UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
INSTITUTO DE CIÊNCIAS BIOLÓGICAS (ICB)
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E EVOLUÇÃO

SARA LODI DE CARVALHO SPACEK

Meta-research in aquatic ecology

Goiânia
2021

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

[ ] Dissertação [x] Tese

2. Nome completo do autor

Sara Lodi de Carvalho Spacek

3. Título do trabalho

Meta-research in aquatic ecology

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento [x] SIM [ ] NÃO[1]

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:
a) consulta ao(à) autor(a) e ao(à) orientador(a);
b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
O documento não será disponibilizado durante o período de embargo.
Casos de embargo:
- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.

---

sei! assinatura eletrônica

Documento assinado eletronicamente por **Luis Mauricio Bini**, **Professor do Magistério Superior**, em 17/08/2021, às 10:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

---

sei! assinatura eletrônica

Documento assinado eletronicamente por **SARA LODI DE CARVALHO SPACEK**, **Discente**, em 17/08/2021, às 20:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

SARA LODI DE CARVALHO SPACEK

Meta-research in aquatic ecology

Tese apresentada ao Programa de Pós-graduação em Ecologia e Evolução do Instituto de Ciências Biológicas da Universidade federal de Goiás (UFG), como requisito para a obtenção do título de Doutor em Ecologia e Evolução.
Área de concentração: Ecologia e Evolução

Orientador: Professor Doutor Luis Mauricio Bini

Goiânia
2021

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

SPACEK, SARA LODI DE CARVALHO
Meta-research in aquatic ecology [manuscrito] / SARA LODI DE CARVALHO SPACEK, LUIS MAURICIO BINI. - 2021.
32250 f.

Orientador: Prof. Dr. LUIS MAURICIO BINI.
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Ciências Biológicas (ICB), Programa de Pós-Graduação em Ecologia e Evolução, Goiânia, 2021.
Apêndice.
Inclui gráfico, tabelas.

1. META-ANALISE. 2. REVISÃO SISTEMÁTICA. 3. USO DO SOLO. 4. VIÉS DE PUBLICAÇÃO. 5. RESULTADOS NEGATIVOS. I. BINI, LUIS MAURICIO. II. BINI, LUIS MAURICIO, orient. III. Título.

CDU 574

UNIVERSIDADE FEDERAL DE GOIÁS

INSTITUTO DE CIÊNCIAS BIOLÓGICAS

ATA DE DEFESA DE TESE

Ata Nº **106** da sessão de Defesa de Tese de **Sara Lodi de Carvalho Spacek** que confere o título de **Doutora em Ecologia e Evolução**, na área de concentração em **Ecologia e Evolução.**

Aos **trinta dias do mês de julho de dois mil e vinte e um (30/07/2021)**, a partir das **14h:00min**, por **videoconferência, seguindo portaria CAPES no. 36 de 16 de março de 2020 e recomendação da UFG**, realizou-se a sessão pública de Defesa de Tese intitulada **"Meta-research in aquatic ecology"**. Os trabalhos foram instalados pelo Orientador, **Prof. Dr. Luis Mauricio Bini (Departamento de Ecologia/ICB/UFG),** com a participação dos demais membros da Banca Examinadora: **Profa. Dra. Priscilla de Carvalho (Departamento de Ecologia/ICB/UFG)**, membro titular interno; **Profa. Dra. Jascieli Carla Bortolini (Departamento de Botânica/ICB/UFG)**, membro titular interno; **Profa. Dra. Thaísa Sala Michelan (ICB/UFPA)**, membro titular externo; **Prof. Dr. Ludgero Cardoso Galli Vieira (FUP/UnB)**, membro titular externo. Durante a argüição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido a candidata **aprovada** pelos seus membros. Proclamados os resultados pelo **Prof. Dr. Luis Mauricio Bini**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao(s) **trinta dias do mês de julho de dois mil e vinte e um (30/07/2021)**.

TÍTULO SUGERIDO PELA BANCA

Documento assinado eletronicamente por **Jascieli Carla Bortolini**, **Professora do Magistério Superior**, em 30/08/2021, às 13:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Luis Mauricio Bini**, **Professor do Magistério Superior**, em 30/08/2021, às 14:03, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Priscilla De Carvalho**, **Professor do Magistério Superior**, em 30/08/2021, às 15:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Thaísa Sala Michelan**, **Usuário Externo**, em 30/08/2021, às 17:56, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Ludgero Cardoso Galli Vieira**, **Usuário Externo**, em 30/08/2021, às 19:29, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Referência: Processo nº 23070.034458/2021-77 SEI nº 2313420

Summary

Agradecimentos

Seguir a carreira acadêmica sempre foi um sonho. Fazer este doutorado é algo que permaneceu nos meus planos por anos, até mais do que eu jamais poderia prever. O meu mestrado foi um momento maravilhoso, quando eu reafirmei que eu adoro trabalhar com pesquisa. Me propus esse desafio que é o doutorado, agora diferente do que foi o mestrado, tendo família e responsabilidades a mais. Afirmo em primeira mão que a realidade é outra. A cabeça é outra. A capacidade de focar unicamente em um tópico, e o trabalhar com profundidade é algo *realmente* desafiador. [Registro meu respeito e admiração por tantas mulheres que são mães solteiras e não desistem dos seus objetivos de vida].

Quero então começar agradecendo ao Bruno, meu melhor amigo, meu parceiro de vida, pessoa que sempre me apoia seja qual sonho doido eu tenha. Se tem alguém que sabe o desafio que foi desenvolver essa tese é ele, pois esteve ao meu lado sem oscilar. Não pensei em desistir, pois sigo amando o que faço, mas muitas foram as vezes que senti falta da rotina de laboratório, triste por não conseguir dar os 100% que eu gostaria nos estudos. Nesses momentos em especial, o Bruno estava presente para me reconfortar e apoiar. Sou imensamente grata. Agradeço também à minha filha, Gabriela, por ser a flor que ela é. Inúmeras vezes foi compreensiva e me deu o espaço que eu precisava para trabalhar. Na reta final, ao brincarmos de fazer desejo com o fio de cílios que caiu ela desejava que "a mamãe consiga terminar a tese". Sempre me apoiando também, na delicadeza e honestidade de uma criança.

Agradeço também aos meus pais. Foram eles que me ensinaram que não tem sonho grande demais que não consigamos alcançar com nosso trabalho honesto e dedicação. Foram eles minha base de apoio todos esses anos que me trouxeram à essa fase da minha vida. Fui criada em uma bolha mágica onde homens e mulheres são iguais e capazes de alcançar os seus objetivos; onde seu sempre pude me expressar e sonhar. Sim, obter o título de doutor é realizar um sonho de criança para mim. Um sonho que a criança registraria como realizado, mas que eu, a adulta que sou, sei que é o início de uma carreira e de uma jornada que começo com alegria.

Não posso falar da minha jornada acadêmica sem falar do meu orientador, o Luis Mauricio Bini. Uma pessoa que transformou minha visão de mundo de uma maneira indescritível. Existe uma Sara antes (entrando na graduação) e depois (defendendo o doutorado) de conhecer o Bini. Já pedi conselho, já pedi abraço (sim, *acredite*), já passei tempos preocupada que o doutorado iria terminar logo e que eu sentiria muita saudade – parte de mim queria parar o tempo para poder apreciar cada momento ainda mais. Uma das lembranças mais legais que tenho do Bini é uma conversa de corredor que tivemos um dia, eu começando minha iniciação científica. Eu na época achando que estava *arrasando*, obedecia cegamente a tudo que ele mandava eu fazer. Até que ele chegou e falou que eu não tinha que obedecer a ele, que eu poderia fazer diferente do que ele falava desde que eu o convencesse (eis a questão). Uma chave foi virada na minha cabeça. Me senti capaz, competente. Claro que raramente consegui o convencer de algo, mas eu estava sempre estimulada a estudar, explorar, ter uma opinião. Recebi relatórios e textos sangrando, como dizem. Recebi críticas e sugestões, típicas de bons orientadores que prezam por ensinar seus alunos. Me considero uma pessoa de sorte por ter tido essa experiência maravilhosa de ter sido orientada por ele (tantas vezes rsrs). Eu jamais terei palavras para descrever o quanto essa pessoa é importante na minha vida, ou o carinho que sinto por ele. Registro minha eterna gratidão pela presença, paciência, e sempre boa vontade.

Não posso falar do Bini sem falar da Pri. Ela tem um jeito mais do que especial de se fazer presente. Lembro de quando comecei a interagir com ela, ainda na graduação. Ela sempre me avisava se era um bom dia para eu conversar com o Bini (afinal, ainda estávamos nos conhecendo). Ela chama para tomar café, escuta, aconselha. Nossa vida em laboratório era sempre a sensação de que a Pri cuidava da gente. O coração do laboratório. Sou muito grata a ela pelas conversas, pelo carinho e pelo apoio de sempre. Sentirei *muitas* saudades. Aproveito para agradecer aos meus colegas de laboratório que são pessoas maravilhosas, sempre dispostas a ajudar uns aos outros. Agradeço em especial a Leticia pelas inúmeras trocas de figurinha, ao Jean por sempre estar disponível para me ajudar, e ao Jaques por topar qualquer troca de ideia, por mais maluca que seja. Agradeço aos meus colegas de turma que tornaram as aulas sempre divertidas. São pessoas competentes que eu sou muito grata por ter conhecido.

Agradeço ao programa e aos professores por fazerem desta uma experiência tão rica, em especial ao Hortal que leu tantas versões deste trabalho desde o início. Agradeço à CAPES pela bolsa que me permitiu ter dedicação exclusiva no Doutorado e por disponibilizar uma biblioteca tão rica de artigos que são os periódicos da CAPES. Agradeço aos funcionários da UFG que sempre possibilitaram um bom funcionamento da estrutura física, e aos colegas do Labeco-UFPA (em especial à Thaisa Michelan e à Ana Paula) por me receberem de braços abertos no laboratório.

Agradeço aos meus irmãos, Miguel, Débora, Ana Celeste e Maria Cecília, que tantas vezes tornaram essa jornada mais leve. Tantas vezes liguei e falei que eu queira conversar atoa para espairecer, e me encontrei com horas de conversas leves, saudosas e cheias de afeto. Agradeço aos meus amigos por aturarem minha distância e seguirem cultivando nossa amizade independente da minha ausência. Meus amigos tão queridos que, muitos à distância, acompanhavam esse momento tão único que foi escrever essa tese. Agradeço à Natácia Lima e ao André Nascimento, pessoas que dividiram teto comigo, e tornaram Goiânia uma segunda casa. Sou grata também pelos novos amigos que fiz durantes estes 4 anos por me aceitarem em uma fase no mínimo, diferente, também sempre me apoiando.

Dedico este trabalho a todos vocês.

Abstract

Systematic reviews value transparency, systematic reporting and reproducibility and follow clear and direct guidelines. It is as simple as systematizing the qualitative review, as the name implies. A quantitative review (or meta-analysis) may also be conducted and goes a step further by providing empirical support for the conclusions. Using this technique, we can combine multiple, independent studies by calculating an average effect size weighted by study precision, which provides the analysis with an increased statistical power. This thesis addresses meta-analysis as its main theme, and is divided into three chapters, all concerning meta-research in freshwater environments. In Chapter 1, I evaluated the quality of meta-analyses in freshwater ecology and made recommendations on how we can increase quality of meta-analyses in our filed. I evaluated another issue of quality in Chapter 2, specifically biases towards positive results. In Chapter 3 I conducted a meta-analysis using studies that evaluate the effects of land use on freshwater environments at regional and local scales. I compared the effect sizes obtained when using local and land scape variables. Interestingly, I also provide an example of a study with negative results, given that my hypotheses were not confirmed. I believe this thesis will expand the view of limnologists on the possibilities of the meta-analytical technique applied to our field of research.

Keywords: Meta-analysis, systematic review, land use, publication bias, negative results, quality of meta-analyses

Introduction

Meta-research is an option to organize and make generalizations from the plethora of articles published in a given field. It may be defined as ‘the study of research itself: its methods, reporting, reproducibility, evaluation, and incentives” (Ioannidis, 2018). Meta-research has been well-developed in fields like psychology and medicine, where steps towards better research practices (Altman, 1994; Ioannidis, 2005, 2008, 2010; Ioannidis & Trikalinos, 2007; Fanelli, 2010b a, 2012; Fanelli, Costas & Larivière, 2015; Page & Moher, 2016; Munafò *et al.*, 2017; Brembs, 2018) and more well-systematized reviews have been taken (Moher *et al.*, 2009b; Liberati *et al.*, 2009; Willis & Quigley, 2011). For example, the Cochrane organization is a reference for producing quality systematic reviews and synthesis for decision makers to reach informed and reliable conclusions in the field of medicine (see https://cochrane.org/ and the Cochrane Handbook of Systematic Reviews; Higgins & Green, 2011). Such advancements have also been reaching ecology (Hillebrand & Cardinale, 2010; Jennions & Lortie, 2013; Koricheva & Gurevitch, 2013a; Koricheva, Jennions & Lau, 2013; Gurevitch *et al.*, 2018; Lodi *et al.*, 2021), although more recently.

Systematic reviews value transparency, systematic reporting and reproducibility (Gurevitch *et al.*, 2018) and follow clear and direct guidelines (Moher *et al.*, 2009a and see Tools for Transparency in Ecology and Evolution as well; https://osf.io/g65cb, 2015). It is as simple as systematizing the qualitative review, as the name implies. A quantitative review (or meta-analysis) may also be conducted (Gurevitch *et al.*, 2018) and goes a step further by providing empirical support for the conclusions. A meta-analysis calculates an overall effect size, its significance, and provides a formal, transparent, and quantitative approach for exploring possible causes of variation in estimated effects (Koricheva & Gurevitch, 2013b; Senior *et al.*, 2016; Gurevitch *et al.*, 2018). Using this technique, we can combine multiple, independent studies by calculating an average effect size weighted by study precision (defined as the inverse variance, see Borenstein *et al.*, 2009), which provides the analysis with an increased statistical power. This is especially important in Ecology as the effect size of individual ecological studies tends to be low (Jennions & Lortie, 2013).

This thesis addresses meta-analysis as its main theme, and is divided into three chapters, all concerning meta-research in freshwater environments. In Chapter 1, I evaluated the quality of meta-analyses in freshwater ecology and made recommendations on how we can increase quality of meta-analyses in our filed. I conducted a standardized search and screening and evaluated meta-analyses in freshwater ecology one by one, seeking relevant quality traits in the studies. I observed that the term meta-analysis is frequently misused. However, despite the lower-than-expected quality of the papers, guidelines seem to be reaching researchers that use the term properly.

I evaluated another issue of quality in Chapter 2, specifically biases towards positive results. This is not a novel issue to be discussed in science but is being often revisited. I evaluated meta-analyses and observed the incidence of positive results; positive being the confirmation of the hypothesis in direction and significance. Positive results did prevail. I discuss the importance of bias-control practices in meta-analyses and possible influence of biased primary studies.

In Chapter 3 I conducted a meta-analysis using studies that evaluate the effects of land use on freshwater environments at regional and local scales. I compared the effect sizes obtained when using local and land scape variables. Interestingly, I also provide an example of a study with negative results, given that my hypotheses were not confirmed. I believe this thesis will expand the view of limnologists on the possibilities of the meta-analytical technique applied to our field of research.

