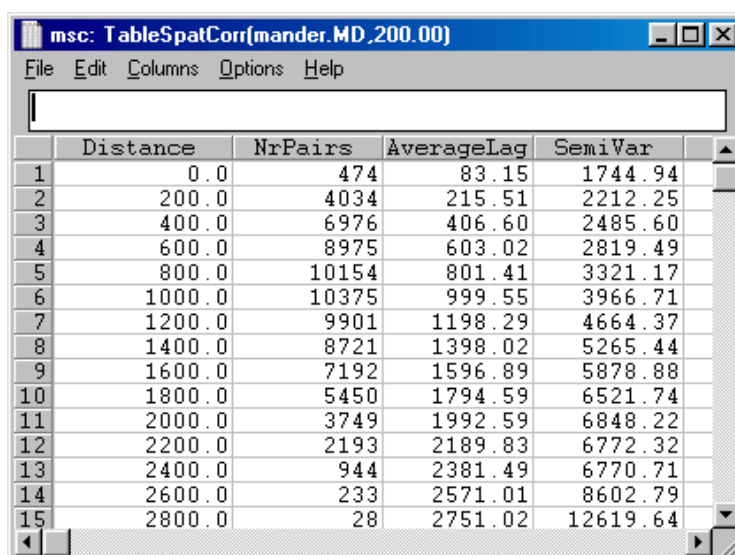
It is defined using all available pairs of point vectors $\mathbf{x_j}$ and $\mathbf{x_k}$ in the map, by:

$$\gamma(h_i) = \Sigma\Sigma [Z(\mathbf{x_j}) - Z(\mathbf{x_k})]^2 / 2N(h_i)$$

The distance h_i ranges over a discrete and finite set of mean lag-length values.

To each h_i value corresponds a set of possible separation vectors $\mathbf{h}$, that have a length belonging to the class of h_i (in fact h_i equals the mean of all lengths in that set). The class of h_i contains $N(h_i)$ point pairs.

For a given i, the summation $\Sigma\Sigma$ is taken over all $\mathbf{x_j} - \mathbf{x_k}$ with $(i - 1).s < \|\mathbf{x_j} - \mathbf{x_k}\| < i.s$, where i = 1 n, with n the number of distance classes (" lags") and s is the so-called lag spacing. (Ilwis-Help1999) The values of $\gamma(h_i)$ are obtained in ILWIS from a point map as input. The point map must have an attached attribute table that contains columns (attributes) with numerical values. The computation is implemented in the operation called " Spatial Correlation" .

Figure 1. Tabular output of point pair statistics

This operation produces a table with several columns as output. (Figure 2) Each pair of value columns can be plotted in an xy graph so that semi-variograms can be displayed and studied in order to find the best choice for the parameters of a variogram model needed for the Kriging interpolation.

Semivariogram values restricted to certain directions with respect to the map North can be computed. In this way so-called **bi-directional** variograms are produced. These can in turn serve as input to the Kriging operation with geometric **anisotropy**.

If anisotropy is expected, it can be visualized quickly by making a **variogram surface**. This is a square raster " map" that shows clearly the direction and strength of possible anisotropy, so that the azimuth and range ratio can be used to choose proper parameters for the bi-directional case.
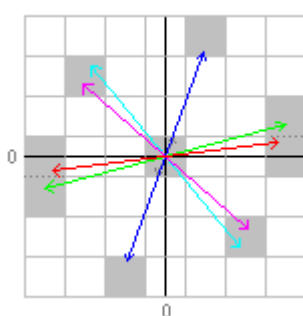
Figure 2. The grey cells show 5 distinct groups of point-pairs found in the point map

4. MODELLING THE SEMIVARIOGRAM

The theoretical semivariogram $\gamma(\mathbf{h})$ is estimated by the described formula, but remains unknown. To account for spatial continuity in the stochastic sense, and use it for

prediction, we need to adopt a **variogram model** that fits best the outcomes of the experimental semivariogram. That model needs to be formulated as a continuous function defined for all distances occurring between the point where the interpolation takes place. ILWIS offers 7 types of functions: Spherical, Exponential, Gaussian, Wave, Circular and Rational Quadratic. Each of these models depends on the choice of 3 parameters: Nugget, Sill and Range. Moreover there is a Power model (including the linear case) which requires Nugget, Slope and Power-exponent. Within certain range limits these parameters can be set by the user in order to match as close as possible the experimental outcomes. (Ilwis-Help, 1999)

