Métodos para encontrar modelos de crecimiento forestal de gran escala espacial y sus aplicaciones prácticas

Methods to detect large-scale spatial patterns of forest growth and their practical applications

***Rüdiger Unseld*, Hans-Peter Kahle, Heinrich Spiecker**
Institute for Forest Growth; University of Freiburg
Tennenbacherstrasse 4; 79085 Freiburg
instww@iww.uni-freiburg.de
Germany

Resumen: se considera que el incremento radial o de altura, parámetros comunes de crecimiento forestal, se ven afectados principalmente por factores ambientales locales a pequeña escala, tales como propiedades de suelo locales o condiciones climáticas locales. En este estudio se presentan métodos geoestadísticos para identificar modelos de crecimiento forestal a gran escala espacial para ayudar en la interpretación de influencias ambientales espaciales utilizando modelos típicos de variaciones de crecimiento. Para conseguir esto, se calcularon correlogramas (cruzados), semivariogramas (cruzados) y regresiones. En un ejemplo, se pudo mostrar la autocorrelación entre parámetros importantes de crecimiento a gran escala. Unos resultados adecuados mostraban la sensibilidad como valor promedio para la variación del incremento radial durante un perído de tiempo largo y el índice de crecimiento anual como valor individual para la variación del incremento radial en un período excepcional de sequía. En una comparación inicial del incremento radial con factores ambientales, se indentificó la precipitación como uno de los principales controladores del crecimiento radial anual a gran escala, independientemente de, por ejemplo, propiedades de suelo. Una vez que se identificaron modelos apropiados, la información puntual fue generalizada a un dominio espacial utilizando Ordinary Kriging. Los resultados se compararon con regresiones lineales y regresiones ponderadas geográficas. Para las regresiones ponderadas geográficas, se encontró que las estimaciones mejoraban de manera significativa. Aplicaciones prácticas de métodos geoestadísticos incluyen la observación de cambios a gran escala en crecimientos forestales y la optimización de diseños experimentales para el monitoring del crecimiento forestal en orden de reducir costos de inventarización y maximizar el output de información.

Palabras claves: crecimiento forestal, geoestadística, regresiones ponderadas geográficas, anillo de crecimiento

Abstract: Radial, or height increment, common parameters of forest growth, are primarily considered to be affected by local environmental factors on a smaller scale such as local homogenous soil properties and local climate conditions. In this study geostatistical methods are presented to identify large-scale spatial patterns of forest growth to assist in the interpretation of spatial environmental influences using patterns of typical growth variations. To achieve this, (cross-) correlograms, (cross-) semivariograms and regressions were calculated. In an example the spatial autocorrelation of important growth parameters could be shown on large scale. Adequate results were obtained for the variation in radial increment over a prolonged time period expressed as sensitivity, and at a point in time to determine annual increment during a drought period. An initial comparison of the radial increment with environmental factors identified precipitation one of the main drivers of annual radial increment on large scale, which is independent of soil properties for example. Once suitable models were identified, the point information was generalised into a spatial domain using Ordinary Kriging. The results were then compared with results of Linear and Geographically Weighted Regression. For Geographically Weighted Regression the estimations were found to improve significantly. Practical applications of the geostatistical methods presented include the observation of large-scale changes in forest growth and the optimisation of experimental designs for forest growth inventories in order to reduce inventory costs and maximise the information output.

Keywords: forest growth, geostatistics, geographically weighted regression, tree-ring

1. Introduction

In many disciplines of ecology, a common practise of analysis is to identify and quantify spatial patterns (CLIFF & ORD, 1981; ROSSI ET AL., 1992; ASPINALL, 1992). In fact, most environmental data are inherently composed of several levels of spatial structure: large scale and small scale patterns together with local random noise (GOOEVAERTS, 1997; WEBSTER & OLIVER, 2001). Spatial patterns arise with the presence of spatial autocorrelation. Spatial autocorrelation is given when the value at any one point in space is dependent on values of the surrounding points. That is, the arrangement of values is not just random. Positive spatial correlation means that similar values tend to be near each other. Large-scale patterns of forest growth are a result of positive spatial autocorrelation, not only between a value and its closest neighbours, but also with values far away. One general opinion about the growth behaviour of trees is that tree growth mainly responds to small scale influences like local weather conditions and soil properties. The response to large scale influences like climate conditions is often not appropriately considered. However, growth reactions on large scales are of great interest, considering possible effects of large scale climate-changes or effects of atmospheric pollutants. This applies to all forested areas, from the tropics up to the boreal zone. In the following text, we present some specific data analysis methods to detect and describe spatial patterns and the main drivers which effect growth on a larger scale.

2. Statistical background

In this study, our focus is on the class of second-order spatial statistics with two main branches. The development of these branches results mainly from the type of spatial data, which are either spatial point pattern data or geostatistical data. Spatial point pattern data are point "events" within a study area at spatial locations. Often, no response variable is taken at the location and the location is the datum. Examples are the distribution of individual trees in a field, or the location of homes of individuals with certain attributes within a defined area. For geostatistical data, the response variable exists at every point in the study region. Examples are the distribution of a metal ore across a landscape, pollution variation across space or acid rain deposition across a region.

For the first type of data, a set of spatial statistics was developed mainly for applications in health or criminalistic studies (FORTIN, 1999). In the case of inferential applications they test spatial significance of correlations between spatial data sets. The set of methods, using geostatistical data, was originally developed for mining purposes and the search of oil. Here, the primary goal was to build exact models or simulations to get a continuous map of the data preferably able to handle sparse data sampling. The main characteristic of these type of data is described, in statistical terms, as non-stationarity or as spatial discontinuity of the data. Hence, the application of point interpolation methods often differs between geological and ecological applications because, in many cases, smooth gradients do not exist for ecological observations like animal abundance, and the principle of stationarity is violated. It is not possible to test the model-assumption of the stationarity with a statistical test. The assumption is only justifiable in regards to content and corresponds to an extra-mathematical experience. Only a detailed explorative analysis at the beginning of a geostatistical study allows conclusions, whether stationarity exists or not. In our study example, we assume stationarity for forest growth at a large scale within an area of more than 30.000 km^2. This is justified by considering tree growth not only as a point event but an expression of the action of continuous environmental factors like soil properties and climate. Tree growth parameters often do not change abruptly across soil boundaries and vary rather smoothly across space.

In former studies dealing with forests and tree growth parameters, stationarity was assumed, e.g. for interpolations between growth parameters of all trees in a stand, or between sample trees in stands of a continuous forest area (RIHA et al. 1986;

RAMIREZ-MALDONADO, 1988; MANDALLAZ, 1992). Based on these thoughts, we picked out methods of spatial statistics and geostatistics to detect and explain large scale patterns of forest growth. Interested readers may find detailed descriptions of the presented tools in GOOEVAERTS (1997), ROGERSON (2001), FOTHERINGHAM et.al. (2000) or WACKERNAGEL (2003). The analyses were performed with commercial software packages ISATIS 5.0 (www.geovariance.fr), GWR 3.0 (http://www.ncl.ac.uk/~ngeog/GWR) and the open source application R (www.r-project.org).

