Spatial analysis of earthworm biodiversity at the regional scale

M. Joschko [a,*], C.A. Fox [b], P. Lentzsch [a], J. Kiesel [c], W. Hierold [d], S. Krück [a], J. Timmer [e]

[a] Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF), Institut für Landschaftsstoffdynamik, Eberswalder Str. 84, D-15374 Müncheberg, Brandenburg, Germany
[b] Agriculture and Agri-Food Canada, Greenhouse and Processing Crops Research Centre, 2585 County Road 20, East, Harrow, Ont., Canada N0R 1G0
[c] Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF), Institut für Landschaftssystemanalyse, Eberswalder Str. 84, D-15374 Müncheberg, Germany
[d] Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF), Abteilung für Landschaftsinformationssysteme, Eberswalder Str. 84, D-15374 Müncheberg, Germany
[e] Freiburger Zentrum für Datenanalyse und Modellbildung, Universität Freiburg, Eckerstr. 1, D-79104 Freiburg, Germany

Abstract

Information about the spatial patterns of soil biodiversity at the regional scale is limited though required, e.g. for understanding regional scale effects of biodiversity on ecosystem processes. This study was conducted to determine whether earthworm biodiversity parameters display spatial patterns at the regional scale and whether they are related to soil properties. A transect study was carried out in 1996, 1997 and 1998 on tilled, sandy to loamy soils in the morainal landscape of northeast Germany. One soil block (50 cm × 50 cm × 20 cm) was sampled at 50 sites for earthworm number, biomass and species composition.

Total earthworm abundance and biomass fluctuated randomly along the transect in all years studied. However, reproducible spatial patterns of earthworm biodiversity were found along the transect, in respect to the occurrence of single earthworm species and in respect to species number and species composition. Species number and species composition displayed increasing trends from south to north, in all years studied. This pattern corresponded to increasing trends of soil properties like pH, soil organic carbon, total nitrogen (N_t) and clay content. By applying state-space analysis to the 1996 data set, the number of earthworm species along the transect could be estimated with parameters pH and total nitrogen.

The results show that spatial patterns of earthworm biodiversity may be identified at the regional scale. State-space analysis was shown to be an effective tool for detecting relationships between earthworm biodiversity and soil properties at this scale.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Earthworms; Species number; Species composition; Sandy soil; Management; State-space analysis

1. Introduction

A challenge for soil biodiversity research is the analysis of biodiversity patterns and their relation to environmental factors at regional scales. For most soil biota groups, regional scale information on species composition and abundance is scarce. Such information is required to enable predictions about potential impacts on soil biota populations

as influenced by changes in land use and management, to evaluate the extent of contributions by soil biota to soil processes and to identify interrelationships with soil attributes (Fox, 2003; Fox and MacDonald, 2003; Fox et al., 2004).

One main reason responsible for the absence of information about regional scale biodiversity is the lack of adequate methods for sampling and analyzing data at this dimension (Groffman et al., 1992). For analysis of soil faunal biodiversity and its relation to soil properties, data are often sampled at "representative" sites and analyzed by

* Corresponding author. Tel.: +49 33432 82 254; fax: +49 33432 82 330.
E-mail address: mjoschko@zalf.de (M. Joschko).

means of multiple regression analysis (Ekschmitt et al., 2003). However, important information may be lost if locations of observations are not considered: approaches which allow for spatial analyses of data (Nielsen and Wendroth, 2003), may provide a promising methodology and opportunity for evaluating soil biota distribution patterns at regional scales.

An adequate approach for the analysis of spatial patterns is a transect study in which samples are taken in a certain order and with a certain distance between samples. Transect sampling is a common method in landscape ecology (Luck and Wu, 2002). If environmental gradients are studied, this approach can be regarded as "unplanned field experiment" in which study plots are arranged in a certain order along the transect (Luck and Wu, 2002; McDonnell and Pickett, 1990). The analysis of transect data allows the detection of "spatial signatures" along environmental gradients (Luck and Wu, 2002). In addition, quantitative analyses of transect data are possible.

If sets of transect data are autocorrelated and cross-correlated, state-space analysis may be carried out (Nielsen and Wendroth, 2003). This analytical approach has a great potential for analyzing relationships between different spatial series of data, such as soil biological characteristics and physico-chemical soil properties (Nielsen and Alemi, 1989).

State-space analysis was developed from time-series analysis and is based on a sequence of observations in time or space (Nielsen et al., 1983, 1994; Wendroth et al., 1992; Shumway and Stoffer, 2000; Timmer et al., 2000; van Kessel and Wendroth, 2001; Nielsen and Wendroth, 2003). This analysis considers the location of observations and spatial relations between different data series. Unlike other techniques of spatial statistics, e.g. geostatistics, state-space analysis is applicable under a wide array of circumstances and does not request stationarity of data (Nielsen and Wendroth, 2003). With this method of analysis, coincidences between soil properties may be found which may remain undetected when conventional statistical methods are applied. In addition, spatial processes may be described even though many factors are not considered (Nielsen et al., 1994). Whereas most studies are focussed on the field scale, the approach is also suited for regional scale data (Nielsen and Wendroth, 2003).

