

Journal of Freshwater Ecology

ISSN: 0270-5060 (Print) 2156-6941 (Online) Journal homepage: www.tandfonline.com/journals/tjfe20

Relevance of pesticides for macroinvertebrate communities in agricultural streams – a re-evaluation

Udo Hommen, Silke Classen , Richard Ottermanns, Oliver Körner, Carola Schriever , Stefan Kimmel & Sophia Müllner

To cite this article: Udo Hommen, Silke Classen , Richard Ottermanns, Oliver Körner, Carola Schriever , Stefan Kimmel & Sophia Müllner (2026) Relevance of pesticides for macroinvertebrate communities in agricultural streams – a re-evaluation, Journal of Freshwater Ecology, 41:1, 2744324, DOI: 10.1080/02705060.2026.2744324

To link to this article: https://doi.org/10.1080/02705060.2026.2744324

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 14 Oct 2026.

Submit your article to this journal

Article views: 326

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tjfe20

JOURNAL OF FRESHWATER ECOLOGY
2026, VOL. 41, NO. 1, 2744324
https://doi.org/10.1080/02705060.2026.2744324

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Relevance of pesticides for macroinvertebrate communities in agricultural streams – a re-evaluation

Udo Hommen[a], Silke Classen[b,1], Richard Ottermanns[c], Oliver Körner[d], Carola Schriever[e], Stefan Kimmel[f] and Sophia Müllner[g]

[a]Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Schmallenberg, Germany; [b]Research Institute gaiac, Aachen, Germany; [c]Institute for Environmental Research, RWTH Aachen University, Aachen, Germany; [d]ADAMA Deutschland GmbH, Cologne, Germany; [e]BASF Agricultural Solutions Deutschland GmbH, Limburgerhof, Germany; [f]Corteva AgriScience, Munich, Germany; [g]Industrieverband Agrar e. V. (IVA), Frankfurt, Germany

ABSTRACT

Attributing ecological impairment in agricultural streams to individual stressors is a common challenge in freshwater monitoring because pesticide exposure often co-occurs with habitat degradation, nutrient enrichment, oxygen depletion, and other environmental pressures. The German 'small water bodies monitoring project' ('Kleingewässermonitoring', KgM) provides an opportunity to examine this challenge using a large-scale field dataset. The KgM was conducted in 2018–2019 to assess the chemical and biological condition of small agricultural watercourses. At 101 stream sites, water samples were analysed for pesticide active substances, pharmaceuticals, and other urban contaminants. Aquatic macroinvertebrates were sampled to determine community composition. The authors of the KgM concluded that pesticide exposure was the dominant stressor affecting vulnerable insects and the macroinvertebrate community based primarily on analyses using the $SPEAR_{Pesticides}$ (SPEcies At Risk) index. Because conclusions about stressor relevance may depend on biological indicators, reference conditions and analytical methods, we re-evaluated the robustness of the original interpretation using the publicly available KgM dataset. Based on our re-evaluation, we argue that the choice of index can strongly influence the interpretation of stressor relevance. In addition, the $SPEAR_{Pesticides}$ results were affected by the inclusion of forest-dominated sites, which differed from agricultural sites not only in pesticide exposure but also in other chemical pollutants and abiotic factors. Analysis of full macroinvertebrate communities using multivariate ordination methods, rather than indices designed to identify sensitivity to single stressors, indicated associations with several environmental factors affecting ecological quality. Our findings demonstrate that inferences about the relative importance of pesticides depend on study design, reference conditions and analytical methods. More broadly, the study highlights challenges that are relevant to freshwater monitoring projects internationally whenever ecological responses are evaluated along correlated multi-stressor gradients and supports the use of integrated multi-stressor assessment frameworks.

ARTICLE HISTORY
Received 30 April 2026
Accepted 30 September 2026

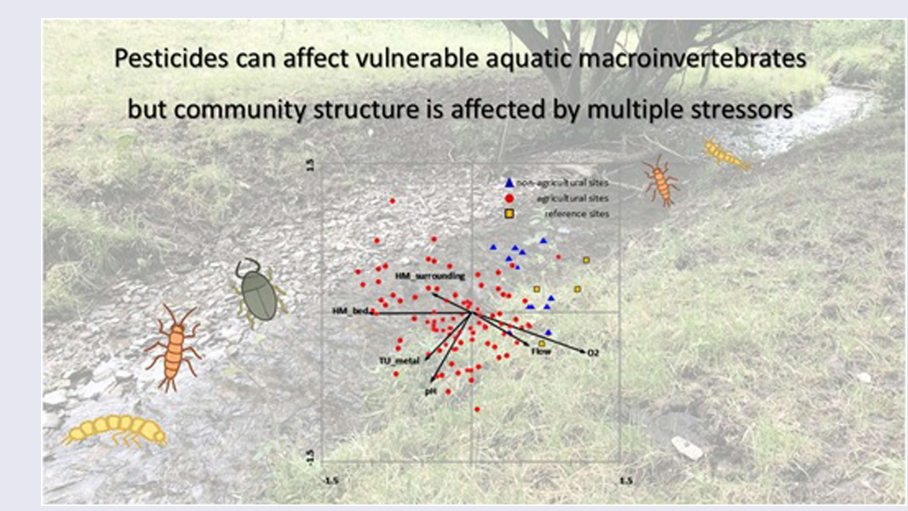
KEYWORDS
Plant protection products; ecological quality; ordination; SPEAR; EPT; reference sites

CONTACT Udo Hommen udo.hommen@ime.fraunhofer.de Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Schmallenberg, Auf dem Aberg 1, 57392, Germany
[1] Present address: Büro für Gewässerökologie Aachen, Aachen, Germany.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02705060.2026.2744324.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.



1. Introduction

Agricultural streams worldwide are exposed to multiple interacting stressors, including pesticide contamination, nutrient enrichment, altered hydrology, habitat degradation, oxygen depletion, and increased water temperatures. Because these pressures frequently co-occur along land-use gradients, attributing ecological impairment to a single stressor remains one of the major challenges in freshwater ecology and environmental monitoring. Identifying the relative importance of different stressors is essential for designing effective management and restoration measures, yet conclusions often depend on the biological indicators and analytical approaches employed. Trait-based indices have been developed to improve the detection of stressor-specific ecological responses. One such index, the $SPEAR_{Pesticides}$ (SPEcies At Risk) index, was developed specifically to respond to pesticide pressure (Liess and von der Ohe 2005; Liess et al. 2022) and is now widely used to assess the impact of pesticides on aquatic environments (Lorenz et al. 2022; von Gönner et al. 2024; BLE 2026).

The 'small water bodies monitoring project' (Kleingewässermonitoring, KgM) was a nationwide German monitoring programme conducted in 2018–2019 by the Helmholtz Centre for Environmental Research (UFZ) on behalf of the German Environment Agency (Umweltbundesamt, UBA) to characterise pesticide exposure, environmental conditions and macroinvertebrate communities in small streams across Germany. The programme included 101 sites distributed across the North German Plain, Central German Uplands and Alpine Foreland regions, with measurements of pesticides, metals, nutrients, hydromorphological characteristics and aquatic macroinvertebrates. The $SPEAR_{Pesticides}$ index was used to assess the ecological status of small streams and ditches potentially affected by non-point sources of pesticides and their metabolites. Liess et al. (2021a) concluded that there was a '*prime relevance of agricultural pesticide pressure for the composition of invertebrate communities*' and called for an additional assessment factor of 20 to be included in environmental risk assessments of plant protection products. These conclusions were primarily based on ordinary and multiple linear regressions of various ecological indices against environmental factors including pesticide stress, oxygen level and hydromorphological characteristics.

Ecological indices summarize the complexity of community composition into a single numerical value such as taxon richness or the Shannon diversity index. Some reflect species traits, such as vulnerability to organic pollution in the Saprobic index. Thus, they rely on assumptions about species traits suitable for grouping. In contrast to trait-based or stressor-specific indices, ordination methods do not require taxa to be assigned *a priori* weights according to their presumed vulnerability to a particular stressor. Instead, occurrence or abundance data for multiple taxa are analysed jointly. Ordination reduces the dimensionality of multivariate community datasets and represents the principal gradients in community composition using a limited number of axes (Legendre and Legendre 2012; Clarke et al. 2014; Braak and Šmilauer 2018).

Key policy highlights

- Stressor-specific macroinvertebrate-based indices support diagnosis but not stressor ranking.
- Pesticide effects should be evaluated within a multi-stressor framework.
- Effective stream restoration requires improvements in habitat quality, oxygen conditions, and other co-occurring stressors as well as pesticide management.

The present study re-evaluates the publicly available KgM dataset (Liess et al. 2021b) and the official project report (Liess et al., 2022) to assess the robustness of conclusions regarding the relative importance of pesticide exposure for macroinvertebrate communities. Specifically, we (1) analyse the influence of including forest sites with expected low pesticide loading in the analyses of pesticide effects in edge-of-field water bodies, (2) discuss the transparency of the $SPEAR_{Pesticides}$ calculation and its underlying assumptions and (3) apply ordination methods to the full macroinvertebrate abundance dataset to identify environmental variables associated with community composition without imposing pesticide-specific trait classifications. Thus, this study addresses questions that are complementary to those examined in the recent reanalysis by Nguyen et al. (2026), focusing on site selection, SPEAR calculation and full-community ordination. By examining how stressor attribution is influenced by indicator choice, reference-site selection, and analytical methodology, this study aims to evaluate the robustness of the conclusion that pesticides are the dominant stressor in the KgM dataset. More broadly, it seeks to provide insights into how methodological decisions influence stressor attribution in freshwater monitoring programs operating under complex multi-stressor conditions.

