Diversity of diversity indices: Which diversity measure is better?

O. M. Kunakh*, A. M. Volkova*, G. F. Tutova**, O. V. Zhukov**

*Oles Honchar Dnipro National University, Dnipro, Ukraine
**Bogdan Khmelnitsky Melitopol State Pedagogical University, Melitopol, Ukraine

The article evaluates the dependence of the most common indices of species diversity on sample size and determines their ability to differentiate between different types of ecosystems, with a special emphasis on discriminating between natural and anthropogenic ecosystems. An approach to adjusting the indices to reduce their dependence on sample size was also proposed. The study was conducted in seven types of ecosystems: four were natural and three were anthropogenically transformed. Samples of soil animals were selected in 2011–2013 and 2021 using the same methods. A total of 20,518 soil animal specimens belonging to 202 species were collected in all study locations. The null alternative was generated by randomly selecting samples containing 2, 3, ..., 110 soil animals from the combined soil animal sample. For each gradation of sample size, 200 sample variants were formed. The density of soil macrofauna in natural ecosystems ranged from 3.6 ± 1.5 to 15.2 ± 6.9 specimens per sample, and in artificial ecosystems – from 13.2 ± 7.6 to 21.0 ± 11.9 specimens per sample. The number of species ranged from 22–80 species, and in artificial ecosystems it was 38–99 species. Indicators of species diversity correlated with each other. A high level of correlation was observed between indicators within groups of indices: indices of species richness and indices of heterogeneity and evenness. Fisher's log-series alpha and the fundamental parameter of biodiversity were highly correlated with each other, as well as with the Margalef, species richness, and Chao's species abundance indices. The log-normal distribution best describes the dominance patterns in terms of abundance in the natural ecosystems, and the Zipf-Mandelbrot distribution best describes the dominance patterns in terms of abundance in the artificial ecosystems. Diversity indices were ordered in the space of two dimensions, one explaining the variation between ecosystems and the other depending on sample size. The ordering of the traditional indices showed that there is a vacancy for the best index in the sense that such an index should best explain differences between ecosystems and differences between natural and artificial ecosystems. It should also be independent of sample size. The Simpson heterogeneity index and the Simpson evenness index were the best of the traditional indices, but they did not explain differences between ecosystems very well, especially when it came to distinguishing between natural and artificial ecosystems. The Margalef index, which is supposed to be independent of sample size, on the other hand, showed a very high level of dependence. Such a dependence was also found for the Menhinick index, though to a lesser extent. Obviously, size dependence negatively affects the differential ability of the indices. The corrected indices of species richness and the Shannon index are practically independent of sample size and have a greater ability to differentiate ecosystems by the level of diversity, with natural ecosystems characterized by consistently higher values of the corrected indices than artificial ecosystems. The dependence on the sample size makes indices from different ecosystems practically incomparable, which makes their use meaningless. Even minor differences in sample size can lead to significant deviations in the values of diversity indices. The application of the Michaelis-Menten model allowed us to suggest a method of correction of species richness indices and the Shannon index. After the correction, the indices are practically independent of the sample size, and their differential ability to characterize individual ecosystems and the level of anthropogenic transformation increases significantly.

Keywords: species richness; Shannon index; evolution; succession; urban park; recultivation; ecosystem comparison.

Introduction

Understanding the reasons for the differences in species richness between territories is a crucial ecological issue (Margalef, 1958). The number of species varies between communities of living organisms and it is difficult to explain, so there are a lot of reasonable hypotheses to interpret patterns of diversity (Palmer, 1994). Some communities include a large number of species, while in other communities individuals are distributed among only a few species, and therefore, given equal abundance in the compared samples, the number of individuals of each species is greater (Lawton, 1999). Diverse living conditions support a greater diversity of species (Stein et al., 2014). Species diversity is controlled by resource quality, which consists of habitat factors and resources that determine the "number of niches" (Dietrich et al., 2021). Resource and consumer abundance is an important driver of diversity, consisting of factors that determine the amount of resources (e.g., area and productivity) and factors that determine the number of consumer individuals (especially body size) that share those resources (Laan & de Polavieja, 2018). Interactions between

species can increase or decrease species diversity by affecting the number of individuals or their adaptability to habitat conditions (Thébault & Loreau, 2005). Dynamic processes such as extinction, immigration and speciation affect species diversity in both equilibrium and non-equilibrium situations (Diamond, 1988). Extreme conditions contribute to the formation of communities consisting of a small number of species, where each of them is represented by many individuals (Merino et al., 2019).

The number of species and the evenness of their abundance are components of the species diversity of a community (Ma, 2005). The causes of the uneven number of species in a community are complex (Jost, 2010). The distribution of species by abundance may seem simple in form (Gaston, 1996), but the different numbers of individuals of different species and their consistent pattern of correlation are the result of the integration of a large number of factors and processes (Whittaker, 1965). The mechanisms of community structure formation also affect the uneven distribution of individuals of different species (Zang et al., 2022). Food chains allow us to identify the separate trophic levels in the community, which can be represented by a pyramid (Elton's pyramid) (Cohen & Łuczak,

1992), because the number of individuals of large species at higher trophic levels is lower than that of species at lower levels (Bonnaffé et al., 2021). Smaller species at lower levels of the trophic pyramid are able to coexist better together, competing more intensively than species at higher levels (Ceia et al., 2023). Habitat complexity is one of the most important factors that structure biotic communities. In each habitat of a certain complexity, the relative size of niches or habitats follows a distribution in which smaller habitats are more abundant and larger habitats are less abundant (Kovalenko et al., 2012). The size of available storage facilities affects the distribution of species of different sizes (Wilson et al., 1995). The restriction on dispersal decreases with increasing body size of animals (Bailey et al., 2018). The distribution of species adapted to certain conditions corresponds to the structure of the habitat, and the number of corresponding individuals reflects the relative size of microhabitats (Cramer & Willig, 2005). The sample from the community includes what represents the main habitat, as well as fragments of other communities represented by individuals migrating outside the native habitat (Rybicki et al., 2020). The same holds true for temporal dynamics. At any given time in the annual population sequence, there is a core of dominant species plus other, less abundant species that represent residues of previous populations or seeds of future populations (Zhukov et al., 2021).

The diversity patterns of similar species groups can be explained by the theory of neutral diversity (Thomas & Foin, 1982). The theory of neutral diversity predicts that very diverse communities of equivalent species arise because random extinctions are balanced by speciation (Kopp, 2010). The assumption of equivalence of adaptation, combined with stochastic or random processes that include death, immigration from the regional species pool, and speciation, can explain the formation of species-rich communities (Milanesi et al., 2020). Random changes in the number of species over time form an ecological drift (Gilbert & Levine, 2017). Random death, dispersal, and speciation are important features of the neutral theory of biodiversity, but its key, essential feature is the assumption of the ecological identity of individuals (Etienne et al., 2007). The theory of neutral diversity assumes that species may have differences, but these differences do not matter because all individuals have the same adaptability and relate to each other in the same way (Zhou & Zhang, 2008). Dispersion processes, stochasticity, and speciation are important, though not unique to neutral theory (Chave, 2004). The processes of dispersal and speciation in neutral models can explain the formation of very diverse communities (Doncaster, 2009). However, the coexistence of species in neutral communities is unstable because there are no mechanisms that force one species to remain dominant or prevent the extinction of rare species (Alonso & McKane, 2004). Since neutral processes are driven by random effects, and since all individuals are competitively identical, their numbers either increase or decrease by random chance (Hubbell, 2001). A stable coexistence, or long-term persistence of a species, is possible in a closed system only when there are niche differences that make individuals compete with individuals of their own species more strongly than with individuals of other species (Zhukov et al., 2018). Niche-based stabilization mechanisms restrict the growth of species when they become too abundant, while stabilization mechanisms allow rare species to increase because they have higher adaptability in the surroundings of neighbors of other species with which they compete less strongly (Julliard et al., 2006). The differential reproduction and differences in the rate at which populations of competing species approach competitive equilibrium are also the causes of unequal numbers of species in the community (Huston, 1979). The reproductive ability and mortality of species depends on an unlimited number of factors (Aubier, 2020). The reproductive and mortality rates of each species are balanced in a community in equilibrium (DeAngelis & Waterhouse, 1987), and there is no correlation between the rate of reproduction and the abundance of the species (Petraitis et al., 1989).

Historically, many biologists believed that evolution and ecology acted independently because evolution occurred over distances too large to affect most ecological patterns. Nowadays, there is evidence to suggest that evolution can operate at different spatial scales, including small spatial ones. Thus, evolutionary divergence in space can often interact with mechanisms that also determine spatial ecological patterns (Urban et al., 2020). Evolutionary effects also affect the uneven distribution of species

abundance in a community (Lu et al., 2019). A biotic community represents, on a small scale, a pattern of natural survival of species (Gregory, 2009). If there is a resource gradient in a community, species evolve to utilize different parts of this gradient. Thus, competition between them is reduced (Whittaker, 1972). In nature, few species are represented by many individuals that are widely distributed, but there are many more localized and restricted species represented by a small number of individuals. This pattern is a consequence of the normal course of evolution. Species with a small number of individuals form smaller populations or decline to a small number of individuals over time, which gives more opportunity for isolation and genetic change. An evolution is a self-accelerating process. The number of species grows exponentially when they are subjected to a trend of progressive specialization and differentiation. The result is a characteristic distribution that is preserved by the probability of extinction, which increases when a species becomes very rare (Margalef, 1958).

The patterns of distribution of individuals among species in a community have been described by several empirical mathematical expressions. The most famous expressions are the "geometric series" or the "niche preemption" model of Motomura (1932), "logarithmic series" by Fisher, Corbett and Williams (Fisher et al., 1943), Preston's "log-normal distribution" (Preston, 1948), MacArthur's broken stick model (Pielou, 1975), the Zipf model (Frank, 2019) and the Zipf-Mandelbrot model (Spatharis & Tsirtsis, 2013). The broken stick metaphor provides a null model in which individuals are randomly distributed among the observed species. The model has no fitted parameters (Pielou, 1975). The preemption coefficient is the only parameter of the Motomura model that shows the rate of decline in numbers by rank (Motomura, 1932). Fisher's logarithmic series model has one parameter, alpha, or also called Williams' alpha, which can be used as an indicator of diversity (Fisher et al., 1943). Fisher's alpha is often considered to be the most reliable measure of alpha diversity (Magurran, 2004). The distribution of resources between species in a community can be described by three types of curves: (a) the preference for a strongly dominant niche expressed as a geometric series, (b) random boundaries between niches expressed as a MacArthur distribution, and (c) the assignment of relative importance to multiple factors, so that species form a frequency distribution based on the logarithmic basis of importance values, i.e., a log-normal distribution. The shapes of the importance-value curves do not allow for a clear conclusion about resource distribution, but are of interest for their expression of species relationships and their impact on diversity measurement (Whittaker, 1972). Preston (1948) was not satisfied with Fisher's model, which assumed infinite species richness, and postulated that rare species were a declining class, with most species in the middle of the frequency scale. This was achieved by breaking down high-frequency classes into increasingly broad "octaves" with doubled class boundaries: 1, 2, 3–4, 5–8, 9–16, etc. occurrences. Obviously, Preston considered the frequencies 1, 2, 4, etc. "bound" between octaves (Williamson & Gaston, 2005). This means that only half of the species with a frequency of 1 are shown in the lowest octave, and the rest are moved to the second octave. Half of the species from the second octave are also moved to the highest octave, but usually not as many species. This practice makes the data more lognormal by reducing the usually high lowest octaves.

For any of the models, the distribution of individuals by species is associated with a certain correspondence between the total number of individuals and the total number of species (Waldock et al., 2022). The change in the relationship between the number of species and the number of individuals with increasing sample size depends on the type of distribution (Matthews & Whittaker, 2015). The characteristic parameter for each community can be calculated on the basis of a specific expression and by finding a correspondence between the natural community and the theoretical distribution, which expresses an intimate description of its structure, manifested in the way individuals are distributed by species (Margalef, 1958). This parameter is called the "diversity index", and the first requirement it must satisfy is that it should not depend on the sample size (Soetaert & Heip, 1990). Diversity indexes reflect the richness of species (Pielou, 1969). It is high in communities that include a large number of species and in which the number of individuals of each species decreases relatively slowly in the transition from more common to less common (Magurran, 2004). The diversity index is low in communities with a small

number of species, where the number of individuals per species decreases rapidly as transition from dominant species to successively less important ones (Margalef, 1958).

Species diversity consists of two components: species richness and species evenness, while species richness is a part of species diversity (Nijs & Roy, 2000). Evenness of species diversity measures how evenly species are represented on the territory (Pielou, 1975). A fundamental weakness of many diversity indices is their dependence on sample size (Sanders, 1968), which makes comparisons between studies difficult. However, the main purpose of quantifying diversity using a numerical index is to provide a means of comparison between different communities. One way to avoid incomparability of measurements obtained on samples of different sizes was proposed by the Sanders rarefaction method (Sanders, 1968). This method calculates the number of species that can be expected from each sample if the sample size is standardized. The rarefaction method was shown to overestimate the expected number of species and introduced an exact calculation formula for this indicator. However, a simple count of the number of species does not capture all the information available in the community because it does not relate to how individuals are distributed among species.