Experimental and model semi-variograms can be placed in one graph with common x and y axis. For a judgement of the Goodness of Fit one should compute the model function values for the distances (lags) that occur really in the output of the experimental semi-variogram table. In ILWIS one can use for this the ColumnSemivatiogram operation.
A criterion of good fit could be to minimize as much as one can the following expression of relative squares of differences.

$$R^2 = 1 - \frac{\sum \left(\hat{\gamma}(h_i) - \gamma(h_i) \right)^2}{\sum \left(\hat{\gamma}(h_i) - \sum \hat{\gamma}(h_i)/N \right)^2}$$

$\gamma(h_i)$ is the experimental result for distance class h_i

$\gamma(h_i)$ is the model semi-variogram value for the same class h_i

The ideal value of R^2 is 1 (a perfect fit):

One can of course also simply evaluate and minimize the sum of squared differences

$$\sum \left(\hat{\gamma}(h_i) - \gamma(h_i) \right)^2$$

possibly modified with weight factors.

It will be 0 in case of perfect fit.

Any other criterion to **optimize the curve fitting** can be formulated and investigated.
This can be repeated many times for varying parameters, either on the command line or in a script as an automated iteration.
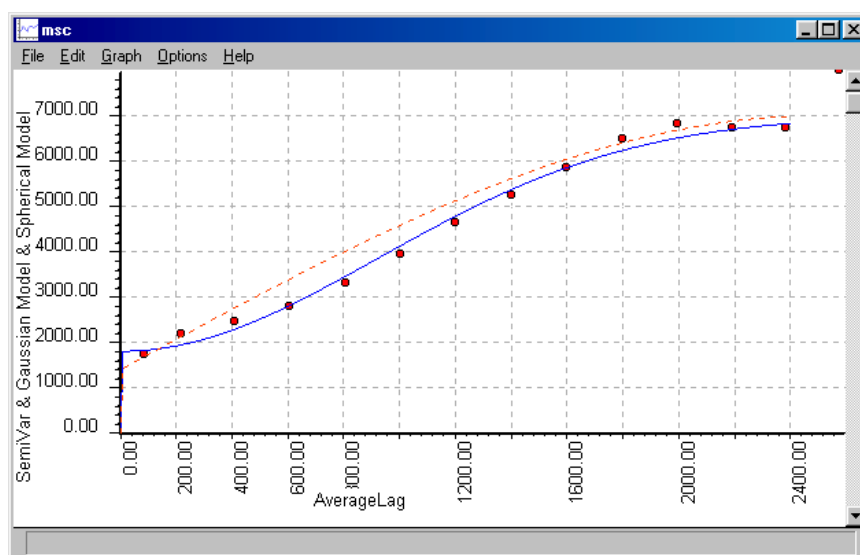
Figure 3. Experimental variogram outcomes (red points) combined with two possible model variograms (smooth curves)

5. POINT MAP INTERPOLATION

ILWIS offers a variety of interpolation possibilities where dimensionless points in a map are used (together with attribute values and an appropriate georeference) such that the program produces a raster map with equidistant pixels (" gridding") that get estimations in a similar value domain. In almost all methods the estimation is a weigthed mean. (except for the methods " Nearest Point" and " Trend Surface").
The methods " Moving Average" and " Moving Surface" are deterministic and assume no model of spatial continuity.
The different Kriging methods, however, need a model function that is based on a good estimation of the covariance or semi-variance in the point data. The function describes how the values in different location stochastically depend on each other.
Five Kriging methods are available:

> **Simple** Kriging
> **Ordinary**, the most widely used method
> **Anisotropic** Kriging (accounting for geometric anisotropy)
> **Universal** Kriging (accounting for local trends)
> **Co-Kriging** (with more than 1 regionalized variable as input)

The ILWIS implementation expects in all cases a semivariogram model with fixed parameters as input. Per interpolation only 1 function is allowed, hence nested models are not (yet) possible.
On the other hand, while varying the search radius in Universal Kriging, local trends in the output (the predictions) can be incorporated
Simple Kriging in ILWIS means that there is no search radius limitation. Perhaps ' global' Kriging would have been a more appropriate term, because many authors refer to a model with known and constant expectation for the regionalized variable if the term ' simple' is used (Cressie 1993).
Anisotropic Kriging must be preceded by proper inspection of variogram surfaces with different pixel sizes. One should be aware of the effect of **scale change**. What appears anisotropic in a certain direction need not to be so at another scale.