3. The data and proposed methods to detect spatial patterns and their causes

Data for forest monitoring generally were gathered on sample plots. Conventional statistics may deliver interesting results of quantities, classifications and correlations, but the observation of spatial variations of data requires special methods. Some of the basic methods, which are suitable for the given application are presented in this chapter.

3.1 The data

The most important data on forest growth is data of height and radial growth. Many other quantities to describe forest productivity and growth like site indexes, standing volume or volume increment are based on these data and the presented methods are also applicable to them. In our study (chapter 4), we selected data on annual radial increment measured retrospectively based on the stem analysis technique. This growth parameter can be used as a proxy for changing environmental conditions over long time periods.

3.2 Statistics to detect large scale autocorrelation and patterns of growth data

The following approaches are part of an intensive explorative data analysis.

3.2.1 Moran's "I"

There are a number of ways to measure spatial autocorrelation of spatial point data. One of the most common is the calculation of Moran's "I" (FOTHERINGHAM ET AL., 2000; ROGERSON, 2001). Moran’s “I” is calculated based on the spatial weight matrix, where the weight *W* is the inverse of the distance between two points. For any continuous variable x_i, a mean can be calculated and the deviation of any one observation from that mean can also be computed. The statistic then compares the value of the variable at any one location with the value at all other locations.

$$I = \frac{N\sum_i \sum_i W_{ij}(X_i - \overline{X})(X_j - \overline{X})}{(\sum_i \sum_j W_{ij})\sum_i (X_i - \overline{X})^2}$$

N: Number of Observations
X_i: Variable value at a particular location *i*
X_j: Variable at another location where $i \neq j$
W_{ij}: Weight applied to the comparison between location *i* and *j*

For data tested by Moran’s “I” the null hypothesis is that there is no spatial autocorrelation. It is usually rejected at the p=0.05 confidence level. Values of “I” greater than the expected “I” indicate clustering, while values of “I” less than the expected “I” indicate dispersion. The significance test indicates whether these differences are greater than what would be expected by chance.

3.2.2 Variograms

Variogram calculations are standard geostatistical methods for measuring patch size in a landscape. We used them as a suitable tool to detect large scale growth patterns. Important varieties of variograms are the semivariogram, indicator variogram, madogram and relative variogram.

Semivariogram

The variogram allows a detailed description of overall spatial patterns. It is the fundamental tool of geostatistics which establishes the correlation as a function of the distance by measuring the average dissimilarity between data on sample points, separated by a distance vector h. It is computed as half of the average difference between the components of every data pair,

$$\gamma(h) = \frac{1}{2N(h)} \sum_{N(h)} [\ z(x_i) - z(x_i + h)\ ]^2$$

where x_i is a data location, h is a lag vector, $z(x)$ is the data value at the location x_i, N is the number of data pairs at distance h and $z(x_i) - z(x_i + h)$ is called the h-increment of the variable z.

With the basic structures of the semivariogram different scales of autocorrelation are visible in different directions and conclusions can be made regarding small- or large-scale behaviour of growth parameters. The basic structures of the semivariogram are the *lag*, the *sill*, the *range* and the *nugget* (Isaaks & Srivastava, 1983). The lag is the chosen distance class interval used for variogram computation. The range indicates the distance over which values are spatially correlated. Values are no longer spatially dependent at the lag distance of the sill. The sill is the variogram value corresponding to the range and reflects the maximum variance of the data. The nugget represents the extent of random variation within the data. According to the equation above, variance at the distance of 0 is theoretically equal to 0. However, sampling errors and small scale variability may often cause sample values within a small separation to be quite dissimilar. Thus, a discontinuity at the origin of the variogram occurs, which is referred as the nugget effect. There are several aspects which have to be considered during the calculation process like the number of data pairs per lag, the maximum distance of computing, the direction of calculation, anisotropy-detection, detection and handling of outliers and the presence of a trend. More detailed information is available in CRESSIE (1993), GOOEVAERTS (1997) or WEBSTER & OLIVER (2001). The advantage of the semivariogram technique is not only its ability to detect scale dependencies in an easy way but also its ability to fit an appropriate theoretical model semivariogram of known mathematical properties, which allows values of unsampled locations to be estimated e.g. by kriging techniques based on the experimental semivariogram. The interpolated data, represented as map, gives further indications about the distribution of large scale patterns.

Indicator variogram

The pattern of continuity (variability) of a category s_k can be characterised by semivariograms defined as an indicator coding of the presence/absence of that category (GOOEVAERTS, 1997). The so called indicator variogram allows to comparing the variogram structures of categorical variables, like rock or soil types and exposition with the semivariogram structures of a growth parameter.

$$i(u_\alpha; s_k) = 1 \quad \text{if} \quad s(u_\alpha) = s_k$$
$$= 0 \quad \text{otherwise}$$

Relative variogram

If forest growth data were gathered with preferential sampling, e.g. in large sample circles with changing numbers of sample trees, a good alternative would be to compute a relative semivariogram instead of the traditional semivariogram (ISAAKS & SRIVASTAVA, 1989). High or low values, which are clustered, may entail undesirable effects like heteroscedasticity and render a semivariogram uninterpretable.

Madogram

The values of the semivariogram are sensitive to extreme data values. One way to handle this problem is to use the madograms (DEUTSCH & JOURNEL, 1992). Instead of taking the mean squared differences (see semivariogram formula above), one can also calculate the mean absolute difference. Avoiding the squared term results in a more robust measure of the correlation structure. Preferentially they can be used to infer different ranges of spatial distributed data, hence they are of special interest in our context.

3.3 Statistics to detect similarities between large scale patterns of growth data and environmental factors

Moran`s “I” and the different variogram approaches detect univariate spatial autocorrelations. In forest ecology it is often desirable to detect the causes of growth phenomena and to describe relations between environmental factors and growth. Suitable models provide insight into the effect-structure of tree-growth.

3.3.1 Cross correlogram and cross semivariogram

One important step in detecting scale relationships between measurements of growth parameters and different environmental variables consists of looking at their cross dependence. A simple possibility to measure similarity of spatial variability is the cross correlogram. It measures how the value at one location, e.g. a growth variable, is related to the value of a distance h apart, e.g. an environmental-variable.
In opposite to the cross correlogram, the cross semivariogram measures the dissimilarity between *h*-increments $z_\alpha(x_i) - z_\beta(x_i + h)$ of growth and environmental variables. The cross semivariogram is defined as half the non-centered covariance between *h*-increments. It can be computed only from those locations where the growth variable and the explanatory variable are gathered. The standard function of the cross variogram is the following

$$\gamma^{p}_{\alpha\beta}(h) = \frac{1}{2N(h)} \sum_{N(h)} \left[z_\alpha(x_i) - z_\alpha(x_i + h) \right] \bullet \left[z_\beta(x_i) - z_\beta(x_i + h) \right]$$

where z_α and z_β are the values of different variables on sample points. If the variables were standardised to zero mean and unit variance, the sill of the experimental cross semivariogram directly reflects the magnitude of the correlation between the variable (GOOEVAERTS, 1997) and the range shows the cross-dependence on small or large scale (see 3.2.2).

