Allgemeines Statistisches Archiv 0, 0–18
 ISSN 0002-6018

An investigation on the spatial correlations for relative purchasing power in Baden–Württemberg

By S. Eckel, F. Fleischer, P. Grabarnik and V. Schmidt *

Summary: The relative purchasing power, i.e. the purchasing power per inhabitant is one of the key characteristics for businesses deciding on site selection. Apart from that it also plays a major role in regional planning, pricing policy and market research. In this study we investigate the spatial correlations for relative purchasing power of the townships in Baden–Württemberg. In particular, changes in relative purchasing power are analysed for three different time intervals, 1987–1993, 1993–1998 and 1998–2004, by means of distance–dependent characteristics like the mark–correlation function, the Simpson indices $\alpha(r)$ and $\beta(r)$ and by tests on random labelling. It is shown that there are positive correlations for small distances between different townships but that these positive correlations are becoming weaker over the years until they are almost non–existent in the sense that hypotheses of random labelling are no longer rejected. A conclusion from this loss of spatial correlations with time is that the relative purchasing power might become more and more purely random. This means that the relative purchasing power in a township is less and less influenced by the relative purchasing power of townships nearby. We further analysed these changes in the regions Bodensee–Oberschwaben and Stuttgart to compare the development of the relative purchasing power in an urban and a rural environment.

Keywords: Purchasing power, spatial correlation, (distance–dependent) Simpson index, mark–correlation function, random labelling, point process modeling. JEL R11, R12.

1. Introduction

One of the key characteristics that is considered with respect to the site selection for businesses is given by the so–called purchasing power in a township. It is defined here as the sum of the wages, profits and income from assets and transfers minus the direct taxes, the social security contributions, transfers, expenditures for habitation and net savings. So, the purchasing power can be considered as a measure for the free net income that is not dedicated to inevitable expenditures like the rent (for more details see for example [5]). Often, in order to facilitate comparisons, the relative purchasing power, given as the total purchasing power divided by the number of inhabitants, is considered instead of the total purchasing power. Note that the purchasing power in this context does not refer to the purchasing power of the money which is formally equal to the reciprocal of the level of prices but as the sum of funds that are available for free consumption.

The purchasing power plays a significant role not only in site selection but also in sales planning, pricing policy or market research. The meaningfulness of the purchasing power is not restricted to businessmen but also applies

Received: / Revised:

* This analysis was done in close cooperation with W. Brachat–Schwarz and W. Walla of the Statistical Office of Baden–Württemberg. S. E. is supported by a grant of the graduate college 1100 at Ulm University.

to people that are involved in regional planning and research as well as to politicians. All these groups use the purchasing power as an indicator for the quality of location and the wealth of a region. Note that in general the purchasing power does not indicate how much money a single person is able to spend but it provides inference about the purchasing potential of a community.

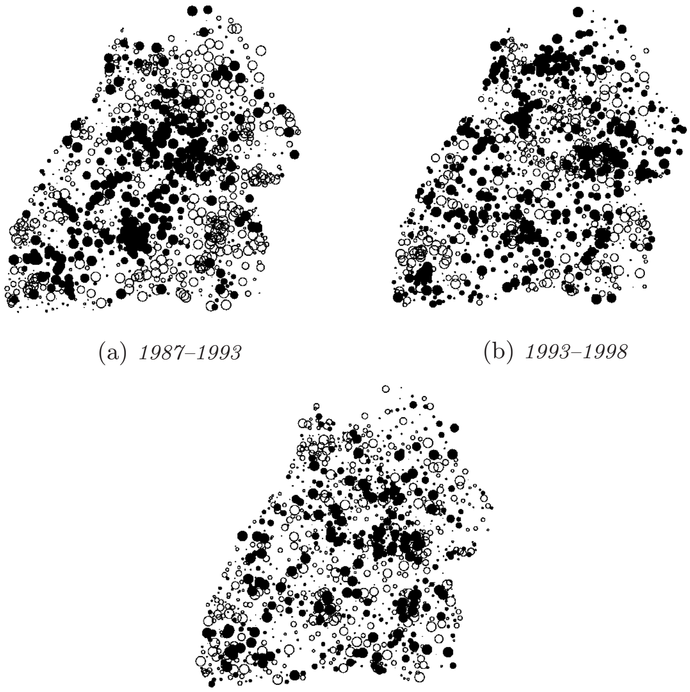
Based on the reasons mentioned above, the Statistical Office of Baden–Württemberg (Germany) is recording the total as well as the relative purchasing power of 1111 townships in Baden–Württemberg, where data are available for four different years (1987, 1993, 1998, 2004) with approximately the same time span between adjacent measurements. This study investigates the behavior of the relative purchasing power over time. Points of interest in this context are the change of the relative purchasing power over time and the spatial correlations associated with this change. For example, if a township has a distinct increase in relative purchasing power for a specific time interval, is it more likely that a township nearby has also distinct increase or is it more probable that the relative purchasing power is distinctly declining for townships that are near to the first. A second interesting topic is the question whether such effects of spatial correlations between different townships are becoming stronger or weaker over the years. This might become important since if the spatial correlations are vanishing then after some years also the relative purchasing power in a township itself might become more or less completely random in the sense that it is not influenced by the relative purchasing power in other townships nearby. So, in summary, this study addresses two questions related to relative purchasing power in Baden–Württemberg. First, is there a significant spatial correlation for the change of purchasing power with respect to different time intervals, as shown in Fig. 1 and second, if the existence of such a spatial correlation is assumed, has it become stronger or weaker over the years. Moreover, one overriding situation is that these correlations might differ between urban and rural regions. For this purpose two regions of Baden–Württemberg, region Stuttgart (population 2.7 millions) and region Bodensee–Oberschwaben (population 1.8 millions), have been analysed separately. We chose these two regions, since the Stuttgart area is a typical representant of an urban region in Baden–Württemberg and Bodensee–Oberschwaben represents a typical rural region. Furthermore we expected stable results, since the number of observations and the population size of these two selected regions were sufficient.