References

Altman D.G. (1994). The scandal of poor medical research. *Bmj* **308**, 283–284. https://doi.org/10.1136/bmj.308.6924.283

Borenstein M., Hedges L. V., Higgins J.P.T. & Rothstein H.R. (2009). Introduction to Meta-Analysis. *Psychotherapy research journal of the Society for Psychotherapy Research* **19**, 421. https://doi.org/10.1002/9780470743386

Brembs B. (2018). Prestigious Science Journals Struggle to Reach Even Average Reliability. *Frontiers in Human Neuroscience* **12**, 1–7. https://doi.org/10.3389/fnhum.2018.00376

Fanelli D. (2010a). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLoS ONE* **5**. https://doi.org/10.1371/journal.pone.0010271

Fanelli D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics* **90**, 891–904. https://doi.org/10.1007/s11192-011-0494-7

Fanelli D. (2010b). "Positive" results increase down the hierarchy of the sciences. *PLoS ONE* **5**. https://doi.org/10.1371/journal.pone.0010068

Fanelli D., Costas R. & Larivière V. (2015). Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. *PLOS ONE* **10**, e0127556. https://doi.org/10.1371/journal.pone.0127556

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

Hillebrand H. & Cardinale B.J. (2010). A critique for meta-analyses and the productivity– diversity relationship. *Ecology* **91**, 2545–2549. https://doi.org/10.1890/09-0070.1

Ioannidis J.P.A. (2010). Meta-research: The art of getting it wrong. *Research Synthesis Methods* **1**, 169–184. https://doi.org/10.1002/jrsm.19

Ioannidis J.P.A. (2018). Meta-research: Why research on research matters. *PLOS Biology* **16**. https://doi.org/10.1371/journal.pbio.2005468

Ioannidis J.P.A. (2008). Why Most Discovered True Associations Are Inflated. *Epidemiology* **19**, 640–648. https://doi.org/10.1097/EDE.0b013e31818131e7

Ioannidis J.P.A. (2005). Why most published research findings are false. *PLoS Medicine* **2**, 0696–0701

Ioannidis J.P.A. & Trikalinos T.A. (2007). An exploratory test for an excess of significant findings. *Clinical Trials* **4**, 245–253. https://doi.org/10.1177/1740774507079441

Jennions M.D. & Lortie C.J. (2013). Using Meta-analysis to Test Ecological and Evolutionary Theory. In: *Handbook of Meta-analysis in Ecology and Evolution*. pp. 381–404.

Koricheva J. & Gurevitch J. (2013a). *Handbook of Meta-Analysis in Ecology and Evolution*. (Eds J. Koricheva, J. Gurevitch & K. Mengersen), Princeton University Press, Princeton and Oxford.

Koricheva J. & Gurevitch J. (2013b). Place of Meta-Analysis among Other Methods of Research Synthesis. *Handbook of Meta-analysis in Ecology and Evolution*, 498. https://doi.org/10.2307/j.ctt24hq6n

Koricheva J., Jennions M.D. & Lau J. (2013). Temporal Trends in Effect Sizes: Causes, Detection, and Implications. In: *Handbook of Meta-analysis in Ecology and Evolution*.

Liberati A., Altman D.G., Tetzlaff J., Mulrow C., Gotzsche P.C., Ioannidis J.P.A., *et al.* (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* **339**, b2700–b2700. https://doi.org/10.1136/bmj.b2700

Lodi S., Godoy B.S., Ortega J.C. gonçalves & Bini L.M. (2021). Quality of meta-analyses in freshwater ecology: A systematic review. *Freshwater Biology* **66**, 803–814. https://doi.org/10.1111/fwb.13695

Moher D., Liberati A., Tetzlaff J. & Altman D.G. (2009a). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, b2535–b2535. https://doi.org/10.1136/bmj.b2535

Moher D., Liberati A., Tetzlaff J., Altman D.G., Altman D., Antes G., *et al.* (2009b). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine* **6**

Moher D., Shamseer L., Clarke M., Ghersi D., Liberati A., Petticrew M., *et al.* (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* **4**, 1. https://doi.org/10.1186/2046-4053-4-1

Munafò M.R., Nosek B.A., Bishop D.V.M., Button K.S., Chambers C.D., Percie du Sert N., *et al.* (2017). A manifesto for reproducible science. *Nature Human Behaviour* **1**, 0021. https://doi.org/10.1038/s41562-016-0021

Page M.J. & Moher D. (2016). Mass Production of Systematic Reviews and Meta-analyses: An Exercise in Mega-silliness? *The Milbank Quarterly* **94**, 515–519. https://doi.org/10.1111/1468-0009.12211

Senior A.M., Grueber C.E., Kamiya T., Lagisz M., O'Dwyer K., Santos E.S.A., *et al.* (2016). Heterogeneity in ecological and evolutionary meta-analyses: Its magnitude and implications. *Ecology* **97**, 3293–3299. https://doi.org/10.1002/ecy.1591

Willis B.H. & Quigley M. (2011). The assessment of the quality of reporting of meta-analyses in diagnostic research: A systematic review. *BMC Medical Research Methodology* **11**, 163. https://doi.org/10.1186/1471-2288-11-163

Chapter I

Quality of meta-analyses in freshwater ecology: A systematic review

Lodi, S., B. S. Godoy, J. C. gonçalves Ortega, & L. M. Bini, 2021. Quality of meta-analyses in freshwater ecology: A systematic review. Freshwater Biology, 66:5 803-814. https://onlinelibrary.wiley.com/doi/10.1111/fwb.13695.

Quality of meta-analyses in freshwater ecology: a systematic review

Sara Lodi[1]*, Bruno Spacek Godoy[2], Jean Carlo Gonçalves Ortega[3], Luis Mauricio Bini[4]

[1]Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, 74690-900, Goiânia, Goiás, Brazil

[2]Núcleo de Ciências Agrárias e Desenvolvimento Rural, Universidade Federal do Pará, Belém, Pará, Brazil

[3]Programa de Pós-Graduação em Ecologia e Manejo de Recursos Naturais, Universidade Federal do Acre, Rio Branco, Acre, Brazil

[4]Departamento de Ecologia, Universidade Federal de Goiás, 74690-900, Goiânia, Goiás, Brazil

*Corresponding author: lodi.sara@gmail.com

Resumo

Revisões sistemáticas e meta-análises em ecologia tem sido cada vez mais usadas em ecologia. Logo, seus protocolos devem ser seguidos à risca para garantir a qualidade. Vários *checklists* (listas de verificação) estão disponíveis para orientar os pesquisadores rumo a uma meta-análise de alta qualidade. Os estudos de ecologia de água doce tradicionalmente realizam estudos experimentais, fornecendo dados ideais para testar hipóteses usando meta-análise. Aqui, avaliamos a qualidade de 114 meta-análises em ecologia de água doce e 86 meta-análises em ecologia e evolução para fins comparativos. Descobrimos que muitos estudos ainda usam o termo meta-análise incorretamente e que esse erro persistiu com o tempo. A qualidade dos estudos que conduziram uma meta-análise formal melhorou. Assim, especulamos que as diretrizes disponíveis estão sendo eficazes na melhoria da qualidade dos estudos meta-analíticos. A qualidade não foi associada ao fator de impacto do periódico onde as meta-análises foram publicadas ou ao número médio de citações. Além do uso incorreto do termo, descobrimos que muitos estudos falharam em: relatar estatísticas de heterogeneidade, avaliar mudanças temporais no tamanho do efeito, conduzir análises de viés de publicação, abordar a colinearidade entre moderadores e fornecer os dados. Em geral, meta-análises em ecologia e evolução têm uma pontuação média ligeiramente melhor do que meta-análises em ecologia de água doce. Embora a qualidade das meta-análises em ecologia de água doce tenha melhorado ao longo do tempo, há muito espaço para melhorias. Os autores não devem nomear seus estudos como meta-análises se esses métodos não forem usados. A conformidade com as listas de verificação deve ser amplamente incentivada, pois as meta-análises estão cada vez mais sendo usadas para resumir resultados em diferentes áreas da ecologia. Autores, revisores e editores devem usar listas de verificação para melhorar a qualidade das meta-análises em ecologia de água doce.

Palavras-chave: sistematização, revisão quantitativa, limnologia, Prisma, checklist

Abstract

The protocols of systematic reviews and meta-analyses in ecology should be closely followed to ensure quality given the increasing use. Several checklists are available to guide researchers towards a high-quality meta-analytic study. Freshwater ecology studies have a tradition of using experimental studies, which provide the ideal data to test hypotheses using meta-analysis. Here, we evaluated the quality of 114 meta-analyses in freshwater ecology and 86 meta-analyses in ecology and evolution for comparative purposes. We found that many studies are still using the term meta-analysis incorrectly and that this error persisted over time. The quality of the studies that did conduct a formal meta-analysis has improved. Thus, we speculate that available guidelines are being effective in improving the quality of meta-analytic studies. Quality was not associated with the impact factor of the journal where the meta-analyses were published or with the average number of citations. In addition to the incorrect use of the term, we found that many studies failed to: report heterogeneity statistics, evaluate temporal changes in effect size, conduct publication bias analyses, address the collinearity among moderators and provide the data. In general, meta-analyses in ecology and evolution have only a slightly better average score than meta-analyses in freshwater ecology. Although the quality of meta-analyses in freshwater ecology has improved over time, there is much room for improvement. Authors should not label their studies as meta-analyses if these methods were not used. Compliance with checklists should be widely fostered as meta-analyses are increasingly being used to summarize findings in different areas of ecology. Authors, reviewers, and editors should use checklists to improve the quality of meta-analyses in freshwater ecology.

Keywords: systematization, quantitative review, limnology, Prisma statement, checklist

Introduction

Most review articles in ecology can be classified as narrative reviews, where a given research theme is qualitatively summarized in a narrative structure (Gates, 2002; Koricheva & Gurevitch, 2013). Narrative reviews are useful to explore perspectives, conceptual frameworks, historical advances, and to provide a broader interpretation of topics (Collins & Fauser, 2005; Gurevitch et al., 2018). However, narrative reviews tend to be plagued with subjectivity and biases (e.g., by not stating the methods used to compile the primary studies reviewed; Koricheva & Gurevitch, 2013; Lortie, 2014). In recent decades, a growing number of studies have shown that systematic reviews and meta-analyses, if properly conducted, can minimize these issues (Gates, 2002; Koricheva & Gurevitch, 2013; Lortie, 2014; Gurevitch et al., 2018). These methods are especially important considering the need for predictive power and generalization in the field of ecology (e.g., Houlahan et al., 2017).

Systematic reviews value transparency, systematic reporting, and reproducibility (Gurevitch et al., 2018). A meta-analysis may be conducted when enough data is available, going a step further by quantitatively summarizing empirical findings (Gurevitch et al., 2018). In a nutshell, a meta-analysis calculates a weighted mean effect size, its statistical significance and, importantly, explores possible causes of variation in effect sizes using moderators (Borenstein et al., 2009). Weighting effect sizes by their precision is a pivotal step (Borenstein et al., 2009), as it ensures that estimates will be more influenced by studies with high precision.

However, as emphasized by Olkin (1992), “doing a meta-analysis is easy, doing one well is hard” (see also Felson, 1992; Berman & Parker, 2002). The devil is in the details, and poorly conducted meta-analysis may do more harm than good. For example, results from meta-analyses, in the medical sciences, are generally regarded as providing stronger evidence than a single randomized control trial (but see Murad et al., 2016). However, if the meta-analysis is conducted poorly, it may deliver misleading conclusions that may be believed as truthful and determine medical treatment regimens or health policies. Similarly, poorly conducted meta-analyses in ecology, in addition to delaying scientific progress, may lead to inappropriate conservation policies. For example, it is common that several primary studies in a meta-analysis contribute more than one effect size (e.g., when the same control is compared with two experimental groups or when the control group is compared with an experimental group at different time-points). Accordingly, a meta-analysis ignoring the fact that multiple effect sizes from the same primary studies are not independent would produce results with inflated type I error rates (see alsoSong et al, 2020; Mengersen, Jennions, & Schmid, 2013; Van den Noortgate et al., 2014). In general, issues associated with lack of transparency (e.g., in selecting studies and reporting), misinterpretation of results (Morrissey, 2016), incorrect methodology (López-López et al., 2018), premature use of meta-analyses where there is insufficient data (Borenstein et al., 2009; Ioannidis, 2010), and lack of systematization (Gurevitch et al., 2018), for example, have prompted critiques and challenged the credibility of meta-analytical results (Whittaker, 2010; Ioannidis, 2016; Morrissey, 2016). This shows a need to better disseminate good practices among different research fields and to improve the quality of meta-analysis studies (Hillebrand & Cardinale, 2010; Gurevitch et al., 2018).

Clear guidelines for conducting a systematic review have been published and updated as a strategy to improve the quality of meta-analysis (Moher et al., 2009; 2015; Fleming et al., 2014; Shamseer et al., 2015). Well-established guidelines for systematic reviews include, but are not limited to, the Cochrane Handbook of Systematic Reviews (Higgins & Green, 2011), PRISMA statement (Moher et al., 2009; 2015), Tools for Transparency in Ecology and Evolution (Parker et al., 2018) and other relevant checklists (Higgins et al., 2013; Nakagawa et al., 2017; see applications in Delaney et al., 2005; Willis & Quigley, 2011). In ecology, Koricheva & Gurevitch (2014) have proposed a checklist of quality criteria for meta-analysis for research synthesists, peer reviewers and editors (see also Nakagawa et al., 2017). However, despite these calls for systematization, misuses and poor reporting are still recurrent in meta-analyses in many fields of research (Gates, 2002; Roberts, Stewart, & Pullin, 2006; Philibert, Loyce, & Makowski, 2012; Vetter, Rucker, & Storch, 2013; Koricheva & Gurevitch, 2014; Senior et al., 2016).

Here, we performed a systematic review of meta-analyses published within the field of freshwater ecology and evaluated the quality of these studies considering established criteria. We used the checklist proposed by Koricheva & Gurevitch (2014) as the main reference to evaluate the quality of meta-analyses in freshwater ecology. For comparative purposes, we also evaluated the quality of meta-analyses in the field of ecology and evolution (Senior et al., 2016). We expected that our measure of quality (see Methods section) would increase over time due to increased knowledge of best practices by authors, reviewers, and editors. We also asked whether quality was associated with citation metrics to test whether better developed articles were published in journals with higher visibility. We also expected that the quality score would be correlated with the average number of citations, assuming that readers would seek well-developed studies as references. We expected that the statistical issues related to phylogenetic relationship would be the least addressed criterion due to the lack of recognition that species cannot be regarded as independent points in statistical analyses (Felsenstein, 1985). Also, phylogenetic data are not readily available for many biological groups in the freshwater realm. We expected that the assessment of temporal changes in effect sizes using, for example, cumulative meta-analysis, would not often be conducted. Despite being a well-established tool (e.g., Leimu & Koricheva, 2004), ecologists have only recently considered its potential and importance (e.g., Ortega, Thomaz, & Bini, 2018; Koricheva & Kulinskaya, 2019). Methods to address the use of multiple outcomes or dependent effect sizes have only recently been published (e.g., Hedges et al., 2010; Van den

Noortgate et al., 2014) and used in ecological studies (e.g., Stein, Gerstner, & Kreft, 2014). Thus, we predict that this quality criterion would also often be neglected.

