Biol Fertil Soils (2004) 39:383–390
DOI 10.1007/s00374-003-0714-2

ORIGINAL PAPER

Violette Geissen · Christian Kampichler

Limits to the bioindication potential of Collembola in environmental impact analysis: a case study of forest soil-liming and fertilization

Received: 20 February 2003 / Accepted: 8 November 2003 / Published online: 4 March 2004

Abstract We tested the suitability of the collembolan community as a bioindicator for assessing the effects of forest soil-liming and fertilization on the belowground decomposer community. Our investigation was based on a 5-year survey that took place in a German oak-hornbeam and spruce forest in which amelioration measures took place in 1988, 1994 and 1995, with chemical parameters and Collembola being sampled between 1993 and 1997. To address these questions, we applied new methods which have not yet, to our knowledge, been used in biomonitoring studies on forest soils. We used a time-lag analysis for the detection of directional change and a regression-tree induction to show the relationship between Collembola and soil factors. Soil parameters changed considerably after the onset of liming and fertilization. However, no change was detected in community composition over time, nor was there a relationship between Collembola and soil parameters that would make possible the development of a model with at least a moderate predictive success. Taking into account the effort invested in this study (5-year sampling period, identification of 35,000 Collembola, 99 species, 1,170 chemical analyses), we question the suitability of collembolan communities as a bioindicator for forest disturbance.

V. Geissen (✉)
El Colegio de la Frontera Sur,
Administración Correos 2, Apartado Postal 1042,
86100 Villahermosa, Tabasco, Mexico
e-mail: vgeissen@web.de
Fax: +52-993-3511861

C. Kampichler
Soil Zoology and Ecology Laboratory,
Free University Berlin,
Grunewaldstrasse 34, 12165 Berlin, Germany

Present address:
C. Kampichler, División Académica de Ciencias Biológicas,
Universidad Juarez Autónoma de Tabasco,
Carretera Villahermosa-Cárdenas Km 0.5 s/n,
Entronque a Bosques de Saloya,
86150 Villahermosa, Tabasco, Mexico

Keywords Ecological monitoring · Indicator species · Community composition change · Regression-tree induction

Introduction

Most European forests have been directly or indirectly influenced by human activity. On the one hand, humans have directly affected natural forests by silvicultural management measures (thereby modifying the composition of wood species), by increasing soil compaction and by other interventions. On the other hand, indirect cross-regional influences, such as airborne emissions of sulfuric acid and N, have also affected forest ecosystems across large parts of Europe. Acidification was regarded as a particular problem for forest health (Veerhoff et al. 1996), so that various forest soil amelioration measures such as liming were carried out in many European countries over the past 15 years (Ulrich 1986; Kautz and Topp 1998). Liming not only increases soil pH and nutrient concentrations, but also inevitably influences the whole forest soil subsystem. This leads to the following questions: How can these changes be characterized? What indicators can be used to determine the effects on the soil organisms? The influence of liming on soil chemical and physical parameters of forest soils has been described by various authors (Kreutzer et al. 1991; De Vries et al. 1995). However, they described effects only on the abiotic state of the soil, but gave no information on the impact on the community of soil organisms. It is difficult to analyze the effects of human amelioration measures on soil flora and fauna because of their immense diversity. Thus, soil ecologists tend to select a particular taxonomic group—mostly governed by their personal taxonomic expertise—and analyze their relationships with soil chemical and physical parameters, hoping that the selected group might turn out to serve as a “surrogate for larger communities” (Markert et al. 2003). Lumbricids, for example, are good indicators to show the reactions of soil fauna to forest soil-liming (Makeschin 1991; Geissen et al. 1998; Kautz and Topp 1998). However, what can be said about the indicator function of other soil animals, such as mesofauna? In many studies, the abundance and community composition of Collembola have been used as indicators to describe changes in forest soils after liming or fertilization (Hågvar 1984; Hågvar and Abrahamsen 1984; Vilkamaa and Huhta 1986; Van Straalen et al. 1988; Fritsch 1994; Illmann et al. 1997; Chagnon et al. 2001; Haimi and Mätäsniemi 2002). In many cases, Collembola showed a reaction to these treatments. However, does this fact alone necessarily suggest their use as bioindicators?

According to Markert et al. (2003), a bioindicator is an organism, a part of an organism or a community of organisms which contains information on the quality of the environment. Hence, any indicator used for a temporal change in environmental quality—as has to be expected after the application of soil melioration measures—must not show any random fluctuations which bear no relevance to the factor to be indicated (Van Straalen 1997). This simply means that the composition of a collembolan community, if chosen as a bioindicator, should remain stable as long as environmental conditions do not change and should show directed change as soon as environmental conditions begin to change, e.g. as a consequence of liming. Van Straalen (1997) presented data showing that soil arthropods actually remain quite stable over time, a fact reported earlier by Bengtsson (1994). Kampichler and Geissen (in preparation) confirmed this constancy over time by analyzing Collembola time-series in stable forest ecosystems. Directed community change after the onset of liming, however, has not been demonstrated yet for Collembola.