3.3.2 Regression models

Cross correlogram and cross semivariogram provide insight into the spatial connectivity between the variables. To gain further information about significant relationships and their spatial behaviour we used linear and geographically weighted regression.

Linear regression

Linear regression is the most common method to detect global relationships by fitting linear equations to observed data. A multivariate linear regression line has an equation of the form

$$y = a_0 + \sum_k a_k x_{ik} + \varepsilon_i$$

where *y* is the dependent variable and *x* are explanatory variables. The slope of the line is a_k, a_0 is the intercept (the value of y when x_{ik} = 0) and *ε* is the random error of prediction. Spatially sampled data often does not fit the independence assumption required for classical linear regression. An important fact is that, if spatial autocorrelation is present in the residuals but not properly acknowledged, the standard errors of regression coefficients from linear regression seriously underestimate the actual variability of the regression coefficients. Only if we include autocorrelation in the analysis of spatially sampled tree growth data, it is possible to prevent misinterpretations of the results. One simple possibility to avoid autocorrelation is to use more sparse data samples, which are taken at a sufficient distance so that the values are not correlated. This may not be practical unless we have a large data set and autocorrelation exists on a large scale. However, if we want to use linear regression to understand the relationship between the explanatory variables and the dependent variable, or to estimate which coefficients are globally significant, then we can trust the results. Hence we used in the first step stepwise linear regression, ignoring spatial autocorrelation, to narrow the set of explanatory variables to a manageable size. After model fitting, autocorrelation of the residuals needs to be tested, e.g. with Moran´s “I”, and visualised by semivariograms. Further on the spatial patterns of the residuals can be interpolated with ordinary kriging and mapped. Visualisation of the residuals is helpful to identify effects of potentially omitted variables in the data set.

Geographically weighted regression (GWR)

With GWR we describe a statistical model which accounts for spatial autocorrelation of the residuals and which allows us to calculate a continuos map of interpretable prediction errors and spatial autocorrelated regression coefficients. This supports the interpretation of potential causes of large scale patterns. The applied method of GWR was developed by Fotheringham, Brunsdon and Charlton (e.g. BRUNSDON et al. 1996, FOTHERINGHAM et al. 2000; 2002).

As mentioned in Chapter 2, one of the basic assumptions of spatially sampled data is, that point pattern data like data of tree growth are stationary over the entire sample area. For global analysis, like linear regression, another assumption is that the relationships between tree growth and environmental growth factors are also stationary over space. But especially in forest growth analysis a global regression equation can be misleading and relationships might be intrinsically different across space. The geographically weighted regression (GWR) is based upon local views of regression as observed from any location. GWR permits the parameter estimates to vary locally by

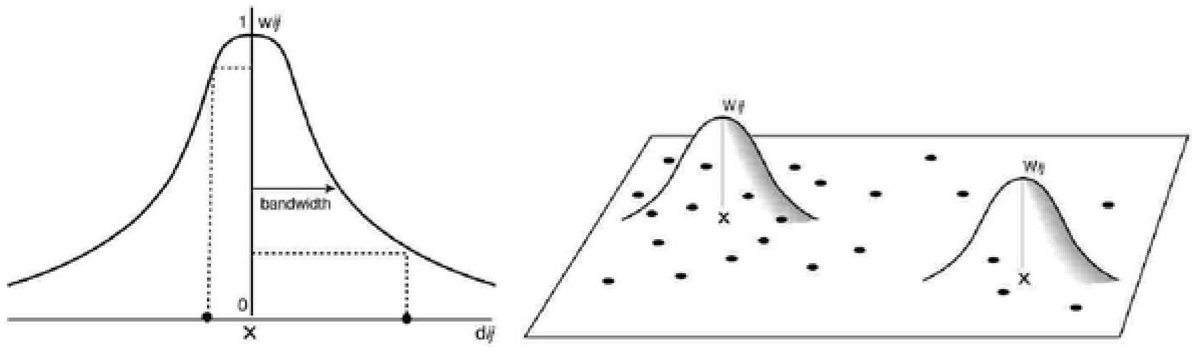
$$y(g) = a_0(g) + \sum_k a_k(g) x_{ik} + \varepsilon_i$$

where (g) indicates that the parameters are to be estimated at a location whose coordinates are given by the vector *g*. Weights are attached to observations surrounding the location. The weights are chosen such that those observations near the point in space, where the parameter estimates are desired, have more influence on the result than observations further away. The weighting scheme is

$$w_{ij} = \exp\left[\frac{\frac{-d_{ij}}{h^2}}{2}\right]$$

where h is known as the bandwidth that controls the degree of distance-decay (Fig.1).

Fig.1: Weights and definition of bandwidth in GWR (FOTHERINGHAM et al. 2002)



x = regression points
• = data points
w_{ij} = weight of data points j at regression point i
d_{ij} = distance between regression point i and data point j

One characteristic of GWR that is not immediately obvious, is that the locations at which parameters are estimated need not be the ones at which the data have been collected. As *h* tends towards infinity with no distance decay, the weights tend towards one and GWR becomes equivalent to classical linear regression. Weights of dependent variables with distances greater than one and a half of the bandwidth apart from the regression point are zero. Hence, the choice of bandwidth has a large impact on the results and result interpretation. As the bandwidth becomes large, the estimate of the explanatory variable depends on observations far away from the regression point, which e.g. allows conclusions about the influence of environmental factors on a large scale. One possibility of selecting an optimal bandwidth is to choose *h* on a least-square criterion and via cross validation. Detailed descriptions of this method are published in BRUNSDON et al. (1996) and FOTHERINGHAM et al. (2000, 2002). After the selection of an appropriate bandwidth, a Monte Carlo-Test selects, which of the coefficients of independent variables vary significantly over space. A spatial variation of increasing negative or positive coefficients of explanatory variables shows the increasing slope and influence respectively on the dependent variable. Significantly varying values of intercept give valuable information about omitted variables, which should be included in order to improve the regression equation.

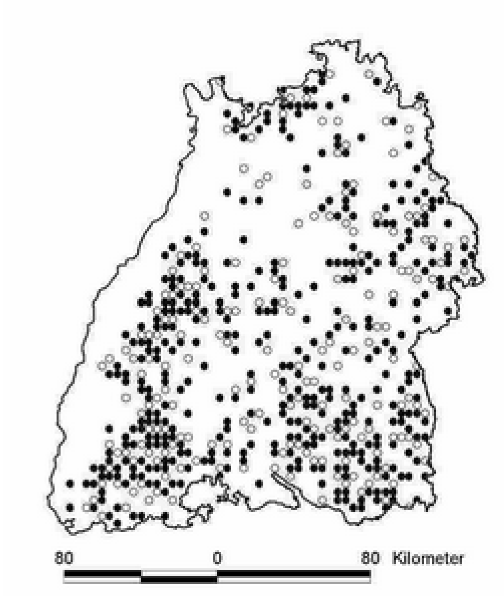
4. Example

In this chapter, the various tools are illustrated using a multivariate data set related to tree ring data.

