Assignment_4

Part 1: Variography with the Swiss Precipitation Dataset

To examine and understand the given dataset it is important to choose the right method by using the geostatistical analyst tool in ArcGIS. With the geostatistical analyst you can "explore data variability and spatial relationships, look for unusual data values, and examine global and local trends." (ESRI, 2012). Different methods like histograms, a voronoi map or a trend analysis can be created to analyze the data or to simulate different scenarios (i.e. an environmental process). So an exploratory spatial data analysis (ESDA) should be done before.

Starting an ESDA with the histogram (Fig. 1) is always a good idea because diverse statistics can be seen here with plenty of the values in the data. "The Histogram tool in ESDA provides a univariate (one-variable) description of your data. The tool displays the frequency distribution for the dataset of interest and calculates summary statistics." (ARCGIS 9.2 DESKTOP HELP, 2010).

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

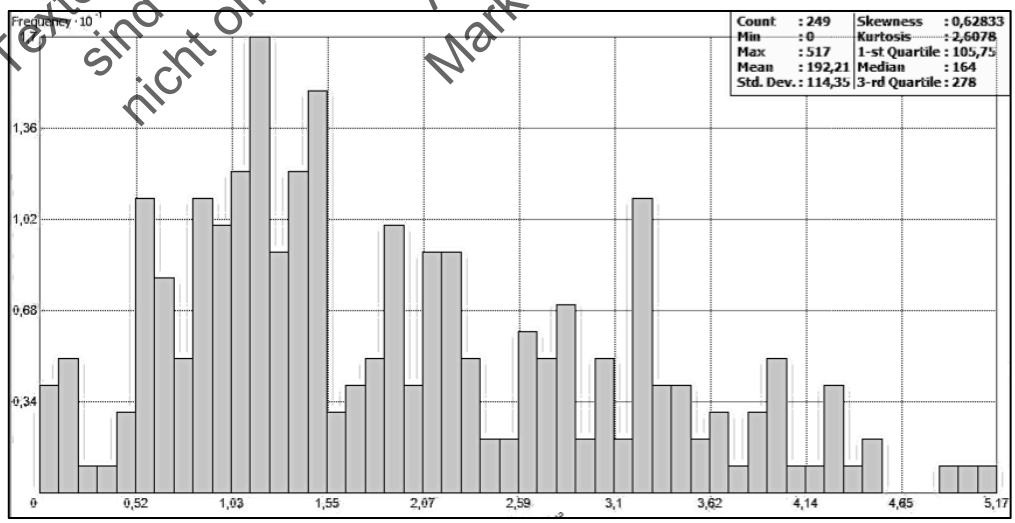


Fig. 1: Histogram of data

A second opportunity to explore the data with the geostatistical analyst wizard is the "Normal-Q-Q-Plot" (Fig. 2). You can use this plot to picture the quantiles from two different allocations, which are plotted against each other. In case the allocations are similar, the points in the QQ-Plot will be shown as a rectilinear line. So a Q-Q-Plot is used to associate the shapes of distributions, offering a graphical interpretation of how properties like position or measure/scale are comparable or distinctive in the two divisions. Q-Q plots can be used to compare groups of data, or hypothetical distractions.

Fig. 2: Normal Q-Q Plot

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

The next tools I already mentioned before is the "Voronoi Map" (Fig. 3). "Voronoi maps are constructed from a series of polygons formed around the location of a sample point." (ARCGIS 9.2 DESKTOP HELP, 2010). Voronoi polygons (very similar to Thiessen Polygons) are formed so that each location inside a polygon is closer to the sample point in that polygon than any other sample point. After the polygons are built, neighbors of a example point are distinct as any other sample point whose polygon shares a border with the chosen sample point. Long story short: Each polygon contains only one measured point.

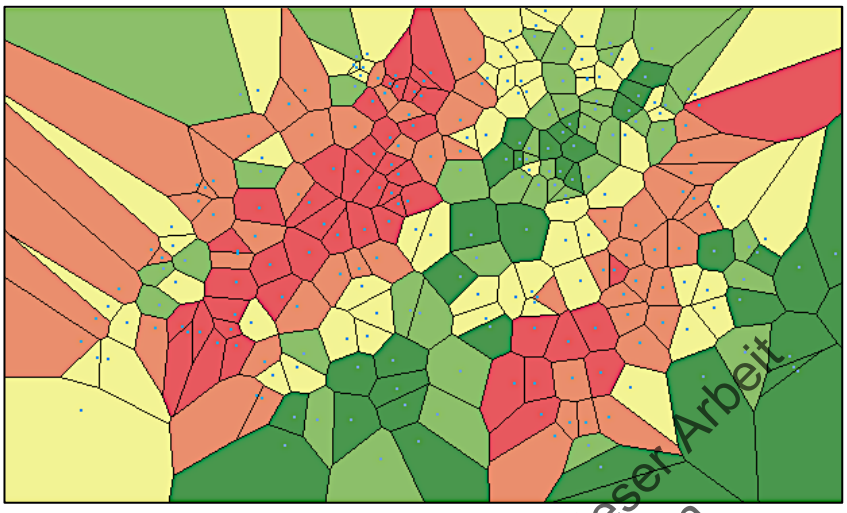


Fig. 3: Voronoi Map

The last tool to have a short look into data could be the "Trend Analysis" (Fig. 4). This ESDA tool is part of the polynomial interpolation, which distinguishes and envisages the data by using mathematical functions to the input sample points. The trend surface changes progressively and apprehends coarse-scale patterns in the data.