Methods

Search strategy and criteria

We searched for meta-analyses in freshwater ecology using the Web of Science (WoS) and Scopus databases (from 1990 to August 8th, 2017). Our search string (see details in Table 1) resulted in 443 and 382 hits for WoS and Scopus, respectively. We removed duplicates using the function “mergeDbSources” from the R package Bibliometrix (Aria & Cuccurullo, 2017) and screened the articles following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009; 2015). The PRISMA statement is a checklist, including a flow diagram, designed to help authors improve the reporting of systematic reviews (Moher et al., 2009; 2015). The first step consisted of reading the abstracts to exclude obviously irrelevant literature. Then, we read the papers for eligibility (full-text assessment). We included studies that contained formal meta-analyses and we excluded all reviews that were not ‘quantitative reviews using meta-analysis procedures’. Also, we only included studies with results from freshwater ecosystems and excluded studies that did not focus on freshwater organisms (Figure 1.1,Table S1. 1).

Table 1.1 Search string used to obtain meta-analytical studies on freshwater ecology published between 1991 to 2017 in the Web of Science (WoS) and Scopus databases. The Boolean codes were adjusted to each search database, but the same words and wildcard functions were used in both.

Terms	Search string
Meta-analysis terms	(meta-anal* OR metanal* OR "quantitative review")
Freshwater ecology terms	((freshwater OR aquatic* OR limnol* OR "inland water*" OR river* OR stream* OR creek* OR reservoir* OR lake* OR lagoon* OR pond* OR mere* OR loch* OR lakelet*) NOT (ocean* OR marine))

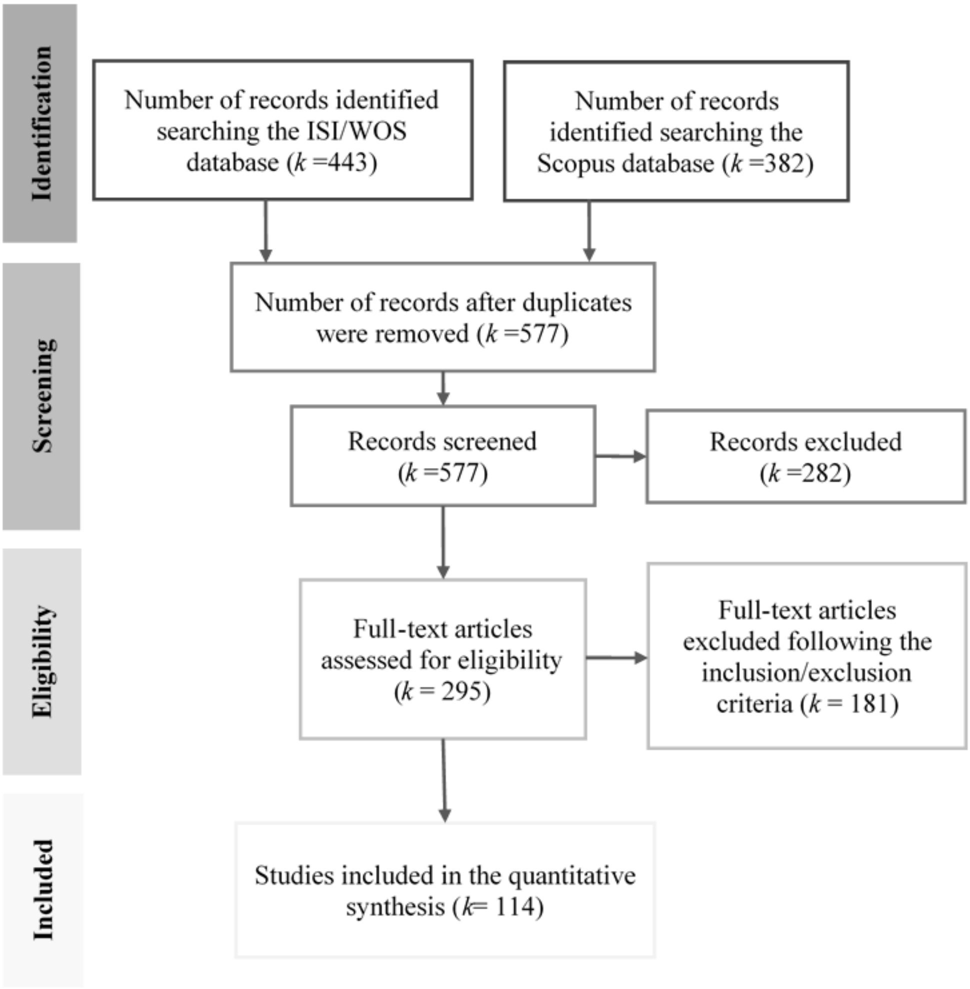


Figure 1.1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram summarizing study inclusion and exclusion phases (adapted from Moher et al., 2009)

A second reviewer re-evaluated 30% of the articles collected from the databases (174 out of 577 papers, selected using the function sample() in the software R; R Core Team, 2020). We then calculated the agreement rate (AR; Orwin & Vevea, 2009) and Cohen's Kappa (Cohen, 1960) to assess the reliability of the screening procedures conducted in this study.

Meta-analysis quality

We assigned scores to each study following the checklist of criteria proposed by Koricheva & Gurevitch (2014), which consisted of 16 criteria with references to support them (see their Table 1. 3). We altered the criterion "use of a Meta-Regression to appraise the issue of exploring existent heterogeneity" to "exploring the heterogeneity among studies" (Table 1. 2). In this case, the criterion was met whenever the authors explored causes of heterogeneity among studies using moderators irrespective of the approach (e.g., meta-regression or subset analyses). For each criterion met, we assigned one point to the study. Half points were assigned when the information was incomplete (e.g., lack of the specific search terms). To fulfil the third criterion, the studies had to have calculated weighted effect sizes before modelling. We also assigned one point to the third criterion when effect sizes were weighted by the number of observations (see Gurevitch et al., 2018). Some studies only sought to explain the variation among studies (without estimating a weighted mean effect size). Consequently, we assumed that studies reporting variance partitioning results met the requirements for criterion 3. Some criteria were not relevant for some studies depending on their design (Table 1. 2). For example, taking phylogenetic relatedness into account is not relevant for studies focused on a single species or ecosystem-level variables. Therefore, we calculated the ratio between the sum of points and the maximum possible number of points for the study. Accordingly, the maximum rating is no longer 16 as initially proposed by Koricheva & Gurevitch (2014) but ranged from 0 (lowest) to 1 (highest quality).

We tested for a temporal trend in the rating scores using the Spearman rank correlation between scores and publication year. Similarly, we tested for a correlation between scores and impact factor, and between scores and Normalized Citation Impact Index (NCII) using the Spearman rank correlation. The NCII is given by the ratio between the total number of citations and the time elapsed (in years) since publication (see Coursaris & Van Osch, 2014 and references therein).

Table 1. 2 Quality criteria used to assign scores to the meta-analysis papers selected and observations explaining the criteria used and exceptions in punctuation. The quality criteria were obtained from Koricheva & Gurevitch (2014).

	Quality criteria	**Observations**
1	Are details of bibliographic search (electronic data bases used, keyword combinations, years) reported in sufficient detail to allow replication?	Half a point when the search terms are not provided; Not relevant for meta-analysis conducted with primary data.
2	Are criteria for study inclusion/exclusion explicitly listed?	Half a point when the criteria provided are not clear enough for reproducibility. Not relevant for meta-analysis conducted with primary data.
3	Have effect sizes been weighted by study precision or has the rationale for using unweighted approach been provided?	
4	Has the statistical model for meta-analysis been described?	
5	Has heterogeneity of effect sizes between studies been quantified?	
6	Have the causes of existent heterogeneity in effect sizes been explored?	
7	If effects of multiple moderators have been tested, have potential non-independence of and interactions between moderators been considered?	Not relevant if moderators were not considered
8	Have tests for publication bias been conducted?	
9	Have sensitivity analysis been performed to test the robustness of results?	
10	If meta-analysis combines studies published over considerable time span, have possible temporal changes in effect size been tested?	
11	If meta-analysis combined studies conducted on different species, has phylogenetic relatedness of species been considered?	Not relevant for studies considering single species or focused on community (*e.g.*, diversity or species richness) or ecosystem responses
12	Has the software used been described?	
13	Have full bibliographic details of primary studies included in a meta-analysis been provided?	Not relevant for meta-analysis conducted with primary data.
14	Has the data set used for meta-analysis, including effect sizes and variances/sample sizes from individual primary studies and moderator variables, been provided as electronic appendix?	
15	Have standard metrics of effect size been used or, if nonstandard metrics have been employed, is the distribution of these parameters known and have the authors explained how they calculated variances for such metrics?	
16	If more than one estimate of effect size per study was included in the analysis, has potential non-independence of these estimates been considered?	Not relevant for studies that used one estimate of effect size per study

We also rated the 86 meta-analyses in ecology and evolution compiled by Senior et al. (2016), to compare with our results. Senior et al. (2016) only included studies that quantified heterogeneity among effect sizes. Thus, for comparative purposes, we rated each of the 86 meta-analyses in ecology and evolution without considering the following criteria: "Has heterogeneity of effect sizes between studies been quantified?" (5th criterion), "Have the causes of existent heterogeneity in effect sizes been explored?" (6th criterion) and "Have standard metrics of effect size been used (…)" (15th criterion). We checked whether freshwater studies were included by Senior et al. (2016) and found that a total of 9 studies also evaluated the freshwater realm, none of which were included in our study, thus ensuring independence between datasets.

Results

Most of the 577 studies were excluded in abstract and full text screening because they did not conduct a formal meta-analysis (224 studies) or because they were not focused on freshwater systems (58 studies). The studies that were excluded during the analysis of the abstracts (282 papers) were clearly out of the objectives of our review. Specifically, 118 and 87 studies were not focused on freshwater systems and on ecological questions, respectively. Also, 76 studies did not conduct a formal meta-analysis and one was clearly methodological.

After reading the full texts, we excluded 147 and 34 studies for not being formal meta-analyses or focused on aquatic ecosystems, respectively. Most of these studies consisted of a re-analysis of primary data, without calculating an effect size per study. We found 114 meta-analytic studies that met our criteria (Figure 1.1) after following the different steps to select the studies. We found an agreement rate (AR) of 0.90 and a Cohen's Kappa of 0.73 (95% CI = 0.62-0.84), indicating that most of the studies were consistently selected by both reviewers during the screening of the abstracts.

The number of meta-analytic studies focused on freshwater ecology increased over time, from the first study published by Wooster (1994), among the studies we selected, to 11 studies published in 2016 and 8 in 2017 (until the search date; Figure 1.2a). These papers were published mostly in Ecology, Oikos, Freshwater Biology and other specific journals within the field of freshwater ecology (Figure 1.2b).

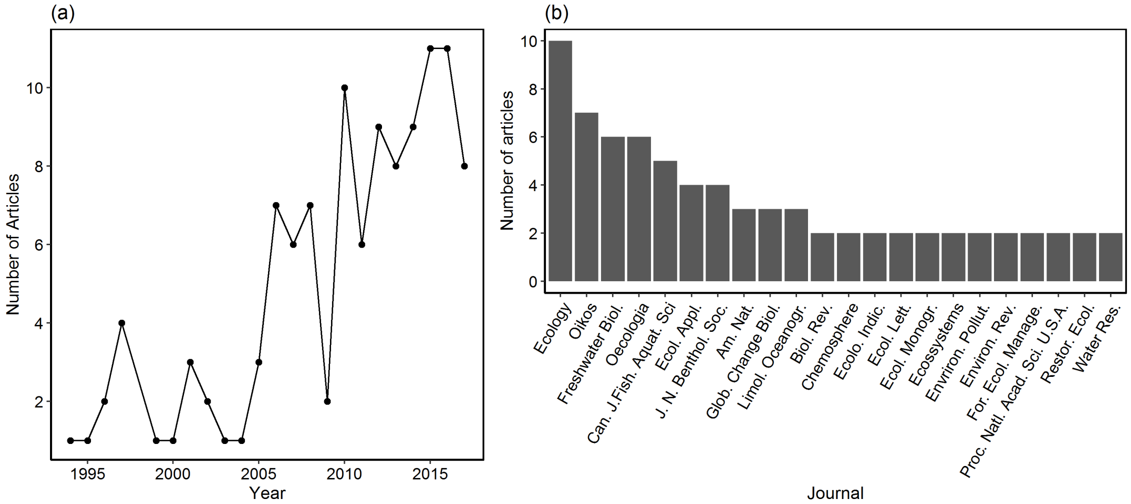


Figure 1.2 Annual number of studies on freshwater ecology using meta-analysis (a). Number of studies per journal (b). The dataset included a total of 114 studies published in 61 journals.

We found an average rating score of 0.56 ± 0.20 (mean ± SD). Rating scores increased over time (Spearman correlation $r = 0.43$; $P < 0.001$) but were not related to journal impact factor ($r = 0.03$; $P = 0.720$) or to the Normalized Citation Impact Index ($r = 0.15$; $P = 0.121$). Most studies (86.8%, out of 114) did not test for temporal changes in effect sizes (10th criterion). Also, most studies did not perform sensitivity analyses (9th criterion), evaluate the dependence among moderators (7th criterion) and did not provide heterogeneity statistics (5th criterion; Table 1. 3). Nearly half of the studies did not check for publication bias (8th criterion) nor provided the data used for the meta-analysis (14th criterion). The issues of using multiple effect sizes per primary study (16th criterion) were also often ignored and approximately 36 % of the studies did not provide details of the full literature search (1st criterion; Table 1. 3). However, other important criteria were met more frequently; for example, most studies used standard metrics of effect size (15th criterion), explored the heterogeneity in effect sizes (6th criterion) and used the inverse-variance method to weight effect sizes (3rd criterion; Table 1. 3).

Table 1. 3 Compliance with the criteria for reporting and following methodological standards assigned to freshwater ecology meta-analyses. This survey encompasses 114 studies published from 1991 to 2017. NR = Not Relevant for the study. A detailed description of the criteria is shown in Table 2.

	Quality Criteria	NR (%)	Yes (%)	No (%)	Partial (%)
1	Searching details	7.02	28.95	35.96	28.07
2	Inclusion/exclusion criteria	7.02	57.02	19.3	16.67
3	Weighted effect sizes	-	75.44	24.56	-
4	Meta-analytical model	-	71.93	28.07	-
5	Heterogeneity in effect sizes	-	41.23	58.77	-
6	Causes of heterogeneity	0.88	82.46	16.67	-
7	Collinearity analysis	15.79	22.81	61.4	-
8	Publication bias	6.14	39.47	54.39	-
9	Sensitivity analysis	-	37.72	62.28	-
10	Changes in effect size	5.26	7.89	86.84	-
11	Controlling for phylogeny	71.05	3.51	25.44	-
12	Software	-	69.3	30.7	-
13	Bibliographic details	6.14	75.44	18.42	-
14	Data	-	52.63	47.37	-
15	Standard effect size	-	93.86	6.14	-
16	Multiple effect sizes	27.19	26.32	39.47	7.02

Most of the criteria were met in ecological and evolutionary studies (average rating score of 0.61 ± 0.21 SD) with a higher frequency than in limnological meta-analyses (Table 1. 4). This was especially so for criteria 1, 3, 4, 8, 9, 11, 13 and 16 (Figure 1.3). In general, criteria 1, 7, 10, 11 and 16 were more frequently unmet (< 50% of the studies) in both fields.