In addition, the usefulness of a bioindicator depends on the strength of the relationship between the causative environmental factor and the ecological endpoint (e.g. species composition; Van Straalen 1997). Only a distinct effect of soil factors on soil microarthropod communities would allow for the conversion of the causal relationship into a bioindication (cf. Van Straalen 1997, Fig. 10.1). Liming affects soil chemistry by changes in pH and concentration of exchangeable elements (Kreutzer 1995; Geissen 2000). Effects of soil chemistry on collembolan communities have been repeatedly investigated (Van Straalen and Verhoef 1997; Chagnon et al. 2000; Cassagne et al. 2003). However, the question remains: do soil chemical parameters, such as pH or element concentrations, have a sufficiently strong and predictive effect on collembolan community characteristics (e.g. abundance of the constituent species)? If the answer to this question is no, then Collembola have no potential for indicating these parameters nor—most of all—for the calibration of the bioindication system (Markert et al. 2003).

Despite some authors advocating the potential of microarthropods for bioindication and for monitoring pollutant effects (Van Straalen 1997; Cortet et al. 1999), we feel that uncertainties remain about the usefulness of collembolan community composition for the bioindication of changes in soil chemical parameters, due to a lack of strength in the relationship to the causative factors. We tested our assumptions based on soil chemistry data derived from a 5-year study on liming and fertilization of acidified forest soils and asked two questions:

1. Is a directed change in collembolan community composition detectable after the onset of amelioration measures?
2. Do the changes in environmental factors, namely pH and concentration of various plant-available elements, have any predictive power for total abundance or abundance of dominant species?

To address these questions, we applied new methods which, to our knowledge, have not yet been used in bioindication studies on forest soils. We used a time-lag analysis for the detection of directional change and a regression-tree induction to show the relationship between Collembola and soil factors.

Materials and methods

Study area and treatment patterns

The study was conducted in the Kottenforst (a forest near Bonn, Germany), consisting of stands of oak-hornbeam and spruce on a plain at 150 m above sea level. Typical soil types found in this region are silty loamy Stagnic Alisols from loess as parent material. In the deciduous forest (150 years old), the humus type is moder with an Of horizon and a very thin (<1 cm) Oh horizon. There is a sparse cover of small balsam (*Impatiens parviflora*). The spruce forest (80 years old) has typically moder to raw humus with fairly thick Of and Oh layers. In the spruce forest, the ground cover is very dense and is dominated by blackberry (*Rubus fruticosus*) and bracken (*Pteridium aquilinum*).

Six plots were selected in both the oak-hornbeam (plots O1–O6) and the spruce (S1–S6) forests. Each plot was 2,500 m^2. Different quantities of lime, partly combined with fertilization using P and K, were applied to these plots in 1988, 1994 and 1995 (Table 1).

Table 1 Liming and fertilization treatments in plots O1–O6 (oak-hornbeam forest) and plots S1–S6 (spruce forest). P was applied as P_2O_5 (partially acidified rockphosphate) and K as K_2SO_4

Time	Application	Plot (treatment code)					
		O1	O2	O3	O4	O5	O6
		S1	S2	S3	S4	S5	S6
October 1988	Dolomite (t ha^{-1})	0	3	3	3	3	3
January 1994	Dolomite (t ha^{-1})	0	0	3	3	6	6
June 1994	P_2O_5 (kg ha^{-1})	0	0	0	200	200	200
	K_2O (kg ha^{-1})				150	150	150
January 1995	Dolomite (t ha^{-1})	0	0	0	0	0	6

Sampling and analysis

Sampling for Collembola and soil chemical analyses were carried out in October 1993 before liming in January 1994 and then each October during 1994–1997. Thus, we studied the collembolan fauna over a period of five consecutive years on 12 plots in total. Twelve soil cores (diam. 5 cm) were taken in the center of each plot (100 m^2) at each sampling date. The minimum distance between single cores or between cores and tree stems was 1 m. We extracted Collembola from ten cores before soil chemical analysis took place. The two additional cores were used only for soil chemical analysis. The soil cores were divided into their organic (L/Of, Oh) and mineral layers (0–5 cm, 5–10 cm).

After the extraction of Collembola, using a modified Tullgren–Berlese extractor, four out of the 12 single samples from each plot and depth were mixed to form one sample. These three mixed samples from each plot and depth were used for soil chemical analyses.

In addition to pH analysis ($CaCl_2$), NH_4NO_3 extraction (Zeien and Brümmer 1989) was used to determine the content of mobile nutrients (Na, K, Ca, Mg, P), heavy metals (Fe, Mn, Pb, Cd) and Al in the different layers of the soil. Further, the content of organic C and total N was determined by standard methods.