State-space analysis has been used seldom in soil biology studies. Nielsen and Alemi (1989) showed how state-space analysis may successfully be applied to spatial variability pattern of cotton yield and abundances of a nematode population (*Paratrichodorus*). Sampson (1996) studied earthworm numbers and N distribution patterns in California orchard systems and analyzed them with state-space modelling techniques.

In this paper, results are presented from a 3-year project studying biodiversity parameters along a 150 km landscape transect through northeastern Germany. The study focussed on tilled sandy to loamy soils of Pleistocene origin (Lentzsch et al., 1999). The basic question was whether there are detectable patterns, like gradients, at this scale, whether they are stable, and whether they can be related to physico-chemical properties of the sampled soils (Lentzsch et al., 1999). With respect to microbiological parameters, Wirth (2001) observed stable patterns, i.e. increasing trends that related to soil texture and soil organic matter content.

The aim of this study was to (1) analyze spatial patterns of earthworm biodiversity along the regional scale transect and (2) relate them to soil properties. Parameters of earthworm biodiversity comprised earthworm abundances, species number and species composition. It was unclear whether stable spatial patterns of earthworm biodiversity were to be expected at this scale. Given the ample literature about spatial patterns at the field scale, spatial structures of earthworm populations at the regional scale appeared to be likely (Poier and Richter, 1992; Boag et al., 1994; Rossi et al., 1997; Cannavacciuolo et al., 1998; Nuutinen et al., 1998). According to the literature, connections between earthworm biodiversity and soil properties are strong in regions and landscapes (Michaelsen, 1903; Graff, 1950; Briones et al., 1995; Wills and Abbott, 2003), and were therefore to be expected for the regional scale transect. Since soil properties along the transect were characterized by increasing trends (Lentzsch et al., 1999; Wirth, 2001), this trend should favour the occurrence of spatial patterns of earthworm populations in our study. For the analysis of relationships between earthworm biodiversity and soil properties along the transect state-space analysis was applied.

2. Materials and methods

2.1. Sites

This study was part of a ZALF research group project, entitled: Biological parameters of soils and plant phyllospheres along a transect across the northeastern German lowland: Studies of the spatial continuum and spatial covariance structure with special reference to selected site properties (Lentzsch et al., 1999). The 151 km transect traversed a landscape near Berlin (northeast of Brandenburg and southwest of Pommerania, Germany, Fig. 1) that was glaciated in the latest glacial stage of the Pleistocene (Weichsel glaciation, Liedtke and Marcinek, 1995). The sampling sites were located more or less on a line beginning southeast of Berlin and running north, principally perpendicular to the direction of the ice movement during glaciation. However, parts of the transect ran parallel to the borders of the ice lobes resulting from different states of glaciation. The transect comprised originally 89 sampling points, with a mean distance of 1.7 km. The orientation of the transect was chosen to include the most common soil types under agriculture in the region in approximately the same frequency (Lentzsch et al., 1999).