4.1 The data

Data on annual radial increment was retrospectively derived via the analysis of stem cross sections of Norway spruce (*Picea abies* L. Karst.) and more than 20 environmental variables of a 30.000 km^2 area in the Southwest of Germany (State of Baden-Württemberg) with a N-S-extension of 230 km and a W-E-Extension of 190 km. Data were recorded at 568 sample plots scattered on a systematic 4 km x 4 km grid. (Fig.2). The large sample size allows the data to be divided into a modelling set (379 samples) and a validation set (189 samples). The validation samples are used to check the results provided by the estimation methods proposed. Tree-ring- and foliar-nutrient-analyses of dominant trees have been conducted in 1983 and 1988 by the Forest Research Centre Baden-Württemberg, FVA to monitor forest health status (EVERS & SCHÖPFER, 1988; HILDEBRAND & SCHÖPFER, 1993,). Our objective was to detect and quantitatively describe large scale spatial structures of growth variations. To achieve this, the mean sensitivity of annual radial increment was chosen as a meaningful parameter to characterise the behaviour of the diameter growth of the trees. The mean sensitivity was calculated over a period from 1951 to 1983. The mean sensitivity is a relative measure for changes in increment from year to year. High values mean rapid and strong changes in radial increment between subsequent years (see SCHWEINGRUBER, 1988).

Fig.2: Model samples (black) and validation samples (white)



Moreover, the annual radial increment of trees in the extraordinarily dry year 1976 was examined. The radial increments were transformed into relative values (indexation). The index is calculated as the ratio between the radial increment in the year 1976 to the average annual increment over the 10 preceding years and is called in the following as “radial increment index”. The idea behind this transformation is a standardisation of the increment values with respect to age- and/or dimension-related trends as well as to account for different levels of growth due to differences in stand and site conditions (COOK & KAIRIUKSTIS, 1990). Before entering the modelling analyses all candidate explanatory variables have been standardized to a mean of zero and unit variance.

The climate variables were computed from a 1 km x 1 km gridded, spatially interpolated data set produced and delivered by the German Weather Service (DWD) (MÜLLER-WESTERMEIER, 1999). Seasonal values were aggregated either by averaging (temperature) or summing (precipitation). The seasonal aggregates from May to September were considered as growing seasonal figures. Monthly precipitation (relative values; %) and air temperature (absolute values; °K) anomalies were calculated for the year 1976 referring to the mean long-term climatic conditions (1961-1990). These anomalies are further on called “precipitation index” and “temperature index”.

Indexation	Standardisation
$$I_t = \frac{y_t}{\hat{y}_t}$$	$$Z = \frac{x-\mu}{\sigma}$$
I_t - index of radial increment y_t - real increment $\hat{y}$ - equal value responding to age trend	Z - standardised variable $x-\mu$ - value subtracted with the mean σ - standard deviation

4.2 Large scale autocorrelation and patterns

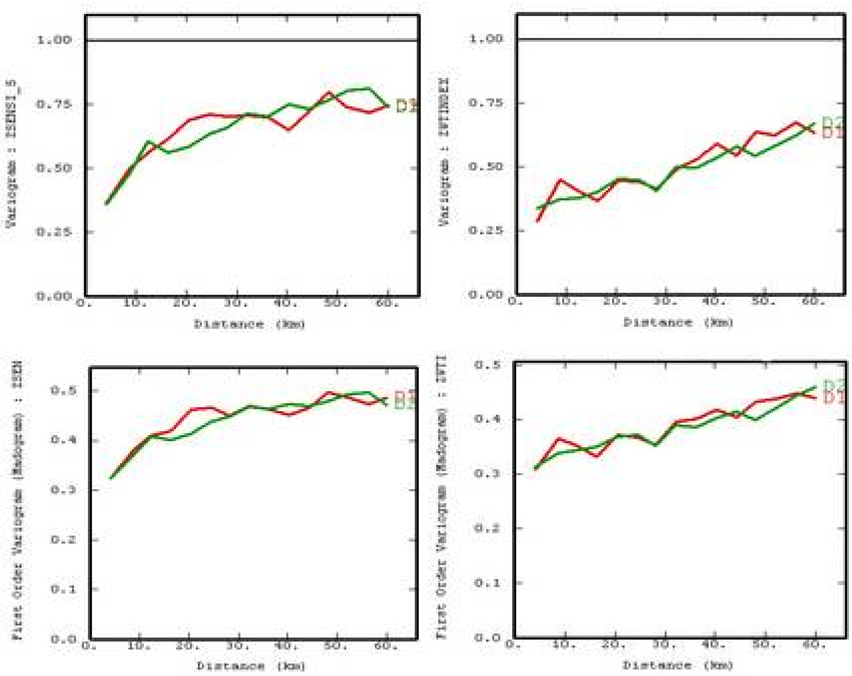
First we investigated the presence and relevance of spatial autocorrelation in the growth data. The global test of Moran`s “I” can confirm, whether there is statistical evidence of clusters (Tab.1).

Tab.1: Autocorrelation of mean sensitivity (MSENS) and radial increment index (RINDEX) during a drought period calculated by Moran`s "I" (n = 378; p ≤ 0.05)

	Moran's "I"	Spatially random (expected) "I"	Normality significance (Z)	Randomisation significance (Z)
MSENS	0.126	-0.00265	24.2	24.2 ***
RINDEX	0.203	-0.00265	38.6	38.6 ***

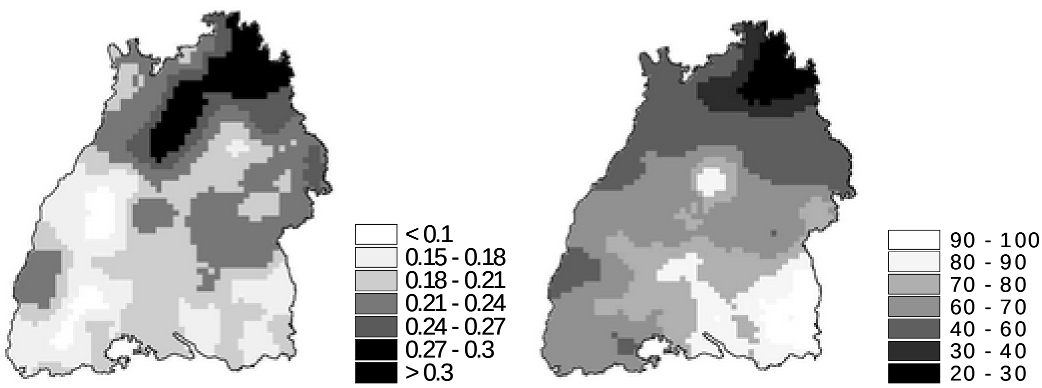
Both types of growth data show evidence of clustering indicated by significant positive autocorrelation. Trees with similar behaviour with respect to sensitivity in radial increment and relative growth response in the exceptionally dry year 1976 globally tend to be located close to each other. The following semivariograms and the madograms give more detailed information about the spatial dependence (Fig.3).