2. Materials and methods

This re-evaluation is based on three complementary sources from the KgM programme: Liess et al. (2021a) presents the principal scientific analysis and interpretation; Liess et al. (2021b) provide the publicly accessible site-level dataset; and Liess et al. (2022) describe the broader monitoring programme and its methods in the official report. Although Moore and Rathjens (2025) criticised the documentation and reproducibility of the analyses reported by Liess et al. (2021a), we used the complete dataset available from Liess et al. (2021b) together with the methodological information provided in Liess et al. (2021a, 2022), for this re-evaluation. Our analysis focuses on the calculation and interpretation of the $SPEAR_{Pesticides}$ index, the consequences of including forest-dominated sites as low-pesticide reference sites, and the use of multivariate analysis as an alternative approach for investigating community structure and the relative importance of environmental stressors.

2.1. KgM sampling sites and measurements

The KgM comprised 101 stream sites distributed across Germany (Figure S1), eleven of which were monitored in both 2018 and 2019, yielding 112 site-year combinations. The present re-evaluation distinguishes three nested site groups used throughout the manuscript. The 'agricultural sites' are the 86 sites whose catchments contained more than 20% agricultural land cover; their median catchment composition was 82% agriculture and 16% forest. Only four of these catchments contained grassland, accounting for 2–8% of their area. The remaining 15 sites, termed 'forest sites', had predominantly forested catchments, with median land cover of 99% forest and 1% agriculture. Liess et al. (2021a) selected five of these forest sites as low pesticide 'reference sites' for normalising the SPEAR Pesticides index; thus, reference sites are a subset of the forest sites rather than a third independent site category. According to the German stream-type map (Briem 2003), 45 sites belonged to ecoregions in the Central German Uplands, 40 to the North German Plain and eleven to the Alps and Alpine Foreland, while five belonged to stream types independent of an ecoregion. Consequently, the title description 'lowland streams' used by Liess et al. (2021a) does not represent the full dataset, and inclusion of multiple ecoregions probably increased environmental variability.

At each KgM site, chemical exposure was characterised using regular grab samples collected at three-week intervals from April to July and event-driven samples triggered by rising water levels to capture runoff-related concentration peaks. Samples were analysed for pesticides, metals, urban pollutants such as pharmaceuticals and wastewater markers, and nutrients. Water temperature, pH, dissolved oxygen, flow

velocity and hydromorphological quality were also recorded. Aquatic macroinvertebrates were sampled using standardised protocols, and taxon abundances were determined for ecological assessment and statistical analysis (Liess et al. 2021a, 2022). This combined design allowed chemical, habitat and biological conditions to be compared across the same monitoring sites.

2.2. Characterisation of chemical stress

Measured concentrations of pesticides were converted into Toxic Units (TUs) for aquatic invertebrates by dividing each concentration by the lowest acute LC_{50} value reported for either *Daphnia magna* or *Chironomus* sp. (Table SI 2–4 in Liess et al. 2021a). Pesticide stress in the analyses reproduced here was represented by the maximum single-substance TU recorded at each site. Although Liess et al. (2021a) reported only the log-transformed maximum TU, termed TU_{max}, Nguyen et al. (2026) found that summed toxic units across pesticide mixtures were generally more strongly associated with $SPEAR_{Pesticides}$ and that combining event and grab sampling results provided a more complete characterisation of exposure than either sampling approach alone. The present analysis nevertheless retains TU_{max}, hereafter denoted TU_{ppp}, because its purpose is to re-evaluate the metric underlying the original dominant-stressor claim. Results based on TU_{ppp} should therefore not be interpreted as a general comparison of alternative mixture-exposure metrics.

For metals, concentration addition was assumed and the substance-specific TUs were summed; the maximum summed value per site and year was reported as TU_{metal}. TUs were not reported for urban pollutants. We therefore attempted to calculate TU_{urban} using the approach described for pesticides. However, verification of the LC_{50} values reported by Liess et al. (2021a) showed that at least one value was a NOEC and differed by a factor of 1,000 from the reported value (see Section 2 of the Appendix). We therefore did not calculate or evaluate TU_{urban} and used only the reported log-transformed TU_{ppp} and TU_{metal} values.

2.3. Calculation of $SPEAR_{Pesticides}$

The SPEARPesticides index was designed to indicate effects of pesticides on aquatic macroinvertebrates: an '*index that relates to the toxic pressure of pesticides on invertebrates*' (Liess et al., 2021b). It is based on the categorization of species into pesticide-vulnerable 'species at risk' (SPEAR) and 'species not at risk' (SPEnotAR) based on four criteria: 'sensitivity' (relative physiological sensitivity), 'generation' (long generation time), 'exposed' (potential exposure) and 'refuge' (no migration ability from refuge areas). A taxon is classified as a 'species at risk' if all four criteria are met. The 'Indicate' software (Liess 2026) facilitates the classification of taxa and the calculation of the SPEARPesticides index. Both the classification and calculation procedures have changed multiple times since the software was released and since the publication of Liess et al. (2021a), see the change log (Liess 2026).

For the KgM data, Liess et al. (2021a) used the following equation:

$$SPEAR_{pesticides} = \frac{\sum_{i=1}^{n} \log(4x_i + 1) \times y}{\sum_{i=1}^{n} \log(4x_i + 1)}$$

with n = total number of taxa, x_i = abundance of taxon i, y = 1 for SPEAR taxa and y = 0 for SPEnotAR taxa. The factor 4 was introduced by Knillmann et al. (2018) and is cited in the $SPEAR_{Pesticides}$ Version 2018.05 in 'Indicate' (Liess 2026). No clear justification for this adjustment is provided.

In the KgM, the resulting ratio of SPEAR taxa abundance to total abundance was normalised by the median $SPEAR_{Pesticides}$ value (0.44) derived from five reference sites in the dataset, with the aim to establish quality classes equivalent to the ecological status classes of the Water Framework Directive (WFD, DIRECTIVE 2000/60/EC). Liess et al. (2021a) defined the following three criteria for reference sites: low pesticide toxicity ($\log(TU_{max}) < -3$), only minimally altered hydromorphological structure (main index hydromorphology value > 2.7), and a high proportion of EPT taxa (Ephemeroptera, Plecoptera and Trichoptera), exceeding the median of the 15 sites with predominantly forested catchments (% EPT > 41). Although Liess et al. (2021a) did not explicitly name the reference sites, applying their criteria

identifies five forest sites (S105, S107, S108, S111, and S115) with widely varying catchment areas (0.9–22 km^2, see Liess et al. (2021a).

We recalculated $SPEAR_{Pesticides}$ based on the macroinvertebrate abundance data provided by Liess et al. (2021b) using the 'Indicate' tool ($SPEAR_{Pesticides}$ Version 2023.03) to evaluate the impact of software and classification updates since the publication by Liess et al. (2021a) and to examine the relevance of the various underlying assumptions of the $SPEAR_{Pesticides}$ index calculation.

2.4. Other indices

Liess et al. (2021a) reported Saprobic Index values and % EPT (% of Ephemeroptera, Plecoptera and Trichoptera) without specifying the software or calculation methods used. Moore and Rathjens (2025) regard this as a lack of transparency, since they were unable to reproduce the % EPT values given in Liess et al. (2021a). In the present study, the macroinvertebrate data were reanalysed using the official online tool Perlodes (Hering 2026), and we found that the % EPT reported by Liess et al. (2021a) correspond to '[%] EPT (HK)' in Perlodes, i.e. the relative abundance of EPT taxa based on abundance classes. We were also able to reproduce the Saprobic Index values using Perlodes. We used Perlodes to calculate the Fauna Index (Lorenz et al. 2004) and the KLIWA Index (Sundermann et al. 2022), both of which were mentioned but not reported by Liess et al. (2021a) (see Section 3 in the Appendix).

2.5. Statistical analysis

Univariate analyses were performed to check for significant differences in variables between agricultural sites ($n = 86$), and forest sites ($n = 15$) or reference sites ($n = 5$, a subset of the forest sites). Normality was tested using the Shapiro-Wilk Test and the Brown-Forsythe Test to assess variance homogeneity. For normally distributed and variance-homogeneous data, the Student's t-test was used to test for differences between groups, e.g. differences in water temperature or nutrient levels between forest and agricultural sites. If the normality test failed, the Mann-Whitney Rank Sum Test was applied to test for differences between groups. For both parametric and non-parametric testing, p-values < 0.05 were considered to indicate a significant difference. Correlations between two variables (e.g. environmental variables or indices describing the macroinvertebrate community) were calculated using Spearman Rank Order Correlation (if the normality test failed, as was mostly the case) or Pearson Product Moment Correlation (for normally distributed data). Ordinary linear regression was used to explain $SPEAR_{Pesticides}$ based on TU_{PPP} both including and excluding the forest sites.

Environmental variables and macroinvertebrate abundances were evaluated by multivariate statistics (Legendre and Legendre 2012; Braak and Šmilauer 2018). Similarities among sites in environmental variables were examined by Principal Component Analysis (PCA) based on inter-sample distances after centring and standardising the variables. Macroinvertebrate community structure was examined by Correspondence Analysis (CA) after (log(abundance + 1) transformation without down-weighting of rare species and secondary projection of environmental variables (indirect gradient analysis). Effects of environmental variables on the community structure were assessed by Canonical Correspondence Analysis (CCA) also after (log(abundance + 1) transformation. Variables were selected by automatic forward selection (maximum of 8 variables ranked by marginal effects). Statistical tests were conducted under the full model with alpha = 0.05.

Statistical comparisons of means between groups, correlation and regression analyses were performed with SigmaPlot 14 (Grafiti LLC). Ordination analyses were conducted with CANOCO 4.56 (Biometrics Plant Research International).

3. Results and discussion

3.1. Influence of index selection on stressor interpretation

Liess et al. (2021a) compared several macroinvertebrate indices with environmental variables using ordinary and multiple linear regressions. Their interpretation of pesticide dominance relied primarily on the relationship between pesticide toxic pressure, expressed as the maximum pesticide Toxic Unit per

site, and SPEAR Pesticides, an index designed to indicate pesticide-related vulnerability. Other indices represented different ecological dimensions, including the relative abundance of Ephemeroptera, Plecoptera and Trichoptera taxa and the Saprobic Index. The following comparison revisits these regression results to determine whether the apparent importance of a stressor depends on the ecological index selected.