The transect could be differentiated into two parts. The southern part, characterized by sandy deposits, was not
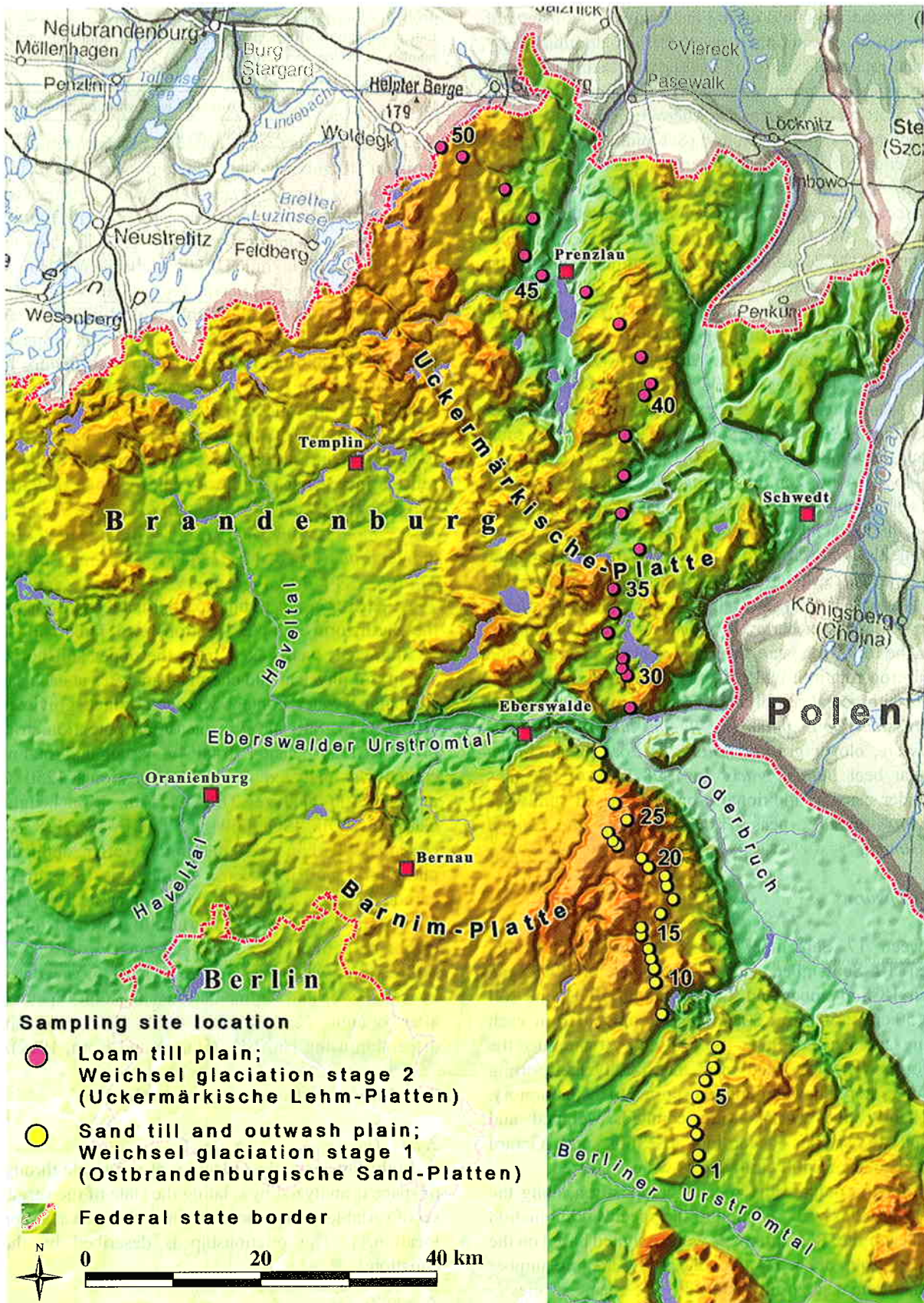
Fig. 1. Location of 50 samples along a transect through the northeastern German agrolandscape in the federal state of Brandenburg near the German-Polish border, crossing two large landscape units (data source: Landesvermessungsamt Brandenburg, 2001).

covered by the ice shield during the Pommeranian stage of the Weichsel glaciation. It is therefore dominated by periglacial material consisting of sandy surface layers over loam to a depth of about a meter. This is in contrast to the northern portion of the transect, where deep, loamy soils of ground moraine origin are found (Schmidt, 1991) (Fig. 1). The direction of sampling thus corresponded to the age of the sediments, with the older deposits in the southern part of the transect being about 5000 years older than those in the northern part and subjected to permafrost. Mean annual precipitation was between 540 and 700 mm, mean annual temperature was between 7.7 and 9.0 °C (Deutscher Wetterdienst, 1999). From the original 89 sites, 50 were selected for earthworm sampling (Fig. 1).

The mean distance between earthworm sampling sites was approximately 3 km (i.e. 2.92 km, range 0.65–7.38 km). There was an increase in distances in the northern half of the transect (Fig. 1). Overall, the soil material ranged from sand to loamy sand with 4–22% clay, 0.47–1.45% C_{org}, and pH 4.2–7.4 in the A-horizon. In summary the soils were mainly variants of Cambi- and Luvisols from glacial and periglacial, but also fluvioglacial, sediments (FAO, 1998).

The fields were cropped with cereals in 1996 (winter wheat *Triticum aestivum* 30%, winter rye *Secale cereale* 38%, triticale X *Triticosecale* 18%, winter barley *Hordeum vulgare* 6%, summer barley *H. vulgare* 4%, corn *Zea mays* 4%). In the following years the range of crops increased due to local crop rotation: in 1997 only 76% of the sites were cropped with cereals; additional crops were winter oilseed rape *Brassica napus*, potatoes *Solanum tuberosum*, lupins *Lupinus sp.*, clover grass *Trifolium sp.*, oat *Avena sativa* and sugar beet *Beta vulgaris*. In 1998 74% of the sites were under cereals; additional crops comprised sunflower *Helianthus annuus*, peas *Pisum sativum* and *Linum usitatissimum*.

2.2. Earthworms

Between 17 September and 5 November in 1996, between 10 September and 3 November in 1997 and between 18 September and 22 October in 1998, one soil block (50 cm × 50 cm × 20 cm) was removed from each sampling site located along the transect. At the site, the earthworms were extracted from the soil by handsorting procedures, stored in water and transported to the laboratory. Subsequently, the worms were counted, weighed and identified to species according to Sims and Gerard (1985). A total of 2400 earthworms were analyzed.

