Global Change Biology WILEY

Classifying ecosystem stressor interactions: Theory highlights the data limitations of the additive null model and the difficulty in revealing ecological surprises

Benjamin J. Burgess[1] 🟢 | Drew Purves[2] | Georgina Mace[1] | David J. Murrell[1]

[1]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, UK

[2]DeepMind, London, UK

Correspondence
Benjamin J. Burgess, Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK.
Email: benjamin.burgess.17@ucl.ac.uk

Abstract

Understanding how multiple co-occurring environmental stressors combine to affect biodiversity and ecosystem services is an on-going grand challenge for ecology. Currently, progress has been made through accumulating large numbers of smaller-scale empirical studies that are then investigated by meta-analyses to detect general patterns. There is particular interest in detecting, understanding and predicting 'ecological surprises' where stressors interact in a non-additive (e.g. antagonistic or synergistic) manner, but so far few general results have emerged. However, the ability of the statistical tools to recover non-additive interactions in the face of data uncertainty is unstudied, so crucially, we do not know how well the empirical results reflect the true stressor interactions. Here, we investigate the performance of the commonly implemented additive null model. A meta-analysis of a large (545 interactions) empirical dataset for the effects of pairs of stressors on freshwater communities reveals additive interactions dominate individual studies, whereas pooling the data leads to an antagonistic summary interaction class. However, analyses of simulated data from food chain models, where the underlying interactions are known, suggest both sets of results may be due to observation error within the data. Specifically, we show that the additive null model is highly sensitive to observation error, with non-additive interactions being reliably detected at only unrealistically low levels of data uncertainty. Similarly, plausible levels of observation error lead to meta-analyses reporting antagonistic summary interaction classifications even when synergies co-dominate. Therefore, while our empirical results broadly agree with those of previous freshwater meta-analyses, we conclude these patterns may be driven by statistical sampling rather than any ecological mechanisms. Further investigation of candidate null models used to define stressor-pair interactions is essential, and once any artefacts are accounted for, the so-called 'ecological surprises' may be more frequent than was previously assumed.

KEYWORDS
environmental drivers, food chain, freshwater, Lotka-Volterra, meta-analysis, multiple stressors, observation error, theoretical ecology

1 | INTRODUCTION

Ecological communities are being subjected to a wide variety of external stressors (Halpern et al., 2015) that act across terrestrial, freshwater and marine biomes and threaten ecosystems and their services (Scheffers et al., 2016). These stressors, also termed drivers, factors or perturbations (Orr et al., 2020), are frequently anthropogenic in origin (Geldmann et al., 2014; Vörösmarty et al., 2010), but are capable of being abiotic or biotic (Przeslawski et al., 2015), and are able to act at the local to global scales (Ban et al., 2014; França et al., 2020). While individual stressors (e.g. climate change, habitat alteration or pollution) are themselves capable of inducing changes in biodiversity or ecosystems and their services (Dirzo et al., 2014; Newbold et al., 2015; Tittensor et al., 2014), ecosystems are frequently, if not predominately, acted upon by multiple stressors simultaneously (Crain et al., 2008). Despite the negative connotations surrounding the term *stressor*, stressors are capable of inducing effects that are either beneficial or detrimental to the affected ecosystem (Kroeker et al., 2017). One of the grand challenges facing ecologists is to be able to detect, understand and predict how these different types of ecosystem stressors interact to affect biodiversity and ecosystem services (Hodgson & Halpern, 2019); although the challenge is more difficult since the observed interactions can substantially deviate from what is anticipated (Christensen et al., 2006). Ultimately, knowledge of how stressors interact is important in guiding conservation and management initiatives, and in helping to prevent remediation measures from being ineffective, or even potentially harming those ecosystems they are intended to preserve (Brown et al., 2013; Côté et al., 2016).

Aquatic ecosystems and communities are particularly threatened by multiple stressors (Birk et al., 2020). For instance, Halpern et al. (2008) describe how every marine area is subjected to human influence, with 41% of these areas being impacted by multiple stressors. Moreover, freshwaters represent some of the most at-risk ecosystems and are frequently exposed to a wide range of stressors (He et al., 2019; Hecky et al., 2010; Ormerod et al., 2010; Woodward et al., 2010), with freshwater biodiversity declining at rates exceeding even those of the most impacted terrestrial ecosystems (Sala et al., 2000), and potentially endangering vital ecosystem services (Malaj et al., 2014). While stressors often interact to impact freshwater ecosystems (Birk et al., 2020), their presence in freshwater ecosystems is not a new phenomenon, with some freshwater bodies having been subjected to stressors for several centuries (Dudgeon et al., 2006). However, the stressors that freshwater systems are currently facing has expanded, with the introduction of novel stressors, such as nanomaterials, while existing ones are continuing to have severe impacts (Reid et al., 2019). The cumulative impact of multiple stressors has been identified as one of the most pressing and emerging threats to freshwater biodiversity, but despite this, our current understanding of both how stressors interact, and the severity of their effects, is poor (Reid et al., 2019).

The term *ecological surprise* (sensu Paine et al., 1998) is often used to describe the changes in a biological response variable that

contrast those anticipated when multiple stressors interact (e.g. Christensen et al., 2006; Jackson et al., 2016). Although an ecological surprise may be defined as an interaction that is either greater than, or less than, the expected magnitude from a null model, particular focus has been on interactions of stressors which interact synergistically; that is, where the combined effect is greater than the sum of the individual effects. Synergistic interactions of multiple stressors are important to document, firstly due to their potential to have a dramatic effect on ecological communities, and secondly because the presence of a synergistic interaction means management strategies can potentially have a large effect by mitigating against just one of the interacting stressors (Brown et al., 2013; Côté et al., 2016; Haller-Bull & Bode, 2019). Because of their potential impact, there has been a great deal of effort in recording the frequency of synergy in stressors across different ecosystems and communities (Côté et al., 2016). However, there is always a danger that an emphasis on their importance could lead to over-estimating the frequency of synergisms or other forms of ecological surprise (e.g. antagonisms) within the multiple stressor literature and, as highlighted by Côté et al. (2016), there is little evidence to suggest that stressors predominately interact in a synergistic manner. A pertinent question which has yet to be addressed is whether ecological surprises should be expected, or whether the prevalence of these interactions are skewed in some way by reporting biases, statistical sampling or both.

There is relatively little ecological theory that generates expectations of when and how often the cumulative effects of pairs of stressors should be synergistic, or indeed any other type of interaction. This is in contrast to other ecological interactions, such as the effects of multiple predators on prey density and biomass, where a much richer body of theoretical knowledge that has been used to generate a number of hypotheses for testing (Schmitz, 2007; Sih et al., 1998). Instead, progress on ecosystem stressor interactions has been made largely by meta-analyses across a number of experiments, realms, trophic levels, measured traits, taxonomic groups and stressor types (e.g. Crain et al., 2008; Darling & Côté, 2008; Jackson et al., 2016; Wu et al., 2011). Within ecosystem stressor research, the most popular approach is to use the additive null model where the stressor interaction is predicted to be simply the sum of their individual effects (e.g. Crain et al., 2008; Darling & Côté, 2008; Jackson et al., 2016; Strain et al., 2014), although the multiplicative null model, the log-transformed version of the additive model, is also relatively common (e.g. Bancroft et al., 2008; Gruner et al., 2008; Harvey et al., 2013; Rosenblatt & Schmitz, 2014). These null models classify interactions as either being null (i.e. the additive or multiplicative effect of interacting stressors), synergisms (i.e. greater than the null) or antagonisms (i.e. less than the null). While distinctions are increasingly being made for various forms of antagonistic interactions in this simple scheme (e.g. Jackson et al., 2016), there exists a range of other classification schemes (Orr et al., 2020), and these have been implemented across a number of studies (e.g. Piggott et al., 2015; Travers-Trolet et al., 2014). The profusion of null models can make it difficult to generalize results across different studies. A 'synergistic' or 'antagonistic' interaction may have contrasting

definitions depending on the scheme being used leading to the same interactions being labelled differently under contrasting schemes; hence the biological and statistical interpretation is therefore dependent on the null model being applied. One way round this issue is to pool published data together to harness increased statistical power and conduct a meta-analysis to search for generalities under a particular null model (examples listed in Côté et al., 2016). However, despite their potential, these meta-analyses have to date not identified any general covariates capable of explaining the broad patterns of multiple stressor interactions, meaning we still lack general predictions of the consequences of multiple stressors (Côté et al., 2016).