In the present study, $SPEAR_{Pesticides}$, % EPT and the Saprobic Index were each significantly correlated with land cover variables, hydromorphology, bed habitat index, oxygen concentration, ammonium, TU_{ppp} and TU_{metal} (Table S4). Correlations exceeding 0.5 occurred only between TU_{ppp} and $SPEAR_{Pesticides}$ and between % EPT and the main hydromorphology index. These findings are consistent with the multiple-regression results reported by Liess et al. (2021a). Their model explained 61% of the variance in $SPEAR_{Pesticides}$, with TU_{ppp} alone accounting for 32%. For % EPT, structural variables including hydromorphology, bed habitat structure and flow velocity explained 38% of the variance, whereas the full model explained 60%. Oxygen deficiency explained 20% of the variance in the Saprobic Index, while the full model explained 45% (Figure 1). Thus, the inferred relevance of environmental stressors depends strongly on the selected ecological index: each trait-based index was predominantly associated with the stressor that it was designed to represent. We agree with Liess et al. (2021a) that such indices can indicate specific stressors more effectively than integrated measures such as total abundance, taxon richness, or diversity. However, the SPEAR and % EPT models had almost identical explanatory power while identifying different stressors as the most relevant. The conclusion that pesticides dominate macroinvertebrate community structure is therefore dependent on the choice of index and should be interpreted in that context.

3.2. Environmental differences between agricultural and forest sites

Forest sites were included in the KgM as sites with low pesticide pressure. In an ideal case, such reference sites should be as similar as possible to the other sites except for the stressor under investigation. In this section, we summarize whether environmental factors other than pesticide pressure differ between agricultural sites, forest sites, and reference sites (a subset of the forest sites) (see Table 1).

As intended, the forest sites were significantly less exposed to pesticide entries, but also to metals that at agricultural sites reached concentrations close to the LC_{50} of *D. magna* (used as a benchmark for TU, see supplementary materials in Liess et al. (2021a)). TUs were not provided for urban pollutants, but Liess

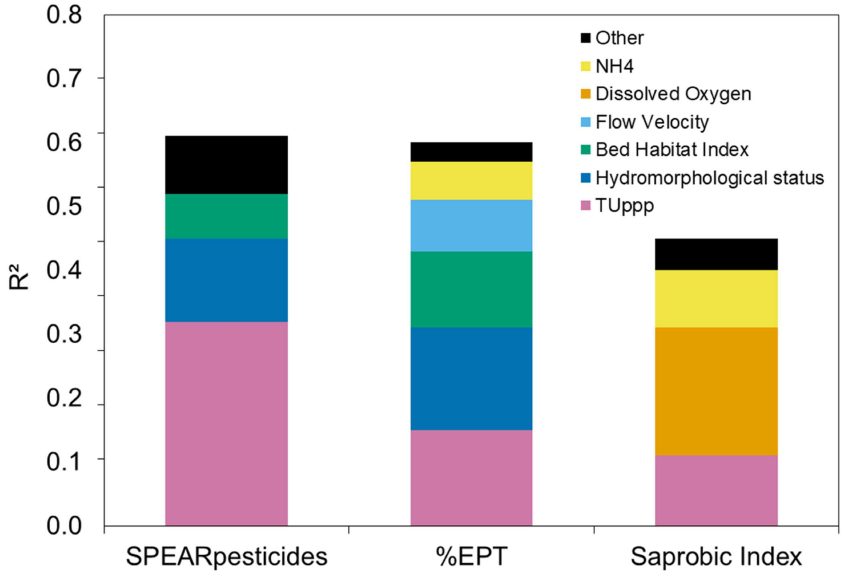


Figure 1. Coefficients of determination (R^2) calculated by Liess et al. (2021a) for multiple regression models of three macroinvertebrate indices.

Table 1. Environmental variables and macroinvertebrate indices from the agricultural, forest and reference sites.

	Label in figures	Agricultural sites $n = 86$	Forest sites $n = 15$	Reference sites $n = 5$
Environmental variables				
Maximum Toxic Unit for pesticides [log]	**TU_{ppp}**	**−2.21 ± 0.84**	−3.40 ± 1.18*** (a)	−4.13 ± 0.95*** (a)
Maximum Toxic Unit for metals [log]	**TU_{metal}**	**−0.32 ± 0.56**	−1.24 ± 0.85*** (b)	−1.17 ± 0.92** (b)
Catchment size [km^2]		13.6 ± 6.8	7.8 ± 8.8*** (b)	8.0 ± 8.6 (n.s.) (b)
Agricultural land cover (catchment) [%]	Cat_agri	75 ± 22	3 ± 6*** (b)	4 ± 8*** (b)
Forest land cover (catchment) [%]	Cat_fore	22 ± 4	97 ± 6*** (b)	96 ± 8*** (b)
Urban land cover (catchment) [%]	Cat_urba	3 ± 4	0 ± 0*** (b)	0 ± 0*** (b)
Main hydromorphological status (subindices for bed, bank, and surrounding land)	**HM_main** (HM_bed/HM_bank/ HM_surrounding)	**4.2 ± 0.9**	2.6 ± 0.8*** (a)	2.0 ± 0.5*** (b)
Bed Habitat Index	BHI	50 ± 27	63 ± 27 (n.s.) (a)	78 ± 18* (b)
Temperature [°C]	Temp	18.5 ± 1.8	19.0 ± 2.3 (n.s.) (a)	19.9 ± 2.0 (n.s.) (a)
Flow velocity [m/s]	Flow	0.26 ± 0.13	0.28 ± 0.12 (n.s.) (a)	0.32 ± 0.13 (n.s.) (a)
Dissolved Oxygen [mg/L]	**O2**	**7.2 ± 1.4**	8.5 ± 0.6*** (b)	8.8 ± 0.7** (b)
pH [-]	pH	7.7 ± 0.4	7.4 ± 0.6** (b)	7.5 ± 0.6 (n.s.) (b)
Nitrite NO_2 [mg/L]	**NO2**	**0.45 ± 0.46**	0.15 ± 0.22*** (b)	0.16 ± 0.21* (b)
Nitrate NO_3 [mg/L]	**NO3**	**30.4 ± 59.3**	4.7 ± 4.4*** (b)	3.8 ± 4.3** (b)
Ammonium NH_4 [mg/L]	**NH4**	**0.67 ± 0.78**	0.08 ± 0.08*** (b)	0.06 ± 0.08** (b)
Total nitrogen N_{total} [mg/L]	**N_total**	**8.51 ± 5.90**	2.49 ± 2.93*** (b)	1.93 ± 0.76*** (b)
Phosphate PO_4 [mg/L]	PO4	1.62 ± 2.49	1.18 ± 2.19 (n.s.) (b)	2.01 ± 3.62 (n.s.) (b)
Total phosphorus P_{total} [mg/L]	**P_total**	**0.52 ± 0.65**	0.17 ± 0.37*** (b)	0.12 ± 0.18** (b)
Macroinvertebrate indices				
$SPEAR_{Pesticides}$		**0.45 ± 0.17**	0.74 ± 0.26*** (a)	0.96 ± 0.16*** (a)
% EPT		**24.5 ± 11.8**	41.7 ± 15.1*** (a)	53.3 ± 14.3*** (a)
Saprobic Index		**2.1 ± 0.2**	1.7 ± 0.1*** (b)	1.8 ± 0.1*** (b)
KLIWA Index		**14.9 ± 1.7**	12.3 ± 1.5*** (a)	12.0 ± 1.5*** (a)
Fauna Index		**−0.14 ± 0.76**	0.26 ± 0.90* (a)	0.63 ± 0.42* (a)

***: $p < 0.001$, **: $p < 0.01$ and *: $p < 0.05$, n.s. = not significant. a = Student's t-Test, b = Mann Whitney Rank Sum Test.

Nitrate was excluded from multivariate analysis due to an extreme value in one agricultural site (site 86 with 553 mg/L, five times higher than the next highest value) which would have dominated the ordination.

Note that the number of sites (n) does not necessarily correspond to the number of data points, because some sites and variables were sampled in two years.

Shown are the mean, standard deviations, and the statistical significance of differences between agricultural and forest sites as well as between agricultural and reference sites (a subgroup of forested sites). Bold values indicate significantly poorer conditions at agricultural sites ($p < 0.05$).

Table 2. Ranking of environmental variables based on Canonical Correspondence Analysis (CCA) for agricultural sites only (Figure 7).

Variable	Canonical eigenvalue	% of explained variance covered	% of total variance covered	p-value
O_2	0.1531	17.88	2.16	0.002
HM_bed	0.1212	14.16	1.71	0.002
pH	0.1138	13.29	1.60	0.002
flow	0.1072	12.52	1.51	0.002
Cu	0.0973	11.37	1.37	0.154
HM_land	0.0896	10.47	1.26	0.074
TU_{metal}	0.0889	10.38	1.25	0.090
TU_{ppp}	0.0850	9.93	1.20	0.122
Sum	0.8561	100	12.06	

The table presents the relevance of the variables selected through automatic forward selection (sum of all eigenvalues: 7.086, sum of all canonical eigenvalues: 0.856). A p-value above 0.05 indicates no significant further improvement in the model. For definitions abbreviations, see Table 1.

et al. (2021a) report that they were much lower than those of pesticides. Forest streams exhibited significantly higher structural quality on average (Table 2), which is known to positively influence macroinvertebrate communities (Effert-Fanta et al. 2019; Sargac et al. 2021). Indeed, the German Fauna Index (FI), reflecting the effects of morphological degradation on the macroinvertebrate community (Lorenz et al. 2004), indicates better hydromorphological conditions in the forest sites than in the agricultural sites, even if the difference is not significant.