In order to compare the species composition along the transect, an index was assigned to the species combination found at each site. Rank numbers were assigned based on the ecological group of species (Bouché, 1977) and the number of species from each group (Table 1). The anecic species *L. terrestris* was ranked higher than endogeic species populations. Indices ranged from one (one single endogeic species [either *Aporrectodea caliginosa* (SAVIGNY 1826), *A. rosea*

Table 1
Indices for earthworm species composition at the sampling sites, based on number of endogeic (*Aporrectodea* sp., *Allolobophora chlorotica*) and anecic species (*Lumbricus terrestris*)

Index	Composition of earthworm community	
	Endogeic species	Anecic species
1	1	–
2	2	–
3	3	–
4	–	1
5	1	1
6	2	1
7	3	1

(SAVIGNY 1826), or *Allolobophora chlorotica* (SAVIGNY 1826)]) to seven (four species found with three endogeic species and the anecic *Lumbricus terrestris* LINNAEUS 1758) (Table 1). Although this somewhat quantitative index is based on an artificial scale, it adequately quantifies and reflects increasing complexity and ecological demands of the earthworm species community.

2.3. Soil analyses

At each transect point, approximately 1 m distant from the earthworm sample location, two bulk soil samples (1.5 kg each) were taken from the upper soil horizon (0–15 cm) for physical and chemical analyses at each sampling campaign. Water content was determined gravimetrically (105°, 24 h). Soil pH was measured potentiometrically in 0.1 M KCL suspension. Total soil carbon and total nitrogen content was analyzed after dry combustion at 1250 °C using a CNS-2000 analyzer (LECO, Ltd., Mönchengladbach, Germany) (DIN ISO 10694, 1996; DIN ISO 13878, 1998). Carbonate carbon was determined after application of phosphoric acid by electric conductivity measurement of carbon dioxide evolution (Schlichting et al., 1995). Organic C content was calculated as the difference of total C and carbonate carbon. C/N was determined from C_{org} and N_t. Soil texture, which was analyzed in 1996, was determined by wet-sieving and sedimentation with Köhn-Pipette method after organic C destruction with H_2O_2 and chemical dispersion using $Na_4P_2O_7$ (Hartge and Horn, 1992).

2.4. Autoregressive state-space analysis

2.4.1. Theory

In this analysis, the evolution of a variable through time or space is analyzed by relating the state of the variable or a set of variables at a given $location_i$ to the state at the previous $location_{i-1}$. The relationship is described by the state equation:

$$Z_i = \Phi Z_{i-1} + \omega_i, \tag{1}$$

where Φ is a matrix of state coefficients and ω_i is an error term (Nielsen and Alemi, 1989; Wendroth et al., 1992, 1997;

Nielsen and Wendroth, 2003). The state variables represent filtered versions of the observed variables, which are described by the observation equation:

$$Y_i = M_i Z_i + v_i, \tag{2}$$

where Y_i is the observed variable or the variable set, M_i is a matrix of transition coefficients characterizing the relationship between true state and observation, and v_i is an error term quantifying the measurement uncertainty or noise (Nielsen and Alemi, 1989; Wendroth et al., 1992, 1997; Nielsen and Wendroth, 2003; Timm et al., 2003). For integrating of variables and updating between prediction and observation, a Kalman filter with an EM (expectation-maximization) algorithm is used (Nielsen and Alemi, 1989; Nielsen and Wendroth, 2003).

2.4.2. Application

The state-space analysis was based on the book by Nielsen and Wendroth (2003). Details of the analysis may be found there or in Stevenson et al. (2001).

The data were first normalized. Variables manifesting autocorrelation and cross correlations with each other were selected for state-space analysis. These were earthworm species number and species composition in 1996, and five sets of soil properties (% clay, pH, N_t, soil moisture, C/N).

The distribution of variables was analyzed with Q–Q plots (Timmer, 1998). All parameters showed reasonable consistency with normal distribution.

Earthworm biodiversity and soil data from consecutive sites were analyzed as if they were equally spaced and belonging to the same distance class (1–7 km). The direction of statistical analyses followed the sampling pattern from south to north corresponding to the age of the soils, with the 'older' sediments in the south preceding the 'younger' sediments in the northern part of the transect.

For the specific state-space analyses a custom designed program STATE (R. Shumway, provided by courtesy of Ole Wendroth, ZALF, Germany) was used.

To evaluate model quality, the following procedure was carried out (Nielsen and Wendroth, 2003). The results of the linear state-space model were compared with models, in which every other value or two out of three values for the dependent variable (species number, species composition) were omitted. The first value of each series was always considered in the analysis.

