Ecological Indicators 122 (2021) 107277

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Bait-lamina test for assessment of polluted soils: Rough vs. Precise scales

Evgenii L. Vorobeichik *, Igor E. Bergman

Institute of Plant and Animal Ecology, Ural Branch, Russian Academy of Sciences, 8 Marta Str. 202, Yekaterinburg 620144, Russia

ARTICLE INFO

Keywords:
Feeding activity
Soil fauna
Detritivores
Earthworms
Heavy metals
Copper smelter

ABSTRACT

The bait-lamina test is one of the few available methods for measuring the functional activity of soil animals, which was standardized for soil health assessment by ISO 18311. The bait consumption is measured visually on discrete scales. A two-point scale (0 or 1) is used more often, but the threshold for bait consumption scores varies from study to study: 1 point is assigned either to a hole that has been perforated to any extent, or to a hole that has been at least half perforated, or to a fully empty hole. Less often, a three-point scale (0, 0.5, and 1) or a five-point scale (0, 0.25, 0.5, 0.75, and 1) can be used. We investigated how much the use of different scales can influence statistical inferences.

We have established that a point on a five-point scale is directly proportional to the proportion of the mass consumption of bait, so in this scale the value of feeding activity is an acceptable surrogate for the bait consumption rate. We analyzed outcomes of the bait-lamina test for 7 measurement rounds at the control area and an area highly polluted with metals near the copper smelter. We assumed the most accurate five-point scale to be the reference and compared it to six less precise scales with fewer points (two three-point and four two-point scales). The difference between the reference and such rough scales depends on the percentage of intermediate points (0.25, 0.5, and 0.75). It turned out that intermediate values are by no means rare (from 11% to 100% of all non-zero values), and in the polluted area they are almost twice as common as in the control area (66% and 28%, respectively). Because of this, the relative difference between the reference scale and rough scales is 2–4 times greater at the polluted site than at the control site. This can lead to a bias in the effect size index: for the rough scales, the log Response Ratio can either double or decrease by one-third relative to the reference scale. This increases the probability of type I and II errors in statistical hypothesis testing.

Thus, the outcomes of the bait-lamina test are not invariant relative to the measuring procedure. The three-point scale and ISO 18311 two-point scale are the least biased relative to the reference scale. The use of other rough scales carries a risk of artifacts and should be avoided.

1. Introduction

The bait-lamina test (BLT), proposed by Von Törne (1990) 30 years ago, is very popular for estimating the feeding activity of soil detritivores. Despite its exceptional simplicity, or thanks to it, it has found wide application in many areas: when comparing the activity of soil detritivores in soil fertility gradients (Geissen et al., 2007; Rożen et al., 2010; Spehn et al., 2000), when studying the dynamics of soil animals (Eisenhauer et al., 2018; Musso et al., 2014; Thakur et al., 2018), when comparing modes of agricultural land management (e.g., Birkhofer et al., 2011; Graenitz and Bauer, 2000) and forest practices (Römbke et al., 2006), when assessing the effect of plant invasion (Pehle and Schirmel, 2015), fires (Musso et al., 2014; Podgaiski et al., 2014), forest fragmentation (Simpson et al., 2012), urbanization (Bergman et al., 2017), pesticide application (Förster et al., 2011; Larink and Sommer, 2002; Niemeyer et al., 2018), and soil pollution (e.g., André et al., 2009; Boshoff et al., 2014; Filzek et al., 2004; Vorobeichik and Bergman, 2020), in assessing remediation processes (Van Gestel et al., 2001), in addition to standard ecotoxicological tests for earthworms (Jänsch et al., 2017) and potworms (Bart et al., 2018), etc. The use of the BLT for evaluating the quality of contaminated soils is standardized by ISO 18311 (ISO, 2016). The BLT has been repeatedly recommended for inclusion in the minimum set of soil health indicators as one of the few available methods for estimating the functional activity of soil animals (Griffiths et al., 2016; Ritz et al., 2009; Römbke, 2014). Recently, interest in the BLT has increased due to the need to incorporate soil detritivores in carbon cycle models (Siebert et al., 2019; Thakur et al., 2018).

* Corresponding author.
E-mail addresses: ev@ipae.uran.ru (E.L. Vorobeichik), 5554505@mail.ru (I.E. Bergman).

https://doi.org/10.1016/j.ecolind.2020.107277
Received 5 August 2020; Received in revised form 7 December 2020; Accepted 13 December 2020
Available online 21 December 2020

The BLT has been described in detail in ISO 18311 (ISO, 2016). Methodological aspects of the BLT have repeatedly attracted attention: e.g., different composition of the bait (Helling et al., 1998; Simpson et al., 2012), the choice of the exposure duration (Gongalskii et al., 2003; Gongalsky et al., 2008; Van Gestel et al., 2003; Vorobeichik and Bergman, 2020), and the required sample size for statistical tests (Welsch et al., 2019) has been discussed.

The BLT was originally proposed as an express method, which implied a rapid visual measurement of bait consumption on a discrete scale. It was initially declared that although the inevitable coarsening can lead to loss of information and its distortion, such losses are compensated by the speed and rigor of the test (Von Törne, 1990). However, different coarsening options are possible, in terms of both the scale roughness, i.e., the number of the scale points, and the algorithm of assigning the consumption value to this or that point. As far as we know, the BLT outcomes obtained using different scales and/or different algorithms have not been compared before.

All algorithms are equivalent to each other in two extreme points, i.e. absolutely untouched and fully empty holes, but differ in intermediate points, i.e. partially perforated holes. In the original BLT, Von Törne (1990) suggested evaluating the feeding activity of soil detritivores on a two-point scale: 0 – the bait is not perforated, 1 – the bait is perforated to any extent (provided that artifacts, i.e. cracks in the bait, are excluded if it dries or deforms). Many have used this algorithm (e.g., Joschko et al., 2008; Pehle and Schirmel, 2015). The ISO 18311 recommends coarsening the two-point scale from the other side – with a decrease: 1 point corresponds to the consumption of the bait by at least half, and in all other cases the consumption is assumed to be 0 (ISO, 2016). This recommendation was also followed by many (e.g., André et al., 2009; Bart et al., 2018; Spehn et al., 2000), or the threshold for a score of 1 was slightly moved from 50% to 30% of bait consumption (Niemeyer et al., 2018). Sometimes an even more decreasing algorithm was used: a score of 1 corresponds to a fully empty hole, and in all other cases the consumption is 0 (Podgaiski et al., 2014). In several works, instead of a two-point scale, a three-point scale was used, adding an intermediate gradation of 0.5 (bait was partially consumed) (Siebert et al., 2019; Thakur et al., 2018). In our previous research we used a five-point scale in which points corresponded to the approximate proportion of the area of bait consumed: 0 – untouched; 0.25 – consumed about 25%; 0.50 – about 50%; 0.75 – about 75%; and 1 – fully empty (Bergman et al., 2017; Vorobeichik and Bergman, 2020; Vorobeichik et al., 2007).

The authors did not always report which algorithm was used to attribute bait consumption to a score (e.g., Geissen et al., 2007; Larink and Sommer, 2002) or insufficient detail in the description would not allow for its unambiguous reconstruction (e.g., Birkhofer et al., 2011; Graenitz and Bauer, 2000). A variety of coarsening algorithms may be an additional source of uncertainty when summarizing the results of many studies in subsequent meta-analyses. Consequently, it is important to know how large an error can be caused by combining data obtained using different scales and different algorithms.

For a more accurate estimate of bait consumption, colorimetric measurements of plant protein concentrations in bait residues after exposure were suggested instead of a visual score (Godwin and O'Neill, 2007). This approach, as well as the more laborious residual mass measurement, is certainly the most accurate possible and the least influenced by the operator. However, we do not know of any detailed description of this modification, nor of any attempts to apply it. To the best of our knowledge, all BLT applications are based solely on visual estimates of bait consumption in discrete scales.

Obviously, the coarsening of the scale may shift the estimate of the bait consumption rate, i.e. understate or overestimate it. Of course, the absolute values of feeding activity are not important *per se*, but only when comparing control ('good') and experimental ('bad') sites (Römbke, 2014). However, if the value or, moreover, the direction of the bias differs in comparable sites, it may lead to a shift in the final value of the difference between the sites and, consequently, to erroneous conclusions. Taking into account that under strong impacts, for example metal pollution, not only the abundance, but also the composition of soil macroinvertebrate communities change dramatically (Vorobeichik et al., 2019), such shifts are possible. A question arises: how much can algorithms of scale coarsening influence the inferences about the differences between control and experimental sites? The aim of our work is to find an answer to this question.

In this paper, we compare a five-point scale and the coarser scales. We consider a five-point scale to be the reference scale that provides the most precise information on bait consumption compared to other scales. To characterize it, we have analyzed the relationship between the score, i.e. the proportion of bait consumed by area, and the proportion of bait consumed by mass, which most accurately characterizes the rate of consumption.