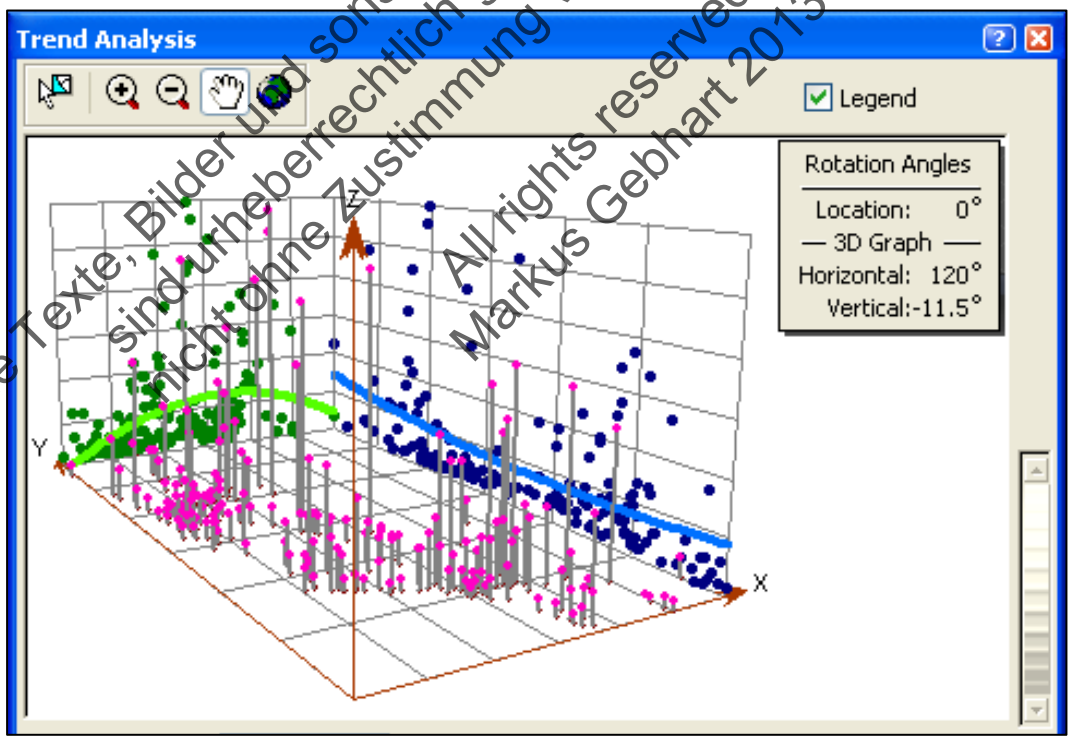
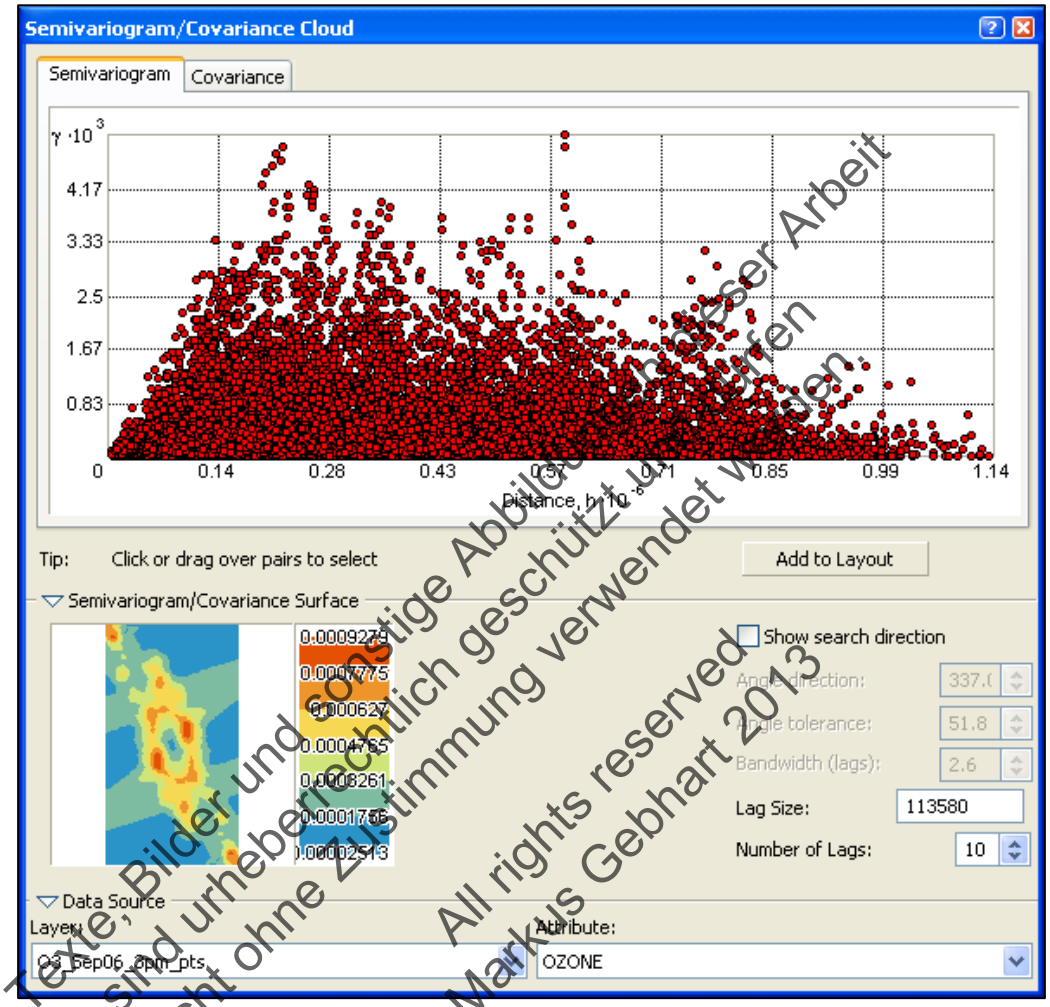


Fig. 4: Trend Analysis

At a last step it is also possible to create a "Semivariogram/Covariance Cloud" (Fig. 5). It "shows the empirical semivariogram and covariance for all pairs of locations within a dataset and plots them as a function of the distance between the two

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

locations." (ARCGIS 9.2 DESKTOP HELP, 2010). The semivariogram/covariance cloud can be used to observe the local features of spatial autocorrelation within a dataset and look for local outliers. The semivariogram cloud looks like this:



Alle Texte, Bilder und sonstige Abbildungen sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

Fig. 5: Semivariogram

Each red dot is the observed semivariogram (the modification shaped plotted comparative to the separation distance) between a pair of locations in the dataset.

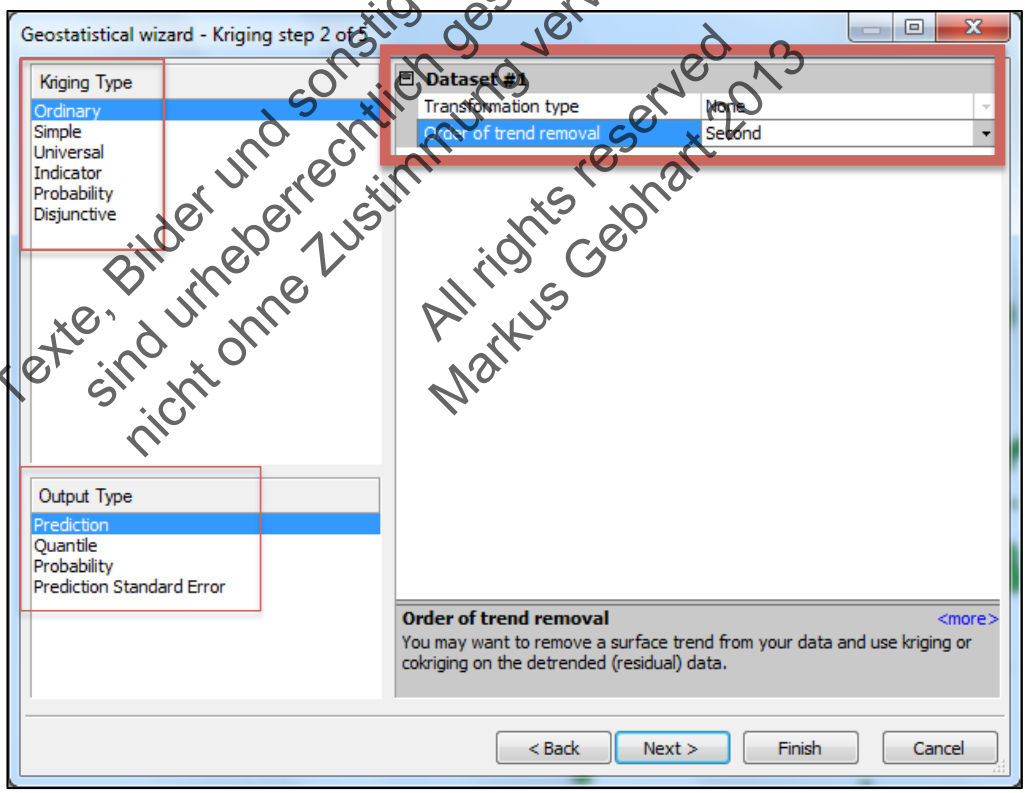
Commonly it is highly endorsed to convert the data if there is no normal distribution given – neither in the "Normal Q-Q-Plot" nor in the "Histogram". Consequently the data should be transformed before the "Kriging" to correspond to a normal distribution.