Forest sites also showed significantly higher oxygen levels and lower nutrient loads than agricultural sites. As reported by Liess et al. (2022), low oxygen levels, often resulting from elevated nutrient inputs, can be critical stressor for macroinvertebrates (Birk et al. 2020), and adverse effects on sensitive taxa cannot be excluded. The photographic documentation in Liess et al. (2022) illustrates that agricultural streams often lacked riparian shading vegetation, suggesting that they may experience higher water temperatures. Nevertheless, the mean of the reported temperature values per site, represented by the 75th quantile of each measurement series, was slightly lower than in the forest sites, although the difference was not

significant. Moore and Rathjens (2025) found that gaps in the reported temperature and oxygen data had been filled using an undocumented procedure and classified many of these values as failing their quality criteria. In contrast to the reported temperature measurements, the macroinvertebrate-based KLIWA index (Sundermann et al. 2022) suggests greater temperature stress in agricultural sites than in forest sites (Table 1). The data in Table 1 demonstrate that besides pesticide pressure (TU_{ppp}), the agricultural stream sites differed from the forest sites with respect to several other factors (lower oxygen, poor structural quality, more nutrients, probably higher temperatures as demonstrated by the KLIWA Index), all of which impact macroinvertebrate communities. These findings align with Liess et al. (2022) who mention that agricultural sites differ from forest sites in terms of non-chemical stressors, e.g. that 70% of the agricultural sites, oxygen levels had low oxygen levels below 8 mg/L and were therefore considered unacceptably low. The observed differences indicate that forest sites cannot be regarded as differing from agricultural sites solely with respect to pesticide exposure. Instead, they represent a distinct set of environmental conditions associated with multiple stressors and habitat characteristics.

3.3. Environmental similarity of reference sites

Because the five reference sites used for $SPEAR_{Pesticides}$ normalisation were selected from the forest-site group, their comparability with agricultural sites is critical for the interpretation of pesticide effects. We therefore further evaluated similarities and differences among agricultural, forest, and reference sites using Principal Component Analysis (PCA). PCA confirmed the results of the univariate analyses. To prevent land use and pesticide exposure from predetermining the ordination, land-use variables and TU_{ppp} were treated as supplementary variables. This approach allowed us to assess whether the reference sites differed from agricultural sites only in terms of pesticide exposure or also with respect to broader environmental conditions. The agricultural sites and the forest sites form two clusters with little overlap (Figure 2A) along a gradient of environmental variables that is strongly correlated with chemical pollution (TU_{ppp}). The cluster of the reference sites does not overlap with the agricultural sites at all, indicating that environmental conditions including pesticide concentrations are quite different. In a second analysis, only the

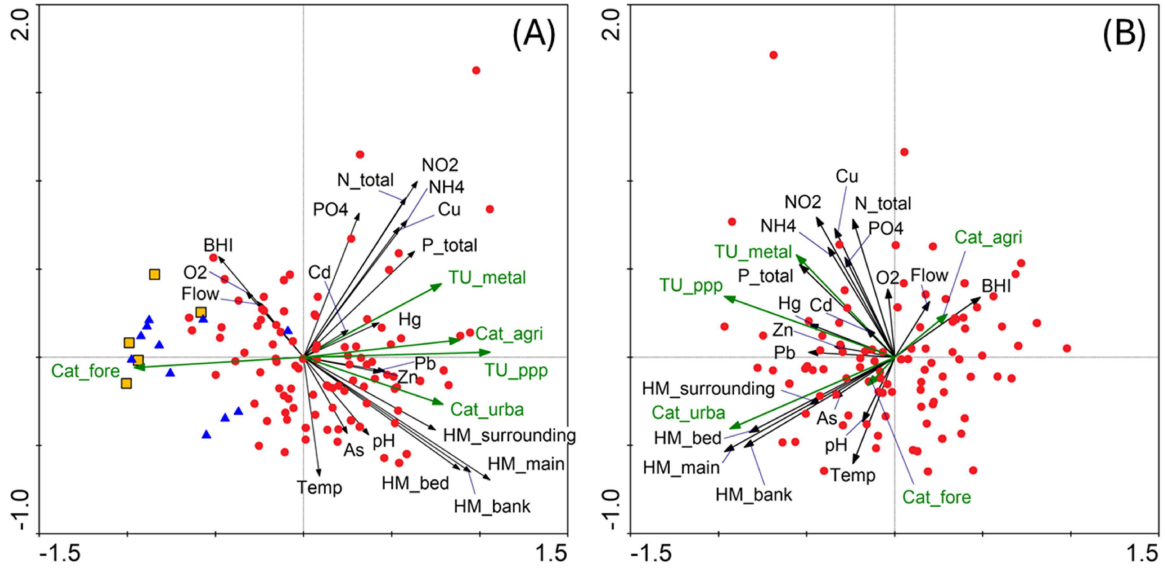


Figure 2. Principal component analysis (PCA) of environmental variables for the initial set of 112 site-year combinations. Panel (A) shows all sites ($n = 111$ site-year combinations, one forest site omitted because land use information was missing). The first and the second axes explained 21.5% and 11.6% of variance, respectively. Panel (B) shows only agricultural site-year combinations ($n = 95$). The first and second axes explained 16.5% and 13.6% of the variance, respectively. Black arrows: environmental variables used for ordination. Green arrows: land use variables (Cat_x) and TU_{ppp} as supplementary variables. Red circles: agricultural sites. Blue triangles: forest sites excluding reference sites; yellow squares: reference sites. For definition of labels, see Table 1.

agricultural sites were ordinated (Figure 2B). The results are very similar to those of the first ordination, but the eigenvalues are considerably smaller (0.165 and 0.136 for 1st and 2nd axis compared to 0.215 and 0.117, respectively, in the PCA including forest sites) because the variability within the agricultural sites alone is much smaller than in combination with the non-agricultural sites.

In summary, the variance of chemical pollution (TU_{metal} and TU_{ppp} as supplementary variable) contributes strongly to the dissimilarity between the investigated sites. However, agricultural and forest sites and especially the reference sites do not differ only with respect to pesticide pressure as intended but to several other stressors. Thus, the suitability of forest reference sites for assessing pesticide stress is questionable. On the other hand, it should be noted that it is probably impossible to find reference sites which only differ with respect to the stress of pesticides while other stressors show less variability or at least are not correlated with pesticide stress. Therefore, a multi-stressor situation cannot be avoided in monitoring studies, but this limitation should be carefully considered when interpreting ecological assessments and drawing conclusions about the relevance of individual stressors.

3.4. Influence of reference-site selection on SPEAR normalisation

Liess et al. (2021a) normalised the $SPEAR_{Pesticides}$ values against reference sites to define ecological quality classes analogous to those in the Water Framework Directive (WFD; Directive 2000/60/EC). The normalisation approach was first implemented in the 'Indicate' tool with $SPEAR_{Pesticides}$ Version 2018.05, using the method of Knillmann et al. (2018). A reference value of 0.31 was applied, corresponding to the mean $SPEAR_{Pesticides}$ value at sites with $\log(TU_{max}) \leq -4.5$. Consequently, reference conditions were defined solely by toxic pressure. In Version 2019.10, the normalisation value was updated to 0.27 without explanation or revised criteria. In the KgM project, the reference value for streams in agricultural landscapes was set to 0.44 based on the median of five forest sites selected for low pesticide pressure, good hydromorphological structure and the presence of EPT taxa (see Section 2.3). As shown in the previous section, the selected reference sites differed substantially from the agricultural sites with respect to several environmental characteristics, including land use, hydromorphological condition, nutrient status and oxygen concentrations. These differences imply that the resulting quality classes are contingent on the choice of reference sites and their selection criteria. Although the taxon classifications were updated in $SPEAR_{Pesticides}$ Versions 2023.03 and 2023.08, the reference value of 0.44 was retained. For methodological consistency, the reference value should be recalculated when underlying trait data are updated.

This assignment of values to certain quality classes based on normalisation differs from procedures used for other indices, such as % EPT, where reference sites are predefined for each stream type (European Commission 2003). The specific index values of these reference sites serve as a benchmark for determining the quality class thresholds. Given the availability of macroinvertebrate data for these reference streams (Haase et al. 2004), stream-type-specific quality class definitions could also be developed for the $SPEAR_{Pesticides}$ index. This would ensure independence from individual monitoring campaigns, account for ecological variation among stream types, and align $SPEAR_{Pesticides}$ with established WFD assessment methods.

3.5. Effects of forest-site inclusion on stressor interpretation

Ordinary linear regression relating normalized $SPEAR_{Pesticides}$ to TU_{ppp} explained 43% of the variance in $SPEAR_{Pesticides}$ at the 101 agricultural and forest sites (Liess et al. 2021a) (Figure 3A). Liess et al. (2021a) reported that replacing exceptionally high TUs at approximately 20% of sites with the next-lower value increased the explained variance from 32% to 43%. They conclude that this indicates the higher relevance of several smaller pesticide pulses compared to a single high pulse. Nguyen et al. (2026) reproduced this pattern with summed toxic units and interpreted it as evidence that very short external concentration peaks may not translate proportionally into internal exposure because uptake and elimination kinetics differ among substances. This toxicokinetic explanation is plausible, but the replacement rule remains empirically defined rather than mechanistically parameterised. The resulting regression should therefore be presented as a sensitivity analysis of exposure aggregation choices rather than as evidence that high-concentration pulses are ecologically unimportant.

More generally, selecting or replacing exposure values according to their agreement with $SPEAR_{Pesticides}$ introduces a risk of circularity, compresses the upper exposure gradient, and may artificially increase the apparent linearity and explanatory power of the regression. A non-proportional decline in $SPEAR_{Pesticides}$ at extreme pesticide pulses does not necessarily establish ecological irrelevance, but may instead reflect index saturation, temporal mismatch between exposure and biological sampling, delayed recovery, or effects on taxa not captured by the index. In addition, high pulses could affect a larger part of the macroinvertebrate community, since less sensitive species would become affected, too. It seems questionable to ignore such short-term events of very high exposure just because they reduce the coefficient of determination in the regression.