We compare control area with sites polluted by copper smelter emissions. At polluted sites, the abundance of detritivores is sharply reduced due to high soil toxicity (see *3.1*). Control and polluted sites should therefore vary greatly regardless of the scale used for BLT. Consequently, we are solving an 'exercise' where we already know the answer, namely, the direction of the difference between the control and polluted sites and the fact that the difference is quite large. Based on this, we can determine which of the ways to solve this 'exercise' is correct and which is not, i.e. which of the algorithms of coarsening yields the smallest error. Since it is obvious that such an error depends on the cases of partial bait consumption, we analyze how common they are.

We consider BLT outcomes for several years, which differ in weather conditions and therefore in the factors influencing the feeding activity, i.e. soil temperature and humidity (Gongalsky et al., 2008; Joschko et al., 2008; Rożen et al., 2010). In this case, these factors should be considered as confounding because pollution is the main factor under study. Accordingly, such a comparison makes it possible to assess how stable the conclusions about the difference between the sites are when using different scales.

2. Theoretical background

Without losing in general, let us consider the calculation of the feeding activity value for one bait-lamina strip with a set of holes. The value of feeding activity in the five-point scale (S5) can be expressed as follows:

$$S5 = 0f_0 + 0.25f_{0.25} + 0.5f_{0.5} + 0.75f_{0.75} + 1f_1 \quad (1)$$

where f_0, $f_{0.25}$, $f_{0.5}$, $f_{0.75}$, and f_1 are frequencies of holes with 0, 0.25, 0.5, 0.75, and 1, respectively. The value of feeding activity in rough scales can also be represented through these frequencies, but with other weights. The algorithms of coarsening differ depending on what exactly substitutes this or that intermediate score of the five-point scale, i.e. 0.25, 0.5, and 0.75. Further, in the designations for the three-point (S3) and two-point (S2) scales on each of the three intermediate positions the sign '+' indicates an overestimation relative to the five-point scale, the sign '−' – an underestimation, the sign '0' – no difference. The algorithms of coarsening are summarized in Table 1.

Two algorithms of coarsening are possible for S3. In one case all intermediate values of S5 are replaced by 0.5, i.e., S3(+0−). In the second case, an intermediate value closer to zero is replaced by 0, closer to one – by 1, and the value of 0.5 is not changed, i.e., S3(−0 +). Since the authors who used a three-point scale did not specify the subtleties of the procedure (Siebert et al., 2019; Thakur et al., 2018), we considered both variants of S3.

Four algorithms of coarsening are possible for S2. If all intermediate values are replaced by 1, this is the strongest overestimation of the rough scale compared to the reference one, S2(+++). This algorithm was suggested by Von Törne (1990). If all intermediate values are replaced by 0, this is the strongest underestimation, S2(− − −). This algorithm was used by Podgaiski et al. (2014). Finally, there are two transitional

Table 1
Algorithms that convert the reference scale (S5) into the rough scales and formulas for the absolute difference between S5 and rough scales.

Rough scale	Points of reference scale					Formula for the absolute difference between S5 and rough scale
	0	0.25	0.50	0.75	1	
S3(+0−)	0	0.5	0.5	0.5	1	$-(0.25\ f_{0.25} - 0.25\ f_{0.75})$
S3(−0 +)	0	0	0.5	1	1	$0.25\ f_{0.25} - 0.25\ f_{0.75}$
S2(+++)	0	1	1	1	1	$-0.75\ f_{0.25} - 0.5\ f_{0.5} - 0.25\ f_{0.75}$
S2(−++)	0	0	1	1	1	$0.25\ f_{0.25} - 0.5\ f_{0.5} - 0.25\ f_{0.75}$
S2(−−+)	0	0	0	1	1	$0.25\ f_{0.25} + 0.5\ f_{0.5} - 0.25\ f_{0.75}$
S2(− − −)	0	0	0	0	1	$0.25\ f_{0.25} + 0.5\ f_{0.5} + 0.75\ f_{0.75}$

variants, where one intermediate value is replaced by 0 and the others by 1, or vice versa, one is replaced by 1 and the others by 0. The first of them, S2(−++), is a variant recommended by the ISO 18311 (ISO, 2016), while the second, S2(−−+), as far as we know, has not been used.

Each of the rough scales can be easily expressed through the frequency of intermediate values, similar to formula (1). For example,

$$S3(+0-) = 0f_0 + 0.5f_{0.25} + 0.5f_{0.5} + 0.5f_{0.75} + f_1 \tag{2}$$

Such a representation makes it very easy to calculate the absolute difference between the reference and rough scale. For example,

$$S5 - S3(+0-) = 0.25f_{0.25} + 0.5f_{0.5} + 0.75f_{0.75} + f_1 - (0.5f_{0.25} + 0.5f_{0.5} + 0.5f_{0.75} + f_1) = \\ = -0.25f_{0.25} + 0.25f_{0.75} \tag{3}$$

Table 1 presents formulas for the absolute difference between S5 and other scales. Obviously, in all cases, the difference is determined exclusively by intermediate values. For the four scales, S3(+0−), S3(−0 +), S2(−++), and S2(−−+), depending on the specific combination of frequency of intermediate values, the difference may either be absent or negative or positive. For the two scales, regardless of the frequency combination, the difference is always different from zero and either always negative (S2(+++)) or always positive (S2(− − −)). For the two variants of the three-point scale, the difference is the same for the module, but differs in sign: [S5–S3(−0 +)] = – [S5–S3(+0−)].

It is convenient to evaluate the impact of pollution on feeding activity using the effect size index, which allows characterizing not only the statistical significance, but also the magnitude of differences between the polluted and control sites. As an effect size index, environmental applications often use log Response Ratio (RR), which has several useful properties, in particular, additivity (Hedges et al., 1999). To take advantage of this property, let us imagine the feeding activity measured on the rough scale through the value in S5 and the relative difference between the reference and rough scale (D). For S3 we have (for other scales in the same way)

$$S3 = S5\,D3, \text{ where } D3 = 1 - (S5 - S3)/S5 \tag{4}$$

Then RR for the polluted and control sites can be represented as follows:

$$RR(S3) = log\left(S3_{polluted}\ /\ S3_{control}\right) = log\left(S5_{polluted}\ D3_{polluted}\ /\ S5_{control}\ D3_{control}\right) = \\ log\left(S5_{polluted}\ /\ S5_{control}\right) + log\left(D3_{polluted}\ /\ D3_{control}\right) = RR(S5) + RR(D3) \tag{5}$$

Thus, the RR of the feeding activity measured in the rough scale is the sum of the RR of the activity measured in the reference scale and the RR of the relative difference between the reference and rough scale. In other words, if Ds are the same at the polluted and control sites, even if they are very large, then the second term in the formula (5) is zero, and RR for the rough and reference scale are equal to each other. In this case, it does not matter which scale to use, whether it is a more precise or a less precise one involving overestimating or underestimating, as the impact of pollution will still be estimated equally. Otherwise, i.e., when the second term in formula (5) is not equal to zero and its value is at least somehow comparable with the first term, we can come to various inferences about the pollution impact using the reference and rough scales.

It is necessary to pay attention to the fact that at averaging, i.e. at transition from one strip to their set, the specified decomposition into components is fair only when the geometric mean rather than the arithmetic mean is used.

So, for the existence of such considerable bias of the rough scale relative to the reference scale, which can lead to erroneous inferences, three conditions must be met: 1) intermediate points exist, i.e. bias can occur; 2) the relative differences between the reference and rough scale in polluted and control sites are not equal to each other, i.e. the bias can affect the statistical inference; 3) the terms of sum (5) are comparable by module, i.e. the bias is comparable to the effect under study. These conditions can be formalized as follows:

$$IP = (f_{0.25} + f_{0.5} + f_{0.75}) > 0 \tag{6}$$

$$D_{polluted} \neq D_{control} \quad \text{or} \quad RR(D) \neq 0 \tag{7}$$

$$|RR(D)| > k|RR(S5)|, \quad \text{where } 1 > k > 0 \tag{8}$$

Conditions (6) and (7) should be considered necessary but not sufficient, and condition (8) necessary and sufficient. The strict finding of the k value represents a separate problem, far beyond our consideration. Probably, the lower boundary of k lies within the range of 0.1–0.3. Thus, the aim of our work can be reformulated as an empirical check of the fulfillment of conditions (6), (7), and (8).