2.5. Multiple regression analysis

In addition multiple regression analysis (STATISTICA®) was carried out to determine correlations between earth-
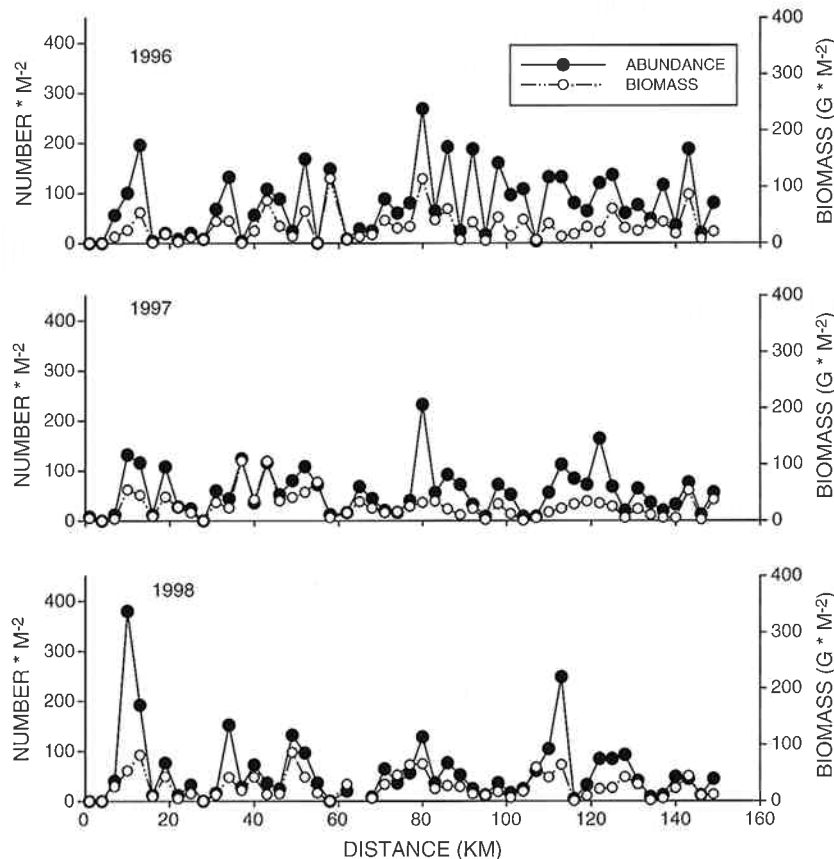

Fig. 2. Total earthworm abundances and biomass along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km).

worm biodiversity (species number, species composition in 1996) and soil variables (% clay, pH, N_t, soil moisture, C/N) (stepwise regression with forward selection). In addition, standard regression analyses were carried out with those variables identified as important predictors in state-space analysis.

3. Results

3.1. Earthworm biodiversity patterns

Earthworm total numbers for the 50 observations are shown according to a calculated mean distance interval (3 km) between sampling points (Fig. 2). Abundances ranged between 0 and 400 individuals m^{-2}, corresponding to 0 and 120 g m^{-2} biomass. Earthworm population density varied considerably from one location to the other. No trend was detected; abundances fluctuated randomly in all years. The dominant species was the endogeic *Aporrectodea caliginosa*, which was found more or less at all sites along the transect (Fig. 3). Similarly *Lumbricus terrestris*, the only anecic earthworm species present, was observed throughout the transect distance. Its occurrence was however, more variable during the three sampling years (Fig. 4). In contrast, the two endogeic species *Allolobophora chlorotica* and *Aporrectodea rosea* were mainly found in the northern part of the transect (Figs. 5 and 6). Thus, two different spatial signatures could be distinguished in the four earthworm species found: one extending along the whole transect, the

Fig. 4. Abundances of *Lumbricus terrestris* along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km).
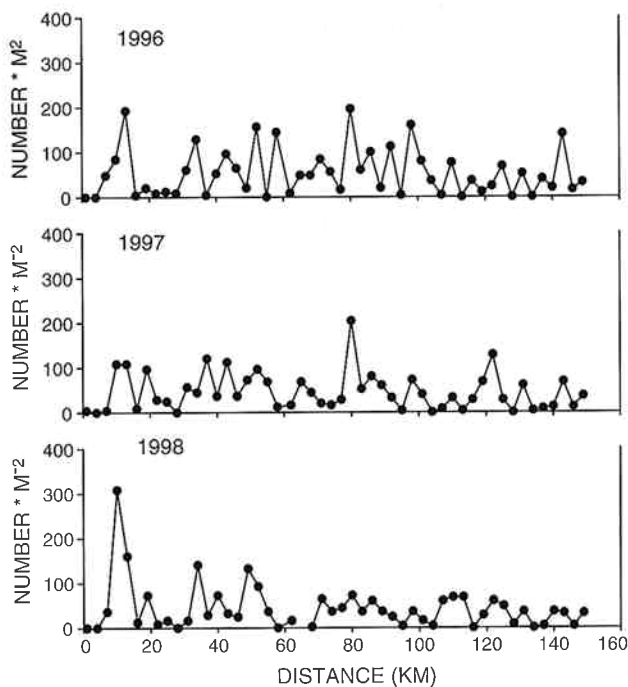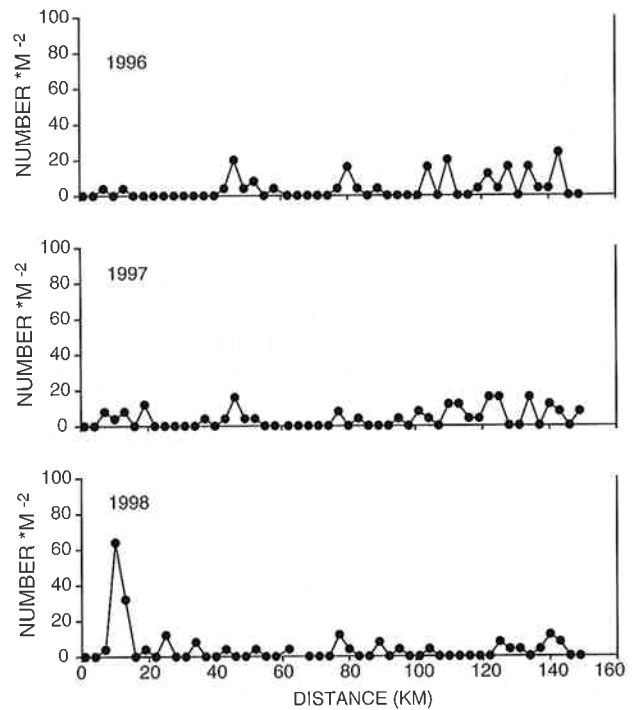