Species diversity can only be meaningfully compared between communities that have been sampled equivalently in some sense (Chao & Jost, 2012). There are several ways to standardize samples, and the choice of standardization method can greatly affect the results. There are three main approaches to standardizing samples to obtain comparable results: equalizing effort, equalizing sample size, and equalizing coverage (Roswell et al., 2021). Two factors determine how well a sample reflects the true diversity of a community: how carefully it is selected and how many species are in it and in what relative abundance (Chao, 1987). The main problem with equal-effort sampling is that sample size usually varies across communities for the same effort, and sample size partly determines how well the observed abundance distribution matches the true abundance distribution of species in the community. For example, a small sample is likely to contain only a few species, all of which are common. When the sample contains more individuals, the number of species increases and the diversity of the sample increases (Preston, 1948). The second way ecologists standardize samples is through sample size. The rarefaction is the removal of individuals from larger samples until all samples contain the same number of individuals (Stier et al., 2016). However, rarefaction also does not ensure sampling bias because it still does not take into account the distribution of relative abundance throughout a large community (Willis, 2019). The standardization of sample size leads to a greater underestimation of diversity for more diverse communities (Chao & Jost, 2012).

The number of species in a sample is a very imperfect measure of diversity. There is a view that the number of species in a community as an estimate of possible similarity between two different communities is not correct (Margalef, 1958). Species richness is closely related to the number of individuals in a sample, especially in the early stages of sampling. As the sample size increases, the accumulation curves representing different communities often overlap (Lande et al., 2000). This means that the relative richness of two communities measured with a smaller sample size does not predict their relative richness with a larger sample size correctly (Haegeman et al., 2013). This is the case even when estimates such as Chao1 are used to predict true diversity. Species richness is very sensitive to sample size and relative abundance, and its estimate may depend on how the samples are standardized. Even the best asymptotic richness estimates, such as Chao1 (Gotelli & Colwell, 2011), cannot reliably predict the true diversity of a community (Chao & Jost, 2015). The problem is that both sample richness and sample-based richness estimates are strongly influenced by the rarest species. The species richness is not recommended by any of the authors who have systematically tested diversity metrics (Haegeman et al., 2013), because it is difficult to estimate accurately outside of an experimental context. The sample richness varies dramatically depending on the sample size and the method of sample equalization. This is because this index is very sensitive to the rarest species. The same problem applies to asymptotic estimates of richness (Chao & Jost, 2015). Species richness is best used in special cases, such as when the community is fully known, or perhaps when there is sufficient information to parameterize the occupancy model (Guillera-Arroita et al., 2019).

The Margalef index is the simplest, and by no means the weakest, of the various diversity indices (Margalef, 1958). The index is based on the assumption of a linear relationship between the number of species and the logarithm of the area or number of individuals. Despite attempts to adjust for sample size, the Margalef and Menhinick indices remain strongly influenced by sample size (Cazzolla Gatti et al., 2020). The data for the Margalef index should be organized as absolute numbers, not as a matrix of density data (Gamito, 2010). Nevertheless, these are intuitive indices and can play a useful role in biodiversity research. (van Loon et al., 2018).

Ecologists most often measure the species diversity of communities using the traditional Shannon and Simpson indices (Magurran & McGill, 2011). Traditional diversity indices, which explicitly include the relative abundance (Magurran & McGill, 2011), such as the Shannon (Shannon & Weaver, 1949) and Simpson (Simpson, 1949) indices, are more resistant than richness indices to the sampling problems described above. However, their use creates a new set of problems: these indices have different units of measurement and do not scale intuitively with increasing and decreasing species abundance (Tuomisto, 2010b). These problems have led to the assumption that diversity lacks a conceptual framework (Hurlbert, 1971). The first problem with traditional diversity indices is that they measure very different things (Tuomisto, 2010a). Species richness, of course, measures the number of species. The Shannon index measures uncertainty about the identity of species in a sample, and its units measure quantitative information in bits (Hurlbert, 1971). The Gini-Simpson index measures the probability that two individuals randomly drawn from a sample will belong to different species (Hurlbert, 1971). Because species richness, the Shannon index, and the Gini-Simpson index measure different metrics, it is particularly difficult to justify choosing one of them to represent diversity. Another problem is that the Shannon and Gini-Simpson indices behave in ways that do not make sense for a diversity metric. Obviously, significant changes in the number of species in a community are reflected in disproportionate changes in the indices (Roswell et al., 2021).

The Simpson diversity index may be a good choice for a research question that is mainly concerned with patterns of relative abundance of common species, requires confidence that the expected diversity will not change significantly with additional sampling, or concerns the probability that two randomly selected individuals are of the same species (Hurlbert, 1971). The inverse scale used to calculate Simpson's diversity spreads low values of rarity in different directions and brings high values closer together. Thus, Hill-Simpson diversity is the most sensitive to differences in low rarity values, i.e., the relative abundance of common species. The expected value of Simpson's sample diversity tends to be robust to sample standardization and changes little with increasing sample size. In addition, true Simpson's diversity can be estimated with a small error (Grabchak et al., 2017).

Shannon's diversity is between that of Simpson's and the number of species in terms of its properties and can be the "right" indicator for many applications (Kempton, 1979). The index can react strongly to both very high and very low values of rarity. Another argument in favour of the Shannon index is that the distribution of many species is approximately lognormal (Williamson & Gaston, 2005), and hence their central tendency can be well described by the geometric mean, which forms the basis for calculating the index itself. The observed Shannon diversity begins to stabilize at achievable sample sizes, and asymptotic estimates for Shannon diversity work quite well (Beck & Schwanghart, 2010). Shannon's diversity retains some of the sensitivity of the species richness index, and also retains the robustness to sample size and sampling standardization that is characteristic of Simpson's diversity. As a result, the Hill-Shannon method can be a good choice for characterizing biodiversity gradients in an ecologically meaningful way (Roswell et al., 2021).

Neutral models can only predict realistic patterns of species diversity in a few ways (Chisholm & Pacala, 2010). One of the parameters is the fundamental parameter of biodiversity, which is higher the greater the number of individuals in the meta-community and the higher the rate of speciation (He & Hu, 2005). Neutral models are able to predict important biodiversity parameters, such as the number of species and their relative abundance in different systems, based on the fundamental parameter of biodiversity and the assessment of dispersal (Missa et al., 2016). An analytical link was established between the fundamental parameter of biodiver-

sity and the well-known Simpson diversity index. This link helps to connect the evolutionary aspect of biodiversity with the ecological and statistical aspect of diversity (He & Hu, 2005).

A unified method for measuring diversity was developed by Hill (1973) and reintroduced to ecologists by Jost (2006). This method assumes that both the number and relative abundance of species are components of diversity, and that these components cannot be completely separated from each other. The diversity metric developed by Hill (1973), consists of a single equation that, depending on the value of a single parameter, can vary from counting all species equally, even if they are endangered, to heavily emphasizing the most common species. Hill's diversity behaves in a way that is logically justified for a measure of diversity (Hurlbert, 1971).

There is no clear answer to the question of which diversity index is best to use (Roswell et al., 2021). As Southwood (1977) noted: "There can be no universal "best option", although there is ample scope for misuse". Thus, resistance to sample sizes is applied to assess the advantage of diversity indices. The dependence of a diversity index on sample size does not allow for the use of diversity itself as a criterion for comparing different ecosystems. However, the differential ability of the index remains an underestimated criterion for characterizing the degree of species diversity. Therefore, the aim of our article was to evaluate the dependence of the most common species diversity indices on sample size and determine their ability to differentiate between different types of ecosystems, with a special emphasis on distinguishing between natural and anthropogenic ecosystems, and to propose an approach to correcting the indices to reduce their dependence on sample size.

Material and methods

Ecosystems studied. The research was conducted in seven types of ecosystems: four were natural, three were anthropogenically transformed (Fig. 1). In the ecosystem on reclaimed lands (47°39'01" N 34°08'36" E), soils were represented by technosols on red-brown clays (105 samples), grey-green clays (105 samples), loess-like loams (105 samples), and with a top humus layer (105 samples) (Zhukov & Gadorozhnaya, 2016). All samples were considered as a combined sample in the article. The vegetation cover was represented by 135 species of herbaceous plants, dominated by *Bromus squarrosus* L., *Medicago sativa* L., *Lactuca tatarica* (L.) C. A. Mey., *Lactuca tatarica* (L.) C. A. Mey.

The broadleaf forest was studied on the right bank of the Samara River (48°45'56" N 35°26'00" E). The soil was represented by Luvic Chernozem (Siltic, Hyperhumic, Pachic) (Yakovenko & Zhukov, 2021). The vegetation cover was represented by an oak forest. The plant community was represented by 35 species. The tree stand was dominated by *Quercus robur* L., *Acer campestre* L., *Fraxinus excelsior* L. The understory was dominated by *Euonymus europaeus* L., *Eu. verrucosus* Scop., *Crataegus fallacina* Klokov, *Acer tataricum* L. The herbaceous layer was dominated by *Viola odorata* L., *Aegopodium podagraria* L., and *Glechoma hederacea* L.

The psammophytic steppe was located in the Dnipro-Orel Nature Reserve (48°30'47" N 34°49'35" E). The soil was represented by Eutric Arenosol (Aeolic, Ochric, Thaptoochric). The herbaceous vegetation cover was represented by 57 plant species, among which *Artemisia campestris* L., *Festuca beckeri* (Hack.) Trautv., *Secale sylvestre* Host. prevailed. The moss cover was represented by *Syntrichia ruralis* (Hedw.) F. Weber & D. Mohr.

The steppe on forest-like loams was located on the slope of Kamianuvasta gully (48°21'51" N 35°08'09" E). The soil was represented by Calcic Chernozem (Siltic, Tonguic). The herbaceous vegetation cover was represented by 106 plant species among which *Stipa capillata* L., *Silene wolgensis* (Hornem.) Otth, *Euphorbia stepposa* Zoz. ex Prokh., *Teucrium polium* L., *Festuca valesiaca* Schleich. ex Gaudin, *Galatella villosa* (L.). prevailed.

The meadow ecosystem was studied in the Dnipro-Orel Nature Reserve (48°31'25" N 34°49'19" E) (Tutova et al., 2022). The soil was represented by Calcic Mollic Gleyic Solonetz (Fluvic, Humic, Loamic). The herbaceous vegetation cover was represented by 91 plant species, dominated by *Aster sedifolius* L. subsp. *dracunculoides* (Lam.) Merxm., *Festuca*

valesiaca Schleich. ex Gaudin, *Limonium gmelinii* (Willd.) Kuntze, *Ranunculus repens* L., *Taraxacum campylodes* G. E. Haglund.

The city park after reconstruction was studied in the Botanical Garden of Oles Honchar Dnipro National University (48°25'55" N 35°02'41" E). The soil was represented by Calcic Chernozem (Siltic, Tonguic). The vegetation cover was represented by 57 plant species. The stand was dominated by *Robinia pseudoacacia* L., *Acer platanoides* L., *Fraxinus excelsior* L. Undergrowth was practically absent. The herbage was dominated by *Geum urbanum* L., *Taraxacum campylodes* G. E. Haglund, *Chelidonium majus* L., *Galium aparine* L., *Stellaria media* (L.) Vill.

The city park without reconstruction was studied in the Botanical Garden of Oles Honchar Dnipro National University (48°25'58" N 35°02'28" E). The soil was represented by Calcic Chernozem (Siltic, Tonguic). The vegetation cover was represented by 47 plant species. The stand was dominated by *Acer platanoides* L., *A. negundo* L., *Robinia pseudoacacia* L. The understory was dominated by *Ulmus laevis* Pall., *Clematis vitalba* L., *Parthenocissus quinquefolia* (L.) Planch. The grass stand was dominated by *Geum urbanum* L., *Viola odorata* L., *Impatiens parviflora* DC., *Galium aparine* L., *Stellaria media* (L.) Vill.

Assessment of the macroinvertebrate community. Soil animals were sampled in the 2011–2013 and 2021 sampling series using the same methods. Macrofauna was collected according to ISO 23611-5 (Anderson & Ingram, 1993). Each sampling plot consisted of 105 sampling points. The points were located along 7 transects of 15 points each. The distance between points in a transect, as well as the distance between transects, was 3 m. At each of the 4 sampling points, soil blocks measuring 25 × 25 cm and 30 cm deep were collected. Macroinvertebrates visible to the naked eye were collected manually by sorting through leaf litter and soil. They were preserved in 75% alcohol, and earthworms in 4% formaldehyde. Animals were identified to the species level, if possible. The results of quantitative studies were presented as the number of individuals per sample with a soil surface area of 25 × 25 cm (Zhukov et al., 2023).

Assessment of species richness indicators. Assuming that S is the number of species in the sample, $N = \sum N_i$ is the total number of individuals in the community, *i* is the rank of the taxon (conditional ordinal number), N_i is the number of individuals of the *i*-th taxon, $p_i = N_i/N$ is the relative proportion of the *i*-th taxon in the community, then the diversity indices can be calculated as follows.