In order to investigate these questions different tools from spatial statistics are applied, in particular methods for spatial point process modeling are used (cmp. [1]). A novelty of our approach is that these methods which are typical for point pattern analysis are applied to this type of real socio–economic data in Baden–Württemberg for the first time. Note that there are also different methodological approaches for spatio–temporal analysis, e.g. temporal GIS and spatio–temporal modelling (cmp. [4], [9]), but here we focus on an explorative analysis using characteristics of marked point processes. First of all the change of the relative purchasing power in different time intervals is investigated using the mark–correlation function (cmp. [8]). Another technique that is used for the detection of spatial correlations and their behavior over time are distance–dependent Simpson indices (cmp. [6]) based on the multivariate pair–correlation function and the multivariate K–function (cmp. [2]). These indices are applied to a discretized version of the differences of relative purchasing power, where the given data is divided into suitable categories (distinct increase, distinct decrease, no distinct change in relative purchasing power). Based on these categories a null hypothesis of random labelling is checked or in other words whether there are any spatial correlations with respect to the change in relative purchasing power is examined. The results for all three methods (mark–correlation function, distance–dependent Simpson indices, tests for random labelling) show evidence that for small distances between townships there is a positive spatial correlation and that the strength of this spatial correlation is decreasing in time meaning that for later years it is almost non–existent. Furthermore by considering the mark–correlation functions for the two regions, these effects mainly seem to derive from the changes in the rural region, since for the urban region no significant correlations can be observed over the years.

2. Data description

The data investigated was provided by the Statistisches Landesamt Baden–Württemberg. It consists of a set of the locations (i.e. the midpoints) for townships in Baden–Württemberg together with the relative purchasing power, i.e. the total purchasing power divided by the number of inhabitants for the years 1987, 1993, 1998 and 2004. In total there are 1111 townships in the observed region.

The data was preprocessed such that the mean relative purchasing power is constant for the four different years. Afterwards differences between two adjacent years of measurement were computed. Note that due to the preprocessing it is assured that the mean difference for adjacent years of measurement in relative purchasing power is equal to zero. For later investigations we considered two different types of characteristics for each location. First of all the differences in relative purchasing power in such time intervals, 1987–1993, 1993–1998, 1998–2004, leading to a continuous mark space for the point pattern. Apart from this we classified the differences in relative purchasing power into three different categories: townships whose relative purchasing power has increased in the time interval, townships that had a decreasing relative purchasing power and townships whose relative purchasing power stayed almost constant. As a criterion for the categorization we considered $1/4$ times the standard deviation σ of the difference in relative purchasing power for the regarded time interval. Since the mean change in relative purchasing power is scaled to 0, a township that has a change in relative purchasing power that is less than $-\sigma/4$ is considered to have a dis-



(a) *1987–1993* (b) *1993–1998*

(c) *1998–2004*

Figure 1. Changes of relative purchasing power of townships in Baden–Württemberg for different time intervals (decrease=empty circle, increase=full circle, size of circle=magnitude of change)

tinct decrease, whereas if the change in relative purchasing power is more than $\sigma/4$ it has a distinct increase the change of relative purchasing power.

In summary, we obtain two different types of marked point patterns, where the locations of the points are fixed in each case. For the first type, the actual differences in relative purchasing power, the mark space is continuous, while the second type represents a discretization into three different possible marks, increase, decrease or almost constant. A visualization of the point pattern together with the borders of the sampling window is displayed in Figure 2.

For the analysis of the regions Bodensee–Oberschwaben and Stuttgart we preprocessed the data analogously to the data of whole Baden–Württemberg. Since in this case we only consider the mark–correlation function, there is no need for a categorization of the marks.

3. Multivariate spatial point processes and Simpson indices

Let $X^{(1)}, ..., X^{(m)}$ be m univariate spatial point processes in $\mathbb{R}^2$. This means that for any $i = 1, ..., m$ and for any sampling window $B \subset \mathbb{R}^2$, the ran-

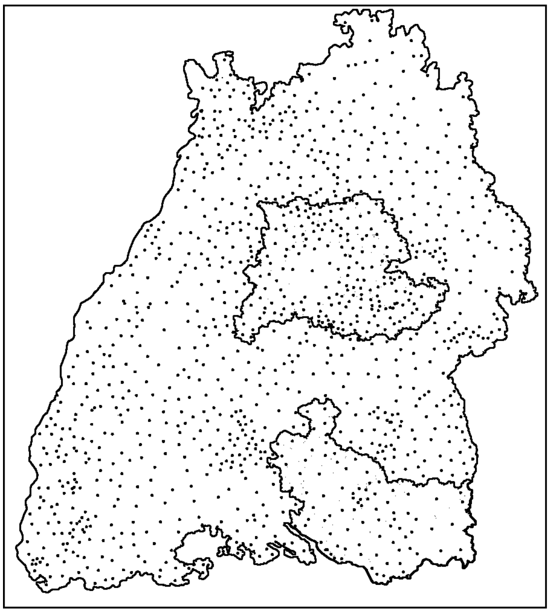


FIGURE 2. Analysed point pattern (lower region=Bodensee–Oberschwaben, upper region=Stuttgart)

dom variable $X^{(i)}(B) = \#\{n : X_n^{(i)} \in B\}$ denotes the number of points $X_n^{(i)}$ of type i observed in the window B. Then, we call $X = \sum_{i=1}^{m} X^{(i)}$ a multivariate point process in $\mathbb{R}^2$ and $X^{(i)}$ a component of X. Note that it is also possible to define X as a marked point process with mark space $\mathbb{M} = \{1, ..., m\}$. In the following we will consider two–dimensional multivariate point processes with $m = 3$ which are considered to be stationary and isotropic, meaning that their distribution is invariant with respect to translations and rotations around the origin.

3.1. INTENSITY. Let $\{X_n\}$ denote the points of a (stationary and isotropic) multivariate point process X in $\mathbb{R}^2$. The intensity measure $\Lambda(B)$ of X which represents the mean number of points in a sampling window B, is given as

$$\Lambda(B) = EX(B) = \lambda|B|, \tag{1}$$

where $X(B) = \#\{n : X_n \in B\}$ denotes the number of points of X in B and $|B|$ means the area of B. The constant λ is called the intensity of X. A natural estimator for λ is given by

$$\widehat{\lambda} = \frac{X(B)}{|B|}. \tag{2}$$

Based on the recommendation in [8], λ^2 has been estimated by

$$\widehat{\lambda^2} = \frac{X(B)(X(B)-1)}{|B|^2}. \tag{3}$$

Note that estimates for the intensity λ_i of the ith component $X^{(i)}$ of X can be obtained in a natural way by replacing X with the relevant $X^{(i)}$ in the estimators above.