Fig. 3. Abundances of *Aporrectodea caliginosa* along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km).
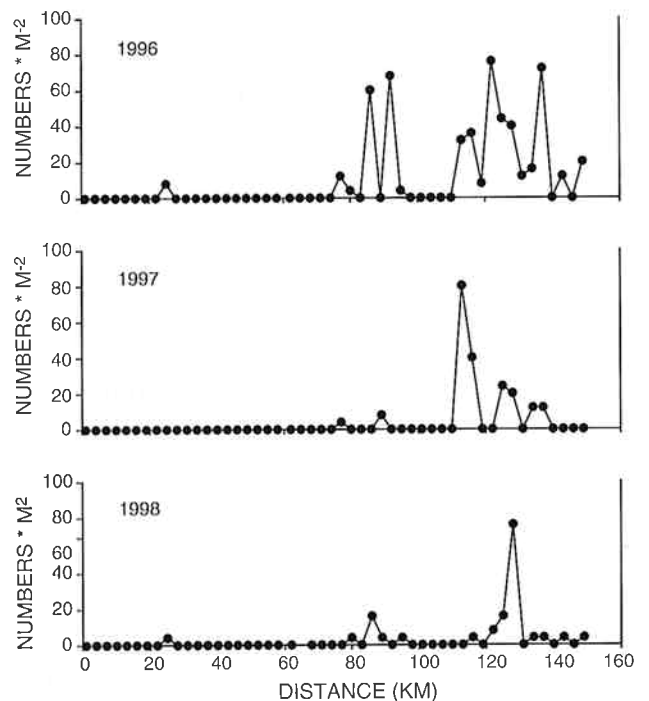

Fig. 5. Abundances of *Allolobophora chlorotica* along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km).
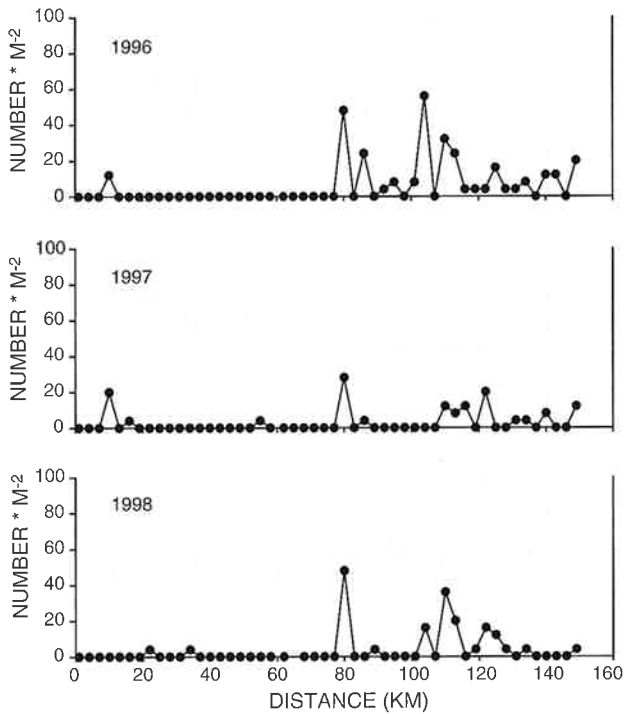
Fig. 6. Abundances of *Aporrectodea rosea* along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km).
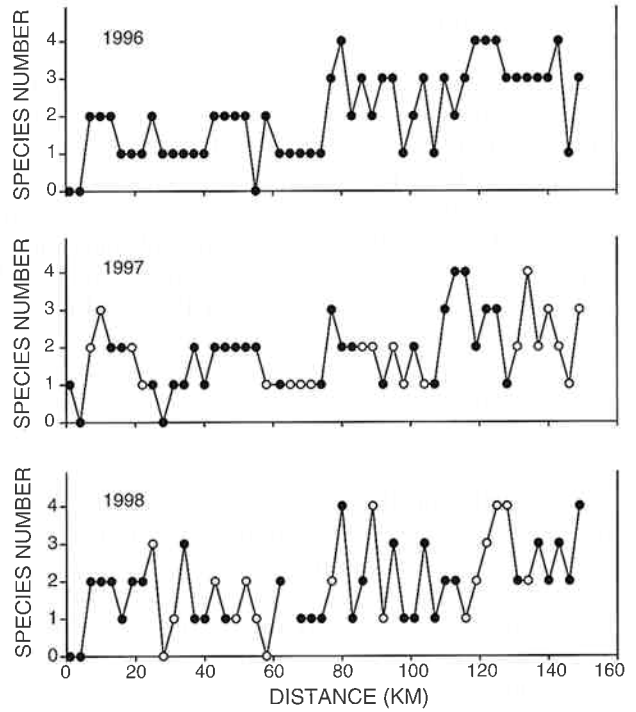