3. Materials and methods

3.1. Study area

The BLT is carried out in the southern taiga, in the spruce-fir forest, in the area affected by long-term emissions from the Middle Ural Copper Smelter (Revda, Sverdlovsk region). Until recently, this factory was the largest source of air pollution in Russia: in the 1980s, its emissions exceeded 220,000 tons of pollutants per year (Vorobeichik and Kaigorodova, 2017). Control sites are located 20 km (one study site) and 30 km (3 study sites) west of the smelter, while polluted sites are located 1 km (one study site) and 2 km (3 study sites) west of the smelter, a total of 8 study sites. The soil contamination did not differ considerably among sites within the control area or sites within the polluted area. The study sites were permanent, with a size of 10 × 10 m and a distance of 100–150 m between them.

Long-term (since 1940) deposition of metal(loid)s (Cu, Pb, Cd, Zn, Fe, As, etc.) near the smelter resulted in a 10–100-fold excess of their background concentrations in the upper soil horizons (Korkina and Vorobeichik, 2018; Vorobeichik and Kaigorodova, 2017). The combination of such high concentrations with soil acidification due to sulfur dioxide emissions, with naturally slightly acidic soils, has had particularly dramatic consequences for terrestrial ecosystems and especially soil biota (Korkina and Vorobeichik, 2018; Mikryukov et al., 2020; Vorobeichik et al., 2019, 2014). Soil, vegetation and soil fauna indicators that are important for interpretation of the results are presented in Table 2.

It should be noted that in comparison with the control sites,

Table 2
Characteristics of the studied areas.

	Area and distance to copper smelter, km			
	Control		Polluted	
	30	20	2	1
Location	N 56°47′51″ E 59°25′03″	N 56°49′11″ E 59°34′33″	N 56°50′50″ E 59°51′41″	N 56°50′37″ E 59°52′44″
Altitude, m a.s.l.	400	380	415	370
Landscape description	Spruce-fir forest on the flattened lower part of the eastern slope of the Kirgishan Ridge	Spruce-fir forest on the gentle western midslope of a small mountain (440 m a.s.l.)	Spruce-fir forest on the lower gentle part of the eastern slope of the Shaitan Ridge	
Stand description [1]				
composition	*Ab.s.* – 50%, *Pic.o.* – 20%, *Pop.t.* – 20%, *Bet.* – 10%.	*Ab.s.* – 40%, *Pic.o.* – 30%, *Bet.* – 20%, *Pop.t.* – 10%.	*Pic.o.* – 50%, *Ab.s.* – 40%, *Bet.* – 10%.	*Pic.o.* – 50%, *Ab.s.* – 30%, *Bet.* – 10%, *Sal.* – 10%.
age, year	100	100	77	77
stock, m^3/ha	413	523	200	113
Dominant species of herbaceous layer [2]	*Oxalis acetosella*, *Dryopteris* spp., *Calamagrostis arundinacea*, *Aegopodium podagraria*, *Ajuga reptans*		*Agrostis capillaris*	
Soil type [3]	Albic Retisol (Cutanic)		Stagnic Retisol (Cutanic, Toxic)	
Prevailing humus form [4]	Dysmull		Eumor	
Thickness of forest litter* [4], cm	1.5 ± 0.1		5.8 ± 0.2	
Acid-soluble concentration in forest litter* [4], μg/g:				
Cu	37.3 ± 4.3		3484.3 ± 543.1	
Pb	67.3 ± 8.3		2462.5 ± 327.0	
Cd	2.4 ± 0.2		16.6 ± 2.6	
pH (water) in forest litter * [4]	5.9 ± 0.1		4.9 ± 0.1	
Abundance**, ind./m^2:				
earthworms (excluding cocoons) [5]	238 ± 25		1.1 ± 1.2	
enchytraeids [5]	1005 ± 138		4.1 ± 2.6	
diplopods [5]	11.3 ± 5.3		3.4 ± 1.9	
collembolans [6]	35333 ± 3403		12376 ± 2349	

Note: (a) * – ±SE, n = 5, study site as replication, ** – ±SE, n = 9–11 study site × year as replication; (b) *Ab.s.* – *Abies sibirica*, *Pic.o.* – *Picea obovata*, *Pop.t.* – *Populus tremula*, *Bet.* – *Betula* spp., *Sal.* – *Salix* spp.; (c) References: [1] – Bergman and Vorobeichik (2017), [2] – Mikryukov and Dulya (2017), [3] – Vorobeichik and Kaigorodova (2017), [4] – Korkina and Vorobeichik (2018), [5] – Vorobeichik et al. (2019), [6] – Kuznetsova (2009).

earthworms, i.e. the key group of soil decomposers, are absent in polluted areas (Vorobeichik et al., 2019, 2020); the abundance of other soil detritivores has been sharply reduced, in particular enchytraeids, diplopods, mollusks (Vorobeichik et al., 2019), and collembolans (Kuznetsova, 2009). In our previous studies, we have found a decrease in the feeding activity of soil animals by the BLT (Vorobeichik and Bergman, 2020; Vorobeichik et al., 2007). The composition and diversity of soil microflora have also been changed (Mikryukov and Dulya, 2017; Mikryukov et al., 2020) and the rate of cellulose decomposition has been reduced (Vorobeichik and Pishchulin, 2011). Adverse effects of metal pollution can be seen with the naked eye: in polluted sites, the accumulation of a thick layer of forest litter without signs of its processing by soil animals is observed (Korkina and Vorobeichik, 2018), which in turn negatively affects the diversity and recovery of herbaceous vegetation (Vorobeichik et al., 2014).

Although emissions have almost ceased since 2010 as a result of the factory reconstruction, there has been no reduction in soil metal content in highly polluted areas (Vorobeichik and Kaigorodova, 2017), or revegetation (Vorobeichik et al., 2014). However, with lower levels of pollution due to the normalization of soil acidity (Vorobeichik and Kaigorodova, 2017), the recovery of soil macroinvertebrates has started in the last few years: in particular, earthworms (Vorobeichik et al., 2019, 2020) have advanced closer to the smelter. However, they are still practically absent in the polluted sites studied in this work.

3.2. Data collection

BLTs were performed twice every growing season (May-June and August-September) for 5 years (2015–2019) and only once (September) in 2019. Following the recommendation of ISO 18311 (ISO, 2016) in this paper we included in the analysis only those rounds where the average feeding activity in the layer with maximum activity (the upper half of the strip, see 3.4) was higher than 30% on at least one study site of the control area. The rounds that met this condition turned out to be 7 out of 9 (Table 3).

During this time, the technical performance of the BLT did not change. Plastic strips, 160 mm × 6 mm × 1.5 mm, with 16 bi-conical apertures 1.5 mm in diameter arranged every 5 mm were used. A mixture of nettle leaf powder and microcrystalline cellulose (3:7 w/w) was used as bait-material. The strips filled with the bait were dried at room temperature for two days.

On each study site 25 strips were exposed at a distance of 0.5–1.0 m from each other. The locations of the strips within the study site were chosen randomly, excluding areas around trunks of large trees with a radius of about 1 m, as well as areas of visible soil disturbance. In different years, the locations were not the same.

At the installation point incisions in the forest litter and mineral horizon were made previously with a sharp knife. The strips were installed strictly vertically; the upper hole corresponded to a depth of 0.5 cm from the litter surface, the lower hole – 8.0 cm. Within the round, the strips were installed and removed in one day at all study sites. The duration of exposure was 7 days, except for two rounds in 2015, when it was equal to 9 and 11 days (see Table 3).

3.3. Measurements in the laboratory

In order to prevent the bait-material from drying out before measurements, strips were stored in plastic bags in a refrigerator at 5°Celsius (for not more than 7–8 days). Immediately after the extraction from the refrigerator, the perforation of each hole was visually assessed. The

Table 3
Dates of BLTs and maximal feeding activity (%) in control area (upper half of strip).

Year/round	Start date	End date	Duration, days	Maximal feeding activity	Inclusion in the analysis
2015/1	10.08	21.08	11	86	yes
2015/2	01.09	10.09	9	69	yes
2016/1	25.05	01.06	7	41	yes
2016/2	29.08	05.09	7	24	no
2017/1	22.05	29.05	7	34	yes
2017/2	01.09	08.09	7	58	yes
2018/1	04.06	11.06	7	65	yes
2018/2	04.09	11.09	7	16	no
2019/1	02.09	09.09	7	64	yes

following five-point scale was used: 0 – untouched; 0.25 – approximately 25% of the hole area eaten; 0.5 – 50%; 0.75 – 75%; 1 – hole is fully empty. All measurements for all rounds were made by one operator.

Some of the strips exposed in 2019 in the control (12 strips) and polluted (6 strips) sites were used to analyze the relationship between the score and the weight of the remaining bait (18 strips in total). The strips were selected randomly, excluding 'homogeneous' variants, i.e. when all holes in the strip were with a score of 0 or 1. After the evaluation of the perforation, the strips were dried at room temperature for two days. After that, the remaining bait was removed from each hole with a dissecting needle. In addition, to determine the initial mass of the bait, it was similarly extracted from 60 holes of unexposed strips, which were prepared in a standard manner. The extracted bait was weighed individually for each hole on analytical scales HR-120 (A&D Company, Japan) with an accuracy of 0.1 mg.

