Citizen science for assessing pesticide impacts in agricultural streams

Julia von Gönner [a,b,c,*], Diana E. Bowler [b,c,d], Jonas Gröning [e], Anna-Katharina Klauer [f], Matthias Liess [e], Lilian Neuer [g], Aletta Bonn [a,b,c]

[a] Helmholtz Centre for Environmental Research - UFZ, Department Ecosystem Services, Permoserstr. 15, 04318 Leipzig, Germany
[b] Friedrich Schiller University Jena, Institute of Biodiversity, Dornburgerstr.159, 07743 Jena, Germany
[c] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, 04103 Leipzig, Germany
[d] UK Centre for Ecology & Hydrology, Benson Lane, Wallingford OX10 8BB, UK
[e] Helmholtz Centre for Environmental Research - UFZ, Department System-Ecotoxicology, Permoserstr. 15, 04318 Leipzig, Germany
[f] Saxony State Foundation for Nature and the Environment (LaNU), Riesaer Str. 7, 01129 Dresden, Germany
[g] Friends of the Earth Germany e.V. (BUND), Kaiserin-Augusta-Allee 5, 10553 Berlin, Germany

HIGHLIGHTS

- Analysis of citizen science data accuracy for bio-indicator based pesticide monitoring.
- Macroinvertebrate data were used to derive the biological indicator SPEAR$_{pesticides}$.
- Citizen science and professional data for SPEAR and hydromorphology agreed well.
- Citizen science SPEAR values and measured pesticide concentrations correlated well.
- Citizen science stream data is suitable to complement official monitoring programs.

GRAPHICAL ABSTRACT
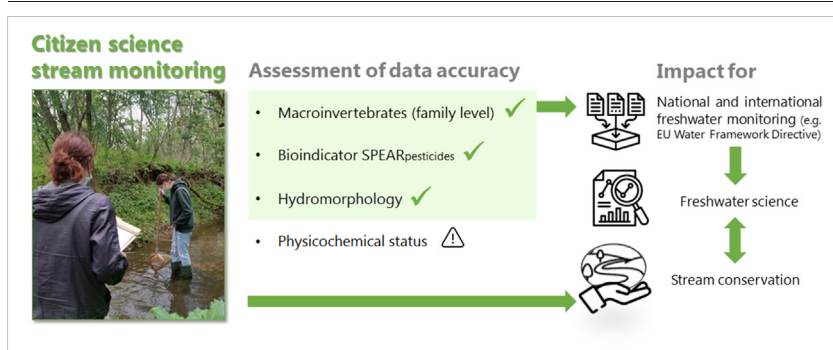
ARTICLE INFO

ABSTRACT

The majority of central European streams are in poor ecological condition. Pesticide inputs from terrestrial habitats present a key threat to sensitive insects in streams. Both standardized stream monitoring data and societal support are needed to conserve and restore freshwater habitats. Citizen science (CS) offers potential to complement international freshwater monitoring while it is often viewed critically due to concerns about data accuracy. Here, we developed a CS program based on the Water Framework Directive that enables citizen scientists to provide data on stream hydromorphology, physicochemical status and benthic macroinvertebrates to apply the trait-based bio-indicator SPEAR$_{pesticides}$ for pesticide exposure. We compared CS monitoring data with professional data across 28 central German stream sites and could show that both CS and professional monitoring identified a similar average proportion of pesticide-sensitive macroinvertebrate taxa per stream site (20 %). CS data were highly correlated to the professional data for both stream hydromorphology and SPEAR$_{pesticides}$ ($r = 0.72$ and 0.76). To assess the extent to which CS macroinvertebrate data can indicate pesticide exposure, we tested the relationship of CS generated SPEAR$_{pesticides}$ values and measured pesticide concentrations at 21 stream sites, and found a fair correlation similar to professional results. We conclude that given appropriate training and support, citizen scientists can generate valid data on the ecological status and pesticide contamination of streams. By complementing official monitoring, data from well-managed CS programs can advance freshwater science and enhance the implementation of freshwater conservation goals.

* Corresponding author at: Department Ecosystem Services, German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, D-04103 Leipzig, Germany.
E-mail address: julia.vongoenner@idiv.de (J. von Gönner).

1. Introduction

National and international freshwater policies have been adopted in recent decades to improve the ecological status of surface waters and reverse their past degradation. The United Nations Sustainable Development Goal (SDG) 6 on Clean Water and Sanitation (UN, 2018) seeks to improve water quality by reducing pollution and use of hazardous chemicals (SDG target 6.3) and to protect water-related ecosystems (SDG target 6.6). In the European Union, the Water Framework Directive (WFD, 2000/60/EC) was established in 2000 and requires the member states to restore or maintain a "good ecological status" in all surface water bodies. About 60 % of European rivers, however, still fail to achieve a good ecological status (EEA, 2018). This percentage is even higher in central European countries such as Germany, where only 7 % of rivers reach a good ecological status (UBA, 2017, 2022). Recent studies show that pesticide inputs induced by run-off from terrestrial, agricultural environments are among the main risk factors for stream ecosystems and pesticide-sensitive aquatic invertebrates (Liess et al., 2021; Wolfram et al., 2021). In addition, nutrient inputs and habitat degradation pose a considerable threat to stream ecological status (Malaj et al., 2014; Vörösmarty et al., 2010).

Despite the Europe-wide WFD surface water monitoring, there is a lack of systematic large-scale monitoring data on the ecological status of small streams, so that it remains difficult to track the impacts of land use and the effectiveness of environmental management and conservation policies. WFD monitoring covers rivers and streams with catchment areas $\geq 10 km^2$ and focuses on larger rivers ($>100 km^2$), while small streams ($<10\ km^2$) are not included (apart from exceptional cases, Wick et al., 2019, Weisner et al., 2022). Two thirds of the entire river network, however, consist of small streams below $10 km^2$ (BfN, 2021; Meyer et al., 2007), which play an important role for the conservation of plant, bird, amphibian and insect diversity. Many insects have aquatic larval stages and depend on good freshwater habitat quality (Dijkstra et al., 2014). Due to their small water volume and often proximity to agricultural areas, small streams can be particularly affected by agricultural pesticide inputs (Halbach et al., 2021; Szöcs et al., 2017). The current WFD monitoring strategy, however, leads to limited knowledge about pesticide contaminations and overall ecological status of small streams (Weisner et al., 2022). Refined pesticide monitoring approaches suitable to directly measure pesticide exposure (e.g. the German monitoring of plant protection products in small streams) are restricted to a limited number of sampling sites (Weisner et al., 2022).

To effectively monitor streams and reduce pesticide and nutrient inputs as well as habitat degradation, active support from civil society actors as well as citizen engagement and compliance is essential (Cardoso et al., 2011; Jackson et al., 2016). In line with this, the WFD aims to involve citizens and stakeholders into water resource and ecosystem management (EPA, 2006). Citizen science (CS) not only has the potential to generate important large-scale data to assess ecological trends (Bowler et al., 2021), but also to foster environmental learning, civic engagement (Turrini et al., 2018) and social license for conservation (Kelly et al., 2019). In the freshwater realm, CS programs offer potential to fill gaps in official monitoring schemes and to enhance research on ecological pressures and management measures (Carvalho et al., 2019; Jackson et al., 2016; Maasri et al., 2022). Simultaneously, CS can raise public awareness and harness expertise from society to implement freshwater conservation measures (Brooks et al., 2019; Huddart et al., 2016).

Macroinvertebrates are widely used as biological indicators of freshwater health (Brooks et al., 2019; Chessman et al., 2007; Moffett and Neale, 2015). Due to their life spans from several months to years and their sensitivity to various environmental factors, macroinvertebrate communities provide an integrated assessment of water and habitat quality over time (Friberg et al., 2011). The biological indicator SPEAR_{pesticides} is a trait-based indicator based on the relative abundance of pesticide-sensitive macroinvertebrate taxa at a stream site (Liess and Von Der Ohe, 2005). Depending on the occurrence of functional traits (physiological sensitivity to pesticides, generation time, life cycle or hatching time, and ability to migrate and recolonize), each taxon is categorized as either "SPEcies At Risk" (SPEAR) or "SPEcies not At Risk" (SPEnotAR). The SPEAR$_{pesticides}$ index mainly reacts to pesticide exposure and is mostly independent of other stressors such as oxygen deficiency or nutrient load (Knillmann et al., 2018; Liess et al., 2008, 2021). Therefore, the indicator is a suitable method for identifying pesticide exposure and establishing dose-effect relationships at large spatial scales. It is used for pesticide indication in the German National Stream monitoring (Liess et al., 2021) and in the German WFD stream assessment (LAWA and UBA, 2022). The indicator has also been shown to provide accurate results with macroinvertebrate data whose taxonomic resolution is limited to family level (Beketov et al., 2009; Liebmann et al., 2022) and is therefore well suited for a participatory CS stream monitoring.

To assess the ecological status of rivers and streams according to WFD standards, European freshwater monitoring examines three components: biological communities, hydromorphology and physicochemical status. Hence, to gain a comprehensive picture about stream ecological status, CS programs should also consider those three monitoring components. Several existing CS water monitoring data sets have been shown to provide useful insights into the physicochemical status of streams and the impact of land use change over wide temporal and spatial scales (Abbott et al., 2018; Albus et al., 2020; Safford and Peters, 2018). Studies from the US (Fore et al., 2001; Nerbonne and Vondracek, 2003), from New Zealand (Moffett and Neale, 2015; Storey and Wright-Stow, 2017) and the UK (Brooks et al., 2019) found that if citizen scientists are provided with appropriate training and robust protocols, they can provide valid macroinvertebrate data. To our knowledge, however, there is no evidence yet on the applicability of CS for quantifying stream pesticide contamination based on sampled macroinvertebrate community composition. Moreover, the applicability of CS for official freshwater monitoring is still often questioned due to concerns about data accuracy (Albus et al., 2019; Quinlivan et al., 2020a; Safford and Peters, 2018). In addition, many CS programs are not yet properly aligned with standard monitoring and reporting processes (Fritz et al., 2019; Stepenuck and Genskow, 2018). As a result, many existing CS freshwater datasets remain unused or cannot be taken into account by political and environmental decision makers (Carlson and Cohen, 2018).

To examine the applicability of CS for assessing the ecological status and pesticide contamination of small streams, we launched the CS stream monitoring program FLOW in Germany (https://flow-projekt.de). The FLOW program provides training, support and field equipment for citizen scientists to generate data on macroinvertebrate community composition and taxa abundance to calculate the bio-indicator SPEAR$_{pesticides}$, stream hydromorphology, and physicochemical status. As a criterion for CS data accuracy (i.e. the "degree to which data are correct overall", Kosmala et al., 2016:552) we applied the concept of fitness for use (Bowser et al., 2020). To be considered accurate enough to meet our research goals, the CS stream data should provide assessments of stream ecological status that are comparable and highly correlated with professional data under various environmental conditions. Consequently, we investigated (i) to what extent data on macroinvertebrate communities, stream hydromorphology and physicochemical status collected by trained citizen scientists compared to professionally gathered data. More specifically, we assessed (ii) how CS generated macroinvertebrate data compared to professional data in terms of identification and counting accuracy as well as recording of SPEAR$_{pesticides}$ functional traits. Since error and variation exist in both CS and professional data, it is also important to assess both CS and professional data against a common, known reference (Kosmala et al., 2016; Specht and Lewandowski, 2018). Therefore, we analyzed (iii) to what extent CS and professionally generated SPEAR$_{pesticides}$ index values aligned with event-driven measurements of pesticide exposure at our stream sample sites.

2. Material and methods

2.1. Study design and site selection

We selected a total of 30 lowland and highland stream sample sites distributed over Central Germany with catchment sizes up to 30 km².