The direction of the 2 main axes of anisotropy can be found from the variogram surface, but nugget, sill and ranges should be estimated from the graphs obtained from the bi-directional ' spatcorr' operation

In ILWIS geometric anisotropy is handled by applying an affine transformation on the matrix of distances between the points within the search circle. These transformed distances are fed into the Kriging system of equations both on left and right hand side of the equations. The ' elliptic' distance transformation consists of a rotation of coordinate axes towards the axes of anisotropy , followed by a differential scaling with scale ratio equal to the ratio of the two principal ranges. This ratio is one of the parameters entered by the user, together with the semivariogram model parameters.

All methods ask from the user to decide how to treat almost **co-inciding points**, These points create possibly a singularity in solving the Kriging system This yields undefined estimations.

6. KRIGING ESTIMATION AND ERROR MAP

The Kriging system produces n weights w_1 w_n as solutions if there are n points in the search circle (ellipse). Apart from that, an extra unknown, the ' Lagrange multiplier'λ is computed.

The weights are used to compute the Kriging estimation.

λ is used to compute the error variance of that estimation. (Ilwis Help, 1999)

$$
\begin{pmatrix}
0 & \gamma(h_{12}) & \gamma(h_{13}) & \dots & \gamma(h_{1n}) & 1 \\
\gamma(h_{21}) & 0 & \gamma(h_{23}) & \dots & \gamma(h_{2n}) & 1 \\
\gamma(h_{31}) & \gamma(h_{32}) & 0 & \dots & \gamma(h_{3n}) & 1 \\
\dots & \dots & \dots & \dots & \dots & \dots \\
\gamma(h_{n1}) & \gamma(h_{n2}) & \gamma(h_{n3}) & \dots & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0
\end{pmatrix}
\bullet
\begin{pmatrix}
W_1 \\ W_2 \\ W_3 \\ \dots \\ W_n \\ \lambda
\end{pmatrix}
=
\begin{pmatrix}
\gamma(h_{p1}) \\ \gamma(h_{p2}) \\ \gamma(h_{p3}) \\ \dots \\ \gamma(h_{pn}) \\ 1
\end{pmatrix}
$$

The left-hand matrix depends on the mutual distances between the various input sample points and on the variogram function γ selected for the variable to be interpolated.

The position of the pixel (point) where the estimation takes place, influences only the right hand column in the equation. This column must be recomputed for each pixel that is interpolated. The left matrix needs only to be recomputed if the set of sample points inside the limiting circle is different from the previous one.

One could improve the matrix inversion by making the diagonal dominant with respect to the off-diagonal elements, thus avoiding too much pivoting when solving the system. This is particularly relevant for large matrices caused by dense sample points combined with a large search radius.

It can be implemented by setting up the Kriging system using the covariance function in stead of the experimental variogram outcomes as input for an ordinary Kriging algorithm as described in (Deutsch and Journel, 1992) and (Isaaks and Srivastava, 1989). It will be considered in future releases of ILWIS because of its numerical advantages.

The map of square roots of **error variances** (in ILWIS simply called " error map") is a raster map with geometric properties equal to those of the interpolation map (the map with Kriging estimates).

It has the same number of rows and columns and the same real world coordinates connected to it by means of its georeference.

Therefore all ILWIS map calculations can be performed on combinations of estimates and errors. One may produce in this way probability maps that show confidence intervals around the estimates. (Pebesma, 1996)

The georeference of the output raster maps ensures that they have a coordinate system in common with the input sample points. The ILWIS feature of dependency between source

objects and their resulting ' target' objects makes recalculation after slight modifications in the samples an easy task,
This offers a very handy tool to investigate and optimize sampling patterns and strategies, especially if one automates the interpolation of many different input point samples in a script.
(Van Groenigen, 1999)
It makes ILWIS a flexible and open end **research tool**.
Another example of making use of the integrated functionality of ILWIS is the modelling of variograms and subsequent Kriging with **indicators**. The indicators are data transformed into the Boolean domain based on a given cut-off value. This is simply done by creating an extra boolean column in the attribute table of the point map, using an iff-statement. The new column is used as input for variogram modelling and fitting. The variogram parameters are used in turn, to perform so-called **Indicator Kriging**.
For **Block Kriging** a feasible approach is to make proper use of the filter and aggregate operations in the ILWIS raster domain. It offers many standard filters, many user-defined filters and a variety of aggregation methods (Ilwis Help, 1999)