3.2. Multivariate K–function. The multivariate K–function $K_{ij}(r)$ for a multivariate point process X and its components $X^{(i)}$ and $X^{(j)}$ is defined such that $\lambda_j K_{ij}(r)$ is the mean number of points of $X^{(j)}$ located in a disc with radius r around a randomly chosen point of $X^{(i)}$. Mathematically, we can define $K_{ij}(r)$ by

$$K_{ij}(r) = \mathbb{E} \sum_{X_n^{(i)} \in B} \frac{X^{(j)}(b(X_n^{(i)}, r) \setminus \{X_n^{(i)}\})}{\lambda_i \lambda_j |B|}, \tag{4}$$

where $b(x, r)$ is the disc with midpoint $x \in \mathbb{R}^2$ and radius $r > 0$. Note that $K_{ij}(r) = K_{ji}(r)$, also if $i \neq j$. Furthermore, since we want to regard only simple point processes we assume that $P(X_n^{(i)} \in X^{(j)}) = 0$ for all n and $i \neq j$, i.e. the probability that two points originating from different components have the same location is 0. In the case $i = j$ it is easy to see that $K_{ii}(r)$ is the univariate K–function for the component $X^{(i)}$. As an estimator for $K_{ij}(r)$ we used

$$\widehat{K_{ij}}(r) = \frac{1}{\widehat{\lambda}_i \widehat{\lambda}_j} \sum_{X_n^{(i)}, X_l^{(j)} \in B : X_n^{(i)} \neq X_l^{(j)}} \frac{\mathbb{1}_{[0,r]}(|X_n^{(i)} - X_l^{(j)}|)}{|B_{X_n^{(i)}} \cap B_{X_l^{(j)}}|}, \tag{5}$$

where $|x|$ is the length of the vector $x \in \mathbb{R}^2$ and where $|B_{X_n^{(i)}} \cap B_{X_l^{(j)}}|$ denotes the area of the intersection of the two shifted sampling windows $B_{X_n^{(i)}}$ and $B_{X_l^{(j)}}$; $B_x = \{x + y : y \in B\}$.

3.3. Multivariate pair–correlation function. In order to define the multivariate, or better cross pair–correlation function we have to define the (bivariate) second factorial moment measure $\alpha_{ij}^{(2)}$ for two components $X^{(i)}$ and $X^{(j)}$ of X first. It is given by

$$\alpha_{ij}^{(2)}(B_1 \times B_2) = \mathbb{E}\Big(\sum_{n,l : X_n^{(i)} \neq X_l^{(j)}} \mathbb{1}_{B_1}(X_n^{(i)}) \mathbb{1}_{B_2}(X_l^{(j)}) \Big) \tag{6}$$

and can often be expressed using a density $\varrho_{ij}^{(2)}$ as follows

$$\alpha_{ij}^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} \varrho_{ij}^{(2)}(x_1, x_2) dx_1 dx_2. \tag{7}$$

The density function $\varrho_{ij}^{(2)}$ is called the (bivariate) second product density. If one takes two balls C_1 and C_2 with infinitesimal volumes dV_1 and dV_2 and midpoints x_1 and x_2 with $x_1 \neq x_2$, respectively, the probability for having in C_1 at least one point of $X^{(i)}$ and for having in C_2 at least one point of $X^{(j)}$ is approximately equal to $\varrho_{ij}^{(2)}(x_1, x_2) dV_1 dV_2$. In the stationary and isotropic case we replace $\varrho_{ij}^{(2)}(x_1, x_2)$ by $\varrho_{ij}^{(2)}(r)$, where $r = |x_1 - x_2|$. As an estimator

$$\widehat{\varrho_{ij}^{(2)}}(r) = \frac{1}{2\pi r} \sum_{X_n^{(i)}, X_l^{(j)} \in B : X_n^{(i)} \neq X_l^{(j)}} \frac{k_h(r - |X_n^{(i)} - X_l^{(j)}|)}{|B_{X_n^{(i)}} \cap B_{X_l^{(j)}}|} \tag{8}$$

has been used, where $k_h(x)$ denotes the Epanechnikov kernel

$$k_h(x) = \frac{3}{4h}(1 - \frac{x^2}{h^2} \mathbb{1}_{(-h,h)}(x)) \tag{9}$$

with bandwidth $h = 0.15\lambda_i^{-\frac{1}{2}}$.

Using the second product density $\varrho_{ij}^{(2)}(r)$ the (bivariate) pair–correlation function $g_{ij}(r)$ can be defined as

$$g_{ij}(r) = \frac{\varrho_{ij}^{(2)}(r)}{\lambda_i \lambda_j}. \tag{10}$$

Note that if $i = j$ we obtain the ordinary univariate pair–correlation function $g_{ii}(r)$ for the process $X^{(i)}$.

3.4. Mark–correlation function for marked point processes. For a stationary and isotropic marked point process $X = \{(X_n, M_n)\}_{n \geq 1}$ we regard the mark–correlation function $\kappa(r)$ given by

$$\kappa(r) = \mathbb{E}_{o,x}(M(o)M(x)), \tag{11}$$

where $\mathbb{E}_{o,x}$ is expectation subject to the conditioning that X has points at the positions o and x, where $r = |x|$. Heuristically, $\kappa(r)$ denotes the mean of the product of marks at two positions that have a distance of r, conditional to the event that there are indeed two points of X at these two locations. In the following, we assume (without loss of generality) that $\mathbb{E}_o M(o) = 0$, i.e., the mean mark of a randomly selected point of X equals zero.

The marked point process $X = \{(X_n, M_n)\}_{n \geq 1}$ is called randomly labelled if the sequences $\{X_n\}$ and $\{M_n\}$ are independent and if $\{M_n\}$ consists of independent and identically distributed random variables. Note that for random labelling of the marks we obtain that $\kappa(r) = 0$ for all $r > 0$, since we assume that the mean mark is 0.

The mark–correlation function can be estimated very similar to the pair–correlation function using the following estimator $\widehat{\kappa}(r)$ given by

$$\widehat{\kappa}(r) = \frac{\sum_{X_n, X_l \in B, n \neq l} \frac{k_h(r - |X_n - X_l|)}{|B_{X_n} \cap B_{X_l}|} M_l M_n}{\sum_{X_n, X_l \in B, n \neq l} \frac{k_h(r - |X_n - X_l|)}{|B_{X_n} \cap B_{X_l}|}}. \tag{12}$$

For the construction of r–wise acceptance intervals of the mark–correlation function $\kappa(r)$ under random labelling, a resampling technique w.r.t. the sampling window B is applied. The observed values of the marks are assigned completely randomly (without repetition) to the given realisation of locations of the point process X. For this set of locations with newly assigned marks the mark–correlation function is estimated at distances $r_i, i = 1, .., s$. The resulting values for each r_i from k (e.g. $k = 100$ for $\alpha = 5\%$) simulations have to be arranged in ascending order. Then the values at ranks $k\alpha$ and $k(1-\alpha)$ constitute the lower and upper bounds of the acceptance interval at distance r_i, respectively.