Catchments were characterized by a gradient of agricultural land (mean agricultural land cover 67 % ±30 %) and <5 % of urban areas to focus on agricultural diffuse source pollution (for detailed site characteristics see appendix Table A1.1–2, Fig. A1). Each site was sampled once by a group of 8 to 15 trained citizen scientists between April and early July 2021, the main pesticide application time for most crops (Szöcs et al., 2017). During the same time period, professional sampling was conducted by ecotoxicologists from Helmholtz Center for Environmental Research - UFZ as part of the German monitoring of plant protection products in small streams (Liess et al., 2021). For the comparison of CS and professional data (Sections 3.1–3.3), we excluded two stream sites in which CS and professional sampling conditions were not comparable because streams dried out between the CS and professional monitoring days (sample size $n = 28$). To analyze the relationship between CS and professional SPEAR$_{pesticides}$ index values and measured peak pesticide toxicity (Section 3.4), we excluded all sites in which macroinvertebrate communities were severely affected by low flow velocity (<0.05 m/s) or drought in the period from April to July, so that accurate pesticide bioindication was not possible (sample size $n = 21$, see Liess et al., 2021, Table A1.1).

2.2. Citizen science training

We recruited 30 citizen scientist groups (13 regional groups of Friends of the Earth Germany, 8 senior high school classes, 5 angling clubs, and 4 groups consisting of students and agency employees), with a total of 303 participants aged 15 to 65 years (mean 32.3 ± 12 years).

In this study, the term "citizen scientists" encompasses a variety of individual backgrounds (Eitzel et al., 2017). A large majority of citizen scientists were interested newcomers with little to no prior experience in ecological stream assessment. At the same time, eight local freshwater experts (with in-depth taxonomic or ecological knowledge gained through long-term voluntary engagement) participated in the CS monitoring as group leaders (for details on participants see Table A2.1). In contrast, the term "professionals" is used in this study to refer to experienced ecotoxicologists who acquired expertise in limnology as part of their profession as full-time researchers.

Before the monitoring events, all citizen scientists participated in a half-day training led by the FLOW team in methods of stream assessment and in macroinvertebrate identification to family level. After looking at distinguishing features of the most important macroinvertebrate families, the citizen scientists practiced sorting and identifying voucher specimens with a stereomicroscope. To consolidate learning content from the training, all citizen scientists received a project booklet with field protocols and further learning material (i.e. video tutorials, identification booklet, online quiz on macroinvertebrate identification and the assessment of stream hydromorphology). For details on the adaptation of official monitoring guidances to the CS context, see Table A2.2–3.

2.3. Data collection

At each site, two representative 50 m stream sections were chosen for a. CS and b. professional sampling (with 20 m distance between them to avoid the sampling events influencing each other).

For macroinvertebrate sampling, citizen scientists first recorded stream bed substrates (Meier et al., 2006) to ensure standardized multi-habitat sampling according to WFD. For each stream section, substrate type distribution was documented on a percentage basis (smallest unit 5 %). Based on this, a total of 20 subsamples were divided proportionally between the occurring substrate types: Each subsample substrate unit (5 %) was sampled by kick sampling ten times using a net with a surface of 0.0625m^2 and a mesh size of 0.5 mm (Liess et al., 2021). The sampled organisms were separated from the coarse organic debris using a column sieve set. Individuals were sorted into white trays with tweezers. Then, the citizen scientists identified the sampled macroinvertebrates alive and on site at least to family level and counted them using stereomicroscopes with 20-fold magnification. The taxonomic and abundance data were entered into the SPEAR

calculator (https://www.systemecology.de/indicate/) to determine a SPEAR$_{pesticides}$ value and a corresponding biological status class for each site. Afterwards, the CS samples were preserved in 90 % ethanol and re-identified in the laboratory by professionals to examine CS identification accuracy.

In the professional monitoring, macroinvertebrates were sampled twice per site (once in April and once in June 2021 as explained above), preserved in 90 % ethanol and identified in the laboratory. While the citizen scientists had only one afternoon per sample site to complete the sorting, counting and identification, professionals had more time and worked with high resolution microscopes. As such, counting and identification effort was higher in the professional monitoring than in the CS monitoring. For the pairwise comparisons of CS and professional macroinvertebrate samples, we selected the professional samples that were closest in time to the CS samples. The average time interval between the CS and professional macroinvertebrate samples was 14.4 days (sd 9.3).

Stream hydromorphology was recorded according to the official protocol by the German Water Working Group of the Federal States (LAWA, 2019) in both CS and professional monitoring. Citizen scientists used an illustrated and annotated version of the official protocol (see Table A2.3). Like in the professional monitoring, they quantified all hydromorphological criteria required under the WFD, including meandering of the watercourse, variation in stream depth and width, flow diversity as well as bed habitat structure, riparian conditions and land use within a 100 m river stretch of the sample site (European Commission, 2000). Citizen scientists determined hydromorphology index values using prepared excel spreadsheets for stream type-specific index calculation according to WFD. Index values were classified into one of the five WFD hydromorphology status classes.

As an additional information, citizen scientists measured physicochemical water parameters (i.e. nitrite, nitrate, pH, water temperature, dissolved oxygen, electrical conductivity) once per site in the afternoon of the CS sampling day. In the professional monitoring, pH values and nutrient concentrations were measured at intervals of three weeks (five times per site). Dissolved oxygen, water temperature and electrical conductivity were continuously measured from April to July in a three-minute interval using multi-parameter probes. For information about measuring devices, sampling days and methods to compare the CS and professional physicochemical data, see appendix Table A7 and Fig. A3.1–2.

The professional monitoring included two sampling methods for pesticide detection (Liess et al., 2021). Grab samples were taken regularly in a three-week cycle (following governmental monitoring practices under the WFD, regardless of weather conditions). Automatic rain Event-Driven Samplers (EDS) captured runoff-induced exposure peaks associated with heavy rainfall (Liess et al., 1999). Pesticide concentrations were determined by liquid chromatography-electrospray ionization-mass spectrometry (Halbach et al., 2021, see Table A8 for details on measured target substances). Based on the 50 % lethal concentration (LC50) in acute standard laboratory test systems (*Daphnia magna* or *Chironomus* sp.), measured pesticide concentrations were converted to macroinvertebrate toxicity (Toxic Units, TU). The peak pesticide exposure (TUmax) was calculated according to Liess et al. (2021) and describes the highest single substance toxicity measured in the water samples per site.

2.4. Statistical analysis

For each type of variable (biological, hydromorphological and physicochemical), we compared the CS data to the professional data to assess CS data accuracy. We analyzed two major components of accuracy: 1) data precision, i.e. the amount of variation in the data, using Pearson's or Spearman's correlation coefficient, and 2) data bias, i.e. systematic under- or overestimation of variables, using the concordance correlation coefficient *CCC* (Lin, 1989). To quantify bias, *CCC* supplements Pearson's correlation coefficient with a bias correction factor. The bias correction factor measures how far the best-fit line deviates from a line at 45 degrees, with a value of 1 indicating no deviation. *CCC* was calculated using the *epi*.ccc

function in the R package "epiR" (Version 2.0.41, Steven and Sergeant, 2022). Because the CS and professional data could lie on the 1:1 line but still differ in absolute values, we also used linear and general linear mixed effect models, including site as a random effect, to compare CS and professional sampled data. Both LMM and GLMM were calculated with the R package "lme4" (Version 1.1–29, Bates et al., 2015).

To analyze macroinvertebrate identification accuracy, we calculated the mean number of macroinvertebrate taxa recorded per site (on order, family, genus and species levels) by citizen scientists and by professionals. We tested for differences between taxon numbers per site recorded in the CS and professional monitoring with a linear mixed effect model (LMM). By re-identifying the preserved CS macroinvertebrate samples, we analyzed which proportion of taxa (on order, family, genus and species level) had been identified correctly by the citizen scientists.

Differences between CS and professional macroinvertebrate total abundance counts (summed across taxa) were tested using a generalized linear mixed effect model (GLMM), assuming a negative binomial error structure because of the overdispersion in the abundance counts. For those macroinvertebrate taxa recorded by both citizen scientists and professionals, we assessed the agreement of CS and professional abundance counts on family and order level (on relative scales) using Pearson's rank correlation coefficient and the concordance correlation coefficient.

We also examined whether the distribution of functional traits in the macroinvertebrate communities recorded by citizen scientists and professionals differed. For this, we extracted trait information for each recorded taxon from the SPEAR calculator (https://www.systemecology.de/indicate/). Then, we calculated community weighted means (mean of trait values weighted by their log abundance) with the site-level data for each of the four SPEAR$_{pesticides}$ macroinvertebrate traits: (i) Physiological sensitivity to pesticides (measured as the 50 % lethal concentration for each taxon on a log scale, normed by reference organism *Daphnia magna*); (ii) Generation time in years; (iii) Pesticide exposition (classified sensitive if taxon has aquatic stages in spring and early summer); and (iv) Refuge (classified sensitive if taxon cannot migrate and recolonize from refuge habitats).

To quantify the agreement between CS and professional SPEAR$_{pesticides}$ or hydromorphology or physicochemical values, we used Pearson's or Spearman's (when the data were not normally distributed) correlation coefficients and the concordance correlation coefficient CCC as a measure of bias. We tested for differences in the CS and professional community weighted mean values and index values (on absolute scales) using LMMs (non-normally distributed, proportional variables were arcsine-transformed).

To examine which factors could have influenced the accuracy of CS SPEAR$_{pesticides}$ and hydromorphology index values, we first calculated the site-specific differences between CS and professional index values. We then fitted multiple linear regression models including three numeric predictors (CS group size, average age of the CS group, number of days gap between the CS and the professional monitoring) and two categorical predictors (CS group category- see Section 2.2.- and prior experience of CS group - FLOW training excluded). Model residuals were checked with diagnostic plots and intercorrelation of the predictors was examined using the variance inflation factor. We also determined which proportion of stream sites had been assigned to the same WFD status class by citizen scientists and professionals with regards to hydromorphology and SPEAR$_{pesticides}$.

Finally, we used single linear regression models to analyze the relationship between CS or professional SPEAR$_{pesticides}$ values and peak pesticide concentrations (TUmax) measured directly at the stream sites. All statistical analyses were done with R (R Core Team, 2022, Version 4.1.2).

3. Results

3.1. Macroinvertebrate sampling and identification

The average number of taxa recorded per site ($n = 28$) was significantly higher in the professional monitoring (mean 27.3, sd 9.5) than in the CS monitoring (mean 17.9, sd 5.2; estimate of fixed effects 9.36, SE 1.68,

df $= 27$, $p < 0.001$). In the CS taxa lists, genus level was the most precise level of identification for most taxa (51 %), while 26 % were recorded at family level, 22 % at species level and 1 % on order level. In the professional monitoring, most taxa could be identified to species level (46 %), while 25 % each were recorded at genus and family level and 4 % at order level (Fig. 1).

The re-identification of conserved CS macroinvertebrate samples in the laboratory showed that the rate of correct CS identifications was very high at the order and family level (99 % and 91 % respectively). At genus and species level, correct CS identification rates were lower (65 % and 61 % respectively). On average, 42 % of the macroinvertebrate families recorded per sample site were documented in both the CS and professional taxa lists (sd 9.6 %). 25 % of the families were only recorded by the citizen scientists (sd 9.7 %), whereas 34 % (sd 12.0 %) were unique to the professional taxa lists.