Magalef index (Margalef, 1958):
$$R_1 = \frac{S-1}{\ln N}.$$

Menhinick index (Menhinick, 1964):
$$R_2 = \frac{S-1}{\sqrt{N}}.$$

Estimation of the real number of species in the Chao community (Chao, 1982; Chao & Lee, 1992):
$$S_{Chao1} = \begin{cases} S + \frac{N-1}{N}\frac{s_1^2}{2s_2^2} & if \ s_2 > 0 \\ S + \frac{N-1}{N}\frac{s_1(s_1-1)}{2} & if \ s_2 = 0 \end{cases}$$

Assessment of indicators of community heterogeneity. Shannon index (Shannon-Wiener, Shannon-Wiener-Weaver) (Shannon, 1948):
$$H = -\sum_{i=1}^{S} p_i log_2 p_i.$$

The diversity index was proposed by Gini (1912) and adapted for biological research by Simpson (1949), so it is most often called the Gini-Simpson index (Guiasu & Guiasu, 2010):
$$D_{S1} = 1 - \sum_{i=1}^{S} p_i^2$$

Brillouin diversity index (Brillouin, 1956):
$$H = \frac{log_2(N!) - \sum_{i=1}^{S} log_2(n_i!)}{N}.$$

Berger-Parker diversity index (Berger & Parker, 1970):
$$D_{BP} = \frac{1}{p_{max}}.$$

Indicators of community evenness. Shannon evenness (Pielou, 1975):
$$E = \frac{H}{H_{max}} = \frac{H}{log_2 S}.$$

Simpson's evenness index:
$$E = \frac{1}{S\sum_{i=1}^{S} p_i^2}.$$

Fig. 1. Ecosystem types studied: *a* – reclaimed land; *b* – deciduous forest; *c* – steppe; *d* – sandy steppe; *e* – meadow; *f* – urban park after reconstruction; *g* – urban park without reconstruction

The McIntosh index (McIntosh, 1967) expresses sample heterogeneity in geometric terms. It describes a sample as a point in an S-dimensional hypervolume and uses the Euclidean distance of that point from the origin.

Dominance:

$$D = \frac{N - U}{N - \sqrt{N}}.$$

Evenness:

$$D = \frac{N - U}{N - \frac{N}{\sqrt{S}}}.$$

where U is the distance of the sample from the origin in the S-dimensional hypervolume:

$$U = \sqrt{\sum_{i=1}^{s} n_i^2}.$$

The fundamental biodiversity parameter θ was defined as Hubbell (2001):

$$\theta = 2J\rho,$$

where J is the number of total individuals in the (meta)community and ρ is the speciation rate.

Statistical calculations. The indices of species diversity were calculated using the tabula package (Frerebeau, 2019) for the R statistical software. The fundamental biodiversity parameter θ was calculated using the untb package (Hankin, 2007). The Fisher's log-series parameter α and RAD model parameters were calculated using the vegan package (Oksanen et al., 2019). The null alternative was generated by randomly selecting samples containing 2, 3, ..., 110 soil animal individuals from the pooled sample of soil animals. For each gradation of sample size, 200 sample variants were generated.

Results

A total of 20,518 specimens of soil animals belonging to 202 species were collected in all studied locations. The density of soil macrofauna in natural ecosystems ranged from 3.6 ± 1.5 to 15.2 ± 6.9 ind. per sample, and in artificial ecosystems the density ranged from 13.2 ± 7.6 to 21.0 ± 11.9 ind. per sample (Table 1). In natural ecosystems, the number of species varied in the range of 22–80 species, and in artificial ecosystems this value was 38–99 species.

The indices of species diversity were correlated with each other (Table 2). There was a high level of correlation between indicators within the groups of indices: the indices of the species richness, and the indices of heterogeneity and evenness. Fisher's log-series alpha and the fundamental parameter of biodiversity were strongly correlated with each other, as well as with Margalef indices, species richness, and Chao's estimate of species number. The log-normal distribution and the Zipf-Mandelbrot distribution were the best alternatives among the Rank–abundance dominance (RAD) models (Table 3). The log-normal distribution best described the RAD models of natural ecosystems, and the Zipf-Mandelbrot distribution was the best for describing the RAD models of artificial ecosystems (Fig. 2).

Table 1
Indicators of the abundance and diversity of communities

Parameter	Reclaimed soil	Deciduous forest	Steppe	Sand Steppe	Meadow	Reconstructed park	Undamaged park
Number of the samples and abundance indicators							
Number of the samples	420	210	105	105	210	210	210
Number of the individuals	5535	3122	853	375	2972	3259	4402
Community density (ind./sample)	13.2 ± 7.6	15.2 ± 6.9	5.3 ± 3.1	3.6 ± 1.5	14.2 ± 8.4	15.5 ± 10.0	21.0 ± 11.9
Indicators of species richness							
Margalef diversity index	11.37	7.58	4.89	3.54	9.88	4.57	5.36
Species richness	99	62	34	22	80	38	46
Menhinick diversity index	1.33	1.11	1.16	1.13	1.47	0.67	0.69
Chao species richness	133.03	74.10	42.09	22.50	85.33	40.00	53.14
Indicators of heterogeneity							
Shannon diversity index	2.18	2.85	2.41	2.47	3.40	1.75	1.60
Simpson diversity index	0.79	0.91	0.86	0.89	0.94	0.73	0.69
Brillouin diversity index	2.15	2.81	2.34	2.37	3.35	1.73	1.58
Berger-Parker diversity index	2.83	5.40	4.20	5.19	5.46	2.48	2.17
Indicators of evenness							
Shannon evenness index	0.47	0.69	0.68	0.80	0.78	0.48	0.42
Brillouin evenness index	0.47	0.69	0.68	0.80	0.78	0.48	0.42
McIntosh evenness index	0.61	0.79	0.75	0.85	0.85	0.57	0.52
Simpson evenness index	0.05	0.17	0.21	0.42	0.21	0.10	0.07
Other diversity indicators							
α parameter of Fisher's log-series	17.12	10.96	7.08	5.09	15.13	6.04	7.16
Fundamental Biodiversity parameter θ	17.02	10.85	6.94	4.94	15.02	5.94	7.07
γ Zipf-Mandelbrot	–2.5	–3.2	–120210758	–379.6	–2.4	–2.9	–3.2
ß Zipf-Mandelbrot	2.1	9.1	515223563	1737.7	11.0	2.2	2.2

Table 2
Spearman correlations between diversity indices (only coefficients statistically significant for P < 0.05 are shown)

Diversity indices*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00	0.99	–	0.97	–	–	–	–	–	–	–	–	1.00	1.00
2	0.99	1.00	–	0.98	–	–	–	–	–	–	–	–	0.98	0.98
3	–	0.57	1.00	–	0.83	0.79	0.82	–	–	–	–	–	–	–
4	0.97	0.98	–	1.00	–	–	–	–	–	–	–	–	0.96	0.96
5	–	–	0.83	–	1.00	0.95	1.00	0.90	0.83	0.83	0.88	–	–	–
6	–	–	0.79	–	0.95	1.00	0.94	0.98	0.93	0.93	0.97	–	–	–
7	–	–	0.82	–	1.00	0.94	1.00	0.89	0.81	0.81	0.86	–	–	–
8	–	–	–	–	0.90	0.98	0.89	1.00	0.96	0.96	0.98	–	–	–
9	–	–	–	–	0.83	0.93	0.81	0.96	1.00	1.00	0.99	0.86	–	–
10	–	–	–	–	0.83	0.93	0.81	0.96	1.00	1.00	0.99	0.86	–	–
11	–	–	–	–	0.88	0.97	0.86	0.98	0.99	0.99	1.00	0.80	–	–
12	–	–	–	–	–	–	–	–	0.86	0.86	0.80	1.00	–	–
13	1.00	0.98	–	0.96	–	–	–	–	–	–	–	–	1.00	1.00
14	1.00	0.98	–	0.96	–	–	–	–	–	–	–	–	1.00	1.00

Note: * 1 – Margalef diversity index; 2 – Species richness; 3 – Menhinick diversity index; 4 – Chao species richness; 5 – Shannon diversity index; 6 – Simpson diversity index; 7 – Brillouin diversity index; 8 – Berger-Parker diversity index; 9 – Shannon evenness index; 10 – Brillouin evenness index; 11 – Mcintosh evenness index; 12 – Simpson evenness index; 13 – α parameter of Fisher's log-series; 14 – Fundamental Biodiversity parameter θ.
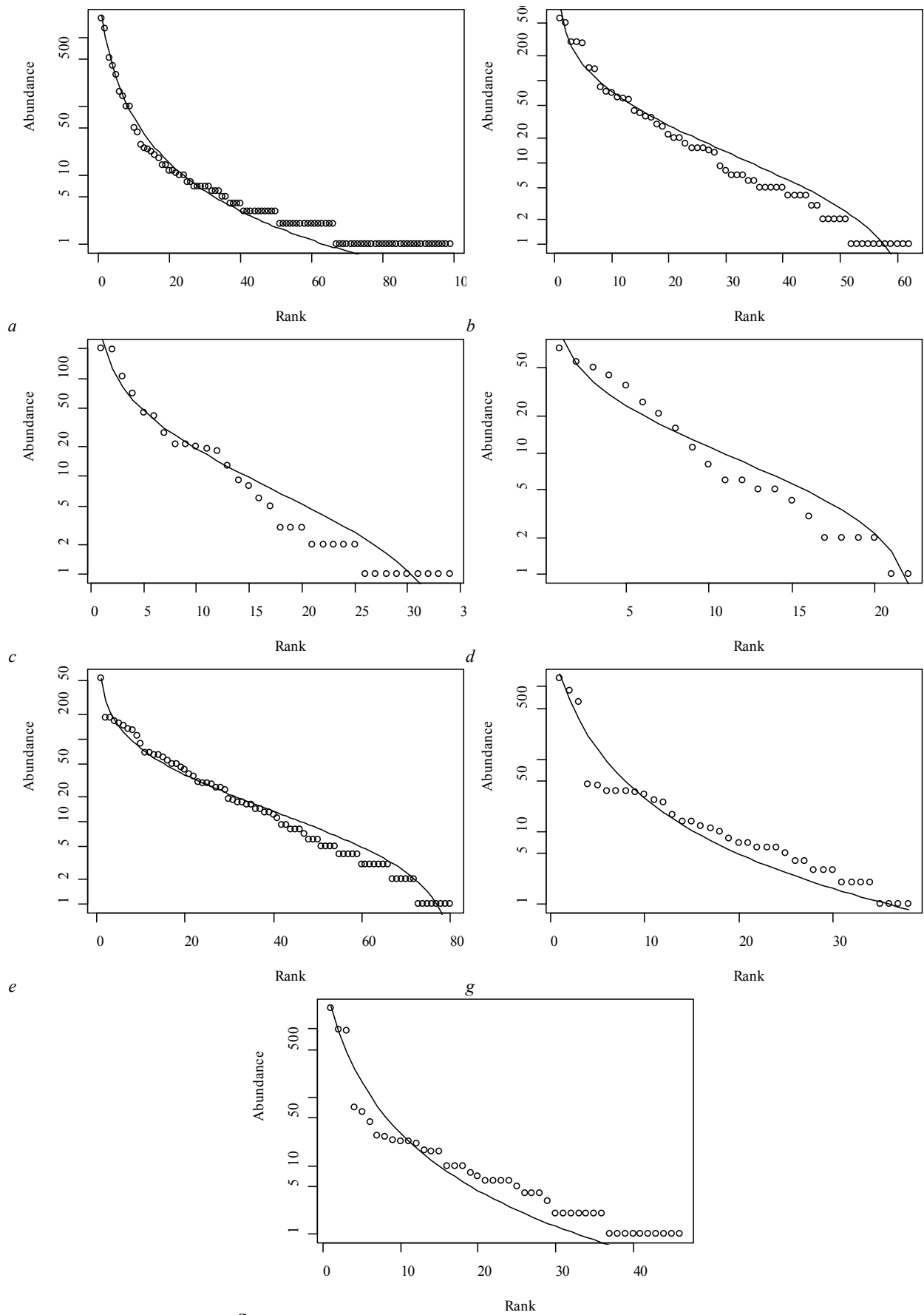
Fig. 2. Rank–Abundance Dominance (RAD): the abscissa is the order of species, and the ordinate is the number of species (experimental data are marked with dots); lines indicate the best models: *a* – reclaimed land (Zipf-Mandelbrot model); *b* – deciduous forest (log-normal model); *c* – steppe (log-normal model); *d* – sandy steppe (log-normal model); *e* – meadow (log-normal model); *f* – urban park after reconstruction (Zipf-Mandelbrot model); *g* – urban park without reconstruction (Zipf-Mandelbrot model)

Table 3

Deviations of observed data from Rank-abundance dominance models (bold indicates the smallest deviations corresponding to the best models)

Model		Reclaimed soil	Deciduous forest	Steppe	Sand Steppe	Meadow	Reconstructed park	Undamaged park
Null		28.9	11.1	6.4	5.6	13.3	30.8	50.5
Preemption		12.6	3.2	6.8	9.1	6.7	28.9	42.0
Log-normal		1.9	2.3	0.3	0.5	0.6	8.7	10.7
Zipf		2.2	4.6	0.9	0.6	1.8	10.0	12.1
	Zipf-Mandelbrot	**0.9**	**2.7**	**0.4**	0.8	**1.5**	**7.0**	**8.2**

The ecosystem type and sample size were able to explain 28–82% of the variation in diversity indicators (Table 4). The species richness indicators were most strongly determined by these predictors (52–82%). About 28–79% of the variation in heterogeneity indicators could be explained by ecosystem type and sample size. The evenness indicators were least dependent on the considered predictors (28–42%). The indicators of species richness and community heterogeneity increased with increasing sample size, while the indicators of community evenness decreased (Table 5). The Margalef diversity index, species richness, and Brillouin index were the most sensitive to sample size. The Simpson diversity index did not depend statistically significantly on the sample size. Also, the Simpson evenness index was the least sensitive to sample size.