3.5. Distance–dependent Simpson indices. The diversity index called the Simpson index D (cmp. [7]) is given by

$$D = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2}{\lambda^2}. \tag{13}$$

In the case of random labelling, it can easily be shown that D is the probability that a randomly selected point pair of $\{X_n\}$ has different marks, i.e. the points belong to different components. An estimator $\widehat{D}$ for D is given by

$$\widehat{D} = 1 - \sum_{i=1}^{m} \frac{X^{(i)}(B)(X^{(i)}(B)-1)}{X(B)(X(B)-1)}. \tag{14}$$

A first generalization of D (see also [6]) is given by $\beta(r)$ which is defined by

$$\beta(r) = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2 g_{ii}(r)}{\lambda^2 g(r)}. \tag{15}$$

Note that $\beta(r)$ can be interpreted as the conditional probability that two randomly selected points of X belong to different components, given that the two points have a distance of r. As an estimator for $\beta(r)$ we used

$$\widehat{\beta}(r) = 1 - \sum_{i=1}^{m} \frac{\widehat{\lambda_i^2}\widehat{g_{ii}}(r)}{\widehat{\lambda^2}\widehat{g}(r)}. \tag{16}$$

Another natural extension of D is given by $\alpha(r)$ which is defined by

$$\alpha(r) = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2 K_{ii}(r)}{\lambda^2 K(r)}. \tag{17}$$

The function $\alpha(r)$ can be interpreted as the conditional probability that two randomly selected points of X belong to different components, given that the distance of the two points is smaller or equal r. An estimator $\widehat{\alpha}(r)$ for $\alpha(r)$ is given by

$$\widehat{\alpha}(r) = 1 - \sum_{i=1}^{m} \frac{\widehat{\lambda_i^2}\widehat{K_{ii}}(r)}{\widehat{\lambda^2}\widehat{K}(r)}. \tag{18}$$

An important property of the distance–dependent generalizations $\alpha(r)$ and $\beta(r)$ of the Simpson index D that will be used in the following is that for random labelling we have that $\alpha(r) = \beta(r) = D$ for all $r \geq 0$, i.e. if the marks are purely randomly distributed we obtain that both characteristics become independent of the distance r.

3.6. A simulation based test for random labelling. Consider a marked point process $X = \{(X_i, M_i)\}$ with a discrete mark space $\mathbb{M} = \{1, ..., m\}$. The null hypothesis being investigated in the following is that the points X_i of X are randomly labelled, i.e. the labelling for each point X_i is performed independent of the labels for all other points. In such a case we obtain that

$$K_{11}(r) = K_{22}(r) = ... = K_{mm}(r). \tag{19}$$

Note the subtle difference to a scenario where X consists of m component point processes $X^{(1)}, ..., X^{(m)}$ and where the component point processes are mutually independent. In case of independence it is true that $K_{ij}(r) = \pi r^2$ for any marks $i \neq j$. Obviously the two scenarios of random labelling and independent component point processes are only identical if the component point processes are stationary Poisson point processes that are mutually independent (see also [2]).

Equation (19) can be used in order to construct a simulation based test for random labelling. Note that in the following $m = 3$, but the same test can be performed for arbitrary number of different mark types with only small modifications. For a sequence of distances $\{r_1, ..., r_s\}$ we regard the test statistic

$$T = \sum_{l=1}^{s} \omega_{12}(r_l)(\widehat{K_{11}}(r_l) - \widehat{K_{22}}(r_l))^2 + \omega_{13}(r_l)(\widehat{K_{11}}(r_l) - \widehat{K_{33}}(r_l))^2 \\ + \omega_{23}(r_l)(\widehat{K_{22}}(r_l) - \widehat{K_{33}}(r_l))^2, \tag{20}$$

where $\omega_{ij}(r_l) = (Var(\widehat{K_{ii}}(r_l) - \widehat{K_{jj}}(r_l)))^{-1}$ for $l = 1, ..., s$, $i = 1, ..., 3$ and $i < j \leq 3$.

For the computation of the test statistic k simulations of random labelling of the marks are necessary, i.e., the observed realisation of the marks is assigned k times completely randomly (without repetition) to the given realisation of locations of the point process X. For each set of newly assigned marks the K_{ii}-functions ($i = 1, ..., 3$) are estimated at distances $\{r_1, ..., r_s\}$. For each distance r_l, $l = 1, ..., s$ the variance $\omega_{ij}^{-1}(r_l)$ of the difference $\widehat{K_{ii}}(r_l) - \widehat{K_{jj}}(r_l)$ is estimated from the k simulations. Having determined the weights $\omega_{ij}(r_l)$ for $l = 1, ..., s$, $i = 1, ..., 3$ and $i < j \leq 3$, the values $t^*, t_1, ..., t_k$ of the test statistic T corresponding to the observed realisation and the k simulations, respectively, can be computed. These values $t^*, t_1, ..., t_k$ have to be arranged in ascending order and the null hypothesis is rejected if the rank of t^* becomes too large, i.e., given $k = 999$ and $\alpha = 0.05$, if the rank of t^* is between 951 and 1000.

4. Results

Data analysis was done using the GeoStoch library, which has been developed by the Institute of Applied Information Processing and the Institute of Stochastics of Ulm University ([3], http://www.geostoch.de).

4.1. Mark–correlation function.

Figure 3 displays the estimated mark–correlation function $\widehat{\kappa}(r)$ for the difference in relative purchasing power in the time intervals 1987–1993, 1993–1998 and 1998–2004. Recall that the differences are scaled such that the mean difference is equal to 0. A first observation is that the values for $\widehat{\kappa}(r)$ are above 0 for all $r < 50$ and for all regarded time intervals suggesting a positive correlation in the difference in relative purchasing power between townships of such a distance r. Also, with respect to different time intervals, the estimated values of $\kappa(r)$ are decreasing for later times. This can be understood as a weakening of the positive correlations over the years. In order to investigate these two effects in more detail, we classified the data into three categories (distinct increase, distinct decrease and no distinct changes in the relative purchasing power) and computed different Simpson indices (distance–dependent as well as distance–independent) for it.