Fig.3: Semivariograms (above) and madograms (below) from standardised values of mean sensitivity (left) and radial increment index(right) in two directions (green line = N-S; red line = W-O)



The presented semivariograms and madograms show similar structures. The semivariograms have a relatively small nugget effect (approx. 30% of total variance) which results from measurement errors and small scale variability (*h* < 4 km) and the pattern of increase does not show directional dependence. Both growth parameter values are spatially autocorrelated in different scales. A short-range structure (< 10 km) and long-range structure (> 30 km) is latently present, however, a regional scale (range 20-30 km) is explicitly observable.. The shape of the radial increment index is dominated by a linear long-range structure (> 50 km). Theoretical variogram models were fitted to the experimental variograms and the coefficients of the models were used to assign optimal weights for a interpolation using ordinary kriging to get a continuous map with interpolated data (Fig.4).

Fig. 4: Mapped mean sensitivity (left) and radial increment index (right) using ordinary kriging



The growth data show distinct special patterns. The mean sensitivity is highest in the north-east of the study area. In this part the radial increment index also decreased drastically during the drought year. The reduction is by more than 70 % of the reference value. The mean sensitivity is low in most parts of the south. However radial increment is not substantially reduced in the south-west, whereas in large areas radial increment decreased by around 30-50 %.

4.3. Large-scale cross dependencies

To detect spatial correlation between growth data and environmental factors the correlation-matrix, cross correlograms and cross semivariogram were computed first. The results provide insight into the spatial effect-structure between of growth data and explanatory data. Table 1 gives the linear correlation coefficients computed among the growth parameters and five of the strongest correlated environmental factors.

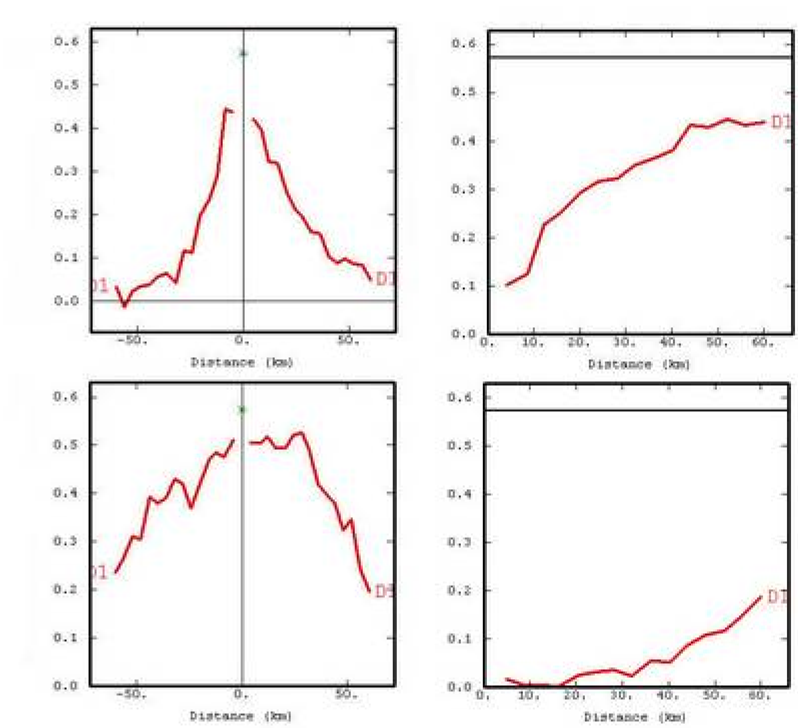
Tab.1: Linear correlation coefficients ρ (n = 379)

Explanatory variable	Sensitivity	Radial increment index
Precipitation index (1976)	-	0,58
Mean precipitation (1961-90)	-0,53	0,28
Altitude	-0,48	0,25
Mean temperature (1961-1990)	0,44	-0,10
Foliar Calcium concentration	0,57	-0,19

Sensitivity is strongly negatively correlated with mean seasonal precipitation and altitude and positively correlated with mean seasonal temperature and foliar calcium concentration. The radial increment index 1976 shows a close correlation with seasonal precipitation (precipitation index1976).

The next step to detect large scale relationships consists of looking at the cross dependence of variables with strong correlation coefficients. To measure similarity of spatial variability of growth and pre-selected predictor variables, cross correlograms and cross semivariograms were calculated (Fig.2).

Fig.2: Cross correlograms (left) and cross semivariograms (right) of sensitivity versus foliar concentration [Ca] (above) and growth index versus precipitation index (below)



The cross-correlogram value (y-axes) at a distance of 0 km and the sill of the cross-semivariogram reflect the magnitude of correlation between sensitivity and foliar [Ca] (see table 1; ρ = 0,57) and radial increment index and precipitation index (ρ = 0,58) respectively. The cross correlograms and cross semivariograms of mean sensitivity versus foliar [Ca] as well as radial increment index versus precipitation index show a large-range cross-dependence. However the cross-dependence of sensitivity versus foliar [Ca] strongly decreases with increasing distance (range 10 - 30 km). The cross dependence of radial increment index versus precipitation index decreases strongly at a range larger than 40 km.

4.4 Large-scale autocorrelation of residuals

The significance of more than 20 explanatory variables were tested by stepwise linear regression. The following models were computed:

mean sensitivity = 0,21 + 0,0125 CA - 0,000068 ALTD – 0,000059 PREC	r^2 = 0,43
radial increment index 1976 = -24,7 + 0,81 PRE_INDEX + 0,053 PREC	r^2 = 0,38

These global results of linear regression suggest that the mean sensitivity of trees was negatively related to altitude (ALTD) and mean precipitation during the growing season (PREC) and was positively related to calcium concentrations in the needles (CA). Mean temperature was strong intercorrelated with the altitude ($\rho > 0,9$) and was excluded from the analysis. Mean sensitivity does not appear to be related to other nutrient variables. Only 43% of the variance is explained by the model. The results for the radial increment index suggest that the radial increments were positively related to the precipitation anomaly but also to the average of precipitation.