Statistics

We examined the response of collembolan community composition to amelioration treatments over time by the time-lag analysis (TLA) of Collins et al. (2000), a method developed for the short time-series that are often required in ecological studies. Briefly, for all possible pairs of years (*j* with *k*), we calculated the Euclidean distance *ED* between their communities as:

$$1.\ ED = \left[\sum_{i=1}^{S}\left(x_{ij} - x_{ik}\right)^2\right]^{\frac{1}{2}}$$

Here, x_{ij} indicates the abundance of the *i*th species in the *j*th year, x_{ik} indicates the abundance of the *i*th species in the *k*th year and *S* indicates the total number of species.

We included those species in our analysis which showed a numerical dominance of >3% in at least one of the plots during at least one of the study years. We plotted the *ED* values against time-lag for all lags and calculated a linear regression for *ED* as a function of the square root of the time-lag. The square root transformation corrects for the fact that the smaller number of points at larger time-lags potentially biases the regression analysis. By means of TLA, it is possible to discriminate between different patterns of community change: significantly positive regression slopes ($P<0.05$) imply directional change (*ED* increases with time-lag), non-significant slopes ($P>0.05$) imply stability or stochastic variation (*ED* does not change with time-lag) and significantly negative slopes ($P<0.05$) imply convergence (*ED* decreases with time-lag).

We analyzed the relationship between soil chemical properties and the abundance of dominant collembolan species by applying an automated regression-tree induction. The data used for these analyses were average values per plot ($n_{\text{plots}}=6$) and per year ($n_{\text{years}}=5$), yielding a data set with 30 cases per forest type and soil horizon. In the spruce forest, the analyses were performed only for the Of and Oh horizons, as collembolan abundance was too low in the Ah horizon. Regression-trees are a tool for data-exploration developed by artificial intelligence researchers (Breiman et al. 1984). Briefly, tree-structured regression is based on the assumption that the relationship between independent and dependent variables is not uniform over the entire range of variable values but can be approximated in smaller sub-domains, thus operating as a piecewise linear regression. The tools developed for model tree induction identify these sub-domains automatically by splitting the database in all possible ways into two subsets and selecting the portion that maximizes error reduction in the data subsets. The "branches" of the tree characterize the identified sub-domains with regression functions for the dependent variable. Regression-tree induction is gaining increasing recognition in ecological data analysis and the development of predictive models (Kompare and Dzeroski 1995; Andersen et al. 2000; De'ath and Fabricius 2000; Kampichler et al. 2000; De'ath 2002). We applied the Cubist program (provided by Rulequest Research, http://www.rulequest.com) for regression-tree induction and constructed the following models for the dominant species of Collembola.

The global model

This model includes all data for every chosen dependent variable (abundance of total Collembola, abundance of the three most dominant species in both oak-hornbeam and spruce forests). Performance of the model is characterized by:

1. The mean absolute error (MAE), which is the average of the absolute difference between prediction and observation of the number of individuals per core.
2. The relative absolute error (RAE), which is the ratio between MAE and the error magnitude that would result from always predicting the mean value).
3. The correlation coefficient (*r*) between predicted and observed values.

Validation (comparing prediction with observation) is performed based on the data that were used for model induction, the so-called "training data". A sound model produces:

- A small MAE. The smaller the MAE, the closer together are prediction and observation. Ideally, the MAE should be a small fraction of the average number of Collembola.
- A small RAE. The smaller the RAE, the larger is the error reduction in comparison with using the mere average number of Collembola as a predictor. For example, RAE=0.5 means an error reduction by 50%, RAE=0.1 by 90%.
- A correlation coefficient *r* that is close to 1.0.

A two-fold cross-validation

In this case, the data set is split in two; and one half of the data at a time is used for regression-tree induction (training data), while the other half is held back for validation of the model (test data). This allows two models to be derived and makes possible an examination of their correspondence. The performance of these models is characterized by MAE, RAE and *r* for both the training data and test data. Cross-validation thus gives an estimate of the model's performance for new cases, that is to say, its predictive power.

Results and discussion

Soil chemical properties

The amelioration treatments led to significant changes in soil chemical properties in the organic layers and the Ah horizon in both the spruce and the deciduous forests (Geissen 2000). The pH and basic cations increased significantly in the treated plots, while the content of Al and heavy metals decreased. The effect of the treatment in the spruce forest was mainly limited to the organic horizons (Of, Oh; Geissen 2000). Only the S5 and S6 plots showed this effect in the Ah horizon. In the deciduous forest, significant changes in soil chemical parameters in the Ah horizon were found in all plots (Geissen 2000). The development of soil properties in all

Table 2 Ranges of soil chemical parameters [pH ($CaCl_2$), concentration of plant-available nutrients, Al, heavy metals, organic C (C_{org}), total N (N_{tot})] in autumn 1997 in the Of and Ah horizons of the oak-hornbeam forest and in the Of and Oh horizons of the spruce forest (ranges given for all six treatments). *Asterisk* (*) indicates significant effect of treatment ($P<0.05$) on soil chemical parameters during the years 1994–1997 (for details, see Geissen 2000)