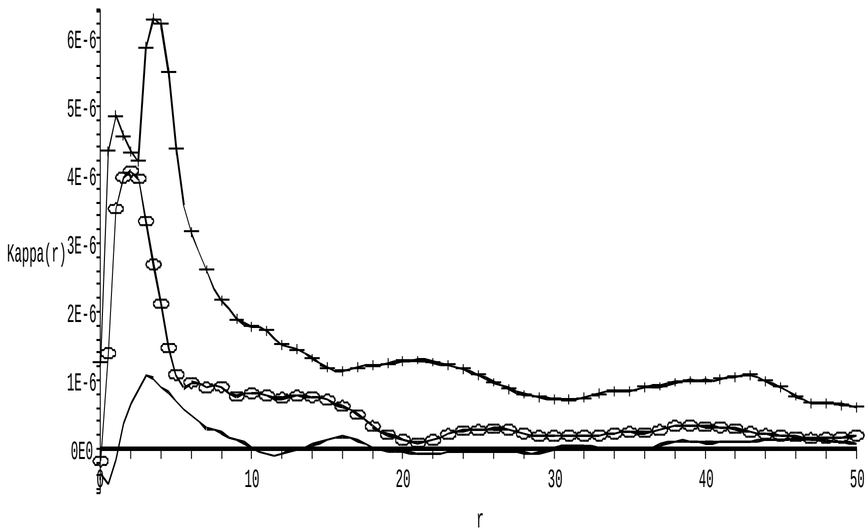


Figure 3. Estimated mark-correlation function $\widehat{\kappa}(r)$ for different time intervals (1987–1993 [×], 1993–1998 [∘], 1998–2004 [–])

TABLE 1. Number of points and estimated distance–independent Simpson index $\widehat{D}$ for each time interval

	1987-1993	1993-1998	1998-2004
Number of points - decrease	440	328	428
- increase	324	476	318
- almost constant	347	307	365
Simpson index $\widehat{D}$	0.6611	0.6535	0.6623

4.2. DISTANCE–INDEPENDENT SIMPSON INDEX. Table 1 gives a summary of results for the number of points and the estimated distance–independent Simpson index $\widehat{D}$ with respect to the three time intervals 1987–1993, 1993–1998 and 1998–2004. Recall that we regard three different types of marks, one for which the relative purchasing power is distinctly increasing in the investigated time interval, one for which the relative purchasing power is distinctly decreasing and a third type where no distinct changes in relative purchasing power occur.

A first conclusion that one can draw is that, although the number of points are slightly varying between the different time intervals, the Simpson index is not able to detect any changes in the correlation structure. Its value is almost stable for all three different cases. Hence as a next step distance–dependent Simpson indices as described in Section 3.5 are regarded.

4.3. DISTANCE–DEPENDENT SIMPSON INDEX. In order to investigate the change of correlation structure in more detail, the two distance–dependent Simpson indices $\alpha(r)$ and $\beta(r)$ are estimated (see Section 3.5). The corresponding graphs are displayed in Figures 4 and 5. Recall that the estimated distance–independent Simpson index $\widehat{D}$ was almost identical for all three time intervals (cmp. Section 4.2). If we look at the indices $\alpha(r)$ and $\beta(r)$ for specific distance values of r we observe that almost everywhere $\widehat{\alpha}(r) < \widehat{D}$ and $\widehat{\beta}(r) < \widehat{D}$ which can be interpreted as a sign for a positive correlation between the different types. We also note that the two distance–dependent Simpson indices seem to approximate the Simpson index $\widehat{D}$ more rapidly with respect to r for later time intervals. This is a first indication that in the course of the years this positive spatial correlation, although still existing, has become much weaker.

4.4. TEST FOR RANDOM LABELLING. In order to investigate the weakening of the spatial correlation for later time intervals in more detail a test for random labelling is performed (see Section 3.6). Recall that under the null hypothesis of random labelling we have that $K_{11}(r) = K_{22}(r) = ... = K_{mm}(r)$ (in our case $m = 3$). A visualisation of the three different estimated $\widehat{K_{ii}}$–functions for the different time intervals regarded is given in Figure 6. These graphs already indicate that for the later time intervals the three

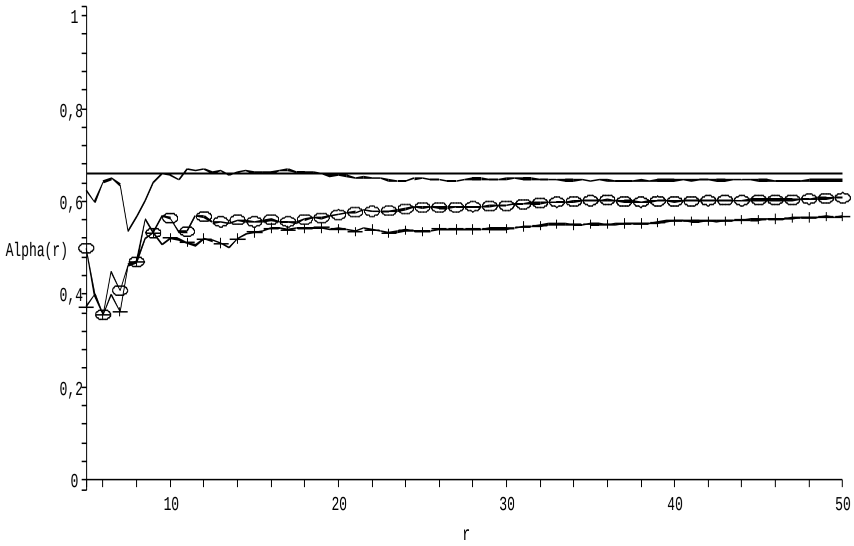


FIGURE 4. Distance–dependent index $\widehat{\alpha}(r)$ and $\widehat{D}$ for different time intervals (1987–1993 [$\times$], 1993–1998 [$\circ$], 1998–2004 [–])

different $\widehat{K_{ii}}$–functions seem to be more similar. The simulation based tests for random labelling confirm this in that for the first time interval (1987–1993) the null hypothesis of random labelling is clearly rejected (rank 1000 in a total size of 1000 each, $p < 0.001$), whereas for the time intervals 1993–1998 and 1998–2004 the same null hypothesis can not be rejected for $\alpha = 0.05$ (rank 932 and rank 597 in a total size of 1000, $p = 0.068$ and $p = 0.403$).