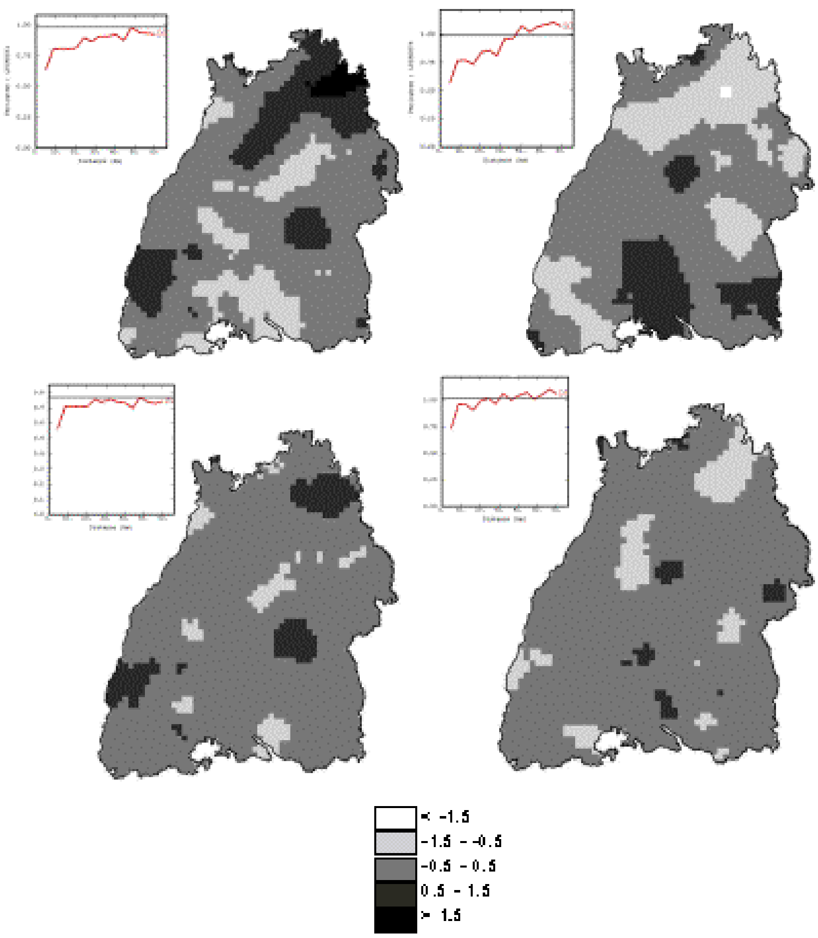
To model the data with geographically weighted regression, the same explanatory variables were used as in the global model of linear regression. The fixed kernel bandwidth (see Fig.1) was determined by cross validation. In case of mean sensitivity the chosen bandwidth was approximately 43 km, so the weighting function was zero at a distance of more than 65 km. With more than 51 % of explained variance, the model replicates the data much better than the linear regression model (43%). The F-test (ANOVA) suggests that the GWR model was a significant improvement. In case of the radial increment index in the drought year 1976 the chosen bandwidth was approximately 21 km, so the weighting function was zero at a distance of more than 31 km. With the GWR-Model 57 % of variance was explained. This was significantly more as with linear regression (38%). Global autocorrelation of residuals was tested with Morans “I” (Table1).

Tab.2: Autocorrelation of standardised residuals from linear regression (LREGR) and geographically weighted regression (GWR) tested with Moran`s "I" (n = 378; p ≤ 0.05)

		Moran's "I"	Spatially random (expected) "I"	Normality significance (Z)	Randomization significance (Z)
LREGR	MSENS	0,232	-0,0027	4,85	4,85
	GINDEX	0,045	-0,0027	8,85	8,86
GWR	MSENS	0,0034	-0,0026	1,13	1,13
	GINDEX	0,0046	0,0026	0,36	0,37

All residuals were significantly autocorrelated. The results of GWR display very moderate, but still significant, global spatial autocorrelation. To get further imformations, the residuals of the two models were represented afterwards as semivariogram and spatially interpolated values (Fig.3).

Fig.3: Semivariograms and interpolated values (OK) of standardised residuals after linear regression (above) and geographically weighted regression (below). Left: sensitivity, right: radial increment index 1976

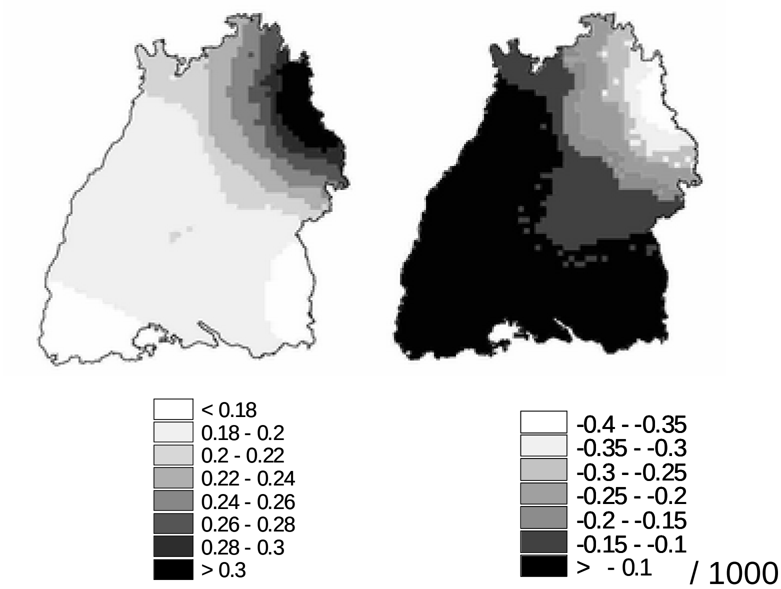


The residuals of linear regression show autocorrelation at small (< 10 km) and larger scales (10 – 40 km) whereas geographically weighted regression prevented large scale autocorrelation and reduced autocorrelation on a small scale. The map shows for linear regression residuals extended areas where sensitivity data was underestimated in the north-east and overestimated in the centre and in the south. The opposite tendency can be detected for the radial increment index. With geographically weighted regression, only small areas of under- or overestimated data occur in the map.

With geographically weighted regression we investigated whether the relationships between growth data and explanatory variables are intrinsically different across space or not. Significantly varying values of intercept provide insight into the effects of potential candidate variables which were omitted from the regression analyses. A significant variation of negative or positive coefficients of explanatory variable indicates an increasing influence of explanatory variable on the dependent variable.
Running the geographically weighted regression with sensitivity data, for the intercept and coefficient of mean precipitation significant spatial variations were computed.