Assignment_4

Part 2: Kriging

ArcGIS proposes three different ways to convert the data because a normal distribution is compulsory for some geostatistical experts. This transformation can be solved with the Kriging and Co/Kriging tool of the Geostatistical Wizard. Kriging can be used to interpolate a surface from points. "Kriging is a processor-intensive process. The speed of execution is dependent on the number of points in the input dataset and the size of the search window." (ARCGIS 9.2 DESKTOP HELP, 2010).

As we see in Fig. 6, the transformation type can be chosen. It is not necessary to use a transformation type for the Swiss precipitation data. So I specified an *Ordinary* kriging with no *Transformation* type and a *Prediction* output.



Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

Fig. 6: Include transformation type

As we can see in the dialog box below there is an order to explore the possibilities for a trend removal, which allows us to inspect the trend surface we were using – but this can be done later in an additional window of the wizard.

The next step is the setup of the Method Properties (Fig. 7). It is important to choose an own properties setup because by default the Geostatistical Analyst settings just picture an global trend of the map, which has to be adopted to the needs of the data. Then the model should be optimized for bandwidth, spatial condition number and the searching neighborhood values.

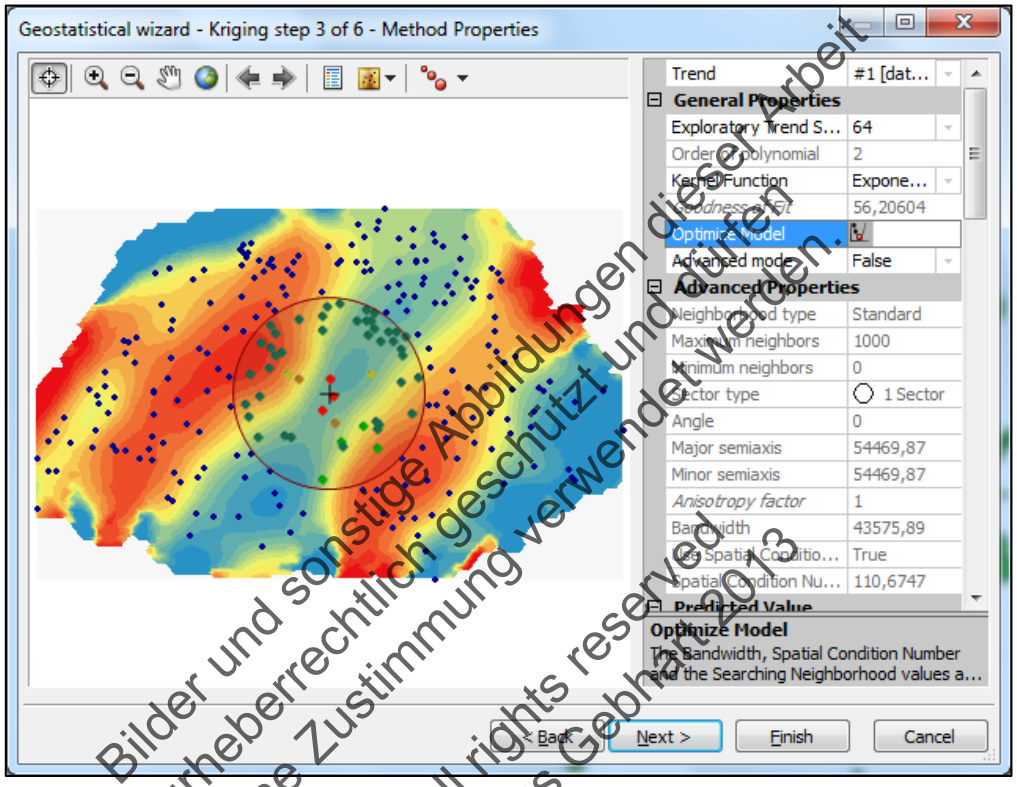


Fig. 7: Method Properties

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

Step 4 of 6 is the Semivariogram/Covariance Modeling Cloud where the semivariogram plot can be seen. In general you can say that kriging is an interpolator that can be exact or smoothed depending on the measurement error model. It is very flexible and allows investigating graphs of spatial auto- and cross-correlation. Some diverse functions from the available list of functions were suggested to display the graphics. I tried some functions and experimented with them but sticked with the default settings by now at last. “By removing the trend, the semivariogram will model the spatial autocorrelation among data points without having to consider the trend in the data.” (ARCGIS 9.2 DESKTOP HELP, 2010). The trend will be automatically added back to the calculations before the final surface is produced.

The color scale, which symbolizes the analyzed semivariogram value, delivers a direct connection between the first-hand semivariogram values on the diagram and those on the semivariogram surface. The significance of each cell in the semivariogram surface is color-coded, with higher values shown in orange and red and lower values shown in blue and green.

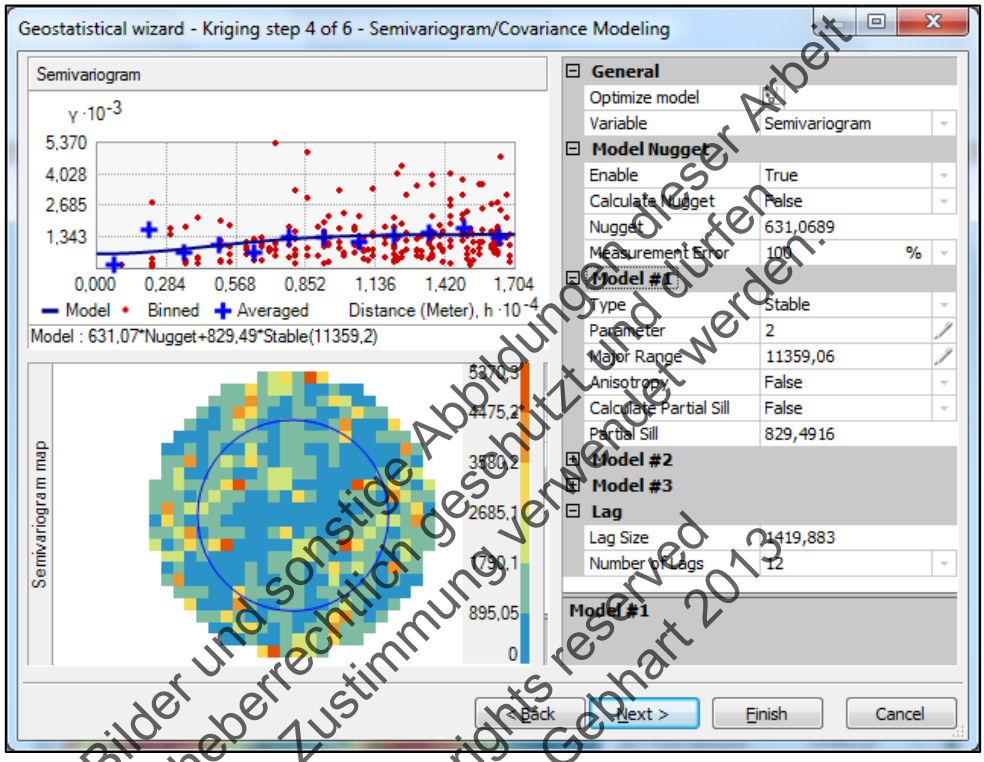


Fig. 8: Semivariogram/Covariance Modeling

The standard value for each cell of the semivariogram surface is contrived on the semivariogram graph and represented as a red point. The average value for each interval (which comprehends various cells) is plotted on the semivariogram graph as well, and is shown as a blue cross.