3.4. Data analysis

The relationship between the feeding activity score and the mass of the remaining bait was approximated by linear regression. The difference between the initial mass of the bait and the mass at zero point after exposure was estimated by the Student *t*-test.

Feeding activity was considered both for the whole strip (16 holes, layer 0.5–8.0 cm) and separately for its two halves: the top (8 top holes, layer 0.5–4.0 cm) and the bottom (8 bottom holes, layer 4.5–8.0 cm). From the obtained array of values in a five-point scale six arrays with values in rough scales according to the algorithms presented in Table 1 were formed.

Relative differences D were calculated by the formula (4) and compared between the control and polluted areas using the Mann-Whitney test, while multiple comparisons between the scales were performed using the Dunn test (nonparametric alternative to the Tukey test). The correlation between the average feeding activity and the total percentage of intermediate points (IP) was estimated with the Spearman coefficient. The use of ordinal tests was dictated by the lack of information about the theoretical distribution of the variables under consideration.

The decomposition of RR into components was done by formula (5) using geometric mean. In 11 cases out of 168 (7 rounds × 8 study sites × 3 variants, i.e. the whole strip, the upper and lower halves) the geometric mean was not determined because there were no values with a score of 1 on the five-point scale. In these cases, the number of holes with the point 1 equal to 0.1 was conditionally accepted. Since the variance of RR using the geometric mean is not known, RR was also calculated in a standard way through the arithmetic mean. We used unbiased estimator RR^{Δ} (Lajeunesse, 2015), ARPobservation v.1.2.0 in R package 3.6.3 (Pustejovsky, 2019). The feeding activity values were preliminarily transformed as $\arcsin\sqrt{x}$. The effect sizes index for which the 95% confidence interval (CI) did not include zero was considered as significant.

The study site, i.e. the average activity value of 25 strips, was counted as a replicate. Since the distances between the strips within the study site are considerably (100–300 times) smaller than the distances between the sites, it is correct to consider the study site as a true replicate.

4. Results

The mass of the bait in the holes with score 0 after exposure (±SD, 1.51 ± 0.22 mg) was significantly lower than the initial bait mass (1.93 ± 0.33 g), t = 11.0, p≪0.0001, weight loss was 21.8%. There is a clear linear relationship between the points and remaining bait mass: F (1;268) = 2394, p≪0.0001, R^2 = 0.90 (Fig. 1). Slopes and intercepts are practically equal to each other by modulus (±SE, −1.498 ± 0.031 and 1.497 ± 0.015, respectively), i.e. the transition from absolute mass values to the fraction of the remaining mass, indicates a direct proportionality between the fraction of the bait disappearance, estimated by the area decrease, and the fraction estimated by the mass decrease.

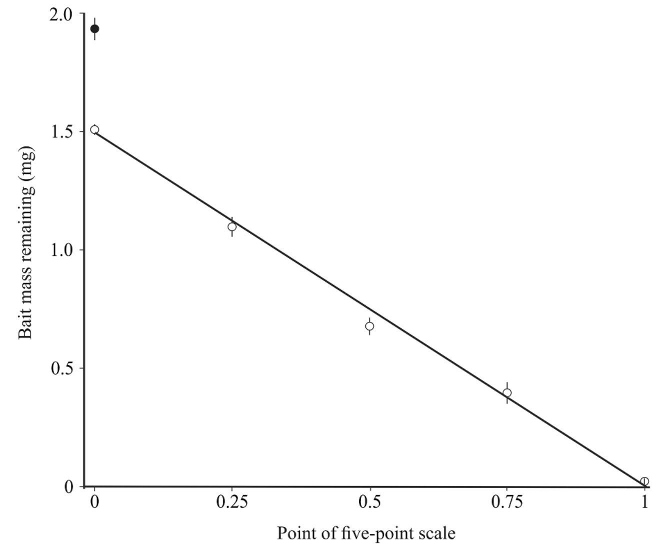

Fig. 1. Relationships between bait mass remaining (average ± standard error) and score of five-point scale (for scores 0, 0.25, 0.5, 0.75, and 1 *n* = 147, 31, 30, 21, and 41, respectively). The filled dot represents the initial bait mass before exposure (*n* = 60); empty dots represent bait mass after exposure. The linear regression curve obtained for bait mass after exposure is illustrated.

Intermediate points occupy an essential part of all non-zero values (Fig. 2). In the control area, IP is on the average (±SE, n = 28, 7 rounds × 4 sites) 28 ± 3% (11 to 68%), while in the polluted area it is almost twice as much: 66 ± 4% (21 to 100%). The smaller the average activity value, the more IP there is: for the control area Spearman's coefficient is –0.73, p < 0.0001, for the polluted area correlation is –0.50, p = 0.007.

The value of the relative difference between the reference and rough

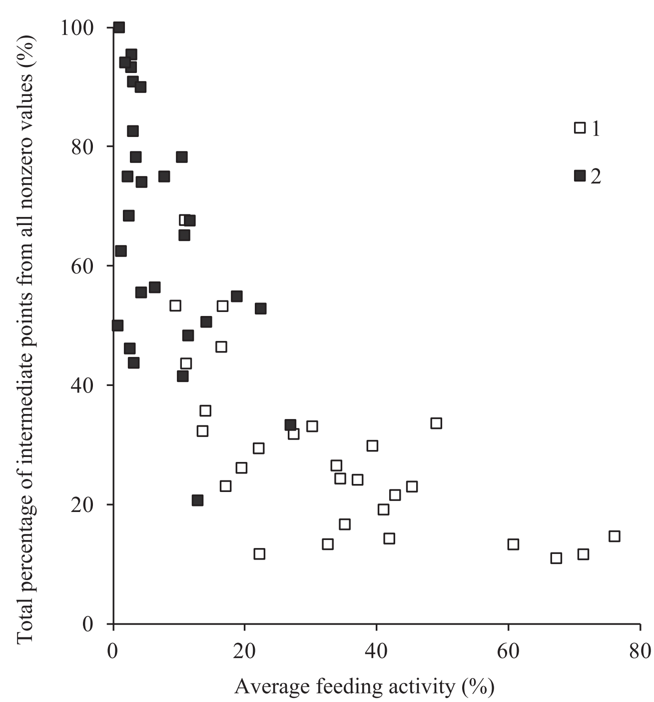

Fig. 2. Relationships between total percentage of intermediate points (0.25, 0.5, and 0.75) from all nonzero values and average feeding activity for the whole strip: (1) control area; (2) polluted area.

scale differs between the scales; also it is different in the control and polluted areas (Fig. 3). According to |D|, the scales are lined up in the ascending order as follows:

S3(+0−) = S3(−0 +) ≈ S2(−++) > S2(− −+) > S2(− − −) ≈ S2 (+++).

The first extreme group of the series, i.e., S3 and S2(−++), significantly differs from the second extreme group, i.e., S2(− − −) and S2 (+++), by Dunn's test (at least with $p = 0.0134$), and the differences within these groups are insignificant. This series is the same for both control and polluted areas, though the conspecific values are far from the diagonal: in almost all cases, the relative difference is higher in the polluted sites compared to the control ones. The difference is particularly large for the extreme scales of the series (average for all rounds, ±SE, $n = 7$): |D| for S2(+++) in the control area is 18 ± 4% (maximum 38%), while in the polluted area it is 66 ± 12% (118%), and for S2 (− − −) it is 16 ± 3% (33%) and 50 ± 6% (71%), respectively. |D| for S3 and S2(−++) are small: in the control area it is 1 ± 1% and 3 ± 1% (maximum 4% and 6%), while in the polluted area it is 8 ± 3% and 7 ± 2% (24% and 15%). The differences between the control and polluted areas are significant for all scales (Mann-Whitney test, at least $p = 0.0127$, $n = 7$), except for S2(−++).

The decomposition by formula (5) allows us to identify the reasons for the bias of the effect size index of the rough scales relative to the reference scale (Fig. 4). RR(S5), i.e. the first term of the sum, is in all cases negative, which indicates an adverse effect of pollution on feeding activity. However, RR(D), i.e. the second term of the sum, is not equal to zero, and shifts the effect size index of the rough scales to one side or another. The bias is minimal for S3 and S2(−++) scales: by module (average for all rounds and variants, ±SE, $n = 21$) it is only 4 ± 1% (from + 0.1% to –15.8%). For the other scales, the bias is quite considerable. For S2(− − −) the bias is always negative, i.e. it overestimates the effect size, by module it is 48 ± 9% (from –2.2% to –167.3%). For S2(− −+) the bias is also negative, but less large (23 ± 4%, from 0 to 70%). For S2(+++), on the contrary, due to different signs of the terms of the sum (5), the effect size is underestimated by 20 ± 3% (from 1.3% to 73.1%).