The average total macroinvertebrate abundance recorded per site ($n = 28$) was much higher in the professional monitoring (mean 1488.8, sd 1351.6) than in the CS monitoring (mean 511.1, sd 281.5; estimate of fixed effects $= 1.07$, SE $= 0.19$, z $= 5.46$, $p < 0.001$). For the macroinvertebrate taxa recorded in both CS and professional taxa lists, the CS and professional abundance counts were significantly correlated when taking into account the differences in sampling effort by calculating relative abundance counts (Fig. 2). Nonetheless, the line of best fit deviated systematically from the 1:1 line due to less variation in relative abundance counts in the CS data than in the professional data (Fig. 2). The concordance of CS and professional abundance counts CCC was higher for common macroinvertebrate families (i.e. families recorded at numerous sites) than for rare macroinvertebrate families (recorded at fewer sites, Fig. 3). The (log-transformed) commonness of macroinvertebrate taxa was a significant predictor for the concordance of CS and professional abundance counts on family level ($R^2 = 0.21$, F $(1,56) = 14.79$, estimate 0.30, $p < 0.001$).

3.2. Distribution of macroinvertebrate functional traits

Testing for differences in macroinvertebrate functional traits, we observed no significant difference between CS and professional community weighted means for physiological sensitivity to pesticides (estimate $= -0.03$, SE $= 0.3$, df $= 27$, $t = -0.86$, $p = 0.4$) and migration ability (i.e. refuge taxa; estimate $= 0.01$, SE $= 0.02$, df $= 27$, $t = 0.65$, $p = 0.52$; see Fig. 4A + C). Community weighted means for the traits generation time (estimate $= -0.04$, SE $= 0.02$, df $= 27$, $t = -2.12$, $p = 0.04$) and
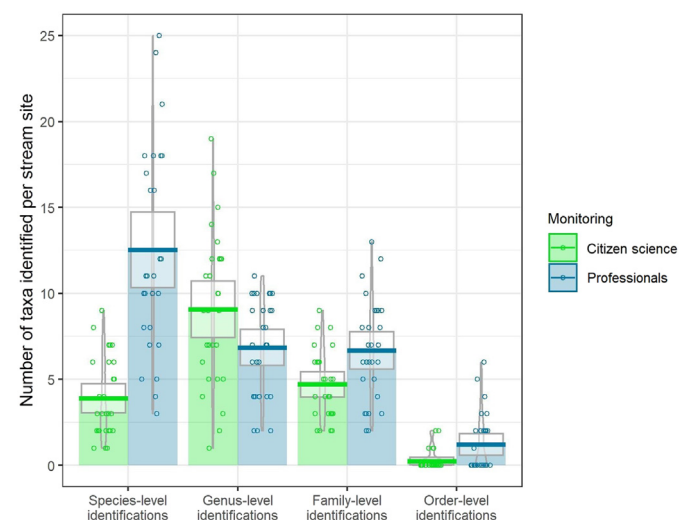

Fig. 1. Number of species, genus, family and order level identifications per stream site ($n = 28$) in the citizen science and professional monitoring (non-cumulative representation, i.e. taxa determined up to species level are not counted in the superordinate categories genus, family and order). For details on differences between CS and professional order level identifications, see appendix Table A3.
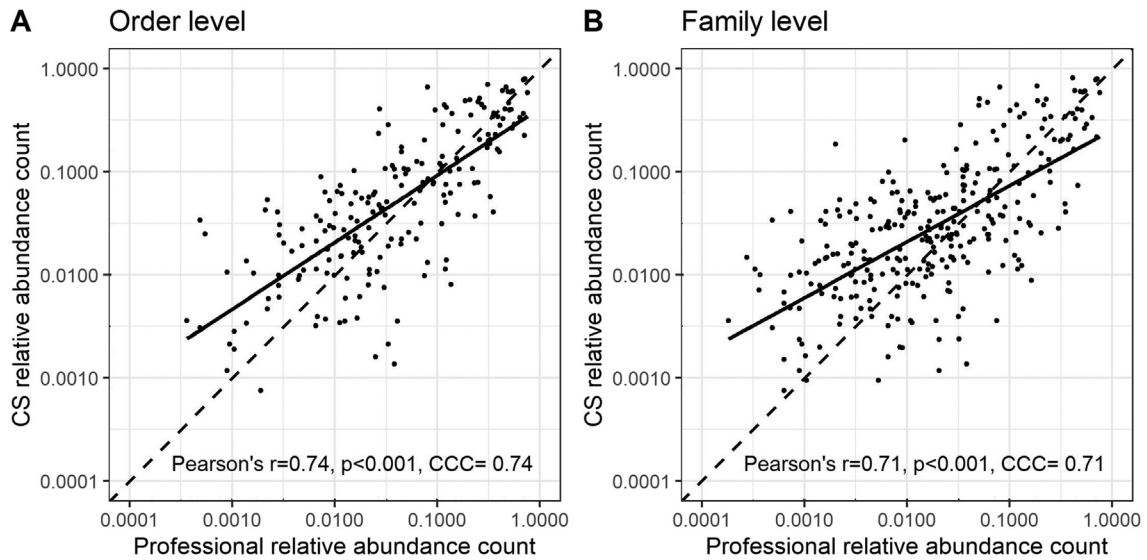
Fig. 2. Correlation of citizen science and professional macroinvertebrate abundance counts on order and family level. To account for differences in sampling effort, taxa specific abundances were divided by the total macroinvertebrate abundance counts recorded by citizen scientists and professionals at the respective stream sites. Abundances were log10-transformed. Correlation coefficients are based on those taxa which were recorded at the respective stream sites in both CS and professional monitoring (dashed line = 1:1 line, black line = regression line).
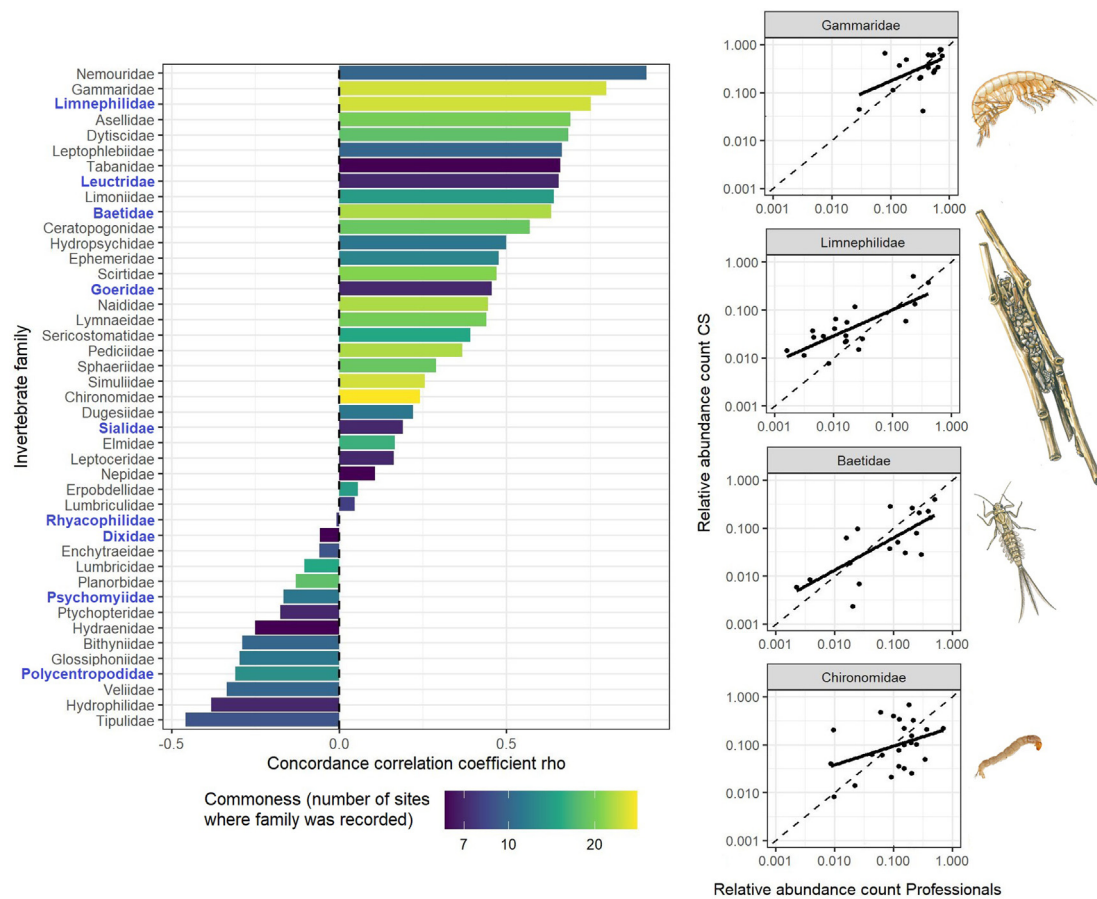

Fig. 3. Agreement between citizen science and professional macroinvertebrate family abundance counts visualized in the form of concordance correlation coefficient (*CCC*). *CCC* values of ±1 indicate perfect concordance (or discordance), values near zero indicate very low concordance. Bars are colored according to commonness of the respective macroinvertebrate families. Pesticide-sensitive macroinvertebrate families are indicated with blue letters. Families which occurred at <6 sites were excluded. For 95 % confidence intervals of CS and professional relative abundance counts, see appendix Table A4. Macroinvertebrate images © Franckh-Kosmos Verlag (Engelhardt et al., 2020). Chironomidae © Cyril Bennett.

pesticide exposition (estimate = −0.04, SE = 0.02, df = 27, t = −1.95, p = 0.06) were slightly higher in the macroinvertebrate communities recorded by citizen scientists than in the communities recorded by professionals (Fig. 4B + D). This did not result in a higher number of recorded pesticide-sensitive taxa, as only macroinvertebrate taxa classified sensitive with regards to all four SPEAR$_{pesticides}$ traits (Fig. 4A-D) are counted as pesticide-sensitive SPEAR taxa (Liess and Von Der Ohe, 2005). In total, both CS and professional monitoring recorded an average of 20 % pesticide-sensitive SPEAR taxa at the n = 28 stream sites. Overall, we found no significant difference in CS and professional community weighted means for SPEAR taxa (estimate = −0.008, SE = 0.03, df = 27, t = −0.27, p = 0.78; see Fig. 4E).

3.3. Agreement of SPEAR$_{pesticides}$, hydromorphology and physicochemical data

We found that CS and professional SPEAR$_{pesticides}$ values (n = 28 stream sites) were highly correlated (Pearson's r = 0.76, p < 0.001, Fig. 5A). Bias was small (CCC = 0.75) and we observed no significant difference between average CS and professional SPEAR$_{pesticides}$ values (CS mean 0.4, sd 0.23, professional mean 0.4, sd 0.26; estimate of fixed effects = −0.005, SE = 0.03, df = 27, t = −0.17, p = 0.86). The results also showed that 61 % of the stream sites had been rated with the same SPEAR status class by both citizen scientists and professionals, while 32 % of the sites were rated one SPEAR class apart and 7 % were rated two SPEAR classes apart (appendix Fig. A2.1). As such, CS and professional SPEAR assessments agreed in 90 % of the cases on whether a stream achieved a good status in terms of pesticide exposure (i.e. classification as SPEAR status class I or II). The multiple regression modeling showed that none of the analyzed predictors was significantly related to the difference between CS and professional SPEAR$_{pesticides}$ values (appendix Table A5).

CS and professional hydromorphology assessments agreed on whether a stream achieved a good ecological status according to WFD in 82 % of the cases. In detail, we found that 50 % of the stream sites (n = 28) were rated with the same status class by both monitoring teams, while the other 50 % were rated one status class apart (appendix Fig. A2.2). The CS and professional hydromorphology index values were highly correlated (Pearson's r = 0.72, p < 0.001, see Fig. 5B). Bias was moderate (CCC = 0.68) and we found that citizen scientists assessed stream hydromorphology slightly more negatively than professionals did (CS mean 4.13, sd 0.97; professional mean 3.77, sd 0.99; estimate of fixed effects = −0.36, SE = 0.14, df = 27, t = −2.62, p = 0.01). Differences between CS and professional hydromorphology index values were not explained by hypothesized variables as CS group category, time between CS and professional monitoring day or prior experience of the CS group, but there was marginal evidence for an effect of CS group size (appendix Table A6).