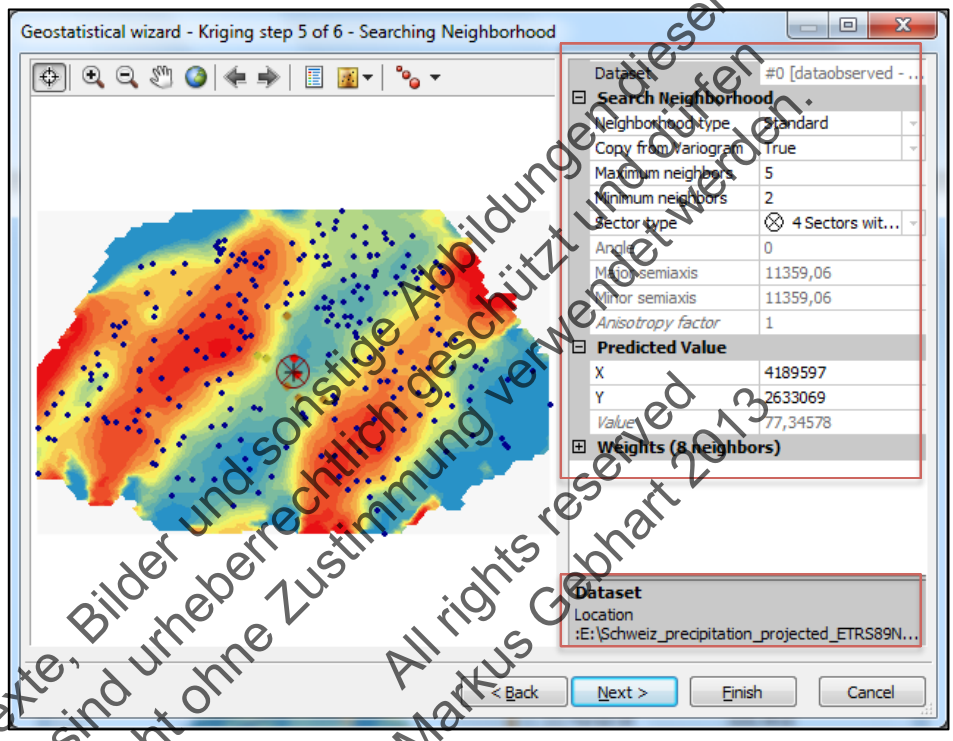
The next step, 5 of 6, is the important step of “Searching Neighborhood” (Fig. 9). For this setting there are three different setting-tags (Search Neighborhood, Predicted Value and Weights) available.

In the first tag you can see previews of what the interpolated surface will look like with the currently given selected parameters. You can also specify how many neighbors to include using the “*Method: Neighborhood*” options. You can decide how many neighbors to include and also how they are distributing around the

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2012

location to be estimated using the *Sector Type* options. The several 'pie slice' options, describe numerous regions around the location which are each necessary to comprise the needed amount of neighboring control points as indicated by the minimum and maximum neighbors options.

A wide collection of neighborhood types can be applied through these options and it is good to experiment. Frequently, it takes intense changes to the neighborhood pattern to make much more evident distinction to the outcome surface.



Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

Fig. 9: Searching Neighborhood

After that step, the “Cross Validation” dialog box shows up, where you can execute diagnostics on the parameters to control how good the model will be. The objective of this step is to give “you an idea of how well the model predicts the unknown values.” (ARCGIS 10 DESKTOP HELP, 2011). The Cross Validations of different settings will be compared later on, so the next figure (Fig. 10) just shows the result of the optimized kriging with default settings in Geostatistical Analyst.

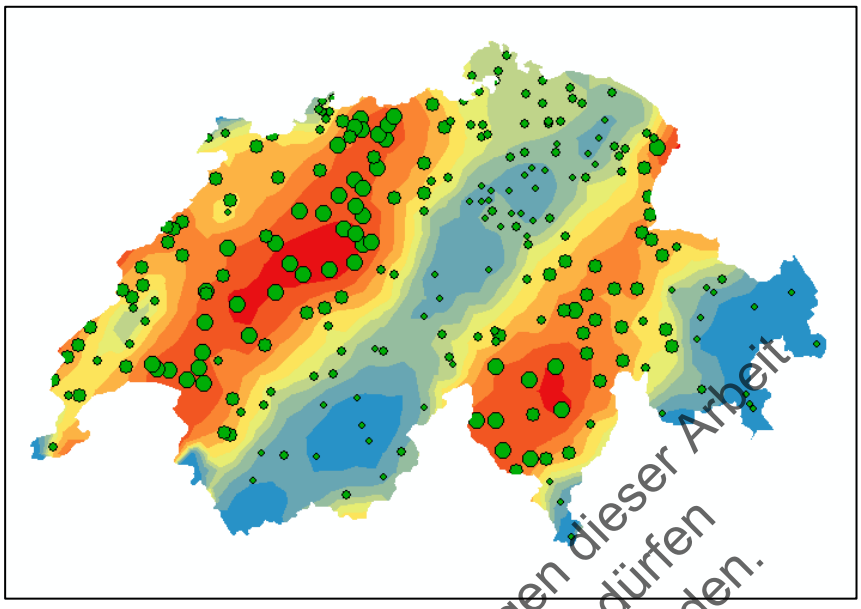


Fig. 10: Optimized Kriging (default settings)

I will now show some comparisons of different optimized settings in step 5 of 6 (Searching Neighborhood) and their results and later on some own values:

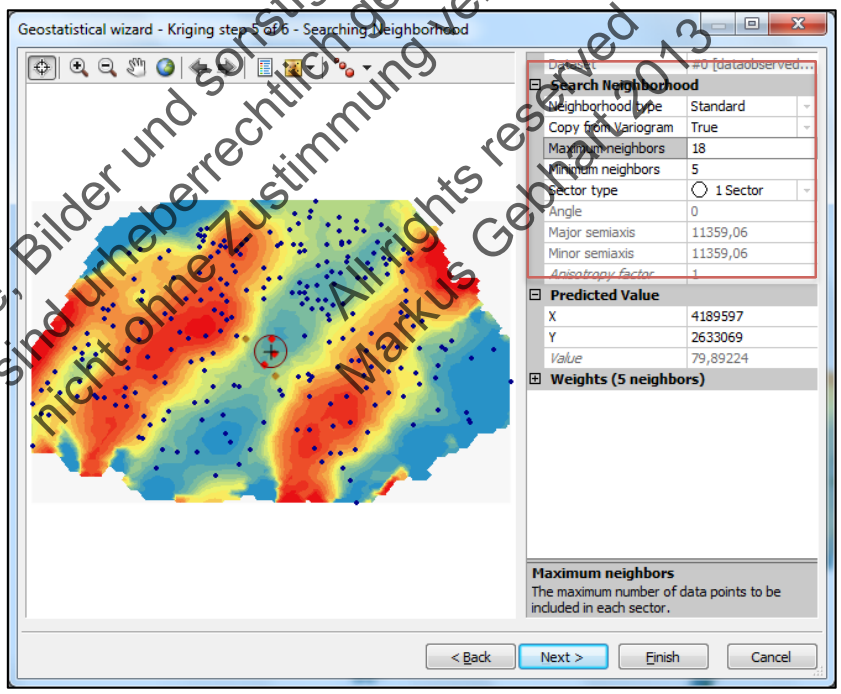


Fig. 11: Optimized Neighborhood Setting #1

As we can see in Fig. 11 I choose the maximum neighbors to 18 and kept the minimum neighbors with 5. I also changed the Sector type to only 1 Sector. The result can be seen in Fig. 12:

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart, 2013

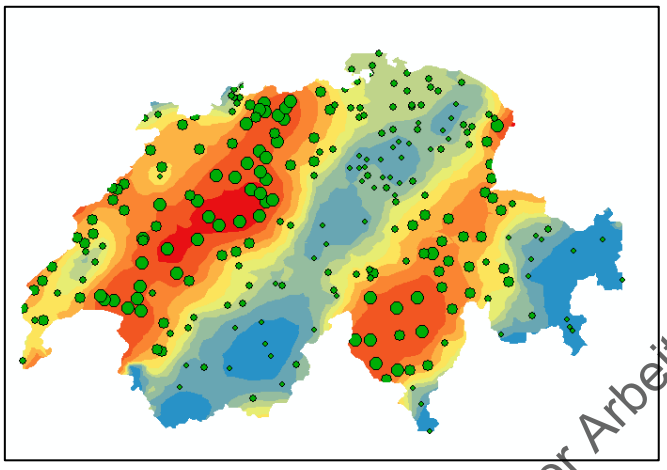
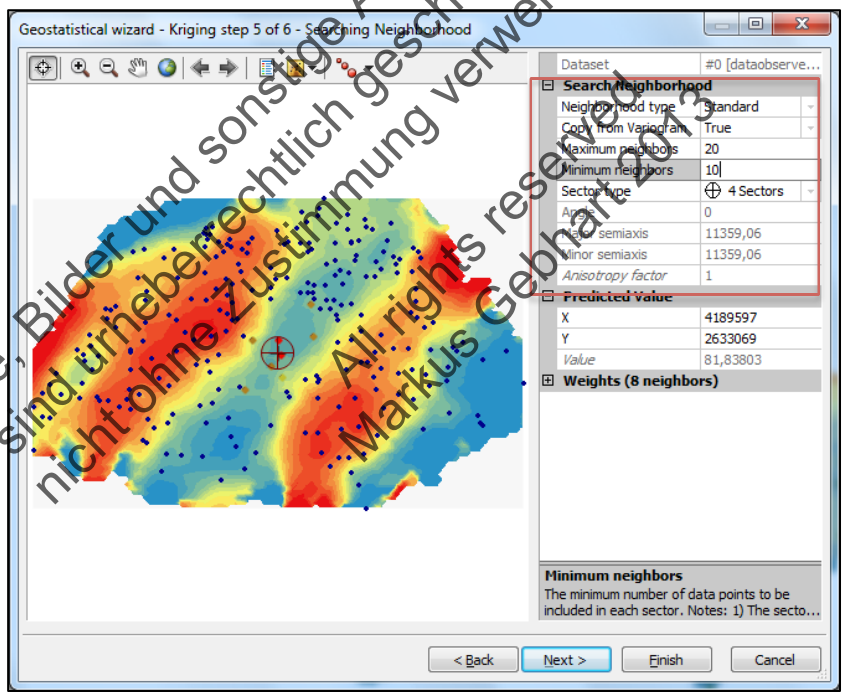


Fig. 12: Result of optimized neighborhood setting #1

In Fig. 13 I choose the maximum neighbors with 20 and changed the minimum neighbors to 10. I also changed back the Sector type to 4 Sectors. The result is nearly similar to Fig. 11 (Optimized Neighborhood Setting #1) and can be seen in Fig. 14



Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2012

Fig. 13: Optimized Neighborhood Setting #2

This result is not quite surprising because I did not change anything in the Method Properties in step 3 of 6. I just changed the maximum and minimum number of neighbors that were used to estimate the value at the unknown location on the one hand and on the other hand the Sector Type, which refers to the geometry of the

neighborhood. (one sector – single ellipse, four sectors – ellipse divided in four sectors).

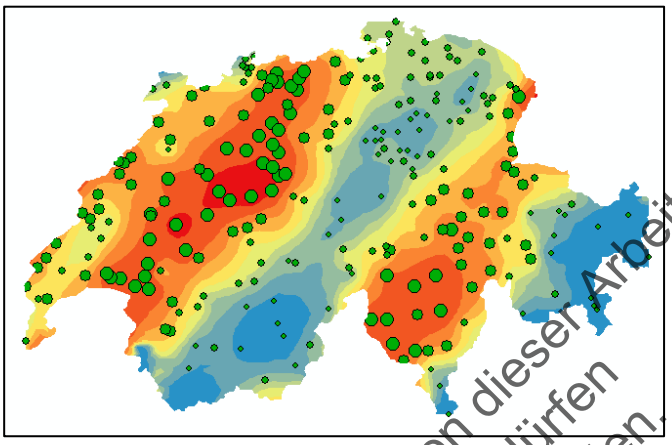


Fig. 14: Result of optimized neighborhood setting #2

Now I want to do an ordinary kriging by using own values. So first I changed some values in the Method Properties. I downscaled the Exploratory Trend Surface from 64 to 47 and also changed the Kernel Function from exponential to polynomial5. This causes an adjustment from proportionally growth or decrease to a fifth-order polynomial function.