Given the lack of consistent generalities from empirical studies, the development of ecological theory within multiple stressor research may represent an approach capable of providing novel insights. Some theory has been developed for particular case studies (e.g. Brown et al., 2013; Galic et al., 2018), but only a few studies (e.g. Haller-Bull & Bode, 2019) have so far investigated more general insights. Of primary interest is the generation of theory which can provide a mechanistic underpinning to the field, and potentially allow for an increased understanding of multiple stressor interactions, compared to that provided solely by a null model approach (De Laender, 2018). However, theory could also be used to better understand the results obtained from the null model approach to empirical classification of stressor interactions. In particular, we know of no study that has investigated how robust the null models are to noisy data (i.e. sampling uncertainty and/or process variation); yet understanding this is important before we can draw strong conclusions from the empirical analyses. This knowledge is also important for evaluating the relative performances of the profusion of null models, and is therefore something which may help guide the end-user to decide which null model may be both appropriate and likely to yield important results in the face of what is often noisy and/or limited data.

Here, we begin to close these gaps in understanding by testing for the prevalence of non-additive effects of co-occurring pairs of stressors in freshwater ecosystems. We first develop classical community ecology models based on Lotka–Volterra consumer-resource dynamics in order to simulate data from biologically simple food webs impacted by pairs of stressors. This provides us with 'data' where we know the underlying stressor-pair interactions. We then use these simulated data to investigate the ability of the additive null model to recover interactions under a range of different levels of data uncertainty which we model as observation error. With a better understanding of the statistical null model, we then review the experimental literature to compile and analyse the largest (in terms of the number of interactions) dataset for the effects of co-occurring stressor interactions on the biomasses and densities of freshwater organisms. In particular, we ask whether ecological surprises are common in freshwater stressor interactions. The simulation experiments allow considerable insights into our empirical analyses and help prevent over-interpretation of our results.

2 | MATERIALS AND METHODS

2.1 | Theoretical models

In order to provide a better understanding of the empirical results that follow, we built food chain models based on the classical Lotka-Volterra consumer resource equations (Heath et al., 2014). We chose these models since we believe stressors may act directly on population- and trophic-level patterns, but also indirectly via trophic cascades (e.g. a species may be indirectly affected if its primary resource is directly affected by a stressor). This approach is also broadly in line with our empirical data which focuses on population and community-level metrics as the responses to stressor treatments (see below). To increase the robustness of our conclusions, we considered two forms of the model; one where (within trophic level) density dependence affects the death rates of each trophic level, and the other where consumer uptake is density regulated (Table 1). Both these scenarios were analysed by Heath et al. (2014) to investigate the roles of different types of density dependence on trophic cascades (see details therein). In both models, the basal level of the chain describes the dynamics of a key nutrient that limits the productivity of the food chain, and we assumed nutrients are added

TABLE 1 Equations used to establish theoretical food chains. The equations, sets and a brief description of the equivalent ecological trophic are shown

	Equation type	Equation	Description
1a	Density dependence	$\frac{dx_n}{dt} = \alpha_n \varepsilon_n x_{n-1} x_n - \delta_n x_n - \lambda_n x_n^2$	Change in density of apex consumer (x_n)
1b	Density dependence	$\frac{dx_i}{dt} = \alpha_i \varepsilon_i x_{i-1} x_i - \alpha_{i+1} x_i x_{i+1} - \delta_i x_i - \lambda_i x_i^2$	Change in density of non-apex consumer (x_i)
1c	Density dependence	$\frac{dx_0}{dt} = \omega - \alpha_1 x_0 x_1$	Change in density of nutrients (x_0)
2a	Consumer uptake regulation	$\frac{dx_n}{dt} = \frac{\alpha_n \varepsilon_n x_{n-1} x_n}{1 + \nu_n x_n} - \delta_n x_n$	Change in density of apex consumer (x_n)
2b	Consumer uptake regulation	$\frac{dx_i}{dt} = \frac{\alpha_i \varepsilon_i x_{i-1} x_i}{1 + \nu_i x_i} - \frac{\alpha_{i+1} x_i x_{i+1}}{1 + \nu_{i+1} x_{i+1}} - \delta_i x_i$	Change in density of non-apex consumer (x_i)
2c	Consumer uptake regulation	$\frac{dx_0}{dt} = \omega - \frac{\alpha_1 x_0 x_1}{1 + \nu_1 x_1}$	Change in density of nutrients (x_0)

at a constant rate, ω (Heath et al., 2014). Each subsequent equation then describes a different type of consumer. The first level is wholly dependent on the nutrients and may represent a primary producer such as an algal species that requires a key mineral such as silica. The second level consumes the first trophic level and is in turn consumed by a third trophic level, and so on until the apex consumer is reached. In the density dependence model (Equation 1; Table 1), the consumer (trophic level i) exploits the resource (trophic level $i - 1$) with a constant consumption/attack rate, α_i, and the conversion efficiency parameter, ε_i, determines the proportion of the consumed resource that is converted into new consumers (Heath et al., 2014). Under density dependence, the consumer is self-regulated by the intraspecific density dependence parameter λ_i, which leads to an increase in death rate as the consumer density increases (Heath et al., 2014). In contrast, the consumer uptake regulation model (Equation 2; Table 1), assumes the effect of increasing consumers is to slow down the consumption of the resource, perhaps due to increased interference (Heath et al., 2014). In this case, the parameter v_i, determines the consumer density at which the maximum per capita uptake rate is halved, defined as the density $x_i = 1/v_i$ (Heath et al., 2014).

Using these equations, we established food chains comprising either three, four or five trophic levels, and the equation for each trophic level models how the biomass or density changed over time. For simplicity, we assumed all key parameters (nutrient input ω; consumption rates α_i; conversion efficiencies ε_i; uptake regulators v_i; density independent δ_i, and dependent death rates λ_i, for trophic level i) do not vary over time, and we investigated the effect of stressors on equilibrium biomasses/densities. The models also ignore spatial structure in the community, which also remain closed to immigration from outside apart from the constant input of the nutrient. Hence these models represent the simplest form of community dynamics that can be used to investigate the effects of multiple stressors as well as the manner in which they interact.

Stressors to the food chains were modelled by changing the values for parameters and comparing the resultant equilibrium densities across all trophic levels to the equilibria for a set of baseline parameter values. Equations 1 and 2 are not mechanistic models for specific stressors (e.g. pollution, temperature) but instead capture the net effect of stressors on the ecological processes of the food web species. For simplicity, we assumed each stressor had either a positive or negative effect on one model parameter (i.e. ω, α_i, ε_i or δ_i), and we investigated how pairs of stressors interact to affect community densities. The baseline parameters were drawn from uniform distributions with ranges given in Table 2. Therefore, for a given food chain, the baseline parameters for all trophic levels were independently sampled from the distribution of values given in Table 2. Similarly, the processes (parameters) affected by each stressor were randomly selected from the possible candidates, and the intensity of its effect on the baseline rate was drawn from a uniform distribution with the ranges shown in Table 2. The baseline parameter set therefore represented the control community, and as in experimental studies that employ the factorial design approach (e.g. Davis et al., 2018; Matthaei et al., 2010), we manipulated our model communities by investigating the effect of each stressor acting alone, as well as the stressors acting in combination. From these cases, we then computed the type of stressor interaction and how they combined to alter the community densities (see below for definitions of how stressor interactions are computed). We therefore chose one trophic level at random from the entire food chain, excluding the nutrient level. We focused on this population/trophic level and mirrored it in our selection of empirical data (see below). This also means the species or trophic levels under scrutiny were not always directly affected by the stressor but could be affected solely due to a trophic cascade effect. It is also important to note that a stressor could have led to either an increase or a decrease in parameter value relative to the baseline; and that multiple stressors could have acted

TABLE 2 Explanation of the different parameters within Equations (1) and (2), with the mechanism they reflect, alongside the minimum and maximum values for the ranges of baseline parameter values. Parameter values were drawn from a uniform distribution U~(a, b) with lower limit, a, and upper limit, b, with the limits differing between the baseline and stressed parameters. The method for determining stressed parameter values is detailed in Supplementary Information S1