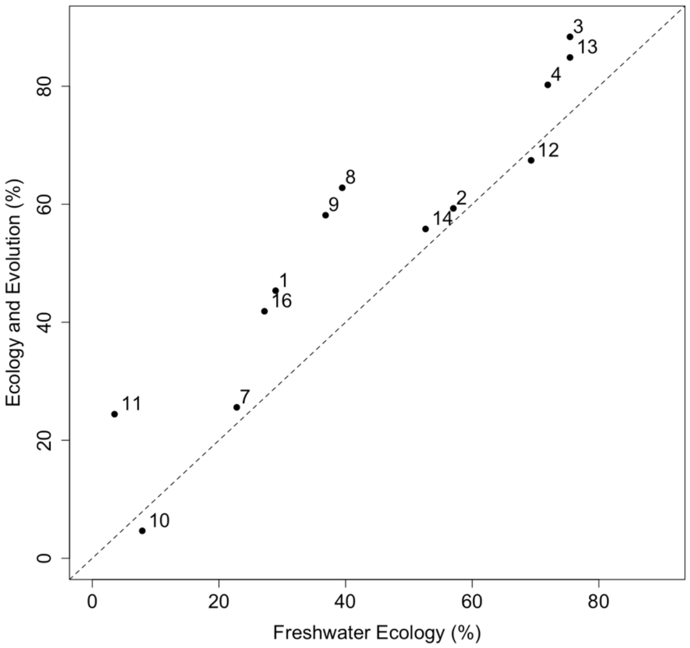


Figure 1.3 Relationship between percentages of meta-analytical studies complying with different criteria for both freshwater ecology (this study) and ecology and evolution (compiled by Senior et al., 2016). Numbers indicate the quality criteria (Table 1. **2**). Values above the line represent criteria that were met more frequently by ecological and evolutionary meta-analyses, while values below the line were met more frequently by freshwater ecology meta-analyses.

Table 1. 4 Compliance with the criteria for reporting and following methodological standards assigned to 86 ecological and evolutionary meta-analyses reviewed by Senior et al. (2016). A detailed description of the criteria is shown in Table 2. * indicates criteria that were disregarded for ecological and evolutionary meta-analyses.

	Quality Criteria	Yes (%)	No (%)	Partial (%)
1	Searching details	49.41	27.06	20.00
2	Inclusion/exclusion criteria	63.53	11.76	21.18
3	Weighted effect sizes	92.94	7.06	0.00
4	Meta-analytical model	84.71	15.29	0.00
5	Heterogeneity in effect sizes	*	*	*
6	Causes of heterogeneity	*	*	*
7	Collinearity analysis	27.06	63.53	0.00
8	Publication bias	65.88	30.59	0.00
9	Sensitivity analysis	61.18	38.82	0.00
10	Changes in effect size	5.88	90.59	0.00
11	Controlling for phylogeny	24.71	45.88	0.00
12	Software	71.76	28.24	0.00
13	Bibliographic details	89.41	7.06	0.00
14	Data	58.82	41.18	0.00
15	Standard effect size	*	*	*
16	Multiple effect sizes	42.35	35.29	4.71

Discussion

Meta-analysis is considered a powerful method for summarizing results of independent studies in different research areas (e.g., Lortie, 2014; Ellis, 2010; Gurevitch et al., 2019). However, the higher credibility of meta-analysis results may be a double-edged sword, especially if reliability is questionable (Ioannidis, 2010; 2016; 2017; Nakagawa et al., 2017). In this context, it was disturbing to find many studies stating that they conducted meta-analyses despite the use of different approaches (e.g., re-analysis of primary data, vote counting, scientometrics, qualitative reviews). Other studies failed to mention the criteria used to select studies, disregarded publication bias analyses (Koricheva & Gurevitch, 2014; Nakagawa et al., 2017) or focused mostly on statistical significance, often neglecting to model heterogeneity (Higgins & Thompson, 2002; Koricheva & Gurevitch, 2014; Senior et al., 2016). Many of the studies failed to address non-independence issues with the use of multiple effect sizes from the same studies (e.g., Hedges, Tipton, & Johnson, 2010), or at least to conduct a sensitivity analysis to demonstrate the robustness of their results (Gurevitch et al., 2018). The analysis of temporal stability in effect sizes was also disregarded in several meta-analyses, which is of concern as it may lead to misleading conclusions (e.g., Koricheva & Kulinskaya, 2019).

According to Koricheva & Gurevitch (2014), meta-analysis is a "set of statistical methods for combining outcomes (effect sizes) across different data sets addressing the same research question to examine patterns of response across these data sets and sources of heterogeneity in outcomes". Here, 224 studies were excluded solely for not being a formal meta-analysis. Some of these exclusions occurred after reading the abstract (77) because the studies stated that their results were important for the development of future meta-analysis, while others cited the term meta-analysis in other contexts. The most striking result concerns the exclusion of 154 studies, after reading the full texts, for not using meta-analytic approaches (e.g., vote-counting and re-analysis of primary data), despite being described as meta-analyses by the authors. This erroneous use of the term meta-analysis has also been observed in other areas of ecology (Vetter et al., 2013; Koricheva & Gurevitch, 2014; Senior et al., 2016), despite many studies urging authors to "apply the term 'meta-analysis' consistently and correctly and not to confuse it with other summary techniques" (e.g., Vetter et al., 2013). Indeed, we found that the number of studies erroneously self-classified as meta-analysis is still increasing (Spearman's r = 0.74; P = 0.0002; n = 20 years, (1998-2017)).

We found that meta-analyses in freshwater ecology require improvements, in line with previous studies from different fields (Delaney et al., 2005; Koricheva & Gurevitch, 2014; Ioannidis, 2016). Meta-analysis quality was not correlated with journal impact factor, indicating that the most well-conducted meta-analyses are not necessarily published in journals with high impact factors (Brembs, 2018). The lack of correlation between the average number of citations an article received and its quality score, in turn, suggests that other factors (e.g., themes and specific results) are more important to "citability" than the quality of the meta-analysis itself (see also Padial et al., 2010 for a general analysis of citation frequency of ecological articles). However, we found that our rating scores were significantly correlated with publication year, suggesting that published guidelines to improve the use of meta-analyses (e.g., Moher et al., 2009; Koricheva & Gurevitch, 2014; Nakagawa et al., 2017) are being effective.

Reporting the search strings and databases (e.g., Web of Science and Scopus) used to find articles that provide data to a meta-analysis, is far from being a mere bureaucratic step (Nakagawa et al., 2007; Gusenbauer & Haddaway, 2020). For example, researchers may avoid wasting time through the repetition of an inclusive meta-analysis when search strings are reported or may detect the need for a new meta-analysis when/if the one published was too restrictive (see Babic et al., 2020). However, our results parallel those of Koricheva &

Gurevitch, (2014) by showing that only a few studies included full details of bibliographic searches. The problem becomes more pronounced when the inclusion/exclusion criteria are not reported (for a discussion on this theme, see Lortie & Callaway, 2006). Even the easy-to-meet quality criterion of providing the list of primary studies included in the meta-analysis was not met by 21 studies, despite being essential for reproducibility testing. Specifically, if the list of primary studies used in a published meta-analysis differs substantially from a list of primary studies retrieved by using a different search string (e.g., due to the use of synonyms or different technical terms), then, all else being equal, a new meta-analysis, considering the different set of primary studies, may produce different results and new insights (e.g. compare Westgate et al., 2014 to de Morais et al., 2018 and Stein, Gerstner, & Kreft, 2014 to Ortega, Thomaz, & Bini, 2018; see also Page & Moher, 2016 for a discussion about the problem of redundant meta-analyses). In addition, only nearly half of the meta-analyses in freshwater ecology provided the dataset used (60 studies). Proportionally, this figure was even lower in plant ecology (Koricheva & Gurevitch, 2014). Providing the dataset used is crucial for cumulative meta-analysis (Leimu & Koricheva, 2004) and to explore the same dataset using different moderators (e.g., Menegotto, Kurtz & Lana, 2019).

A critical step in a meta-analysis is to weight effect sizes by the inverse variance method (Borenstein et al., 2009; Koricheva & Gurevitch, 2013). Many of the studies we reviewed weighted the effect sizes (86 out of 114). Still, 28 of the studies included in our review did not use the inverse variance method. In some cases, the authors stated that the variance (or the data needed to estimate it) was not available in the primary studies. However, other weighting strategies are available, for example, estimating variance from partial data and weighting by n (Koricheva & Gurevitch, 2014). As the quality of reporting and transparency increases in primary studies, the frequency of these issues is expected to decrease (Gerstner et al., 2017).

Thirty-two studies (out of 114) did not report the meta-analytical model used to analyze the data. However, different models (fixed or random effects) have different assumptions, test different hypotheses, and may provide contrasting results (Borenstein et al., 2009; Koricheva & Gurevitch, 2013). In general, the use of a random-effects model is the best alternative because of the restrictive assumptions of a fixed-effect model (mainly that effect sizes are homogeneous across primary studies). Similarly, thirty-five studies did not report the software used, although this is important because it may influence results (Koricheva & Gurevitch, 2014).

Heterogeneity measures are crucial for a comprehensive interpretation of the weighted mean effect size (Senior et al., 2016; Nakagawa et al., 2017; Forstmeier, Wagenmakers & Parker, 2017). Interpreting the weighted mean effect size without the heterogeneity measures is comparable to drawing conclusions from the mean without showing the standard deviation. Few studies have quantified heterogeneity (47 out 114; heterogeneity quantified as T^2 and Q). Still, many explored possible causes of variation (94 out of 114), because most meta-analyses are more interested in exploring variation in effect sizes among studies and in the causes of this variation (Koricheva & Gurevitch, 2014). However, not reporting the heterogeneity or variation of effect sizes is concerning for ecological meta-analyses, given the high complexity of natural systems (Vetter et al., 2013).

The basic strategy to explore causes of heterogeneity is to conduct a meta-regression with several moderators, which corresponds to a weighted multiple regression. Thus, the same assumptions of multiple regression also apply to meta-regression (e.g., checking for collinearity among moderators; Borenstein et al., 2009; Koricheva & Gurevitch, 2013; Nakagawa et al., 2017). While many studies explored the causes of heterogeneity in effect sizes, less than a quarter checked for collinearity. Issues such as instability of parameter estimates, inflated standard errors and biased inferences (Dormann et al., 2013) arise when moderators are collinear. As in any other research approach, the issues with multicollinearity highlight the need for careful planning to conduct a meta-analysis (Berman & Parker, 2002). For example, many studies are screened to compile data on moderators hypothetically related to the effect size. However, without proper planning and theoretical reasoning, this hard work may be in vain if different pairs of moderators are highly correlated to each other or even worse, if redundant moderators are compiled at the expense of ignoring other relevant moderators.

Publication bias is said to exist when the results of a study influence the decision of authors, reviewers, and editors to publish a manuscript. This subject has been extensively debated (Møller & Jennions, 2001; Rothstein, Sutton & Borenstein, 2006; Dwan et al., 2013) due to its effects on systematic reviews and meta-analyses (and also on narrative reviews). Positive results are more likely to be published (Fanelli, 2012; Fanelli, Costas & Ioannidis, 2017; Forstmeier, Wagenmakers, & Parker, 2017). In this context, positive results may be over-represented in the pool of articles available for meta-analysis potentially leading to overestimation of the weighted mean effect size (Møller & Jennions, 2001). Although researchers highlight the need to study publication bias (e.g., Møller & Jennions, 2001; Tomkins & Kotiaho, 2004; Jennions, 2013), over half of the studies (62 out of 114) did not assess it. The lack of concern with publication bias is a recurrent issue in ecology (Roberts, Stewart, & Pullin, 2006; Nakagawa & Santos, 2012; Koricheva & Gurevitch, 2014), as are requests to analyze this issue (Nakagawa & Poulin, 2012).

Twenty-seven meta-analytical studies, out of the 45 that tested for publication bias, used only one method. However, the use of multiple approaches is advisable (Lin et al., 2018). The funnel plot, which may indicate publication bias through asymmetric distribution of effect sizes, was the most used approach (28 papers). The funnel-plot has been criticized on several grounds (Terrin, Schmid & Lau, 2005; Jennions, 2013; Lin, 2019). For instance, asymmetry may be the consequence of other factors when the number of primary studies (k) is small (Egger et al., 1997). Consequently, many studies also used the funnel plot associated with other tests (19 studies – Fail-safe numbers, Quantile-plot, and correlation). The 'trim and fill' method is often used to "correct" for publication bias (Duval & Tweedie, 2000; Peters et al., 2007; and see Weinhandl & Duval, 2012 for an alternative method applied to meta-regression). However, this method tests how robust the results are to publication bias (Jennions & Møller, 2002a; Sutton, 2002) and, therefore, is more of a sensitivity analysis than a method to correct for publication bias (Rothstein, Sutton & Borenstein, 2006).

Conducting sensitivity analyses is also important to check the robustness of results. Basically, it consists of double-checking decisions made throughout the study (Noble et al., 2017). The use of multiple effect sizes, models, publication bias methods and the evaluation of temporal changes in effect sizes are all issues that can be evaluated. However, the best strategy needs to be determined on a case-by-case basis.

A useful strategy to analyze the results is to evaluate whether the effect sizes vary over time (e.g., using a cumulative meta-analysis; Leimu & Koricheva, 2004; Ortega, Thomaz, & Bini, 2018) or to use publication year as a moderator when exploring sources of heterogeneity (Wood & Eagly, 2009; Lehmann et al., 2012). Cumulative sum charts may also be used to detect possible outliers and temporal tendencies (Kulinskaya & Koricheva, 2010; Dogo, Clark & Kulinskaya, 2017). Causes of variation include the tendency of earlier studies to publish higher effect sizes (i.e., time-lag bias), publication bias, Proteus phenomenon (the attraction to contradictory results), true heterogeneity or even chance (Jennions & Møller, 2002b; Trikalinos & Ioannidis, 2006; Ioannidis, 2008). A decrease in effect size over time can be also attributed to the test of a "valid theory beyond its domain of application" (Wilson, Harris, & Wixted, 2020). Despite evidence that effect sizes may change over time (see Koricheva, Jennions, & Lau, 2013 for detailed examples) and the practical implications of this change (Koricheva & Kulinskaya, 2019), only nine meta-analytical studies in freshwater ecology explored this topic.

Independent data points are essential for statistical testing (Forstmeier, Wagenmakers, & Parker, 2017). Non-independence of effect sizes occurs when the data exhibit a correlated or clustered structure, which may be driven by several factors (Lajeunesse, Rosenberg, & Jennions, 2013). For example, in a cross-species meta-analysis, one can find that the effect sizes are phylogenetically structured. Thus, by not accounting for the phylogenetic structure, we would violate the assumption of independence of effect sizes, which has been shown to change the results of meta-analyses (Chamberlain et al., 2012; Nakagawa & Santos, 2012). We found that 29 out of 33 studies for which this criterion was relevant did not control for phylogenetic structure. Nevertheless, methods to incorporate phylogeny have been extensively developed in evolutionary biology and are also available for different meta-analysis models (Adams, 2008; Lajeunesse, 2009; Lajeunesse et al., 2013).

Primary studies often report more than one effect sizes. For example, a primary study may test the relationship between a response variable and different explanatory variables (across the same sampling units). However, if multiple effect sizes per study are used, the total sample size of a meta-analysis will be inflated (Jennions & Schmid, 2013), resulting in increased type I error rates (Hedges et al., 2010; Noble et al., 2017; Forstmeier, Wagenmakers, & Parker, 2017). Only 30 freshwater ecology studies addressed this issue, primarily by comparing the weighted mean effect size obtained using multiple effect sizes to the weighted mean effect size obtained by using one (selected or averaged) effect size per study (e.g., following the approaches described in López-López et al., 2018). There are several statistical methods available to address effect size multiplicity (e.g., Jennions & Schmid, 2013; Tanner-Smith & Tipton, 2014; López-López et al., 2018). For example, Hedges et al. (2010) proposed a method called Robust Variance Estimate (RVE). Given the availability of software (e.g., Fisher & Tipton, 2015), we think that the issue of multiple effect sizes per study will be addressed in most future studies.