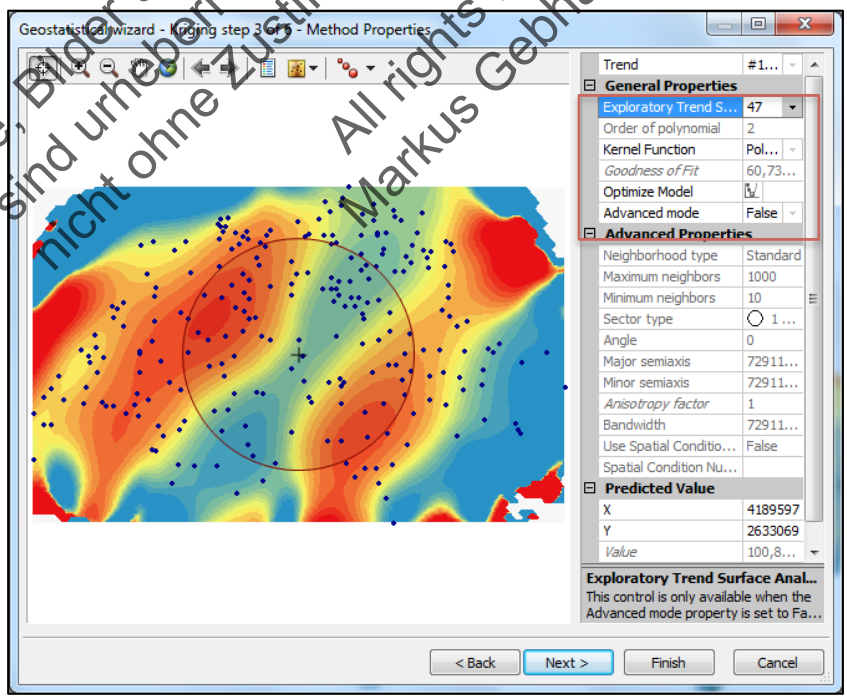


Fig. 15: Method Properties modification - own values

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

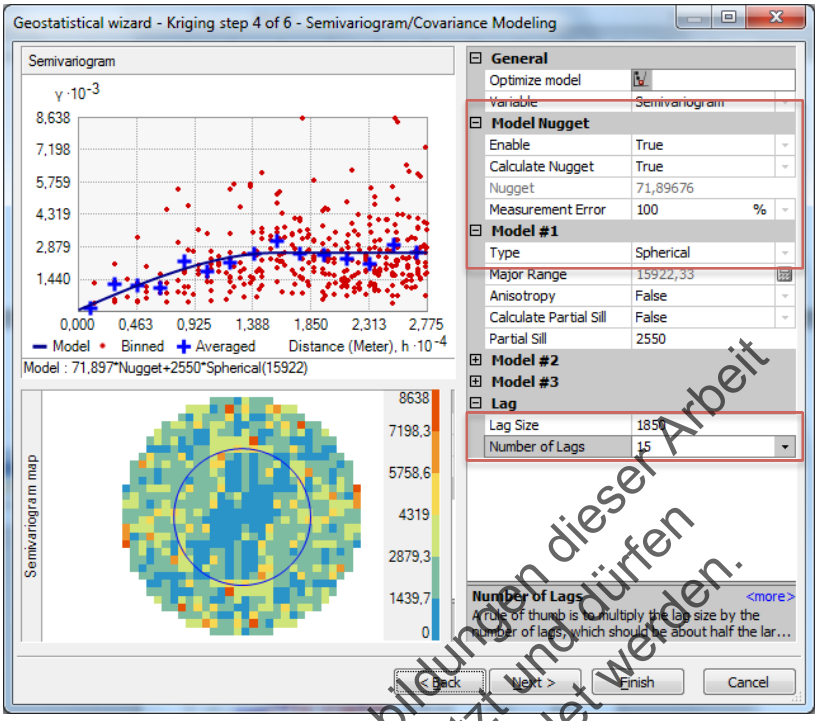


Fig. 16: Semivariogram/Covariance modification

In the next step, the Semivariogram/Covariance Model, I calculated the Nuggets and switched from a stable type to a spherical type. The number of lags were also increased from 12 to 15 and the size from 1419,883 to 1850. "The nugget represents measurement error and/or microscale variation (variation at spatial scales too fine to detect." (ARCGIS 10 DESKTOP HELP, 2011). The change of lags size simulates the "size of a distance class into which pairs of locations are grouped to reduce the large number of possible combinations." (ARCGIS 10 DESKTOP HELP, 2011). So this means that I zoomed out of the model a little. Therefore I had to increase the number of lags a rule of thumb is to multiply the lag size by the number of lags. I kept the Searching Neighborhood settings as shown in Fig. 9 – so the default values. As Fig. 17 shows there are already differences according to the default setting (Fig. 9)

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

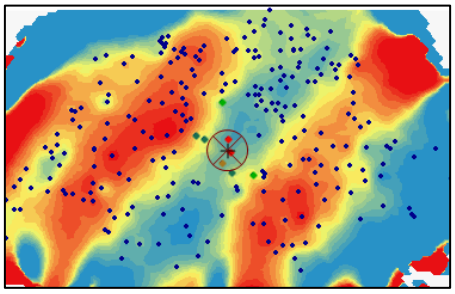


Fig. 17: Searching Neighborhood - default settings

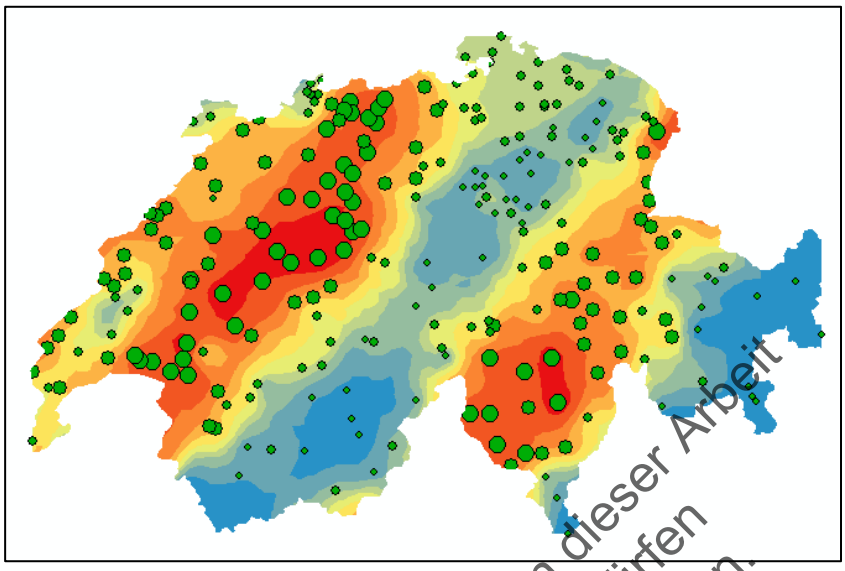


Fig. 18: Result own values - default searching neighborhood

By changing the Sector Types from 4 Sectors (Fig. 17) to only 1 Sector (Fig. 19) and leaving all other settings the same, we see a huge difference in the two figures. Also the result diverges from the result before (Fig. 20).