Parameter	Ecological mechanism	Baseline value range	Stressed value range
α	The rate at which a trophic level predates upon the trophic level directly below.	α_b = U~(0.25, 0.75)	U~(0.01, 0.99)
ε	The efficiency at which a trophic level can transform consumed matter into new individuals.	ε_b = U~(0.25, 0.75)	U~(0.01, 0.99)
δ	The density independent mortality rate of a trophic level.	δ_b = U~(0.25, 0.75)	U~(0.01, 0.99)
ω	The constant rate at which a resource (x_0) is input into the food chain.	ω_b = U~(25, 75)	U~(1, 99)
λ	The density dependent mortality rate of a trophic level.	λ_b = U~(0.00625, 0.025)	The parameter was not under selection for alteration by a stressor
v	A limit to the uptake rate of a consumer through a trait-mediated response that may be behavioural or otherwise.	v_b = U~(0.05, 0.15)	The parameter was not under selection for alteration by a stressor

on the same, or a different trophic level. We chose to model the scenario where each stressor affected only one parameter (and therefore one biological process); hence within our model communities, stressors did not interact at the parameter level. However, relaxing this assumption to allow two stressors to affect a single process (parameter) did not alter our results (Supplementary Information S1).

Overall, 1,320,000 different combinations, of equations, food chain lengths, stressors pairs and randomly selected baseline values were generated. Equilibrium densities, for each of these combinations, were calculated using Mathematica 10.4 (Wolfram Research & Inc., 2016). We only considered cases where the equilibria were all stable, and feasible (i.e. all densities were positive), and only equilibrium densities for trophic levels x_1 and above were included in the stressor interaction results (i.e. we excluded the nutrient level from our stressor interaction analyses). Stability was assessed by determining the Jacobian matrix for each community and calculating the corresponding eigenvalues. For every community, all eigenvalues had a negative real part with the equilibria being point attractors.

Across all 1,320,000 combinations, 79.9% of the parameter sets resulted in the determination of equilibrium densities that were both stable and feasible, with the discarded 20.1% parameter sets resulting in at least one biologically unfeasible density. From the full set of stable and feasible communities, we randomly selected 360,000 theoretical interactions, and for each community, we randomly selected a single trophic level for the focus of our estimation of the stressor interaction. All subsequent analyses of the theoretical data were performed on this group of 360,000 theoretical interactions. This subsetting was required as there was a negative relationship between the number of trophic levels and the likelihood of the community being both stable and feasible, which biased the full dataset towards communities with only three trophic levels. The final 360,000 stressor interactions were selected with weighted probabilities to ensure approximately one third (i.e. ~120,000) were from each of the three food chain lengths, and that each model (Table 1) was also approximately equally represented.

Unlike the empirical studies used in the meta-analyses below, the food chain models are purely deterministic, meaning that there are no random fluctuations around the equilibrium densities. In effect, for any given pair of stressors, there is no uncertainty (observation error) in the theoretical data. Clearly, this differs from the empirical data where observation error leads to an estimate of the densities/biomasses under investigation in the control and treatment replicate communities, and this observation error may lead to some stressor interactions being misclassified. For a better comparison with the empirical data, and to test the robustness of the additive null model to observation error, we modelled observation error by taking the 360,000 theoretical interactions from our original analyses and then multiplying the density of each trophic level by a random number drawn from a Gaussian distribution with a mean of 1.00 and standard deviation of σ. This process was repeated between three and six times for each treatment, analogous to the number of replicates per

treatment found in our empirical data (see below). Thus, larger values for σ led to larger deviations around the equilibrium biomasses, and therefore a larger observation error, with an increased likelihood that the stressor interaction was misclassified. Standard deviations, σ, were from one of 50 different levels, ranging from 1×10^{-6} to 5×10^{-1}, in consistent logarithmic increments (e.g. 8×10^{-6}, 9×10^{-6}, 1×10^{-5}, 2×10^{-5}, etc.). The interpretation of σ is straightforward, as we would expect 99.7% of all observations to fall within 3σ's of the 'true' stressor effect (i.e. the biomass/density in the absence of any observation error). Supplementary Information S1 details a complete overview of how observation error was incorporated into the theoretical data.

2.2 | Collation of empirical data

Through Web of Science we searched the primary scientific literature, for papers published before 1st January 2019, which investigated the impacts of multiple stressors on freshwater communities. In order to be incorporated, papers needed to report results where there was a factorial design, namely: (i) a control (without stressors); (ii) each stressor acting individually; and (iii) the stressors acting simultaneously. We required papers to report the mean value of the response, the number of replicates, and standard deviation or standard error for each treatment in the factorial design; failure to report any of this information led to the study being excluded from our analysis. Additionally, papers were required to report at least one of the following untransformed metrics: biomass, abundance, density or chlorophyll-*a* of one or more groups of organisms within the stressed community. Hence, and in line with our trophic models, the focus of our effort was directed towards studies that report the effects of stressors acting at the population and community levels. Papers often reported the impacts of stressors on multiple different groups of organisms within a community; when this occurred, the responses of all different groups of organisms were included within the overall dataset. The different groups of organisms could comprise: populations of a single species (e.g. *Daphnia pulex*); a group of organisms within the same feeding guild (e.g. herbivores); a group of taxonomically similar organisms (e.g. taxa within the genera *Ephemeroptera*, *Plecoptera* and *Trichoptera*); or a group of similar organisms (e.g. macroinvertebrates or algae). To be included within our dataset, papers had to investigate communities comprising a minimum of two different groups of organisms. Studies investigated a wide range of different stressors, although these were subsequently grouped into broader stressor categories, such as temperature, contamination and habitat alteration.

Previous analyses have frequently focused on collating data for only the greatest single intensity of a stressor (e.g. Jackson et al., 2016). In contrast, where studies reported the responses of communities to multiple intensities of different stressors, data for all of the different intensities were collated. All interactions considering the different intensities of stressors were included in the overall dataset, although covariation in data due to repeated experiments across

different stressor intensities was accounted for in the final meta-analyses (see section *below*).

Some studies reported multiple different response metrics for the same group of organisms, included the same species within multiple different groups, or reported data for the same experiment over multiple different time points. Accordingly, in order to reduce the correlation/covariance within the overall dataset, these interactions were removed from our analyses. For instance, where the effects of stressor interactions on multiple different traits were reported, those considering density as the biological response metric were prioritized over abundances, which were in turn prioritized over biomasses, or those considering chlorophyll-a, respectively. Similarly, where papers reported data for stressor interactions over multiple different time points, only the final time point was used as this best matched our equilibrium assumption for the theoretical models.

Supplementary Information S2 gives a complete overview of the different search terms used to find studies, the method used to determine whether the data for a study could be collated, the processes for extracting and collating the data, and the process for removing interactions to prevent covariance.

2.3 | The determination of effect sizes and the classification of interactions

Across both the theoretical and empirical datasets, we used the same method to determine the classification of an interaction, using the factorial form of the effect size metric, Hedges' *d* (Gurevitch et al., 2000). Hedges' *d* is frequently used to investigate the impacts of multiple stressors as it estimates the standardized mean difference between the means of stressed and control samples and is unbiased

by small sample sizes (Hedges & Olkin, 1985). It is calculated by comparing the effect of the interaction on ecological communities to the sum of effects of the stressors acting individually; namely, an additive null model. In line with current methods, we inverted the sign of the interactions when the expected effect of the additive null model was negative (Piggott et al., 2015). This method allowed for interaction effect sizes to be compared regardless of their directionality. We therefore focused on the classification of the interaction as opposed to the absolute magnitude/polarity of the effects. Supplementary Information S3 gives a complete breakdown of the equations used for calculating Hedges' *d*.

