Check for updates

Received: 18 January 2024 | Revised: 26 May 2024 | Accepted: 30 May 2024

DOI: 10.1111/ele.14463

SYNTHESIS

ECOLOGY LETTERS cnrs WILEY

Studying interactions among anthropogenic stressors in freshwater ecosystems: A systematic review of 2396 multiple-stressor experiments

James A. Orr[1,2] | Samuel J. Macaulay[1] | Adriana Mordente[1] | Benjamin Burgess[3] | Dania Albini[1] | Julia G. Hunn[4] | Katherin Restrepo-Sulez[1] | Ramesh Wilson[1] | Anne Schechner[5,6] | Aoife M. Robertson[7] | Bethany Lee[1] | Blake R. Stuparyk[8] | Delezia Singh[9] | Isobel O'Loughlin[1] | Jeremy J. Piggott[7] | Jiangqiu Zhu[10] | Khuong V. Dinh[11] | Louise C. Archer[12] | Marcin Penk[7,13] | Minh Thi Thuy Vu[11] | Noël P. D. Juvigny-Khenafou[14,15] | Peiyu Zhang[10] | Philip Sanders[1] | Ralf B. Schäfer[16,17] | Rolf D. Vinebrooke[8] | Sabine Hilt[5] | Thomas Reed[18] | Michelle C. Jackson[1]

Correspondence
James A. Orr and Michelle C. Jackson, Department of Biology, 11a Mansfield Road Oxford, OX1 3SZ, UK.
Email: james.orr@uq.edu.au; michelle.jackson@biology.ox.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/V001396/1

Editor: Akira S Mori

Abstract

Understanding the interactions among anthropogenic stressors is critical for effective conservation and management of ecosystems. Freshwater scientists have invested considerable resources in conducting factorial experiments to disentangle stressor interactions by testing their individual and combined effects. However, the diversity of stressors and systems studied has hindered previous syntheses of this body of research. To overcome this challenge, we used a novel machine learning framework to identify relevant studies from over 235,000 publications. Our synthesis resulted in a new dataset of 2396 multiple-stressor experiments in freshwater systems. By summarizing the methods used in these studies, quantifying trends in the popularity of the investigated stressors, and performing co-occurrence analysis, we produce the most comprehensive overview of this diverse field of research to date. We provide both a taxonomy grouping the 909 investigated stressors into 31 classes and an open-source and interactive version of the dataset (https://jamesaorr.shinyapps.io/freshwater-multiple-stressors/). Inspired by our results, we provide a framework to help clarify whether statistical interactions detected by factorial experiments align with stressor interactions of interest, and we outline general guidelines for the design of multiple-stressor experiments relevant to any system. We conclude by highlighting the research directions required to better understand freshwater ecosystems facing multiple stressors.

KEYWORDS
antagonism, ecology, ecotoxicology, global change biology, research synthesis, synergism

For affiliations refer to page 15.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2024 The Author(s). *Ecology Letters* published by John Wiley & Sons Ltd.

INTRODUCTION

Despite being home to a disproportionately high diversity of life and providing essential ecosystem services, freshwater ecosystems are being degraded by anthropogenic activities (Almond et al., 2020; Reid et al., 2019). Lakes, rivers, ponds and wetlands are directly impacted by stressors such as overharvesting and hydrological modifications, can accumulate stressors like fine sediment and toxic chemicals arising from activities in the surrounding landscape, and are increasingly being impacted by climate change (Reid et al., 2019; Sala et al., 2000). These diverse physical, chemical and biological stressors co-occur and therefore have the potential to interact, which complicates ecosystem management and environmental risk assessment (Côté et al., 2016; Folt et al., 1999). For instance, global warming can enhance the spread and impact of many non-native species (Rahel & Olden, 2008), environmental stressors can alter the sensitivity of an organism to chemical stressors through physiological trade-offs (Moe et al., 2013), and toxic pollutants can become attached to suspended sediment or plastic particles, allowing them to spread much further from their source (Naasz et al., 2018; Viers et al., 2009). A mechanistic understanding of such stressor interactions is crucial for accurate predictions of global change impacts and for effective management of freshwater ecosystems (Spears et al., 2021).

In response to increasing concerns over the cumulative ecological impacts of multiple novel or extreme anthropogenic environmental changes, there has been a proliferation of research on the interactions among stressors (Orr et al., 2020). Indeed, studies that explore these antagonistic (reducing combined effects) and synergistic (increasing combined effects) interactions have become very common in the freshwater sciences (Jackson et al., 2016; Ormerod et al., 2010; Reid et al., 2019). While observational research (e.g., Birk et al., 2020; Outhwaite et al., 2022) and advances in ecological theory (e.g., Beauchesne et al., 2021; De Laender, 2018; Vinebrooke et al., 2004) have been critical for generating knowledge on stressor interactions, the primary tool used by multiple-stressor researchers has, without doubt, been the factorial experiment. However, factorial experiments, in which stressors are manipulated individually and in combination, involve trade-offs between ecological realism, replication, temporal scale, and spatial scale (Schindler, 1998; Thomas & Ranjan, 2024), so they are conducted in very different ways, resulting in considerable context-dependency of results (Orr et al., 2020; Simmons et al., 2021). The observed interactive effects of two stressors can be species- or system-specific or can depend on factors including the exposure duration, the stressors' intensities, the response variable being measured, and even the value of other environmental factors that can vary across space and time (Dinh et al., 2023; Jackson et al., 2021; Kefford et al., 2023; Turschwell et al., 2022). The first general synthesis of multiple-stressor experiments in freshwater systems was based on 88 publications and was conducted in 2014/15 (Jackson et al., 2016). The >700 citations accumulated by that original review (as of December 2023) and the subsequent, more targeted syntheses of multiple-stressor studies on freshwater fish (Lange et al., 2018, $n=28$) and on freshwater multispecies assemblages (He et al., 2023, $n=167$) suggest that research efforts in the field are intensifying. Comprehensive reviews across all freshwater types are becoming increasingly difficult to perform due to the diversity of stressors, systems, and response combinations, but they are required to effectively guide future research and freshwater management.

Here we used a novel machine learning framework to perform the largest systematic review in freshwater multiple-stressor research, and one of the largest in global change biology in general, to provide a synthesis of how interactions between stressors are experimentally studied (Figure 1a). We started with an extremely broad initial *Web of Science* search that returned over one quarter of a million publications. By using a machine learning framework that makes abstract screening far more efficient, we identified 4085 potentially relevant publications, from which 2396 multiple-stressor experiments in freshwater ecosystems were found during manual full-text screening. There are three main practical outputs from this research that will be useful to ecologists and evolutionary biologists interested in the effects of anthropogenic stressors on freshwater biodiversity. Firstly, we provide an analysis of this new dataset to give an overview of how these experiments were conducted and to describe trends in the popularity and co-occurrence of the investigated stressors. Secondly, we outline a taxonomy of anthropogenic stressors in freshwater systems, where the 909 specific stressors tested across the experiments are grouped into 31 classes within six broader groups. Thirdly, we provide this new dataset as an open-source, interactive web application (https://jamesaorr.shinyapps.io/freshwater-multiple-stressors/) that can be used by students, scientists and stakeholders to rapidly find publications relevant to their work, to identify knowledge gaps, and even to contribute additional knowledge to the dataset. More broadly, however, we use the results of our synthesis and recent discussions in the literature to address two general questions relevant to all researchers interested in multiple stressors (Figure 1b). We first propose a framework to better understand whether statistical interactions observed by experiments reflect stressor interactions of interest. Based on the confounding variables blurring the connection between stressor interactions and statistical interactions, and by comparing the diverse experimental designs employed by the studies in our systematic review, we then ask if there is an 'ideal' multiple-stressor experiment and we provide

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

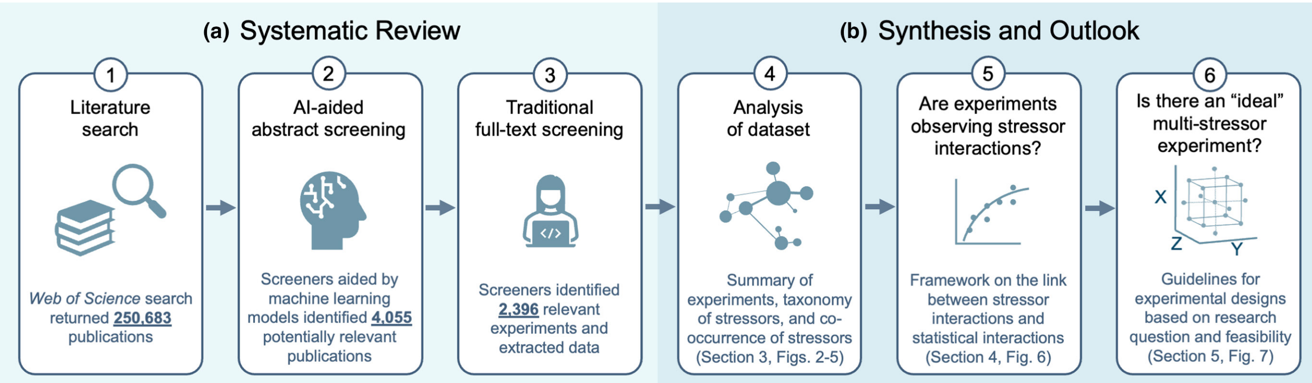

FIGURE 1 Flow chart outlining the approach of our synthesis. (a) Initially, we performed a systematic review where a very broad literature search (1) was used in combination with AI-aided abstract screening (2) and traditional full-text screening (3) to compile a novel dataset of 2396 freshwater multiple-stressor experiments. (b) This dataset was then used to perform a quantitative synthesis of experimental work on multiple stressors in freshwater systems (4). The results of this analysis, and group discussions during the screening process, led us to ask if multiple-stressor experiments are directly observing stressor interactions or if these stressor interactions of interest are being confounded by other factors (5). Based on the challenges of observing stressor interactions and the diversity of experimental designs within our compiled dataset, we then asked if there was an 'ideal' multiple-stressor experiment, and we produced practical guidelines for future experimental work on multiple stressors (6).

practical guidelines for their design. Finally, we conclude by recommending directions for future research on the cumulative effects of anthropogenic stressors that are necessary to optimize the management and protection of freshwater ecosystems.

METHODS

The diversity of stressors and systems being studied in multiple-stressor research means that comprehensive reviews require very broad search queries that return a daunting number of records. To overcome this challenge, which has greatly limited previous syntheses of the field, we employed a novel machine learning framework that uses active learning to efficiently screen vast quantities of abstracts. We followed the PRISMA-EcoEvo guidelines (O'Dea et al., 2021) and provided extensive supporting information with detailed descriptions of each step from study identification to abstract screening to full-text screening and finally to data analysis (see Figure S1 for a decision tree and see Figure S2 for a PRISMA-style flowchart). Our goal was to collect a large and unbiased (based on the focus of the experiments) sample of published multiple-stressor experiments conducted in freshwater ecosystems. We considered an experiment relevant if it: (i) included multiple anthropogenic stressors; (ii) quantified biological responses in a freshwater system; and (iii) quantified both individual and combined effects of stressors.