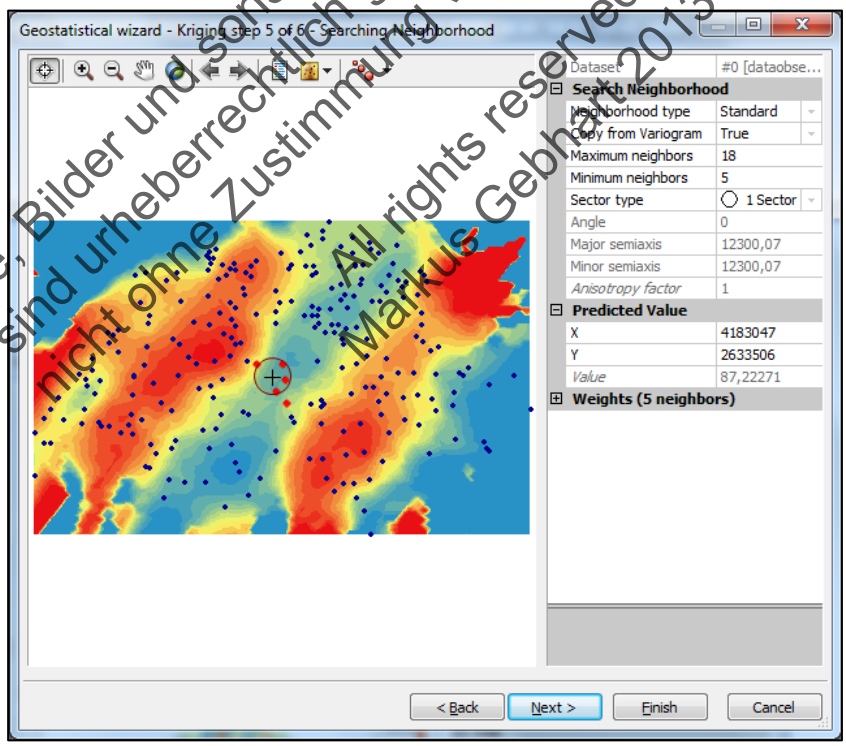


Fig. 19: Changed Sector Type from 4 to 1

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

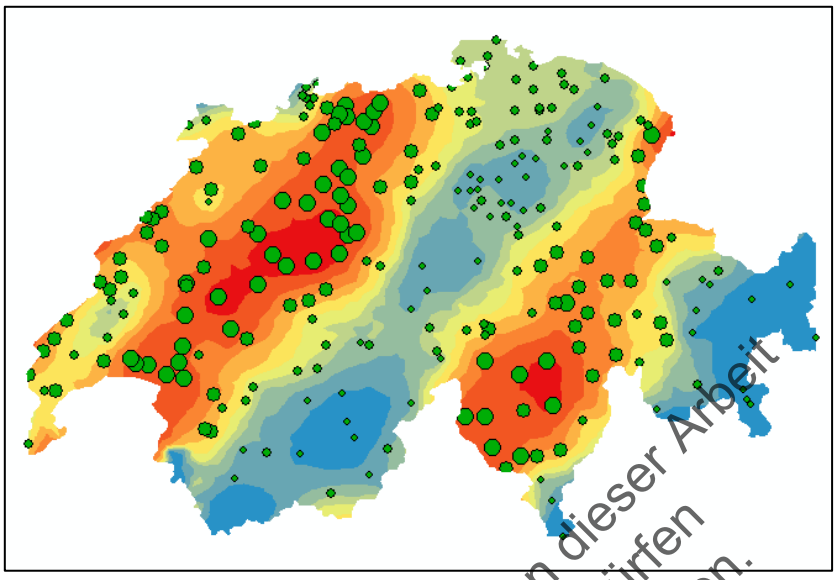


Fig. 20: Result different Sector Types

I did the same procedure for the example from Fig. 13 (Optimized Neighborhood Setting #2). The result can be seen below.

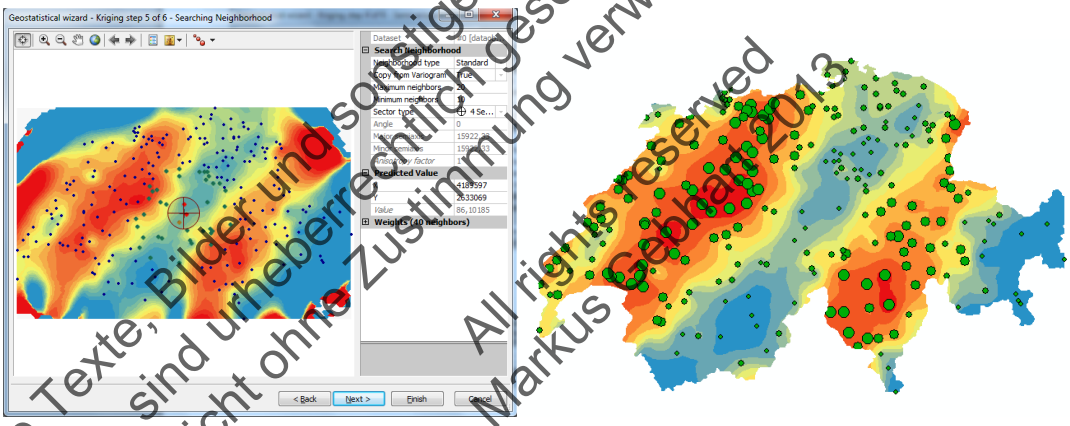


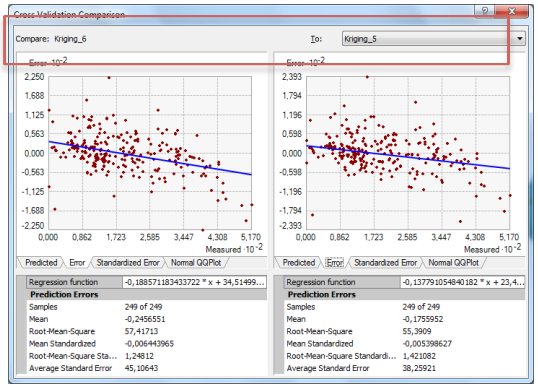
Fig. 21: Optimized Neighboorhood Setting #2 and result with own values in Method Properties and Semivariogram

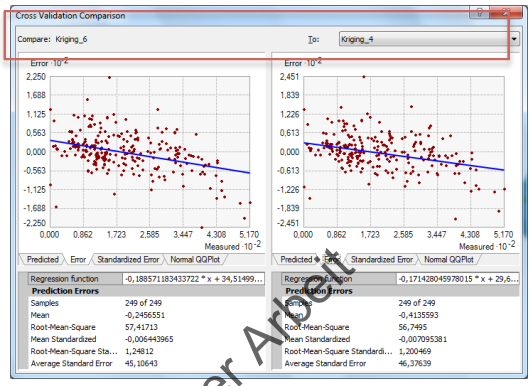
Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

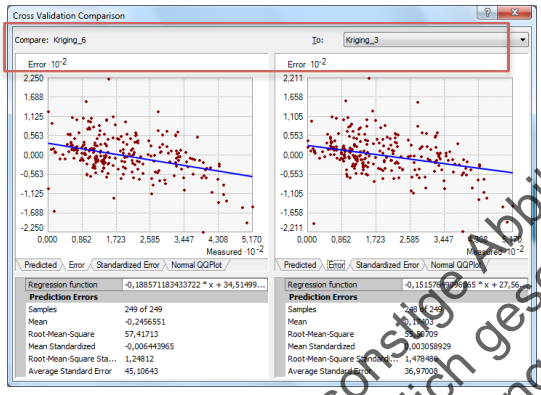
As I mentioned before I wanted to compare the Cross Validation of all the krigings I've done. I decided to compare each Kriging with each other. The krigings shown here are just compared to Kriging_6, but you can see very well, that the result was not satisfactory. The distribution of dots along the standard line is quite widespread.