As shown in Section 3.2, forest sites differ significantly from agricultural sites not only in pesticide concentrations but also in other stressors such as nutrient levels, oxygen concentration, and structural quality. When forest sites are excluded from the dataset used by Liess et al. (2021a), the explanatory power of TU_{ppp} was reduced to explaining only 22% of the variance (Figure 3B). In both regressions, the $SPEAR_{Pesticides}$ values in sites with TUs between −1 and −3, i.e. around the regulatory acceptable concentration (RAC) in a Tier 1 assessment ($TU_{ppp} = -2$, corresponding to EC50/100), show large variability. Consequently, the regressions are mainly driven by sites with very low and very high TU values (Figure 3). Without the five sites with $\log(TU_{ppp}) < -4$ or > -1 (i.e. concentrations at least 10 times above the acute Tier 1 RAC), only 10% of the variance in $SPEAR_{Pesticides}$ is explained by TU. In addition, several agricultural sites with $\log(TU_{ppp})$ up to −0.1 are classified as of good quality according to $SPEAR_{Pesticides}$ despite relatively high pesticide concentrations. This suggests that at agricultural sites, other environmental stressors may play a larger role in shaping macroinvertebrate communities than pesticide toxicity alone.

3.6. SPEAR classification and reproducibility

A taxon is classified as a SPEAR, if all four traits (sensitivity, generation time, exposure and refugium), indicate vulnerability to pesticides. In the following, we discuss uncertainties regarding these traits and their consequences for the classification of taxa as SPEAR.

In 'Indicate' ($SPEAR_{Pesticides}$ Version 2023.08), sensitivity reflects a taxon's relative physiological sensitivity to organic pollutants, including pesticides, standardized to the acute sensitivity (LC_{50}) of *Daphnia magna* (von der Ohe and Liess 2004). Species with sensitivity higher than the median were regarded as sensitive for the SPEAR classification (Liess and von der Ohe 2005). Brinke et al. (2017) mention an update of the sensitivity trait using acute toxicity values for insecticides only but provide no details on how individual taxon's sensitivity was revised, e.g. number and type of datasets used. Also, no

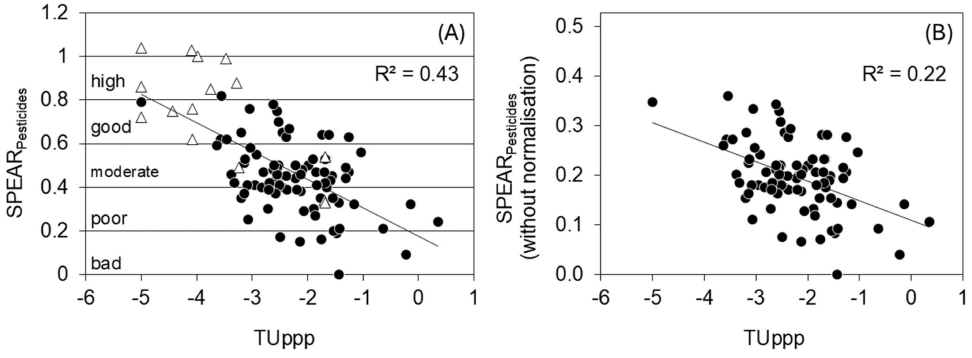


Figure 3. Relationship between $SPEAR_{Pesticides}$ and TU_{ppp}. Panel (A): all sites ($n = 112$ site-year combinations, forest sites marked as open triangles) as in Liess et al. (2021a). Panel (B): agricultural sites ($n = 95$ site-year combinations). In panel (A), normalised $SPEAR_{Pesticides}$ values are shown, based on the median $SPEAR_{Pesticides}$ value of the five forest reference sites; horizontal lines indicate quality class boundaries. Panel (B) displays non-normalized $SPEAR_{Pesticides}$ values without quality classes, since the forest reference sites are not included in the regression analysis; the relative vertical position of the non-standardized values remains unchanged compared to panel (A). R^2 denotes the coefficient of determination.

such update is documented in the $SPEAR_{Pesticides}$ version history, suggesting continued reliance on the sensitivities reported in von der Ohe and Liess (2004), which were derived from toxicity data for organic pollutants. If sensitivities were revised, the threshold for sensitivity classification should also have been adjusted, as it was originally set to the median sensitivity (Liess and von der Ohe 2005).

The extent to which the sensitivity classification in $SPEAR_{Pesticides}$ driven specifically by insecticide toxicity cannot be verified. Most sensitivity values reported by Liess et al. (2021a) are identical to those published by von der Ohe and Liess (2004) for organic compounds (see Table S2 in the appendix). Only a few taxa, such as Simuliidae and Megaloptera, do not have corresponding sensitivity values in the earlier publication. In addition, sensitivity values for some crustaceans, particularly *Asellus*, have been modified since von der Ohe and Liess (2004), although the basis for these revisions is unclear. Because all taxa classified as SPEAR are insects, and many belong to the EPT group, the classification appears to be strongly influenced by assumptions regarding the comparatively low sensitivity of non-insect taxa such as Molluscs or Oligochaetes. However, these groups are often sensitive to fungicides or indirectly affected by herbicides. Consequently, the extent to which $SPEAR_{Pesticides}$ sensitivity classifications reflect responses specifically to insecticides remains.

Taxa with a generation time of 0.5 years or longer are considered vulnerable to pesticides. According to Liess et al. (2021a), most taxa meet this criterion. Exceptions include certain Diptera (Ceratopogonidae, Chironomidae, Simuliidae, Stratiomyidae and Syrphidae) and Oligochaeta from the family Lumbricidae, which are listed with a generation time shorter than 0.5 years. We did not verify the accuracy of these assignments.

Exposure is defined as the likelihood of a taxon being present during the main pesticide application period. According to Liess et al. (2021b), see '5_1_Invertebrate_Abundances.txt'), most taxa at the 101 sites are classified as exposed. For taxa classified as not exposed, this seems to be due to an assumed emergence before the beginning of the pesticide application period. Although we did not conduct a detailed review of emergence timing, initial literature comparisons suggest that this assumption may not always be accurate. Additionally, the time between emergence, egg-laying and larval development should be considered, as it may affect actual exposure in water bodies. Furthermore, the classification of gammarids as not exposed is inconsistent. As exclusively aquatic organisms, they are continuously present in water bodies and thus likely to be exposed during pesticide application periods. In fact, their exclusion from the exposure category is based on migration ability (Liess and von der Ohe 2005), a trait more relevant to recovery potential. Conflating recovery with exposure introduces conceptual confusion.

The refuge trait, introduced into $SPEAR_{Pesticides}$ by Knillmann et al. (2018) replaced the former migration criterion. Refuge taxa are considered as not vulnerable to pesticides due to their dispersal-based resilience, enabling recolonization via drift or adult flight. In their study, Knillmann and co-workers identified eleven refuge taxa based on their nearly exclusive presence at sites with nearby refuge areas not exposed to toxic pressure. It should be noted that the nationwide categorization of refuge taxa is based only on empirical data for a limited region and not on a general analysis of species traits. Our study did not assess individual taxon classification in detail. Yet some classification decisions are difficult to follow. Notably, *Silo* sp. shows limited drift and low adult dispersal, challenging its refuge status (Tachet et al. 2010; Knillmann et al. 2018). Conversely, taxa with strong dispersal abilities (e.g. *Baetis* sp., Heptageniidae, Hydropsychidae, Odontoceridae) are not classified as refuge taxa. While integrating recolonization is conceptually sound, classification decisions as described above introduce inconsistencies. Further refinement and alignment with ecological data are needed.

Changes in taxon classification and software versions also affect the reproducibility of $SPEAR_{Pesticides}$ results. Using $SPEAR_{Pesticides}$ Version 2023.03 in the 'Indicate' software, the resulting $SPEAR_{Pesticides}$ values differ from those presented in Liess et al. (2021a), with several sites now assigned to different ecological status classes, likely due to changes in the SPEAR classification of individual taxa (Figure 4). These differences are likely attributable to revisions of individual taxon classifications and demonstrate that even relatively small changes in the underlying trait database can influence $SPEAR_{Pesticides}$ values and the resulting ecological assessments. Although the overall relationship between sites remains similar, the observed shifts highlight the importance of transparent documentation of classification updates and software versions when $SPEAR_{Pesticides}$ results are interpreted or compared across studies.

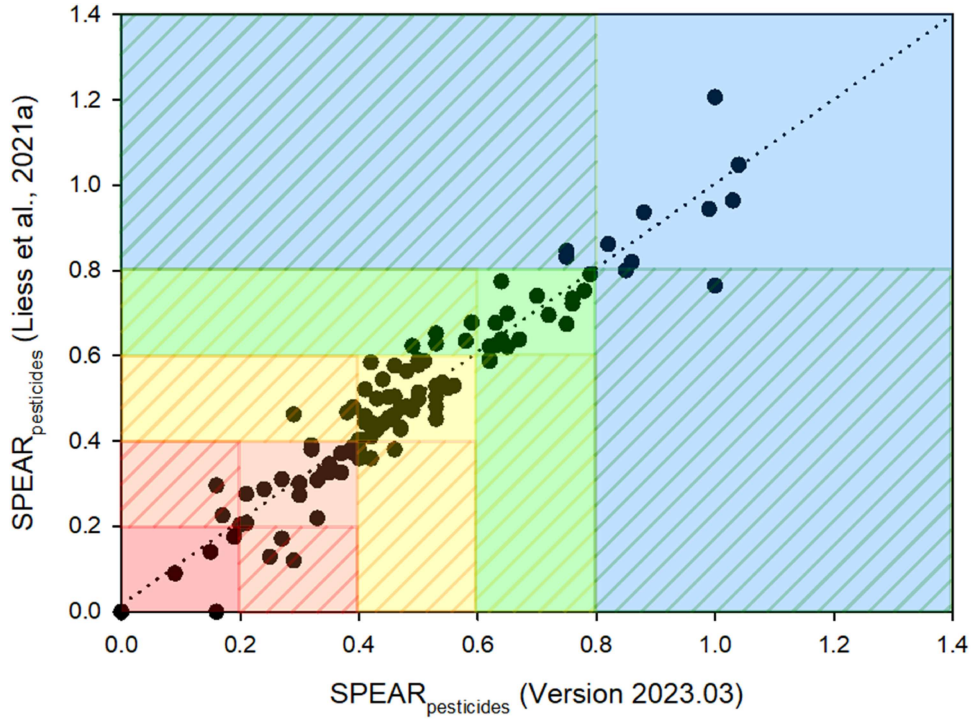


Figure 4. Comparison of $SPEAR_{Pesticides}$ values reported by Liess et al. (2021a) with values recalculated in this study using $SPEAR_{Pesticides}$ Version 2023.03 ($n = 112$ site-year combinations). The diagonal dashed line represents the 1:1 relationship. Background colours indicate ecological status classes: red = bad, orange = poor, yellow = moderate, green = good, blue = high. Data points located within the hatched areas represent sites assigned to different ecological status classes in the two datasets.