Regarding physicochemical water parameters (n = 28 sites, appendix Fig. A3.1–2), CS and professional data for water temperature (r = 0.94, p < 0.001) were strongly correlated, whereas other parameters such as pH (rho = 0.57, p < 0.01), dissolved oxygen (rho = 0.57, p < 0.05), electrical conductivity (r = 0.53, p < 0.05), flow velocity (r = 0.42, p < 0.05) and nutrient measurements (nitrate rho = 0.46, p < 0.05, nitrite rho = 0.45, p < 0.05) were only moderately to weakly correlated. Here, measuring bias (quantified as CCC values) was moderate to high (appendix Fig. A3.1–2).

3.4. Relationship between citizen science generated SPEAR$_{pesticides}$ values and measured pesticide toxicity

For both the CS and the professional SPEAR$_{pesticides}$ values (n = 21), we found a clear negative relationship between SPEAR$_{pesticides}$ and measured
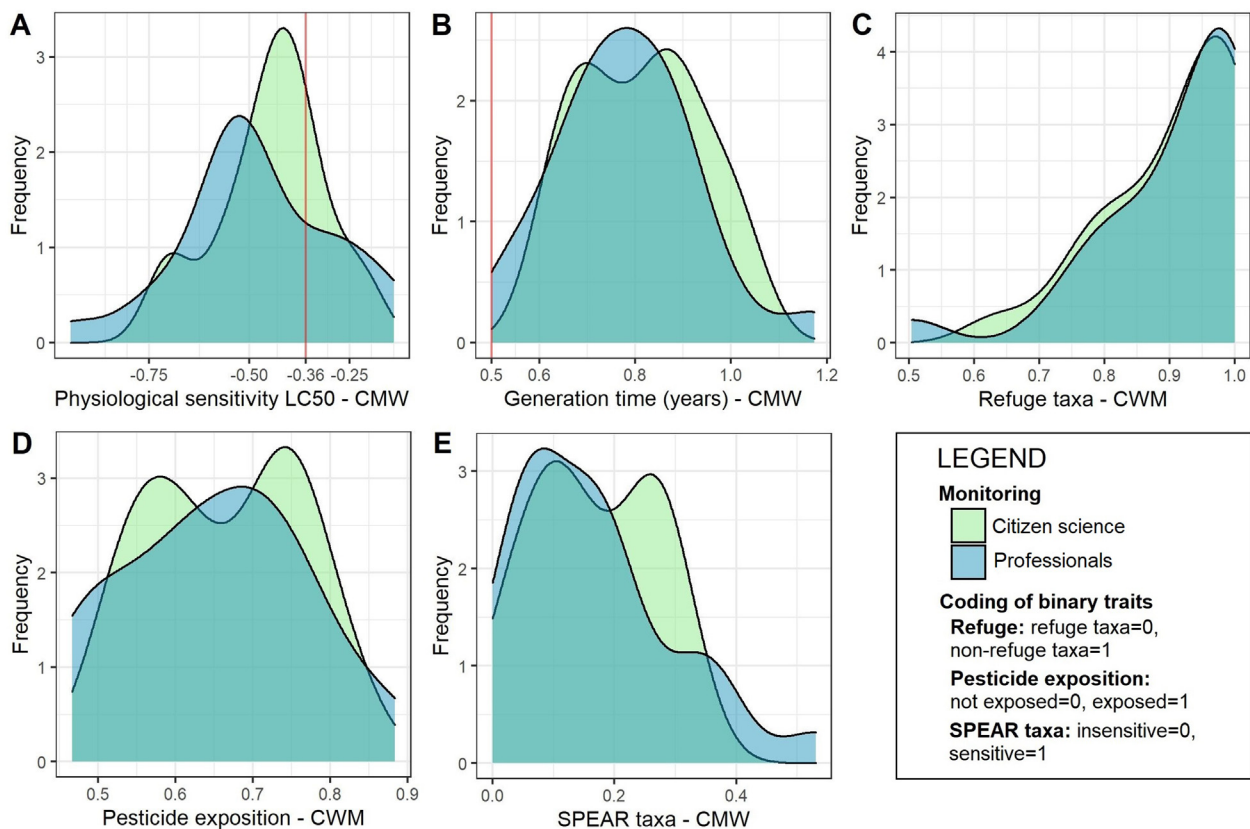

Fig. 4. Distribution of community weighted means (CWM) for SPEAR$_{pesticides}$ functional traits recorded in the citizen science and professional macroinvertebrate taxa lists (n = 28 stream sites). Red lines indicate threshold values for (A) physiological sensitivity (LC50 > −0.36) and for (B) reproductive sensitivity (generation time > 0.5 years). (C) Migration ability from refuge habitats, (D) pesticide exposition of aquatic stages in spring and early summer and (E) SPEAR classification of taxa as either pesticide-sensitive or insensitive.
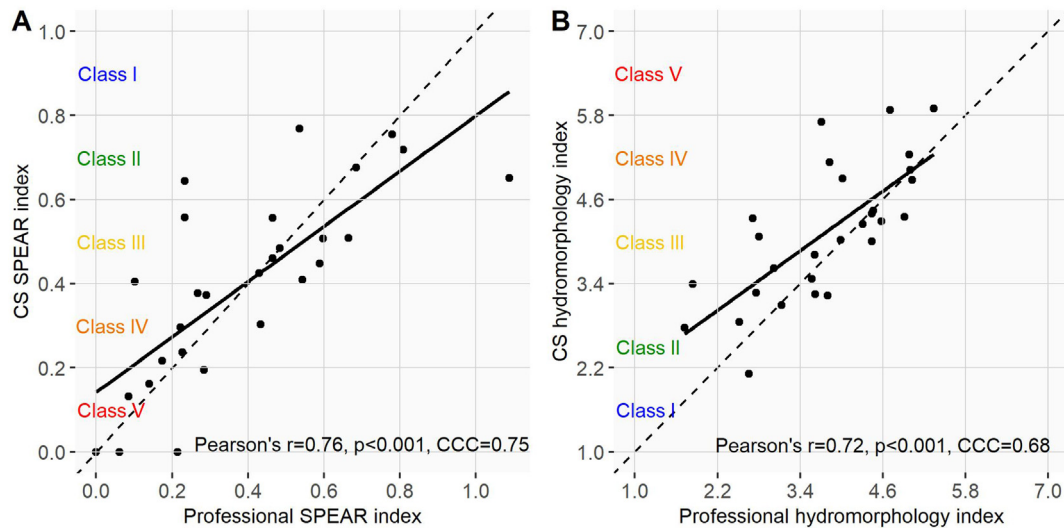
Fig. 5. Correlation of citizen science and professional (A) SPEAR$_{pesticides}$ index values and (B) hydromorphology index values ($n = 28$). The regression line is shown in black, the dashed line represents the 1:1 line. CCC = concordance correlation coefficient. Status classes according to Water Framework Directive (WFD): I: "high", II: "good", III: "moderate", IV: "poor", V: "bad".

peak pesticide concentrations (TUmax, Fig. 6). The CS generated SPEAR$_{pesticides}$ values explained 28 % of the variance in TUmax values ($R^2 = 0.28$, F(1,19) = 7.21, estimate = -0.12, SE = 0.04, $t = -2.68$, $p = 0.01$, res. SE = 0.16) while the professional SPEAR$_{pesticides}$ values explained 33 % of the variance in TUmax values ($R^2 = 0.33$, F(1,19) = 9.28, estimate = -0.18, SE = 0.06, $t = -3.05$, $p = 0.007$, res. SE = 0.21). We also analyzed to which degree professional re-identification to species level of the CS invertebrate samples in the laboratory changed the relationship between CS SPEAR$_{pesticides}$ values and measured pesticide toxicity. The resulting corrected CS SPEAR$_{pesticides}$ values explained 37 % of the variance in TUmax (see appendix Fig. A4).

4. Discussion

Through an in-depth assessment of citizen science (CS) monitoring data accuracy and comparison with professional monitoring data for 28 stream
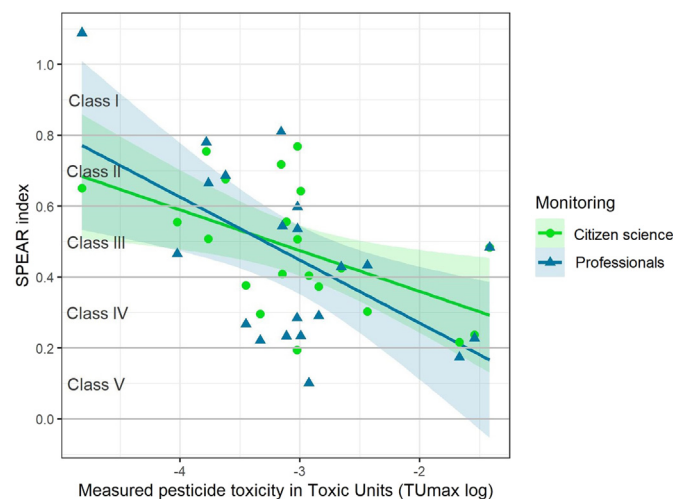

Fig. 6. Relationship between citizen science and professional SPEAR$_{pesticides}$ index values and measured peak pesticide toxicity (TUmax) at $n = 21$ stream sites. Status classes according to Water Framework Directive (WFD): I: "high", II: "good", III: "moderate", IV: "poor", V: "bad".

sites, we show that trained citizen scientists can provide accurate data on pesticide exposure and hydromorphology that compare well with professional assessments. Citizen scientists were able to correctly identify macroinvertebrates at order and family level, which was sufficient for calculation of the bio-indicator SPEAR$_{pesticides}$. CS and professional relative macroinvertebrate abundance counts agreed especially well for common or easy-to-identify taxa. Further, SPEAR$_{pesticides}$ values derived from CS taxa lists performed almost as well as professionally generated SPEAR$_{pesticides}$ data when related to measured pesticide toxicity (TUmax). Based on these findings, we discuss opportunities, limitations and specific requirements of CS in stream monitoring.

4.1. Accuracy of citizen science macroinvertebrate identification

Identification accuracy of citizen scientists depended on taxonomic resolution and commonness of the invertebrate taxa. Correct CS identification rates were especially high at order (99 %) and family level (91 %). Several experienced participants were eager to record taxa at genus or species level, although we indicated that family level was sufficient. Yet, the lower correct CS identification rates at genus and species level (65 % and 61 %, respectively) clearly show a trade-off between taxonomic resolution and identification accuracy (Moffett and Neale, 2015). Similar to results presented by Fore et al. (2001) and Reid et al. (2016), average taxon richness and invertebrate abundance counts recorded per site were significantly lower in the CS than in the professional monitoring. Differences between the CS and professional taxa lists most likely occurred during macroinvertebrate identification: Accurately sorting out macroinvertebrates from a net sample and correctly identifying them to genus or species level usually requires months of intensive practice and experience. Moreover, for the citizen scientists, time for invertebrate sorting, identification and counting was restricted to one afternoon in the field and CS identification equipment (simple identification booklet, stereomicroscopes with 20-fold magnification suitable for field use) was not designed to equal professional lab equipment.

Consequently, CS projects and researchers should consider that with growing complexity and taxonomic precision of the required identification task, identification error rates are likely to increase. Regarding the field sampling method, the re-identification of CS invertebrate samples by professionals in the laboratory showed that the process of kick sampling does not seem to have caused relevant differences between the CS and professional taxa lists (see Fig. A4). Similarly, Fore et al. (2001) showed that

macroinvertebrate samples taken in a standardized way by trained citizen scientists did not differ significantly from professionally taken samples.