Chemical parameter	Oak-hornbeam forest		Spruce forest	
	Of	Ah	Of	Oh
pH ($CaCl_2$)	4.2–6.2*	3.6–5.9*	3.7–6.5*	3.9–6.6*
Ca (mg kg^{-1})	3,661–6,920*	174–1168*	3,480–7,683*	4,427–8,000*
Mg (mg kg^{-1})	535–1,747*	36–354*	192–1,721*	879–1,975*
K (mg kg^{-1})	766–1,121*	61–160*	340–707*	212–308*
P (mg kg^{-1})	108–170*	1.1–3.3*	28–65*	8–29*
Al (mg kg^{-1})	3.9–15.5*	1.7–126.0*	1.8–33.3*	1.6–60.6*
Fe (mg kg^{-1})	2.3–4.8	1.0–21.0*	3.3–8.0*	2.7–12.6*
Mn (mg kg^{-1})	130–873*	11–76*	74–609*	27–277*
Zn (mg kg^{-1})	0.8–11.5*	0.2–5.5*	0.8–38.2*	0.4–22.9*
Cd (mg kg^{-1})	0.01–0.12	0.01–0.13*	0.02–0.34*	0.01–0.28*
Pb (mg kg^{-1})	0.13–0.47	0.03–3.55*	0.13–2.5	0.2–4.1*
C_{org} (g in 100 g)	27–33	4.5–7.1	30–38	28–36
N_{tot} (g in 100 g)	13–19	2.1–3.5	15–18	14–17

years was described by Geissen (2000). Table 2 shows the ranges of soil chemical properties in the last year of the investigations and presents significant changes in soil chemical parameters after treatment.

Collembolan fauna

During 5 years of investigation, we extracted 34,941 Collembola (12,199 specimens from oak-hornbeam forest, 22,742 from spruce forest). They ranged from a minimum of 4,357 individuals m^{-2} in plot O3 in 1996 to a maximum of 115,120 individuals m^{-2} in S2 in 1994. We identified 99 collembolan species in total, 73 of them appearing in the oak-hornbeam forest and 76 in the spruce forest (Geissen 2000).

Collembolan community response

Thirteen species from the oak-hornbeam forest and 12 species from the spruce forest were sufficiently abundant to be included in the TLA analysis. This corresponded to an inclusion of approximately 80% of all collembolan individuals.

There were no indications of a directional change in the collembolan community; and no single TLA yielded a regression line statistically different from zero, i.e. $P<0.05$, neither in the oak-hornbeam nor in the spruce forest. *P* ranged between 0.17 and 0.95 (Fig. 1), which is what we would expect for the control plots O1 and S1, which did not receive any liming or fertilization treatment (Fig. 1a, g). However, the treated plots did not show any response to liming or fertilization, as regards community structure (Fig. 1b–f, h–l). The Euclidian distances showed enormous year-to-year fluctuations and any signal at all in the community data induced by treatments was entirely masked by the noise of the stochastic population fluctuations of the constituent species. This means that either the collembolan community composition did not respond to the treatments in a directed way, or even exhaustive monitoring over five consecutive years (with sampling and identification of 35,000 individual collembolans) was insufficient for detecting responses at the community level. Neither conclusion supports the potential of Collembola as practical indicators for habitat quality-change following liming or fertilization in such forest. Hågvar (1984) showed long-term changes in the dominance structure of Collembola after the liming of acidic forest soils. However, he did not consider the time factor and only compared different plots with different treatments. Other studies only investigated some dominant species as indicators for changes in the collembolan community (Fromm 1998; Mebes 1998; Geissen 2000). These studies were carried out over a period of 3–5 years.

Relationship between soil chemical properties and Collembola

Regression trees (Fig. 2 presents the models for total Collembola) are read from left to right. For any data case (that is, a value of the dependent variable, with the values of its corresponding independent variables), the tree runs from the left (its roots) to the right (its branches). At every bifurcation, the results of the related values are compared with two alternatives and the corresponding branch is chosen. Every case leads to a branch and to a prediction for the value of the dependent variable. If a case is represented by more than one branch, the average is calculated as a prediction of the dependent variable.

Tables 3 and 4 sum up the statistics for all constructed regression trees. Induction of regression trees for total Collembola (Fig. 2) produced models with a high correlation between prediction and observation (r=0.87 for each of the forest systems), suggesting considerable predictive success if the data used for tree generation are also used for tree validation. The precision of the predictions, however, was only moderate. In the oak-hornbeam forest, for example, the average number of Collembola per core was 197.2 and MAE amounted to 84.9 (43% of the average; Table 3). This means that, despite a favorable correlation between predictions and observations, there is a considerable deviation among their values. Similar conditions are valid for total