Fig. 7. Earthworm species number along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km); crops at the sites: cereals (black), non-cereals (white).
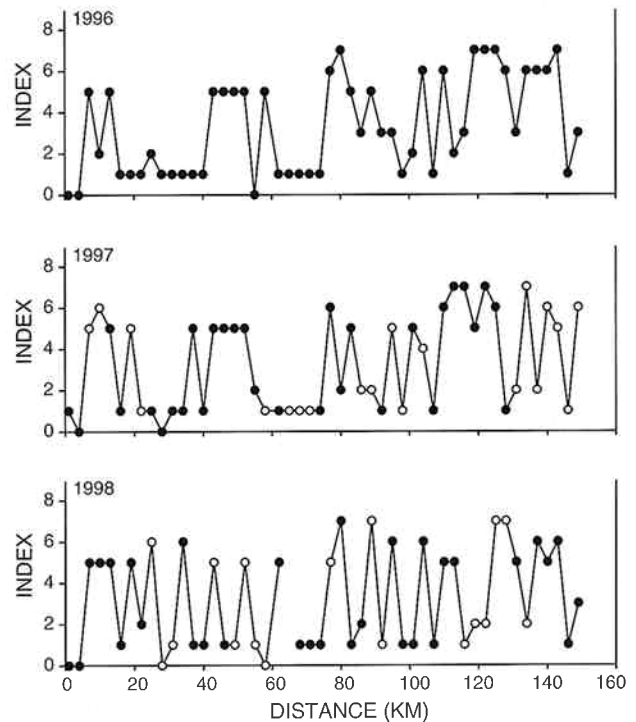

Fig. 8. Earthworm species composition (Table 1) along the transect 1996–1998 plotted against distance (calculated mean distance interval: 3 km); crops at the sites: cereals (black), non-cereals (white).

other one being confined more or less to the northern transect part.

As consequence of species distribution along the transect, there was an increasing trend in respect to number of earthworm species from position 1 to 50. Species number, varying between 0 and 4, increased along the transect (Fig. 7). Although the trend was found in each year, spatial autocorrelation over several lags was found only in 1996, when cereals were grown at all sampling sites (data not shown).

The second biodiversity parameter, earthworm species composition, based on an artificial scaling of earthworm species combinations (Table 1), displayed a similar spatial pattern, with increasing values along the transect (Fig. 8). In 1998, however, this trend was less clear compared to the preceding years. Again, species composition was spatially autocorrelated only in 1996; the 1997 and 1998 data fluctuated randomly (data not shown).

In 1996, when species number and species composition were autocorrelated, successions of identical earthworm biodiversity character (species numbers or species composition) were found, which related to maximally five consecutive sampling points corresponding to approx. 15 km, mainly in the southern part of the transect (Figs. 7 and 8).

Species number and species composition seem to vary with high point to point variation, from position 25 to 50, corresponding to that portion of the transect where in the original sampling interval distances between points were greater than the mean.

3.2. Soil properties

All soil variables showed clear spatial patterns, consisting in increasing trends along the transect (Figs. 9 and 10) (see also Wirth, 2001). Only C/N decreased gradually in course of the transect (Fig. 10). Differences between years were generally small (see also Wirth, 2001). All variables were characterized by spatial autocorrelation, with only minor exceptions from this general pattern (e.g. no spatial autocorrelation in C_{org} in 1996, pH in 1997, and C/N in 1998).

3.3. Relationship between earthworm biodiversity and soil properties

Subsequently multivariate state-space analysis was carried out with the 1996 earthworm biodiversity data and five soil properties, which were characterized by spatial autocorrelation and crosscorrelation. State variables were selected which were suited to estimate the alterations of species number or species composition along the transect.