Table 4

General linear model of the dependence of species diversity indices on ecosystem type (F = 396, P < 0.001) and sample size (F = 9798, P < 0.001)

Index	R_{adj}^2	Model sum of squares	Df model	Model mean of squares	Residual sum of squares	Df residual	Residual mean of squares	F-ratio	P-level
				Indicators of species richness					
Margalef diversity index	0.80	53602	8	6700	13112	23512	0.56	12014.4	<0.001
Species richness	0.82	14807	8	1851	3300	23512	0.14	13185.1	<0.001
Menhinick diversity index	0.52	4546	8	568	4116	23512	0.18	3245.7	<0.001
Chao species richness	0.57	19720	8	2465	14713	23512	0.63	3939.2	<0.001
				Indicators of heterogeneity					
Shannon diversity index	0.70	4865	8	608	2106	23512	0.09	6788.7	<0.001
Simpson diversity index	0.41	52	8	6	75	23512	0.00	2032.6	<0.001
Brillouin diversity index	0.79	5578	8	697	1505	23512	0.06	10895.5	<0.001
Berger-Parker diversity index	0.28	12503	8	1563	32463	23512	1.38	1131.9	<0.001
				Indicators of evenness					
Shannon evenness index	0.41	39	8	5	57	23512	0.00	2011.2	<0.001
Brillouin evenness index	0.42	50	8	6	69	23512	0.00	2127.9	<0.001
Mcintosh evenness index	0.28	30	8	4	75	23512	0.00	1166.9	<0.001
Simpson evenness index	0.32	416808	8	52101	890663	23512	37.88	1375.4	<0.001

Table 5

Beta regression coefficients obtained from the results of a General Linear Model of the dependence of biodiversity indices on ecosystem type and sample size

Effect	Sample size	Ecosystem types (regression coefficients calculated based on comparison with the null alternative)						
		reclaimed soil	deciduous forest	steppe	sand steppe	meadow	reconstructed park	undamaged park
				Indicators of species richness				
Margalef diversity index	0.78 ± 0.003	–0.05 ± 0.007	0.04 ± 0.008	–	–	0.07 ± 0.008	–0.12 ± 0.008	–0.17 ± 0.008
Species richness	0.78 ± 0.003	–0.02 ± 0.007	0.10 ± 0.008	–0.11 ± 0.009	–0.16 ± 0.010	0.10 ± 0.008	–0.08 ± 0.008	–0.11 ± 0.008
Menhinick diversity index	0.44 ± 0.005	–0.09 ± 0.011	0.08 ± 0.013	–	–0.04 ± 0.016	0.13 ± 0.013	–0.21 ± 0.013	–0.30 ± 0.013
Chao species richness	0.57 ± 0.004	–	0.10 ± 0.012	–0.10 ± 0.013	–0.16 ± 0.016	0.11 ± 0.012	–0.13 ± 0.012	–0.16 ± 0.012
				Indicators of heterogeneity				
Shannon diversity index	0.63 ± 0.004	–0.08 ± 0.009	0.13 ± 0.010	–0.13 ± 0.011	–0.19 ± 0.013	0.16 ± 0.010	–0.12 ± 0.010	–0.17 ± 0.010
Simpson diversity index	–	–0.42 ± 0.012	0.11 ± 0.015	0.18 ± 0.016	0.17 ± 0.018	0.29 ± 0.015	–0.40 ± 0.015	–0.55 ± 0.015
Brillouin diversity index	0.75 ± 0.003	–0.05 ± 0.007	0.10 ± 0.009	–0.13 ± 0.009	–0.17 ± 0.011	0.12 ± 0.009	–0.06 ± 0.009	–0.10 ± 0.009
Berger-Parker diversity index	0.29 ± 0.006	–0.12 ± 0.014	0.03 ± 0.016	–0.06 ± 0.017	–0.08 ± 0.020	0.15 ± 0.016	–0.13 ± 0.016	–0.17 ± 0.016
				Indicators of evenness				
Shannon evenness index	–0.63 ± 0.005	–0.38 ± 0.013	–	0.16 ± 0.016	0.26 ± 0.018	0.16 ± 0.015	–0.15 ± 0.015	–0.34 ± 0.015
Brillouin evenness index	–0.64 ± 0.005	–0.39 ± 0.012	–	0.17 ± 0.016	0.30 ± 0.018	0.16 ± 0.014	–0.17 ± 0.014	–0.36 ± 0.014
Mcintosh evenness index	–0.42 ± 0.006	–0.49 ± 0.014	–	0.18 ± 0.017	0.29 ± 0.020	0.21 ± 0.016	–0.20 ± 0.016	–0.40 ± 0.016
Simpson evenness index	–0.06 ± 0.006	–0.41 ± 0.013	–0.40 ± 0.016	0.97 ± 0.017	1.09 ± 0.020	–0.13 ± 0.016	–0.44 ± 0.016	–0.43 ± 0.016

The number of soil invertebrates found in fixed-size samples ranged from 1 to 108 specimens (Table 6). The natural and anthropogenic forest ecosystems did not differ in this indicator (planned comparison F = 1.32, P = 0.25). The herbaceous ecosystems under reclamation conditions also did not differ in sample size from the natural meadow communities (planned comparison F = 0.14, P = 0.71). The sample size in the steppe community was the smallest among all the studied ecosystem types (planned comparison F = 28.5, P < 0.001), and this indicator did not differ in the sandy steppe and in the steppe on loess-like loams (planned comparison F = 0.20, P = 0.66). The value of the Margalef index was the lowest in the city park and in reclaimed soils compared to natural ecosystems (planned comparison F = 353.1, P < 0.001). Among the natural ecosystems, the Margalef index was highest in the meadow ecosystems (planned comparison F = 32.3, P < 0.001). The steppe ecosystems were characterized by a lower Margalef index than forest ecosystems (planned comparison F = 3.30, P < 0.001). The number of species per sample was higher in forests and meadows (planned comparison F = 731.7, P < 0.001). The lowest value was in the sandy steppe and it was statistically significantly different from the steppe on loess-like loams (planned comparison F = 19.9, P < 0.001). The number of soil animal species in one sample did not differ significantly between park plantings after reconstruction and without reconstruction. The forests and meadows had the highest Menhinick index values (planned comparison F = 130.1, P < 0.001). The steppe ecosystems did not differ from each other in this indicator (planned comparison F = 0.78, P = 0.38) and reclaimed lands prevailed (planned comparison F = 6.88, P < 0.001). The lowest value of the Menhinick index was found for animal communities in the city park. The Chao species richness score indicates that the highest species richness was characteristic of meadow and forest ecosystems, with a slightly lower value in reclaimed lands. The steppe ecosystems and park plantation ecosystems differed little in terms of Chao's species richness.

Table 6
Sample size and species richness indices of soil macrofauna communities

Ecosystem type	N	Mean ± standard deviation	Minima	Maxima
Sample size, individuals per sample of 0.25 × 0.25 meter in size				
Reclaimed soil	420	13.2 ± 7.6	2	68
Deciduous forest	205	15.2 ± 6.9	4	41
Steppe	161	5.3 ± 3.1	2	18
Sand Steppe	105	3.6 ± 1.5	2	9
Meadow	210	14.1 ± 8.5	2	58
Reconstructed park	210	15.5 ± 10.0	2	108
Undamaged park	210	21.0 ± 12.0	2	84
Random	22000	55.5 ± 31.8	1	110
Total	23521	52.8 ± 32.5	1	110
Margalef diversity index				
Reclaimed soil	420	1.6 ± 0.6	0.4	3.9
Deciduous forest	205	2.1 ± 0.6	0.6	4.0
Steppe	161	1.5 ± 0.4	0.7	2.6
Sand Steppe	105	1.4 ± 0.5	0.5	2.7
Meadow	210	2.3 ± 0.7	0.0	5.1
Reconstructed park	210	1.3 ± 0.5	0.0	3.0
Undamaged park	210	1.2 ± 0.4	0.4	2.3
Random	22000	5.0 ± 1.5	0.0	9.2
Total	23521	4.8 ± 1.7	0.0	9.2
Species richness				
Reclaimed soil	420	2.5 ± 0.4	1.6	3.6
Deciduous forest	205	2.9 ± 0.5	1.6	3.9
Steppe	161	2.0 ± 0.3	1.6	2.8
Sand Steppe	105	1.8 ± 0.3	1.6	3.0
Meadow	210	2.9 ± 0.5	1.0	4.3
Reconstructed park	210	2.3 ± 0.4	1.0	3.3
Undamaged park	210	2.4 ± 0.4	1.6	3.3
Random	22000	4.3 ± 0.8	1.0	5.5
Total	23521	4.2 ± 0.9	1.0	5.5
Menhinick diversity index				
Reclaimed soil	420	1.4 ± 0.4	0.5	3.2
Deciduous forest	205	1.8 ± 0.4	0.6	3.0
Steppe	161	1.5 ± 0.3	0.7	2.3
Sand Steppe	105	1.4 ± 0.3	0.7	2.3
Meadow	210	1.9 ± 0.4	0.6	3.3
Reconstructed park	210	1.2 ± 0.4	0.5	2.5
Undamaged park	210	1.0 ± 0.3	0.4	1.9
Random	22000	2.9 ± 0.5	0.7	4.4
Total	23521	2.8 ± 0.6	0.4	4.4
Chao species richness				
Reclaimed soil	420	3.0 ± 0.8	1.6	5.7
Deciduous forest	205	3.5 ± 0.8	1.6	5.8
Steppe	161	2.4 ± 0.6	1.6	4.1
Sand Steppe	105	2.1 ± 0.6	1.6	3.8
Meadow	210	3.5 ± 0.8	1.0	5.9
Reconstructed park	210	2.6 ± 0.6	1.0	4.7
Undamaged park	210	2.6 ± 0.6	1.6	4.3
Random	22000	5.5 ± 1.0	1.0	8.6
Total	23521	5.3 ± 1.2	1.0	8.6

The steppe communities were in the zone of the zero alternative according to the Margalef index (Fig. 3), which was also proved by the regression coefficients, which were not statistically significant. This indicator for the city park communities was the farthest from the null alternative, despite the fact that the sample size was quite high. The null alternative for the Margalef index was significantly dependent on the sample size ($r = 0.86$, $P < 0.001$). The regression model with sample size as a predictor was able to explain 75% of the variation in the Margalef index, which was calculated from a random sample. In its features, the species richness resembles the Margalef index.

The meadow and forest ecosystems were superior in terms of Shannon's index compared to other ecosystems (planned comparison $F = 871.4$, $P < 0.001$, Table 7). The reclaimed soil communities showed rather high values of the Shannon index, which exceeded the diversity indicators in the meadow and steppe natural ecosystems, as well as in the artificial forest ecosystems (planned comparison $F = 77.5$, $P < 0.001$). The diversity of steppe ecosystems was higher than that of the sandy steppe (planned comparison $F = 12.4$, $P < 0.001$), and the diversity of the soil macrofauna community in the park plantation after reconstruction was higher than in

the area that did not undergo reconstruction (planned comparison $F = 12.2$, $P < 0.001$).

The Simpson index was very sensitive to anthropogenic impact. The anthropogenic ecosystems had significantly lower diversity scores than the natural ecosystems (planned comparison $F = 2656.9$, $P < 0.001$). The meadow ecosystems had the highest Simpson's index value compared to other natural ecosystems (planned comparison $F = 45.4$, $P < 0.001$). The steppe and sandy steppe communities did not differ from each other in terms of the Shannon index (planned comparison $F = 0.0068$, $P = 0.94$). The diversity of the forest ecosystem was lower than that of the steppe ecosystem (planned comparison $F = 13.4$, $P < 0.001$).

The Brillouin diversity index indicated the highest diversity of soil macrofauna communities in the natural meadow and forest ecosystems compared to all other ecosystems (planned comparison $F = 714.5$, $P < 0.001$). The anthropogenic ecosystems occupy an intermediate position between steppe ecosystems on the one hand and forest and meadow ecosystems on the other. The differences in the Brillouin diversity index between different types of anthropogenic ecosystems are not significant.

The meadow ecosystems are characterized by the highest value of the Berger-Parker index. A slightly lower value of this index was found for forest ecosystems. The community of soil macrofauna of reclaimed soils did not differ from steppe ecosystems by the Berger-Parker index (planned comparison $F = 0.29$, $P = 0.58$), as well as from park plantations (planned comparison $F = 1.10$, $P = 0.31$). Thus, the Berger-Parker index does not differentiate well between natural ecosystems and anthropogenically transformed ecosystems.