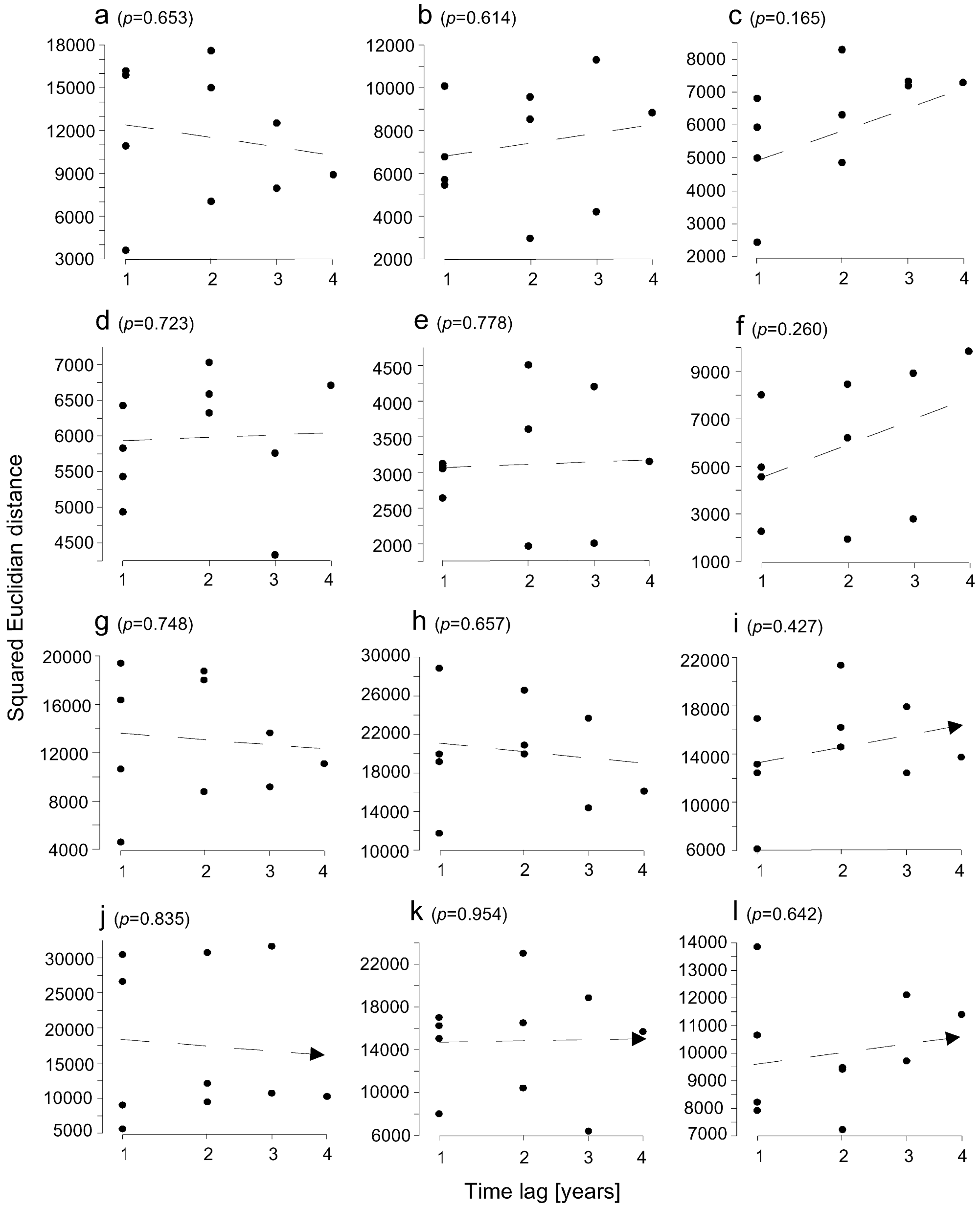

Fig. 1 Relationships between time-lag (years) and the squared Euclidian distance between collembolan communities over five consecutive years in plots in oak-hornbeam forest (**a–f**) and spruce forest (**g–l**), with different liming and fertilization treatments. The significance levels (p) of the linear regressions are *given in parentheses*. **a** Plot O1, **b** O2, **c** O3, **d** O4, **e** O5, **f** O6, **g** S1, **h** S2, **i** S3, **j** S4, **k** S5, **l** S6. See Table 1 for an explanation of the treatment in each plot

Collembola in spruce forest with MAE amounting to 41% of the average number of Collembola (Table 4).

Predictive success varied widely for the dominant species. The regression trees for *Folsomia quadrioculata* led to considerable error reduction (RAE=0.20 in oak-hornbeam, RAE=0.29 in spruce; Tables 3, 4). For other species, modeling success was weak (e.g. *Friesea mirabilis* in the spruce stand) or completely absent (*Mesaphorura macrochaeta* in the oak-hornbeam stand; Tables 3, 4).

The two-fold cross-validation revealed only a weak predictive power for the tree models. RAE never achieved values smaller than 0.57 when unseen test data were used for validation; and 12 out of 16 RAE values were between 0.94 and 1.28 (Tables 3, 4), which meant that no sensible predictive model could be generated. Using this model was no better than simply using collembolan averages as predictors. Correlation coefficients between observed and predicted values never rose above r=0.60, with most of the values of r ranging around 0.30 (Tables 3, 4). This indicates that, for some species and for total Collembola, a relationship with environmental data could be described only when the entire data set was used. Soil chemical parameters, however, had no predictive power for new situations not originally included in the model generation.