3.7. Ordination of sites based on macroinvertebrate data

SPEAR relies on *a priori* trait-based sensitivity classifications that incorporate assumptions about species vulnerability and recovery potential, which may not fully reflect context-dependent ecological responses (van den Brink and ter Braak 2012). Unlike multivariate approaches that analyse complete taxonomic datasets directly, SPEAR is constrained by these predefined classifications and may be less adaptable when stressor combinations differ from those under which the index was developed or calibrated. Such limitations may become more pronounces in large-scale monitoring datasets characterized by heterogeneous habitat conditions, variable taxonomic resolution, and incomplete trait information.

To evaluate community patterns without relying on predefined trait classifications, we therefore applied Correspondence Analysis (CA) to the macroinvertebrate dataset. In this re-evaluation, macroinvertebrate data were ordinated using CA, without considering environmental variables. The resulting biplot (Figure 5) visualises patterns in macroinvertebrate community composition based on species abundance data. Environmental variables were then correlated with the ordination scores and added as vectors.

The eigenvalues and explained variance of the first (x) and second (y) ordination axes were low (eigenvalues = 0.331 and 0.260; 4.5% and 3.5%, respectively), indicating that the ordination result reflects only a small proportion of variance within the abundance data. Most of the remaining variance is random noise that cannot be explained by the unimodal model of the CA. Consequently, the biplot summarizes only a relatively small part of the total variance (8%). Nevertheless, the first axis reflects a consistent separation between most agricultural sites (positive scores) and forest and reference sites (negative scores), suggesting a land-use gradient. However, some agricultural sites cluster with forest sites, indicating overlapping community structures. The second axis shows weaker separation, with forest sites generally located in the upper part of the biplot (positive scores).

Environmental variables align with the ordination patterns, consistent with earlier results (Figure 2). Ammonium concentration (NH_4) correlates strongly with the first ordination axis, indicating higher

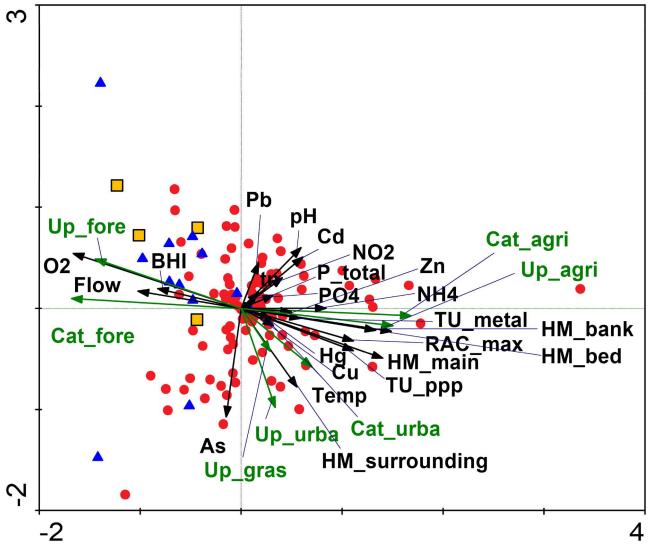


Figure 5. Correspondence analysis (CA) of log-transformed macroinvertebrate data ($n = 110$ site-year combinations; one agricultural site excluded due to the extreme abundance of a single taxon and one forest site excluded due to missing land use information). The first and second axes explained 4.5% and 3.5% of the variance, respectively. Red circles: agricultural sites; blue triangles: forest sites; yellow squares: reference (forest) sites. Environmental variables (black arrows) were not included in the ordination but correlated post-hoc with the ordination scores. Supplementary land use variables are shown as green arrows. For definitions of labels, see Table 1.

concentrations at agricultural sites. Similarly, chemical concentrations (TU_{ppp}, TU_{metal}) and structural quality indices (HM_bed, HM_bank, HM_main) point towards agricultural sites, indicating higher concentrations and lower structural quality. In contrast, oxygen concentration (O_2), bed habitat quality (BHI), and flow velocity point towards reference and forest sites, indicating higher oxygen levels and better bed habitat quality there. TU_{ppp}, hydromorphological status (HM_main) and abiotic parameters like Ammonium (NH_4) strongly correlate with macroinvertebrate communities associated with agricultural land use (Cat_agri, Up_agri). Arsenic (As) correlates strongly with the second ordination axis but is unlikely to be ecologically relevant, as the measured concentrations were below ecotoxicologically relevant levels for macroinvertebrates.

In conclusion, the CA reveals a weak overall structure in the macroinvertebrate dataset. A clear land-use gradient separates agricultural sites from forest/reference sites. Environmental variables like ammonium and oxygen levels align with this gradient, but the low explained variance suggests limited interpretive power.

3.8. Environmental drivers of community composition

A Canonical Correspondence Analysis (CCA) with automatic forward selection of variables was conducted to rank the environmental variables according to their relevance to explain the community structure. As outlined in Section 3.7, CCA analyses the complete macroinvertebrate assemblage directly and allows multiple environmental stressors to be evaluated simultaneously without relying on predefined sensitivity classifications.

The ordination separated communities from agricultural and forest streams, although some overlap remained (Figure 6). Together, all environmental variables explained 18.3% of the total variance. The six variables selected as significant at $p < 0.05$ were oxygen content (O_2), hydromorphology of the streambed (HM_bed), pH, metal stress (TU_{metal}), hydromorphology of the surrounding (HM_land), and flow velocity (Flow); together these variables explained 11.1%. The eigenvalues of the first and the second axis were 0.211 and 0.126, respectively, and the displayed axes accounted for 4.5% of the total variance. Pesticide contamination (TU_{ppp}) was not automatically selected as a significant explanatory variable. Temperature (Temp) was also selected during the procedure but did not explain a significant proportion of the variance.

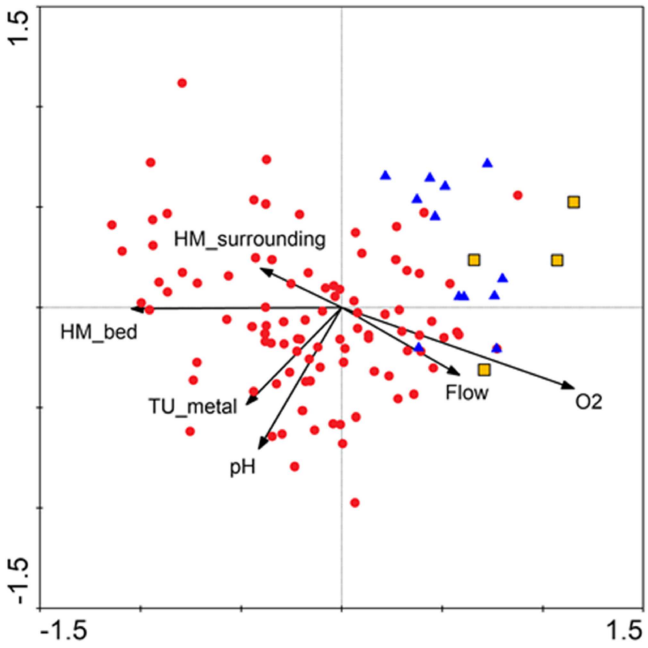


Figure 6. Canonical Correspondence Analysis (CCA) of all sites, with automatic forward selection of variables ($n = 110$ site-year combinations; one agricultural site excluded due to the extreme abundance of a single taxon and one forest site excluded due to missing land use information). The first and second axes explained 2.8% and 1.7% of the variance, respectively. Red circles: agricultural sites; blue triangles: forest sites; yellow squares: reference (forest) sites. Arrows indicate environmental variables. For definitions of labels, see Table 1.

Because the separation between agricultural and forest sites may primarily reflect fundamental differences in land use and associated environmental conditions rather than pesticide exposure alone, the CCA was repeated for agricultural sites only (Figure 7). This analysis removes the influence of forest reference sites and evaluates variation within agricultural sites. Sites in good SPEAR quality classes 1 and 2 (blue circles and green squares, respectively) and sites in bad SPEAR classes 4 and 5 (red triangles and red circles, respectively) overlap considerably (Figure 7). Only oxygen content (O_2), bed structure (HM_bed), pH, and flow velocity (flow) explain a significant proportion of the variation in macroinvertebrate communities at agricultural sites ($p < 0.05$, Table 2).

Table 2 presents the forward-selection ranking for agricultural sites corresponding to Figure 7. The sum of all eigenvalues was 7.086 and the sum of the canonical eigenvalues was 0.856. Oxygen, streambed hydromorphology, pH and flow velocity significantly improved the model at $p < 0.05$, whereas TU_{ppp} did not provide a significant additional contribution at the step at which it entered the selection procedure. The marginal effect of TU_{ppp} nevertheless accounted for 1.20% of total community variance and 9.93% of the variance explained by the selected variables. Accordingly, these results do not support identifying pesticides as the dominant independent predictor of full community composition under this CCA specification, but they also do not exclude pesticide toxicity as a contributor to the correlated multi-stressor gradient. This interpretation is consistent with Nguyen et al. (2026), who found pesticide toxicity to be robustly influential across univariate ecological metrics while concluding that no single stressor dominated across responses.