Table 7
Indices of diversity of soil macrofauna communities

| Ecosystem type	N	Mean ± standard deviation	Minima	Maxima
Shannon diversity index				
Reclaimed soil	420	1.23 ± 0.41	0.29	2.30
Deciduous forest	205	1.62 ± 0.37	0.56	2.40
Steppe	161	1.03 ± 0.28	0.51	1.75
Sand steppe	105	0.88 ± 0.30	0.50	1.89
Meadow	210	1.66 ± 0.43	0.00	2.73
Reconstructed park	210	1.16 ± 0.32	0.00	2.06
Undamaged park	210	1.12 ± 0.29	0.41	1.70
Random	22000	2.55 ± 0.45	0.00	3.30
Total	23521	2.46 ± 0.54	0.00	3.30
Simpson diversity index				
Reclaimed soil	420	0.69 ± 0.19	0.17	1.00
Deciduous forest	205	0.81 ± 0.11	0.35	1.00
Steppe	161	0.83 ± 0.17	0.27	1.00
Sand steppe	105	0.83 ± 0.17	0.40	1.00
Meadow	210	0.85 ± 0.12	0.00	1.00
Reconstructed park	210	0.68 ± 0.12	0.00	0.96
Undamaged park	210	0.65 ± 0.13	0.20	1.00
Random	22000	0.92 ± 0.04	0.00	1.00
Total	23521	0.91 ± 0.07	0.00	1.00
Brillouin diversity index				
Reclaimed soil	420	0.89 ± 0.30	0.21	1.73
Deciduous forest	205	1.20 ± 0.31	0.35	1.92
Steppe	161	0.64 ± 0.21	0.35	1.12
Sand steppe	105	0.51 ± 0.20	0.32	1.27
Meadow	210	1.21 ± 0.38	0.00	2.14
Reconstructed park	210	0.89 ± 0.25	0.00	1.54
Undamaged park	210	0.89 ± 0.25	0.28	1.49
Random	22000	2.09 ± 0.48	0.00	2.86
Total	23521	2.01 ± 0.55	0.00	2.86
Berger-Parker diversity index				
Reclaimed soil	420	2.22 ± 0.95	1.09	10.00
Deciduous forest	205	2.90 ± 0.97	1.24	6.50
Steppe	161	2.31 ± 0.74	1.17	5.00
Sand steppe	105	2.17 ± 0.79	1.25	5.00
Meadow	210	3.43 ± 1.30	1.00	8.00
Reconstructed park	210	2.14 ± 0.69	1.00	7.50
Undamaged park	210	2.02 ± 0.52	1.12	3.67
Random	22000	4.93 ± 1.25	1.00	15.00
Total	23521	4.77 ± 1.38	1.00	15.00
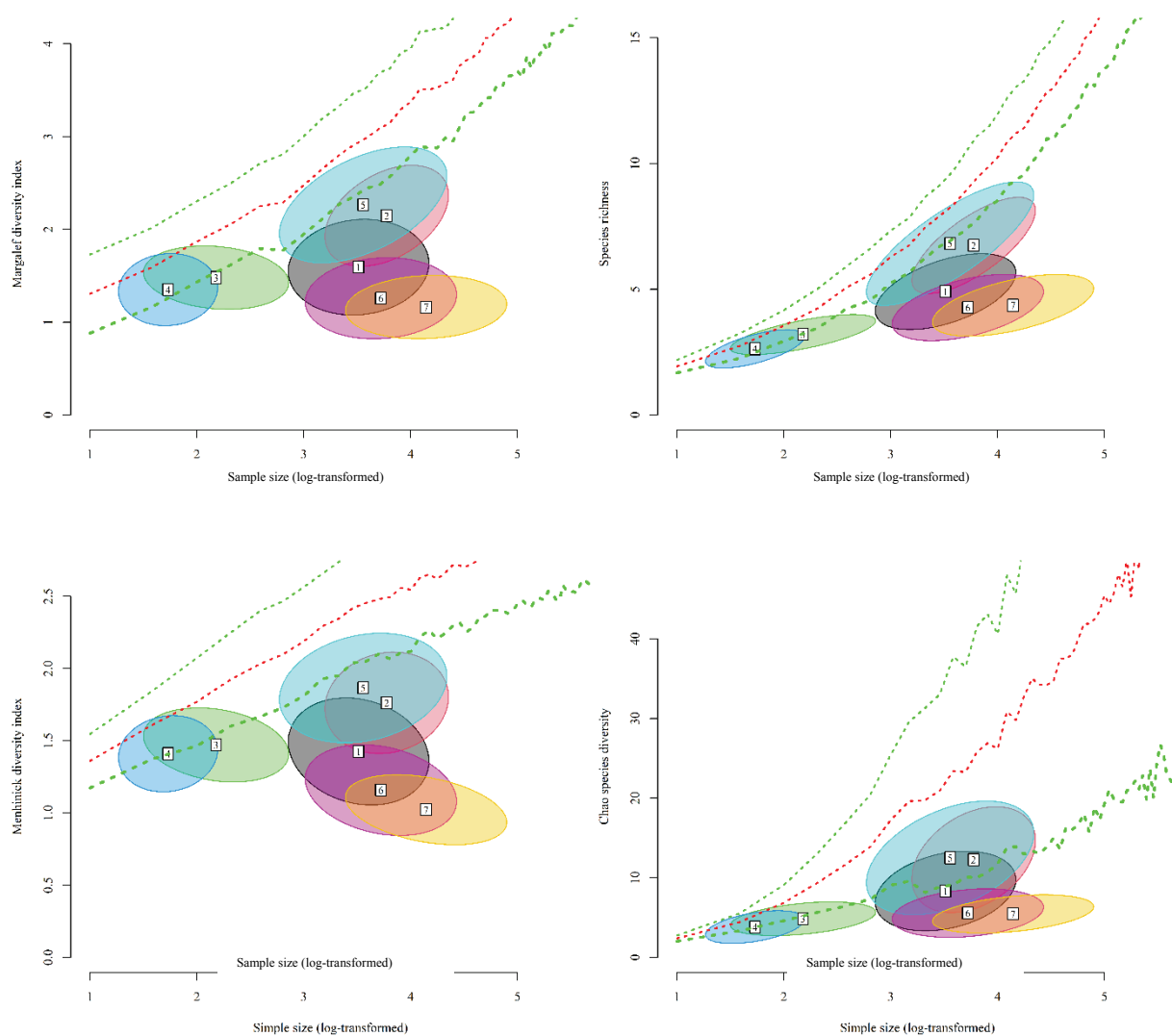
Fig. 3. Dependence of species richness indices on sample size: the abscissa is the sample size (number of individuals in a 0.25 × 0.25 m sample), and the ordinate is the diversity index: 1 – reclaimed soil, 2 – deciduous forest, 3 – steppe, 4 – sand steppe, 5 – meadow, 6 – reconstructed park, 7 – undamaged park; the red dotted line indicates the average value of the corresponding index for the community, generated randomly from the list of species with the corresponding abundance, which is typical for the community as a whole; the green dashed line indicates the 95% confidence interval of the null hypothesis

The Shannon and Brillouin indices showed a significant dependence on the sample size (Fig. 4). The number of individuals in the sample significantly influenced the differences in these indices. The dependence of Berger-Parker on the sample size is not significant, and the Simpson index was practically independent of the sample size. The general linear model, as well as the graph analysis, showed that the heterogeneity indices were statistically significantly different from the null alternative. The natural ecosystems were characterized by a greater approximation to the zone of the zero alternative, although they could have lower absolute values of the diversity indices. This was clearly seen in relation to the Shannon and Brillouin indices. The strong dependence of these indices on the sample size significantly distorts the observed estimates of the diversity of communities. The Simpson index was practically independent of sample size, and according to it, natural ecosystems differed from anthropogenically transformed ones by higher values of diversity estimates. However, the approximation of the diversity of natural ecosystems to the maximum possible values of the Simpson index made it impossible to distinguish between different natural ecosystems. Anthropogenically transformed ecosystems are also poorly distinguished from each other. The Berger-Parker index occupied an intermediate position in terms of its properties between the Shannon and Brillouin indices on the one hand and the Simpson index on the other. The Berger-Parker index had the same disadvantages caused by the dependence on sample size as the Shannon and Brillouin indices, although this dependence was somewhat less. Therefore, it was found that

the forest and meadow ecosystems differed significantly from all other ecosystems, while other ecosystems did not differ from each other in terms of the Berger-Parker index.

The Shannon index of evenness was highest for the natural systems compared to the anthropogenically transformed ones (planned comparison F = 1215.4, P < 0.001, Table 8). The highest value of community evenness was found for the sandy steppe, which was statistically significantly higher than this indicator for the steppe communities on loess-like loams and meadows (planned comparison F = 20.5, P < 0.001). The differences in Shannon's index of evenness between the steppe communities on loess-like loams and meadows were not statistically significant (planned comparison F = 0.19, P = 0.66). The evenness of forest communities was the lowest among all natural ecosystems (planned comparison F = 123.8, P < 0.001). The park plantation community after the park reconstruction had a higher Shannon's evenness index compared to other anthropogenically transformed ecosystems (planned comparison F = 101.1, P < 0.001). The differences between the communities of reclaimed soils and the park without reconstruction were not statistically significant (planned comparison F = 0.15, P = 0.69). The Brillouin and McIntosh evenness indices followed the pattern that was characteristic of the Shannon evenness index. The steppe community was significantly superior to all other communities by Simpson's index of evenness (planned comparison F = 7403.8, P < 0.001). Other ecosystems differed slightly in the Simpson index of evenness.
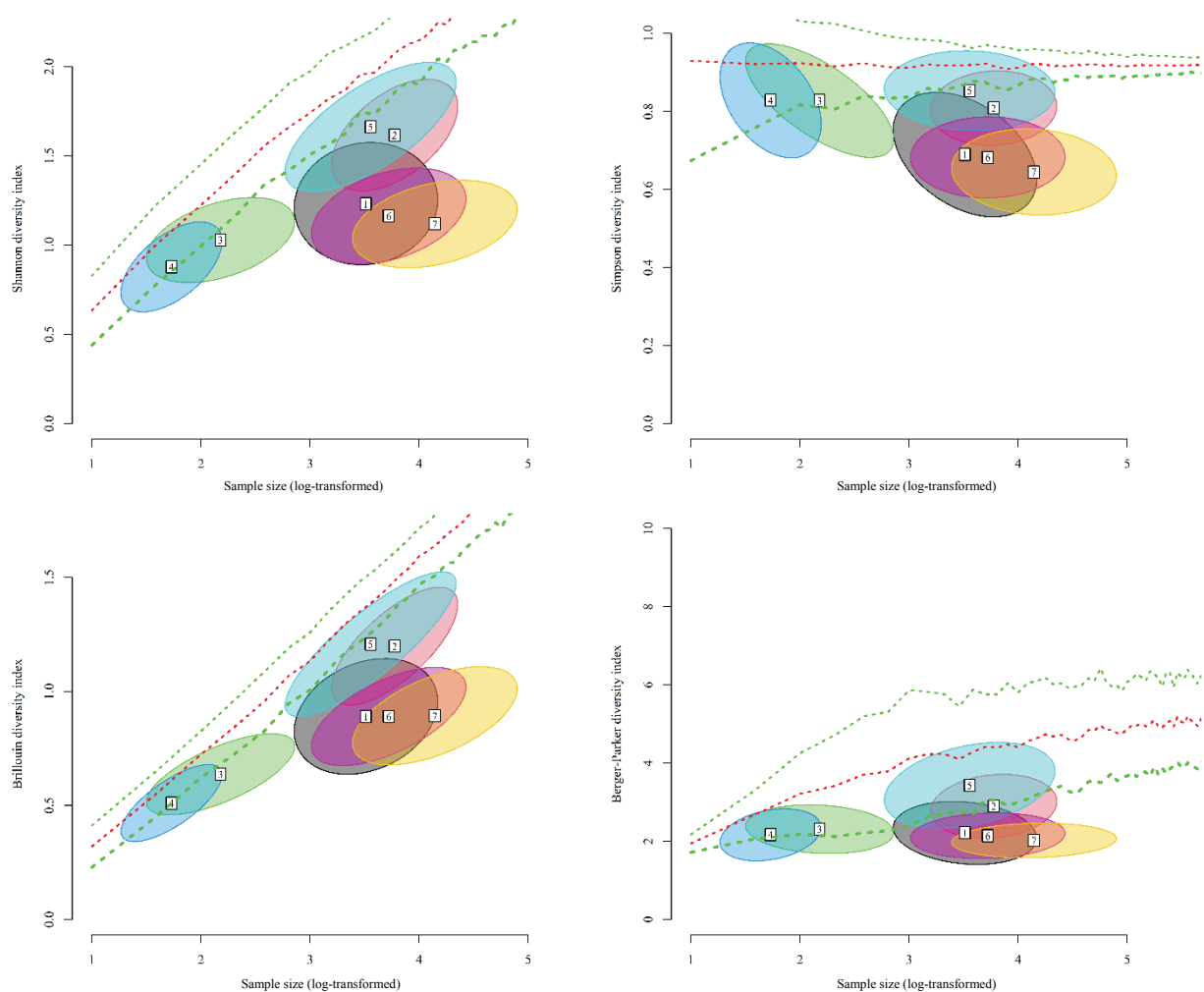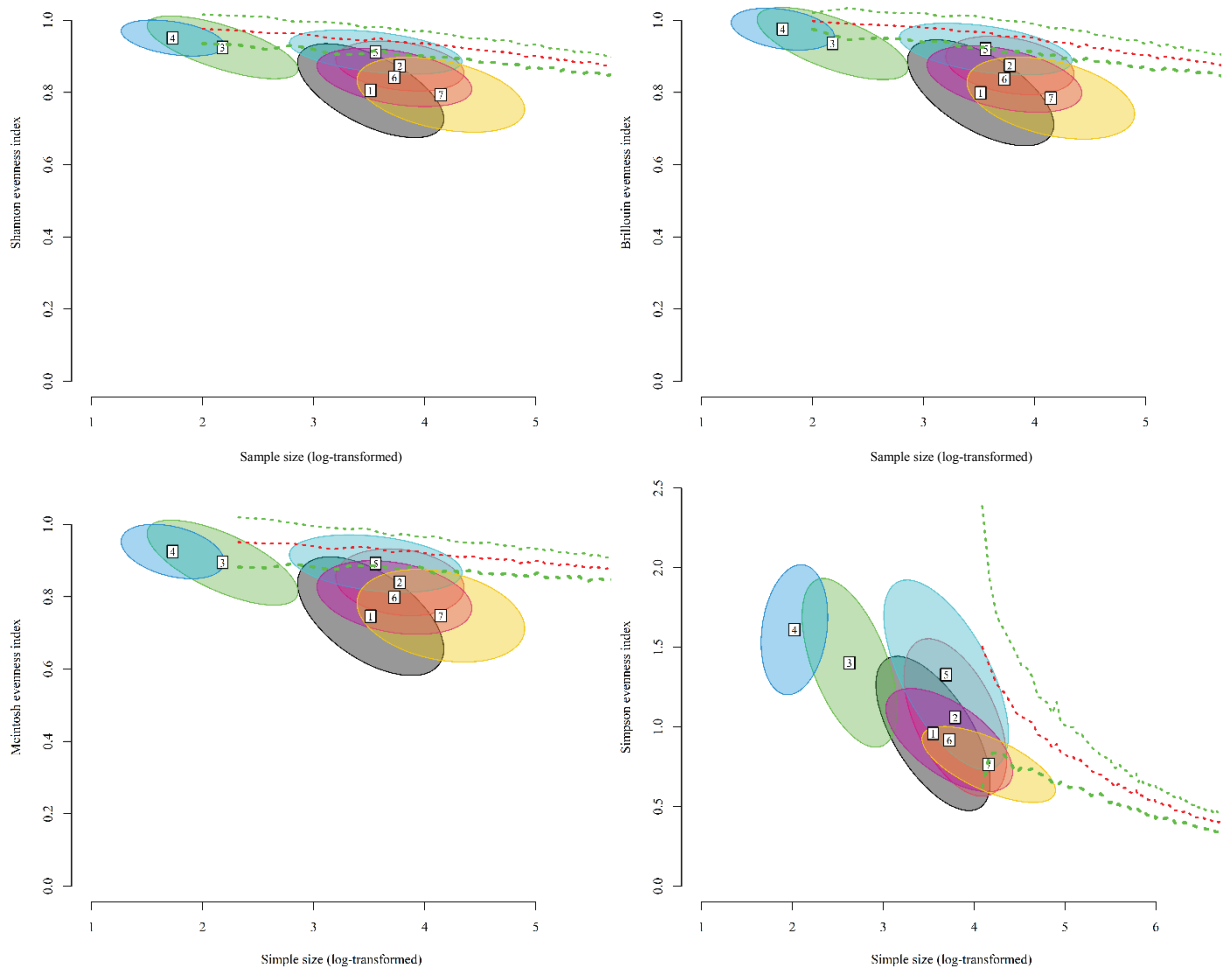
Fig. 4. Dependence of community heterogeneity indices on sample size: the abscissa is the sample size (number of individuals in a 0.25 × 0.25 m sample), and the ordinate is the diversity index: 1 – reclaimed soil, 2 – deciduous forest, 3 – steppe, 4 – sand steppe, 5 – meadow, 6 – reconstructed park, 7 – undamaged park; the red dotted line indicates the average value of the corresponding index for the community, generated randomly from the list of species with the corresponding abundance, which is typical for the community as a whole; green dashed line indicates the 95% confidence interval of the null hypothesis

