SAMPLE SIZE SIMULATION STUDY FOR INVESTIGATION OF TESTATE AMOEBAE SPECIES DIVERSITY

1 Chizhikova N.A., 2 Bobrov A.A., 3 Saveliev A.A.

1,3 Ecology faculty, Kazan State University, Kremlevskaya Street 18, 420008, Kazan, Russia
2 Soil Faculty of Moscow State University, Vorob’evi Gori, 119899, Moscow, Russia
3 chizhikova_n@rambler.ru

This work is devoted to the fundamental methodological problem of protozoology, to the sample size selection. The question is: “How much of testacea individuals must be examined to recover species diversity at a site?”. Dependence of sample size from species abundance structure of community was investigated using simulation of communities.

Testate amoeba (Rhizopoda) are a group of free-living protozoans that have an organic shell (Fig. 1). They live within water films on the soil particles. Well-defined ecological preferences of Rhizopod amoeba and the relatively good preservation of their shells in peats, lake sediments, and buried soils form the basis for the development of rhi-zopod analysis as a method for reconstruction of climate and environmental changes (Tolstoy, 1986). The number of rhizopod species in the samples varied between the diversity of species depends not only on ecological conditions, but on climate as well. It may be suggested that the environmental conditions of the paleoenvironments with higher rhizopod diversity were the most climatically favourable (Bobrov, 2004).

1 gram of soil contains from 100 in humic-accumulative soil horizons to 1000 in organogenic horizons. Researchers prepare soil extract and then examine species in diluted extract. The general agreement is that direct counting from a sample only a small fraction of species diversity (reviewed by Fois-sner, 1987). Ideally, the bigger the number of examined species, the larger the probability that all species presented in assemblage were captured. But this approach entails an enormous sampling effort. 100 observations in the standard quantity customary in protozoology (Woodland, 1998), but this amount isn’t suited for all communities. This quantity is almost inaccessible for communities with low density of rhizopod individuals (par example in desert soils), or researcher will spend several days for inspecting of every sample. Sample size, or the number of observed testacea individuals depends on many factors: (i) on original community species richness, the bigger number of species calls for the bigger size of sample; (ii) on original community heterogeneity structure, because presence of dominant species bring down the possibility to recover full species structure; (iii) on density of testaeas organisms in soil, (iv) on amount of shells damaged owing to fossilisation process. Also patchy distribution of testate amoebae reflectinh habitat heterogeneity may affect the process of species recovering.

Customary approach to define sample size is to plot species accumulation curve during examination of a sample (Fig. 2) and then extrapolate its behaviour. But this approach also calls for a very laborious work. In this study there was investigated using simulations behaviour of a sample size depending on two factors: on original species richness and species abundance structure at a site (Fig. 3). Impact of damaged shells and habitat heterogeneity wasn’t taken into consideration. Also it was considered the only amount of individuals that is sufficient to recover some ratio of species true composition, but not to recover proper species abundance structure. This task is a challenging study, as an investigation how improper species abundance structure may affect prognosis or estimates of palaeo-environmental conditions at a site. There are many models of species abundance distribution: lognormal, broken stick, geometric series and others. In this work simulations of communities were carried out for pre-emption model of species abundance distribution (geometric series or Motomura model) because of its simplicity:

\[
d(r) = \sigma e^{(\alpha r - 1)}
\]

where \(\sigma\) is expected abundance of species at rank \(r\), \(\alpha\) – pre-emption factor which decrease the yield of abundance per rank, and \(r\) – overall amount of individuals at a sample from community. Preemption factor \(\alpha\) is a characteristic of community abundance structure. Big preemption coefficient indicates more pronounced disproportional of abundance between dominant and non-dominant species.

We investigate the possibilities to improve quality of climate paleo-reconstruction by using different information about community (species abundance, presence of key indicator species, species pool), and by using some mathematical models. We also compared the predictions of expert systems and expert systems based on fuzzy logic.