We first performed a very broad search on *Web of Science* on 9 January 2022, with three groups of terms informed by recent reviews of multiple-stressor research (Orr et al., 2020) and of freshwater biology (Lodi et al., 2021) that aligned with our three core inclusion criteria (Figure S1). This search returned 250,683 records, and after filtering out review articles, publications that were not written in English (a requirement of our screening approach, which unfortunately introduces a bias towards English-language publications (Amano et al., 2023)), and publications without abstracts in their *Web of Science* record (which were screened manually), 236,075 potentially relevant records progressed to abstract screening. We used an open-source machine learning tool called ASReview (version 0.19), which allows for the rapid, transparent and reproducible screening of large amounts of bibliometric text (Van De Schoot et al., 2021). In practice, an abstract screener uploads their dataset of bibliometric records to the software and provides training data, with some records marked as relevant and others as irrelevant based on prior knowledge. The active learning model then uses a machine learning classifier (e.g., Naïve Bayes) to give relevancy scores to different words (and combinations of words) based on the training data. The model then searches through the entire dataset and identifies the publication it considers to be most relevant. The abstract screener then reads the abstract of this paper and decides whether to mark it as relevant or irrelevant. The entire model is then updated based on the decision of the abstract screener, as this is an active learning process with interaction between model and human, and the model recommends the paper that it now considers most relevant. If the model works well, the abstract screener will quickly find most of the relevant papers while only having manually screened a fraction of the entire dataset; plotting the number of papers reviewed against the number of relevant papers found will return a saturating curve.

We split the abstract screening process among eight researchers who simultaneously worked with eight different active learning models over 6 months to screen ca. 30,000 abstracts. These eight active learning models were initialized using the same training data and the same

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Naïve Bayes machine learning classifier. The training data were 20 relevant records randomly chosen from Jackson et al. (2016) and 30 randomly chosen records from the *Web of Science* search, which we screened and labelled as either relevant or irrelevant. One of the randomly chosen *Web of Science* records was relevant, so the final training dataset had 21 relevant records and 29 irrelevant records. The relevant records in the training dataset contained a diverse range of studies; all three types of experimental systems were represented (5 lab studies, 12 mesocosm studies and 4 field studies), responses from all levels of biological organization were quantified, and 29 different stressor identities from 12 different stressor classes were tested. Given that we were using active learning models that were updated whenever a classification was made by a human, we did not expect the choice of training data to strongly influence the screening process. However, to test the intuitive idea that a biased training dataset would lead to less efficient screening, we performed a simulation study (described in detail in Section 3.6 in the supporting information) where we compared models initialized with (i) our 'diverse' training set, (ii) a 'biased' training set made up of lab studies focusing on insecticides, and (iii) effectively no training data (just one relevant record and one irrelevant record). This simulation study showed that the choice of training data had no effect on the screening process (Figure S5) and demonstrates the robustness of the AI-aided screening framework.

Each of the eight subsets of abstracts also contained the same set of ca. 1000 abstracts that were used as benchmark data to assess the consistency and accuracy of the active learning models and screeners (Figure S4). Our predetermined stopping criterion, which was established following discussions with the authors of ASReview, was that we would screen at least 5% of the records (ca. 1500 abstracts each) and then continue screening until ASReview suggested 50 irrelevant papers in a row. After screening 5% of records with ASReview, no screener had identified 50 irrelevant papers in a row, so we continued to screen an additional 1% of the records at a time until the stopping criterion had been met. Once 8% of the records had been manually screened (a total of 19,421 abstracts: 2427.6 +/− 13 abstracts per screener), four of the eight screeners had identified at least 50 irrelevant papers in a row (a median of 50 with a range between 28 and 65), so as per our pre-determined stopping criterion, we ended the abstract screening process. The curve of the total number of papers screened against the total number of relevant papers found was clearly saturating, and the rate of identifying relevant records had fallen below what would be expected by chance alone (Figure S3). Indeed, based on our simulation study that tested the effect of the training data, a very high percentage of relevant records (>95%) can be found after screening just 8% of the relevant records in such a large dataset (Figure S5). After combining all results and removing duplicates, 14,158 records had been manually excluded by the screeners based on the decision tree (Figure S1), 217,074 records had been excluded for not being suggested by the active learning models before the stopping criterion was met, and 4085 records progressed to full-text screening (Figure S2).

The full-text screening, which was split among 25 researchers, resulted in a total of 1768 publications being excluded, with 950 publications not meeting the first criterion (i.e., 'not anthropogenic'), 342 publications not meeting the second criterion (i.e., 'not freshwater'), 415 publications not meeting the third criterion (i.e., 'not interactive'), and 61 publications not meeting multiple criteria. The remaining 2302 publications were included in our systematic review. In the dataset, each row is a unique experiment; therefore, publications were split across multiple rows if they described multiple distinct experiments. Similarly, publications were merged into the same row if they described the same experiment. From the 2302 relevant publications, there were 2396 distinct experiments. For each experiment, a wide range of data was collected from the text and figures in the publications that summarized bibliometric information, experimental design information, response information and stressor information (Table S1). During the full-text screening, we collectively built a 'stressor taxonomy' (Table S2) where we grouped the stressors tested in the experiments based on their intrinsic traits. The taxonomy was inspired by recent work from terrestrial systems (Rillig et al., 2021) and groups stressors based on whether they are physical, chemical or biological in nature. Within these broad groups, there are multiple 'classes' of stressors, and within those classes, there are specific stressor 'identities' (which describe the specific treatments in the experiments). For example, a stressor treatment that increases temperature by a constant amount would be a 'warming' stressor identity, within the 'temperature' stressor class, within the 'physical' stressor nature. Overall, there were 31 classes of stressors, comprising a total of 909 stressor identities. The taxonomy is by no means definitive and is only intended to help us record the information in these experiments and to illustrate general trends in the literature. Analyses of the final dataset, including the quantification of publication rates, the relative popularity of stressors, and the co-occurrence of stressors, were performed in R (version 4.3.1) with the packages listed in Table S3. All data and code used for these analyses are available at https://zenodo.org/doi/10.5281/zenodo.11100467.

RESULTS

Overview of experiments

Although the first freshwater multiple-stressor experiment found by our systematic review was published almost 60 years ago (Yankow, 1965), it was not until the 1990s that the number of these published experiments

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

began to increase exponentially (Figure 2a). In fact, the growth rate of publications on freshwater multiple-stressor experiments has outstripped the growth rate of academic publishing in general since the 1990s (Figure 2a inset). There were experiments from 73 countries across all seven continents in our dataset, but over 50% of the experiments were conducted in only four countries: China, the USA, Canada and Spain (Figure 2b). A total of 79% of the experiments (n=1896) were conducted in laboratories, but 500 experiments were performed in more natural experimental settings, either in outdoor mesocosms (n=387, ca. 16%) or in experimental field systems (n=113, ca. 5%). Most of these mesocosm and field experiments were conducted in either lentic (n=317) or lotic (n=145) systems but there are also 36 wetland experiments and two tank bromeliad experiments. Forty percent of the experiments were performed in less than 1 week (acute 24-hour and 48-hour standardized tests on model organisms were very common), while 36 experiments (only 1.5%), which were mostly mesocosm or field experiments, lasted for over 1 year (Figure S6). Experiments focusing on physiological (e.g., gene expression, tissue health), individual (e.g., behaviour, feeding rate), and population (e.g., reproduction, abundance) level responses were more common than experiments focusing on community (e.g., diversity, stability) or ecosystem (e.g., total biomass, decomposition) level responses. Indeed, ca. 77% of experiments quantified responses only at the population level or below, while only ca. 12% of experiments focused exclusively on the community or ecosystem levels. Less than 2% of studies (n=37) quantified responses at four or five levels of biological organization (Figure 2c). These results were mirrored by most studies (77%) having focused on a single species. Indeed, more single-species experiments using fish have been conducted than multi-species experiments (Figure 2f). Most experiments (73%) simulated only two stressors, but 21 experiments studied ten or more stressors (Figure 2d). Half of the 424 experiments simulating three stressors were fully factorial, 20% of the 122 experiments simulating four stressors were fully factorial, and three out of the 52 experiments simulating five stressors were fully factorial. Half of stressor treatments were applied as presence/absence treatments, but many experiments, particularly lab studies on chemicals, followed gradient designs, with 22% of stressor treatments being applied at five or more levels (Figure 2e). Finally, in relation to the assessment of the temporal dynamics of stressors and their effects, 50% of experiments quantified responses at multiple timepoints, while only 11% of studies applied stressors sequentially rather than simultaneously, and even less (7%) quantified the recovery following the removal of a stressor. Many of these results will be intuitive to researchers in the field, but this is the first time that these characterizations of multiple-stressor research (e.g., increasing popularity, dominance of presence-absence designs, focus on short-term experiments and responses from lower levels of biological organization) have been robustly quantified.

Diversity of stressors

There were a total of 6118 stressor treatments across the 2396 experiments. These stressor treatments were classified in our dataset into 909 stressor identities within 31 stressor classes based on our taxonomy (Table S2). Previous taxonomies of anthropogenic stressors have either been specific to certain types of stressors (Eisenberg & McKone, 1998) or have not contained information about the popularity of stressors (Rillig et al., 2021). 423 stressor identities were manipulated only once, while 215 stressor identities were manipulated at least five times (Figure 3). The five most common stressor identities—warming (n=441), copper (n=260), cadmium (n=248), nutrient enrichment (n=218) and zinc (n=152)—accounted for over 20% of all stressor treatments. Chemical stressors accounted for almost 75% of all stressor treatments (n=4508), while only 17% and 3% of stressor treatments were purely physical or biological stressors, respectively. Moreover, chemical stressors accounted for even more of the overall diversity of stressor identities in the taxonomy, with 82% of all stressor identities belonging to chemical classes. Although physical stressors made up a greater proportion of stressor treatments compared to biological stressors, they had a lower diversity in the taxonomy (44 stressor identities are physical and 75 stressor identities are biological). Up until and during the 1990s, almost all stressor treatments were chemical stressors, but during the 2000s, 20% of stressor treatments were physical, and around 5% were either biological or a mixture of different stressor natures (Figure 4a). These proportions remained quite consistent up until the 2020s with the notable exception that since 2010, the proportion of stressor treatments that were chemical gradually declined to about 60% and the proportion of stressor treatments that were a mixture of chemical and physical stressors (i.e., nanoparticles and microplastics) increased from 0% to ca. 10%. However, there was much more variability in the relative popularity of stressor treatments when examining trends at the level of stressor classes (Figure 4a, Figure S7). Some physical stressors like ultraviolet radiation, habitat alteration, and visible light have greatly decreased in relative popularity, while the proportion of temperature stressors rose from ca. 5% to ca. 15%. Metals accounted for up to 60% of stressor treatments in the 1990s but have gradually dropped to about 15% of stressor treatments. Aside from temperature, the only other stressor classes that have seen major increases in relative popularity are microplastics and nanoparticles, which have gone from 0% of stressor treatments before 2010 to ca. 5% each in the early 2020s. The interactive dataset (https://jamesaorr.shinyapps.io/freshwater-multiple-stressors/) can be used to explore how these relative frequencies of stressors also vary across experimental system types and across habitat types.