Comparison of 95% CI RR$^{\Delta}$ for S5 and for rough scales shows the same picture in the overwhelming majority of cases (Fig. 5). However, in several cases the transition from S5 to rough scales shifts the statistical inference: a significant effect ceases to be significant (S2(+++) and S2 (−++) for the whole strip, S2(− − −) and S2(−++) for the lower half) or, conversely, an insignificant effect becomes significant (S2(− − −) for the upper half, S2(+++) for the lower half).

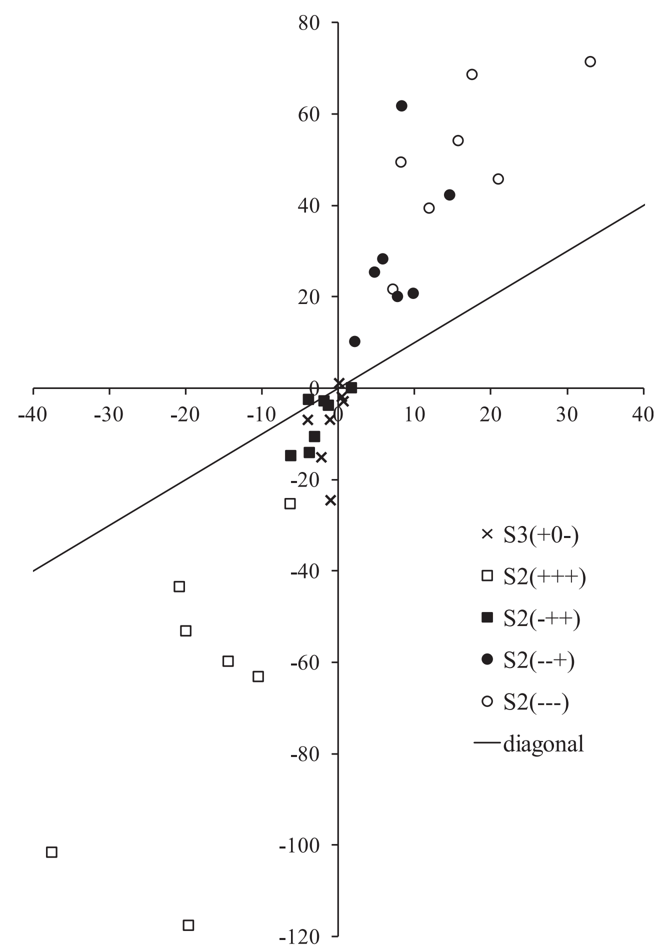

Fig. 3. Relative difference (%) between the reference and rough scales on control (X-axis) and polluted (Y-axis) area (for whole bait-lamina strip). S3 (−0+) is not shown because it is symmetrical to S3(+0−).

5. Discussion

A clear linear relationship between a five-point scale and the bait mass remaining after the exposure is very important for the interpretation of the results (see Fig. 1). Firstly, this dependence allows interpreting BLT outcomes specifically as the rate of consumption of plant material by soil invertebrates and not just as a conventional index. In other words, the average activity value of a set of strips on a five-point scale is equivalent to the average mass of bait consumed in a fixed time. It should be noted that when using a rough two-point scale, i.e. in terms of the proportion of empty holes, this interpretation is not obvious. Considering the portion of the bait consumed as the rate of its disappearance allows directly integrating the BLT results into the conceptual framework for the analysis of organic matter decomposition (Berg and McClaugherty, 2008). Secondly, it makes it reasonable to apply standard statistical methods suitable for variables measured on a ratio scale, such as mass. In particular, it is possible to use Response Ratio instead of Risk Ratio or Odds Ratio as would be necessary in the case of a variable measured in an ordinal scale. Thirdly, it ensures the satisfactory accuracy of the five-point scale for estimating bait consumption rate giving grounds for assuming it as a reference scale. Fourthly, it indicates a good reproducibility of the visual measurement procedure, since the mass variation of the remaining bait within a single point was not very large: it rather arises from the discreteness of the five-point scale applied to the measurement of a continuous value, rather than from visual evaluation errors.

The reasons for the significant difference between the initial mass of the bait and the mass after exposure at a zero point of ~20% in 7 days of exposure can be as follows: 1) microbial degradation of the material, 2) leaching of soluble compounds and/or mechanical losses, 3) bait consumption by invertebrates that did not result in perforation. It is well known that in litter bag tests, even under exclusion of soil fauna, the mass of plant material decreases exponentially over time, i.e., the greatest losses occur during the early stages of decomposition. This is due to both chemical leaching of easily soluble substances and microbial degradation of easily biodegradable compounds such as sugars and proteins (Berg and McClaugherty, 2008). However, such losses are clearly insufficient to be visually assessed by bait perforation in such a short exposure. In defaunated soil, but with active microflora, the perforation of holes is absent (Gongalsky et al., 2008; Helling et al., 1998; Von Törne, 1990). On the other hand, and in the absence of perforation, i.e., when there is no opening in the entire thickness of the bait, traces of its consumption by soil invertebrates can be seen, and sometimes the volume of bait consumed may even exceed the volume at minimum but not zero perforation (Von Törne, 1990). Without additional experiments, we cannot estimate the contribution of these causes. However, it was logical to take the mass of the exposed bait at zero point as the initial mass when analyzing the dependence of the remaining mass on the score rather than the initial mass before the exposure.

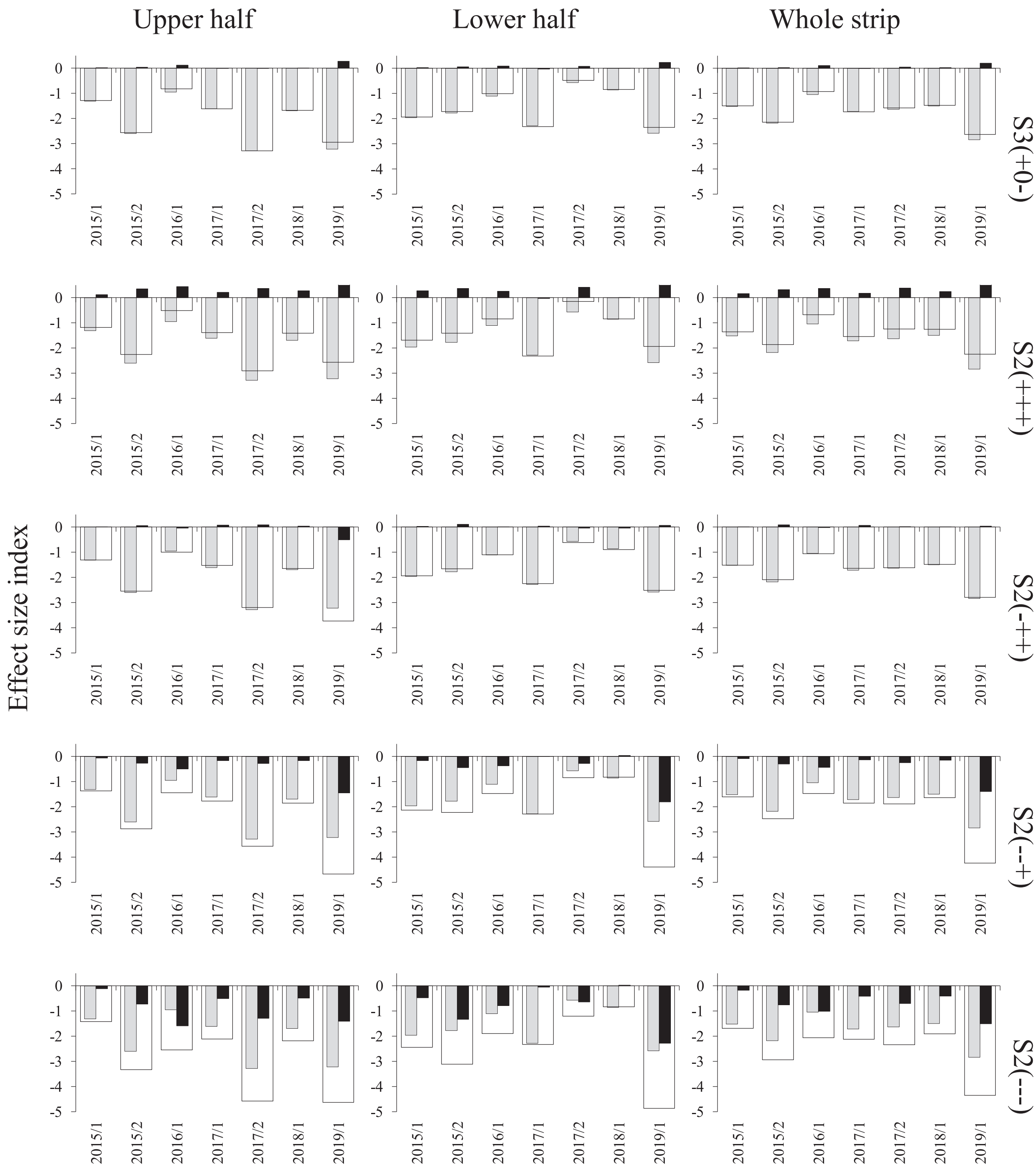

Fig. 4. Effect size index (ln Response Ratio using the geometric mean) for reference scale (gray), relative difference between the reference and rough scale (black) and rough scale (empty) for upper half of strip, lower half of strip, and whole strip.