For recent decades there are originated studies which assessed habitat preferences of some testacea species quantitatively (Chuman et al., 2000; Mitchell et al., 2000). Thus rhizopod analysis can be used as a method for reconstruction of climate and hydrological regimes of bog ecosystems (Chuman D.J., Warner B.G. 1997; Mitchell et al., 2001). Analysis of species community composition is concerned with some obstacles. Main problem is the inconsistencies of species composition and disturbance species abundance structure owing to damage of testaeas shells during process of fossilisation. Such information about species as presence-absence is less sensitive to fossilisation in comparison with species abundance. Nevertheless there are some delicate species which can disappear from the species assemblages completely. Also except species abundance and presence-absence one use species pool (Reeves, 1991). In the Masul & Dahlgren model (1997) this approach was proposed to overcome “zero-translantion problem” in sparse community matrices. Absence of some species is sample doesn’t mean that environmental conditions aren’t favourable for them, but this may be due to destruction of amoebae shells during fossilisation process, or due to patchy distribution of species, errors of species identification.

For every combinations of species richness (the maximum possible number of species) and preemption coefficient there were simulated a community which comprises (i) 5000 and (ii) 10000 individuals in soil dilution. Then individuals were rearranged (“dilution was shaked up”). Then there was assessed the minimum sample size which is sufficient for determining (a) 90% and (b) 100% of total species composition. “Shaking” was done 1000 times and distribution of sample size of 90% and 100% composition were obtained.

Main findings:

1. Consequence of self-similarity law – the same species richness and preemption coefficient of initial communities from which dilution was obtained calls for almost the same sample size regardless of the density of individuals in soil dilution or density of individuals in soil.

2. Sample size for determining species composition up to 100% increases more rapidly with rise of community species richness and coefficient of preemption.

In this study there were assessed the possibilities to improve quality of climate paleo-reconstruction by using different information about community (species abundance, presence of key indicator species, species pool), and by using some mathematical models. We also compared the predictions of expert systems and expert systems based on fuzzy logic.

We investigate the possibilities to improve quality of climate paleo-reconstruction by using different information about community (species abundance, presence of key indicator species, species pool), and by using some mathematical models. We also compared the predictions of expert systems and expert systems based on fuzzy logic.

For recent decades there are originated studies which assessed habitat preferences of some testacea species quantitatively (Chuman et al., 2000; Mitchell et al., 2000). Thus rhizopod analysis can be used as a method for reconstruction of climate and hydrological regimes of bog ecosystems (Chuman D.J., Warner B.G. 1997; Mitchell et al., 2001). Analysis of species community composition is concerned with some obstacles. Main problem is the inconsistencies of species composition and disturbance species abundance structure owing to damage of testaeas shells during process of fossilisation. Such information about species as presence-absence is less sensitive to fossilisation in comparison with species abundance. Nevertheless there are some delicate species which can disappear from the species assemblages completely. Also except species abundance and presence-absence one use species pool (Reeves, 1991). In the Masul & Dahlgren model (1997) this approach was proposed to overcome “zero-translantion problem” in sparse community matrices. Absence of some species is sample doesn’t mean that environmental conditions aren’t favourable for them, but this may be due to destruction of amoebae shells during fossilisation process, or due to patchy distribution of species, errors of species identification.

For every combinations of species richness (the maximum possible number of species) and preemption coefficient there were simulated a community which comprises (i) 5000 and (ii) 10000 individuals in soil dilution. Then individuals were rearranged (“dilution was shaked up”). Then there was assessed the minimum sample size which is sufficient for determining (a) 90% and (b) 100% of total species composition. “Shaking” was done 1000 times and distribution of sample size of 90% and 100% composition were obtained.

Main findings:

1. Consequence of self-similarity law – the same species richness and preemption coefficient of initial communities from which dilution was obtained calls for almost the same sample size regardless of the density of individuals in soil dilution or density of individuals in soil.

2. Sample size for determining species composition up to 100% increases more rapidly with rise of community species richness and coefficient of preemption.