Our overall results (mean quality) may be generalizable among different fields in ecology (i.e., considering the results for freshwater ecology, ecology and evolution (Senior et al., 2016) and plant ecology (Koricheva & Gurevitch, 2014)). However, ecological and evolutionary meta-analyses had a higher compliance with the quality criteria than freshwater ecology meta-analyses. Its noteworthy that few studies explored temporal dependence in effect sizes, checked for dependence among moderators, or controlled for phylogenetic dependence in either areas, suggesting this may be a recurrent issue. A similar pattern was observed for plant ecology (Koricheva & Gurevitch, 2014). Thus, we think that further studies should be especially aware of these issues, once lack of independence (temporal, phylogenetic or in explanatory variables), temporal trends in effect sizes and collinearity among moderators are issues of primary concern.

Our findings point to the need for a greater adherence to guidelines for systematic reviews. Thus, we reiterate the importance of applying the checklists, which are widely available, to improve reporting of meta-analyses in different areas of ecology (e.g., Parker et al., 2018; Roberts, Stewart, & Pullin, 2006; Moher et al., 2009; 2015; Koricheva & Gurevitch, 2014). We expect that the use of checklists will increase the quality of meta-analyses, as has been reported, for example, for biomedical research (Han et al., 2017). The need to enhance quality should not be taken lightly, given that their results are potentially used by decision makers as decisive evidence (Koricheva & Kulinskaya, 2019).

Conclusion

We found that the quality of meta-analyses in freshwater ecology has increased over time; however, there is much room for improvement in complying with important quality criteria. Firstly, authors should not mislabel their studies as meta-analyses when, in fact, the method was not used. Secondly, compliance with checklists should be widely fostered as meta-analyses are increasingly being used to summarize findings in different areas of ecology. Thus, authors, reviewers and editors should comprehensively use checklists to improve the quality of meta-analyses in freshwater ecology.

References

Adams, D. C. (2008). Phylogenetic meta-analysis. *Evolution, 62*, 567–572. DOI: 10.1111/j.1558-5646.2007.00314.x

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*, 959–975. DOI: 10.1016/j.joi.2017.08.007

Babić, A., Pericic, T.P., Pieper, D., Puljak, L., (2020). How to decide whether a systematic review is stable and not in need of updating: Analysis of Cochrane reviews. *Research Synthesis Methods*, *11*, 6. DOI: 10.1002/jrsm.1451

Berman, N. G., & Parker, R. A. (2002). Meta-analysis: neither quick nor easy. *BMC Medical Research Methodology*, *2*, 10. DOI: 10.1186/1471-2288-2-10

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (1st ed.). Chichester, UK: John Wiley & Sons. DOI: 10.1002/9780470743386

Brembs, B. (2018). Prestigious science journals struggle to reach even average reliability. *Frontiers in Human Neuroscience*, *12*, 1–7. DOI: 10.3389/fnhum.2018.00376

Chamberlain, S. A., Hovick, S. M., Dibble, C. J., Rasmussen, N. L., Van Allen, B. G., Maitner, B. S., … Whitney, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters*, *15*, 627–636. DOI: 10.1111/j.1461-0248.2012.01776.x

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46. DOI: 10.1177/001316446002000104

Collins, J. A., & Fauser, B. C. J. M. (2005). Balancing the strengths of systematic and narrative reviews. *Human Reproduction Update*, *11*, 103–104. DOI: 10.1093/humupd/dmh058

Coursaris, C. K., & Van Osch, W. (2014). A scientometric analysis of social media research (2004–2011). *Scientometrics*, *101*, 357-380. DOI: 10.1007/s11192-014-1399-z

Delaney, A., Bagshaw, S. M., Ferland, A., Manns, B., Laupland, K. B., & Doig, C. J. (2005). A systematic evaluation of the quality of meta-analyses in the critical care literature. *Critical Care*, *9*, R575–R582. DOI: 10.1186/cc3803

Dogo, S. H., Clark, A., & Kulinskaya, E. (2017). Sequential change detection and monitoring of temporal trends in random-effects meta-analysis. *Research Synthesis Methods*, *8*, 220–235. DOI: 10.1002/jrsm.1222

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46. DOI: 10.1111/j.1600-0587.2012.07348.x

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. DOI: 10.1111/j.0006-341X.2000.00455.x

Dwan, K., Gamble, C., Williamson, P. R., & Kirkham, J. J. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One*, *8*, e66844. DOI: 10.1371/journal.pone.0066844

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*, 629–634. DOI: 10.1136/bmj.315.7109.629

Ellis, P. D. (2010). *The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press, Cambridge. DOI: 10.1017/CBO9780511761676

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, *90*, 891–904. DOI: 10.1007/s11192-011-0494-7

Fanelli, D., Costas, R., & Ioannidis, J. P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National Academy of Sciences*, *114*, 3714–3719. DOI: 10.1073/pnas.1618569114

Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, *125*, 1–15. DOI: 10.1086/284325

Felson, D. T. (1992). Bias in meta-analytic research. *Journal of Clinical Epidemiology*, *45*, 885–892. DOI: 10.1016/0895-4356(92)90072-U

Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. DOI: arXiv:1503.02220v1

Fleming, P. S., Koletsi, D., Seehra, J., & Pandis, N. (2014). Systematic reviews published in higher impact clinical journals were of higher quality. *Journal of Clinical Epidemiology*, *67*, 754–759. DOI: 10.1016/j.jclinepi.2014.01.002

Forstmeier, W., Wagenmakers, E.-J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings - a practical guide. *Biological Reviews*, *92*, 1941–1968. DOI: 10.1111/brv.12315

Gates, S. (2002). Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology*, *71*, 547–557. DOI: 10.1046/j.1365-2656.2002.00634.x

Gerstner, K., Moreno-Mateos, D., Gurevitch, J., Beckmann, M., Kambach, S., Jones H. P., … Seppelt, R. (2017). Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution*, *8*, 777–784. DOI: 10.1111/2041-210X.12758

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*, 175–182. DOI: 10.1038/nature25753

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*, 181–217. DOI: 10.1002/jrsm.1378

Han, S., Olonisakin, T. F., Pribis, J. P., Zupetic, J., Yoon, J. H., Holleran, K. M., … Lee, J. S. (2017). A checklist is associated with increased quality of reporting preclinical biomedical research: a systematic review. *PLoS One*, *12*, e0183591. DOI: 10.1371/journal.pone.0183591

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65. DOI: 10.1002/jrsm.5

Higgins, J. P. T., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. London: The Cochrane Collaboration.

Higgins, J. P. T., Lane, P. W., Anagnostelis, B., Anzures-Cabrera, J., Baker, N. F., Cappelleri, J. C., … Whitehead, A. (2013). A tool to assess the quality of a meta-analysis. *Research Synthesis Methods*, *4*, 351–366. DOI: 10.1002/jrsm.1092

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. DOI: 10.1002/sim.1186

Hillebrand, H., & Cardinale, B. J. (2010). A critique for meta-analyses and the productivity–diversity relationship. *Ecology*, *91*, 2545–2549. DOI: 10.1890/09-0070.1

Houlahan, J. E., McKinney, S. T., Anderson, T. M., & McGill, B. J. (2017). The priority of prediction in ecological understanding. *Oikos*, *126*, 1–7. DOI: 10.1111/oik.03726

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*, 640–648. DOI: 10.1097/EDE.0b013e31818131e7

Ioannidis, J. P. A. (2010). Meta-research: the art of getting it wrong. *Research Synthesis Methods*, *1*, 169–184. DOI: 10.1002/jrsm.19

Ioannidis, J. P. A. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Quarterly*, 94, 485–514. DOI: 10.1111/1468-0009.12210

Ioannidis, J. (2017). Next-generation systematic reviews: prospective meta-analysis, individual-level data, networks and umbrella reviews. *British Journal of Sports Medicine*, *51*, 1456–1458. DOI: 10.1136/bjsports-2017-097621

Jennions, M. D. (2013). Publication and related biases. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.

Jennions, M. D., & Møller, A. P. (2002a). Publication bias in ecology and evolution: an empirical assessment using the "trim and fill" method. *Biological Reviews of the Cambridge Philosophical Society*, *77*, 211–222. DOI: 10.1017/s1464793101005875

Jennions, M. D., & Møller, A. P. (2002b). Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences*, *269*, 43–48. DOI: 10.1098/rspb.2001.1832

Jennions, M. D., & Schmid, C. H. (2013). Statistical models for the meta-analysis of nonindependent data. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.

Koricheva, J., & Gurevitch, J. (2013). Place of meta-analysis among other methods of research synthesis. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.

Koricheva, J., & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, *102*, 828–844. DOI: 10.1111/1365-2745.12224

Koricheva, J., Jennions, M., & Lau, J. (2013). Temporal trends in effect sizes: causes, detection, and implications. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.

Koricheva, J., & Kulinskaya, E. (2019). Temporal instability of evidence base: a threat to policy making? *Trends in Ecology and Evolution*, *34*, 895–902. DOI: 10.1016/j.tree.2019.05.006

Kulinskaya, E., & Koricheva, J. (2010). Use of quality control charts for detection of outliers and temporal trends in cumulative meta-analysis. *Research Synthesis Methods*, *1*, 297–307. DOI: 10.1002/jrsm.29

Lajeunesse, M. J. (2009). Meta-analysis and the comparative phylogenetic method. *The American Naturalist*, *174*, 369–381. DOI: 10.1086/603628

Lajeunesse, M. J., Rosenberg, M. S., & Jennions, M. D. (2013). Phylogenetic nonindependence and meta-analysis. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.

Lehmann, A., Barto, E. K., Powell, J. R., & Rillig, M. C. (2012). Mycorrhizal responsiveness trends in annual crop plants and their wild relatives-a meta-analysis on studies from 1981 to 2010. *Plant and Soil*, *355*, 231–250. DOI: 10.1007/s11104-011-1095-1

Leimu, R., & Koricheva, J. (2004). Cumulative meta-analysis: A new tool for detection of temporal trends and publication bias in ecology. *Proceedings of the Royal Society B: Biological Sciences*, *271*, 1961–1966. DOI: 10.1098/rspb.2004.2828

Lin, L. (2019). Graphical augmentations to sample-size-based funnel plot in meta-analysis. *Research Synthesis Methods*, *10*, 376–388. DOI: 10.1002/jrsm.1340

Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., … Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of General Internal Medicine*, *33*, 1260–1267. DOI: 10.1007/s11606-018-4425-7

López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. T. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*, *9*, 336–351. DOI: 10.1002/jrsm.1310

Lortie, C. J., & Callaway, R. M. (2006). Re-analysis of meta-analysis: support for the stress-gradient hypothesis. *Journal of Ecology*, *94*, 7–16. DOI: 10.1111/j.1365-2745.2005.01066.x

Lortie, C. J. (2014). Formalized synthesis opportunities for ecology: systematic reviews and meta-analyses. *Oikos*, *123*, 897–902. DOI: 10.1111/j.1600-0706.2013.00970.x

Mengersen K, Jennions MD, Schmid CH. (2013). Statistical models for the meta-analysis of non-independent data. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton

Menegotto, A., Kurtz, M. N., & Lana, P. da C. (2019). Benthic habitats do show a significant latitudinal diversity gradient: A comment on Kinlock et al. (2018). *Global Ecology and Biogeography*, *28*, 1712–1717. DOI: 10.1111/geb.12970

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*, e1000097. DOI: 10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … Stewart, L. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*, 1. DOI: 10.1186/2046-4053-4-1

Møller, A. P., & Jennions, M. D. (2001). Testing and adjusting for publication bias. *Trends in Ecology and Evolution*, *16*, 580–586. DOI: 10.1016/S0169-5347(01)02235-2

de Morais, G. F., Ribas L. G. S., Ortega, J. C. G., Heino, J., & Bini, L. M. (2018). Biological surrogates: a word of caution. *Ecological Indicators*, *88*, 214–218. DOI: 10.1016/j.ecolind.2018.01.027

Morrissey, M. B. (2016). Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *Journal of Evolutionary Biology*, *29*, 1882–1904. DOI: 10.1111/jeb.12950

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evidence-Based Medicine*, *21*, 125–127. DOI: 10.1136/ebmed-2016-110401

Nakagawa, S., Noble, D. W. A., Senior, A. M., & Lagisz, M. (2017). Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biology*, *15*, 18. DOI: 10.1186/s12915-017-0357-7

Nakagawa, S., & Poulin, R. (2012). Meta-analytic insights into evolutionary ecology: an introduction and synthesis. *Evolutionary Ecology*, *26*, 1085–1099. DOI: 10.1007/s10682-012-9593-z

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, *26*, 1253–1274. DOI: 10.1007/s10682-012-9555-5

Noble, D. W. A., Lagisz, M., O'dea, R. E., & Nakagawa, S. (2017). Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Molecular Ecology*, *26*, 2410–2425. DOI: 10.1111/mec.14031

Olkin, I. (1992). Meta-analysis: methods for combining independent studies. *Statistical Science*, *7*, 226–236.

Ortega, J. C. G., Thomaz, S. M., & Bini, L. M. (2018). Experiments reveal that environmental heterogeneity increases species richness, but they are rarely designed to detect the underlying mechanisms. *Oecologia*, *188*, 11–22. DOI: 10.1007/s00442-018-4150-2

Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. In: *The handbook of research synthesis & meta-analysis* (Eds: H. Cooper, L. V. Hedges, & J. C. Valentine). Russell Sage Foundation, New York.

Padial, A. A., Nabout, J. C., Siqueira, T., Bini, L. M., & Diniz-Filho, J. A. F. (2010). Weak evidence for determinants of citation frequency in ecological articles. *Scientometrics*, *85*, 1-12.

Page, M. J., & Moher, D. (2016). Mass production of systematic reviews and meta-analyses: an exercise in mega-silliness? *Milbank Quarterly*, *94*, 515–519. DOI: 10.1111/1468-0009.12211

Parker, T. H., Bowman, S. D., Nakagawa, S., Gurevitch, J., Mellor, D. T., Rosenblatt, R. P., & DeHaven, A. C. (2018). Tools for Transparency in Ecology and Evolution (TTEE). Available in <https://osf.io/g65cb/> DOI: 10.17605/OSF.IO/G65CB

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, *26*, 4544–4562. DOI: 10.1002/sim.2889

Philibert, A., Loyce, C., & Makowski, D. (2012). Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems and Environment, 148*, 72–82. DOI: 10.1016/j.agee.2011.12.003

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roberts, P. D., Stewart, G. B., & Pullin, A. S. (2006). Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation*, *132*, 409–423. DOI: 10.1016/j.biocon.2006.04.034

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester, UK: John Wiley & Sons.