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

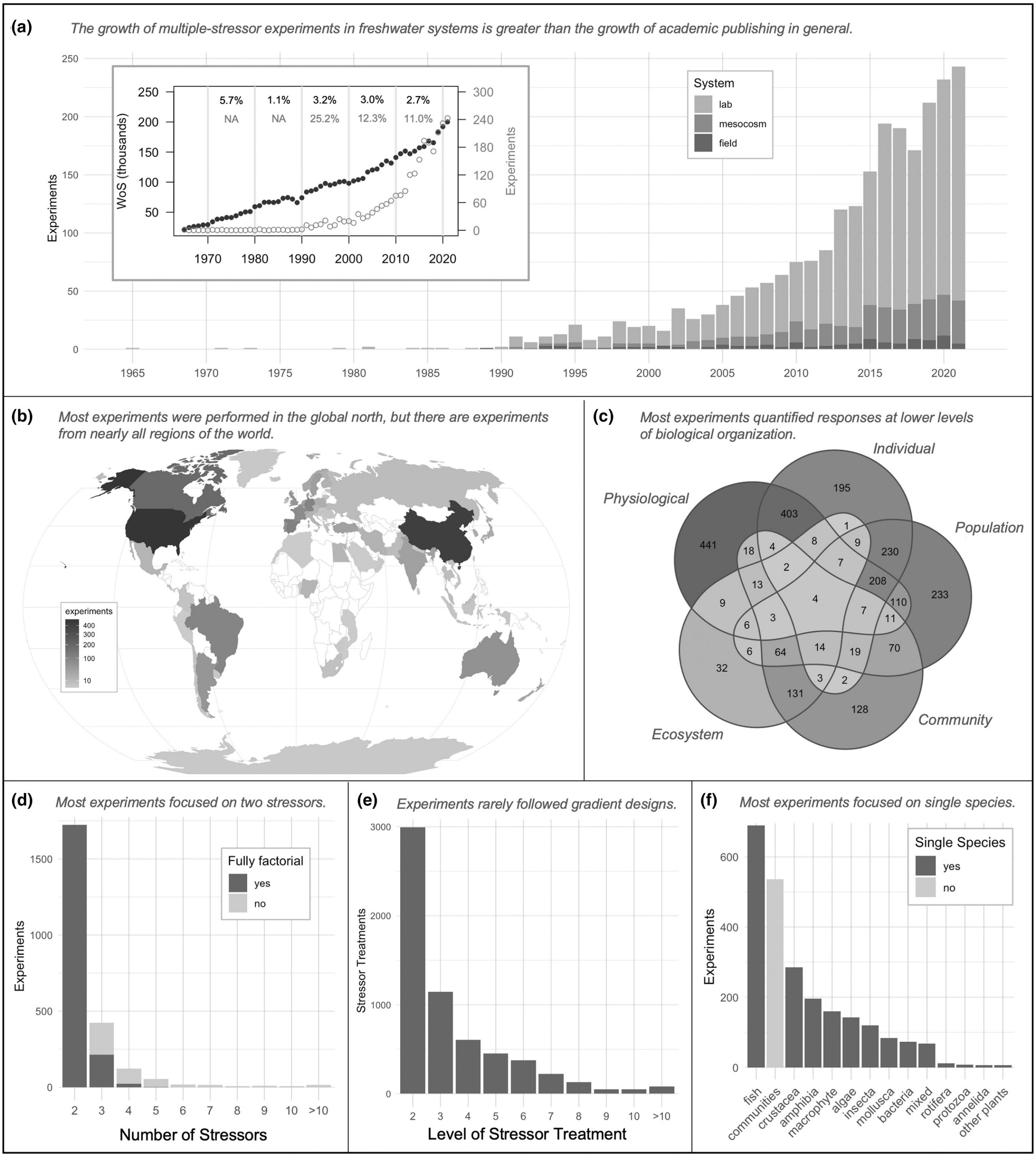

FIGURE 2 Overview of experiments in the dataset. (a) The number of experiments per year in this systematic review is shaded by whether they were lab, mesocosm or field experiments. The inset shows that the growth rate of published freshwater multiple-stressor experiments (pale grey, empty circles) has been far greater than the growth rate of publications in the 'Biology' category within the *Web of Science* database (dark grey, solid circles) since the 1990s. Growth rate values are compound annual growth rates for each decade (undefined when the number of publications in a given year was 0, as was the case in 1970 and 1980). Note the different y-axis scales for the two groups. (b) World map using the Winkel Tripel projection with countries shaded based on the number of experiments conducted there. (c) A five-way Venn diagram showing the number of experiments that quantified responses at each level of the biological hierarchy (physiological, individual, population, community and ecosystem). (d) The frequency of experiments that manipulated different numbers of stressors (with experiments that manipulated more than 10 stressors pooled). Fully factorial and partially factorial studies are in dark and light grey, respectively (e) The frequency of stressor treatments tested at different levels (with stressor treatments with more than 10 levels pooled). (f) The frequency of experiments that focused on single species within specific taxonomic groups (dark grey) or on multi-species communities (pale grey).

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

	Class	Identity
PHYSICAL	habitat alteration (85)	agricultural land-use (5), habitat complexity (6), sediment type (12), sedimentation (46)
	hydrology (103)	drought (20), flow regime (11), flow velocity (31), water level (30)
	radiation (11)	gamma radiation (9)
	temperature (581)	cooling (9), heat wave (20), temperature range (86), temperature variability (22), warming (441)
	UV light (143)	UV light (68), UVA (15), UVB (58)
	visible light (113)	artificial light (6), light intensity (49), shading (51)
	water clarity (21)	brownification (13)
	microplastics (71)	microplastic beads (18), microspheres (15), polystyrene (19)
	nanoparticles (206)	carbon nanotubes (12), cerium nanoparticles (5), copper nanoparticles (17), fullerene (6), graphene oxide (9), nanoplastics (5), polystyrene nanoparticles (8), silver nanoparticles (21), titanium nanoparticles (59), zinc nanoparticles (33)
CHEMICAL	acidity (102)	dissolved humic materials (5), humic acid (15), low pH (50), pH range (31)
	alkalinity (43)	bicarbonate (5), calcium (24), hardness (11)
	antibiotic (186)	antibiotic mixture (7), chloramphenicol (5), ciprofloxacin (12), clarithromycin (6), enrofloxacin (5), erythromycin (13), florfenicol (8), norfloxacin (6), ofloxacin (5), roxithromycin (5), sulfadiazine (5), sulfamethoxazole (21), tetracycline (10), triclocarban (6), trimethoprim (9)
	carbon dioxide (39)	increased carbon dioxide (35)
	fungicide (150)	azoxystrobin (7), chlorothalonil (11), dichlofluanid (5), difenoconazole (5), fungicide mixture (5), ketoconazole (6), prochloraz (12), propiconazole (6), pyrimethanil (6), sea nine 211 (5), tebuconazole (9)
	general biocide (97)	benzalkonium chloride (6), biocide mixture (14), irgarol (6), pentachlorophenol (7), tributyltin (10), triclosan (38)
	herbicide (416)	2,4-d (22), alachlor (6), atrazine (84), butachlor (5), diquat (6), diuron (26), glyphosate (80), herbicide mixture (13), isoproturon (5), linuron (7), metolachlor (9), paraquat (6), prometryn (5), s-metolachlor (7), simazine (5), terbuthylazine (14)
	insecticide (494)	alpha-cypermethrin (8), azinphos-methyl (6), bifenthrin (9), carbaryl (38), carbofuran (9), chlorpyrifos (105), clothianidin (5), cypermethrin (22), deltamethrin (10), diazinon (25), endosulfan (15), esfenvalerate (14), fipronil (10), imidacloprid (33), insecticide mixture (5), lambda-cyhalothrin (10), lindane (5), malathion (35), methyl parathion (7), permethrin (10), thiacloprid (9), thiamethoxam (11)
	metals (1082)	aluminium (32), arsenate (6), arsenic (28), arsenite (7), cadmium (248), cadmium chloride (9), chromium (30), cobalt (8), copper (260), copper sulphate (6), iron (18), lead (84), magnesium (5), manganese (15), mercury (36), metal mixture (29), methylmercury (5), nickel (56), silver (11), zinc (152)
	nutrients (714)	ammonia (28), ammonium (13), ammonium nitrate (12), carbon (6), dissolved organic carbon (27), dissolved organic matter (16), natural organic matter (5), nitrate (48), nitrite (13), nitrogen (103), nutrient enrichment (218), nutrient limitation (7), nutrient ratio (17), phosphate (25), phosphorus (129), selenium (13), silicon (9), sodium nitrate (5)
	other chemicals (582)	2,4-dichlorophenol (5), 4-nonylphenol (8), 4-tert-octylphenol (9), benzo[a]pyrene (19), beta-naphthoflavone (5), bisphenol a (41), caffeine (6), dbp (6), decabromodiphenyl ether (5), fluoranthene (10), iodopropynyl butylcarbamate (5), naphthalene (5), nonylphenol (21), octocrylene (6), other chemical mixture (15), oxybenzone (7), pah mixture (5), perfluorooctanesulfonic acid (17), perfluorooctanoic acid (5), phenanthrene (13), polychlorinated biphenyls (6), pyrene (7), selenomethionine (6), sodium dodecyl benzene sulfonate (5), tetrabromobisphenol-a (6)
	oxygen (85)	low oxygen concentration (79)
	pharmaceuticals (400)	17 beta trenbolone (5), bezafibrate (6), carbamazepine (14), citalopram (7), clofibric acid (6), diclofenac (21), diphenhydramine (7), estradiol (41), estrone (13), ethinylestradiol (44), fluoxetine (32), ibuprofen (13), ketoprofen (6), levonorgestrel (8), pharmaceutical mixture (9), propranolol (7), sertraline (10), venlafaxine (8)
	salinity (117)	increased salinity (74), salinity variation (8), sodium chloride (13)
	cyanotoxin (36)	microcystin-lr (24)
BIOLOGICAL	biological alterations (70)	community composition (7), cyanobacteria bloom (22), leaf litter type (10), resource availability (7), species abundance (5), species loss (12)
	disease (32)	*Batrachochytrium dendrobatidis* (18)
	non-native species (93)	blue gum (*Eucalyptus globulus*) (7), golden mussel (*Limnoperna fortunei*) (5), mosquitofish (*Gambusia holbrooki*) (5), rainbow trout (*Oncorhynchus mykiss*) (5), western mosquitofish (*Gambusia affinis*) (8), zebra mussels (*Dreissena polymorpha*) (8)
	composite stressors (35)	agricultural effluent (8), industrial effluent (5), wastewater effluent (11), water grade (9)

FIGURE 3 Taxonomy of stressors with only the stressor identities that were manipulated in at least five experiments ($n=215$). The number in parentheses after each class and identity shows their occurrences across the entire dataset. The occurrence values for each class also include identities from that class that had less than five occurrences, which can be found in the full stressor taxonomy in Table S2. Microplastics and nanoparticles are considered a mixture of physical and chemical stressors, while cyanotoxins are considered a mixture of chemical and biological stressors. Stressors that are a combination of all three natures are classified as 'composite' stressors.