The results of the state-space analysis for species number are shown in Fig. 11. The spatial pattern of species number could be estimated with a combination of pH and N_t (Fig. 11a). That is, the development of species number along the transect could be estimated based on values for species number at previous or neighbouring locations$_{i-1}$ together with values for pH and N_t at previous locations$_{i-1}$. The relative contribution of neighbouring values of species
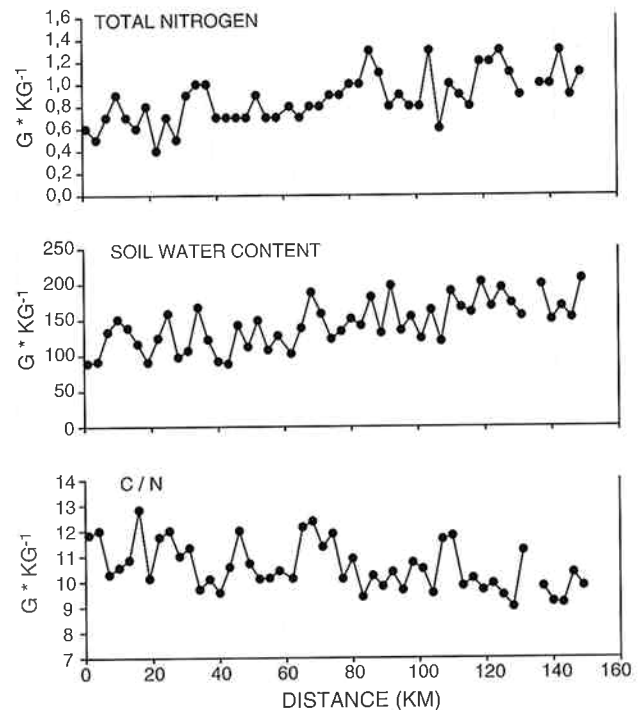

Fig. 10. Total nitrogen content, water content at sampling and C/N along the transect in 1996 plotted against distance (calculated mean distance interval: 3 km).
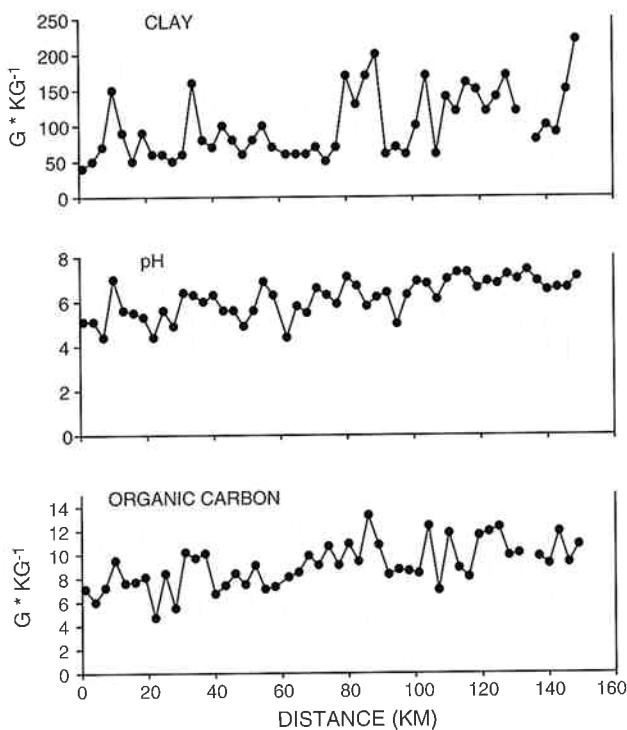

Fig. 9. Clay content, pH and content of organic carbon along the transect in 1996 plotted against distance (calculated mean distance interval: 3 km).

number along the transect was about 20%; soil properties at previous locations were more important for estimating species numbers, as the transition coefficients reveal (Fig. 11a). Most important predictor was soil pH contributing almost 50% to estimated values. The relationship between estimated and measured values of species number was characterized by a coefficient of determination r^2 equal to 0.97. Other combinations of species number with soil properties yielded less satisfying results (data not shown).

In state-space analysis the estimation of parameters is affected by the updating process within the Kalman filter (Wendroth et al., 2001). Therefore the model quality was tested by omitting successively originally measured values from the estimation (Wendroth et al., 1997, 2001; Nielsen and Wendroth, 2003). Thus the number of updating steps along the transect was reduced, and the predictive power of the model for not observed locations was tested by this procedure.

The results of this procedure are shown in Fig. 11b and c. When only every other value of species number along the transect was considered in the estimation, the transition coefficient, reflecting the relative weight of state variables in estimating species number along the transect, changed. The estimations were, however, still sufficiently close to measured values, with a coefficient of determination (r^2) of 0.78 between measured and estimated values of species number (Fig. 11b). Even when two out of three values of