Thus, based on a typical sampling protocol for soil chemical and microarthropod data, we do not see any possibility for predicting collembolan numbers over the range in our data set. In addition, a closer inspection of the alternative models of the two-fold cross-validation shows that they deviated considerably both from each other and from the global model. For example, in the

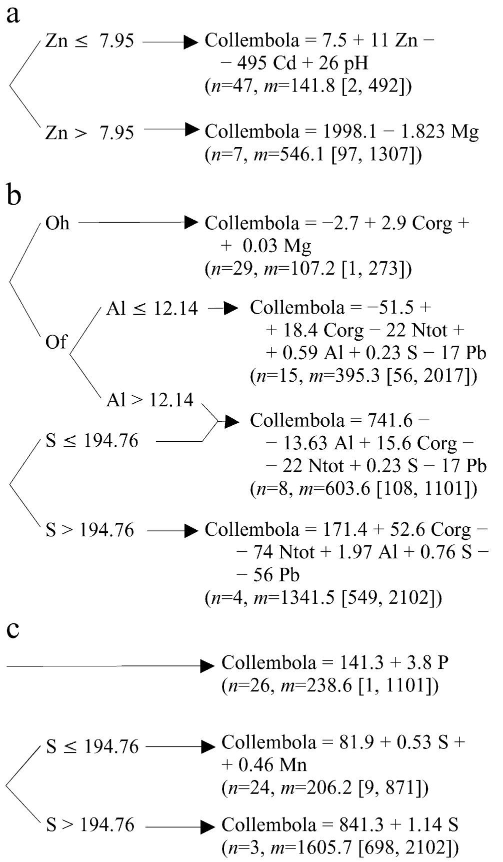

Fig. 2 Global regression trees for total Collembola in oak-hornbeam forest (**a**) and spruce forest (**b**) and regression trees yielded by two-fold cross-validation for total Collembola in spruce forest (**c**). The branches of the trees show multiple regressions that predict the value of the dependent variable. Independent variables in the equation are ordered according to their effect on the dependent variable. The data in parentheses signify the number of cases (*n*) covered by the respective regression equation and the average (*m*) of the dependent variable in these cases (with minimum and maximum values in square brackets).

spruce stand, one half of the data for total Collembola led to a single linear model relating collembolan abundance to P, while the other half identified S as the element with best predictive power (Fig. 2c). Both trees differed drastically from the global model (Fig. 2b). Regression-tree induction for the dominant species produced similar results (data not shown). Thus, the models depended to a large degree on the idiosyncrasies of single sample units, yielding models of different structure with little potential for generalization. This can certainly be put down to the small sample size. However, this point of criticism has to be measured against the degree of feasibility for carrying out such a study. Here, the identification of Collembola from 300 cores per forest system (6 plots, 5 sampling dates, 10 replicates) and 1,170 chemical analyses per forest system (6 plots, 5 sampling dates, 3 replicates, 13 chemical parameters) did not allow for the development of a reasonably successful predictive model; and no single indicator species could be identified.

We are convinced that studies already published would have produced a similar picture, if purely correlative data analysis had been augmented by tests for predictive power. Some authors seem to be content with regression analysis and an estimate of r^2. In addition, conventional correlation analyses have often been performed (Hågvar 1984; Hågvar and Abrahamsen 1984; Fritsch 1994; Geissen et al. 1997; Fromm 1998; Mebes 1998) and sometimes used for the characterization of species as indicators for certain soil conditions and for labeling them as “acidophilic” or the like. Applying this approach to the data in this study yielded 11 statistically significant Spearman correlation coefficients (n=30, P<0.05) in the oak-hornbeam forest and nine in the spruce forest between single collembolan species and single soil chemical parameters (cf. Geissen 2000), a fact which contrasts quite obviously with the meager results of regression-tree induction. Thus, we conclude that large values for r^2 in global regression models or significant correlation coefficients between single species and soil parameters are not sufficient for justifying the use of collembolan communities as bioindicator systems, unless

Table 3 Predictive accuracy of regression-trees for the dominant collembolan species in the oak-hornbeam forest. To assist interpretation of the statistics, the average number of individuals per core is given in the first line. *MAE* Mean absolute error, *RAE* relative absolute error, *r* correlation coefficient (for explanation of statistical data, see Materials and methods)

	Folsomia quadrioculata		*Isotomiella minor*		*Mesaphorura macrochaeta*		Total Collembola	
Average number per core	52.4		19.6		16.2		197.2	
Global model								
MAE	10.2		14.4		15.6		84.9	
RAE	0.20		0.68		1.00		0.59	
r	0.99		0.75		0.00		0.87	
Cross-validation training data								
MAE	12.7	29.9	4.7	21.5	17.0	12.5	75.5	81.9
RAE	0.50	0.36	0.90	0.63	1.00	0.86	0.37	0.92
r	0.77	0.90	0.45	0.77	0.00	0.36	0.94	0.27
Cross-validation test data								
MAE	67.3	52.0	27.2	25.6	16.1	15.8	155.4	175.0
RAE	0.99	1.11	0.96	1.03	1.00	1.04	1.28	0.99
r	0.05	0.48	0.36	0.45	0.00	0.20	0.12	0.03