Co-occurrence of investigated stressors

Co-occurrence analysis revealed modular patterns at the levels of stressor class and stressor identity. There were high frequencies of co-occurrence between stressors within the same class, particularly for chemical stressors (Figure 4b). Although there was widespread and relatively strong co-occurrence within and between chemical stressor classes, and there was a particularly strong increase in the co-occurrences between temperature and all other classes in the past decade (Figure S8), there were many pairs of stressor classes whose interactive effects were not tested by the experiments in our dataset. There were low co-occurrences between biological stressor classes, between physical (except temperature) and physical–chemical stressor

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

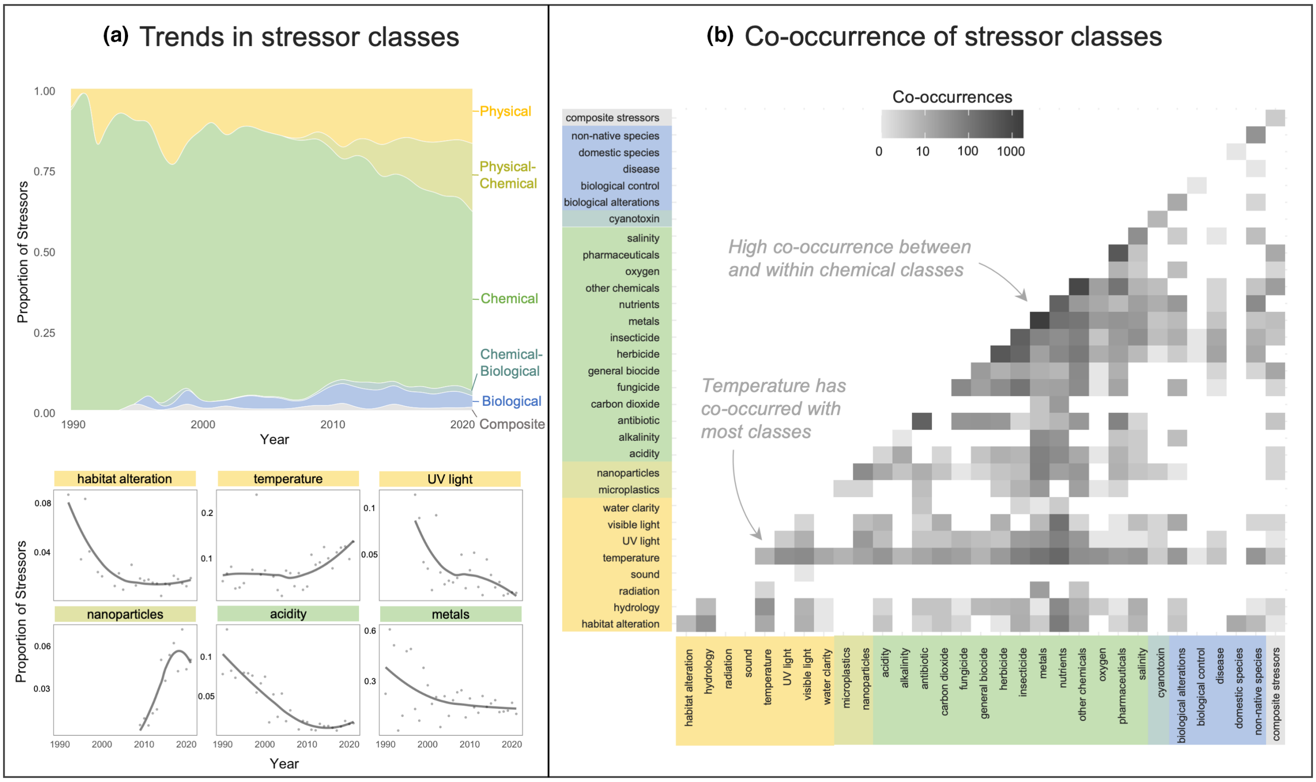

FIGURE 4 (a) Trends in the proportions of stressor nature categories (physical, physical–chemical, chemical, chemical-biological, biological, and composite stressors) as a timeseries proportion plot and trends in the proportions of six stressor classes (habitat alteration, temperature, UV light, nanoparticles, acidity and metals) as timeseries scatter plots with Loess curves that have an α smoothing parameter set to 1. (b) Co-occurrence matrix of stressor classes ordered by stressor nature (and then ordered alphabetically within a nature category) as in the stressor taxonomy. Only one half of the matrix is required because co-occurrences are undirected. The cells along the diagonal represent the number of times stressors from the same class co-occurred.

classes, and between biological stressor classes and all the other stressor classes (except some of the pesticide classes). Visualizing the co-occurrence patterns between actual stressor identities as a force-directed network produced clusters of stressor identities that revealed compartmentalization between the different research disciplines studying multiple-stressor interactions in freshwaters (Figure 5). The stressor identities within the metal class formed one distinct cluster, while stressor identities in the physical classes and the nutrient, salinity, and oxygen classes formed another. Pesticide stressor identities (fungicides, insecticides, and herbicides) grouped together near the cluster with physical stressors while biological stressors had a broader distribution in the network. Microplastics, nanoparticles, and stressor identities in the acidity class were found between the two main clusters, while pharmaceuticals and other synthetic chemicals formed a loose collection of nodes away from the other clusters. Stressor identities more associated with ecological research (particularly mesocosm and field studies) were found in the top-left of the network, while the right and lower parts of the network contained chemical stressors more associated with ecotoxicological research. The divisions between research disciplines are well documented, but it seems that these divisions also influence which combinations of stressors are tested.

ARE MULTIPLE-STRESSOR EXPERIMENTS OBSERVING STRESSOR INTERACTIONS?

The 2396 experiments in our dataset were all designed to test if some combination of anthropogenic stressors interact. Understanding if one stressor influences the intensity or the effect of another stressor (i.e., a stressor interaction) improves predictions of combined effects and can guide environmental risk assessments and management strategies. However, stressor interactions were not directly observed by these experiments; rather, interactive effects from statistical models (traditionally ANOVA) that were fitted to the response data were used to infer if stressors interacted. Whether the statistical interactions observed in a multiple-stressor experiment reflect the stressor interactions the researchers were originally interested in depends on many factors that are often overlooked. This is particularly true for presence-absence designs, which were the most common experiments in our systematic review; at least one of the stressors was manipulated at only two levels in ca. 64%

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

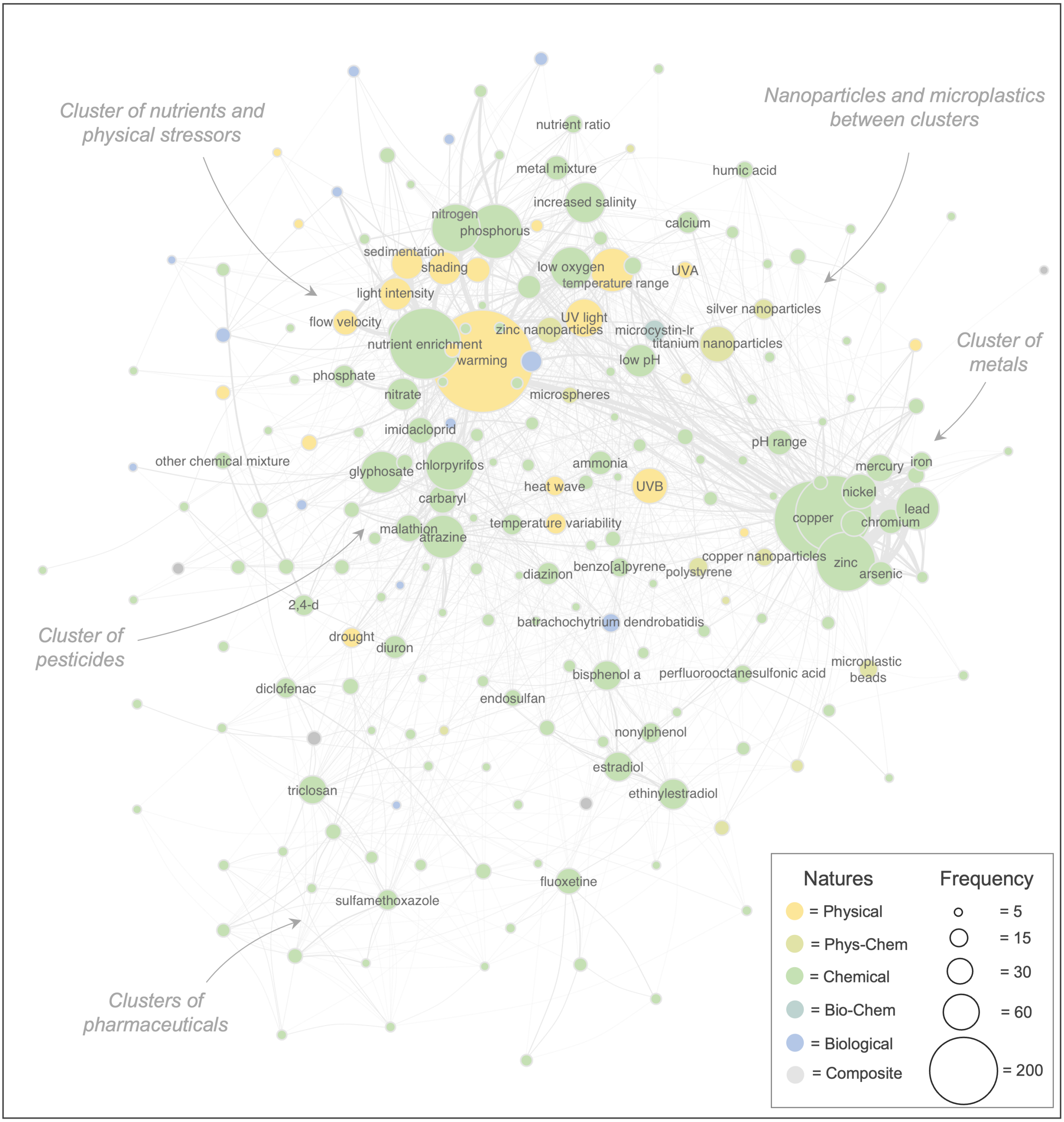

FIGURE 5 Co-occurrence network of stressor identities that occurred at least five times as a force-directed graph. Each node is an individual stressor identity that is coloured by its stressor nature category and has a size proportional to the number of experiments where it was manipulated. The width of the links represents the number of co-occurrences between two stressor identities. The Distributed Recursive Graph Layout algorithm used to generate the network groups frequently co-occurring stressors and aims to maximize interpretability by reducing the amount of overlap between nodes and links. The clustering of stressor identities shows compartmentalization of the different research disciplines studying multiple stressors, with stressors more associated with ecological studies found in the top left part of the network and stressors more associated with ecotoxicology studies in the lower and right-hand sides of the network.

of experiments. Indeed, our systematic review revealed common challenges and emerging trends in experimental multiple-stressor research that can be summarized to inform future research. Here, we outline these recent methodological developments from multiple-stressor research and from the study of interactions (or 'context-dependency') in ecology and evolution more generally, to provide a framework clarifying the blurry connection between stressor interactions and statistical interactions (Figure 6). Our key message is that detecting the stressor interactions of