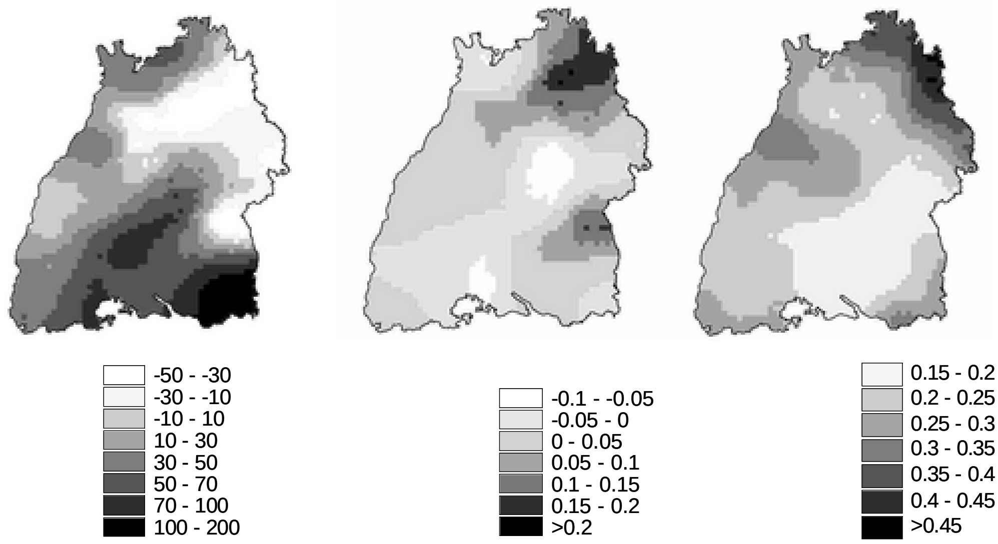
Figure 5 shows the spatial distribution of the intercept and the particular relationship of the mean sum of precipitation in the vicinity of the sample points in space.

Fig.5: Spatial variation of intercept (left) and the regression coefficient of the mean precipitation (right). Dependent variable: sensitivity.



The high intercept values that occur mainly in the north-east of the research area suggest a raised mean sensitivity of radial increment in these areas (see also fig.4). The coefficients corresponding to mean precipitation are negative in the whole study area. In the north-east low precipitation is inversely related to high sensitivity whereas the effect of other factors (foliar [Ca]; altitude) remains constant.

Fig.5:: Spatial variation of intercept (left) and the regression coefficients of mean precipitation (centre) and precipitation index (right). Dependent variable: radial increment index.



The results of GWR with the radial increment index (fig.5) suggest strong spatial non-stationarity of intercept and the coefficients of the precipitation index and mean precipitation. The coefficients differ largely in comparison to the coefficients of the global linear regression model.

The spatial pattern of the intercept values is quite similar to the structure of the coefficient of mean precipitation and the structures of global residuals in fig. 3 (upper, right). They represent the variability of the radial increment index much better than global linear regression and leads to a drastically reduction of the error residual. The coefficient of precipitation index indicates an increasing importance in the north-east. Obviously the annual variation in the reference year 1976 led here to stronger effects on radial increment as in other parts of the study area.

4.5 Comparison of estimation results with validation-grid values

For evaluation of the estimation results the measured values of the validation grid (see fig.2) were compared with the predicted values estimated by ordinary kriging, linear regression and geographically weighted regression. The residuals from measured and predicted values are presented in scatter-plots (fig.6). Since measured and predicted values are theoretically independent, this cloud should have no preferential shape. All tested models have the tendency to underestimate low values for sensitivity as well as the radial increment index and to overestimate higher values. The largest part of the values is within the range of +/- 2 standard deviations suggesting that all techniques are acceptable. There are only few values far outside this range. These are caused by outliers in the measured values.

Fig.6: Scatterplots of measured values versus standardised residuals of predicted values. Estimations of sensitivity (left) and radial increment index (right)in a validation grid.

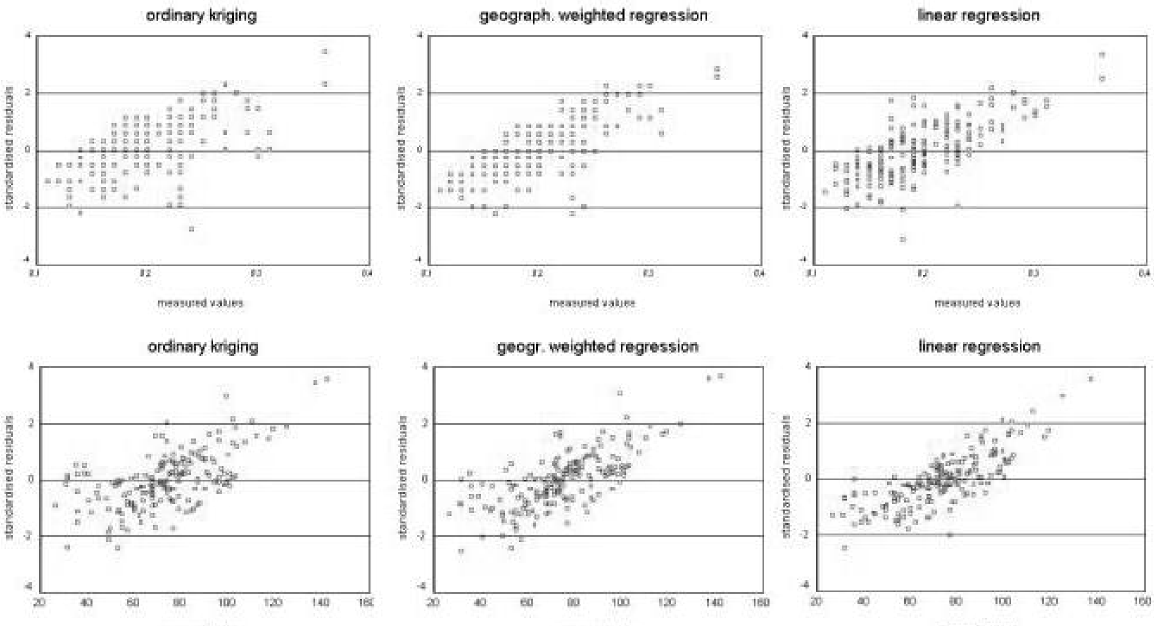


Table 3 shows the normalised residual index produced by each of the applied estimation methods. The normalized residual index is the ratio of the average unsigned residuals to the measured average of the variable. He can be regarded as a relative measure of the prediction error.

Tab.3: Normalised residual index of estimated values for mean sensitivity and radial increment index calculated with different estimation methods

	Linear Regression	GWR	OK
Sensitivity	0,140	0,136	0,141
Radial increment index	0,136	0,129	0,154

The GWR performs better than the other estimation methods. Larger prediction errors, especially for the radial increment index, are obtained for Ordinary Kriging which only considers information provided by surrounding sample values without taking into account information from explanatory variables. Results are improved with those methods which solely include information of explanatory variables.

5. Conclusions

The presented methods offer suitable tools to detect spatial autocorrelation on various spatial scales. In the example considered here autocorrelation existed for two examined growth-parameters. On the contrary, results were highly structured and the spatial growth patterns became visible with interpolation-techniques. As investigations with other growth-parameters on smaller scale have shown (RIHA et al. 1986; RAMIREZ-MALDONADO, 1988; JOST, 1993; SCOZ, 2002) autocorrelation existed for most of them and the presented methods might also be transferable to detect their potential large-scale noise. The spatial correlation between growth parameters and explanatory variables was analysed by the calculation of the cross-dependencies and examination of regression residuals. The methods are not only mathematical tools to describe the spatial variations, the analyses also sufficiently explain the causes. The gaps of explanation in the used data sets became exposed and it was possible to determine, which influence factors were important and which were omitted.