Ecosystem type	N	Mean ± standard deviation	Minima	Maxima
Reconstructed park	210	0.80 ± 0.12	0.22	1.00
Undamaged park	210	0.75 ± 0.15	0.25	1.00
Random	22000	0.88 ± 0.05	0.00	1.00
Total	23521	0.88 ± 0.07	0.00	1.00
Simpson evenness index				
Reclaimed soil	420	0.03 ± 0.15	0.002	1.00
Deciduous forest	205	0.02 ± 0.10	0.003	1.00
Steppe	161	0.38 ± 0.48	0.004	1.00
Sand steppe	105	0.42 ± 0.49	0.008	1.00
Meadow	210	0.09 ± 0.27	0.003	1.00
Reconstructed park	210	0.01 ± 0.00	0.004	1.00
Undamaged park	210	0.01 ± 0.07	0.003	1.00
Random	22000	0.01 ± 0.01	0.002	1.00
Total	23521	0.01 ± 0.07	0.002	1.00

Table 8

Indices of evenness of soil macrofauna communities

Ecosystem type	N	Mean ± standard deviation	Minima	Maxima
Shannon evenness index				
Reclaimed soil	420	0.81 ± 0.15	0.31	1.00
Deciduous forest	205	0.87 ± 0.08	0.49	1.00
Steppe	161	0.92 ± 0.10	0.46	1.00
Sand steppe	105	0.95 ± 0.06	0.72	1.00
Meadow	210	0.91 ± 0.07	0.57	1.00
Reconstructed park	210	0.84 ± 0.10	0.35	1.00
Undamaged park	210	0.79 ± 0.12	0.37	1.00
Random	22000	0.88 ± 0.06	0.00	1.00
Total	23521	0.88 ± 0.06	0.00	1.00
Brillouin evenness index				
Reclaimed soil	420	0.80 ± 0.17	0.28	1.00
Deciduous forest	205	0.87 ± 0.10	0.46	1.00
Steppe	161	0.94 ± 0.12	0.42	1.00
Sand steppe	105	0.98 ± 0.07	0.70	1.00
Meadow	210	0.92 ± 0.09	0.53	1.00
Reconstructed park	210	0.84 ± 0.11	0.31	1.00
Undamaged park	210	0.78 ± 0.13	0.34	1.00
Random	22000	0.88 ± 0.06	0.00	1.00
Total	23521	0.88 ± 0.07	0.00	1.00
Mcintosh evenness index				
Reclaimed soil	420	0.75 ± 0.19	0.20	1.00
Deciduous forest	205	0.84 ± 0.11	0.36	1.00
Steppe	161	0.89 ± 0.14	0.32	1.00
Sand steppe	105	0.92 ± 0.09	0.60	1.00
Meadow	210	0.89 ± 0.09	0.44	1.00

The evenness indices decreased with increasing sample sizes (Fig. 5). The patterns for the Simpson, Shannon, and McIntosh evenness indices were almost identical. The differences between the evenness indices were significantly affected by the sample size. Therefore, the degree of proximity to the zone of the zero alternative was a better indicator of the level of anthropogenic transformation than absolute indicators of community evenness. If the situation is viewed from this angle, then the natural communities were characterized by greater proximity to the highest levels of evenness, which were indicated by the zone of the zero alternative. Of course, a negative feature of these evenness indices was their poor differential ability to characterize the different types of ecosystems. The Simpson evenness index was also characterized by a very wide range of the zero alternative zone at small and moderate sample sizes. Obviously, for the

vast majority of cases, the observed values of the Simpson evenness index were difficult to distinguish from the null hypothesis.

Thus, the tendency for biodiversity indices to depend on sample size significantly degrades the properties of indices as measures of diversity. The Margalef diversity index includes a normalization to the logarithm of the sample size, and the Menhinick diversity index includes a normalization to the logarithm of the sample size. However, it is obvious that such normalization does not allow us to extract the component of variation that is independent of the sample size.

Fig. 5. Dependence of community equality indices on sample size: the abscissa is the sample size (number of individuals in a 0.25 × 0.25 m sample), and the ordinate is the diversity index: 1 – reclaimed soil, 2 – deciduous forest, 3 – steppe, 4 – sand Steppe, 5 – meadow, 6 – reconstructed park, 7 – undamaged park; the red dashed line indicates the average value of the respective index for the community, generated randomly from the list of species with the corresponding abundance, which is typical for the community as a whole; green dashed line indicates the 95% confidence interval of the null hypothesis
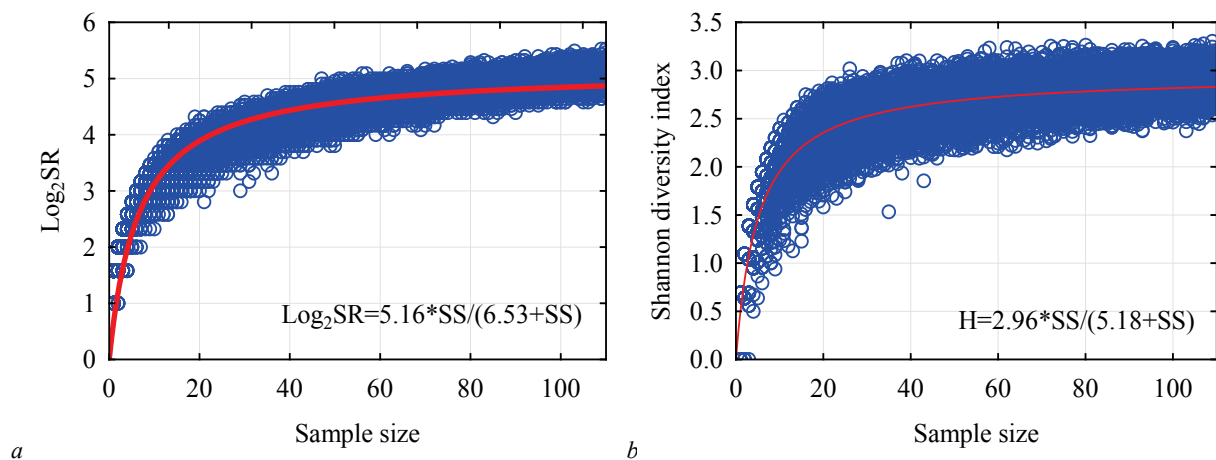

Fig. 6. Relationship between sample size and diversity indices: the abscissa is the sample size (number of individuals in one sample), the ordinate is the logarithm of species richness (*a*) and Shannon's index (*b*); the red line indicates the Michaelis-Menten model with saturation

The relationship between sample size and the logarithm of species richness and Shannon's index can be approximated by the Michaelis-Menten model with saturation (Fig. 6). The model allows us to calculate the correction factor depending on the sample size: it will be the largest for the smallest size and will approach zero as the sample size increases.

For the logarithm of the number of species, the corrected diversity index will be as follows:

$$Log_2 SR_c = Log_2 SR + \left(5.16 - \frac{5.16 \times SS}{6.53 + SS}\right) = Log_2 SR + \frac{33.7}{6.53 + SS},$$

where $Log_2 SR_c$ is the adjusted value of the logarithm of the number of species, taking into account the sample size; SS is the observed value of the number of species; SS is the sample size.

The corrected Shannon index for diversity will be as follows:

$$H_c = H + \left(2.96 - \frac{2.96 \times SS}{5.18 + SS}\right) = H + \frac{15.33}{5.18 + SS},$$

where H_c is the adjusted value of the Shannon index taking into account the sample size; SS is the sample size.

The traditional and corrected diversity indices were ordered in the space defined by the coefficients of determination for the nested ANOVA and the sample size trend (Fig. 7).
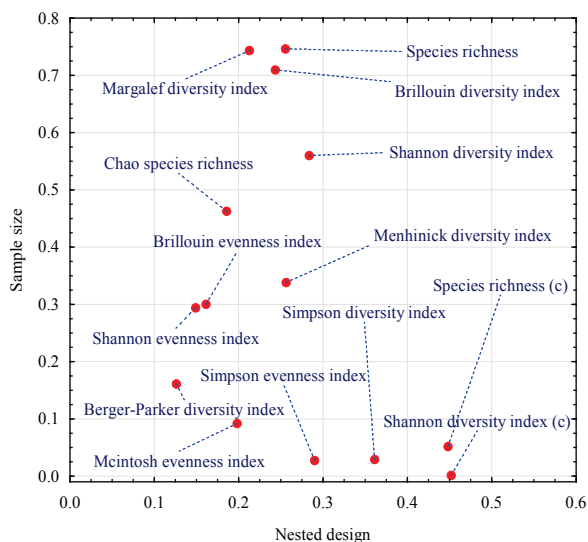
Fig. 7. Ordination of community diversity indices in the space of ability to distinguish between ecosystem types and levels of anthropogenic transformation of ecosystems and dependence on sample size: the abscissa is the explained variance of the nested ANOVA with ecosystem types nested in levels of anthropogenic transformation as predictors for sample size as a conditional variable (R_{adj}^2), the abscissa is the explained variance by sample size (R_{adj}^2)

Indices should be located in the lower right corner to have the best properties as a measure of diversity. Such indices are independent of sample size effects and are best able to differentiate between ecosystems, especially between natural and anthropogenic ecosystems. None of the traditional indices are in the lower right corner. Simpson's indices (heterogeneity and equality) are the least dependent on sample size, but they have a very moderate ability to differentiate ecosystems. Very popular indicators of diversity, such as species richness or Shannon's index, are highly dependent on sample size and their ability to differentiate ecosystems is very low. The Chao index, which refines the estimate of the number of species, is somewhat less scale-dependent than the number of species *per se*, but this comes at the expense of lowering the index to differentiate ecosystems. The corrected species richness index and the corrected Shannon index are practically independent of sample size and have the greatest ability to differentiate ecosystems.