When taking into account the differences in CS and professional sampling effort (in terms of sorting, identification and counting), abundance counts of invertebrate orders and families recorded at the same sites by both citizen scientists and professional were highly correlated (Fig. 2), especially for common families (Fig. 3). We assume that this is partly because several of the common invertebrate families are relatively large and/or easy to recognize through their conspicuous body shape or type of movement, and were pointed out specifically in the CS identification trainings (e.g. *Gammaridae, Limnephilidae, Asellidae, Dytiscidae, Baetidae*). Moreover, for common families with numerous abundance data points, deviating CS and professional abundance counts on the site level can be compensated by averaging across sample sites. In contrast, spotting small and sometimes motionless insect larvae in the debris requires practice and a well-trained eye. Beginners tend to focus on picking out large (>10 mm), moving individuals, while smaller (1-10 mm), immobile specimens are sometimes overlooked (Nerbonne and Vondracek, 2003; Storey and Wright-Stow, 2017). Besides, some invertebrate traits relevant for identification (like mouthparts or the shape of gill filaments) are often difficult to recognize in small organisms with a 20-fold magnification. For rare taxa, the averaging effect of CS and professional abundance counts does not apply as much as for common taxa. Accordingly, we observed lower concordance (CCC) values for CS and professional abundance counts for many of the small invertebrate families (e.g. *Diptera* such as *Simuliidae, Chironomidae, Dixidae*) and for rare taxa only recorded at few sites (e.g. *Sialidae, Rhyacophilidae, Dixidae* and *Hydraenidae*, see Fig. 3).

Thus, CS is a particularly well suited monitoring approach if the indicator system in question is mainly based on common taxa and does not depend too much on rare taxa. For some rare families, however, we still found comparatively high concordance values (Fig. 3). These families have typical identification features that make them easy to distinguish (e.g. *Nemouridae*: small but broad-bodied, often dark colored and bristly stonefly larvae with divergent wing pads; *Leptophlebiidae*: mayfly larvae with finely divided, tree- or thread-shaped gill filaments; *Goeridae*: caddisfly larvae with small lateral ballast stones attached to its case; *Tabanidae*: large, segmented wormlike horsefly larvae with fleshy rings and pseudopods circling the body).

4.2. Assessment of stream biological status and pesticide exposure with citizen science generated invertebrate data

The strong correlation between CS and professional SPEAR$_{pesticides}$ index values ($r = 0.76$) demonstrates that citizen scientists can accurately capture macroinvertebrate functional traits and community composition in different ecological conditions. Storey et al. (2016) who compared CS and professional macroinvertebrate samples taken at the same sites and dates and observed a slightly stronger correlation between macroinvertebrate metrics ($r = 0.85$). Meanwhile, Moffett and Neale (2015) found a weaker correlation ($r = 0.54$) for CS and professional samples taken at slightly different times and dates. These results indicate that it is difficult to compare correlation coefficients across different study designs, macroinvertebrate metrics and data ranges. Since our comparison of CS and professional invertebrate community weighted means for SPEAR$_{pesticides}$ traits (Fig. 4) is based on a relatively small sample size ($n = 28$), future studies might re-examine potential differences on a larger scale. Because of the observed inconsistencies between CS and professional SPEAR$_{pesticides}$ status class assignments (in 39 % of the sample sites), we propose to additionally consider the more general and largely consistent classification of sites into "good" or "unsatisfactory" biological status (appendix Fig. A2.1B) or to apply more relaxed, "fuzzy" status class boundaries as demonstrated by Storey and Wright-Stow (2017).

Invertebrate communities are affected by multiple stressors, including pesticide and nutrient inputs and altered hydromorphology, and many taxa are sensitive to several different stressors. Therefore it is often challenging to quantify the effect of individual stressors based on

biological metrics (Lemm et al., 2019). Various studies have shown that SPEAR$_{pesticides}$ reliably indicates pesticide pressure across different biogeographical regions, but the bio-indicator also responds to a small degree to deficient hydromorphology (Liess et al., 2021; Liebmann et al., 2022).

The CS generated SPEAR$_{pesticides}$ values in our study explained the measured pesticide toxicity almost equally well as the professional data (28 % and 33 % of explained variance, Fig. 6). Thus, the coarser CS taxonomic identification levels and lower average CS taxon richness only slightly reduced the ability of the bio-indicator to indicate pesticide stress. These results corroborate findings from Beketov et al. (2009) and Liebmann et al. (2022) showing that SPEAR$_{pesticides}$ works well with family level data. Identification to family level is sufficient for determining valid SPEAR$_{pesticides}$ values in many cases since traits are characterized at family level for most freshwater invertebrate taxa and traits determining pesticide sensitivity are thought to show little variation within invertebrate families (Liebmann et al., 2022). Moreover, our results confirm findings from Liebmann et al. (2022) who observed that SPEAR$_{pesticides}$ also relates well to pesticide pressure when it is based on less accurate abundance data (e.g. abundance class data). Still, potential data users should take into account the trade-off between identification effort and pesticide indication accuracy when working with CS generated SPEAR$_{pesticides}$ data.

In our study, R^2 values for the association between CS or professional SPEAR$_{pesticides}$ and TUmax are slightly lower than in a previous study where professionally generated SPEAR$_{pesticides}$ values for $n = 101$ lowland stream sample sites explained 43 % of the variance in measured pesticide toxicity ($R^2 = 0.43, p < 0.001$, Liess et al., 2021). Our sample size was comparatively smaller with $n = 21$ per monitoring team (initially $n = 30$ each, excluding nine sites each because of drought and very low flow velocity). With a smaller sample size, probability increases that the invertebrate samples and corresponding SPEAR$_{pesticides}$ values from a given total population randomly yield a different R^2 (smaller or higher). Additionally, changes in pesticide use may have led to a slight underestimation of toxicity in our pesticide measurements: In April 2018, the EU adopted a comprehensive restriction of use for neonicotinoid insecticides. Since then, pyrethroids, which exert toxic effects on insects even in small quantities, have been increasingly used as substitute active substances. In our study, pyrethroids could not be analyzed, so that the determined TUmax values could be too low to represent the actual toxic effects on the macroinvertebrate communities, resulting in slightly weaker correlations between SPEAR index and measured toxicity.

4.3. Assessment of stream hydromorphology

Hydromorphology assessments require a high level of judgment and the variability of hydromorphological stream classifications has been shown to be quite high even among professional observers (Clapcott, 2015; Storey et al., 2016). Against this background, the observed correlation coefficient between CS and professional index values ($r = 0.72$) and the 82 % agreement rate of stream classifications into good or unsatisfactory ecological status (appendix Fig. A2.2B) is very satisfactory. Storey et al. (2016) found a similar correlation between CS and professional hydromorphology assessments ($r = 0.7$), while observing stronger correlations for some aspects of stream hydromorphology (e.g. technical changes in waterway, riparian components), and lower correlations for some other aspects which appear to be more difficult to assess for beginners (e.g. bank erosion and stability, sediment deposition and flow type). This demonstrates the importance of specific training and clarification of technical terms to prepare citizen scientists for the assessment of stream hydromorphology. Bias for CS hydromorphology data was stronger than for SPEAR$_{pesticides}$ data when compared to professional data. Storey et al. (2016) showed that CS data accuracy could be improved by averaging repeated CS hydromorphology assessments for each site and season. In addition, cross-validation among different CS observers could reduce bias and possibly optimize the agreement of CS and professional hydromorphological stream classifications into the five WFD status classes.

4.4. Assessment of physicochemical status

Unlike the hydromorphology and SPEAR$_{pesticides}$ assessments, the CS physicochemical water data (except for water temperature) were only moderately to weakly correlated to the professional data (appendix Fig. A3.1–2, possible explanations are listed in appendix Table A7).

To reduce bias caused by differences in CS and professional measuring equipment, CS test kits and measuring devices should be calibrated with professionally used systems before use. Moreover, internationally recognized protocols or "Standard Operating Procedures" for CS physicochemical water measurements (Quinlivan et al., 2020b) and a standardized set of user-friendly, inexpensive test kits should be provided that comply with the assessment ranges of international reporting systems (e.g. WFD or SDG 6.3.2). Newly developed, easy-to-use technologies for evaluating nutrient tests (as described in Zheng et al., 2022, based on a smartphone app and digital image colorimetry) could also help to improve CS physicochemical data accuracy.

Several studies that compared large-scale CS and professional data sets with numerous data points for each site and parameter found good agreement between CS and professional physicochemical water data despite differences in measuring times and devices (Albus et al., 2020; Dyer et al., 2014; Safford and Peters, 2018; Shupe, 2017). For CS to be a useful and reliable approach to physicochemical water status assessment, citizen scientists need to be incentivized to visit each sample site multiple times per season to conduct repeated measurements of each physicochemical parameter (i.e. in three-week intervals as done in the professional monitoring). This was not feasible in our study because citizen scientists had to travel up to 40 km to visit the sample sites preselected for the pesticide and professional measurements. For CS practitioners in structured monitoring programs, we recommend to provide key criteria for sample site selection to ensure data comparability and usability (McGoff et al., 2017). At the same time, citizen scientists should be encouraged to suggest suitable sample sites themselves to maintain motivation and feasibility and thereby ensure sufficient sampling effort (Scott and Frost, 2017). CS programs that need or want to focus monitoring effort on one to two comprehensive stream assessments per site and season and are capable of providing appropriate CS training should focus on biological indicators, such as SPEAR$_{pesticides}$, that are less variable within a season.

4.5. Opportunities, limitations and future outlook for citizen science stream monitoring

Innovative, large-scale monitoring approaches are needed to advance research on pesticide exposure and the ecological status of rivers and streams, and to efficiently track the effects of EU environmental policies and freshwater management interventions (Bieroza et al., 2021; Carvalho et al., 2019).

As we demonstrate, well-designed, appropriately managed CS stream monitoring programs can generate accurate macroinvertebrate and hydromorphology data that agree well with professional data. Therefore, CS data as produced in our study are well suited to fill data gaps and complement official stream monitoring programs, especially for small streams under 10km^2 that are not covered by WFD monitoring and reporting (Wick et al., 2019). This could be facilitated by new methods for integrating different data streams together to produce biodiversity indicators (Isaac et al., 2020).

Since pesticide inputs are a dominant stressor for sensitive insects in streams and of raising concern for environmental management, regular pesticide-specific biomonitoring is necessary (Liess et al., 2021). Therefore, using SPEAR$_{pesticides}$ in a CS context with family level data is a valuable approach to enable citizens to quantify pesticide pressure at local and large spatio-temporal scales and support the implementation of pesticide-specific management measures. Experiences from the UK Anglers' Riverfly Monitoring Initiative (Brooks et al., 2019) show that setting catchment-specific "trigger levels" for biological metrics is a successful approach for environmental authorities to use and benefit from CS data. For instance,

the responsible authorities could be notified if CS SPEAR$_{pesticides}$ index values within a certain catchment systematically fall into status classes IV (poor) or V (bad). In addition, as SPEAR$_{pesticides}$ is not designed for assessing overall ecological status, researchers and other stakeholders may also use the standardized CS invertebrate data to calculate integrative indices such as the biological monitoring working party (BMWP) index or the Average Score Per Taxon (ASPT). Providing a holistic ecological assessment, these indices have been shown to identify the community's response to multiple stressors based on family-level data (Moolna et al., 2020). Additionally, CS hydromorphology data could be used to flag stream sites where degraded habitat quality severely affects stream ecological functioning. Thus, if it was expanded to other regions, the CS stream monitoring could produce up-to-date information on the regional realization of WFD goals (in terms of pesticide exposure and overall ecological status) with higher spatio-temporal resolution than is currently possible in official monitoring schemes.