4.5. Spatial correlations in the regions Bodensee–Oberschwaben and Stuttgart. The effects which have been detected so far in Baden–Württemberg, namely a positive correlation in the change of the relative purchasing power between neighboring townships and the weakening of such a correlation over the years, may differ for rural and urban regions. In order to investigate the spatial correlation for a representant in each case, the mark–correlation function $\kappa(r)$ is estimated for the regions Bodensee–Oberschwaben and Stuttgart, respectively. The corresponding graphs are displayed in Figures 7 and 8.

Furthermore we calculated 90% r-wise acceptance intervals for the mark–correlation functions (see Figures 9 and 10). It can be seen that the spatial correlation is only significant for the region Bodensee–Oberschwaben in the time periods 1987–1993 and 1993–1998. This finding corresponds well to the behavior of the spatial correlation in Baden–Württemberg. For the region Stuttgart it is visible that there are no significant spatial correla-

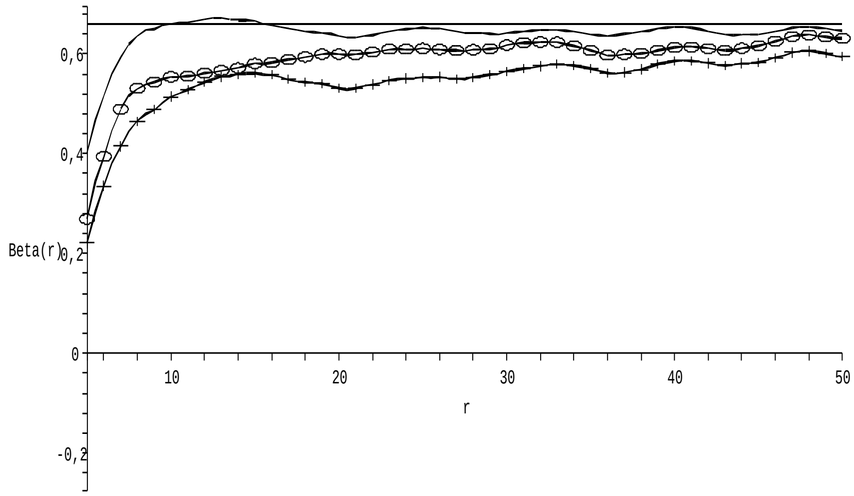


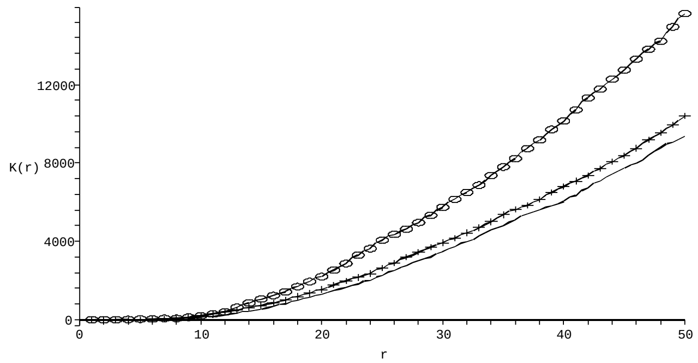
FIGURE 5. Distance–dependent index $\widehat{\beta}(r)$ and $\widehat{D}$ for different time intervals (1987–1993 [$\times$], 1993–1998 [$\circ$], 1998–2004 [–])

tions between neighboring townships and thus no significant development of the relative purchasing power in urban regions. We can conclude that the spatial correlation of the change in the relative purchasing power in Baden–Württemberg derives from the spatial correlation in rural regions.

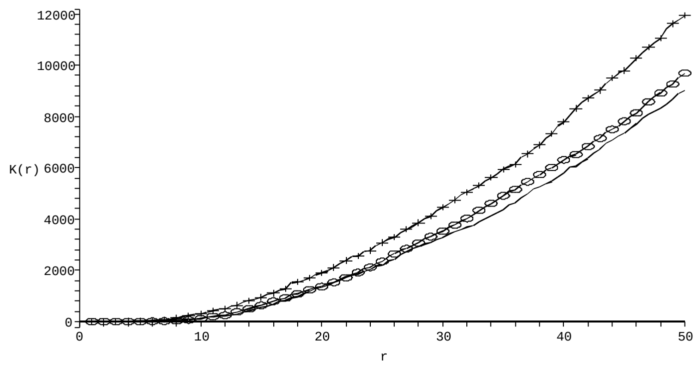
5. Discussion

The main focus of this study is on the detection and the temporal development of spatial correlations with respect to the change in relative purchasing power for townships in Baden–Württemberg. Here we applied three different methods from spatial statistics: the mark–correlation function $\kappa(r)$, distance–dependent Simpson indices $\alpha(r)$ and $\beta(r)$ as well as simulation–based tests on the hypothesis of random labelling for the marks.

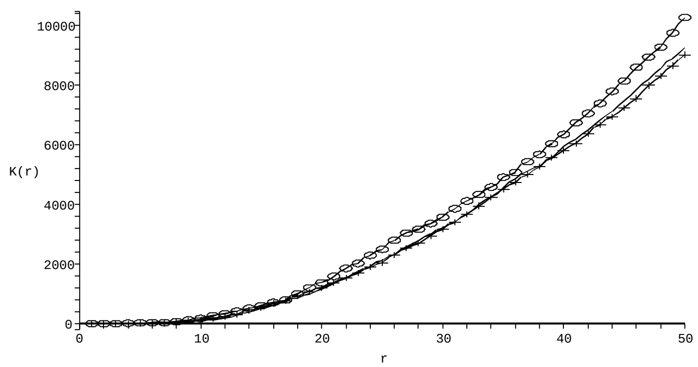
All the results shown in Section 4 provide clear indications that (at least) for small interpoint distances there is a positive correlation given with respect to the change in relative purchasing power and that this positive correlation is becoming weaker if later time intervals are considered (see Figure 3). Here, all values of $\kappa(r)$ are above 0, a sign for a positive correlation due to the fact that the mean change of relative purchasing power is scaled to 0. Furthermore, it can be deduced that these positive correlations are becoming weaker since the functions $\widehat{\kappa}(r)$ are decreasing towards 0 for later time intervals and constant interpoint distance. Investigations for the distance–dependent Simpson indices applied to categorized data (distinct increase, distinct decrease and no distinct change in relative purchas-



(a) *1987–1993*



(b) *1993–1998*



(c) *1998–2004*

FIGURE 6. Estimated $\widehat{K_{ii}}$–functions for different time intervals ($\widehat{K_{11}}$ [$\times$], $\widehat{K_{22}}$ [$\circ$], $\widehat{K_{33}}$ [–])