Discussion

The choice of the "best" species diversity index should be based on certain criteria (Roswell et al., 2021). Certainly, the species diversity index should be proportional to the number of species (Stirling & Wilsey, 2001). The species diversity index should be used to compare communities by diversity (Staudhammer et al., 2018), so it should be able to characterize diversity itself, not another property of the community. The index should be comparable regardless of the conditions of the experiment and should be able to differentiate between communities. Thus, independence from

sample size and differential ability can be considered the most important criteria for assessing the quality of a diversity index.

The unequal distribution of the number of individuals among species is also a source of variability in community diversity. The laws of distribution can be described by certain models, and some of the parameters of the models can be considered as indices of species diversity. Our research has shown that Fisher's log-series is the best model for describing the structure of the soil macrofauna community in all natural ecosystems. The Zipf-Mandelbrot model is the best for artificial ecosystems, although the descriptive power of Fisher's log series is also high. The parameter of Fisher's log series is the alpha constant, which can be considered as an index of diversity. This index actually reflects the number of species in the community and, accordingly, has the corresponding disadvantages of species richness as an indicator of diversity, namely, a strong dependence on sample size. The same is true for the fundamental parameter of diversity. It was estimated based on the assumptions of the neutral diversity theory about the ecological equivalence of species, and this indicator is also strongly correlated with the number of species. Multifactorial impacts on a community are usually considered to generate a distribution of species in the community that can be best described by Fisher's log series. The assumption of neutral diversity theory is the opposite: all species in a community are equivalent. But diversity indices based on the opposite assumptions have similar properties: they are strongly correlated with species richness. Our empirical data do not support the assumption that the fundamental parameter of biodiversity is correlated with the Simpson index. This index is practically the only one of the considered ones that has the best properties of a diversity index and is least dependent on sample size. Obviously, the fundamental parameter of biodiversity does not have this property. However, it is difficult to verify this assumption, since the accuracy of calculating indices that are parameters of distributions is very low for small sample sizes. Therefore, we have considered these indices only for the communities as a whole without detailing them for individual samples.

The finding that soil macrofauna communities in the artificial ecosystems can be best described by the Zipf-Mandelbrot model is promising. The Zipf-Mandelbrot model (Zipf, 1949; Mandelbrot, 1977) is a deterministic model with a strong theoretical foundation and many applications in various fields of science. The Zipf-Mandelbrot model describes community structure using two parameters, beta and gamma. A good fit of the model can support the hypotheses about the underlying processes that link the requirements of different species to the probability of encountering optimal growth conditions in the environment. The Zipf-Mandelbrot model has been found to provide a good statistical description of the structure and transformation of natural communities, provided that the habitat can be considered hierarchical in structure (Barangé & Campos, 1991). According to the theory, the gamma parameter of the Zipf-Mandelbrot model takes on low values in highly organized systems with complex interactions between species (Frontier, 1985). These systems, due to their complexity, require a large amount of information to describe and are thus less predictable than less structured systems. The information added by a species in a complex system is low, as is the probability of its emergence, since most of the available niche is already occupied. Our data confirm these theoretical assumptions. In soil macrofauna communities from natural ecosystems, the beta parameter is usually lower than in anthropogenic ecosystems. The Zipf-Mandelbrot parameter beta reflects the potential diversity of the environment or niche diversification, taking on higher values when the environment provides space for more alternatives. Empirical evidence suggests that natural ecosystems do indeed differ from artificial ecosystems in terms of environmental diversity.

The species richness indices are highly dependent on sample size. The zone of the null hypothesis is marked by a nonlinear dependence on sample size, indicating that with a small increase in sample size, species richness can grow disproportionately fast only due to random causes. One of the disadvantages of such indices is the inability to compare different ecosystems that differ in sample size, even when these differences are not very significant. The natural ecosystems typically have the diversity indices that are close to the zone of the null alternative, which can be viewed as the state of highest entropy for a given sample size, which is proportional to the density of the community. Accordingly, the artificial

ecosystems are more distant from the zero alternative. But due to the curvature of the sample size/diversity space, there are situations where the diversity of natural ecosystems is less than that of artificial ones.

The mechanisms of controlling the number of individual species in the anthropogenic ecosystems are disrupted and significant temporary outbreaks of species and community abundances can occur (Zhukov et al., 2021). The variability in abundance can lead to species turnover, which is a factor in increasing the number of species in a community. The species richness can often be higher under conditions of anthropogenic impact (Elo et al., 2018). However, the anthropogenic impacts can change the trophic composition of ecosystems, which can stimulate the overall productivity of communities. However, the traditional indices may not be able to capture the fact that the increase in diversity in the anthropogenic conditions is disproportionately less than what would be expected for a natural ecosystem with a similar level of abundance or productivity. Therefore, we observe a situation where the anthropogenic ecosystems may have more diversity than their natural counterparts according to diversity indices.

The Menhinick and Margalef indices were designed to normalize the species richness to sample size. However, they still depend on the sample size because they do not take into account the real shape of the dependence of the diversity index on the sample size. The normalization based on the logarithm or root of the square of the sample size should be adjusted by the additional coefficients that are obviously unique to a particular meta-community. The range of values of both the logarithmic and quadratic functions is infinite, so the correction factor for estimating the effect of sample size will be uncertain. The Michaelis-Menten model with saturation has a parameter to which the function approaches, but never exceeds. This model describes the relationship between the diversity index and sample size well. The difference between the saturation level and the observed value of the function at a certain sample size is the correction factor. This coefficient approaches zero as the sample size increases. Thus, the calculation of the correction factor allows us to estimate the diversity regardless of the sample size, or more precisely, to bring the diversity estimate to a sample size of infinite dimension.

Obviously, the specific values of the model of the null alternative's dependence on the sample size are not universal and depend on the properties of the particular community from which the samples are generated to calculate the null alternative. Therefore, we can say that it is impossible to propose universal coefficients for correcting the traditional diversity indices. These correction indices are context-dependent. The context is the specific group of living organisms, the way they are accounted for, the specific environmental conditions in which these living organisms live, and the range of ecosystems covered by the study. The ultimate goal of correction coefficients is to provide an adequate comparison of ecosystems in terms of diversity. Obviously, the tools for adequate comparison can only be developed for specific conditions. However, the principles of development of such tools are universal. The universality lies in the asymptotic growth of diversity indicators with increasing sample size.

The diversity indices were ordered in the space of two dimensions, one of which was the explanatory variation between ecosystems and the other was the dependence on sample size. The ordination of the traditional indices showed that there is a vacancy for the best index in the sense that such an index should best explain differences between ecosystems and differences between natural and artificial ecosystems. It should also be independent of sample size. The Simpson heterogeneity index and the Simpson evenness index were the best among the traditional indices, but they did not explain differences between ecosystems very well, especially when it came to differentiating between natural and artificial ecosystems. The Margalef index, which is supposed to be independent of sample size, on the contrary, showed a very high level of dependence. To a lesser extent, but such a dependence was also found for the Menhinick index. Obviously, the size dependence negatively affects the differential ability of the indices. The corrected indices of species richness and Shannon's index were practically independent of sample size and had a great ability to differentiate ecosystems by the level of diversity, with the natural ecosystems characterized by consistently higher values of the corrected indices compared to artificial ecosystems.

Recording animals is usually a matter of counting individuals, so the number of animals is certainly a marker of sample size. For plants, projective cover is a very common marker. Therefore, the most adequate analog of sample size for plants may be the area of the survey site. Therefore, in the future, it will be very interesting to investigate the value of plant area for correcting the indicators of species diversity of plant communities.

Conclusion

Biodiversity indices differ in their sensitivity to the number of species and the evenness of their abundance. The species richness indices are formally insensitive to the distribution of species abundance, and the evenness indices are formally insensitive to the number of species. The heterogeneity indices respond to both the number of species and the distribution of species abundance. In reality, species richness and evenness of species abundance in a community depend on the fitness of the distribution of species abundance to a certain model, and the model statistics are also indices of diversity. The best model for the natural ecosystems was Fisher's log series, and the best model for the artificial ecosystems was the Zipf-Mandelbrot law. The dependence on sample size is typical for almost all of the considered diversity indices. To the least extent, it is characteristic of the Simpson's indices of heterogeneity and evenness. The dependence on sample size makes indices from different ecosystems practically incomparable, which makes their use meaningless. Even minor differences in sample size can lead to significant deviations in the values of diversity indices. The application of the Michaelis-Menten model allowed us to propose a method for correcting species richness indices and the Shannon index. The indices are practically independent of sample size after correction, and their differential ability to characterize individual ecosystems and the level of anthropogenic transformation increases significantly.

References

Alonso, D., & McKane, A. J. (2004). Sampling Hubbell's neutral theory of biodiversity. Ecology Letters, 7(10), 901–910.

Anderson, J., & Ingram, J. (1993). Tropical soil biology and fertility. A handbook of methods. CAB International, Wallingford.

Aubier, T. G. (2020). Positive density dependence acting on mortality can help maintain species-rich communities. ELife, 9, e.57788.

Bailey, R. I., Molleman, F., Vasseur, C., Woas, S., & Prinzing, A. (2018). Large body size constrains dispersal assembly of communities even across short distances. Scientific Reports, 8(1), 10911.

Barangé, M., & Campos, B. (1991). Models of species abundance: a critique of and an alternative to the dynamics model. Marine Ecology Progress Series, 69, 293–298.

Beck, J., & Schwanghart, W. (2010). Comparing measures of species diversity from incomplete inventories: An update. Methods in Ecology and Evolution, 1(1), 38–44.

Berger, W. H., & Parker, F. L. (1970). Diversity of planktonic foraminifera in deep-sea sediments. Science, 168(3937), 1345–1347.

Bonnaffé, W., Danet, A., Legendre, S., & Edeline, E. (2021). Comparison of size-structured and species-level trophic networks reveals antagonistic effects of temperature on vertical trophic diversity at the population and species level. Oikos, 130(8), 1297–1309.

Brillouin, L. (1956). Science and information theory. Academic Press, New York.

Cazzolla Gatti, R., Amoroso, N., & Monaco, A. (2020). Estimating and comparing biodiversity with a single universal metric. Ecological Modelling, 424, 109020.

Ceia, F. R., Cherel, Y., Silva, A. V., Garrido, S., Angélico, M. M., da Silva, J. M., Laranjeiro, M. I., & Ramos, J. A. (2023). Drivers of niche partitioning in a community of mid-trophic level epipelagic species in the North Atlantic. Hydrobiologia, 850(7), 1583–1599.

Chao, A. (1982). Nonparametric estimation of the number of classes in a population. Scandinavian Journal of Statistics, 11(4), 265–270.

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. Biometrics, 43(4), 783–791.

Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. Ecology, 93(12), 2533–2547.

Chao, A., & Jost, L. (2015). Estimating diversity and entropy profiles via discovery rates of new species. Methods in Ecology and Evolution, 6(8), 873–882.

Chao, A., & Lee, S.-M. (1992). Estimating the number of classes via sample coverage. Journal of the American Statistical Association, 87(417), 210–217.

Chave, J. (2004). Neutral theory and community ecology. Ecology Letters, 7(3), 241–253.

Chisholm, R. A., & Pacala, S. W. (2010). Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological

communities. Proceedings of the National Academy of Sciences, 107(36), 15821–15825.

Cohen, J. E., & Łuczak, T. (1992). Trophic levels in community food webs. Evolutionary Ecology, 6(1), 73–89.

Cramer, M. J., & Willig, M. R. (2005). Habitat heterogeneity, species diversity and null models. Oikos, 108(2), 209–218.

DeAngelis, D. L., & Waterhouse, J. C. (1987). Equilibrium and nonequilibrium concepts in ecological models. Ecological Monographs, 57(1), 1–21.

Diamond, J. (1988). Factors controlling species diversity: Overview and synthesis. Annals of the Missouri Botanical Garden, 75(1), 117–129.

Dietrich, P., Cesarz, S., Liu, T., Roscher, C., & Eisenhauer, N. (2021). Effects of plant species diversity on nematode community composition and diversity in a long-term biodiversity experiment. Oecologia, 197(2), 297–311.

Doncaster, C. P. (2009). Ecological equivalence: A realistic assumption for niche theory as a testable alternative to neutral theory. PLoS One, 4(10), e7460.