By raising awareness and encouraging environmental stewardship among citizen scientists and other project stakeholders, CS monitoring is also a valuable tool to foster civic advocacy for biodiversity and freshwater conservation (Ballard et al., 2017; McKinley et al., 2017; Peter et al., 2021). CS programs can improve the citizens' valuation and scientific understanding of freshwater ecosystems (Ballard et al., 2017) and positively develop their sense of place and connectedness to local rivers and streams (Church et al., 2019; Haywood et al., 2016, 2020). Several CS programs have already made important contributions to uncovering the causes of water pollution, for example the citizen monitoring initiative during the Flint Water crisis (Pieper et al., 2018) or the international program FreshwaterWatch (Earth Watch, 2020). Likewise, by reporting bad or deteriorating ecological conditions to local authorities in a timely fashion, citizen scientists participating in the stream monitoring program FLOW could initiate official investigations and actions to mitigate pesticide pollution hotspots that might otherwise not be detected (Brooks et al., 2019; Edwards, 2016). Similarly, they could be empowered to monitor the effects of recent environmental regulation or of river restoration projects (Edwards et al., 2018; Huddart et al., 2016).

Naturally, CS monitoring also has its limitations and "is no panacea" (Metcalfe et al., 2022:4). Due to the generally lower CS taxon richness and abundance recordings and the coarser level of CS invertebrate identification, CS stream monitoring should not be expected to provide complete invertebrate taxa lists with accurate taxonomic information on genus or species level, especially in near natural streams with high invertebrate diversity. Detecting small differences in invertebrate community composition requires accurate invertebrate identification at genus or species level (Chessman et al., 2007; Fore et al., 2001). Therefore, CS monitoring might not be suitable to record subtle differences in stream biological condition.

To optimize CS data accuracy for hydromorphology and to enable accurate CS physicochemical assessments, FLOW CS monitoring design needs to be refined, e.g. by increasing measuring frequency.

The great potential and also main challenge of CS lays in meeting several complex goals at once: CS programs aim to motivate citizens to actively engage in research to generate new scientific knowledge and promote environmental learning (Turrini et al., 2018). Since project resources and engagement time of citizens are limited, it is important for CS programs to negotiate trade-offs and enhance synergies between these scientific, educational and participatory goals. To produce valid pesticide indication data, CS monitoring design should be closely aligned with the scientific standards for stream assessment. Since the level of CS training and learning progress will affect CS data accuracy (Fore et al., 2001), it is important to invest in sufficient training with experts. Further, quality assurance during fieldwork is essential (Storey et al., 2016). In the FLOW monitoring, data and results were checked and discussed with an experienced citizen scientist or a member of the FLOW team before submission, so that open questions and problems could be clarified. Feedback by experts or scientists helps citizen scientists gain confidence in their abilities and motivates them to continue engaging in the monitoring activity (Storey et al., 2016;

Weeser et al., 2018). We therefore recommend that each CS group should be accompanied by an experienced participant or mentor during the fieldwork. At the same time, for citizens to benefit from their experience and stay motivated, it is important not to overburden them and include the fun aspect of discovering stream biodiversity while being outdoors with their peers. This may differ among stakeholder groups: while experienced NGO members and anglers with prior knowledge might be capable and keen to closely follow scientific standards, newcomers such as high school students usually need more guidance and focus on nature discovery.

Since considerable investments in CS training and guidance are needed, CS programs such as ours should not be expected to be cost-effective (Capdevila et al., 2020) at least in their establishment phase. CS data accuracy will increase over time with the citizen scientists' experience in identifying invertebrates. Therefore, it is important to invest in the long-term retention of trained, experienced participants. If CS programs aim to encourage their participants to engage in freshwater conservation, sufficient project management resources should be allocated to actively promote this goal. Research has shown that social interaction and opportunities for community building are particularly important motivations for citizen scientists to advocate for conservation (Agnello et al., 2022; Asah et al., 2014; Richter et al., 2018). CS projects can leverage this potential by encouraging monitoring in teams or groups as done in the FLOW program, or by organizing opportunities to meet and socialize with other, likeminded (citizen) scientists. It is also important for citizen scientists to experience that they can make a difference (Day et al., 2022; Newman et al., 2017). For instance, inviting citizen scientists and other stakeholders to co-design and implement local stream management measures can create new ecological insights, help to legitimize conservation measures, and reduce conflicts among stakeholders (McKinley et al., 2017). While the FLOW project's participants came from various backgrounds, we were, at this stage, unable to recruit farmers as important stakeholders in the pesticide issue. When comparing the CS to the professional data, however, we found no evidence that the CS data interpretation was influenced by group composition. In future, involving farmers into stream monitoring could enrich citizen dialogue about freshwater protection and support the development of integrative, societally supported water protection measures. By harnessing expertise of citizen scientists and different stakeholders to monitor, evaluate and adjust stream management measures, CS can foster a more sustainable, adaptive management of freshwater ecosystems (Ballard et al., 2017; Nerbonne and Nelson, 2004; Yardi et al., 2019).

5. Conclusions

To reach conservation goals for freshwater ecosystems, large-scale monitoring and reliable data on the ecological condition of streams are essential. In this context, it is important to assess drivers of change at the terrestrial-aquatic interface, such as the effects of terrestrial pesticide pollution on aquatic ecosystems. With the FLOW project, we developed a citizen science (CS) stream monitoring focusing on small streams aligned with European Water Framework Directive (WFD) standards to enable maximum impact for research and policy uptake. Our results demonstrate that macroinvertebrate and hydromorphology data collected by trained citizen scientists are of sufficiently high accuracy to quantify stream ecological stressors such as habitat degradation and pesticide exposure originating from terrestrial application. These CS records are suitable to assess stream health and ecological status according to WFD. CS invertebrate identification accuracy was high at order and family level, while overall taxa detection rates were lower in the CS than in the professional monitoring. Citizen scientists adequately captured the distribution of invertebrate functional traits that are necessary to determine pesticide-sensitive taxa, and thereby provided valid pesticide bio-indicator (SPEAR$_{pesticides}$) values and assessments. The overall accuracy of CS hydromorphology assessments was good, yet we suggest that it could be further optimized through repeated CS assessments for each monitoring site and season. For physicochemical point measurements, the agreement between CS and professional data was, however, only moderate to low. We conclude that CS is only suitable for physicochemical water assessment if citizen scientists can be incentivized to conduct multiple repeated measurements per sample site and season after being trained.

As an outlook for CS in the freshwater realm, we suggest that CS stream monitoring could become an important, valuable tool to augment the existing (inter-)national, scientific and regulatory stream monitoring and management across space and time (Albus et al., 2020; Hadj-Hammou et al., 2017; Moffett and Neale, 2015). To realize this potential, CS freshwater monitoring should be closely aligned with official monitoring standards (e.g. WFD or SDG indicator 6.3.2, see Quinlivan et al., 2020a, 2020b). Supported with appropriate training and guidance as well as data quality assurance, citizen scientists can contribute new insights into stream ecological conditions and effectively support researchers, government agencies and NGOs in achieving freshwater research and conservation goals.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.159607.

References

Abbott, B.W., Moatar, F., Gauthier, O., Fovet, O., Antoine, V., Ragueneau, O., 2018. Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France. Sci. Total Environ. 624, 845–858. https://doi.org/10.1016/j.scitotenv.2017.12.176.

Agnello, G., Vercammen, A., Knight, A.T., 2022. Understanding citizen scientists'willingness to invest in, and advocate for, conservation. Biol. Conserv. 265, 109422. https://doi.org/10.1016/j.biocon.2021.109422.

Albus, K., Thompson, R., Mitchell, F., 2019. Usability of existing volunteer water monitoring data: what can the literature tell us? Citiz. Sci. Theory Pract. 4, 28. https://doi.org/10.5334/cstp.222.

Albus, K.H., Thompson, R., Mitchell, F., Kennedy, J., Ponette-González, A.G., 2020. Accuracy of long-term volunteer water monitoring data: a multiscale analysis from a statewide citizen science program. PLoS One 15, e0227540. https://doi.org/10.1371/journal.pone.0227540.

Asah, S.T., Lenentine, M.M., Blahna, D.J., 2014. Benefits of urban landscape eco-volunteerism: mixed methods segmentation analysis and implications for volunteer retention. Landsc. Urban Plan. 123, 108–113. https://doi.org/10.1016/j.landurbplan.2013.12.011.

Ballard, H.L., Dixon, C.G.H., Harris, E.M., 2017. Youth-focused citizen science: examining the role of environmental science learning and agency for conservation. Biol. Conserv. 208, 65–75. https://doi.org/10.1016/j.biocon.2016.05.024.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67. https://doi.org/10.18637/jss.v067.i01.

Beketov, M.A., Foit, K., Schäfer, R.B., Schriever, C.A., Sacchi, A., Capri, E., Biggs, J., Wells, C., Liess, M., 2009. SPEAR indicates pesticide effects in streams – comparative use of species- and family-level biomonitoring data. Environ. Pollut. 157, 1841–1848. https://doi.org/10.1016/j.envpol.2009.01.021.

BfN, 2021. Gesamtlänge Fließgewässer. Bundesamt für Naturschutz. URL (accessed 03/08/2022) https://www.bfn.de/infothek/daten-fakten/nutzung-dernatur/nutzung-von-binnengewaessern/ii-32-1-gesamtlaenge-fliessgewaessereinzugsgebiet.html.

Bieroza, M.Z., Bol, R., Glendell, M., 2021. What is the deal with the Green Deal: will the new strategy help to improve European freshwater quality beyond the Water Framework Directive? Sci. Total Environ. 791, 148080. https://doi.org/10.1016/j.scitotenv.2021.148080.

Bowler, D.E., Eichenberg, D., Conze, K., Suhling, F., Baumann, K., Benken, T., Bönsel, A., Bittner, T., Drews, A., Günther, A., Isaac, N.J.B., Petzold, F., Seyring, M., Spengler, T., Trockur, B., Willigalla, C., Bruelheide, H., Jansen, F., Bonn, A., 2021. Winners and losers over 35 years of dragonfly and damselfly distributional change in Germany. Divers. Distrib. 27, 1353–1366. https://doi.org/10.1111/ddi.13274.

Bowser, A., Cooper, C., de Sherbinin, A., Wiggins, A., Brenton, P., Chuang, T.-R., Faustman, E., Haklay, M., Meloche, M., 2020. Still in need of norms: the state of the data in citizen science. Citiz. Sci. Theory Pract. 5, 18. https://doi.org/10.5334/cstp.303.

Brooks, S.J., Fitch, B., Davy-Bowker, J., Codesal, S.A., 2019. Anglers'riverfly monitoring initiative (ARMI): a UK-wide citizen science project for water quality assessment. Freshw. Sci. 38, 270–280. https://doi.org/10.1086/703397.

Capdevila, A., Kokimova, A., Sinha Ray, S., Avellán, T., Kim, J., Kirschke, S., 2020. Success factors for citizen science projects in water quality monitoring. Sci. Total Environ. 728, 137843. https://doi.org/10.1016/j.scitotenv.2020.137843.

Cardoso, P., Erwin, T.L., Borges, P.A.V., New, T.R., 2011. The seven impediments in invertebrate conservation and how to overcome them. Biol. Conserv. 144, 2647–2655. https://doi.org/10.1016/j.biocon.2011.07.024.

Carlson, T., Cohen, A., 2018. Linking community-based monitoring to water policy: perceptions of citizen scientists. J. Environ. Manag. 219, 168–177. https://doi.org/10.1016/j.jenvman.2018.04.077.