Ordination analysis considers all taxa without prior weighting and therefore evaluates whole-community composition rather than predefined ecological traits. Consequently, if species A is present in one sample, whereas only a similar, species B, is present in another, this difference may have a large impact on the ordination depending on the taxa's abundance. In contrast, an index may treat the two species equivalently, if they belong to the same trait group. Thus, trait-based indices such as $SPEAR_{Pesticides}$ reduce variance because similar index values can result from communities with substantially different taxonomic compositions. Ordination therefore typically explains a lower proportion of variance in the full multivariate dataset than ordinary or multiple regression applied to a single index.

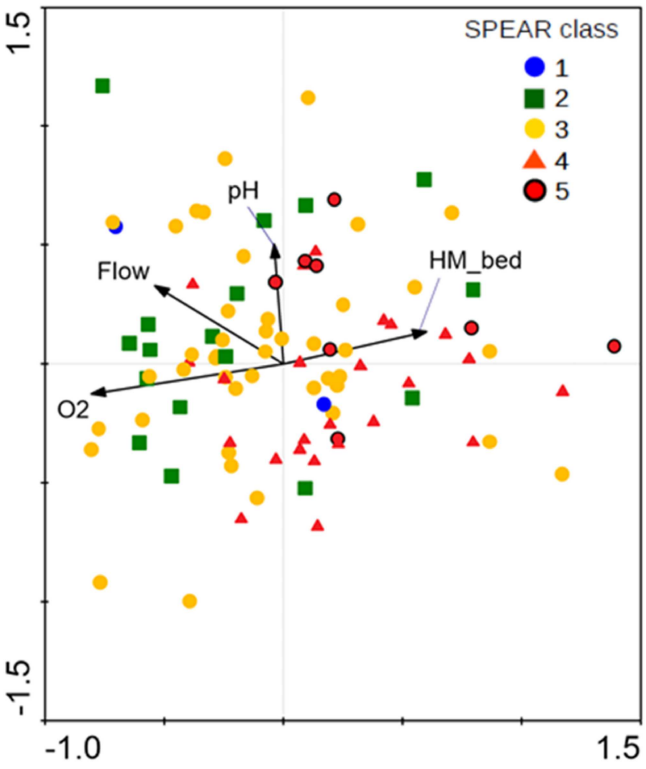


Figure 7. Canonical Correspondence Analysis (CCA) of agricultural sites only, with automatic forward selection of variables ($n = 94$ site-year combinations; one site excluded due to the extreme abundance of a single taxon). The first and second axes explained 2.7% and 2.0% of the variance, respectively. Colours indicate SPEAR classes. Arrows indicate environmental variables. For definitions of labels, see Table 1.

Ordination allows an unbiased ranking of the relative importance of stressors for community structure, whereas the ranking of stressors for an index is necessarily constrained by how the specific index is constructed. However, ordination is not assumption-free: results depend on taxonomic resolution, transformation, scaling, distance model, treatment of rare taxa, predictor collinearity and the variable-selection procedure. Consequently, trait-based indices and ordination consequently answer different questions. Indices test responses along predefined ecological dimensions, whereas ordination examines compositional structure across all recorded taxa. Agreement or divergence between these approaches should be interpreted as complementary evidence rather than as proof that one method provides an unbiased ranking of stressors.

In summary, the CCA of all sites indicates that oxygen levels, hydromorphology, pH, and metal stress, had the strongest detectable independent associations with overall macroinvertebrate community composition under the present model specification. TU_{ppp} was not retained as statistically significant after automatic forward selection and therefore did not provide a significant additional contribution to explaining community composition at the monitored sites. This result complements, rather than overturns, evidence from Nguyen et al. (2026) that pesticide toxicity is a stable predictor of several univariate ecological metrics. Together, these results indicate that the apparent relevance of pesticides depends on the biological response variable, exposure metric and statistical framework. Multiple co-occurring stressors contributed to variation among agricultural streams.

3.9. Stressor interactions and index behaviour

Trait-based indices such as $SPEAR_{Pesticides}$ were developed to improve the diagnostic power of macroinvertebrate community data for identifying pesticide-related effects, particularly under pulsed exposure scenarios typical of agricultural catchments (Liess and von der Ohe 2005; Liess et al. 2008; Becker et al.

2020). Independent studies (Malherbe et al. 2018; Bighiu et al. 2020; Poyntz-Wright et al. 2023) have demonstrated that $SPEAR_{Pesticides}$ can respond sensitively to gradients of agricultural intensity and pesticide use, and in some regional or long-term datasets it has shown clearer responses than more general biotic indices. Consequently, the $SPEAR_{Pesticides}$ has been proposed as a potentially useful tool within weight-of-evidence frameworks linking ecological responses to chemical pressure (Liess et al. 2021a).

Like other indices summarizing the macroinvertebrate community structure, the $SPEAR_{Pesticides}$ index is most responsive to the stressor for which it was optimized. Consequently, pesticide-related stress is expected to be most prominent in taxa vulnerable to general chemical inputs when analysed using SPEAR. However, the dominant stressor at the community level depends on the type of index used. $SPEAR_{Pesticides}$ is a useful tool to indicate potential pesticide effects, but it should not be used in isolation to conclude on the overall relevance of different stressors.

Unlike targeted indices, ordination analysis does not weight species with respect to specific stressor *a priori* (van den Brink and ter Braak 2012). An ordination of the KgM dataset revealed no dominance of pesticide effects but rather that multiple stressors affect the community structure.

Site selection also plays a critical role in interpreting pesticide stress. Using sites with forest-dominated catchments as low-pesticide reference sites, alongside agricultural catchments, strongly influenced the role of pesticide stress in SPEAR analysis for the KgM dataset. Since forest sites differ from agricultural sites not only in pesticide contamination but also in nutrient levels, oxygen levels, other chemical stressors and structural variables, their suitability as reference sites for agricultural streams is questionable.

SPEAR metrics may respond not only to pesticide exposure, but also to other agricultural stressors, including nutrient enrichment, habitat alteration, hydromorphological degradation, and their interactions, as shown by, for examples, van den Brink and ter Braak (2012) and Bray et al. (2021). These findings are particularly relevant for monitoring programmes such as the KgM, where multiple stressors co-vary strongly along gradients of agricultural land use. Consistent with these observations, our KgM results indicate that $SPEAR_{Pesticides}$ is strongly associated with broader environmental gradients characteristic of agricultural intensification, rather than exhibiting a clear, independent relationship with pesticide exposure metrics alone. Similar patterns have been reported in other field studies, where $SPEAR_{Pesticides}$ showed weak or absent correlations with chemically derived toxic units, while simultaneously correlating with general ecological indices such as ASPT (Bighiu et al. 2020). Such findings suggest that SPEAR values may, under certain conditions, reflect integrated community responses to multiple stressors rather than pesticide pressure per se.

4. Conclusion

$SPEAR_{Pesticides}$ is suitable for identifying community patterns consistent with pesticide exposure and can complement indices focussing on other stressors. Nguyen et al. (2026) showed that pesticide toxicity remains a stable and influential predictor across several modelling approaches, but also that nutrients, agricultural land use and hydro-morphological degradation contribute substantially and that inferred rankings depend on toxicity aggregation and ecological metric. Our analysis reaches a convergent conclusion from a complementary perspective: full-community ordination does not identify TU_{ppp} as the dominant independent predictor once correlated environmental variables are considered. Together, these findings suggest $SPEAR_{Pesticides}$ should be used within a multimetric and multimodel assessment framework rather than as a standalone basis for ranking stressors or deriving additional hazard-assessment factors. This broader interpretation also has important implications for regulatory applications that rely on $SPEAR_{Pesticides}$-TU_{ppp} relationships.

Regulatory proposals that rely primarily on $SPEAR_{Pesticides}$-TU relationships should be interpreted with caution such as the proposal by Liess et al. (2021a) to derive an additional assessment factor of 20 for acute toxicity data. The $SPEAR_{Pesticides}$-TU is strongly affected by sites with generally higher ecological quality and sites with pesticide concentrations far exceeding regulatory acceptable concentrations (RACs). The proposal assumes that lowering RAC values would directly increase the abundance of SPEAR taxa, while neglecting the substantial influence of other stressors and of non-label-compliant pesticide entries. Although pesticide exposure above critical levels can adversely affect vulnerable taxa, community composition is also shaped by other factors such as nutrient levels, dissolved oxygen, and habitat structure.

To improve the ecological quality of edge-of-field water bodies a more holistic approach is needed by also considering structural quality of the water bodies and their surroundings. Our analysis shows that a more nuanced evaluation of environmental factors and paths of entry is necessary to identify the main drivers of poor water body status. Proposing an additional safety factor to be applied in the prospective regulatory risk assessment based on an analysis with the SPEAR index will probably not increase the ecological quality of the streams. The reason for this is that a safety factor is not appropriate to mitigate the relevant portion of plant protection products which enters the streams via wastewater- and farmyard-effluents (point sources).

While $SPEAR_{Pesticides}$ provides valuable information on community vulnerability patterns, its responses in the KgM dataset appear to be strongly influenced by co-occurring stressors and methodological constraints inherent to large-scale field monitoring. $SPEAR_{Pesticides}$ appears most informative when applied as one component of an integrated assessment framework that explicitly accounts for multiple stressors, temporal exposure dynamics, and regional ecological context. Within such a framework, $SPEAR_{Pesticides}$ can contribute to hypothesis generation and comparative assessment but should not be interpreted as a uniquely pesticide-specific diagnostic metric in isolation.

Acknowledgements

Andreas Solga, Herbert Ressler, Robert Spatz as well as other members of the IVA-Project Group '*PG Kleingewässermonitoring*' made valuable comments on a draft of this manuscript.

Author contributions

CRediT: **Udo Hommen:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Visualization, Writing – original draft; **Silke Classen:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft; **Richard Ottermanns:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft; **Oliver Körner:** Writing – review & editing; **Carola Schriever:** Writing – review & editing; **Stefan Kimmel:** Writing – review & editing; **Sophia Müllner:** Project administration, Writing – review & editing.