Table 4 Predictive accuracy of regression-trees for the dominant collembolan species in the spruce forest. To assist interpretation of the statistics, the average number of individuals per core is given in the first line (for explanation of statistical data, see Materials and methods)

	Folsomia quadrioculata		*Isotoma notabilis*		*Friesea mirabilis*		Total Collembola	
Average number per core	58.4		43.4		30.0		297.1	
Global model								
MAE	18.3		26.7		21.3		121.5	
RAE	0.29		0.53		0.85		0.46	
r	0.92		0.78		0.48		0.87	
Cross-validation training data								
MAE	18.0	7.0	31.6	22.7	14.9	32.3	188.5	139.2
RAE	0.57	0.12	0.56	0.51	0.85	1.00	0.52	0.80
r	0.59	0.99	0.77	0.78	0.48	0.00	0.76	0.44
Cross-validation test data								
MAE	36.7	70.3	29.3	58.1	27.8	22.2	180.5	265.6
RAE	0.57	1.10	0.60	1.11	0.94	1.00	0.78	0.89
r	0.59	0.21	0.60	0.41	0.29	0.00	0.35	0.47

predictive success is tested by a validation technique, preferably by cross-validation (Olden and Jackson 2000). In particular, we question the approach of “fishing for correlations”, since it is almost inevitable in a species-rich assemblage that there are single species that “respond” to treatments merely by chance.

Conclusions

In our study, we could not detect any directional change of collembolan community composition as a consequence of forest soil-liming and fertilization, nor was it possible to detect any relationship between the abundance of total Collembola or single species and soil chemical parameters that would allow for the development of a model with at least moderate predictive success. Based on the given results, we dare question the suitability of collembolan communities as indicators to determine the influence of amelioration measures on soil ecosystems in European forests. We do not preclude the possibility of a causative relationship between soil parameters and collembolan community and population characteristics that would allow for making predictions. However, from the results of this study, a cost–benefit analysis of the effort undertaken for our study (5-year sampling period, identification of 35,000 Collembola, 99 species, 1,170 chemical analyses) leaves us rather pessimistic whether a relationship, if it existed, could possibly be detected. The input of time and manpower necessary to gather sufficient data by far exceeds any degree of feasibility.

Acknowledgements We are grateful to the Ministerium für Umwelt, Raumordnung und Landwirtschaft of Northrhine–Westfalia, Germany, for financial support. We thank Prof. Brümmer, who supported the analyses of soil chemical parameters in the Department of Soil Science, University of Bonn, Germany. We also thank Anke Flohr, Jutta Illmann and Kirsten Rüsch for their support in field and laboratory work and Enken Tadsen-Duch for language editing of the manuscript.

References

Andersen MC, Watts KM, Freilich JE, Yool SR, Wakefield GI, McCauley JF, Fahnestock PB (2000) Regression-tree modelling of desert tortoise habitat in the central Mojave desert. Ecol Appl 10:890–900

Bengtsson G (1994) Temporal predictability in forest soil communities. J Anim Ecol 63:653–665

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression-trees. Wadsworth, Belmont

Cassagne N, Gers C, Gauquelin T (2003) Relationships between Collembola, soil chemistry and humus types in forest stands (France). Biol Fertil Soils 37:355–361

Chagnon M, Paré D, Hébert C (2000) Relationships between soil chemistry, microbial biomass and the collembolan fauna of southern Québec sugarmaple stands. Ecoscience 7:307–316

Chagnon M, Paré D, Hébert C, Camiré C (2001) Effects of experimental liming on collembolan communities and soil microbial biomass in a southern Québec sugar maple (*Acer saccharum* Marsh.) stand. Appl Soil Ecol 17:81–90

Collins SL, Micheli F, Hartt L (2000) A method to determine rates and patterns of variability in ecological communities. Oikos 91:285–293

Cortet J, Gomot-de Vauflery A, Poinsot-Balaguer N, Gomot L, Texier C, Cluzeau D (1999) The use of invertebrate soil fauna in monitoring pollutant effects. Eur J Soil Biol 35:115–134

De Vries W, Grinsven JJM van, Breemen N van, Leeters EEJM, Jansen PC (1995) Impacts of acid deposition on concentrations and fluxes of solutes in acid sandy forest soils in the Netherlands. Geoderma 67:17–43

De’ath GK (2002) Multivariate regression-trees: a new technique for modelling species–environment relationships. Ecology 83:1105–1117

De’ath G, Fabricius KE (2000) Classification and regression-trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192

Fritsch N (1994) Auswirkungen der Waldkalkung auf die Bodenmesofauna. PhD thesis, University of Mainz, Mainz