7. KRIGING WITH LOCAL TRENDS

A non-stationary regionalized variable can be regarded as having two components: the drift consisting of the average value of the variable within the neighborhood and the residual being the difference between the actual measurements and the drift (Van der Meer, 1993).
Universal Kriging consists of three operations:
First the drift within a limiting distance is estimated and removed, secondly; the stationary residuals are kriged to obtain the needed estimates and finally the estimated residuals are combined with the drift to obtain estimates of the actual surface.
The set of estimates can be seen as a smooth surface defined within the point map limits.
In this ' trend' -Kriging approach, one should choose variogram models with care. The range of the model and the search radius of the interpolation are to be fine-tuned with respect to each other.
ILWIS offers as options both linear and quadratic trend polynomials in two variables x and y (the map coordinates).
The Kriging matrix equation for the linear case is shown below. The estimation is done for the location x_p , y_p . In this example, it is computed using 5 input sample points.
The quadratic case has 3 more rows and columns in the left hand matrix. These contain the 2nd degree terms in the sample coordinates x_i and y_i The same goes for the right hand vector and coordinates x_p and y_p of the estimated point.
The column of unknowns has, in the quadratic case, 3 more Lagrange parameters to be solved: α_3 ,α_4 ,α_5

$$
\begin{pmatrix}
0 & \gamma(h_{12}) & \gamma(h_{13}) & \gamma(h_{14}) & \gamma(h_{15}) & 1 & x_1 & y_1 \\
\gamma(h_{21}) & 0 & \gamma(h_{23}) & \gamma(h_{24}) & \gamma(h_{25}) & 1 & x_2 & y_2 \\
\gamma(h_{31}) & \gamma(h_{32}) & 0 & \gamma(h_{34}) & \gamma(h_{35}) & 1 & x_3 & y_3 \\
\gamma(h_{41}) & \gamma(h_{42}) & \gamma(h_{43}) & 0 & \gamma(h_{45}) & 1 & x_4 & y_4 \\
\gamma(h_{51}) & \gamma(h_{52}) & \gamma(h_{53}) & \gamma(h_{54}) & 0 & 1 & x_5 & y_5 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
x_1 & x_2 & x_3 & x_4 & x_5 & 0 & 0 & 0 \\
y_1 & y_2 & y_3 & y_4 & y_5 & 0 & 0 & 0
\end{pmatrix}
\bullet
\begin{pmatrix}
W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \\ \lambda \\ \alpha_1 \\ \alpha_2
\end{pmatrix}
=
\begin{pmatrix}
\gamma(h_{p1}) \\ \gamma(h_{p2}) \\ \gamma(h_{p3}) \\ \gamma(h_{p4}) \\ \gamma(h_{p5}) \\ 1 \\ x_p \\ y_p
\end{pmatrix}
$$

8. CROSS-VARIOGRAM AND CO-KRIGING

Most data sets with samples points will contain more than only one variable. Besides a primary variable of interest, that has to be interpolated, there can be several secondary variables present.

These secondary variables could be more or less correlated with the primary variable and could serve as supplementary information to better predict the primary one.

ILWIS offers the functionality of estimating a primary variable (the '**predictand**') with one secondary variable (the '**co-variable**').

The underlying theory is well explained in (Cressie 1993), (Deutsch and Journel 1992) and (Isaaks and Srivastava 1989)

For the ILWIS **Cokriging** algorithm see (ILWIS Help,1999) for more information.

In this algorithm we assume two conditions of unbiasedness:

One on the m weights ω_i assigned to predictand A, for which $\Sigma\,\omega_i = 1$

And one on the n weights η_i assigned to covariable B, for which $\Sigma\,\eta_i = 0$.

It leads to a system with one more row and column and an extra Lagrange parameter μ_2 for the covariable.

Thus the system has $n + m + 2$ unknowns:

At this moment we have implemented the variogram model values of predictand A, covariable B and the cross-variogram of A and B as input to the CoKriging system. So the sub-matrices G_{**} are made out of semivariogram values whose parameters are selected by the user. This requires a well chosen combination of semi-variogram models γ_A , γ_B and γ_{AB}

The models should fit the empirical outcomes of the Crossvar operation sufficiently close.