Once Hedges' *d* for a given interaction of stressors was calculated, we then classified the interaction into one of four types as illustrated in Figure 1 and following the convention of Jackson et al. (2016). In brief, the four interaction classifications were: (i) *Additive*, where the effect of the additive null model was statistically indistinguishable from the effect of observed interaction; (ii) *Synergistic*, where the observed interaction effect was greater than the effect of the additive null model; (iii) *Antagonistic*, where the observed interaction effect was less than the effect of the additive null model, but both effects had the same polarity; and (iv) *Reversal*, where the observed interaction effect was less than the effect of the additive null model, but the observed and expected effects had contrasting polarities. The distinction between antagonistic and reversal interactions is relatively recent (e.g. Jackson et al., 2016; Travers-Trolet et al., 2014), with most research still continuing to use the appellation of antagonistic to refer to both antagonistic and reversal interactions (e.g. Gomez Isaza et al., 2020; Velasco et al., 2019).

In our method, if Hedges' *d* was positive, the interaction was classed as synergistic, and if negative, the interaction was classed as either an antagonistic or reversal interaction, although this could only
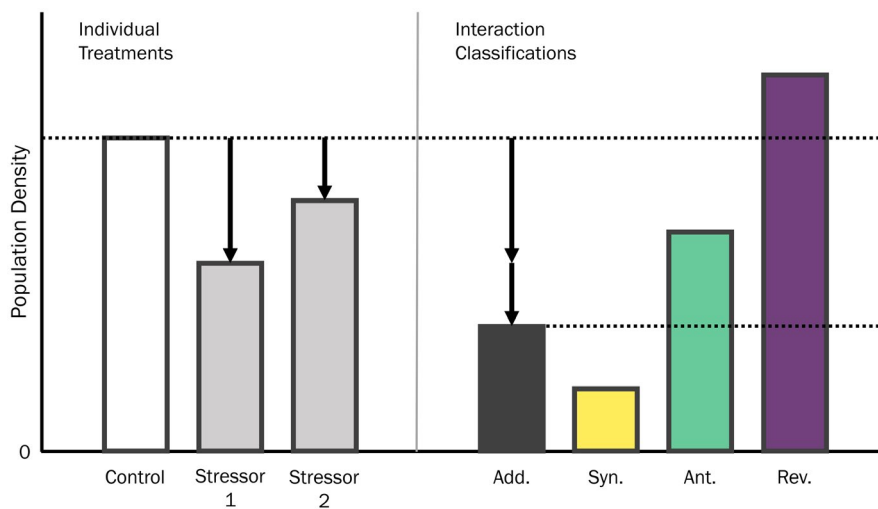

FIGURE 1 Graphical depiction of interaction types using population density as a response metric. White and grey bars denote densities under control and single stressors, respectively. The black bar denotes the additive (Add.) interaction classification (i.e. the sum of the effects for the individual stressors shown by the black arrows). The yellow bar denotes a synergistic (Syn.) interaction classification (i.e. a decrease in population density greater than the additive effect). The green bar denotes an antagonistic (Ant.) interaction classification (i.e. a decrease in population density less than the additive effect). The purple bar denotes a reversal (Rev.) interaction classification (i.e. a change in population density in the opposite direction to that of the additive effect)

be determined by comparing the effect of the additive null model to the observed effect (as outlined above). Each value of Hedges' *d* had corresponding 95% confidence intervals; if these confidence intervals incorporated 0, then an interaction was deemed to be indistinguishable from additive. The classification scheme outlined above is one of a number of possible choices (e.g. Crain et al., 2008; Jackson et al., 2016), and Supplementary Information S4 details a comparison of how these different schemes contrast each other.

2.4 | Vote-counting

Following the classification of all interactions, we implemented a vote-counting method to determine the relative proportions of the interaction classes across both the theoretical and empirical datasets. To consider the effect of different strengths of observation error on the ability to detect the 'true' stressor interaction in the modelled data, we computed the frequency of interaction types for each level of observation error investigated.

2.5 | Summary effect sizes

Alongside the vote-counting method, we calculated summary effect sizes across both the theoretical and empirical datasets. The calculation of summary effect sizes represents one of the key components that defines a formal meta-analysis (Koricheva & Gurevitch, 2014), allowing for the collation of the individual effect sizes of multiple independent experiments or studies and determining a single summary effect. Pooling the data in this way increases the statistical power of our analyses, and therefore leads to a greater probability of correctly rejecting the null hypothesis that stressor interactions are additive. Meta-analyses and summary effect sizes are both useful and well-established within the field of multiple stressors (e.g. Crain et al., 2008; Jackson et al., 2016; Seifert et al., 2020), and give higher weightings to individual effect sizes with lower uncertainties (i.e. lower variances) which lead to more precise estimates of the overall summary effect size (Koricheva & Gurevitch, 2014).

For the empirical analysis, summary effect sizes were determined by using a weighted random effect model and implemented in the *metafor* package (Viechtbauer, 2010) in R. Random effects were specified as being the identity (ID) of the study group of organisms nested within the ID for study. The random effects were specified in order to account for both within- and between-study variation contained within the empirical dataset. Additionally, some empirical studies considered multiple intensities of one or more stressors, and as such, calculations of the interaction class for each intensity of stressor used the same control. To account for any covariance between the different intensities of a single stressor, we incorporated covariance-variance matrices within the meta-analytical models. For the empirical dataset, mixed effect models were also conducted with the fixed effects of stressor pair or organism group, alongside the previously described random effects

(see Supplementary Information S5). The summary effect size for the theoretical dataset was also determined using a similar process. However, due to computational limitations caused by the number of interactions under analysis (360,000 interactions at each level of observation error), fixed effect models for the theoretical data were fitted using the *lm* function. The models applied to both the theoretical and empirical datasets are explained in further detail within Supplementary Information S5. While we detail the results of both the vote-counting and summary effect size methods, our results primarily focus on summary effect sizes, in line with recommendations for meta-analyses (Gurevitch et al., 2018).
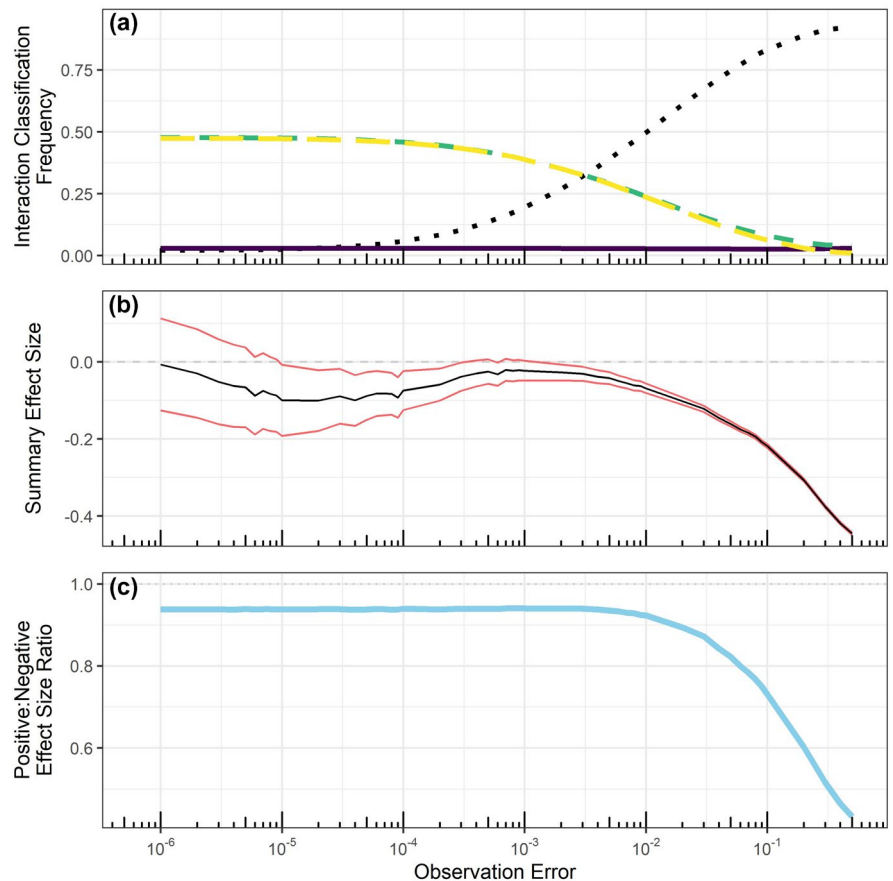
The overall effect from a meta-analysis needs to be checked for consistency among effect sizes, termed as heterogeneity (Nakagawa et al., 2017). We used the I^2 statistic, which is bounded between 0% and 100%, with 25%, 50% and 75% being suggested as levels for, low, medium and high heterogeneity, respectively (Higgins et al., 2003). Ecological meta-analyses often report high levels of heterogeneity (Senior et al., 2016), perhaps due to the variation in study organisms common to the questions being addressed, and we may have expected a high value here due to both range of study organism and range of stressor type. To explore the potential causes of heterogeneity within the empirical meta-analysis, we conducted separate meta-analyses on a sub-group of the dataset, a similar process to running a meta-regression (Nakagawa et al., 2017), using organism group (i.e. producer or consumer) as the categorical moderator to explore heterogeneity (see Supplementary Information S6). We also considered publication bias (see Supplementary Information S6); although it should be noted that common tests for publication bias within meta-analyses can be limited by high heterogeneity (Nakagawa et al., 2017).