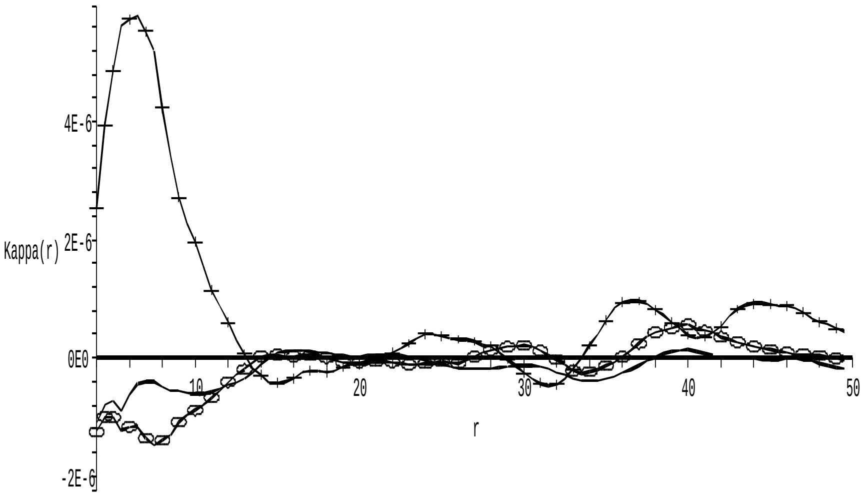


FIGURE 7. Estimated mark-correlation function $\widehat{\kappa}(r)$ for different time intervals (1987–1993 [×], 1993–1998 [∘], 1998–2004 [–]) for the region Bodensee–Oberschwaben

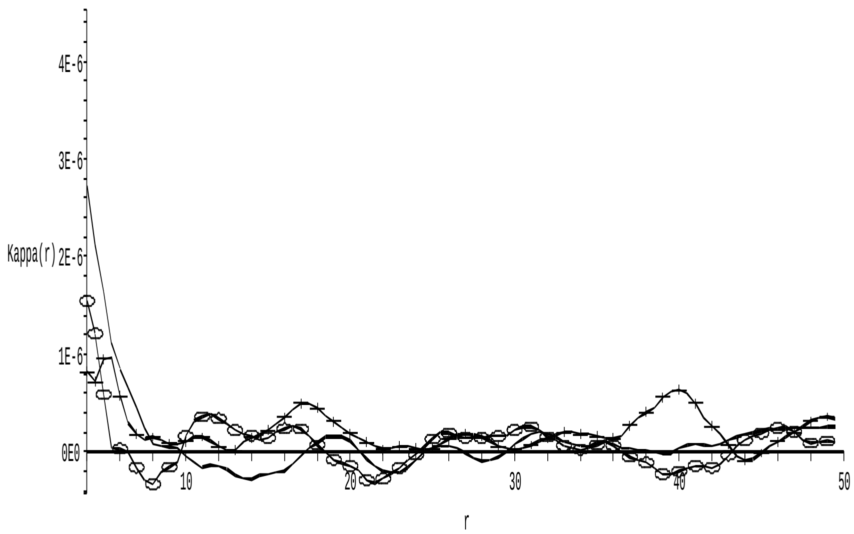
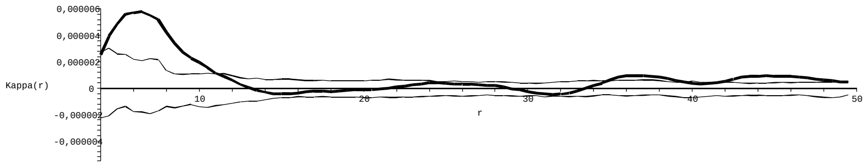
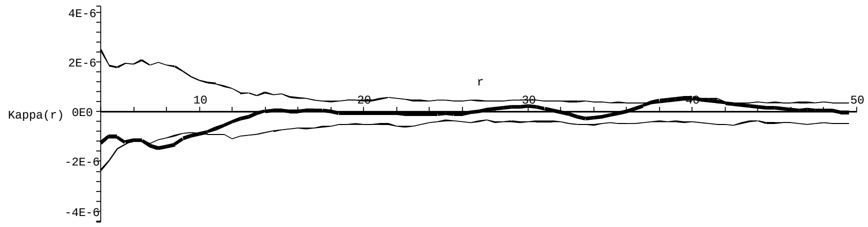


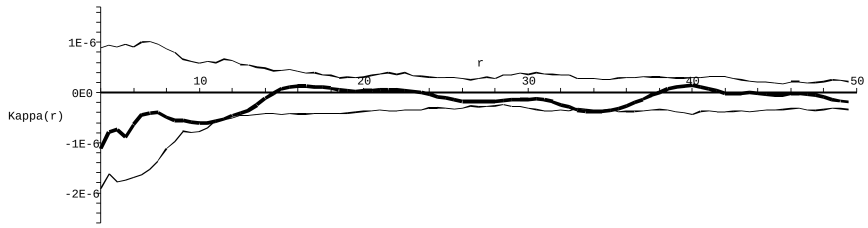
FIGURE 8. Estimated mark-correlation function $\widehat{\kappa}(r)$ for different time intervals (1987–1993 [×], 1993–1998 [∘], 1998–2004 [–]) for the region Stuttgart



(a) *1987–1993*



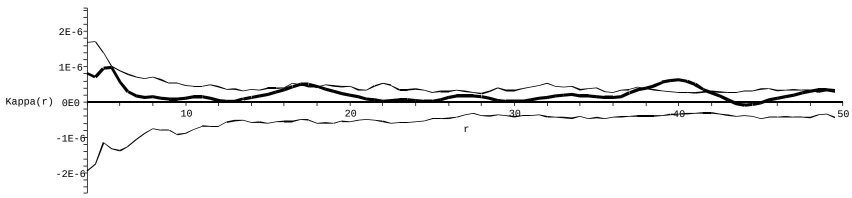
(b) *1993–1998*



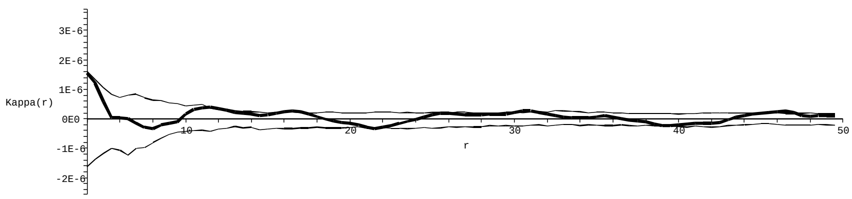
(c) *1998–2004*

FIGURE 9. r-wise acceptance intervals for the mark–correlation function $\kappa(r)$ of different time intervals for region Bodensee–Oberschwaben