Elo, M., Alahuhta, J., Kanninen, A., Meissner, K. K., Seppälä, K., & Mönkkönen, M. (2018). Environmental characteristics and anthropogenic impact jointly modify aquatic macrophyte species diversity. Frontiers in Plant Science, 9, 1001.

Etienne, R. S., Apol, M. E. F., Olff, H., & Weissing, F. J. (2007). Modes of speciation and the neutral theory of biodiversity. Oikos, 116(2), 241–258.

Fisher, R. A., Corbet, A. S., & Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. The Journal of Animal Ecology, 12(1), 42–58.

Frank, S. A. (2019). The common patterns of abundance: The log series and Zipf's law. F1000Research, 8, 334.

Frerebeau, N. (2019). Tabula: An R package for analysis, seriation, and visualization of archaeological count data. Journal of Open Source Software, 4(44), 1821.

Frontier, S. (1985). Diversity and structure in aquatic ecosystems. Oceanography and Marine Biology, 23, 253–312.

Gamito, S. (2010). Caution is needed when applying Margalef diversity index. Ecological Indicators, 10(2), 550–551.

Gaston, K. J. (1996). The multiple forms of the interspecific abundance-distribution relationship. Oikos, 76(2), 211–220.

Gilbert, B., & Levine, J. M. (2017). Ecological drift and the distribution of species diversity. Proceedings of the Royal Society B: Biological Sciences, 284, 20170507.

Gini, C. (1912). Variabilità e mutabilità [Variability and mutability]. In: Pizetti, E., & Salvemini, T. (Eds.). Memorie di metodologica statistica [Memoirs of statistical methodology]. Libreria Eredi Virgilio Veschi, Rome (in Italian).

Gotelli, N. J., & Colwell, R. K. (2011). Estimating species richness. In: Magurran, A. E., & McGill, B. J. (Eds.). Biological diversity: Frontiers in measurement and assessment. Oxford University Press, Oxford. Pp. 39–54.

Grabchak, M., Marcon, E., Lang, G., & Zhang, Z. (2017). The generalized Simpson's entropy is a measure of biodiversity. PLoS One, 12(3), e0173305.

Gregory, T. R. (2009). Understanding natural selection: Essential concepts and common misconceptions. Evolution: Education and Outreach, 2(2), 156–175.

Guiasu, R. C., & Guiasu, S. (2010). The Rich-Gini-Simpson quadratic index of biodiversity. Natural Science, 2(10), 1130–1137.

Guillera-Arroita, G., Kéry, M., & Lahoz-Monfort, J. J. (2019). Inferring species richness using multispecies occupancy modeling: Estimation performance and interpretation. Ecology and Evolution, 9(2), 780–792.

Haegeman, B., Hamelin, J., Moriarty, J., Neal, P., Dushoff, J., & Weitz, J. S. (2013). Robust estimation of microbial diversity in theory and in practice. The Microbial Population and Community Ecology, 7(6), 1092–1101.

Hankin, R. K. S. (2007). Introducing untb, an R package for simulating ecological drift under the unified neutral theory of biodiversity. Journal of Statistical Software, 22(12), 1–15.

He, F., & Hu, X. S. (2005). Hubbell's fundamental biodiversity parameter and the Simpson diversity index. Ecology Letters, 8(4), 386–390.

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. Ecology, 54(2), 427–432.

Hubbell, S. (2001). A unified theory of biodiversity and biogeography. не хватает данных

Hubbell, S. P. (2001). The unified neutral theory of biodiversity and biogeography. Princeton University Press. не хватает данных

Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. Ecology, 52(4), 577–586.

Huston, M. (1979). A general hypothesis of species diversity. The American Naturalist, 113(1), 81–101.

Jost, L. (2006). Entropy and diversity. Oikos, 113(2), 363–375.

Jost, L. (2010). The relation between evenness and diversity. Diversity, 2(2), 207–232.

Julliard, R., Clavel, J., Devictor, V., Jiguet, F., & Couvet, D. (2006). Spatial segregation of specialists and generalists in bird communities. Ecology Letters, 9(11), 1237–1244.

Kempton, R. A. (1979). The structure of species abundance and measurement of diversity. Biometrics, 35(1), 307–321.

Kopp, M. (2010). Speciation and the neutral theory of biodiversity. BioEssays, 32(7), 564–570.

Kovalenko, K. E., Thomaz, S. M., & Warfe, D. M. (2012). Habitat complexity: Approaches and future directions. Hydrobiologia, 685(1), 1–17.

Laan, A., & de Polavieja, G. G. (2018). Species diversity rises exponentially with the number of available resources in a multi-trait competition model. Proceedings of the Royal Society B: Biological Sciences, 285, 20181273.

Lande, R., DeVries, P. J., & Walla, T. R. (2000). When species accumulation curves intersect: Implications for ranking diversity using small samples. Oikos, 89(3), 601–605.

Lawton, J. H. (1999). Are there general laws in ecology? Oikos, 84(2), 177–192.

Lu, H.-P., Yeh, Y.-C., Shiah, F.-K., Gong, G.-C., & Hsieh, C. (2019). Evolutionary constraints on species diversity in marine bacterioplankton communities. Microbial Population and Community Ecology, 13(4), 1032–1041.

Ma, M. (2005). Species richness vs evenness: Independent relationship and different responses to edaphic factors. Oikos, 111(1), 192–198.

Magurran, A. E. (2004). Measuring biological diversity. Blackwell Science Ltd., Oxford.

Magurran, A. E., & McGill, B. J. (2011). Biological diversity: Frontiers in measurement and assessment. Oxford University Press, Oxford.

Mandelbrot, B. B. (1977). Fractals: Form, chance and dimension. W. H. Freeman & Company, San Francisco.

Margalef, R. (1958). Information theory in ecology. General Systems, 3, 36–71.

Matthews, T. J., & Whittaker, R. J. (2015). On the species abundance distribution in applied ecology and biodiversity management. Journal of Applied Ecology, 52(2), 443–454.

McIntosh, R. P. (1967). An index of diversity and the relation of certain concepts to diversity. Ecology, 48(3), 392–404.

Menhinick, E. F. (1964). A comparison of some species-individuals diversity indices applied to samples of field insects. Ecology, 45(4), 859–861.

Merino, N., Aronson, H. S., Bojanova, D. P., Feyhl-Buska, J., Wong, M. L., Zhang, S., & Giovannelli, D. (2019). Living at the extremes: Extremophiles and the limits of life in a planetary context. Frontiers in Microbiology, 10, 780.

Milanesi, P., Della Rocca, F., & Robinson, R. A. (2020). Integrating dynamic environmental predictors and species occurrences: Toward true dynamic species distribution models. Ecology and Evolution, 10(2), 1087–1092.

Missa, O., Dytham, C., & Morlon, H. (2016). Understanding how biodiversity unfolds through time under neutral theory. Philosophical Transactions of the Royal Society B: Biological Sciences, 371, 20150226.

Motomura, I. (1932). On the statistical treatment of communities. Zoological Magazine, 44, 379–383.

Nijs, I., & Roy, J. (2000). How important are species richness, species evenness and interspecific differences to productivity? A mathematical model. Oikos, 88(1), 57–66.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-6.

Palmer, M. W. (1994). Variation in species richness: Towards a unification of hypotheses. Folia Geobotanica et Phytotaxonomica, 29(4), 511–530.

Petraitis, P. S., Latham, R. E., & Niesenbaum, R. A. (1989). The maintenance of species diversity by disturbance. The Quarterly Review of Biology, 64(4), 393–418.

Pielou, E. (1969). An introduction to mathematical ecology. Wiley, New York.

Pielou, E. C. (1975). Ecological diversity. John Wiley & Sons, New York..

Preston, F. W. (1948). The commonness, and rarity, of species. Ecology, 29(3), 254–283.

Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. Oikos, 130(3), 321–338.

Rybicki, J., Abrego, N., & Ovaskainen, O. (2020). Habitat fragmentation and species diversity in competitive communities. Ecology Letters, 23(3), 506–517.

Sanders, H. L. (1968). Marine benthic diversity: A comparative study. The American Naturalist, 102(925), 243–282.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. The University of Illinois Press, Urbana.

Simpson, E. H. (1949). Measurement of diversity. Nature, 163(4148), 688.

Soetaert, K., & Heip, C. (1990). Sample-size dependence of diversity indices and the determination of sufficient sample size in a high-diversity deep-sea environment. Marine Ecology Progress Series, 59, 305–307.

Southwood, T. R. E. (1977). Habitat, the templet for ecological strategies? The Journal of Animal Ecology, 46(2), 336–365.

Spatharis, S., & Tsirtsis, G. (2013). Zipf–Mandelbrot model behavior in marine eutrophication: Two way fitting on field and simulated phytoplankton assemblages. Hydrobiologia, 714(1), 191–199.

Staudhammer, C. L., Escobedo, F. J., & Blood, A. (2018). Assessing methods for comparing species diversity from disparate data sources: The case of urban and peri-urban forests. Ecosphere, 9(10), e02450.

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecology Letters, 17(7), 866–880.

Stier, A. C., Bolker, B. M., & Osenberg, C. W. (2016). Using rarefaction to isolate the effects of patch size and sampling effort on beta diversity. Ecosphere, 7(12), e01612.

Stirling, G., & Wilsey, B. (2001). Empirical relationships between species richness, evenness, and proportional diversity. The American Naturalist, 158(3), 286–299.

Thébault, E., & Loreau, M. (2005). Trophic interactions and the relationship between species diversity and ecosystem stability. The American Naturalist, 166(4), E95–E114.

Thomas, W. R., & Foin, T. C. (1982). Neutral hypotheses and patterns of species diversity: fact or artifact? Paleobiology, 8(1), 45–55.

Tuomisto, H. (2010a). A consistent terminology for quantifying species diversity? Yes, it does exist. Oecologia, 164(4), 853–860.

Tuomisto, H. (2010b). A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography, 33(1), 23–45.

Tutova, G. F., Zhukov, O. V., Kunakh, O. M., & Zhukova, Y. O. (2022). Response of earthworms to changes in the aggregate structure of floodplain soils. IOP Conference Series: Earth and Environmental Science, 1049(1), 012062.

Urban, M. C., Strauss, S. Y., Pelletier, F., Palkovacs, E. P., Leibold, M. A., Hendry, A. P., De Meester, L., Carlson, S. M., Angert, A. L., & Giery, S. T. (2020). Evolutionary origins for ecological patterns in space. Proceedings of the National Academy of Sciences, 117(30), 17482–17490.

van Loon, W. M. G. M., Walvoort, D. J. J., van Hoey, G., Vina-Herbon, C., Blandon, A., Pesch, R., Schmitt, P., Scholle, J., Heyer, K., Lavaleye, M., Phillips, G., Duineveld, G. C. A., & Blomqvist, M. (2018). A regional benthic fauna assessment method for the Southern North Sea using Margalef diversity and reference value modelling. Ecological Indicators, 89, 667–679.

Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., Tjiputra, J., & Pellissier, L. (2022). A quantitative review of abundance-based species distribution models. Ecography, 2022(1), e05694.

Whittaker, R. H. (1965). Dominance and diversity in land plant communities. Science, 147(3655), 250–260.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. Taxon, 21(2–3), 213–251.

Williamson, M., & Gaston, K. J. (2005). The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. Journal of Animal Ecology, 74(3), 409–422.

Willis, A. D. (2019). Rarefaction, alpha diversity, and statistics. Frontiers in Microbiology, 10, 2407.

Wilson, J. B., Sykes, M. T., & Peet, R. K. (1995). Time and space in the community structure of a species-rich limestone grassland. Journal of Vegetation Science, 6(5), 729–740.

Yakovenko, V., & Zhukov, O. (2021). Zoogenic structure aggregation in steppe and forest soils. In: Dmytruk, Y., & Dent, D. (Eds.). Soils under stress. Springer International Publishing. Pp. 111–127.

Zang, Z., Zeng, Y., Wang, D., Shi, F., Dong, Y., Liu, N., & Liang, Y. (2022). Species-abundance distribution patterns of plant communities in the Gurbantünggüt desert, China. Sustainability, 14(20), 12957.

Zhou, S., & Zhang, D. (2008). Neutral theory in community ecology. Frontiers of Biology in China, 3(1), 1–8.

Zhukov, A., & Gadorozhnaya, G. (2016). Spatial heterogeneity of mechanical impedance of a typical chernozem: The ecological approach. Ekológia (Bratislava), 35(3), 263–278.

Zhukov, O. V., Kunah, O. M., Dubinina, Y. Y., & Novikova, V. O. (2018). The role of edaphic, vegetational and spatial factors in structuring soil animal communities in a floodplain forest of the Dnipro river. Folia Oecologica, 45(1), 8–23.

Zhukov, O., Kunah, O., Fedushko, M., Babchenko, A., & Umerova, A. (2021). Temporal aspect of the terrestrial invertebrate response to moisture dynamic in technosols formed after reclamation at a post-mining site in Ukrainian steppe drylands. Ekológia (Bratislava), 40(2), 178–188.

Zhukov, O., Kunakh, O., Yorkina, N., & Tutova, A. (2023). Response of soil macrofauna to urban park reconstruction. Soil Ecology Letters, 5(2), 220156.

Zipf, G. K. (1949). Human behavior and the principle of least effort. Addison-Wesley Press, Inc., Cambridge.