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

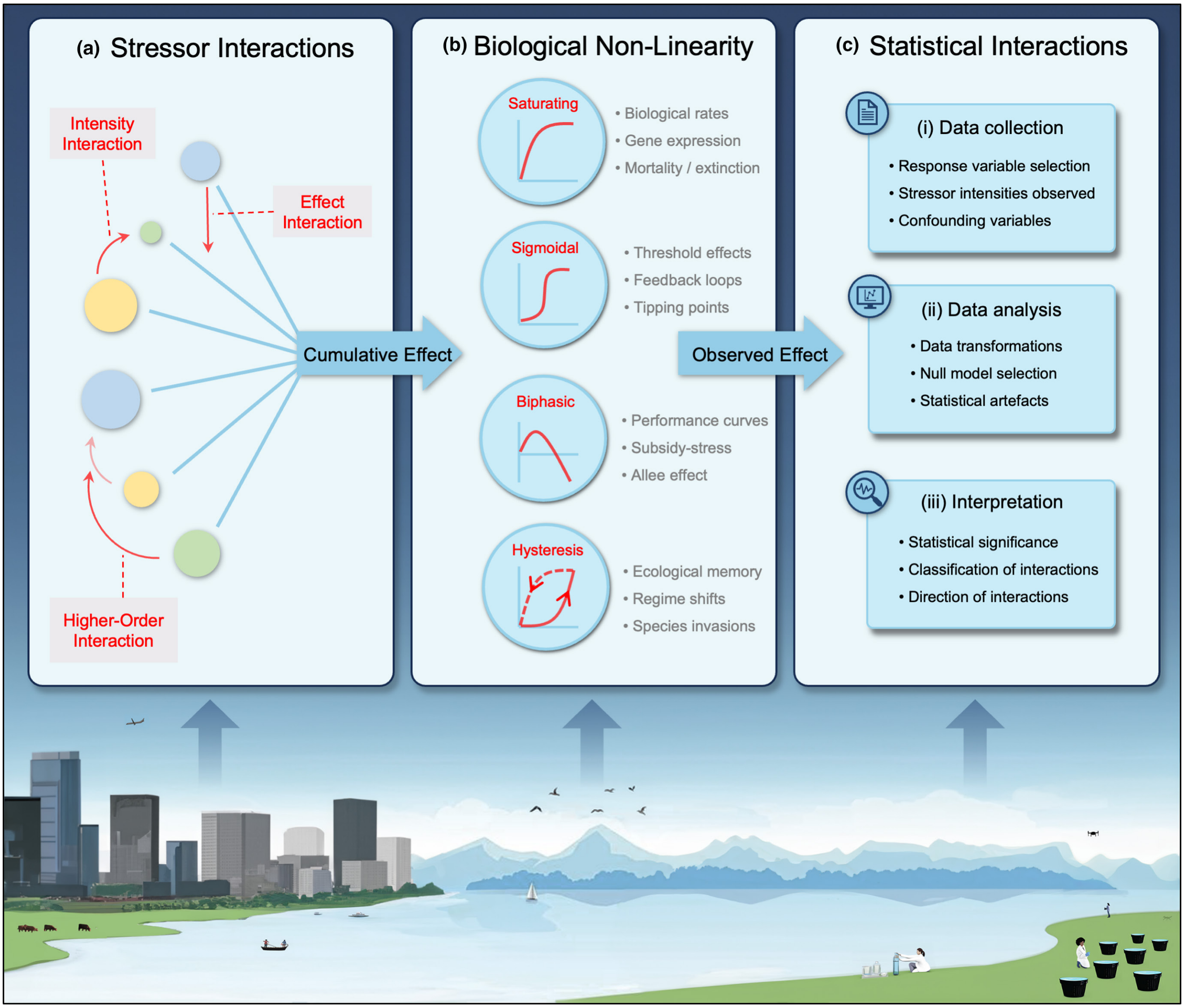

FIGURE 6 Framework for understanding the blurry connection between the intrinsic stressor interactions of interest and the statistical interactions that are ultimately tested by an experiment. (a) Stressors, which are represented by the circles, can influence the intensity or the effect of other stressors (red arrows show intrinsic stressor interactions). (b) The cumulative effects of stressors may be non-additive due to non-linear biological responses. Non-linear biological responses to increasing levels of stress are caused by a variety of effects that can occur at different levels of biological organization. (c) The detected statistical interactions depend on methodological decisions made by the researchers during data collection, analysis and interpretation. The cartoon below the boxes illustrates the three stages at which these effects occur: (a) when stressors interact with each other, (b) when ecological systems respond to the cumulative effects, and (c) when these cumulative effects are studied.

interest (Section 4.1) requires knowledge of non-linear biological responses (Section 4.2) and of the methodological decisions that influence the statistical interactions that are ultimately detected and interpreted (Section 4.3).

Stressor interactions

Our definition of stressor interactions—when stressors influence the intensity or effect of other stressors—focuses on the intrinsic interactions between stressors and the biological systems they are impacting (sensu Didham et al., 2007; Schäfer et al., 2023) rather than on statistically significant deviations from null models. These intrinsic stressor interactions have a mechanistic underpinning, whereas deviations from null models (i.e., statistical interactions) can be highly context-dependent. It is useful to conceptualize three broad types of intrinsic stressor interactions (Figure 6a). (1) *Intensity interactions* occur when one stressor changes the intensity of another stressor. How temperature influences the amount of oxygen that can be dissolved in water, how urbanization increases the warming caused by climate change (Grey et al., 2023), or how drought and floods can change the concentration of chemical pollutants in a water body are all examples of intensity interactions. (2) *Effect interactions* occur when one stressor

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

changes the effect of a given intensity of another stressor. For instance, organisms' responses to temperature are influenced by other stressors (and vice versa), which typically lower both the optimum and maximum performance temperatures (Litchman & Thomas, 2023). Furthermore, pyrethroid insecticides (that inactivate nerve cell sodium channels) but not organophosphate insecticides (whose mode of action is not related to sodium channels) can interact through physiological mechanisms with road salts in zooplankton communities (Lewis et al., 2021). (3) *Higher-order interactions* occur when multiple intensity interactions and/or effect interactions combine. In other words, the interactions between two stressors can be influenced by a third (Diamant et al., 2023). Examples of these high-order interactions are rare in the literature—e.g., temperature-dependent interactions between pesticides (Delnat et al., 2019) and complex suppressive interactions in three-drug combinations (Beppler et al., 2017)—but they could be more prevalent in natural systems.

Biological non-linearities

Without knowledge of stressor–response relationships—how the biological response changes with increasing levels of stress—we do not know if statistical interactions from a factorial experiment will accurately reflect the intrinsic stressor interactions outlined above. Expecting that the combined effects of stressors will be additive—the default of most multiple-stressor experiments—assumes that the biological response of interest will change linearly with increasing levels of stress (Pirotta et al., 2022; Schäfer & Piggott, 2018). If a stressor–response relationship is non-linear, however, even adding the same stressor twice would result in a non-additive combined effect (see Figure 2 in Schäfer et al., 2023). Combining stressors will typically increase overall stress intensity (the x-axis of a stressor–response relationship), so if biological responses to increasing stress are non-linear, additive models can detect statistical interactions that do not reflect the actual interactions between stressors. Non-linear stress responses are widespread in nature due to a range of effects that play out across all levels of biological organization (Figure 6b). For instance, many response variables, like the expression of genes or population abundances, have limits beyond which increasing stress levels will have no further effects (i.e., saturating stressor–response relationships). Some stressors, like nutrients, can have positive effects at some stressor levels but negative effects at others (i.e., biphasic stressor–response relationships, Odum et al., 1979). Finally, stressor–response relationships can themselves vary over time, with processes like evolutionary adaptation or phenotypic plasticity leading to reductions in the effects of stressors over time (Bell, 2017). A robust understanding of individual stressor–response relationships and multiple stressor–response surfaces is not only essential for accurate prediction of individual and combined effects of stressors (Pirotta et al., 2022; Rosenfeld et al., 2022; Van Moorsel et al., 2023), but it can also guide the selection of statistical models that preserve the link between stressor interactions and statistical interactions.

Statistical interactions

The statistical interactions ultimately detected by a multiple-stressor experiment do not only depend on the intrinsic interactions between stressors and the non-linearity of biological responses, but also on methodological decisions made throughout the scientific process (Figure 6c). Firstly, during data collection, researchers should be aware that the choice of biological response variable used to quantify stressor effects, the range of stressor intensities observed or tested, and the measurement of potentially confounding variables can all affect the magnitude and direction of detected statistical interactions (Duncan & Kefford, 2021; Mack et al., 2022; Turschwell et al., 2022). Secondly, during data analysis, researchers must ensure that they use appropriate null models that are mechanistically informed by stressor–response relationships (Schäfer & Piggott, 2018; Tekin et al., 2020). They must also be aware of how common data transformations change the detectability and meaning of statistical interactions (Duncan & Kefford, 2021; Spake et al., 2023) and even how statistical artefacts associated with more complex modelling approaches can introduce non-additivity unrelated to the study system (Orr et al., 2021; Thompson et al., 2018). Finally, when researchers are interpreting model outputs, they must consider if they have enough statistical power, based on the sample sizes in their experiment, to be able to statistically detect meaningful interactions (Burgess et al., 2022). If relevant, it is also important for researchers to consider the symmetry of the interactions (i.e., is stressor A influencing the effect of stressor B, is stressor B influencing the effect of stressor A, or do they both influence each other's effects?) by examining marginal effects plots for the different relationships (Spake et al., 2023). When researchers studying the same systems make different methodological decisions in their studies, there can be 'apparent context-dependency' in the types of statistical interactions that are observed (Catford et al., 2022), which limits our ability to gain mechanistic insights into stressor interactions.

IS THERE AN 'IDEAL' MULTIPLE-STRESSOR EXPERIMENT?

Given these challenges of linking the statistical interactions detected by an experiment to the actual stressor interactions of interest, it is natural to ask if there is an

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

'ideal' multiple-stressor experiment. Indeed, which of the great diversity of experimental designs employed by the studies in our systematic review is best suited for disentangling the effects of multiple stressors? In this section, we argue that the answer to this question depends on the specific research topic and on the feasibility of an experimental design, which depends not only on the number and required control of treatments but also on the complexity of the experimental system and on the resources available to the researchers.

Design options by research topic

Perhaps the simplest setting for a multiple-stressor study is when stressors can be considered categorical factors, so researchers are interested in whether the *presence or absence* of one stressor influences the intensity or the effect of another (Figure 7a). For instance, if a stressor only occurs at a specific intensity in natural systems or if a stressor's intensity varies due to ecological dynamics (e.g., non-native species), then simple factorial experimental designs are appropriate options (Box et al., 2005). Almost one-quarter of the experiments in our dataset (24.2%, $n=580$) were 2-by-2 factorial designs that manipulated two stressors at two levels (e.g., presence and absence). However, based on the diversity of investigated stressors, most of these experiments did not use 2-by-2 factorial designs because their stressors were categorical factors, as described above, but because of feasibility challenges. Indeed, compared to the full dataset, a higher proportion of field and mesocosm studies were 2-by-2 factorial experiments (194 of 500, 38.8%), suggesting that this simple experimental design is more commonly used in more complex systems. If one stressor is categorical but the other is continuous (i.e., variations in its intensity are of interest), then 2-by-n factorial designs are more informative. 2-by-n experiments make up another large portion of the dataset (21.3%, $n=511$) but are more challenging to perform as they require more treatments and greater control over the intensity of a stressor. Increasing the number of levels of the continuous stressor reduces the feasibility of the design; 277 experiments in the dataset used 2-by-3 designs, but just eight studies manipulated 10 or more levels of the second stressor. Designs that include different treatments for various exposure timings and sequences of the presence or absence of stressors (e.g., MacLennan & Vinebrooke, 2021) allow for the detection of temporal-context dependency of stressor interactions (Jackson et al., 2021), but these experiments require more treatments and even more control over stressor intensities.