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S. A., & Nakagawa, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology*, *97*, 3293–3299. DOI: 10.1002/ecy.1591

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … The PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*, *349*, g7647. DOI: 10.1136/bmj.g7647

Song, S., Peacor, D. S., Osenberg, C. W., Bence, J. R. (2020). An assessment of statistical methods for nonindependent data in ecological meta-analyses. *Ecology,* 101(12). DOI:10.1002/ecy.3184

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17, 866–880. DOI: 10.1111/ele.12277

Sutton, A. J. (2002). Empirical assessment of effect of publication bias on meta-analyses. *BMJ*, *320*, 1574. DOI: 10.1136/bmj.320.7249.1574

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*, 13–30. DOI: 10.1002/jrsm.1091

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, *58*, 894–901. DOI: 10.1016/j.jclinepi.2005.01.006

Tomkins, J. L., & Kotiaho, J. S. (2004). Publication bias in meta-analysis: seeing the wood for the trees. *Oikos*, *104*, 194–196. DOI: 10.1111/j.0030-1299.2004.12357.x

Trikalinos, T. A., & Ioannidis, J. P. A. (2006). Assessing the evolution of effect sizes over time. In: *Publication bias in meta-analysis: prevention, assessment and adjustments* (Eds: H. R. Rothstein, A. J. Sutton & M. Borenstein). John Wiley & Sons, Chichester.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2014). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*, 1274–1294. DOI: 10.3758/s13428-014-0527-2

Vetter, D., Rucker, G., & Storch, I. (2013). Meta-analysis: a need for well-defined usage in ecology and conservation biology. *Ecosphere*, *4*, 1–24. DOI: 10.1890/ES13-00062.1

Weinhandl, E. D., & Duval, S. (2012). Generalization of trim and fill for application in meta-regression. *Research Synthesis Methods*, *3*, 51–67. DOI: 10.1002/jrsm.1042

Westgate, M. J., Barton, P. S., Lane, P. W., & Lindenmayer, D. B. (2014). Global meta-analysis reveals low consistency of biodiversity congruence relationships. *Nature Communications*, *5*, 3899. DOI: 10.1038/ncomms4899

Whittaker, R. J. (2010). Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness-productivity relationship. *Ecology*, *91*, 2522–2533. DOI: 10.1890/08-0968.1

Willis, B. H., & Quigley, M. (2011). The assessment of the quality of reporting of meta-analyses in diagnostic research: a systematic review. *BMC Medical Research Methodology, 11*, 163. DOI: 10.1186/1471-2288-11-163

Wilson, B. M., Harris, C. R., & Wixted, J. T. (2020). Science is not a signal detection problem. *Proceedings of the National Academy of Sciences, 117*, 5559-5567. https://doi.org/10.1073/pnas.1914237117

Wood, W., & Eagly, A. H. (2009). Advantages of certainty and uncertainty. In: *The handbook of research synthesis and meta-analysis* (Eds: H. Cooper, L. V. Hedges & J. C. Valentine). Russell Sage Foundation, New York. DOI: 9780871541635

Wooster, D. (1994). Predator impacts on stream benthic prey. *Oecologia*, *99*, 7–15. DOI: 10.1007/BF00317077

Supplementary Information

Table S1. 1. Complete reference of the studies included (n=114; see methods section for details on the selection procedure)

Number	Complete Reference
1	Albertson, L. K., D. C. Allen, & J. C. Trexler, 2015. Meta-analysis: Abundance, behavior, and hydraulic energy shape biotic effects on sediment transport in streams. Ecology .
2	Ballari, S. A., S. E. Kuebbing, & M. A. Nuñez, 2016. Potential problems of removing one invasive species at a time: a meta-analysis of the interactions between invasive vertebrates and unexpected effects of removal programs. PeerJ .
3	Barrett, T. J., R. B. Lowell, M. A. Tingley, & K. R. Munkittrick, 2010. Effects of pulp and paper mill effluent on fish: A temporal assessment of fish health across sampling cycles. Environmental Toxicology and Chemistry .
4	Bartels, P., J. Cucherousset, K. Steger, P. Eklov, L. J. Tranvik, & H. Hillebrand, 2012. Reciprocal subsidies between freshwater and terrestrial ecosystems structure consumer resource dynamics. Ecology .
5	Bell, T., W. E. Neill, & D. Schluter, 2003. The effect of temporal scale on the outcome of trophic cascade experiments. Oecologia .
6	Blanchet, S., O. Rey, R. Etienne, S. Lek, & G. Loot, 2010. Species-specific responses to landscape fragmentation: Implications for management strategies. Evolutionary Applications .
7	Blenckner, T., R. Adrian, D. M. Livingstone, E. Jennings, G. A. Weyhenmeyer, D. G. George, T. Jankowski, M. Järvinen, C. N. Aonghusa, T. Nõges, D. Straile, & K. Teubner, 2007. Large-scale climatic signatures in lakes across Europe: A meta-analysis. Global Change Biology 13: 1314–1326.
8	Borer, E. T., E. W. Seabloom, J. B. Shurin, K. E. Anderson, C. A. Blanchette, B. Broitman, S. D. Cooper, & B. S. Halpern, 2005. What determines the strength of a trophic cascade?. Ecology .
9	Brenden, T. O., B. R. Murphy, & E. M. Hallerman, 2006. Effect of Discharge on Daytime Habitat Use and Selection by Muskellunge in the New River, Virginia. Transactions of the American Fisheries Society .
10	Brett, M. T., & C. R. Goldman, 1996. A meta-analysis of the freshwater trophic cascade. Proc. Natl. Acad. Sci., USA 93: 7723–7726.
11	Brett, M. T., & C. R. Goldman, 1997. Consumer versus resource control in freshwater pelagic food webs. Science .
12	Bundschuh, M., M. O. Gessner, G. Fink, T. A. Ternes, C. S??gding, & R. Schulz, 2011. Ecotoxicologial evaluation of wastewater ozonation based on detritus-detritivore interactions. Chemosphere .
13	Cardinale, B. J., D. S. Srivastava, J. E. Duffy, J. P. Wright, A. L. Downing, M. Sankaran, & C. Jouseau, 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. Nature .
14	Chapman, J. M., C. L. Proulx, M. A. N. Veilleux, C. Levert, S. Bliss, M. È. André, N. W. R. Lapointe, & S. J. Cooke, 2014. Clear as mud: A meta-analysis on the effects of sedimentation on freshwater fish and the effectiveness of sediment-control measures. Water Research. .
15	Cheruvelil, K. S., P. A. Soranno, J. D. Madsen, & M. J. Roberson, 2002. Plant architecture and epiphytic macroinvertebrate communities: the role of an exotic dissected macrophyte. Journal of the North American Benthological Society .
16	Cusens, J., S. D. Wright, P. D. McBride, & L. N. Gillman, 2012. What is the form of the productivity-animal-species-richness relationship? a critical review and meta-analysis. Ecology. .
17	Dahl, J., & L. Greenberg, 1996. Impact on Stream Benthic Prey by Benthic vs Drift Feeding Predators: A Meta-Analysis. Oikos .
18	Darch, T., M. S. A. Blackwell, J. M. B. Hawkins, P. M. Haygarth, & D. Chadwick, 2014. A meta-analysis of organic and inorganic phosphorus in organic fertilizers, soils, and water: Implications for water quality. Critical Reviews in Environmental Science and Technology. .
19	Daufresne, M., & P. Boët, 2007. Climate change impacts on structure and diversity of fish communities in rivers. Global Change Biology .
20	Davenport, J. M., B. R. Hossack, & W. H. Lowe, 2014. Partitioning the non-consumptive effects of predators on prey with complex life histories. Oecologia .
21	Deines, A. M., D. B. Bunnell, M. W. Rogers, T. D. Beard, & W. W. Taylor, 2015. A review of the global relationship among freshwater fish, autotrophic activity, and regional climate. Reviews in Fish Biology and Fisheries. .

22 Drakare, S. . b, J. J. . Lennon, & H. . d Hillebrand, 2006. The imprint of the geographical, evolutionary and ecological context on species-area relationships. Ecology Letters .

23 Earl, J. E., & H. H. Whiteman, 2015. Are Commonly Used Fitness Predictors Accurate? A Meta-analysis of Amphibian Size and Age at Metamorphosis. Copeia .

24 Englund, G., O. Sarnelle, & S. D. Cooper, 1999. The importance of data-selection criteria: Meta-analyses of stream predation experiments. Ecology. .

25 Evans-White, M. A., & H. M. Halvorson, 2017. Comparing the ecological stoichiometry in green and brown food webs - A review and meta-analysis of freshwater food webs. Frontiers in Microbiology. .

26 Faithfull, C. L., A. K. Bergström, & T. Vrede, 2011. Effects of nutrients and physical lake characteristics on bacterial and phytoplankton production: A meta-analysis. Limnology and Oceanography .

27 Feminella, J. W., & C. P. Hawkins, 1995. Interactions between Stream Herbivores and Periphyton: A Quantitative Analysis of past Experiments. Journal of the North American Benthological Society .

28 Ferreira, V., B. Castagneyrol, J. Koricheva, V. Gulis, E. Chauvet, & M. A. S. Graça, 2015. A meta-analysis of the effects of nutrient enrichment on litter decomposition in streams. Biological Reviews .

29 Ferreira, V., & F. Guérold, 2017. Leaf litter decomposition as a bioassessment tool of acidification effects in streams: Evidence from a field study and meta-analysis. Ecological Indicators .

30 Ferreira, V., J. Koricheva, S. Duarte, D. K. Niyogi, & F. Guérold, 2016a. Effects of anthropogenic heavy metal contamination on litter decomposition in streams - A meta-analysis. Environmental Pollution. .

31 Ferreira, V., J. Koricheva, J. Pozo, & M. A. S. Graça, 2016b. A meta-analysis on the effects of changes in the composition of native forests on litter decomposition in streams. Forest Ecology and Management .

32 Francoeur, S. N., 2001. Meta-analysis of lotic nutrient amendment experiments: detecting and quantifying subtle responses. Journal of the North American Benthological Society .

33 Fukushima, M., 2001. Salmonid habitat-geomorphology relationships in low-gradient streams. Ecology .

34 Gallardo, B., M. Clavero, M. I. Sánchez, & M. Vilà, 2016. Global ecological impacts of invasive species in aquatic ecosystems. Global Change Biology. .

35 González, A. L., J. S. Kominoski, M. Danger, S. Ishida, N. Iwai, & A. Rubach, 2010. Can ecological stoichiometry help explain patterns of biological invasions?. Oikos .

36 Harris, S., M. Cormican, & E. Cummins, 2012. The effect of conventional wastewater treatment on the levels of antimicrobial-resistant bacteria in effluent: A meta-analysis of current studies. Environmental Geochemistry and Health .

37 Haxton, T. J., & C. S. Findlay, 2008. Meta-analysis of the impacts of water management on aquatic communities. Canadian Journal of Fisheries and Aquatic Sciences .

38 Hidding, B., E. S. Bakker, M. J. M. Hootsmans, & S. Hilt, 2016. Synergy between shading and herbivory triggers macrophyte loss and regime shifts in aquatic systems. Oikos .

39 Higglns, S. N., & M. J. Vander Zanden, 2010. What a difference a species makes: A meta-analysis of dreissenid mussel impacts on freshwater ecosystems. Ecological Monographs .

40 Hillebrand, H., 2002. Top-down versus bottom-up control of autotrophic biomass—a meta-analysis on experiments with periphyton. Journal of the North American Benthological Society .

41 Hillebrand, H., 2005. Light regime and consumer control of autotrophic biomass. Journal of Ecology .

42 Hillebrand, H., 2008. Grazing regulates the spatial variability of periphyton biomass. Ecology .

43 Hillebrand, H., P. Frost, & A. Liess, 2008. Ecological stoichiometry of indirect grazer effects on periphyton nutrient content. Oecologia .

44 Iacarella, J. C., J. T. A. Dick, M. E. Alexander, & A. Ricciardi, 2015. Ecological impacts of invasive alien species along temperature gradients: Testing the role of environmental matching. Ecological Applications .

45 James, J., F. M. Slater, I. P. Vaughan, K. A. Young, & J. Cable, 2015. Comparing the ecological impacts of native and invasive crayfish: could native species' translocation do more harm than good? Oecologia .

46 Jiang, L., & Z. Pu, 2009. Different Effects of Species Diversity on Temporal Stability in Single-Trophic and Multitrophic Communities. The American Naturalist .

47 Kadykalo, A. N., & C. S. Findlay, 2016. The flow regulation services of wetlands. Ecosystem Services .

48 Kail, J., K. Brabec, M. Poppe, & K. Januschke, 2015. The effect of river restoration on fish, macroinvertebrates and aquatic macrophytes: A meta-analysis. Ecological Indicators .

49 Katano, I., H. Doi, B. K. Eriksson, & H. Hillebrand, 2015. A cross-system meta-analysis reveals

coupled predation effects on prey biomass and diversity. Oikos .

50 Kennedy, K. T. M., & R. W. El-Sabaawi, 2017. A global meta-analysis of exotic versus native leaf decay in stream ecosystems. Freshwater Biology .

51 Khan, F., J. Pilz, & S. Ali, 2017. Improved hydrological projections and reservoir management in the Upper Indus Basin under the changing climate. Water and Environment Journal .

52 Kitada, S., H. Kishino, & K. Hamasaki, 2011. Bias and significance of relative reproductive success estimates based on steelhead trout (Oncorhynchus mykiss) data: a Bayesian meta-analysis. Canadian Journal of Fisheries and Aquatic Sciences .

53 Korsu, K., A. Huusko, & T. Muotk, 2010. Impacts of invasive stream salmonids on native fish: Using meta-analysis to summarize four decades of research. Boreal Environment Research .

54 Kozlowsky-Suzuki, B., A. E. Wilson, & A. da Silva Ferrão-Filho, 2012. Biomagnification or biodilution of microcystins in aquatic foodwebs? Meta-analyses of laboratory and field studies. Harmful Algae .

55 Kreye, M. M., D. C. Adams, & F. J. Escobedo, 2014. The value of forest conservation for water quality protection. Forests .

56 Kuiper, J. J., J. H. Janse, S. Teurlincx, J. T. A. Verhoeven, & R. Alkemade, 2014. The impact of river regulation on the biodiversity intactness of floodplain wetlands. Wetlands Ecology and Management .

57 Lecerf, A., & E. Chauvet, 2008. Diversity and functions of leaf-decaying fungi in human-altered streams. Freshwater Biology .

58 Lewandowska, A. M., A. Biermann, E. T. Borer, M. A. Cebrián-Piqueras, S. A. J. Declerck, L. De Meester, E. Van Donk, L. Gamfeldt, D. S. Gruner, N. Hagenah, W. S. Harpole, K. P. Kirkman, C. A. Klausmeier, M. Kleyer, J. M. H. Knops, P. Lemmens, E. M. Lind, E. Litchman, J. Mantilla-Contreras, K. Martens, S. Meier, V. Minden, J. L. Moore, H. O. Venterink, E. W. Seabloom, U. Sommer, M. Striebel, A. Trenkamp, J. Trinogga, J. Urabe, W. Vyverman, D. B. Van de Waal, C. E. Widdicombe, & H. Hillebrand, 2016. The influence of balanced and imbalanced resource supply on biodiversity-functioning relationship across ecosystems. Philosophical Transactions of the Royal Society B: Biological Sciences .

59 Liess, A., & H. Hillebrand, 2004. Invited review: Direct and indirect effects in herbivore – periphyton interactions. Archiv für Hydrobiologie .

60 Liew, J. H., H. H. Tan, & D. C. J. Yeo, 2016. Dammed rivers: impoundments facilitate fish invasions. Freshwater Biology .

61 Lucas, R. W., J. Klaminder, M. N. Futter, K. H. Bishop, G. Egnell, H. Laudon, & P. Högberg, 2011. A meta-analysis of the effects of nitrogen additions on base cations: Implications for plants, soils, and streams. Forest Ecology and Management. .

62 Mancinelli, G., F. Sangiorgio, & A. Scalzo, 2013. The effects of decapod crustacean macroconsumers on leaf detritus processing and colonization by invertebrates in stream habitats: A meta-analysis. International Review of Hydrobiology .