Fromm H (1998) Räumliche und zeitliche Variabilität der Collembolenfauna und ihre Bedeutung für C- und N-Umsatz in einer Agrarlandschaft. Shaker, Aachen

Geissen V (2000) Reaktionen ausgewählter Tiergruppen (Lumbricidae, Annelida; Collembola, Arthropoda) auf Veränderungen der Nähr- und Schadstoffgehalte von Waldböden nach Kalkung und Düngung. (Bonner Bodenkdl Abh 31) University of Bonn, Bonn

Geissen V, Illmann J, Flohr A, Kahrer R, Brümmer GW (1997) Effects of liming and fertilization on Collembola in forest soils in relation to soil chemical parameters. Pedobiologia 41:194–201

Geissen V, Rüdinger C, Schöning A, Brümmer GW (1998) Microbial biomass and earthworm populations in relation to soil chemical parameters in an oak-beech forest soil. Verh Ges Ökol 28:249–258

Hågvar S (1984) Effects of liming and artificial acid rain on Collembola and Protura in coniferous forest. Pedobiologia 27:341–354

Hågvar S, Abrahamsen G (1984) Collembola in Norwegian coniferous forest soils. III. Relations to soil chemistry. Pedobiologia 27:331–339

Haimi J, Mätäsniemi L (2002) Soil decomposer animal community in heavy-metal contaminated coniferous forest with and without liming. Eur J Soil Biol 38:131–136

Illmann J, Geissen V, Brümmer GW (1997) Auswirkungen von Kalkungs- und Düngungsmaßnahmen auf die bodenbewohnende Collembolenfauna in Laub- und Nadelwald. Verh Ges Ökol 27:427–434

Kampichler C, Džeroski S, Wieland R (2000) The application of machine learning techniques to the analysis of soil ecological data bases: relationships between habitat features and Collembola community characteristics. Soil Biol Biochem 32:197–209

Kautz G, Topp W (1998) Nachhaltige waldbauliche Maßnahmen zur Verbesserung der Bodenqualität. Forstwiss Centralbl 117:23–43

Kompare B, Džeroski S (1995) Getting more out of data: automated modelling of algal growth with machine learning. In: Matsuzaki C, Miyaji Y (eds) Proceedings of the international conference on coastal ocean space utilization. University of Hawaii, Honolulu, pp 209–220

Kreutzer K (1995) Effects of forest liming on soil processes. Plant Soil 168/169:447–470

Kreutzer K, Göttlein A, Pröbstle P (1991) Dynamik und chemische Auswirkungen der Auflösung von Dolomitkalk unter Fichte (*Picea abies*). In: Kreutzer K, Göttlein A (eds) Ökosystemforschung Höglwald. Beiträge zur Auswirkung von saurer Beregnung und Kalkung in einem Fichtenaltbestand. Forstwiss Forsch 39:186–203

Makeschin F (1991) Auswirkungen von saurer Beregnung und Kalkung auf die Regenwurmfauna (Lumbricidae: Oligochaeta) im Fichtenaltbestand Höglwald. In: Kreutzer K, Göttlein A (eds) Ökosystemforschung Höglwald. Forstwiss Forsch 39: 117–127

Markert BA, Breure AM, Zechmeister HG (2003) Definitions, strategies and principles for bioindication/biomonitoring of the environment. In: Markert BA, Breure AM, Zechmeister HG (eds) Bioindicators and biomonitors. Elsevier, Amsterdam, pp 3–39

Mebes K-H (1998) Collembolengemeinschaften in Agrarökosystemen: Steuerung durch Umweltfaktoren—Einfluß auf den Stoffumsatz. Shaker, Aachen

Olden JD, Jackson DA (2000) Torturing data for the sake of generality: how valid are our regression models? Ecoscience 7:501–510

Ulrich B (1986) Die Rolle der Bodenversauerung beim Waldsterben: Langfristige Konsequenzen und forstliche Möglichkeiten. Forstwiss Centralbl 105:421–435

Van Straalen NM (1997) Community structure of soil arthropods as a bioindicator of soil health. In: Pankhurst CE, Doube BM, Gupta VVSR (eds) Biological indicators of soil health. CAB International, Wallingford, pp 235–264

Van Straalen NM, Verhoef HA (1997) The development of a bioindicator system for soil acidity based on arthropod pH preferences. J Appl Ecol 34:217–232

Van Straalen NM, Kraak MHS, Denneman CAJ (1988) Soil microarthropods as indicators of soil acidification and forest decline in the Veluwe area, the Netherlands. Pedobiologia 32:47–55

Veerhoff M, Roscher S, Brümmer GW (1996) Ausmaß und ökologische Gefahren der Versauerung von Böden unter Wald. Schmitt, Berlin

Vilkamaa P, Huhta V (1986) Effects of fertilization and pH on communities of Collembola in pine forest soil. Ann Zool Fenn 23:167–174

Zeien H, Brümmer GW (1989) Chemische Extraktionen zur Bestimmung von Schwermetallbindungsformen in Böden. Mitt Dtsch Bodenk Ges 59:505–510