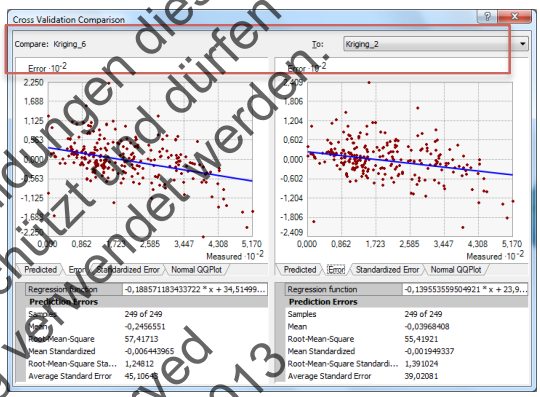
Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

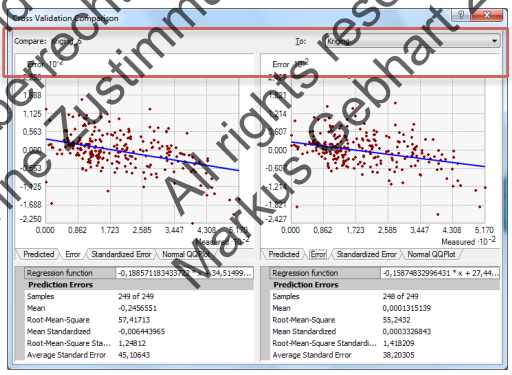
Later I compared the errors of each kriging. The standard errors can be seen below:





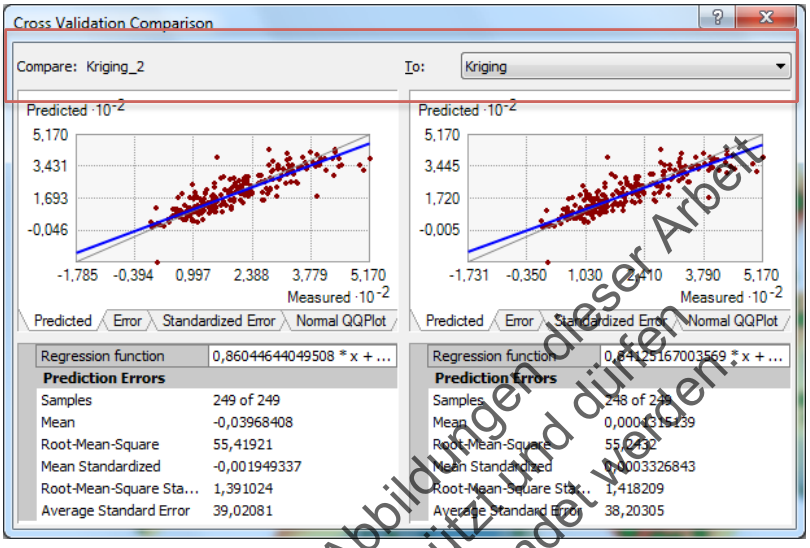






Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

The best kriging model according to the cross validation parameters comparison is Kriging_2 with Kriging. The distribution of points is quite good here, so it would be a good decision to produce a map out of that kriging.



The "Error prediction map" (Fig. 22) shows the differences between the measured and predicted values in the different krigings.

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

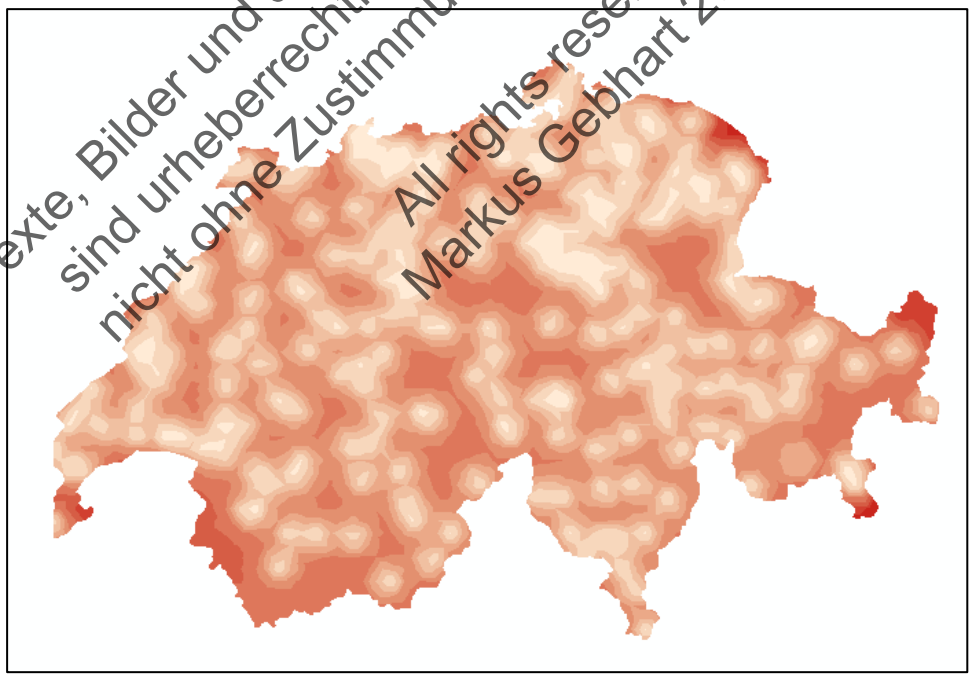


Fig. 22: Error prediction map

The prediction standard errors measure the insecurity for each location in the surface which I created. "A simple rule of thumb is that 95 percent of the time, the true ozone value will be inside the interval formed by the predicted value ± 2 times

the prediction standard error, assuming that the data is normally distributed. Notice in the prediction standard error surface that locations near sample points generally have lower error." (ARCGIS 10 DESKTOP HELP, 2011).

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.

All rights reserved
Markus Gebhart 2013

Table of Figures

FIG. 1: HISTOGRAM OF DATA 1
FIG. 2: NORMAL Q-Q-PLOT 2
FIG. 3: VORONOI MAP 3
FIG. 4: TREND ANALYSIS 3
FIG. 5: SEMIVARIOGRAM 4
FIG. 6: INCLUDE TRANSFORMATION TYPE 5
FIG. 7: METHOD PROPERTIES 6
FIG. 8: SEMIVARIOGRAM/COVARIANCE MODELING 7
FIG. 9: SEARCHING NEIGHBORHOOD 8
FIG. 10: OPTIMIZED KRIGING (DEFAULT SETTINGS) 9
FIG. 11: OPTIMIZED NEIGHBORHOOD SETTING #1 9
FIG. 12: RESULT OF OPTIMIZED NEIGHBORHOOD SETTING #1 10
FIG. 13: OPTIMIZED NEIGHBORHOOD SETTING #2 10
FIG. 14: RESULT OF OPTIMIZED NEIGHBORHOOD SETTING #2 11
FIG. 15: METHOD PROPERTIES MODIFICATION - OWN VALUES 11
FIG. 16: SEMIVARIOGRAM/COVARIANCE MODIFICATION 12
FIG. 17: SEARCHING NEIGHBORHOOD - DEFAULT SETTINGS 12
FIG. 18: RESULT OWN VALUES - DEFAULT SEARCHING NEIGHBORHOOD 13
FIG. 19: CHANGED SECTOR TYPE FROM 4 TO 1 13
FIG. 20: RESULT DIFFERENT SECTOR TYPES 14
FIG. 21: OPTIMIZED NEIGHBOORHOOD SETTING #2 AND RESULT WITH OWN VALUES IN METHOD PROPERTIES AND SEMIVARIOGRAM 14
FIG. 22: ERROR PREDICTION MAP 17

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013

Reference

- ARCGIS 9.2 DESKTOP HELP (2010).
  - http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Voronoi_maps
  - http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Histogram
  - http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=kriging
  -
- ARCGIS 10 DESKTOP HELP (2011).
  - http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//0031000000pz000000.htm
- ESRI (2012). http://www.esri.com/software/arcgis/extensions/geostatistical

Alle Texte, Bilder und sonstige Abbildungen dieser Arbeit sind urheberrechtlich geschützt und dürfen nicht ohne Zustimmung verwendet werden.
All rights reserved
Markus Gebhart 2013