63 Mant, R. C., D. L. Jones, B. Reynolds, S. J. Ormerod, & A. S. Pullin, 2013. A systematic review of the effectiveness of liming to mitigate impacts of river acidification on fish and macro-invertebrates. Environmental Pollution. .

64 Marczak, L. B., R. M. Thompson, & J. S. Richardson, 2007. Meta-analysis: Trophic level, habitat, and productivity shape the food web effects of resource subsidies. Ecology .

65 Martin, R. A., & J. A. Harrison, 2011. Effect of High Flow Events on In-Stream Dissolved Organic Nitrogen Concentration. Ecosystems .

66 Mattson, M. D., P. J. Godfrey, M. F. Walk, P. A. Kerr, & O. T. Zajicek, 1997. Evidence of recovery from acidification in Massachusetts streams. Water, Air, and Soil Pollution .

67 McCarthy, J. M., C. L. Hein, J. D. Olden, & M. Jake Vander Zanden, 2006. Coupling long-term studies with meta-analysis to investigate impacts of non-native crayfish on zoobenthic communities. Freshwater Biology .

68 Mcmullen, L. E., & D. A. Lytle, 2012. Quantifying invertebrate resistance to floods: A global-scale meta-analysis. Ecological Applications .

69 Meissner, K., & T. Muotka, 2006. The role of trout in stream food webs: Integrating evidence from field surveys and experiments. Journal of Animal Ecology .

70 Melvin, S. D., & S. P. Wilson, 2013. The utility of behavioral studies for aquatic toxicology testing: A meta-analysis. Chemosphere .

71 Miller, S. W., P. Budy, & J. C. Schmidt, 2010. Quantifying macroinvertebrate responses to in-stream habitat restoration: Applications of meta-analysis to river restoration. Restoration Ecology. .

72 Monaghan, M. T., S. A. Thomas, G. W. Minshall, J. D. Newbold, & C. E. Cushing, 2001. The influence of filter-feeding benthic macroinvertebrates on the transport and deposition of particulate organic matter and diatoms in two streams. Limnology and Oceanography .

73 Murphy, G. E. P., & T. N. Romanuk, 2012. A Meta-Analysis of Community Response Predictability to Anthropogenic Disturbances. The American Naturalist .

74 Murphy, G. E. P., & T. N. Romanuk, 2014. A meta-analysis of declines in local species richness from human disturbances. Ecology and Evolution .

75 Myers, R. A., M. J. Bradford, J. M. Bridson, & G. Mertz, 1997. Estimating delayed density-dependent mortality in sockeye salmon (Oncorhynchus nerka): a meta-analytic approach. Canadian Journal of Fisheries and Aquatic Sciences 54: 2449–2462, http://www.nrcresearchpress.com/doi/abs/10.1139/f97-146.

76 Orrock, J. L., E. L. Preisser, J. H. Grabowski, & G. C. Trussell, 2013. The cost of safety: Refuges increase the impact of predation risk in aquatic systems. Ecology .

77 Peipoch, M., E. Martí, & E. Gacia, 2012. Variability in δ 15 N natural abundance of basal resources in fluvial ecosystems: a meta-analysis. Freshwater Science .

78 Peng, S., H. Liao, T. Zhou, & S. Peng, 2017. Effects of UVB radiation on freshwater biota: a meta-analysis. Global Ecology and Biogeography .

79 Persson, J., P. Fink, A. Goto, J. M. Hood, J. Jonas, & S. Kato, 2010. To be or not to be what you eat: Regulation of stoichiometric homeostasis among autotrophs and heterotrophs. Oikos .

80 Petrin, Z., G. Englund, & B. Malmqvist, 2008. Contrasting effects of anthropogenic and natural acidity in streams: A meta-analysis. Proceedings of the Royal Society B: Biological Sciences .

81 Poulin, R., 2000. Variation in the intraspecific relationship between fish length and intensity of parasitic infection: Biological and statistical causes. Journal of Fish Biology .

82 Preisser, E. L., D. I. Bolnick, & M. F. Benard, 2005. Scared to death? The effects of intimidation and consumption in predator-prey interactions. Ecology .

83 Presa Abós, C., F. Lepori, B. G. Mckie, & B. Malmqvist, 2006. Aggregation among resource patches can promote coexistence in stream-living shredders. Freshwater Biology .

84 Quesnelle, P. E., K. E. Lindsay, & L. Fahrig, 2015. Relative effects of landscape-scale wetland amount and landscape matrix quality on wetland vertebrates: A meta-analysis. Ecological Applications .

85 Reeves, J. L., & P. D. Lorch, 2012. Biological control of invasive aquatic and wetland plants by arthropods: A meta-analysis of data from the last three decades. BioControl .

86 Richardson, J. S., & S. Béraud, 2014. Effects of riparian forest harvest on streams: A meta-analysis. Journal of Applied Ecology .

87 Rosenfeld, J., S. L. Whiteway, P. M. Biron, A. Zimmermann, O. Venter, & J. W. A. Grant, 2010. Do in-stream restoration structures enhance salmonid abundance? A meta-analysis. Canadian Journal of Fisheries and Aquatic Sciences .

88 Sales, J., 2011. First feeding of freshwater fish larvae with live feed versus compound diets: A meta-analysis. Aquaculture International.

89 Sarà, G., 2007. Ecological effects of aquaculture on living and non-living suspended fractions of the water column: A meta-analysis. Water Research. .

90 Schill, D. J., & R. L. Scarpella, 1997. Barbed Hook Restrictions in Catch-and-Release Trout Fisheries: A Social Issue. North American Journal of Fisheries Management .

91 Shuman-Goodier, M. E., & C. R. Propper, 2016. A meta-analysis synthesizing the effects of pesticides on swim speed and activity of aquatic vertebrates. Science of the Total Environment.

92 Sievers, M., R. Hale, & J. R. Morrongiello, 2017. Do trout respond to riparian change? A meta-analysis with implications for restoration and management. Freshwater Biology. .

93 Sistla, S. A., A. P. Appling, A. M. Lewandowska, B. N. Taylor, & A. A. Wolf, 2015. Stoichiometric flexibility in response to fertilization along gradients of environmental and organismal nutrient richness. Oikos .

94 Smokorowski, K. E., & T. C. Pratt, 2007. Effect of a change in physical structure and cover on fish and fish habitat in freshwater ecosystems – a review and meta-analysis. Environmental Reviews .

95 Smucker, N. J., & N. E. Detenbeck, 2014. Meta-Analysis of Lost Ecosystem Attributes in Urban Streams and the Effectiveness of Out-of-Channel Management Practices. Restoration Ecology.

96 Soria, M., C. Leigh, T. Datry, L. M. Bini, & N. Bonada, 2017. Biodiversity in perennial and intermittent rivers: a meta-analysis. Oikos .

97 Sorte, C. J. B., I. Ibáñez, D. M. Blumenthal, N. A. Molinari, L. P. Miller, E. D. Grosholz, J. M. Diez, C. M. D'Antonio, J. D. Olden, S. J. Jones, & J. S. Dukes, 2013. Poised to prosper? A cross-system comparison of climate change effects on native and non-native species performance. Ecology Letters .

98 Stehle, S., & R. Schulz, 2015. Agricultural insecticides threaten surface waters at the global scale. Proceedings of the National Academy of Sciences .

99 Stewart, G. B., H. R. Bayliss, D. A. Showler, W. J. Sutherland, & A. S. Pullin, 2009. Effectiveness of engineered in-stream structure mitigation measures to increase salmonid abundance: A systematic review. Ecological Applications .

100	Takimoto, G., & D. M. Post, 2013. Environmental determinants of food-chain length: A meta-analysis. Ecological Research .
101	Taranu, Z. E., & I. Gregory-Eaves, 2008. Quantifying relationships among phosphorus, agriculture, and lake depth at an inter-regional scale. Ecosystems .
102	Taylor, M. K., & S. J. Cooke, 2012. Meta-analyses of the effects of river flow on fish movement and activity. Environmental Reviews .
103	Tejedo, M., F. Marangoni, C. Pertoldi, A. Richter-Boix, A. Laurila, G. Orizaola, A. G. Nicieza, D. Álvarez, & I. Gomez-Mestre, 2010. Contrasting effects of environmental factors during larval stage on morphological plasticity in post-metamorphic frogs. Climate Research .
104	Tillmanns, A. R., A. E. Wilson, F. R. Pick, & O. Sarnelle, 2008. Meta-analysis of cyanobacterial effects on zooplankton population growth rate: species-specific responses. Fundamental and Applied Limnology / Archiv für Hydrobiologie .
105	Twardochleb, L. A., J. D. Olden, & E. R. Larson, 2013. A global meta-analysis of the ecological impacts of nonnative crayfish. Freshwater Science .
106	Velghe, K., & I. Gregory-Eaves, 2013. Body Size Is a Significant Predictor of Congruency in Species Richness Patterns: A Meta-Analysis of Aquatic Studies. PLoS ONE 8:.
107	Voje, K. L., A. E. I. Dworkin, & E. Y. Michalakis, 2016. Scaling of Morphological Characters across Trait Type, Sex, and Environment: A Meta-analysis of Static Allometries. The American Naturalist .
108	Ward, J. M., & A. Ricciardi, 2007. Impacts of Dreissena invasions on benthic macroinvertebrate communities: A meta-analysis: Biodiversity research. Diversity and Distributions .
109	Wesner, J. S., 2016. Contrasting effects of fish predation on benthic versus emerging prey: a meta-analysis. Oecologia .
110	Wilson, A. E., O. Sarnelle, & A. R. Tillmanns, 2006. Effects of cyanobacterial toxicity and morphology on the population growth of freshwater zooplankton: Meta-analyses of laboratory experiments. Limnology and Oceanography. .
111	Wittmann, M. E., C. L. Jerde, J. G. Howeth, S. P. Maher, A. M. Deines, J. A. Jenkins, G. W. Whitledge, S. R. Burbank, W. L. Chadderton, A. R. Mahon, J. T. Tyson, C. A. Gantz, R. P. Keller, J. M. Drake, D. M. Lodge, & A. Fisk, 2014. Grass carp in the Great Lakes region: establishment potential, expert perceptions, and re-evaluation of experimental evidence of ecological impact. Canadian Journal of Fisheries and Aquatic Sciences .
112	Wood, K. A., M. T. O'hare, C. McDonald, K. R. Searle, F. Daunt, & R. A. Stillman, 2017. Herbivore regulation of plant abundance in aquatic ecosystems. Biological Reviews .
113	Wooster, D., 1994. Predator impacts on stream benthic prey. Oecologia .
114	Yang, L. H., K. F. Edwards, J. E. Byrnes, J. L. Bastow, A. N. Wright, & K. O. Spence, 2010. A meta-analysis of resource pulse-consumer interactions. Ecological Monographs .

Chapter II

But the emperor is [still] not wearing anything at all! – excessive positive results in limnology

Resumo

O viés de publicação ocorre quando a magnitude ou a direção dos resultados diferem entre estudos publicados e não publicados. Os estudos de meta-análise tendem a incluir principalmente artigos com resultados positivos (ou seja, resultados significativos que apoiam a hipótese), uma vez que são mais propensos a serem aceitos para publicação. Essa tendência aumenta os possíveis problemas com viés de publicação. Avaliei a prevalência de viés de resultados positivos em meta-análises de ecologia de água doce, dividindo os resultados em três conjuntos de dados diferentes. Os resultados positivos prevaleceram sobre os resultados negativos para todos os conjuntos de dados considerados. Resultados possivelmente percebidos como mais relevantes tiveram maior incidência de resultados positivos. Além disso, a incidência de resultados positivos em tamanhos de efeito individuais e resultados de moderadores não foram relacionados ao país dos autores. A incidência de resultados positivos para efeitos globais foi positiva e não-linearmente relacionada ao ano de publicação. A alta prevalência de resultados positivos entre meta-análises pode refletir vieses em duas escalas diferentes, especificamente, em estudos primários e dentro das próprias meta-análises.

Abstract

Publication bias is said to occur when the strength or direction of the results differ between published and unpublished studies. Meta-analysis studies tend to include mostly articles with positive results (i.e., significant results supporting a hypothesis), which are more likely to be accepted by scientific journals. This tendency heightens possible issues with publication bias. I evaluated the prevalence of positive results bias in freshwater ecology meta-analyses, dividing the results into three different datasets covering different scales of results. Positive results prevailed over negative results for all the datasets considered. Results possibly perceived as more relevant had a higher incidence of positive results. Also, the incidence of positive results in individual effect sizes and moderator results was unrelated to the authors' country. The incidence of positive results for summary effects was positively and non-linearly related to publication year. A high prevalence of positive results among meta-analyses may reflect biases in two different scales, specifically, in primary studies and within the meta-analyses itself.

Introduction

The problem with the excessive importance given to statistical significance is by no means a novelty (Bakan, 1966; Cohen, 1994; Møller & Jennions, 2001; Jennions & Møller, 2002; Ioannidis & Trikalinos, 2007a; Fanelli, 2010a, 2012). Bakan (1966) reported that a "*great deal of mischief*" has been associated with the use of *P*-values and acknowledged that stating the existence of such mischief is "*to assume the role of the child who pointed out that the emperor was really outfitted only in his underwear*". It may seem obvious, but the emperor is apparently still running around without his clothes, and the aforementioned 'mischief' may include bias against negative and non-significant results, selective reporting and different types of questionable research practices (Ioannidis, 2005; Ioannidis & Trikalinos, 2007a; Fanelli, 2012; Fanelli et al., 2017; Van Aert et al., 2019).

Publication bias is said to occur when the strength or direction of the results differ between published and unpublished studies (Møller & Jennions, 2001). It may be defined as "the tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or the strength of the study findings" (Dickersin, 1990) or yet as "the selective publishing of research based on the nature and direction of findings, that occurs when studies with significant or favorable results are more likely to be published than those with nonsignificant or unfavorable findings"(Marks-Anglin & Chen, 2020). Studies with significant results that also support their main hypotheses, are more likely to be published than studies reporting non-significant and negative results (Dickersin & Min, 1993) due to several reasons (see below). In general, publication bias may be perceived as an umbrella under which other biases are encountered.

Statistical significance was a major determinant of the decision to submit and of the time it took for a paper in clinical trials to be submitted and published (Ioannidis, 1998), showing that publication bias may occur in all stages involving the publication of an article (Møller & Jennions, 2001). Non-significant results are often considered uninteresting by investigators (Dickersin & Min, 1993), and may never be submitted or published (Knight, 2003), contributing to the file-drawer problem (Rosenthal, 1979; and see Knight, 2003; Fanelli, 2012; Graqvist, 2015). In addition, more prestigious journals tend to publish a lower proportion of non-significant results (Csada et al., 1996; Browman, 1999; Knight, 2003; Lawrence, 2003), directly or indirectly pressing for statistical significance, once funding agencies take into account journal prestige when assessing to whom funds will be directed (Csada et al., 1996). More troubling is when 'significant' results are the product of data fabrication or unreported post-hoc re-interpretation, re-analysis, and selection. Unfortunately, fabrication bias (Borenstein et al., 2009c; Ioannidis, 2010) has been documented in medicine (Ranstam et al., 2000; Al-Marzouki et al., 2005; Gardner et al., 2005) and may occur in other disciplines as well.

Finally, bias towards positive results entails our tendency to confirm expected or desired results (Rosenthal & Fode, 1963). To obtain the expected results authors may change hypotheses after knowing the results (Forstmeier et al., 2017; Fraser et al., 2018), select parts of the results to be reported (selective reporting) and/or try different analytical methods until reaching the desired result (Simmons et al., 2011). As reported by Csada et. al. (1996), "*this raises the question as to how many studies reported to have significant results are nothing more than a researcher's expectations of such results*" and the question "*of how much this contributes to the publishing of fraudulent results*".