The first of the necessary conditions for the existence of bias of the rough scales in relation to the reference one (formula 6) was fulfilled. The percentage of intermediate points of feeding activity turned out to be very large: from one-third to two-thirds of the number of non-zero values, and in some cases up to 100% (see Fig. 2). Firstly, it testifies to the importance of the topic we are discussing. Secondly, it means that the 'Procrustean bed' of the two-point scale is too 'uncomfortable' to measure feeding activity: such a strong discreteness badly corresponds to the continuous value of bait consumption. Thirdly, cases of too much 'freedom of will' for operators in the visual evaluation of bait consumption are quite frequent, which can cause considerable errors and at least requires a clear regulation of measurements.

The total percentage of intermediate points is particularly high at low feeding activity values, which are more numerous in the polluted area, as observed in our study (see Fig. 2). Differences between sites are likely related to changes in the spectrum of bait consumers (Vorobeichik and Bergman, 2020). Near the studied smelter, earthworms play a major role in the control sites, but they are almost absent in the polluted area (Vorobeichik et al., 2019). Obviously, earthworms are more likely to fully eat out the bait in a particular hole once it is detected than other

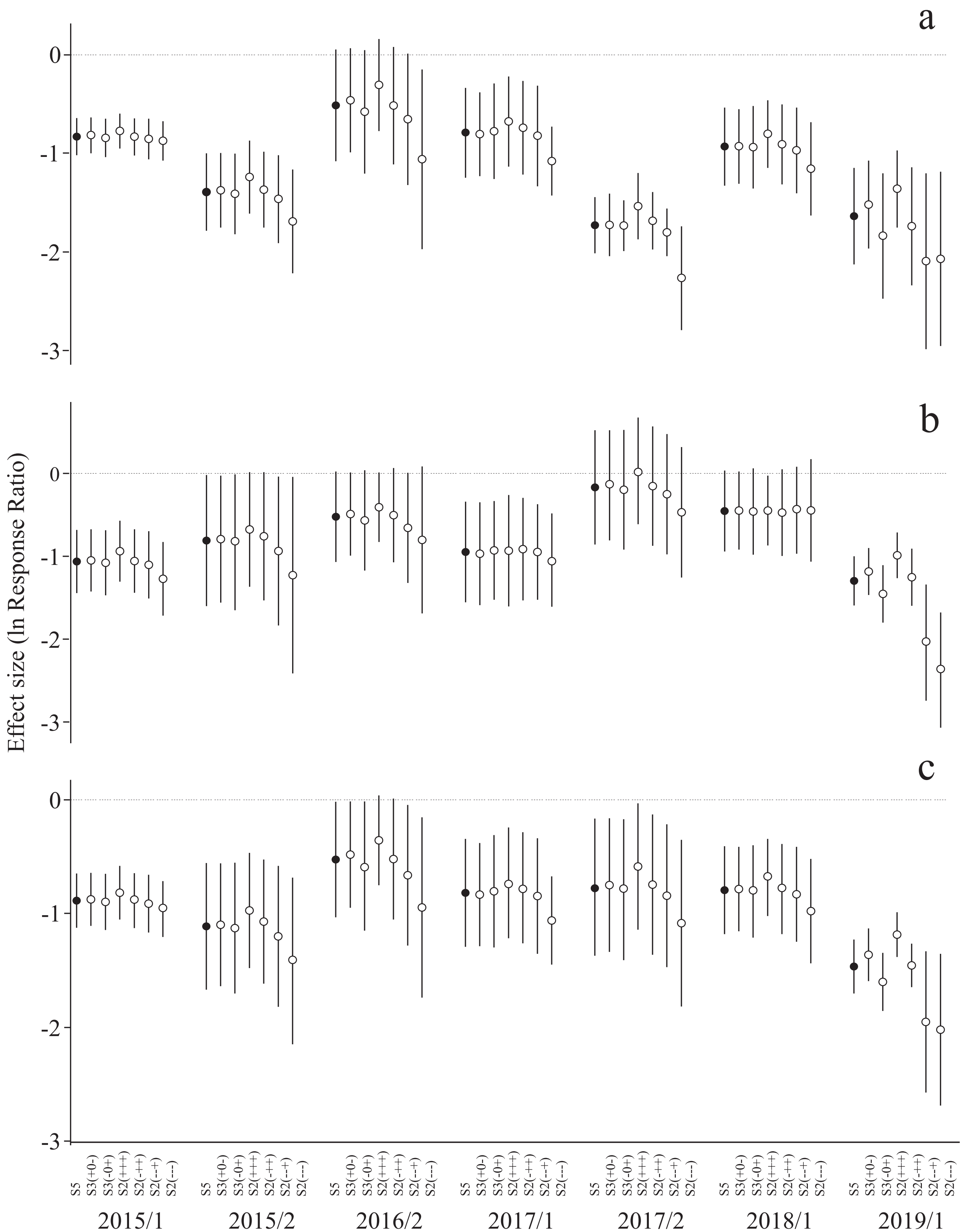

Fig. 5. Effect size index (RR$^{\Delta}$, ±95% CI) for reference scale (empty dot) and rough scales (filled dots): (a) upper half of bait-lamina strip; (b) lower half of bait-lamina strip; (c) whole strip.

potential consumers, e.g. microarthropods, enchytraeids, larvae of Nematoceran flies, etc. at least because of their larger size. Smaller organisms need more time to detect bait and to consume it, which is why there are so many holes with 'unfinished meal' in polluted areas. The frequency distribution of points in the five-point scale can even be an indirect indicator of the decomposer size distribution.

The second necessary condition (formula 7) was also fulfilled. The relative difference between the reference and rough scales turned out to be 2–6 times greater in the polluted area in comparison with the control area (see Fig. 3).

Finally, the third, sufficient, condition (formula 8) was also fulfilled. It turned out that sometimes RR(D) is almost equal to RR(S5) or a little less. In other words, RR(D) can be so large that it can shift the effect size index of the rough scales significantly in comparison with the reference scale. This bias is especially large for S2(−−−) and, to a lesser extent, for S2(−−+), i.e., when using these scales, the effect size index can almost double in comparison with the true value. Thus, the risk of a type I error in statistical hypothesis testing increases: if there is no effect, you may assume that it exists. For another extreme variant, S2(+++), on the contrary, the true value of effect size index is understated, sometimes by almost a third, which increases the risk of a type II error: if there is an effect, it can be considered absent. The situation is rather permissible only for S3 and S2(−++): for them the bias of effect size index relative to the true value is small (on average less than 5%) and they can probably be neglected.

After a transition to the standard variant of RR calculation on the basis of arithmetic average our fears were confirmed: in a number of cases rejection errors are possible (see Fig. 5). It should be noted that we deliberately did not perform *p*-value correction for multiple hypothesis checking (for example, by False Discovery Rate control), because our goal was to imitate the decision making procedure in a single BLT.

Thus, we have demonstrated that the use of rough scales in evaluating the impact of pollution on feeding activity can be a source of artifacts. The main practical conclusion of our work is that the scales are arranged as follows as their preferability decreases: five-point scale or more precise scale, three-point scale, and two-point scale recommended by ISO 18311 (ISO, 2016). For three-point scales, it does not matter which algorithm distributes the partial bait consumption between the points. The use of other scales should be avoided. In any case, it is important to provide comprehensive information about the scale used in the work, i.e. the thresholds for bait consumption corresponding to a specific score.

BLT outcomes are usually characterized by high spatial heterogeneity even in the absence of any adverse impacts (Gongalskii et al., 2003; Irmler, 1998), and damaging effects further increase variability (Joschko et al., 2008; Vorobeichik and Bergman, 2020). Therefore, the already high uncertainty in the estimation of average feeding activity should not be increased using unacceptable rough scales, especially since more adequate scales do not require any additional effort.

6. Conclusions

The amount of bait consumption by soil organisms is a continuous variable that is evaluated as a discrete variable to speed up and facilitate measurements. We have shown that the number of scale points and thresholds corresponding to specific points is very important. And what is important is not that the absolute value of feeding activity depends on these choices – this is trivial. It is crucial that by using different scales one can come to dissimilar conclusions about the influence of the investigated factor on the feeding activity. This means that the choice of scale is not just a secondary technical issue that can be painlessly ignored.