Carvalho, L., Mackay, E.B., Cardoso, A.C., Baattrup-Pedersen, A., Birk, S., Blackstock, K.L., Borics, G., Borja, A., Feld, C.K., Ferreira, M.T., Globevnik, L., Grizzetti, B., Hendry, S., Hering, D., Kelly, M., Langaas, S., Meissner, K., Panagopoulos, Y., Penning, E., Rouillard, J., Sabater, S., Schmedtje, U., Spears, B.M., Venohr, M., van de Bund, W., Solheim, A.L., 2019. Protecting and restoring Europe's waters: an analysis of the future development needs of the Water Framework Directive. Sci. Total Environ. 658, 1228–1238. https://doi.org/10.1016/j.scitotenv.2018.12.255.

Chessman, B., Williams, S., Besley, C., 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. 26, 546–565. https://doi.org/10.1899/06-074.1.

Church, S.P., Payne, L.B., Peel, S., Prokopy, L.S., 2019. Beyond water data: benefits to volunteers and to local water from a citizen science program. J. Environ. Plan. Manag. 62, 306–326. https://doi.org/10.1080/09640568.2017.1415869.

Clapcott, J., 2015. National Rapid Habitat Assessment Protocol Development for Streams and Rivers. Cawthron Rep. No. 2649. Prep. Northl. Reg. Counc, Nelson, New Zealand.

Day, G., Fuller, R.A., Nichols, C., Dean, A.J., 2022. Characteristics of immersive citizen science experiences that drive conservation engagement. People Nat. https://doi.org/10.1002/pan3.10332.

Dijkstra, K.-D.B., Monaghan, M.T., Pauls, S.U., 2014. Freshwater biodiversity and aquatic insect diversification. Annu. Rev. Entomol. 59, 143–163. https://doi.org/10.1146/annurev-ento-011613-161958.

Dyer, F., Harrison, E., Gruber, B., Nichols, S., Tschierschke, A., O'Reilly, W., 2014. Waterwatch data quality: an opportunity to augment professionally collected data sets. In: Vietz, G., Rutherfurd, I.D., Hughes, R. (Eds.), 7th Australian Stream Management Conference - Full Paper, pp. 357–362 Townsville, Queensland.

Earth Watch, 2020. Freshwater Watch Local issues. URL https://freshwaterwatch.thewaterhub.org/our-data/local-issues (accessed 03/08/2022).

Edwards, P.M., 2016. The value of long-term stream invertebrate data collected by citizen scientists. PLoS One 11, e0153713. https://doi.org/10.1371/journal.pone.0153713.

Edwards, P.M., Shaloum, G., Bedell, D., 2018. A unique role for citizen science in ecological restoration: a case study in streams. Restor. Ecol. 26, 29–35. https://doi.org/10.1111/rec.12622.

EEA (European Environment Agency), 2018. European Waters Assessment of Status and Pressures 2018. URL https://doi.org/10.2800/303664 (accessed 03/08/2022) https://www.eea.europa.eu/publications/state-of-water.

Eitzel, M.V., Cappadonna, J.L., Santos-Lang, C., Duerr, R.E., Virapongse, A., West, S.E., Kyba, C.C.M., Bowser, A., Cooper, C.B., Sforzi, A., Metcalfe, A.N., Harris, E.S., Thiel, M., Haklay, M., Ponciano, L., Roche, J., Ceccaroni, L., Shilling, F.M., Dörler, D., Heigl, F., Kiessling, T., Davis, B.Y., Jiang, Q., 2017. Citizen science terminology matters: exploring key terms. Citiz. Sci. Theory Pract. 2, 1. https://doi.org/10.5334/cstp.96.

Engelhardt, W., Martin, P., Rehfeldt, K., Pfadenhauer, J., 2020. Was lebt in Tümpel, Bach und Weiher? Pflanzen und Tiere unserer Gewässer. 18th ed. Kosmos Naturführer. Franckh-Kosmos Verlag, Stuttgart.

EPA (Environmental Protection Agency), 2006. Water Framework Directive Monitoring Programme. URL (accessed 03/08/2022) https://www.epa.ie/pubs/reports/water/other/wfd/EPA_water_WFD_monitoring_programme_main_report.

European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council. Establishing a Framework for Community Action in the Field of Water Policy Brussels.

Fore, L.S., Paulsen, K., O'Laughlin, K., 2001. Assessing the performance of volunteers in monitoring streams. Freshw. Biol. 46, 109–123. https://doi.org/10.1111/j.1365-2427.2001.00640.x.

Friberg, N., Bonada, N., Bradley, D.C., Dunbar, M.J., Edwards, F.K., Grey, J., Hayes, R.B., Hildrew, A.G., Lamouroux, N., Trimmer, M., Woodward, G., 2011. Biomonitoring of Human Impacts in Freshwater Ecosystems, pp. 1–68. https://doi.org/10.1016/B978-0-12-374794-5.00001-8.

Fritz, S., See, L., Carlson, T., Haklay, M., Oliver, J.L., Fraisl, D., Mondardini, R., Brocklehurst, M., Shanley, L.A., Schade, S., Wehn, U., Abrate, T., Anstee, J., Arnold, S., Billot, M., Campbell, J., Espey, J., Gold, M., Hager, G., He, S., Hepburn, L., Hsu, A., Long, D., Masó, J., McCallum, I., Muniafu, M., Moorthy, I., Obersteiner, M., Parker, A.J., Weisspflug, M., West, S., 2019. Citizen science and the United Nations Sustainable Development Goals. Nat. Sustain. 2, 922–930. https://doi.org/10.1038/s41893-019-0390-3.

Hadj-Hammou, J., Loiselle, S., Ophof, D., Thornhill, I., 2017. Getting the full picture: assessing the complementarity of citizen science and agency monitoring data. PLoS One 12, e0188507. https://doi.org/10.1371/journal.pone.0188507.

Halbach, K., Möder, M., Schrader, S., Liebmann, L., Schäfer, R.B., Schneeweiss, A., Schreiner, V.C., Vormeier, P., Weisner, O., Liess, M., Reemtsma, T., 2021. Small streams–large concentrations? Pesticide monitoring in small agricultural streams in Germany during dry weather and rainfall. Water Res. 203, 117535. https://doi.org/10.1016/j.watres.2021.117535.

Haywood, B.K., Parrish, J.K., Dolliver, J., 2016. Place-based and data-rich citizen science as a precursor for conservation action: citizen science and conservation action. Conserv. Biol. 30, 476–486. https://doi.org/10.1111/cobi.12702.

Haywood, B.K., Parrish, J.K., He, Y., 2020. Shapeshifting attachment: exploring multidimensional people–place bonds in place-based citizen science. People Nat. 3, 51–65. https://doi.org/10.1002/pan3.10174.

Huddart, J.E.A., Thompson, M.S.A., Woodward, G., Brooks, S.J., 2016. Citizen science: from detecting pollution to evaluating ecological restoration. Wiley Interdiscip. Rev. Water 3, 287–300. https://doi.org/10.1002/wat2.1138.

Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., Jarvis, S., Lahoz-Monfort, J., Pagel, J., Pescott, O.L., Schmucki, R., Simmonds, E.G., O'Hara, R.B., 2020. Data integration for large-scale models of species distributions. Trends Ecol. Evol. 35, 56–67. https://doi.org/10.1016/j.tree.2019.08.006.

Jackson, M.C., Weyl, O.L.F., Altermatt, F., Durance, I., Friberg, N., Dumbrell, A.J., Piggott, J.J., Tiegs, S.D., Tockner, K., Krug, C.B., Leadley, P.W., Woodward, G., 2016. Recommendations for the next generation of global freshwater biological monitoring tools. Adv. Ecol. Res. 615–636 https://doi.org/10.1016/bs.aecr.2016.08.008.

Kelly, R., Fleming, A., Pecl, G.T., Richter, A., Bonn, A., 2019. Social license through citizen science: a tool for marine conservation. Ecol. Soc. 24, art16. https://doi.org/10.5751/ES-10704-240116.

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.

Kosmala, M., Wiggins, A., Swanson, A., Simmons, B., 2016. Assessing data quality in citizen science. Front. Ecol. Environ. 14, 551–560. https://doi.org/10.1002/fee.1436.

LAWA (Bund/Länder-Arbeitsgemeinschaft Wasser), 2019. LAWA-Verfahrensempfehlung zur Gewässerstrukturkartierung – Verfahren für kleine bis mittelgroße Fließgewässer. URL (accessed 03/08/2022) https://www.kulturbuch-verlag.de/?product=lawa-verfahrensempfehlung-zur-gewaesserstrukturkartierung.

LAWA, UBA, 2022. Software zur Bewertung des Makrozoobenthos gemäß Perlodes. Länderarbeitsgemeinschaft Wasser, Umweltbundesamt. URL https://gewaesser-bewertung.de/index.php?article_id=419&clang=0 (accessed 03/08/2022).

Lemm, J.U., Feld, C.K., Birk, S., 2019. Diagnosing the causes of river deterioration using stressor-specific metrics. Sci. Total Environ. 651, 1105–1113. https://doi.org/10.1016/j.scitotenv.2018.09.157.

Liebmann, L., Vormeier, P., Weisner, O., Liess, M., 2022. Balancing effort and benefit – how taxonomic and quantitative resolution influence the pesticide indicator system SPEARpesticides. Sci. Total Environ. 848, 157642. https://doi.org/10.1016/j.scitotenv.2022.157642.

Liess, M., Von Der Ohe, P.C., 2005. Analyzing the effects of pesticides on invertebrate communities in streams. Environ. Toxicol. Chem. 24, 954. https://doi.org/10.1897/03-652.1.

Liess, M., Schulz, R., Liess, M.H.-D., Rother, B., Kreuzig, R., 1999. Determination of insecticide contamination in agricultural headwater streams. Water Res. 33, 239–247. https://doi.org/10.1016/S0043-1354(98)00174-2.

Liess, M., Schäfer, R.B., Schriever, C.A., 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.

Liess, M., Liebmann, L., Vormeier, P., Weisner, O., Altenburger, R., Borchardt, D., Brack, W., Chatzinotas, A., Escher, B., Foit, K., Gunold, R., Henz, S., Hitzfeld, K.L., Schmitt-Jansen, M., Kamjunke, N., Kaske, O., Knillmann, S., Krauss, M., Küster, E., Link, M., Lück, M., Möder, M., Müller, A., Paschke, A., Schäfer, R.B., Schneeweiss, A., Schreiner, V.C., Schulze, T., Schüürmann, G., von Tümpling, W., Weitere, M., Wogram, J., Reemtsma, T., 2021. Pesticides are the dominant stressors for vulnerable insects in lowland streams. Water Res. 201, 117262. https://doi.org/10.1016/j.watres.2021.117262.

Lin, L.I.-K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255. https://doi.org/10.2307/2532051.