Especially the geographically weighted regression was a useful instrument to find and quantitatively describe large scale variations of forest growth parameters. Validation sample values were estimated with sufficient accuracy and the spatial variation of residuals were significantly reduced. For certain questions of large scale forest monitoring inventories application of these methods can lead to a reduced costs due to a reduction of sample point density, and by estimating growth values only with explanatory variables. The more detailed analysis (results not presented in this paper), large scale patterns did not change significantly by adding the validation data set to the model data set and the prediction error was reduced only moderately. The higher density of data samples is in this case not justified by a corresponding improvement of the results on large scale. Small-scale patterns were explained more accurately with higher sample density, however a clustered and stratified inventory design could produce better results with lower costs in this case.

In our example, variations of precipitation and altitude (resp. temperature) were the main drivers for the variations of radial increment on a large scale. The precipitation had a decisive effect on both of the investigated parameters, whereas small-scale site factors like hillside, aspect or differences in the soil-type were of secondary importance. The foliar nutrient level (here calcium) is significantly related to variations on smaller scale.

In the context of increasing global warming and the continuation of the recent trends in climatic conditions in central Europe, characterised by an increase in summer temperatures and a decrease in summer precipitation, this will lead to higher drought risk for the investigated tree species *Picea abies.* The observed higher sensitivity of radial growth of *Picea abies* indicates that, on selected sites in the study area, the inter-annual variability of precipitation, but especially the occurrence of drought periods, already had a strong impact on growth on a large scale in the past.

Acknowledgement

This study has been financially supported by Ministry for Natural Resources Baden-Württemberg. We gratefully thank the Forest Research Station Baden-Württemberg for providing the growth, nutrition and soil data and the German Weather Service for providing the climate data.

5. Literature

ANSELIN , L. (1995): Local indicators of spatial association - LISA". Geographical Analysis. 27, No. 2; p. 93-115.

ASPINALL, R.J. (1999): GIS and spatial analysis for ecological modelling. In: GOMEZ-HERNANDEZ,J. & SOARES,A. (EDS.): Geostatistics for Environmental Applications: Geoenv II. Kluwer Academic Publishers, Dordrecht; p.377-396

BRUNSDON, C., FOTHERINGHAM, A.S. & CHARLTON, M.E. (1996): Geographically weighted regression – a method for exploring spatial nonstationarity, Geographical Analysis, 28 (4), p.281-298.

CLIFF,A.D. & ORD,J.K. (1981): Spatial Processes: Models, Applications. Pion Limited, London: 266 p.

COOK, E.R. & KAIRIUKSTIS, L.A. (EDS.) (1990): Methods of Dendrochronology: Applications in the Environmental Sciences. Kluwer Academic Publishers, Dordrecht: 394 p.

CRESSIE, N. A. C. (1993): Statistics for Spatial Data. Rev. ed.. - New York : Wiley, 1993. - XX, 900p.

EVERS, F.H., SCHÖPFER,W. (1988): Darstellung der Ernährungs- und Belastungsverhältnisse der Fichte. Ergebnisse der Belastungsinventur Baden-Württemberg 1983. Allgemeine Forst- und Jagdzeitung. 1988, 159: 8, p.146-154

FORTIN, M.-J. (1999): Spatial statistics in landscape ecology. In: Klopatek, J.M. and R.H. Gardner (eds). Landscape Ecological Analysis. Issues and Applications. Springler-Verlag; p. 253-279.

FOTHERINGHAM, A.S.; BRUNSDON, C.; CHARLTON, M. (2000): Quantitative Geography: Perspectives on Spatial Data Analysis. London : Sage, 270p.

FOTHERINGHAM, A.S., BRUNSDON, C., CHARLTON, M.E., (2002), Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Chichester: Wiley.

GOOVAERTS, P. (1997): Geostatistics for Natural Resources Evaluation. New York ; Oxford : Oxford University Press, 483p.

HILDEBRAND,E.E.; SCHÖPFER,W. (1993): Ergebnisse der Belastungsinventur Baden-Württemberg 1988. Ernährung und Belastung von Fichte und Tanne. Dokumentation. Mitteilungen der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg. 1993, No. 172, 157p.

ISAAKS, E. H., SRIVASTAVA, R.M. (1989): Applied Geostatistics. New York : Oxford Univ. Pr., 1989. - XIX, 561p.

JOST,A. (1993): Geostatistische Analyse des Stichprobenfehlers systematischer Stichproben. Mitteilungen der Abt. für Forstliche Biometrie 93-1; Univ. Freiburg; 113p.

LAFFAN, S.W.(1999): Spatially assessing model error using geographically weighted regression.
Proceedings of the 4th International Conference on GeoComputation Mary Washington College Fredericksburg, Virginia, USA 25 - 28 July 1999. Published on CD-ROM. Produced by: J.Diaz, R.Tynes, D.Caldwell, J.Ehlen. Publisher: "GeoComputation CD-ROM". ISBN 0-9533477-1-0

MANDALLAZ, D. (1992): Geostatistical Methods for Combined Forest Inventories. Proceedings of the IUFRO Centennial Meetings. Berlin/Eberswalde 31.8.-6.9.92.

MÜLLER-WESTERMEIER, G., (1999): Numerical production of climatic maps for parameters with significant correlation to topography. Klimastatusbericht: 83-86.

RAMIREZ-MALDONADO, H. (1988): On the Relevance of the Geostatistical Theory and Methods to Forest Inventory Problems. Ph.D.thesis. Athens. University of Georgia. 163p.

RIHA, S.J.; SENESAC, G.; PALLANT, R. (1986): Effects of forest vegetation on spatial variability of surface mineral soil, pH, soluble aluminium and carbon. Water, Air and Soil Pollution 31 (3-4), p.929-940

ROGERSON, P. A. (2001): Statistical Methods for Geography. London : Sage, 236p.

ROSSI,R.E., MULLA,D.J., JOURNEL,A.G. & FRANZ,E.H. (1992): Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62: 277-314.

SCHWEINGRUBER, F.H.(1988): Tree Rings : Basics and Applications of Dendrochronology. 1. Engl. ed.. - Dordrecht : Reidel, 1988. - 276 p.

SCOZ, R. (2002): Weiterentwicklungen der Inventurmethodik für Plantagen. Diss. Univ. Freiburg; Freiburger Dokumentenserver; URL: http://www.freidok.uni-freiburg.de/volltexte/474

WEBSTER, R.; OLIVER, M. (2001): Geostatistics for Environmental Scientists. Chichester: Wiley, 2001. - XI, 271p.

WACKERNAGEL, H., (2003): Multivariate Geostatistics: An Introduction with Applications. 3. ed. Berlin;Heidelberg: Springer, 387p.