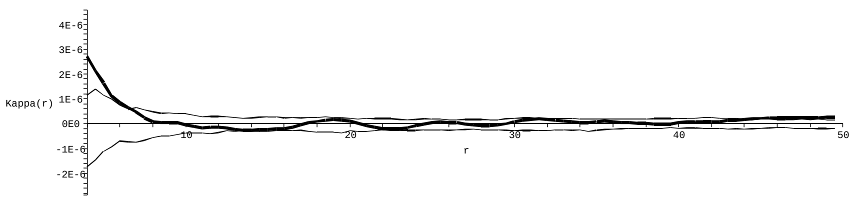
ing power) are confirming the results for the mark–correlation function. Here for small distance values r the estimated functions $\widehat{\alpha}(r)$ and $\widehat{\beta}(r)$ are running below the estimate of the distance–independent Simpson index D which means that for small neighborhoods there is a tendency for pairs of the same type (distinct increase, distinct decrease or no distinct change in relative purchasing power). Regarding the evolution over time, the distance between the value of $\widehat{D}$ and the value for the two functions $\widehat{\alpha}(r)$ and $\widehat{\beta}(r)$ decreases which can be interpreted as a weakening of the spatial correlation. Similar observations hold for the simulation based tests on random labelling. Here, for the first time interval (1987–1993) we show that the hypothesis of random labelling of the marks is clearly rejected ($p < 0.001$), i.e. that there must be a correlation between different labels. For the second time interval (1993–1998) the hypothesis is not rejected for a significance level of $\alpha = 0.05$, but the p–value is still quite small ($p = 0.068$). This might be a hint that there are still spatial correlations but that they are not as



(a) *1987–1993*



(b) *1993–1998*



(c) *1998–2004*

FIGURE 10. r-wise acceptance intervals for the mark–correlation function $\kappa(r)$ of different time intervals for region Stuttgart

strong as for the earlier interval. For the last time interval (1998–2004) the hypothesis is clearly not rejected ($p = 0.403$).

The observation of positive spatial correlations in Baden–Württemberg might not be too surprising. Reasons for this effect might be first of all that often inhabitants of a township are working in a neighboring township and vice versa, thereby positively correlating the change of relative purchasing power for these towns. Another fact that could cause a positive correlation for neighboring townships is that often companies working in a similar sector are located in a relative small distance to each other, for example with regard to metal processing or automobile industry. Hence a boom (or a crisis) in such a sector causes a boom (or crisis) for a whole region, not only for a specific township. A similar argument might hold with respect to the vertical process chain, meaning suppliers and processors for example in the automobile industry.

A behavior that is more interesting and surprising is that the effect of positive correlation for the change in relative purchasing power has become much weaker in later years and that in the meantime this positive correlation is almost non–existent in the sense that tests for random labelling are not rejected. The weakening in the positive correlations of the relative purchasing power might be explained by the increased mobility of the population which causes a larger diversity in the type and the location of work for the population of a township. Also, with respect to companies of a similar sector, mainly of the tertiary sector, in later years there is more and more diversification, meaning that they are more uniformly spread over the state and not that much concentrated as in earlier years.

Furthermore by analysing the spatial correlation of the relative purchasing power in urban and rural regions of Baden–Württemberg, it turns out that the effects for the spatial correlations which appear in rural regions mainly contribute to the effects for the spatial correlations in whole Baden–Württemberg. The same qualitative effects which can be stated for the change of relative purchasing power in Baden–Württemberg as a whole can be observed regarding only rural townships, where lowering of spatial correlations in whole Baden–Württemberg is caused by the lack of such spatial correlations in urban regions. A reason for the absence of spatial correlations between neighboring townships in urban regions might be that the increase of mobility and diversification mentioned above has taken place earlier for urban regions than for rural regions.

A consequence of the results presented in this study might be that if the spatial correlations in the change of relative purchasing power are becoming weaker, the relative purchasing power itself will become more and more random, in the sense that spatial correlations are also weakening for the relative purchasing power itself. This loss of spatial correlations could mean that there will be a large variability with respect to relative purchasing power for neighboring townships in a region. In conclusion this means that businesses as well as politicians have to consider very detailed information with respect to their decisions on a specific township or region, for example regarding the choice of a new location or the distribution of subsidies.

References

1. Baddeley, A., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (eds.): Case studies in spatial point process modeling. Lecture Notes in Statistics 185. New York: Springer 2006
2. Diggle, P. J.: Statistical analysis of spatial point patterns. London: Arnold 2003
3. Mayer, J., Schmidt, V., Schweiggert, F.: A unified simulation framework for spatial stochastic models. Simulation Modelling Practice and Theory 12, 307–326 (2004)
4. Raper, J. F.: Multidimensional geographic information science. London: Taylor & Francis 2000
5. Schwarck, C.: Neue Kaufkraftdaten für die Gemeinden in Baden–Württemberg. Statistisches Monatsheft Baden–Württemberg 5 (2006)
6. Shimatani, K.: Multivariate point processes and spatial variation of species diversity. Forest Ecology and Management 142, 215–229 (2001)

7. Simpson, E. H.: Measurement of diversity. Nature 688, 163 (1949)
8. Stoyan, D., Stoyan, H.: Fractals, random shapes and point fields. Chichester: John Wiley & Sons 1994
9. Worboys, M. F.: A unified model for spatial and temporal information. The Computer Journal 37, 26–34 (1994)

Stefanie Eckel
Institute of Stochastics
Ulm University
Helmholtzstraße 18, 89069 Ulm
Germany
stefanie.eckel@uni-ulm.de

Frank Fleischer
Medical Data Services/Biostatistics
Boehringer Ingelheim Pharma GmbH & Co KG
Birkendorfer Strasse 65, 88397 Biberach a. d. R.
Germany
frank.fleischer@boehringer-ingelheim.com

Pavel Grabarnik
Institute of Physico–chemical and Biological Problems of Soil Science
Russian Academy of Science
Pushchino
Russia
pgrabarnik@rambler.ru

Volker Schmidt
Institute of Stochastics
Ulm University
Helmholtzstraße 18, 89069 Ulm
Germany
volker.schmidt@uni-ulm.de