Other types of designs are more suitable in the applied research setting of *scenario testing* (Figure 7b). Here, researchers are interested in forecasting the biological effects of specific intensities and combinations of stressors that are representative of specific scenarios (e.g., testing temperature and carbon dioxide levels predicted by a climate model, or decreasing the intensity of an agricultural stressor to meet a policy target). The simplest scenario testing designs (e.g., present vs. future) do not manipulate stressors factorially, so interactions cannot

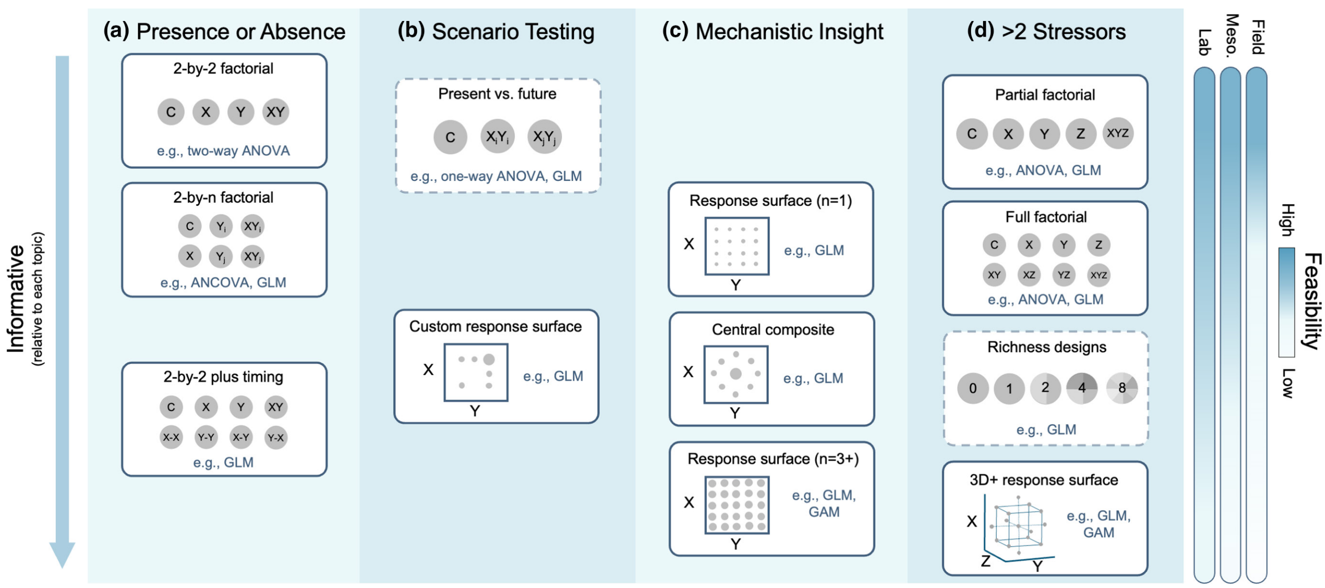

FIGURE 7 Design options for multiple-stressor experiments organized by research topic (a–d). Experimental designs are sorted vertically based on how feasible they are to conduct, which is a function of the experimental system (e.g., laboratory, mesocosm or field). How informative a design is, in the context of a given research topic, is typically negatively correlated with how feasible it is. Grey circles are treatments; X, Y and Z are three stressors; subscripts represent different levels of a stressor; C is for control; and numbers correspond to the number of stressors for the 'richness design'. The design options in the boxes with dashed outlines do not necessarily allow for the quantification of stressor interactions, so they are not found in our dataset.

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

always be quantified. It is very likely that some of the 415 studies that were excluded from our systematic review for not testing the interactive effects of stressors (Figure S2) were designed to test specific scenarios. More complex designs, such as custom response surface designs, can be used to make forecasts that are robust to some level of uncertainty of stressor intensity predictions (Thomas & Ranjan, 2024) and can also characterize the interactions between stressors, but these designs require a high degree of control over the intensities of stressors, which may be difficult to achieve in field or even some mesocosm experiments.

There has been a widespread call for the use of regression-based approaches in multiple-stressor research to improve predictive power and to gain deeper *mechanistic insights* (Figure 7c) on the potentially non-linear responses of biological systems to anthropogenic impacts (Boyd et al., 2018; Collins et al., 2022; Cottingham et al., 2005; Kreyling et al., 2018; Orr et al., 2020; Thomas & Ranjan, 2024). For two continuous stressors, response surface designs can be used to compare biological responses across different points of the state space defined by the intensity of the stressors. Response surface designs come in many different forms, reviewed by Thomas and Ranjan (2024), that vary in their efficacy and feasibility. However, if stressor intensities can be well controlled, these designs can still be feasible for the average research group. If experimental units were limited, for instance, a full factorial response surface design with just one replicate per treatment could be used (e.g., 5-by-5 factorial would require 25 experimental units), and regression models would still be able to provide insights on the mechanisms of the potentially non-linear individual and interactive effects of stressors (but greater replication would be required to detect smaller effect sizes). If experimental units are not limited and a research group has considerable resources (far more likely for laboratory systems), then high resolution response surface designs with replication (e.g., 10-by-10 factorial with three replicates per treatment would require 300 experimental units) could be used in combination with sophisticated modelling techniques such as GAMs to gain deep insights into biological responses to multiple stressors. Here, the uncertainty of predicted non-linear relationships is generally decreased by increasing the sample size or the resolution of tested stressor intensities (Wood, 2017). Response surface experiments were rare in our dataset, with only 194 studies (8%) measuring at least five levels of two stressors. Reflecting feasibility challenges, 181 of these studies were lab experiments, 11 were mesocosm experiments and just two were field experiments. Furthermore, 361 of the 388 stressors studied in these response surface experiments were chemical or chemical–physical in nature, and 91.8% of studies were on single species, suggesting that these designs are more common in ecotoxicological, rather than ecological, research on multiple stressors. There are many alternatives to full-factorial response surfaces, such as central composite designs or space-filling designs that are uncommon in ecology and evolution, which can use prior knowledge of biological responses (e.g., known boundaries) to optimize resources (see simulation studies in Thomas & Ranjan, 2024).

Researchers are becoming increasingly interested in testing the combined and interactive effects of *more than two stressors* (Figure 7d). All the experimental designs discussed above can be extended to three or more stressors, but the number of treatments will increase exponentially with the number of stressors in factorial designs (i.e., the combinatorial explosion problem). Even for three stressors, each tested at just two levels (i.e., presence vs. absence), the number of treatments increases from four to eight. Partial factorial designs (i.e., not testing all possible combinations) can be used to increase the feasibility of a study, but then only the net higher-order interactions (encompassing unmeasured lower-order interactions) can be quantified (Diamant et al., 2023). Indeed, 432 of the 672 experiments in our dataset that tested more than two stressors were partially factorial. Fully factorial experiments with three stressors each at two levels were relatively common ($n=122$) with a slight majority of these being field ($n=16$) or mesocosm ($n=42$) studies. There were even some experiments that built higher-dimensional response surfaces. Although these experiments are very challenging to perform—they require many treatments and a high level of control of stressor intensities—there were 26 studies in our dataset that factorially manipulated at least three stressors each across at least five levels (a minimum of 125 treatments). Unsurprisingly, all these studies were single-species laboratory studies that tested multiple chemical stressors. Finally, it is worth highlighting an interesting form of partially factorial design where treatments are the number (not identity) of stressors (Rillig et al., 2019). Although these studies are not best suited for quantifying pairwise and higher order stressor interactions, they are relatively feasible options for studying the combined effects of many co-occurring stressors.

Recommendations for future experiments

Statistical interactions observed by experiments should ideally reflect the underlying stressor interactions that are important for conservation and ecosystem management. Based on the potential pitfalls outlined in Sections 4.2 and 4.3 and on the trade-off between feasibility and how informative the different experimental designs outlined in Section 5.1 are, we propose several best practices that future multiple-stressor experiments can employ. *(1) Measure the abiotic effects of stressors*. Collecting data on relevant abiotic variables (e.g., temperature, oxygen, pH) and on the concentrations and degradation of chemicals during an

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

experiment can help to identify context-dependency of stressor effects (Kefford et al., 2023). It may be useful to have additional 'abiotic controls' in an experimental design that do not contain biological components so that the physical and chemical interactions between stressors can be independently assessed. *(2) Construct response surfaces*. Response surface designs, where the responses of key biological variables (e.g., growth, mortality and productivity) are quantified across different intensities of stressors, are the best experimental designs for disentangling non-linear biological responses and intrinsic stressor interactions (Collins et al., 2022; Van Moorsel et al., 2023). If stressors are not categorical variables, if stressor intensities can be controlled, and if researchers are not testing a specific scenario, future multiple-stressor experiments should use regression-based approaches that provide far more mechanistic insights than simple presence-absence designs (Boyd et al., 2018; Cottingham et al., 2005; Kreyling et al., 2018; Thomas & Ranjan, 2024). *(3) Focus on mechanisms of interactions rather than deviations from additivity*. Analyses of potential stressor interactions should be theoretically driven and informed by our best mechanistic understanding of the system, rather than being focused on detecting a deviation from an uninformed (e.g., additive) statistical model. Although exploratory analyses with multiple different null models can sometimes be useful, in general, establishing clear-cut predictions followed by transparent and reproducible data analyses will clarify the blurry connection between stressor interactions and statistical interactions. Even for a mechanistically informed null model, experimental designs (e.g., sample sizes) should ideally be informed by power analyses so that biologically meaningful interactions can be detected (Johnson et al., 2015; Lakens & Caldwell, 2021). This approach will help multiple-stressor researchers not to lose sight of a central goal of their experiments: to understand if stressors influence each other's intensities and effects.

FUTURE DIRECTIONS AND CONCLUSIONS

Our systematic review provides resources that will help to inform future work on the impacts of multiple stressors in freshwater ecosystems. More generally, however, our framework for connecting stressor interactions and statistical interactions (Figure 6) and our practical guidelines for the design of multiple-stressor experiments (Figure 7) are relevant to any system and can even be generalized to study other forms of interaction in ecology and evolutionary biology (i.e., Spake et al., 2023). Although performing a formal meta-analysis of the entire dataset of 2396 experiments would be a prohibitively large task—the median and maximum number of studies per meta-analysis were 24 and 369, respectively, in a review of 466 meta-analyses in ecology (Costello & Fox, 2022)—an important next step will be conducting targeted meta-analyses on meaningful subsets of the dataset, such as the 781 stressor combinations that involved temperature or the 1092 stressor combinations that were tested at the community or ecosystem levels. Furthermore, our stressor taxonomy could be coupled to mechanistic hypotheses of stressor interactions and stressor–response relationships for data-rich studies (i.e., studies with many stressor levels) to make predictions and inform ecosystem management. The interactive web application (https://jamesaorr.shinyapps.io/freshwater-multiple-stressors/) can be used by researchers, students and stakeholders to identify combinations of stressors that have not yet been studied, to access studies relevant to their work, and to contribute additional knowledge to the dataset. This freshwater dataset should also eventually be combined with multiple-stressor datasets from other disciplines in ecology and global change biology (e.g., Song et al., 2019; Van Sundert et al., 2023) to enhance cross-fertilization of ideas and perspectives.