3 | RESULTS

3.1 | Stressor interactions within theoretical data

We found no strong difference between the classification of stressor interactions from either form of food chain model (Table 1), or between the different lengths (three, four, five trophic levels) of food chains (see Supplementary Information S1), indicating the frequencies of interaction classifications were robust to these details of the models. For the entire theoretical dataset of 360,000 stressor interactions (comprising both Consumer Uptake Regulation and Density Dependence Equations, and across food chains of three, four and five levels) without observation error, antagonistic and synergistic interactions were the most frequently assigned (0.483 and 0.480, respectively), followed by reversal (0.0288), and finally additive interactions (0.00856). These frequencies represent the 'true' interactions as classified by the additive null model, under no data uncertainty. However, the ability of the additive null model to recover these interaction frequencies is very sensitive to observation error. Increasing observation error led to more interactions being classified as additive

FIGURE 2 The effect of observation error (σ) on the stressor interaction categorization, and summary meta-analytic effect sizes in the theoretical data. (a) Frequency of the different interaction classes for the 360,000 theoretical stressor interactions at each level of observation error: Dotted black line denotes additive interactions; green short-dashed line indicates antagonistic interactions; yellow long-dashed line denotes synergistic interactions; and purple line indicates reversal interactions. (b) Summary effect sizes for the 360,000 theoretical stressor interactions, at each level of observation error - Black line denotes summary effect sizes; and red lines denote 95% confidence intervals. (c) The ratio of positive to negative summary effect sizes at each level of observation error

(the null model) and at higher levels additive interactions were clearly dominant (Figure 2a). This pattern could be generated if our theoretical interactions only weakly deviated from additivity, but checks confirmed that this was not the case, and that over 50% of interactions deviated from additivity by more than 5% (see Supplementary Information S1).

The summary effect size, and summary interaction class as generated from the meta-analytical framework also showed high sensitivity to observation error, although in these analyses the outcome was rather different (Figure 2b). For low levels of observation error, the 95% confidence intervals of the summary effect size overlapped zero, indicative of an additive summary interaction class. This occurred because the frequency and magnitudes of synergistic (positive effect sizes) and antagonistic/reversal (negative effect sizes) interactions were approximately equal for low observation error (Figure 2a), effectively cancelling one another out, and the large confidence intervals were caused by the underlying large variance in effect sizes (see Supplementary Information S1). However, with increasing observation error, the summary effect sizes became increasingly more negative, and confidence intervals for these summary effect sizes did not overlap zero, indicating an antagonistic/reversal summary interaction class. This result may seem surprising since, similar to the case of individual interactions (Figure 2a), we may expect increased observation error to lead to summary effects with larger confidence intervals that overlapped zero. Instead, we found that, although the 'true' stressor interactions (i.e. in the absence of observation error)

were roughly equally divided between synergy and antagonism, the summary effect became increasingly more negative as observation error increased, indicating observation error affected synergistic and antagonistic interactions asymmetrically. Further inspection showed an increase in the proportion of negative effect sizes as observation error increased (Figure 2c), with this being mirrored by a decreasing summary effect size (Figure 2b). Although not so obvious due to the dominance of additive interactions, a similar trend could be observed in the frequencies of interaction classes at higher observation errors, with synergistic interactions heading towards 0 frequency faster than antagonistic interactions (Figure 2a). Hence, analyses of our model results with varying levels of observation error suggested synergies in pairs of ecosystem stressors may be under-reported in many empirical studies.

3.2 | Theoretical expectations

In summary, our theoretical analyses led us to conclude that at biologically plausible levels of observation error (i.e. >0.01), we should expect (i) the empirical data to be dominated by additive interactions for individual interactions (Figure 2a), but (ii) in contrast, the summary effect sizes computed across a large body of such studies may indicate a dominant role for antagonistic, or reversal, interactions. Both of these results may occur even if, as in our simulated data, synergies are common for the 'true' interaction classifications.

3.3 | Stressor interactions within freshwater empirical data

Our literature search within the Web of Science yielded 1805 papers that met our search criteria, 58 of which met our criteria for inclusion. They included 545 interactions summarized in Figure 3 to illustrate the frequency of different interaction classifications and the overall summary effect sizes and interaction classes. Additive interactions were the most frequent (0.829), followed by antagonistic (0.0991), reversal (0.0477) and finally synergistic (0.0239) interactions (Figure 3a). Additionally, the summary effect size for the entire dataset was negative with 95% confidence intervals that did not overlap zero (−0.632 ± 0.260), indicative of an antagonistic/reversal summary interaction class (Figure 3b).

Our meta-analysis showed medium-level heterogeneity (I^2 = 48.5%) although this was considerably lower than the mean heterogeneity (I^2 = 91.7%) found in an analysis of previous ecological meta-analyses (Senior et al., 2016). Furthermore, two additional
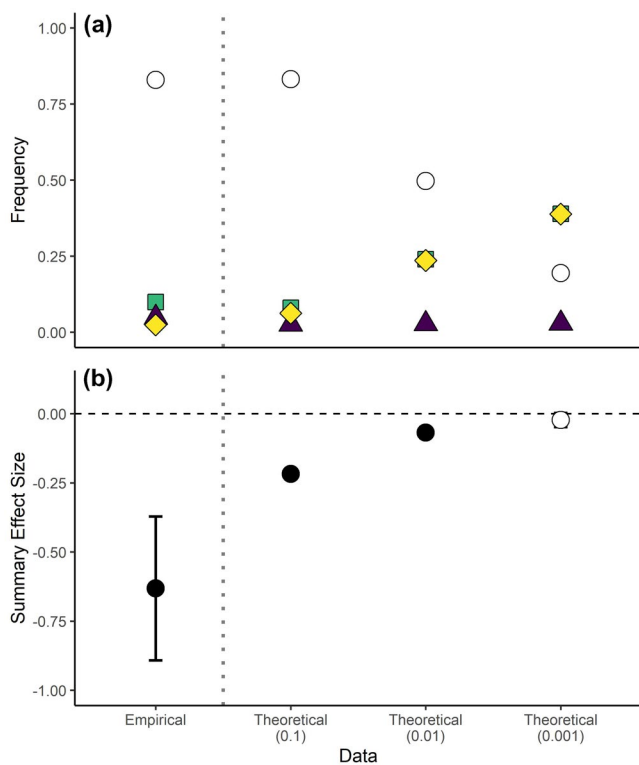
meta-analyses were conducted on sub-groups of the empirical dataset, with the categorical moderator of organism group used to explore this heterogeneity (Nakagawa et al., 2017). However, these additional meta-analyses failed to uncover any source of this heterogeneity, with the meta-analysis for consumers reporting medium-level heterogeneity (I^2 = 42.5%) and the producer meta-analysis reporting high-level heterogeneity (I^2 = 67.7%; see Supplementary Information S6).

3.4 | Comparison of empirical and theoretical interaction classifications

Overall, we found close agreement between our theoretical models with biologically reasonable levels of observation error and the freshwater empirical data (Figure 3). Vote-counting results highlight how individual interactions tended to return an additive classification in the empirical dataset, and that this is expected in the theoretical data when estimates of metrics used to classify the interactions are mostly within 10% of the true value (Figure 3a). Similarly, summary effect sizes were negative in the simulated data under even a very small level of observation error (Figure 3b) despite synergies and antagonisms co-dominating the 'true' interactions. This implies the summary effect size reported in the empirical data for freshwater communities (Figure 3b) may not necessarily be representative of the underlying 'true' stressor interactions.