We have established that a five-point scale satisfactorily estimates the continuous bait consumption value, indicating that it can be accepted as a reference scale. The feeding activity index evaluated on a five-point scale is a good surrogate for the bait's disappearance rate.

Comparison of the five-point scale and rough scale simulates the decision making by the operator when attributing the amount of bait consumption in a particular hole to a point on a three-point or two-point scale. Strictly speaking, the difference between the reference scale and the rough scale depends on the continuous variable domain where the uncertainty of the operator's decision making is particularly high. It turned out that cases corresponding to this domain are not uncommon: they make up from one-third to two-thirds, and sometimes up to 100% of all cases. This suggests that the issue under consideration is not farfetched.

We have shown that two scales are the least dangerous in terms of artifacts, namely the three-point scale and the scale recommended by ISO 18311 (ISO, 2016). Therefore, we recommend using them as an acceptable surrogate for a five-point scale. Other two-point scales are fraught with risk of artifacts, so they should be avoided.

On the one hand, our results relate to a polluted area near a particular smelter. On the other hand, metal pollution is similar to the effects of other adverse factors such as pesticides (Förster et al., 2011) or drought (Siebert et al., 2019; Thakur et al., 2018). Therefore, in our opinion, the conclusions relate to comparison of sites with high and low feeding activity of soil detritivores irrespective of the factors genesis.

In this paper we have not analyzed the reasons for differences between years and rounds and between upper and lower halves of strips, as these are separate tasks. However, consideration of different years, seasons and soil layers has increased the diversity of situations. We have demonstrated that artifacts do not occur frequently, but even if they are relatively rare, their danger should not be neglected.

CRediT authorship contribution statement

Evgenii L. Vorobeichik: Conceptualization, Methodology, Formal analysis, Visualization, Writing - original draft, Funding acquisition. **Igor E. Bergman:** Investigation, Formal analysis, Data curation, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The data collection in 2015–2018 was carried out under the State Assignment of the Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, with funding provided by the Ministry of Science and Higher Education of the Russian Federation. The data collection in 2019, data analysis and preparation of the manuscript were carried out with the financial support of the Russian Foundation for Basic Research (grant number 19-29-05175). We are grateful to Maxim Zolotarev for drawing the graphical abstract, to Vladimir Mikryukov and Olesya Dulya for the discussion and comments on the text of the manuscript.

References

André, A., Antunes, S.C., Gonçalves, F., Pereira, R., 2009. Bait-lamina assay as a tool to assess the effects of metal contamination in the feeding activity of soil invertebrates within a uranium mine area. Environ. Pollut. 157 (8-9), 2368–2377. https://doi.org/10.1016/j.envpol.2009.03.023.

Bart, S., Roudine, S., Amossé, J., Mougin, C., Péry, A.R.R., Pelosi, C., 2018. How to assess the feeding activity in ecotoxicological laboratory tests using enchytraeids? Environ. Sci. Pollut. Res. 25 (34), 33844–33848. https://doi.org/10.1007/s11356-018-1701-3.

Berg, B., McClaugherty, C., 2008. Plant litter. Decomposition, humus formation, carbon sequestration. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-38821-7.