The number and diversity of multiple-stressor experiments found through our systematic review were remarkable and exceeded our expectations. The shared goal of gaining information on all relevant stressor combinations, as well as the pressures surrounding productivity and novelty in research outputs, encourages researchers to design experiments with previously unstudied combinations of stressors and systems. Although this diversity of experiments is undoubtedly a strength of the field, the pursuit of novelty may lead to research efforts becoming spread thin. As an alternative approach, coordinated and distributed experiments (e.g., Harpole et al., 2016; Romero et al., 2020) offer a powerful approach for studying environmental context-dependencies while overcoming experimenter-induced context-dependencies. Although it is vital to fill the many gaps in our knowledge (white cells in Figure 4b), it is important for research efforts to focus on combinations of stressors, and intensities, that typically co-occur in natural systems (Bowler et al., 2020) to most effectively address the urgent threats facing freshwaters. Comparing the co-occurrence of investigated stressors to the current and predicted co-occurrence of stressors in natural systems will be a critical next step. Similarly, focusing on how the effects of dominant stressors in specific systems, such as pesticides in lowland streams (Liess et al., 2021) and nutrient enrichment in lakes (Birk et al., 2020), are modified by other stressors will be particularly informative for conservation and management.

To better understand, predict and manage multiple stressors in natural systems, some experimental designs will be more useful than others. Studies, particularly with communities of species, that report multiple biological responses across levels of organization

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

can help to identify if stressor interactions at one level are causing non-additive responses at another (Rillig, Lehmann, et al., 2023; Simmons et al., 2021). Furthermore, more long-term experiments that study adaptation to stressors or the trajectory of communities following the removal of stressors (only ca. 7% in our dataset) will be essential for determining effective stressor removal strategies (Vos et al., 2023). There is also a risk that too much emphasis is being placed on the identification and classification of stressor interactions rather than on the prediction of their combined effects, which may hinder mitigation or management efforts. Indeed, the diversity of stressors, without knowledge of their interactions, can still provide valuable information about their ecological impacts (Rillig et al., 2019; Rillig, van der Heijden, et al., 2023). Nevertheless, multiple-stressor experiments, particularly response surface designs (Kreyling et al., 2018; Thomas & Ranjan, 2024), are still critical for advancing our fundamental knowledge of the combined effects of multiple stressors in freshwaters.

Of course, not all stressors that impact freshwater ecosystems can be simulated experimentally. There can be ethical issues associated with testing the effects of non-native species and anthropogenically spread diseases. Furthermore, stressors that have direct impacts at the landscape scale, such as habitat fragmentation through damming, are difficult to meaningfully replicate. As such, using empirical evidence to construct process-based models linking physical, chemical, and biological mechanisms (e.g., López Moreira Mazacotte et al., 2023) or to parameterise numerical models of ecosystems (Simmons et al., 2021) will be essential for advancing our understanding of all the threats facing freshwater ecosystems. With this systematic review, we now know the current standing of experimental work on multiple stressors in freshwaters, and the resources provided here will increase clarity and efficiency in the field. The extensive research efforts that have been building over the past decades must continue if we are to gain the knowledge required to effectively conserve and manage freshwater ecosystems in the 21st century.

AUTHOR CONTRIBUTIONS

JAO and MCJ conceived the idea for the project. JAO organized the project, analysed the data, made the figures, wrote the supporting information, and wrote the first draft of the manuscript. JAO and SJM curated the taxonomy of stressors. JAO, SJM, AM, BB, DA, JGH, KR and RW performed abstract screening and full-text screening. MCJ, AS, AMR, BL, BS, DS, IO, JJP, JZ, KVD, LCA, MP, MTTV, NJK, PZ, PS, RBS, RV, SH and TR also performed full-text screening. All authors contributed to the final paper.

AFFILIATIONS

[1]Department of Biology, University of Oxford, Oxford, UK
[2]School of the Environment, University of Queensland, Brisbane, Queensland, Australia
[3]Department of Genetics, Evolution and Environment, University College London, London, UK
[4]Department of Zoology, University of Otago, Dunedin, New Zealand
[5]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany
[6]Ruumi ApS, Svendborg, Denmark
[7]Zoology, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland
[8]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada
[9]Natural Resources Institute, University of Manitoba, Winnipeg, Canada
[10]Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China
[11]Section for Aquatic Biology and Toxicology, Department of Biosciences, University of Oslo, Oslo, Norway
[12]Department of Biological Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada
[13]School of Biology and Environmental Science, University College Dublin, Dublin, Ireland
[14]Institute of Aquaculture, University of Stirling, Scotland, UK
[15]Institute of Environmental Sciences, RPTU Kaiserslautern-Landau, Germany
[16]Research Center One Health Ruhr, University Alliance Ruhr
[17]Faculty of Biology, University Duisburg-Essen, Essen, Germany
[18]School of Biological, Earth & Environmental Sciences, University College Cork, Cork, Ireland

ACKNOWLEDGEMENTS

JAO, SJM and MCJ were supported by a Natural Environment Research Council grant (NE/V001396/1). We are grateful to Lily Adeniji, who worked with JAO to collect the taxonomic group data during her UNIQ+ research internship and to Daniel Leroux Padilla, who worked with JAO to create the interactive web application as part of a British Ecological Society grant (SR23\1202).

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ele.14463.

DATA AVAILABILITY STATEMENT

The dataset produced by the systematic review and the code used to create the figures are publicly available at https://zenodo.org/doi/10.5281/zenodo.11100467.

ORCID

James A. Orr https://orcid.org/0000-0002-6531-5623
Blake R. Stuparyk https://orcid.org/0000-0001-5202-0829
Peiyu Zhang https://orcid.org/0000-0002-9551-3882

REFERENCES

Almond, R., Grooten, M. & Peterson, T. (2020) Living planet report 2020-bending the curve of biodiversity loss. World Wildlife Fund.

Amano, T., Berdejo-Espinola, V., Akasaka, M., de Andrade Junior, M.A., Blaise, N., Checco, J. et al. (2023) The role of non-English-language science in informing national biodiversity assessments. *Nature Sustainability*, 6(7), 845–854.

Beauchesne, D., Cazelles, K., Archambault, P., Dee, L.E. & Gravel, D. (2021) On the sensitivity of food webs to multiple stressors. *Ecology Letters*, 24(10), 2219–2237.

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bell, G. (2017) Evolutionary rescue. *Annual Review of Ecology, Evolution, and Systematics*, 48, 605–627.

Beppler, C., Tekin, E., White, C., Mao, Z., Miller, J.H., Damoiseaux, R. et al. (2017) When more is less: emergent suppressive interactions in three-drug combinations. *BMC Microbiology*, 17, 1–9.

Birk, S., Chapman, D., Carvalho, L., Spears, B.M., Andersen, H.E., Argillier, C. et al. (2020) Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. *Nature Ecology & Evolution*, 4(8), 1060–1068.

Bowler, D.E., Bjorkman, A.D., Dornelas, M., Myers-Smith, I.H., Navarro, L.M., Niamir, A. et al. (2020) Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. *People and Nature*, 2(2), 380–394.

Box, G.E., Hunter, J.S. & Hunter, W.G. (2005) Statistics for experimenters. In: *Wiley series in probability and statistics*. NJ: Wiley Hoboken.

Boyd, P.W., Collins, S., Dupont, S., Fabricius, K., Gattuso, J., Havenhand, J. et al. (2018) Experimental strategies to assess the biological ramifications of multiple drivers of global ocean change—a review. *Global Change Biology*, 24(6), 2239–2261.

Burgess, B.J., Jackson, M.C. & Murrell, D.J. (2022) Are experiment sample sizes adequate to detect biologically important interactions between multiple stressors? *Ecology and Evolution*, 12(9), e9289.

Catford, J.A., Wilson, J.R., Pyšek, P., Hulme, P.E. & Duncan, R.P. (2022) Addressing context dependence in ecology. *Trends in Ecology & Evolution*, 37(2), 158–170.

Collins, S., Whittaker, H. & Thomas, M.K. (2022) The need for unrealistic experiments in global change biology. *Current Opinion in Microbiology*, 68, 102151.

Costello, L. & Fox, J.W. (2022) Decline effects are rare in ecology. *Ecology*, 103(6), e3680.

Côté, I.M., Darling, E.S. & Brown, C.J. (2016) Interactions among ecosystem stressors and their importance in conservation. *Proceedings of the Royal Society B*, 283(1824), 20152592.

Cottingham, K.L., Lennon, J.T. & Brown, B.L. (2005) Knowing when to draw the line: designing more informative ecological experiments. *Frontiers in Ecology and the Environment*, 3(3), 145–152.

De Laender, F. (2018) Community- and ecosystem-level effects of multiple environmental change drivers: beyond null model testing. *Global Change Biology*, 24(11), 5021–5030.

Delnat, V., Tran, T.T., Janssens, L. & Stoks, R. (2019) Daily temperature variation magnifies the toxicity of a mixture consisting of a chemical pesticide and a biopesticide in a vector mosquito. *Science of the Total Environment*, 659, 33–40.

Diamant, E.S., Boyd, S., Lozano-Huntelman, N.A., Enriquez, V., Kim, A.R., Savage, V.M. et al. (2023) Meta-analysis of three-stressor combinations on population-level fitness reveal substantial higher-order interactions. *Science of the Total Environment*, 864, 161163.

Didham, R.K., Tylianakis, J.M., Gemmell, N.J., Rand, T.A. & Ewers, R.M. (2007) Interactive effects of habitat modification and species invasion on native species decline. *Trends in Ecology & Evolution*, 22(9), 489–496.

Dinh, K.V., Albini, D., Orr, J.A., Macaulay, S.J., Rillig, M.C., Borgå, K. et al. (2023) Winter is coming: interactions of multiple stressors in winter and implications for the natural world. *Global Change Biology*, 29(24), 6834–6845. Available from: https://doi.org/10.1111/gcb.16956

Duncan, R.P. & Kefford, B.J. (2021) Interactions in statistical models: three things to know. *Methods in Ecology and Evolution*, 12(12), 2287–2297.

Eisenberg, J.N. & McKone, T.E. (1998) Decision tree method for the classification of chemical pollutants: incorporation of across-chemical variability and within-chemical uncertainty. *Environmental Science & Technology*, 32(21), 3396–3404.

Folt, C., Chen, C., Moore, M. & Burnaford, J. (1999) Synergism and antagonism among multiple stressors. *Limnology and Oceanography*, 44(3), 864–877.

Grey, V., Smith-Miles, K., Fletcher, T.D., Hatt, B. & Coleman, R. (2023) Empirical evidence of climate change and urbanization impacts on warming stream temperatures. *Water Research*, 247, 120703. Available from: https://doi.org/10.1016/j.watres.2023.120703

Harpole, W.S., Sullivan, L.L., Lind, E.M., Firn, J., Adler, P.B., Borer, E.T. et al. (2016) Addition of multiple limiting resources reduces grassland diversity. *Nature*, 537(7618), 93–96.

He, F., Arora, R. & Mansour, I. (2023) *Multispecies assemblages and multiple stressors: synthesizing the state of experimental research in freshwaters*. Wiley Interdisciplinary Reviews: Water, p. e1641.

Jackson, M.C., Loewen, C.J., Vinebrooke, R.D. & Chimimba, C.T. (2016) Net effects of multiple stressors in freshwater ecosystems: a meta-analysis. *Global Change Biology*, 22(1), 180–189.

Jackson, M.C., Pawar, S. & Woodward, G. (2021) The temporal dynamics of multiple stressor effects: from individuals to ecosystems. *Trends in Ecology & Evolution*, 36(5), 402–410.