4 | DISCUSSION

There has been much interest in understanding and cataloguing the joint effects of stressors on ecological communities and ecosystems (Côté et al., 2016; Schäfer & Piggott, 2018; Thompson et al., 2018b), but to date there has been relatively little guidance from ecological theory. Similarly, we know of no demonstration of the abilities of the statistical tools used for classifying interactions to recover known interactions in the face of data uncertainty. Here, our aim was to test the statistical tools used to define stressor interactions using data simulated from ecological theory in order to gain a deeper understanding of a freshwater dataset. Our empirical analyses generate two main results: (1) vote-counting analyses suggest additive interactions to be by far the most dominant stressor interaction types in freshwater community experiments (Figure 3a); but (2) our meta-analysis shows antagonism to be the summary interaction class (Figure 3b). However, the analyses of the simulated data suggest both results should be expected under plausible levels of observation error (i.e. >0.01) in the data (Figure 3), and that only under unrealistically low levels of precision should we expect to recover the 'true' stressor-pair interactions in either individual studies or meta-analyses (Figure 2). We believe that once these statistical aspects are considered, the so-called 'ecological surprises' (sensu Paine et al., 1998) may in fact be more prevalent in both our freshwater dataset, and more widely.

FIGURE 3 Comparisons of the analyses of the freshwater stressor interaction dataset (545 interactions) with the full theoretical stressor interaction dataset for different levels of observation error, σ (given in parentheses on the *x*-axis). Comparisons are for: (a) proportions of the different interaction classes; and (b) summary effect sizes for the empirical and theoretical dataset. Figure 3a; white circles denote additive interactions, green squares denote antagonistic interactions, yellow diamonds denote synergistic interactions, purple triangles denote reversal interactions. Figure 3b; closed circles denote significant summary effect sizes (i.e. 95% confidence intervals do not overlap zero), and open circles denote non-significant summary effect sizes (i.e. 95% confidence do overlap zero)

4.1 | Null model sensitivity to observation error

The choice of the null model is hotly debated within ecological stressor research (Schäfer & Piggott, 2018), and it has been argued that null models should be able to accurately predict the combined effects of stressors (Orr et al., 2020). Our results (Figure 2) are the first attempt to quantify the degree of accuracy for the most commonly used null model, and we conclude that for all but the very lowest levels of observation error it is difficult to correctly reject the additive null interaction (Figure 2a). In other words, we find weak statistical power to recover the underlying stressor-pair interactions. On this basis, and given that most experiments have low sample sizes (we report a mean of 3.83 with a maximum of 16 per treatment in our empirical data), we consider it premature to conclude that most stressor interactions are *truly* additive in the freshwater data we collected. Instead, we should be careful to conclude that in the majority of cases we do not have sufficient evidence to reject the null (additive) interaction. However, it means that we should take notice whenever a non-null interaction is returned by the additive model, since only strong non-additive effects are likely to be detected (see Supplementary Information S1).

Perhaps more surprising is our finding that meta-analyses using the additive null model report antagonism as the summary interaction classification when observation error is non-negligible, despite synergies co-dominating in our simulation data (Figure 2b). A naïve expectation would be for increased observation error to lead to summary effect sizes centred around zero with large confidence intervals making it difficult to rule out an additive summary interaction in our simulation data. Smaller confidence intervals at higher levels of observation error are easily explained by the effect sizes becoming more similar due to the high variances of the response metrics (Supplementary Information S1). Additionally, it is clear that observation error has an asymmetric effect on antagonisms and synergies, with this leading to a shift towards negative effect sizes dominating the distribution of simulated effect sizes (Figure 2c). Hence, although the return of an antagonistic summary interaction for our empirical dataset is mirrored in previous analyses of freshwater stressor experiments (Jackson et al., 2016; Lange et al., 2018), we cannot conclude that this is strong evidence for the dominance of antagonism in freshwater ecosystems. The simulation data therefore add valuable interpretation of our empirical data that would otherwise be missed, and in so-doing highlights the importance of benchmarking statistical tools against data with known attributes.

The high sensitivity to estimation uncertainty may be key reasons why stressor synergies are not as often reported as may be expected (Côté et al., 2016; Darling & Côté, 2008), although other reasons may also contribute, and we can also not rule out that the empirical results do truly reflect the underlying interactions. However, we believe our finding of high sensitivity to observation error in the null model is more general than either our theoretical results, or our freshwater dataset, and we suggest future studies should investigate other null models for their robustness to observation error and sample sizes. Such analyses would build on previous descriptions of the null models (e.g. Folt et al., 1999; Sih et al., 1998, 2004) and would be particularly useful if analyses considered the effect of sample size on statistical power, as this will help guide future empirical studies to improve the detection rate of non-null stressor interactions.

4.2 | Theoretical expectations for interaction frequencies

Our food chain models imply that, given adequate sample sizes (see above), we should expect synergistic and antagonistic interactions to co-dominate at the population and trophic levels, whereas additive interactions and reversals should be relatively rare. It may well be the case that our models are not good descriptors of the data we analyse; certainly, we ignore much important detail that is likely a feature in the data, such as spatial structure and temporal variation in parameters caused by external perturbations not linked to the stressors, and more complex food web structure involving omnivory or parasitism. Unfortunately, the null model sensitivity to observation error implies we do not yet have the tools with which to discern the relative abilities of different theoretical models to capture the empirical data. However, our key theoretical finding for the relative rarity of additive interactions appears to be echoed in the few other theoretical studies on stressor interactions in ecological communities (e.g. Haller-Bull & Bode, 2019; Thompson et al., 2018a; Travers-Trolet et al., 2014). This agreement is despite a variety of key differences in the model assumptions. In particular, Haller-Bull and Bode (2019) focused on populations rather than multispecies communities, but found dominant roles for synergistic and antagonistic interactions, with additive interactions occurring most frequently for stressors affecting the carrying capacity. Similar to our model, Thompson et al. (2018a) also focused on multispecies communities, but they assumed biological interactions were constant, whereas we allow interactions (consumption and conversion rates) to be modified by stressors, an assumption that seems likely to be met on a regular basis. For example, stressors have been shown to influence resource competition (Kroeker et al., 2013); susceptibility to parasitism in oysters (Lenihan et al., 1999); and modify the flow of energy through aquatic food webs by inducing changes in trophic links (Schrama et al., 2017). Despite this difference, Thompson et al. (2018a) found additive interactions were most prominent when species facilitated each other (i.e. positive species interactions), but that synergy or antagonism in combined stressor effects on species richness or community biomass were more common when species interactions are negative (competition or resource use).

The apparent rarity of additive interactions in all of these models may appear at odds with the possible interpretation that two stressors acting on different species within a community could

lead to an additive joint effect (Jackson et al., 2016). However, feedbacks in the food web, like those found in our models, mean that even if a species is unaffected directly by a stressor, it is highly likely that top–down or bottom–up effects will lead to indirect interactions for many species, and as a result, additive interactions are extremely unlikely in the absence of uncertainty (e.g. observation error). Indeed, we anticipate that additive interactions may only truly occur in scenarios where species in different and very weakly interacting sub-communities are affected by different stressors, or, as found by Thompson et al. (2018a), where species interactions are predominantly positive. We believe there will be an increasing role of theory in generating hypotheses for the ways in which stressors interact (De Laender, 2018), and the most progress will be made when the theory is developed so it can be directly compared against empirical data, much as we have done here.

4.3 | Mechanistic understanding of multiple stressors

In this study, we sought to address the question of *how* multiple stressors interact. This approach, when applied across both theoretical and empirical datasets can allow us to discern what may be expected across the interactions of multiple stressors. Future research may seek to address the question of *why* multiple stressors interact in the manner that they do. Undoubtedly, these two questions are entwined, with the answers to each of these questions highly likely to be dependent on the other. However, while the use of null models is essential in determining the combined effect of multiple stressors (Thompson et al., 2018b), the adoption of a mechanistic approach to investigating multiple stressors may provide novel insights which address these joint questions (De Laender, 2018; Schäfer & Piggott, 2018). For instance, a mechanistic understanding may allow for responses such as co-tolerance or co-susceptibility (Todgham & Stillman, 2013) to stressors to be more thoroughly understood from an ecological perspective. Ultimately, as our results imply, such an understanding is likely to require a large amount of empirical data to fully understand; however, there is ample scope for theoretical ecology to help fill this gap in our collective understanding of multiple stressors, and to generate specific hypotheses to be tested. Similarly, a mechanistic understanding of multiple stressor interactions would prove invaluable when mitigating the effects of stressors or implementing conservation initiatives.