Morover they should fulfill the Cauchy-Schwarz condition $\gamma_{AB}{}^2 < \gamma_A\,.\gamma_B$ for all occurring lags (point pair distances) in the sample map.

$$
\begin{bmatrix} G_{AA} & G_{BA} & 1_m & 0 \\ G_{BA} & G_{BB} & 0 & 1_n \\ 1_m' & 0 & 0 & 0 \\ 0 & 1_n' & 0 & 0 \end{bmatrix}
\begin{bmatrix} \omega \\ \eta \\ \mu_1 \\ \mu_2 \end{bmatrix}
=
\begin{bmatrix} \gamma_{pA} \\ \gamma_{pAB} \\ 1 \\ 0 \end{bmatrix}
$$

The estimator is implemented as

$$
\hat{Z} = \sum_i^m \omega_i \bullet A_i + \sum_i^n \eta_i \bullet B_i
$$

And its error variance is implemented as:

$$
Var = \quad \sigma^2 \; = \; \mu_1 \; + \; \Sigma\,\omega_i\,\gamma_A(h_i) \; + \; \Sigma\,\eta_j\,\gamma_{AB}(h_j)
$$

One of the reasons to use a covariable is the fact that the predictand is undersampled, compared to the covariable, for instance if there are large areas in the map where only the more densely sampled variable is measured. In such areas the only information we have about the predictand is its cross-correlation with the covariable. (Stein, 1988, 1991) show examples of simulated and real world data in soil studies.

Another reason to apply Cokriging can be the fact that the covariable measurements have a higher precision or reliabilty.

9. INTEGRATING VECTOR AND RASTER G.I.S. WITH GEOSTATISTICS

ILWIS is a G.I.S package that emphasizes the integration between information in vector format (digitized points, lines, polygons, locaions known from **GPS** readings) and raster format (imagery from remote sensing, aerial photography, scanned maps).

The **georeference** of raster maps is implemented in combination with coordinate system definitions including map projections. The georeference settings define the spatial resolution of interpolation from point to raster maps and keep the link between coordinates in both formats.

Any vector map with coordinates that are convertible to geographic latitude and longitude using a **map projection** and if needed a **geodetic datum**, can be overlaid on a Kriging result without the need to resample the latter again. The coordinate transformation link is established via the coordinate system of the initial point map.

In case of **multivariate interpolation** (" Co-Kriging"), two point maps with different coordinate systems and projections may be combined if they overlap in a common geographical region of interest.

This often happens in real world situations and in research projects.

If the map projection of vector data is unknown, a link can still be defined by entering a minimum number of control points in both **coordinate systems**.

This link ensures a smooth, accurate and very user-friendly display, update, combination and retrieval of geographic information. (see Figure 1)

The information is either stored in the geometrical formats vector or raster, or in tabular form. The tabular storage structure is linked to the various maps by means of common identifiers that make use of the flexible **domain** concept.

In geostatistics the most common domain is of type " value" i.e. numerical values of measurements at point locations are input and interpolated values with the same numerical range are output.

But other domains can be created by the user and used in various contexts.

A boolean domain, for instance, allows for the creation of **indicator** and **probability maps**, if one uses the tabular calculation of critical values and confidence intervals properly.

The filter possibilities of ILWIS in raster maps enable the user to carry out statistical estimations on square sub-areas leading to **block** Kriging.

The ILWIS script language makes the production of **prediction** maps for many **simulated** situations possible. It can help in finding optimal designs of sampling schemes. (Van Groenigen, 1999)