Johnson, P.C., Barry, S.J., Ferguson, H.M. & Müller, P. (2015) Power analysis for generalized linear mixed models in ecology and evolution. *Methods in Ecology and Evolution*, 6(2), 133–142.

Kefford, B.J., Nichols, S.J. & Duncan, R.P. (2023) The cumulative impacts of anthropogenic stressors vary markedly along environmental gradients. *Global Change Biology*, 29(3), 590–602.

Kreyling, J., Schweiger, A.H., Bahn, M., Ineson, P., Migliavacca, M., Morel-Journel, T. et al. (2018) To replicate, or not to replicate–that is the question: how to tackle nonlinear responses in ecological experiments. *Ecology Letters*, 21(11), 1629–1638.

Lakens, D. & Caldwell, A.R. (2021) Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920951503.

Lange, K., Bruder, A., Matthaei, C.D., Brodersen, J. & Paterson, R.A. (2018) Multiple-stressor effects on freshwater fish: importance of taxonomy and life stage. *Fish and Fisheries*, 19(6), 974–983.

Lewis, J.L., Agostini, G., Jones, D.K. & Relyea, R.A. (2021) Cascading effects of insecticides and road salt on wetland communities. *Environmental Pollution*, 272, 116006.

Liess, M., Liebmann, L., Vormeier, P., Weisner, O., Altenburger, R., Borchardt, D. et al. (2021) Pesticides are the dominant stressors for vulnerable insects in lowland streams. *Water Research*, 201, 117262.

Litchman, E. & Thomas, M.K. (2023) Are we underestimating the ecological and evolutionary effects of warming? Interactions with other environmental drivers may increase species vulnerability to high temperatures. *Oikos*, 2023(2), e09155.

Lodi, S., Godoy, B.S., Ortega, J.C. & Bini, L.M. (2021) Quality of meta-analyses in freshwater ecology: a systematic review. *Freshwater Biology*, 66(5), 803–814.

López Moreira Mazacotte, G.A., Polst, B.H., Gross, E.M., Schmitt-Jansen, M., Hölker, F. & Hilt, S. (2023) Microcosm experiment combined with process-based modeling reveals differential response and adaptation of aquatic primary producers to warming and agricultural run-off. *Frontiers in Plant Science*, 14, 1120441.

Mack, L., de la Hoz, C.F., Penk, M., Piggott, J., Crowe, T., Hering, D. et al. (2022) Perceived multiple stressor effects depend on sample size and stressor gradient length. *Water Research*, 226, 119260.

MacLennan, M.M. & Vinebrooke, R.D. (2021) Exposure order effects of consecutive stressors on communities: the role of co-tolerance. *Oikos*, 130, 2111–2121.

Moe, S.J., De Schamphelaere, K., Clements, W.H., Sorensen, M.T., Van den Brink, P.J. & Liess, M. (2013) Combined and interactive effects of global climate change and toxicants on populations and communities. *Environmental Toxicology and Chemistry*, 32(1), 49–61.

Naasz, S., Altenburger, R. & Kühnel, D. (2018) Environmental mixtures of nanomaterials and chemicals: the Trojan-horse phenomenon and its relevance for ecotoxicity. *Science of the Total Environment*, 635, 1170–1181.

O'Dea, R.E., Lagisz, M., Jennions, M.D., Koricheva, J., Noble, D.W., Parker, T.H. et al. (2021) Preferred reporting items for systematic

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biological Reviews*, 96(5), 1695–1722.
Odum, E.P., Finn, J.T. & Franz, E.H. (1979) Perturbation theory and the subsidy-stress gradient. *Bioscience*, 29(6), 349–352.
Ormerod, S.J., Dobson, M., Hildrew, A.G. & Townsend, C. (2010) Multiple stressors in freshwater ecosystems. *Freshwater Biology*, 55, 1–4.
Orr, J.A., Piggott, J.J., Jackson, A.L. & Arnoldi, J. (2021) Scaling up uncertain predictions to higher levels of organisation tends to underestimate change. *Methods in Ecology and Evolution*, 12(8), 1521–1532.
Orr, J.A., Vinebrooke, R.D., Jackson, M.C., Kroeker, K.J., Kordas, R.L., Mantyka-Pringle, C. et al. (2020) Towards a unified study of multiple stressors: divisions and common goals across research disciplines. *Proceedings of the Royal Society B*, 287(1926), 20200421.
Outhwaite, C.L., McCann, P. & Newbold, T. (2022) Agriculture and climate change are reshaping insect biodiversity worldwide. *Nature*, 605(7908), 97–102.
Pirotta, E., Thomas, L., Costa, D.P., Hall, A.J., Harris, C.M., Harwood, J. et al. (2022) Understanding the combined effects of multiple stressors: a new perspective on a longstanding challenge. *Science of the Total Environment*, 821, 153322.
Rahel, F.J. & Olden, J.D. (2008) Assessing the effects of climate change on aquatic invasive species. *Conservation Biology*, 22(3), 521–533.
Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T. et al. (2019) Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873.
Rillig, M.C., Lehmann, A., Orr, J.A. & Rongstock, R. (2023) Factors of global change affecting plants act at different levels of the ecological hierarchy. *The Plant Journal*, 117, 1781–1785.
Rillig, M.C., Ryo, M. & Lehmann, A. (2021) Classifying human influences on terrestrial ecosystems. *Global Change Biology*, 27(11), 2273–2278.
Rillig, M.C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C.A., Buchert, S., Wulf, A. et al. (2019) The role of multiple global change factors in driving soil functions and microbial biodiversity. *Science*, 366(6467), 886–890.
Rillig, M.C., van der Heijden, M.G., Berdugo, M., Liu, Y.-R., Riedo, J., Sanz-Lazaro, C. et al. (2023) Increasing the number of stressors reduces soil ecosystem services worldwide. *Nature Climate Change*, 13(5), 478–483.
Romero, G.Q., Marino, N.A., MacDonald, A.A.M., Céréghino, R., Trzcinski, M.K., Mercado, D.A. et al. (2020) Extreme rainfall events alter the trophic structure in bromeliad tanks across the Neotropics. *Nature Communications*, 11(1), 3215.
Rosenfeld, J., Gonzalez-Espinosa, P., Jarvis, L., Enders, E., Bayly, M., Paul, A. et al. (2022) Stressor-response functions as a generalizable model for context dependence. *Trends in Ecology & Evolution*, 37, 1032–1035.
Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R. et al. (2000) Global biodiversity scenarios for the year 2100. *Science*, 287(5459), 1770–1774.
Schäfer, R.B., Jackson, M., Juvigny-Khenafou, N., Osakpolor, S.E., Posthuma, L., Schneeweiss, A. et al. (2023) Chemical mixtures and multiple stressors–same but different? *Environmental Toxicology and Chemistry*, 42, 1915–1936.
Schäfer, R.B. & Piggott, J.J. (2018) Advancing understanding and prediction in multiple stressor research through a mechanistic basis for null models. *Global Change Biology*, 24(5), 1817–1826.
Schindler, D.W. (1998) Whole-ecosystem experiments: replication versus realism: the need for ecosystem-scale experiments. *Ecosystems*, 1(4), 323–334.
Simmons, B.I., Blyth, P.S., Blanchard, J.L., Clegg, T., Delmas, E., Garnier, A. et al. (2021) Refocusing multiple stressor research around the targets and scales of ecological impacts. *Nature Ecology & Evolution*, 5(11), 1478–1489.
Song, J., Wan, S., Piao, S., Knapp, A.K., Classen, A.T., Vicca, S. et al. (2019) A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change. *Nature Ecology & Evolution*, 3(9), 1309–1320.
Spake, R., Bowler, D.E., Callaghan, C.T., Blowes, S.A., Doncaster, C.P., Antao, L.H. et al. (2023) Understanding 'it depends' in ecology: a guide to hypothesising, visualising and interpreting statistical interactions. *Biological Reviews*, 98(4), 983–1002.
Spears, B.M., Chapman, D.S., Carvalho, L., Feld, C.K., Gessner, M.O., Piggott, J.J. et al. (2021) Making waves. Bridging theory and practice towards multiple stressor management in freshwater ecosystems. *Water Research*, 196, 116981.
Tekin, E., Diamant, E.S., Cruz-Loya, M., Enriquez, V., Singh, N., Savage, V.M. et al. (2020) Using a newly introduced framework to measure ecological stressor interactions. *Ecology Letters*, 23(9), 1391–1403.
Thomas, M.K. & Ranjan, R. (2024) Designing more informative multiple-driver experiments. *Annual Review of Marine Science*, 16, 513–536.
Thompson, P.L., MacLennan, M.M. & Vinebrooke, R.D. (2018) An improved null model for assessing the net effects of multiple stressors on communities. *Global Change Biology*, 24(1), 517–525.
Turschwell, M.P., Connolly, S.R., Schäfer, R.B., De Laender, F., Campbell, M.D., Mantyka-Pringle, C. et al. (2022) Interactive effects of multiple stressors vary with consumer interactions, stressor dynamics and magnitude. *Ecology Letters*, 25(6), 1483–1496.
Van De Schoot, R., De Bruin, J., Schram, R., Zahedi, P., De Boer, J., Weijdema, F. et al. (2021) An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, 3(2), 125–133.
Van Moorsel, S.J., Thébault, E., Radchuk, V., Narwani, A., Montoya, J.M., Dakos, V. et al. (2023) Predicting effects of multiple interacting global change drivers across trophic levels. *Global Change Biology*, 29(5), 1223–1238.
Van Sundert, K., Leuzinger, S., Bader, M.K., Chang, S.X., De Kauwe, M.G., Dukes, J.S. et al. (2023) When things get MESI: the manipulation experiments synthesis initiative—a coordinated effort to synthesize terrestrial global change experiments. *Global Change Biology*, 29(7), 1922–1938.
Viers, J., Dupré, B. & Gaillardet, J. (2009) Chemical composition of suspended sediments in world Rivers: new insights from a new database. *Science of the Total Environment*, 407(2), 853–868.
Vinebrooke, R.D., Cottingham, K.L., Norberg, M.S., Dodson, S.I., Maberly, S.C. & Sommer, U. (2004) Impacts of multiple stressors on biodiversity and ecosystem functioning: the role of species cotolerance. *Oikos*, 104(3), 451–457.
Vos, M., Hering, D., Gessner, M.O., Leese, F., Schäfer, R.B., Tollrian, R. et al. (2023) The asymmetric response concept explains ecological consequences of multiple stressor exposure and release. *Science of the Total Environment*, 872, 162196.
Wood, S.N. (2017) *Generalized additive models: an introduction with R*. New York: Chapman and Hall/CRC. Available from: https://doi.org/10.1201/9781315370279
Yankow, M. (1965) Effect of growth inhibitors and their combinations on the tail regeneration of amphibia. *Transactions of the New York Academy of Sciences*, 27(8 Series II), 878–881.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Orr, J.A., Macaulay, S.J., Mordente, A., Burgess, B., Albini, D., Hunn, J.G. et al. (2024) Studying interactions among anthropogenic stressors in freshwater ecosystems: A systematic review of 2396 multiple-stressor experiments. *Ecology Letters*, 27, e14463. Available from: https://doi.org/10.1111/ele.14463

14610248, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ele.14463 by University Of Stirling Sonia W, Wiley Online Library on [18/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License