4.4 | Future developments

Our analysis represents a novel approach combining both theoretical and empirical methods. While this analysis provides a solid foundation, there are several aspects that could be adjusted in future research. Firstly, there is a clear need to better understand the limitations and data requirements of the null models (e.g. Gurevitch et al., 2000; Lajeunesse, 2011; Thompson et al., 2018b) that are used to classify stressor interactions. Such knowledge would be very useful in guiding experimental design that would maximize the probability of uncovering non-null stressor interactions and would therefore provide a better understanding of their true prevalence. Knowing how many data points are required before we can realistically hope to detect a particular type of pattern, in this case a stressor interaction type of a given strength, is a critical component of experimental design. Moreover, our work has also uncovered some hitherto undescribed biases that lead to meta-analyses potentially over-emphasizing antagonisms, and it is important to investigate other null models for this feature as well as looking for methods to reduce this bias. Secondly, the theoretical communities manipulated here combine multiple populations each on a separate trophic level. While this builds upon similar research conducted on a single population (Brown et al., 2013; Haller-Bull & Bode, 2019), there is scope for this approach to be expanded to consider more complex communities, for instance with multiple populations on a single trophic level (e.g. Thompson et al., 2018a). Finally, the manner in which stressors interact at the parameter or process level can occur in numerous ways, for instance either additively or multiplicatively (Haller-Bull & Bode, 2019). However, whether a process or parameter is impacted in an additive or multiplicative manner, will cause a stressed parameter value to change by differing degrees, with this in turn potentially resulting in contrasting frequencies of interaction classifications at the population level. Accordingly, the manner in which a process or parameter (e.g. feeding rate, mortality) is impacted may be determined by the individual stressors; for instance, if two simultaneously acting stressors are entirely independent of one another then their effect on an ecological process may be additive (Haller-Bull & Bode, 2019). Consequently, allowing stressors to impact the same process undoubtedly represents an area for expansion, particularly when considering how impacts at the parameter level affect population level properties.

5 | CONCLUSIONS

Determining the ways multiple stressors interact is key when attempting to mitigate their effects, with the class of the observed interaction potentially outlining whether the removal of a stressor will have a beneficial, limited or detrimental impact to an ecosystem (Brown et al., 2013; Côté et al., 2016). Our results show the value of developing a theoretical framework which can aid in the interpretation of environmental stressor interactions, and we hope more general theory that makes specific predictions based on ecological mechanisms (e.g. De Laender, 2018; Fu et al., 2018; Thompson et al., 2018a) will be developed *and* tested in future. However, our results also highlight the urgent need to better understand the strengths and limitations of the null models that are used to classify the cumulative effects of community stressors,

and we also believe a unified approach to the meta-analyses of individual studies will increase our understanding of how environmental stressors combine.

AUTHOR CONTRIBUTION

B.J.B. and D.J.M. conceived the idea of the study; B.J.B. conducted the analysis; B.J.B. wrote the first draft and all authors contributed to the editing of the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary information of this article.

ORCID

Benjamin J. Burgess 🔟 https://orcid.org/0000-0003-0439-0061

REFERENCES

Ban, S. S., Graham, N. A., & Connolly, S. R. (2014). Evidence for multiple stressor interactions and effects on coral reefs. *Global Change Biology*, *20*(3), 681–697.

Bancroft, B. A., Baker, N. J., & Blaustein, A. R. (2008). A meta-analysis of the effects of ultraviolet B radiation and its synergistic interactions with pH, contaminants, and disease on amphibian survival. *Conservation Biology*, *22*(4), 987–996.

Birk, S., Chapman, D., Carvalho, L., Spears, B. M., Andersen, H. E., Argillier, C., Auer, S., Baattrup-Pedersen, A., Banin, L., Beklioğlu, M., Bondar-Kunze, E., Borja, A., Branco, P., Bucak, T., Buijse, A. D., Cardoso, A. C., Couture, R.-M., Cremona, F., de Zwart, D., … Hering, D. (2020). Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. *Nature Ecology & Evolution*, *4*(8), 1060–1068.

Brown, C. J., Saunders, M. I., Possingham, H. P., & Richardson, A. J. (2013). Managing for interactions between local and global stressors of ecosystems. *PLoS One*, *8*(6), e65765.

Christensen, M. R., Graham, M. D., Vinebrooke, R. D., Findlay, D. L., Paterson, M. J., & Turner, M. A. (2006). Multiple anthropogenic stressors cause ecological surprises in boreal lakes. *Global Change Biology*, *12*(12), 2316–2322.

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: Biological Sciences*, *283*(1824), 20152592.

Crain, C. M., Kroeker, K., & Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters*, *11*(12), 1304–1315.

Darling, E. S., & Côté, I. M. (2008). Quantifying the evidence for ecological synergies. *Ecology Letters*, *11*(12), 1278–1286.

Davis, S. J., Ó hUallacháin, D., Mellander, P.-E., Kelly, A.-M., Matthaei, C. D., Piggott, J. J., & Kelly-Quinn, M. (2018). Multiple-stressor effects of sediment, phosphorus and nitrogen on stream macroinvertebrate communities. *Science of the Total Environment*, *637*, 577–587.

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.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, *345*(6195), 401–406.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*(2), 163–182.

Folt, C. L., Chen, C. Y., Moore, M. V., & Burnaford, J. (1999). Synergism and antagonism among multiple stressors. *Limnology and Oceanography*, *44*(3part2), 864–877.

França, F. M., Benkwitt, C. E., Peralta, G., Robinson, J. P. W., Graham, N. A. J., Tylianakis, J. M., Berenguer, E., Lees, A. C., Ferreira, J., Louzada, J., & Barlow, J. (2020). Climatic and local stressor interactions threaten tropical forests and coral reefs. *Philosophical Transactions of the Royal Society B*, *375*(1794), 20190116.

Fu, C., Travers-Trolet, M., Velez, L., Grüss, A., Bundy, A., Shannon, L. J., Fulton, E. A., Akoglu, E., Houle, J. E., Coll, M., Verley, P., Heymans, J. J., John, E., & Shin, Y.-J. (2018). Risky business: the combined effects of fishing and changes in primary productivity on fish communities. *Ecological Modelling*, *368*, 265–276.

Galic, N., Sullivan, L. L., Grimm, V., & Forbes, V. E. (2018). When things don't add up: Quantifying impacts of multiple stressors from individual metabolism to ecosystem processing. *Ecology Letters*, *21*(4), 568–577.

Geldmann, J., Joppa, L. N., & Burgess, N. D. (2014). Mapping change in human pressure globally on land and within protected areas. *Conservation Biology*, *28*(6), 1604–1616.

Gomez Isaza, D. F., Cramp, R. L., & Franklin, C. E. (2020). Living in polluted waters: A meta-analysis of the effects of nitrate and interactions with other environmental stressors on freshwater taxa. *Environmental Pollution*, *261*, 114091.

Gruner, D. S., Smith, J. E., Seabloom, E. W., Sandin, S. A., Ngai, J. T., Hillebrand, H., Harpole, W. S., Elser, J. J., Cleland, E. E., Bracken, M. E. S., Borer, E. T., & Bolker, B. M. (2008). A cross-system synthesis of consumer and nutrient resource control on producer biomass. *Ecology Letters*, *11*(7), 740–755.

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*(7695), 175–182.

Gurevitch, J., Morrison, J. A., & Hedges, L. V. (2000). The interaction between competition and predation: A meta-analysis of field experiments. *The American Naturalist*, *155*(4), 435–453.

Haller-Bull, V., & Bode, M. (2019). Superadditive and subadditive dynamics are not inherent to the types of interacting threat. *PLoS One*, *14*(8), e0211444.