Maasri, A., Jähnig, S.C., Adamescu, M.C., Adrian, R., Baigun, C., Baird, D.J., Batista-Morales, A., Bonada, N., Brown, L.E., Cai, Q., Campos-Silva, J.V., Clausnitzer, V., Contreras-MacBeath, T., Cooke, S.J., Datry, T., Delacámara, G., De Meester, L., Dijkstra, K.B., Do, V.T., Domisch, S., Dudgeon, D., Erös, T., Freitag, H., Freyhof, J., Friedrich, J., Friedrichs-Manthey, M., Geist, J., Gessner, M.O., Goethals, P., Gollock, M., Gordon, C., Grossart, H., Gulemvuga, G., Gutiérrez-Fonseca, P.E., Haase, P., Hering, D., Hahn, H.J., Hawkins, C.P., He, F., Heino, J., Hermoso, V., Hogan, Z., Hölker, F., Jeschke, J.M., Jiang, M., Johnson, R.K., Kalinkat, G., Karimov, B.K., Kasangaki, A., Kimirei, I.A., Kohlmann, B., Kuemmerlen, M., Kuiper, J.J., Kupilas, B., Langhans, S.D., Lansdown, R., Leese, F., Magbanua, F.S., Matsuzaki, S., Monaghan, M.T., Mumladze, L., Muzon, J., Mvogo Ndongo, P.A., Nejstgaard, J.C., Nikitina, O., Ochs, C., Odume, O.N., Opperman, J.J., Patricio, H., Pauls, S.U., Raghavan, R., Ramírez, A., Rashni, B., Ross-Gillespie, V., Samways, M.J., Schäfer, R.B., Schmidt-Kloiber, A., Seehausen, O., Shah, D.N., Sharma, S., Sommerwerk, N., Stockwell, J.D., Suhling, F., Tachamo Shah, R.D., Tharme, R.E., Thorp, J.H., Tickner, D., Tockner, K., Tonkin, J.D., Valle, M., Vitule, J., Volk, M., Wang, D., Wolter, C., Worischka, S., 2022. A global agenda for advancing freshwater biodiversity research. Ecol. Lett. 25, 255–263. https://doi.org/10.1111/ele.13931.

Malaj, E., von der Ohe, P.C., Grote, M., Kühne, R., Mondy, C.P., Usseglio-Polatera, P., Brack, W., Schäfer, R.B., 2014. Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. Proc. Natl. Acad. Sci. 111, 9549–9554. https://doi.org/10.1073/pnas.1321082111.

McGoff, E., Dunn, F., Cachazo, L.M., Williams, P., Biggs, J., Nicolet, P., Ewald, N.C., 2017. Finding clean water habitats in urban landscapes: professional researcher vs citizen science approaches. Sci. Total Environ. 581–582, 105–116. https://doi.org/10.1016/j.scitotenv.2016.11.215.

McKinley, D.C., Miller-Rushing, A.J., Ballard, H.L., Bonney, R., Brown, H., Cook-Patton, S.C., Evans, D.M., French, R.A., Parrish, J.K., Phillips, T.B., Ryan, S.F., Shanley, L.A., Shirk, J.L., Stepenuck, K.F., Weltzin, J.F., Wiggins, A., Boyle, O.D., Briggs, R.D., Chapin, S.F., Hewitt, D.A., Preuss, P.W., Soukup, M.A., 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. Biol. Conserv. 208, 15–28. https://doi.org/10.1016/j.biocon.2016.05.015.

Meier, C., Haase, P., Rolauffs, P., Schindehütte, K., Schöll, F., Sundermann, A., Hering, D., 2006. Methodisches Handbuch Fließgewässerbewertung. Handbuch zur Untersuchung und Bewertung von Fließgewässern auf der Basis des Makrozoobenthos vor dem Hintergrund der EG-Wasserrahmenrichtl.

Metcalfe, A.N., Kennedy, T.A., Mendez, G.A., Muehlbauer, J.D., 2022. Applied citizen science in freshwater research. WIREs Water. https://doi.org/10.1002/wat2.1578.

Meyer, J.L., Strayer, D.L., Wallace, J.B., Eggert, S.L., Helfman, G.S., Leonard, N.E., 2007. The contribution of headwater streams to biodiversity in river networks. JAWRA J. Am. Water Resour. Assoc. 43, 86–103. https://doi.org/10.1111/j.1752-1688.2007.00008.x.

Moffett, E., Neale, M., 2015. Volunteer and professional macroinvertebrate monitoring provide concordant assessments of stream health. N.Z. J. Mar. Freshw. Res. 49, 366–375. https://doi.org/10.1080/00288330.2015.1018913.

Moolna, A., Duddy, M., Fitch, B., White, K., 2020. Citizen science and aquatic macroinvertebrates: public engagement for catchment-scale pollution vigilance. Écoscience 27, 303–317. https://doi.org/10.1080/11956860.2020.1812922.

Nerbonne, J.F., Nelson, K.C., 2004. Volunteer macroinvertebrate monitoring in the United States: resource mobilization and comparative state structures. Soc. Nat. Resour. 17, 817–839. https://doi.org/10.1080/08941920490493837.

Nerbonne, J.F., Vondracek, B., 2003. Volunteer macroinvertebrate monitoring: assessing training needs through examining error and bias in untrained volunteers. J. North Am. Benthol. Soc. 22, 152–163. https://doi.org/10.2307/1467984.

Newman, G., Chandler, M., Clyde, M., McGreavy, B., Haklay, M., Ballard, H., Gray, S., Scarpino, R., Hauptfeld, R., Mellor, D., Gallo, J., 2017. Leveraging the power of place in citizen science for effective conservation decision making. Biol. Conserv. 208, 55–64. https://doi.org/10.1016/j.biocon.2016.07.019.

Peter, M., Diekötter, T., Höffler, T., Kremer, K., 2021. Biodiversity citizen science: outcomes for the participating citizens. People Nat. 3, 294–311. https://doi.org/10.1002/pan3.10193.

Pieper, K.J., Martin, R., Tang, M., Walters, L., Parks, J., Roy, S., Devine, C., Edwards, M.A., 2018. Evaluating water lead levels during the flint water crisis. Environ. Sci. Technol. 52, 8124–8132. https://doi.org/10.1021/acs.est.8b00791.

Quinlivan, L., Chapman, D.V., Sullivan, T., 2020a. Applying citizen science to monitor for the sustainable development goal indicator 6.3.2: a review. Environ. Monit. Assess. 192, 218. https://doi.org/10.1007/s10661-020-8193-6.

Quinlivan, L., Chapman, D.V., Sullivan, T., 2020b. Validating citizen science monitoring of ambient water quality for the United Nations sustainable development goals. Sci. Total Environ. 699, 134255. https://doi.org/10.1016/j.scitotenv.2019.134255.

R Core Team, 2022. R: A Language and Environment for Statistical Computing.

Reid, D., Tippler, C., Evans, C., Kotevska, S., 2016. Comparison of stream macroinvertebrate monitoring data from citizen scientists and an aquatic ecologist. 8th Aust. Stream Manag. Conf. Leura.

Richter, A., Hauck, J., Feldmann, R., Kühn, E., Harpke, A., Hirneisen, N., Mahla, A., Settele, J., Bonn, A., 2018. The social fabric of citizen science—drivers for long-term engagement in the German butterfly monitoring scheme. J. Insect Conserv. 22, 731–743. https://doi.org/10.1007/s10841-018-0097-1.

Safford, H., Peters, C.A., 2018. Citizen science for dissolved oxygen monitoring: case studies from Georgia and Rhode Island. Environ. Eng. Sci. 35, 362–372. https://doi.org/10.1089/ees.2017.0218.

Scott, A.B., Frost, P.C., 2017. Monitoring water quality in Toronto's urban stormwater ponds: assessing participation rates and data quality of water sampling by citizen scientists in the FreshWater Watch. Sci. Total Environ. 592, 738–744. https://doi.org/10.1016/j.scitotenv.2017.01.201.

Shupe, S.M., 2017. High resolution stream water quality assessment in the Vancouver, British Columbia region: a citizen science study. Sci. Total Environ. 603–604, 745–759. https://doi.org/10.1016/j.scitotenv.2017.02.195.

Specht, H., Lewandowski, E., 2018. Biased assumptions and oversimplifications in evaluations of citizen science data quality. Bull. Ecol. Soc. Am. 99, 251–256. https://doi.org/10.1002/bes2.1388.

Stepenuck, K.F., Genskow, K.D., 2018. Characterizing the breadth and depth of volunteer water monitoring programs in the United States. Environ. Manag. 61, 46–57. https://doi.org/10.1007/s00267-017-0956-7.

Steven, M., Sergeant, E., 2022. R Package "epiR". URL https://cran.r-project.org/web/packages/epiR/epiR.pdf (accessed 03/08/2022).

Storey, R.G., Wright-Stow, A., 2017. Community-based monitoring of New Zealand stream macroinvertebrates: agreement between volunteer and professional assessments and performance of volunteer indices. N.Z. J. Mar. Freshw. Res. 51, 60–77. https://doi.org/10.1080/00288330.2016.1266674.

Storey, R.G., Wright-Stow, A., Kin, E., Davies-Colley, R.J., Stott, R., 2016. Volunteer stream monitoring: do the data quality and monitoring experience support increased community involvement in freshwater decision making? Ecol. Soc. 21, art32. https://doi.org/10.5751/ES-08934-210432.

Szöcs, E., Brinke, M., Karaoglan, B., Schäfer, R.B., 2017. Large scale risks from agricultural pesticides in small streams. Environ. Sci. Technol. 51, 7378–7385. https://doi.org/10.1021/acs.est.7b00933.

Turrini, T., Dörler, D., Richter, A., Heigl, F., Bonn, A., 2018. The threefold potential of environmental citizen science - generating knowledge, creating learning opportunities and enabling civic participation. Biol. Conserv. 225, 176–186. https://doi.org/10.1016/j.biocon.2018.03.024.

UBA, 2017, 2022. Ökologischer Zustand der Fließgewässer. URL https://www.umweltbundesamt.de/daten/wasser/fliessgewaesser/oekologischer-zustand-der-fliessgewaesser (accessed 03/08/2022).

UN, 2018. Sustainable Development Goal 6 Synthesis Report on Water and Sanitation. URL (accessed 03/08/2022) https://www.unwater.org/publications/sdg-6-synthesis-report-2018-on-water-and-sanitation/.

Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467, 555. https://doi.org/10.1038/nature09440.

Weeser, B., Stenfert Kroese, J., Jacobs, S.R., Njue, N., Kemboi, Z., Ran, A., Rufino, M.C., Breuer, L., 2018. Citizen science pioneers in Kenya – a crowdsourced approach for hydrological monitoring. Sci. Total Environ. 631–632, 1590–1599. https://doi.org/10.1016/j.scitotenv.2018.03.130.

Weisner, O., Arle, J., Liebmann, L., Link, M., Schäfer, R.B., Schneeweiss, A., Schreiner, V.C., Vormeier, P., Liess, M., 2022. Three reasons why the Water Framework Directive (WFD) fails to identify pesticide risks. Water Res. 208, 117848. https://doi.org/10.1016/j.watres.2021.117848.

Wick, A., Bänsch-Baltruschat, B., Keller, M., Scharmüller, A., Schäfer, R., Foit, K., Liess, M., Maaßen, S., Lischeid, G., 2019. Umsetzung des Nationalen Aktionsplans zur nachhaltigen Anwendung von Pestiziden Teil 2: Konzeption eines repräsentativen Monitorings zur Belastung von Kleingewässern in der Agrarlandschaft. Umweltbundesamt Texte 08/2019. URL (accessed 03/08/2022) https://www.umweltbundesamt.de/publikationen/umsetzung-des-nationalen-aktionsplans-zur-0.

Wolfram, J., Stehle, S., Bub, S., Petschick, L.L., Schulz, R., 2021. Water quality and ecological risks in European surface waters – monitoring improves while water quality decreases. Environ. Int. 152, 106479. https://doi.org/10.1016/j.envint.2021.106479.

Yardi, K.D., Bharucha, E., Girade, S., 2019. Post-restoration monitoring of water quality and avifaunal diversity of Pashan Lake, Pune, India using a citizen science approach. Freshw. Sci. 38, 332–341. https://doi.org/10.1086/703440.

Zheng, S., Li, H., Fang, T., Bo, G., Yuan, D., Ma, J., 2022. Towards citizen science. On-site detection of nitrite and ammonium using a smartphone and social media software. Sci. Total Environ. 815, 152613. https://doi.org/10.1016/j.scitotenv.2021.152613.