Disclosure statement

Oliver Körner, Stefan Kimmel and Carola Schriever are employed by companies producing plant protection products. Sophia Muellner works for The Plant Care Industries Association (Industrieverband Agrar e. V., IVA).

Funding

The work of Udo Hommen, Silke Classen and Richard Ottermanns within this project was financed by The Plant Care Industries Association (Industrieverband Agrar e. V. (IVA)), Germany.

ORCID

Udo Hommen https://orcid.org/0000-0002-3301-5968
Richard Ottermanns https://orcid.org/0000-0002-2168-9455
Oliver Körner https://orcid.org/0000-0001-9751-1066
Sophia Müllner https://orcid.org/0000-0001-7263-9389

Data availability statement

The data are available as excel files in the supporting information.

Declaration of generative AI use

Microsoft 365 Copilot (accessed March—September 2026) was used to improve the spelling, grammar, clarity, and flow of the manuscript. ChatGPT (OpenAI; accessed January 2026) was used to assist with reference management.

The authors reviewed and verified all AI-assisted output and take full responsibility for the accuracy, originality, and integrity of the final manuscript.

References

[BLE] Bundesanstalt für Landwirtschaft und Ernährung. 2026. SPEAR-Index (Pflanzenschutzmittel). [place unknown]: [publisher unknown]; [updated 2025 Feb 11; accessed 2026 Mar 27]. https://www.nap-pflanzenschutz.de/indikatorenforschung/indikatoren-und-deutscher-pflanzenschutzindex/pix/spear-index-pflanzenschutzmittel

Becker JM, Russo R, Shahid N, Liess M. 2020. Drivers of pesticide resistance in freshwater amphipods. Sci Total Environ. 735:139264. https://doi.org/10.1016/j.scitotenv.2020.139264

Bighiu MA, Höss S, Traunspurger W, Kahlert M, Goedkoop W. 2020. Limited effects of pesticides on stream macroinvertebrates, biofilm nematodes, and algae in intensive agricultural landscapes in Sweden. Water Res. 174:115640. https://doi.org/10.1016/j.watres.2020.115640

Birk S et al. 2020. Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. Nat Ecol Evol. 4:1060–1068. https://doi.org/10.1038/s41559-020-1216-4

Braak Cter, Šmilauer P. 2018. Canoco reference manual and user's guide: software for ordination (version 5.10). Wageningen University & Research: Wageningen.

Bray JP et al. 2021. Can SPEcies at risk of pesticides (SPEAR) indices detect effects of target stressors among multiple interacting stressors? Sci Total Environ. 763:142997. https://doi.org/10.1016/j.scitotenv.2020.142997

Briem E, editor. 2003. Gewässerlandschaften der bundesrepublik deutschland: morphologische merkmale der Fließgewässer und ihrer auen. In: ATV-DVWK Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall e.V.: Hennef. p 176. (ATV-DVWK-Arbeitsbericht). ISBN: 3924063338.ger.

van den Brink PJ, ter Braak CJF. 2012. Response to "Traits and stress: keys to identify community effects of low levels of toxicants in test systems" by liess and beketov (2011). Ecotoxicology. 21:297–299. https://doi.org/10.1007/s10646-011-0825-8

Brinke M et al. 2017. Umsetzung des Nationalen Aktionsplans zur nachhaltigen Anwendung von Pestiziden: Bestandsaufnahme zur Erhebung von Daten zur Belastung von Kleingewässern der Agrarlandschaft. Texte. ISSN. http://www.umweltbundesamt.de/publikationen

Clarke KR, Gorley RN, Somerfield PJ, Warwick RM. 2014. Change in marine communities: an approach to statistical analysis and interpretation (3rd ed.). [publisher unknown]: Plymouth, UK. PRIMER-E. https://learninghub.primer-e.com/books/change-in-marine-communities

von der Ohe PC, Liess M. 2004. Relative sensitivity distribution of aquatic invertebrates to organic and metal compounds. Environ Toxicol Chem. 23:150–156. https://doi.org/10.1897/02-577

Effert-Fanta EL, Fischer RU, Wahl DH. 2019. Effects of riparian forest buffers and agricultural land use on macroinvertebrate and fish community structure. HyBio. 841:45–64. https://doi.org/10.1007/s10750-019-04006-1

European Commission. 2003. Guidance document no. 10: guidance on typology of reference conditions and classification systems for inland surface waters (Common Implementation Strategy for the Water Framework Directive (2000/60/EC)). Official Publications of the European Communities. https://circabc.europa.eu/sd/a/dce34c8d-6e3d-469a-a6f3-b733b829b691/Guidance%20No%2010%20-%20references%20conditions%20inland%20waters%20-%20REFCOND%20(WG%202.3).pdf.

von Gönner J et al. 2024. Citizen science shows that small agricultural streams in Germany are in a poor ecological status. Sci Total Environ. 922:171183. https://doi.org/10.1016/j.scitotenv.2024.171183

Haase P et al. 2004. Validierung der Fließgewässertypologie Deutschlands, Ergänzung des Datenbestandes und Harmonisierung der Bewertungsansätze der verschiedenen Forschungsprojekte zum Makrozoobenthos zur Umsetzung der Europäischen Wasserrahmenrichtlinie (Modul Makrozoobenthos). Abschlussbericht im Auftrag der LAWA (Projekt O 3.02).

Hering D. 2026. Gewaesser-bewertung-berechnung.de. [publisher unknown]: [place unknown]. [accessed 2026 Mar 27]. https://gewaesser-bewertung-berechnung.de/

Knillmann S, Orlinskiy P, Kaske O, Foit K, Liess M. 2018. Indication of pesticide effects and recolonization in streams. Sci Total Environ. 630:1619–1627. https://doi.org/10.1016/j.scitotenv.2018.02.056

Legendre P, Legendre L. 2012. Numerical ecology. In: Developments in environmental modelling. 24. 3rd ed. Elsevier: Amsterdam, Heidelberg, Boston, London, New York, Oxford, Paris, San Diego, San Francisco, Singpur, Sydney, Tokyo. p 990. ISBN: 9780444538697eng. https://www.sciencedirect.com/science/book/9780444538680.

Liess M et al. 2021a. Pesticides are the dominant stressors for vulnerable insects in lowland streams. Water Res;201:117262. 201:117262. https://doi.org/10.1016/j.watres.2021.117262

Liess M et al. 2021b. The lowland stream monitoring dataset (KgM, Kleingewässer-Monitoring) 2018, 2019. https://doi.org/10.1594/PANGAEA.931673

Liess M et al. 2022. Umsetzung des nationalen aktionsplans zur nachhaltigen anwendung von pflanzenschutzmitteln (NAP): pilotstudie zur ermittlung der belastung von Kleingewässern in der agrarlandschaft mit Pflanzenschutzmittel-Rückständen. Umweltbundesamt: Dessau-Roßlau. (Texte; 07/2022). https://www.umweltbundesamt.de/publikationen/umsetzung-des-nationalen-aktionsplans-zur-1

Liess M. 2026. Indicate 2.3.1. [publisher unknown]: [place unknown]. [accessed 2026 Mar 27]. https://www.systemecology.de/indicate/

Liess M, von der Ohe PC. 2005. Analyzing effects of pesticides on invertebrate communities in streams. Environ Toxicol Chem. 24:954–965. https://doi.org/10.1897/03-652.1

Liess M, Schäfer RB, Schriever CA. 2008. The footprint of pesticide stress in communities--species traits reveal community effects of toxicants. Sci Total Environ. 406:484–490. https://doi.org/10.1016/j.scitotenv.2008.05.054

Lorenz A, Hering D, Feld CK, Rolauffs P. 2004. A new method for assessing the impact of hydromorphological degradation on the macroinvertebrate fauna of five German stream types. HyBio. 516:107–127. https://doi.org/10.1023/B:HYDR.0000025261.79761.b3

Lorenz S, Foit K, Dewenter B, Liess M. 2022. Ist der SPEAR-Index zur Abschätzung der pflanzenschutzmittelbelastung übertragbar auf Standgewässer der agrarlandschaft? Dessau-Roßlau: [publisher unknown]. Texte; 86/2022. https://www.umweltbundesamt.de/publikationen/ist-der-spear-index-zur-abschaetzung-der.

Malherbe W, van Vuren JHJ, Wepener V. 2018. The application of a macroinvertebrate indicator in afrotropical regions for pesticide pollution. J Toxicol. 2018:2581930. https://doi.org/10.1155/2018/2581930

Moore DRJ, Rathjens H. 2025. Are pesticides the dominant stressors in German lowland streams. Integr Environ Assess Manag. 21:739–744. https://doi.org/10.1093/inteam/vjaf038

Nguyen HH, Schreiner VC, Schäfer RB. 2026. Data analysis choices influence the relative importance of multiple stressors on macroinvertebrates in agricultural streams. Environ Sci Technol. 60:25753–25762. https://doi.org/10.1021/acs.est.6c06517

Poyntz-Wright IP, Harrison XA, Johnson A, Zappala S, Tyler CR. 2023. Pesticide pollution associations with riverine invertebrate communities in england. Sci Total Environ. 903:166519. https://doi.org/10.1016/j.scitotenv.2023.166519

Sargac J et al. 2021. Forested riparian buffers change the taxonomic and functional composition of stream invertebrate communities in agricultural catchments. Water. 13:1028. https://doi.org/10.3390/w13081028

Sundermann A, Müller A, Halle M. 2022. A new index of a water temperature equivalent for summer respiration conditions of benthic invertebrates in rivers as a bio-indicator of global climate change. Limnologica. 95:125980. https://doi.org/10.1016/j.limno.2022.125980

Tachet H, Richoux P, Bournaud M, Usseglio-Polatera P. 2010. Invertébrés d'eau douce: systématique. biologie, écologie. Vol. 15. Paris: CNRS éditions. ISBN: 9782271069450.