The **dependency** between source point-map or source table and interpolated (gridded) output is an ILWIS feature that facilitates the update of source and target objects (maps, tables) without the need of re-entering all interpolation parameters again.
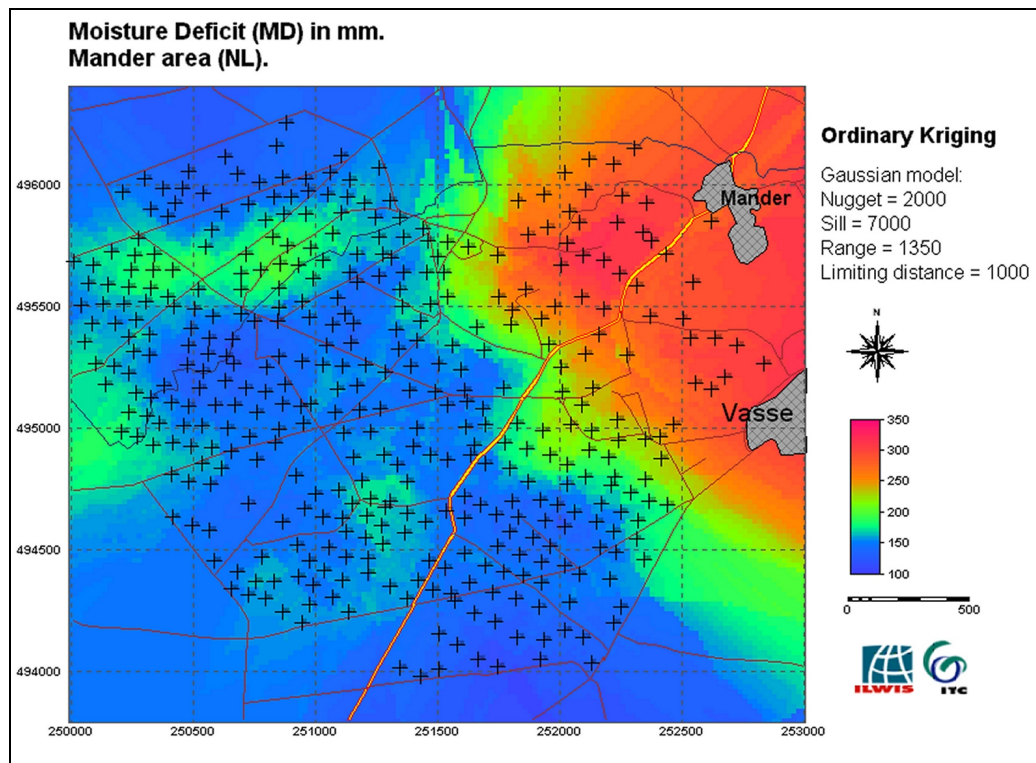
Figure 4. ILWIS Map View showing Kriging results (color) on top of which the point map of input samples and a topographic segment map have been overlaid.

10. CONCLUSIONS

With the release of its youngest version, ILWIS offers a rather complete set of geostatistical functions and tools.
Integrated with its G.I.S and Image Processing capabilities it can serve as a new instrument for research and education, making full use of the smooth linkage between raster, vector and tabular format of data.

11. ACKNOWLEDGEMENTS

12.   REFERENCES

Burrough P.A. and McDonnell. 1997. Principles of G.I.S.; Spatial Information Systems and Geostatistics, Oxford University Press, Oxford

Cressie N.A.C. 1993.   Statistics for Spatial Data, Wiley and Sons, New York

Deutsch and Journel. 1992. Geostatistical Software Library and User' s Guide, Oxford University Press, New York

Ilwis Help. 1999. ILWIS version 2.22 July 1999; On-line Help Web-site: HTTP://www.itc.nl/ilwis

Isaaks and Srivastava. 1989. An introduction to Applied Geostatistics, Oxford University Press, New York

Koolhoven and Wind, 1996. Domains in ILWIS. Proceedings of the 2nd JEC Barcelona, March 96,volume 1 IOS Press, Amsterdam

Krige D.G. 1951. A statistical approach to some basic mine valuation problems.on the Witwatersrand, Journal of the Chemical, Metallurgical and Mining Society of South Africa, 52, 119-139

Matheron G. 1971. The theory of regionalized variables and its applications. Cahiers du Centre de Morphologie Mathematique, no 5, Fontainebleau, France

Pebesma. 1996. Mapping Groundwater Quality in the Netherlands, Dissertation Fac der Ruimtelijke Wetenschappen, University of Utrecht, The Netherlands

Stein A. e.a. 1988. Cokriging Point Data on Moisture Deficit (Soil Science Society of.America.Journal.52)

Stein A. and Corsten L.C.A. 1991. Universal Kriging and Cokriging as a regression Procedure (Biometrics June 91 13 pp)

Stein A. 1998, Spatial Statistics for Soil and the Environment, ITC Lecture Notes, Enschede, The Netherlands

Van der Meer F. 1993. Introduction to Geostatistics, ITC Lecture Notes, Enschede, The Netherlands

Van Groenigen J.W. 1999. Constrained optimisation of Spatial Sampling, Dissertation, University of Wageningen, The Netherlands