Bergman, I.E., Vorobeichik, E.L., Ermakov, A.I., 2017. The effect of megalopolis environment on the feeding activity of soil saprophages in urban forests. Eurasian Soil Sci. 50 (1), 106–117. https://doi.org/10.1134/S1064229317010021.
Bergman, I.E., Vorobeichik, E.L., 2017. The effect of a copper smelter emissions on the stock and decomposition of coarse woody debris in spruce and fir woodlands. Contemp. Probl. Ecol. 10 (7), 790–803. https://doi.org/10.1134/S1995425517070022.
Birkhofer, K., Diekötter, T., Boch, S., Fischer, M., Müller, J., Socher, S., Wolters, V., 2011. Soil fauna feeding activity in temperate grassland soils increases with legume and grass species richness. Soil Biol. Biochem. 43 (10), 2200–2207. https://doi.org/10.1016/j.soilbio.2011.07.008.
Boshoff, M., De Jonge, M., Dardenne, F., Blust, R., Bervoets, L., 2014. The impact of metal pollution on soil faunal and microbial activity in two grassland ecosystems. Environ. Res. 134, 169–180. https://doi.org/10.1016/j.envres.2014.06.024.
Eisenhauer, N., Herrmann, S., Hines, J., Buscot, F., Siebert, J., Thakur, M.P., 2018. The dark side of animal phenology. Trends Ecol. Evol. 33 (12), 898–901. https://doi.org/10.1016/j.tree.2018.09.010.
Filzek, P.D.B., Spurgeon, D.J., Broll, G., Svendsen, C., Hankard, P.K., Parekh, N., Stubberud, H.E., Weeks, J.M., 2004. Metal effects on soil invertebrate feeding: measurements using the bait lamina method. Ecotoxicology 13 (8), 807–816. https://doi.org/10.1007/s10646-003-4478-0.
Förster, B., Boxall, A., Coors, A., Jensen, J., Liebig, M., Pope, L., Moser, T., Römbke, J., 2011. Fate and effects of ivermectin on soil invertebrates in terrestrial model ecosystems. Ecotoxicology 20 (1), 234–245. https://doi.org/10.1007/s10646-010-0575-z.
Geissen, V., Gehrmann, J., Genssler, L., 2007. Relationships between soil properties and feeding activity of soil fauna in acid forest soils. Z. Pflanzenernähr. Bodenk. 170 (5), 632–639. https://doi.org/10.1002/jpln.200625050.
Godwin, H., O'Neill, K., 2007. A modified colorimetric bait-lamina method for estimating litter decomposition by soil microinvertebrates, Management of landscapes and ecosystems. Mid-Atlantic Chapter of the Ecological Society of America, York, Pennsylvania, p. 4.
Gongalskii, K.B., Pokarzhevskii, A.D., Savin, F.A., Filimonova, Z.V., 2003. Spatial distribution of animals and variation in their trophic activity measured using the bait-lamina test in sod-podzolic soil under a spruce forest. Rus. J. Ecol. 34, 395–404. https://doi.org/10.1023/A:1027312501091.
Gongalsky, K.B., Persson, T., Pokarzhevskii, A.D., 2008. Effects of soil temperature and moisture on the feeding activity of soil animals as determined by the bait-lamina test. Appl. Soil Ecol. 39 (1), 84–90. https://doi.org/10.1016/j.apsoil.2007.11.007.
Graenitz, J., Bauer, R., 2000. The effect of fertilization and crop rotation on biological activity in a 90 year long-term experiment. Bodenkultur 51, 99–105.
Griffiths, B.S., Römbke, J., Schmelz, R.M., Scheffczyk, A., Faber, J.H., Bloem, J., Pérès, G., Cluzeau, D., Chabbi, A., Suhadolc, M., Sousa, J.P., Martins da Silva, P., Carvalho, F., Mendes, S., Morais, P., Francisco, R., Pereira, C., Bonkowski, M., Geisen, S., Bardgett, R.D., de Vries, F.T., Bolger, T., Dirilgen, T., Schmidt, O., Winding, A., Hendriksen, N.B., Johansen, A., Philippot, L., Plassart, P., Bru, D., Thomson, B., Griffiths, R.I., Bailey, M.J., Keith, A., Rutgers, M., Mulder, C., Hannula, S.E., Creamer, R., Stone, D., 2016. Selecting cost effective and policy-relevant biological indicators for European monitoring of soil biodiversity and ecosystem function. Ecol. Ind. 69, 213–223. https://doi.org/10.1016/j.ecolind.2016.04.023.
Hedges, L.V., Gurevitch, J., Curtis, P.S., 1999. The meta-analysis of response ratios in experimental ecology. Ecology 80 (4), 1150–1156. https://doi.org/10.2307/177062.
Helling, B., Pfeiff, G., Larink, O., 1998. A comparison of feeding activity of collembolan and enchytraeid in laboratory studies using the bait-lamina test. Appl. Soil Ecol. 7 (3), 207–212. https://doi.org/10.1016/S0929-1393(97)00065-6.
Irmler, U., 1998. Spatial heterogeneity of biotic activity in the soil of a beech wood and consequences for the application of the bait-lamina-test. Pedobiologia 42, 102–108.
ISO, 2016. Method for testing effects of soil contaminants on the feeding activity of soil dwelling organisms — Bait-lamina test. 18311. International Organization for Standardization, Geneva.
Jänsch, S., Scheffczyk, A., Römbke, J., 2017. The bait-lamina earthworm test: a possible addition to the chronic earthworm toxicity test? Euro-Mediterr. J. Environ. Integr. 2 (1), 5 https://doi.org/10.1007/s41207-017-0015-z.
Joschko, M., Oehley, J., Gebbers, R., Wiemer, M., Timmer, J., Fox, C.A., 2008. A spatial approach to soil-ecological experimentation at landscape scale. J. Plant Nutr. Soil Sci. 171 (3), 338–343. https://doi.org/10.1002/jpln.200700088.
Korkina, I.N., Vorobeichik, E.L., 2018. Humus Index as an indicator of the topsoil response to the impacts of industrial pollution. Appl. Soil Ecol. 123, 455–463. https://doi.org/10.1016/j.apsoil.2017.09.025.
Kuznetsova, N.A., 2009. Soil-dwelling Collembola in coniferous forests along the gradient of pollution with emissions from the Middle Ural Copper Smelter. Russ. J. Ecol. 40 (6), 415–423. https://doi.org/10.1134/S106741360906006X.
Lajeunesse, M.J., 2015. Bias and correction for the log response ratio in ecological meta-analysis. Ecology 96 (8), 2056–2063. https://doi.org/10.1890/14-2402.1.sm.
Larink, O., Sommer, R., 2002. Influence of coated seeds on soil organisms tested with bait lamina. Eur. J. Soil Biol. 38 (3-4), 287–290. https://doi.org/10.1016/S1164-5563(02)01161-5.
Mikryukov, V.S., Dulya, O.V., 2017. Contamination-induced transformation of bacterial and fungal communities in spruce-fir and birch forest litter. Appl. Soil Ecol. 114, 111–122. https://doi.org/10.1016/j.apsoil.2017.03.003.
Mikryukov, V.S., Dulya, O.V., Modorov, M.V., 2020. Phylogenetic signature of fungal response to long-term chemical pollution. Soil Biol. Biochem. 140, 107644. https://doi.org/10.1016/j.soilbio.2019.107644.
Musso, C., Miranda, H.S., Soares, A.M.V.M., Loureiro, S., 2014. Biological activity in Cerrado soils: evaluation of vegetation, fire and seasonality effects using the "bait-lamina test". Plant Soil 383 (1-2), 49–58. https://doi.org/10.1007/s11104-014-2233-3.
Niemeyer, J.C., de Santo, F.B., Guerra, N., Ricardo Filho, A.M., Pech, T.M., 2018. Do recommended doses of glyphosate-based herbicides affect soil invertebrates? Field and laboratory screening tests to risk assessment. Chemosphere 198, 154–160. https://doi.org/10.1016/j.chemosphere.2018.01.127.
Pehle, A., Schirmel, J., 2015. Moss invasion in a dune ecosystem influences ground-dwelling arthropod community structure and reduces soil biological activity. Biol. Invas. 17 (12), 3467–3477. https://doi.org/10.1007/s10530-015-0971-7.
Podgaiski, L.R., da Silva Goldas, C., Ferrando, C.P.R., Silveira, F.S., Joner, F., Overbeck, G.E., de Souza Mendonça Jr, M., Pillar, V.D., 2014. Burning effects on detritivory and litter decay in Campos grasslands: grassland fire effects on decomposition. Austral. Ecol. 39 (6), 686–695. https://doi.org/10.1111/aec.12132.
Pustejovsky, J.E., 2019. ARPobservation: Simulating recording procedures for direct observation of behavior. R package version 1 (2).
Ritz, K., Black, H.I.J., Campbell, C.D., Harris, J.A., Wood, C., 2009. Selecting biological indicators for monitoring soils: a framework for balancing scientific and technical opinion to assist policy development. Ecol. Ind. 9 (6), 1212–1221. https://doi.org/10.1016/j.ecolind.2009.02.009.
Römbke, J., 2014. The feeding activity of invertebrates as a functional indicator in soil. Plant Soil 383 (1-2), 43–46. https://doi.org/10.1007/s11104-014-2195-5.
Römbke, J., Höfer, H., Garcia, M.V.B., Martius, C., 2006. Feeding activities of soil organisms at four different forest sites in Central Amazonia using the bait lamina method. J. Trop. Ecol. 22 (03), 313–320. https://doi.org/10.1017/S0266467406003166.
Rożen, A., Sobczyk, Ł., Liszka, K., Weiner, J., 2010. Soil faunal activity as measured by the bait-lamina test in monocultures of 14 tree species in the Siemianice common-garden experiment, Poland. Appl. Soil Ecol. 45 (3), 160–167. https://doi.org/10.1016/j.apsoil.2010.03.008.
Siebert, J., Sünnemann, M., Auge, H., Berger, S., Cesarz, S., Ciobanu, M., Guerrero-Ramírez, N.R., Eisenhauer, N., 2019. The effects of drought and nutrient addition on soil organisms vary across taxonomic groups, but are constant across seasons. Sci. Rep. 9 (1), 639 https://doi.org/10.1038/s41598-018-36777-3.
Simpson, J.E., Slade, E., Riutta, T., Taylor, M.E., 2012. Factors affecting soil fauna feeding activity in a fragmented lowland temperate deciduous woodland. Plos One 7, e29616. https://doi.org/10.1371/journal.pone.0029616.
Spehn, E.M., Joshi, J., Schmid, B., Alphei, J., Körner, C., 2000. Plant diversity effects on soil heterotrophic activity in experimental grassland ecosystems. Plant Soil 224, 217–230. https://doi.org/10.1023/A:1004891807664.
Thakur, M.P., Reich, P.B., Hobbie, S.E., Stefanski, A., Rich, R., Rice, K.E., Eddy, W.C., Eisenhauer, N., 2018. Reduced feeding activity of soil detritivores under warmer and drier conditions. Nature Clim. Change 8 (1), 75–78. https://doi.org/10.1038/s41558-017-0032-6.
van Gestel, C.A.M., Kruidenier, M., Berg, M.P., 2003. Suitability of wheat straw decomposition, cotton strip degradation and bait-lamina feeding tests to determine soil invertebrate activity. Biol. Fertil. Soils 37 (2), 115–123. https://doi.org/10.1007/s00374-002-0575-0.
van Gestel, C.A.M., van der Waarde, J.J., Derksen, J.G.M.(., van der Hoek, E.E., Veul, M. F.X.W., Bouwens, S., Rusch, B., Kronenburg, R., Stokman, G.N.M., 2001. The use of acute and chronic bioassays to determine the ecological risk and bioremediation efficiency of oil-polluted soils. Environ. Toxicol. Chem. 20 (7), 1438–1449. https://doi.org/10.1002/etc.5620200705.
Von Törne, E., 1990. Assessing feeding activities of soil-living animals. 1 Bait-lamina-tests. Pedobiologia 34, 89–101.
Vorobeichik, E.L., Bergman, I.E., 2020. Bait-lamina test in the assessment of polluted soils: choice of exposure duration. Russ. J. Ecol. 51 (5), 430–439. https://doi.org/10.1134/S1067413620050136.
Vorobeichik, E.L., Ermakov, A.I., Grebennikov, M.E., 2019. Initial stages of recovery of soil macrofauna communities after reduction of emissions from a copper smelter. Russ. J. Ecol. 50 (2), 146–160. https://doi.org/10.1134/S1067413619020115.
Vorobeichik, E.L., Ermakov, A.I., Grebennikov, M.E., Golovanova, E.V., Kuznetsov, A.V., Pishchulin, P.G., 2007. Response of the soil macrofauna to emissions from the Middle Ural Copper Smelter. Biological reclamation and monitoring of disturbed lands, Yekaterinburg, pp. 128–148.
Vorobeichik, E.L., Ermakov, A.I., Nesterkova, D.V., Grebennikov, M.E., 2020. Coarse woody debris as microhabitats of soil macrofauna in polluted areas. Biol. Bull. Russ. Acad. Sci. 47 (1), 87–96. https://doi.org/10.1134/S1062359020010173.
Vorobeichik, E.L., Kaigorodova, S.Y., 2017. Long-term dynamics of heavy metals in the upper horizons of soils in the region of a copper smelter impacts during the period of reduced emission. Eurasian Soil Sci. 50 (8), 977–990. https://doi.org/10.1134/S1064229317080130.
Vorobeichik, E.L., Pishchulin, P.G., 2011. Effect of trees on the decomposition rate of cellulose in soils under industrial pollution. Eurasian Soil Sci. 44 (5), 547–560. https://doi.org/10.1134/S1064229311050140.
Vorobeichik, E.L., Trubina, M.R., Khantemirova, E.V., Bergman, I.E., 2014. Long-term dynamic of forest vegetation after reduction of copper smelter emissions. Russ. J. Ecol. 45 (6), 498–507. https://doi.org/10.1134/S1067413614060150.
Welsch, J., Songling, C., Buckley, H.L., Lehto, N.J., Jones, E.E., Case, B.S., 2019. How many samples? Soil variability affects confidence in the use of common agroecosystem soil indicators. Ecol. Ind. 102, 401–409. https://doi.org/10.1016/j.ecolind.2019.02.065.