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., Lowndes, J. S., Rockwood, R. C., Selig, E. R., Selkoe, K. A., & Walbridge, S. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature Communications*, *6*(1), 1–7.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., Bruno, J. F., Casey, E. C., Fox, H., Fujita, R., Heinemann, D., Lenihan, H. S., Madin, E. M. P., Perry, M. T., Selig, E. R., Spalding, M. D., Steneck, R., & Watson, R. (2008). A global map of human impact on marine ecosystems. *Science*, *319*(5865), 948–952.

Harvey, B. P., Gwynn-Jones, D., & Moore, P. J. (2013). Meta-analysis reveals complex marine biological responses to the interactive effects of ocean acidification and warming. *Ecology and Evolution*, *3*(4), 1016–1030.

He, F., Zarfl, C., Bremerich, V., David, J. N. W., Hogan, Z., Kalinkat, G., Tockner, K., & Jähnig, S. C. (2019). The global decline of freshwater megafauna. *Global Change Biology*, *25*(11), 3883–3892.

Heath, M. R., Speirs, D. C., & Steele, J. H. (2014). Understanding patterns and processes in models of trophic cascades. *Ecology Letters*, 17(1), 101–114.

Hecky, R. E., Mugidde, R., Ramlal, P. S., Talbot, M. R., & Kling, G. W. (2010). Multiple stressors cause rapid ecosystem change in Lake Victoria. *Freshwater Biology*, 55, 19–42.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560.

Hodgson, E. E., & Halpern, B. S. (2019). Investigating cumulative effects across ecological scales. *Conservation Biology*, 33(1), 22–32.

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.

Koricheva, J., & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, 102(4), 828–844.

Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., Duarte, C. M., & Gattuso, J.-P. (2013). Impacts of ocean acidification on marine organisms: Quantifying sensitivities and interaction with warming. *Global Change Biology*, 19(6), 1884–1896.

Kroeker, K. J., Kordas, R. L., & Harley, C. D. (2017). Embracing interactions in ocean acidification research: Confronting multiple stressor scenarios and context dependence. *Biology Letters*, 13(3), 20160802.

Lajeunesse, M. J. (2011). On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology*, 92(11), 2049–2055.

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.

Lenihan, H. S., Micheli, F., Shelton, S. W., & Peterson, C. H. (1999). The influence of multiple environmental stressors on susceptibility to parasites: An experimental determination with oysters. *Limnology and Oceanography*, 44(3part2), 910–924.

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. *Proceedings of the National Academy of Sciences of the United States of America*, 111(26), 9549–9554.

Matthaei, C. D., Piggott, J. J., & Townsend, C. R. (2010). Multiple stressors in agricultural streams: Interactions among sediment addition, nutrient enrichment and water abstraction. *Journal of Applied Ecology*, 47(3), 639–649.

Nakagawa, S., Noble, D. W., Senior, A. M., & Lagisz, M. (2017). Meta-evaluation of meta-analysis: Ten appraisal questions for biologists. *BMC Biology*, 15(1), 1–14.

Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Lysenko, I., Senior, R. A., & Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520(7545), 45–50.

Ormerod, S. J., Dobson, M., Hildrew, A. G., & Townsend, C. (2010). Multiple stressors in freshwater ecosystems. *Freshwater Biology*, 55, 1–4.

Orr, J. A., Vinebrooke, R. D., Jackson, M. C., Kroeker, K. J., Kordas, R. L., Mantyka-Pringle, C., Van den Brink, P. J., De Laender, F., Stoks, R., Holmstrup, M., Matthaei, C. D., Monk, W. A., Penk, M. R., Leuzinger, S., Schäfer, R. B., & Piggott, J. J. (2020). Towards a unified study of multiple stressors: Divisions and common goals across research disciplines. *Proceedings of the Royal Society B*, 287(1926), 20200421.

Paine, R. T., Tegner, M. J., & Johnson, E. A. (1998). Compounded perturbations yield ecological surprises. *Ecosystems*, 1(6), 535–545.

Piggott, J. J., Townsend, C. R., & Matthaei, C. D. (2015). Reconceptualizing synergism and antagonism among multiple stressors. *Ecology and Evolution*, 5(7), 1538–1547.

Przeslawski, R., Byrne, M., & Mellin, C. (2015). A review and meta-analysis of the effects of multiple abiotic stressors on marine embryos and larvae. *Global Change Biology*, 21(6), 2122–2140.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873.

Rosenblatt, A. E., & Schmitz, O. J. (2014). Interactive effects of multiple climate change variables on trophic interactions: A meta-analysis. *Climate Change Responses*, 1(1), 1–10.

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. *Science*, 287(5459), 1770–1774.

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.

Scheffers, B. R., De Meester, L., Bridge, T. C. L., Hoffmann, A. A., Pandolfi, J. M., Corlett, R. T., Butchart, S. H. M., Pearce-Kelly, P., Kovacs, K. M., Dudgeon, D., Pacifici, M., Rondinini, C., Foden, W. B., Martin, T. G., Mora, C., Bickford, D., & Watson, J. E. M. (2016). The broad footprint of climate change from genes to biomes to people. *Science*, 354(6313).

Schmitz, O. J. (2007). Predator diversity and trophic interactions. *Ecology*, 88(10), 2415–2426.

Schrama, M., Barmentlo, S. H., Hunting, E. R., van Logtestijn, R. S., Vijver, M. G., & van Bodegom, P. M. (2017). Pressure-induced shifts in trophic linkages in a simplified aquatic food web. *Frontiers in Environmental Science*, 5, 75.

Seifert, M., Rost, B., Trimborn, S., & Hauck, J. (2020). Meta-analysis of multiple driver effects on marine phytoplankton highlights modulating role of pCO2. *Global Change Biology*, 26(12), 6787–6804.

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S., & Nakagawa, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology*, 97(12), 3293–3299.

Sih, A., Bell, A. M., & Kerby, J. L. (2004). Two stressors are far deadlier than one. *Trends in Ecology & Evolution*, 19(6), 274–276.

Sih, A., Englund, G., & Wooster, D. (1998). Emergent impacts of multiple predators on prey. *Trends in Ecology & Evolution*, 13(9), 350–355.

Strain, E. M., Thomson, R. J., Micheli, F., Mancuso, F. P., & Airoldi, L. (2014). Identifying the interacting roles of stressors in driving the global loss of canopy-forming to mat-forming algae in marine ecosystems. *Global Change Biology*, 20(11), 3300–3312.

Thompson, P. L., MacLennan, M. M., & Vinebrooke, R. D. (2018a). Species interactions cause non-additive effects of multiple environmental stressors on communities. *Ecosphere*, 9(11), e02518.

Thompson, P. L., MacLennan, M. M., & Vinebrooke, R. D. (2018b). An improved null model for assessing the net effects of multiple stressors on communities. *Global Change Biology*, 24(1), 517–525.

Tittensor, D. P., Walpole, M., Hill, S. L., Boyce, D. G., Britten, G. L., Burgess, N. D., & Ye, Y. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science*, 346(6206), 241–244.

Todgham, A. E., & Stillman, J. H. (2013). Physiological responses to shifts in multiple environmental stressors: relevance in a changing world. *Integrative and Comparative Biology*, 53(4), 539–544.

Travers-Trolet, M., Shin, Y. J., Shannon, L. J., Moloney, C. L., & Field, J. G. (2014). Combined fishing and climate forcing in the southern Benguela upwelling ecosystem: An end-to-end modelling approach reveals dampened effects. *PLoS One*, 9(4), e94286.

Velasco, J., Gutiérrez-Cánovas, C., Botella-Cruz, M., Sánchez-Fernández, D., Arribas, P., Carbonell, J. A., Millán, A., & Pallarés, S. (2019). Effects of salinity changes on aquatic organisms in a multiple stressor context. *Philosophical Transactions of the Royal Society B*, 374(1764), 20180011.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561.

Wolfram Research. Inc. (2016). *Mathematica (version 10.4)*. Wolfram Research Inc.

Woodward, G., Perkins, D. M., & Brown, L. E. (2010). Climate change and freshwater ecosystems: Impacts across multiple levels of organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1549), 2093–2106.

Wu, Z., Dijkstra, P., Koch, G. W., Peñuelas, J., & Hungate, B. A. (2011). Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. *Global Change Biology*, *17*(2), 927–942.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.