These biases against non-significant results may bring serious consequences to science (Fanelli, 2012). It was regarded by Stephen Jay Gould as "*perhaps the most important effectively undiscussed subject in the entire methodology of science*" (Browman, 1999). The scientific progress relies heavily on unexpected results, without which Popper's falsifiability (Popper, 2005) would never occur, Khuns' paradigm (Kuhn, 1996) would never shift, and nor would Lakatos' auxiliary hypotheses (Lakatos, 2014) be altered or abandoned with the onset of new/different discoveries (see Alatalo et al., 1997 for an example of the importance of negative results for paradigm shifts in science). In medicine, for instance, an ineffective drug could cease to be used by the population by the onset of a published negative result. Of course, the undesirable and harmful consequences are more visible, or perceived by society at large.

Ecology and other fields of knowledge would equally be affected by a possible resistance to the negative results, which is comprehensible. For instance, take climate change, a serious matter that has been evidenced extensively. Even if the evidence was misinterpreted (type I error), we will have increased renewable energy sources, and reduced air pollution, oil exploration in sensitive ecosystems and anthropogenic fires. On the other hand, if we tolerate a high type II error rate and disregard the existence of climate change, the consequences may be catastrophic (Oreskes, 2019). Our resistance to accepting negative results is ultimately a debate between consequences of accepting a false positive result (type I error) or rejecting a true positive result (type II error). The issue with simply accepting results as they come lies in the fact that many published results are type I errors, and these results are too rarely discredited (Parker & Nakagawa, 2014); however, type II error may implicate in more serious consequences (Forstmeier et al., 2017).

Meta-analysis is a powerful tool that is being increasingly used in different areas and consists of weighting the effect sizes obtained from primary studies by the inverse of their variances and of searching for correlates of the heterogeneity among effect sizes (Borenstein et al., 2009a; Koricheva & Gurevitch, 2013; Gurevitch et al., 2018). Tough criticisms have been made to the use of meta-analysis (Berlin & Golub, 2014), among which the 'garbage-in, garbage-out' argument (Ioannidis, 2010; Whittaker, 2010; Berlin & Golub, 2014), once it's reliability depends highly on the quality of primary studies, which in turn are severely biased towards positive results (Fanelli, 2012). Meta-analysis studies tend to include mostly articles with positive results (i.e., significant results supporting a hypothesis), which are more likely to be accepted by scientific journals. This tendency heightens possible issues with publication bias (Hopewell et al., 2006) and raises the question of whether the tools available for dealing with publication bias are being effective (Ioannidis, 2010). Indeed, according to Ioannidis (2010), "it would be naïve to think we are able to filter out or control all the effects of such biases".

Freshwater ecology is a field that brings relevant contributions to the preservation and conservation of waterbodies, especially given the growing issues with water quality throughout the world. As such, evaluating the possibility of bias and consequently assessing reliability of results has important implications which varies among scientific fields and publication environments (Ioannidis, 2010), and therefore must be evaluated specifically for each discipline. In this paper I will evaluate the incidence of positive results in meta-analyses of freshwater ecology. I will focus on the proportions of positive results (the significant results supporting the hypothesis) and negative results (significant results against the hypothesis and non-significant results; following Fanelli, 2012). I expect to (i) find a higher proportion of positive results compared to negative and non-significant results in each meta-analysis evaluated. I also predict that the proportion of positive results per study would be (ii) negatively associated with time, (iii) more associated with countries where science is more competitive, and (iii) negatively so with the number of authors involved in the publication. I expect a negative association with time as an indication of increasing awareness of publication bias, given the increasing literature on the importance of publishing negative findings. In turn, I hypothesized that more competitive countries will tend towards having a higher incidence of bias towards positive results and of selective reporting due to a higher pressure to publish (following Fanelli, 2010b; Fanelli et al., 2015). I also expect that the number of authors involved in a study is inversely related with the proportion of positive results once having more scientists involved could buffer the incidence of some biases (Ioannidis, 2010; Fanelli et al., 2017). Finally, I also expect that the number people holding a doctoral degree in the country of the first author will also be negatively associated with the prevalence of positive results, indicating an increased knowledge and responsibility with the papers published.

Methods

Data collection

I compiled studies from the Web of Science and Scopus databases published from 1991 to 2017. The detailed search string and initial hits are shown in Table 2.1. The Boolean operators were adjusted to each search database, but I used the same words and wildcard functions for both. Subsequently, I screened the studies following the Prisma protocol (Moher et al., 2009), including only formal meta-analyses in freshwater ecology and obtained a total of 113 meta-analyses (Lodi et al., 2021). I included studies that conducted formal meta-analyses about freshwater ecosystems; however, I also included studies that evaluated several systems at once (terrestrial, oceanic, freshwater, etc.) if freshwater environment was evaluated (provided an individual effect size). I excluded studies in any field of expertise other than freshwater ecology and studies carried out within the freshwater environment, but that did study freshwater dynamics or organisms (e.g., Yu and collaborators (2017) conducted soil analysis of a restored wetland). A total of 21 of these papers were also excluded for not including enough information to infer the direction of the hypothesis (see dataset in Table S2.1).

Table 2.1. The search string used in the Web of Science (WoS) platform and the number of hits in both WoS and Scopus databases.

	Search string	**WoS hits**	**Scopus hits**
Meta-analysis terms	(meta-anal* OR metanal* OR "quantitative review")	130.960	199.481
Freshwater ecology terms	((freshwater OR aquatic* OR limnol* OR "inland water*" OR river* OR stream* OR creek* OR reservoir* OR lake* OR lagoon* OR pond* OR mere* OR loch* OR lakelet*) NOT (ocean* OR marine))	1.001.200	1.481.920
(**Meta-analysis terms**) AND (**Freshwater ecology terms**)		443	382

(*) Designates a wildcard function that includes all possible word endings, and quoted words (“”) represent searches of exact phrases.

Classification of the studies

I used a fine-grained measure of positive results (finer than Fanelli, 2012, for example), as suggested by Fanelli (2017), following a more inclusive approach. For each paper, I extracted the hypothesis, and the significance and direction of each result. Results were either for summary effect sizes, effects of moderators, or individual effect sizes. I categorized the results as ‘positive’, ‘negative’ or ‘not significant’. Results were classified as ‘positive’ when they confirmed the hypotheses of the studies and ‘negative’ when the results did not support the hypotheses either because the tests were not significant or because they were significant but in the opposite direction specified by the hypotheses. Some studies did not explicitly present a directional hypothesis. In such cases I read the paper and observed the direction of the hypothesis following the conceptual construct presented by the authors. When this was not possible, I considered that the authors expected to confirm theoretical predictions most often confirmed in the literature. I excluded all papers for which neither of these approaches were feasible. I obtained data in different resolutions instead of including only one hypothesis per paper and considering only clear direct hypothesis (as done by Fanelli, 2012) to obtain data with a finer resolution and broader scope.

I created three datasets (ds), one with the results of summary effect sizes (ds1), a second one with effects of moderators (ds2), and a third with the results of individual effect sizes (ds3). This multi-level approach is more appropriate to evaluate 'longer publications with multiple results' (Fanelli, 2018). Multiple results were presented by most papers, for all datasets (summary effects included). Therefore, I calculated the ratio of positive to negative and non-significant results (positive/(positive+negative+non-significant)) to produce a measure called "proportion of positive results" (PPR). My *sample unit* is the results of a meta-analysis within each ds category, and the *sample size* (n_{ds}) is the number of articles that provided summary effect sizes, moderators or individual effect size results, once multiple results were compiled into one measure per article. I obtained 33 results for summary effect sizes (n_{ds1}=33), 48 for moderators (n_{ds2}=48) and 81 for individual effect sizes (n_{ds3}=81) from a total of 114 selected papers.

To complement the information on the significance and direction of the results, I also obtained the year of publication, first author's country, and the number of authors included in the paper using the function *biblioAnalysis()* available in the R package Bibliometrix (Aria & Cuccurullo, 2017). The first authors' countries were grouped to reduce the number of levels (following Fanelli, 2012) as follows: NAm: United States of America and Canada; EU: Austria, Czech republic, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal, Spain, Sweden, United Kingdom, Norway; AS: China, Hong Kong, India, Japan, Singapore, South Korea, Taiwan; and NZ-AM: New Zealand, Argentina, Chile. I also obtained the proportion of 25-64-year-olds people with a doctorate degree per country from the Organisation for Economic Cooperation and Development (OECD) library (available at https://www.oecd-ilibrary.org/education/education-at-a-glance-2019_f8d7880d-en).

Data analysis

I evaluated whether positive results for summary effect sizes (ds1) prevailed over negative ones, and if the incidence of positive results was associated with publication year, corresponding author's country, number of authors and proportion of doctorate degrees per country, using a nonlinear least squares and a linear model. The variable year exhibited a clear nonlinear pattern in ds1, such that I was not able to fit a reliable linear model including this variable directly. I started with simpler approaches using the arcsine and logarithmic transformations and fitting linear models with different distributions. However, the residues of simpler analyses did not meet the required assumptions, i.e., were not suitable. Finally, I fitted a nonlinear model (function *nls()* in R) with the PPR of summary effects as response variable and publication year as the explanatory variable and extracted the residuals to be used as response variable in the linear model.

The full model for ds1 consisted of the residuals of the nonlinear model as the response variable, and country, number of authors involved in the publication and proportion of doctorate degrees per country as the explanatory variables. The PPR obtained for ds2 and for ds3 were used as response variables in two linear models, one for each dataset, each using publication year, country, number of authors and proportion of doctorate degrees per country as the explanatory variables. All assumptions of heteroscedasticity and residual normality were met.

Results

The average number of positive results surpassed that of negative results for all the datasets considered (Figure 2.1a). In addition, the confidence interval of negative results overlapped zero in all cases (Figure 2.1a). Consequently, the average PPR were greater that 0.5 in the different datasets (Figure 2.1b).

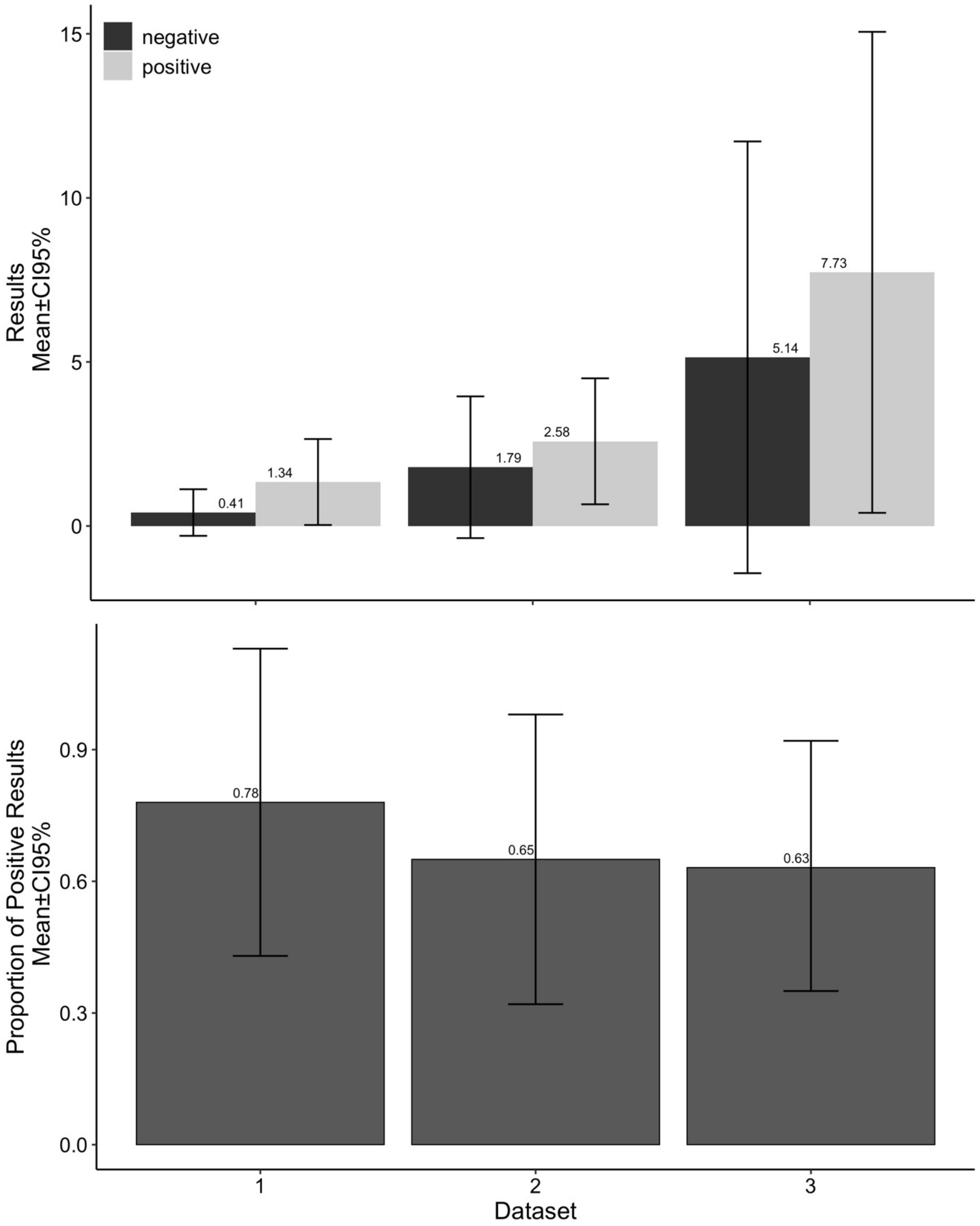


Figure 2.1(a) Mean positive and negative results obtained from the papers evaluated for each dataset and (b) ratio of positive to negative results, per dataset. The ratio varies from 0 to 1. Numbers shown above the bars are the mean PPR values and the error bars are 95% confidence intervals.

The incidence of positive results recorded for summary effects (ds1) was positively and non-linearly related to publication year (Table 2.2). Therefore, I fitted the linear model using the residues of the non-linear model as the response variable. The incidence of positive results varied among countries (Table 2.2), such that studies published in Europe had the highest mean positive results (as estimated in this study) and was followed by studies developed in North America (Figure 2.2). However, North American studies exhibited a higher variation on the incidence of the positive results. The locations with the lowest incidence of positive results were New Zealand and South America.

Table 2.2. Proportion of positive results for summary effect sizes (ds1) modeled in response to publication year, location, and number of authors. A nonlinear regression (PPR ~ year) was first used due to the non-linear effect of the year (first part of the table) and the residuals were used as response variable in the linear model (second part of the table). The full model consisted of the residuals of the nonlinear model as the response variable, and country, number of authors involved in the publication and shares of doctorate degrees per country as the explanatory variables.

DS1		**Estimate**	**Std. Error**	**t-value**	**P**	
	Asymptote	0.83	0.06	13.76	>0.001	
	Mean	2003	1.28	1562.40	>0.002	
	Coefficient	0.63	1.45	0.44	0.67	
	Residual standard error: 0.3247 on 29 degrees of freedom					
	Number of iterations to convergence: 18					
	Achieved convergence tolerance: 9.962e-06					
		df	**SQ**	**MSQ**	**F**	**P**
	Country	2	1.00	0.50	6.92	**0.00**
	Number of authors	1	0.08	0.08	1.06	0.31
	PhD graduates	1	0.04	0.04	0.50	0.49
	Residuals	27	1.95	0.07		

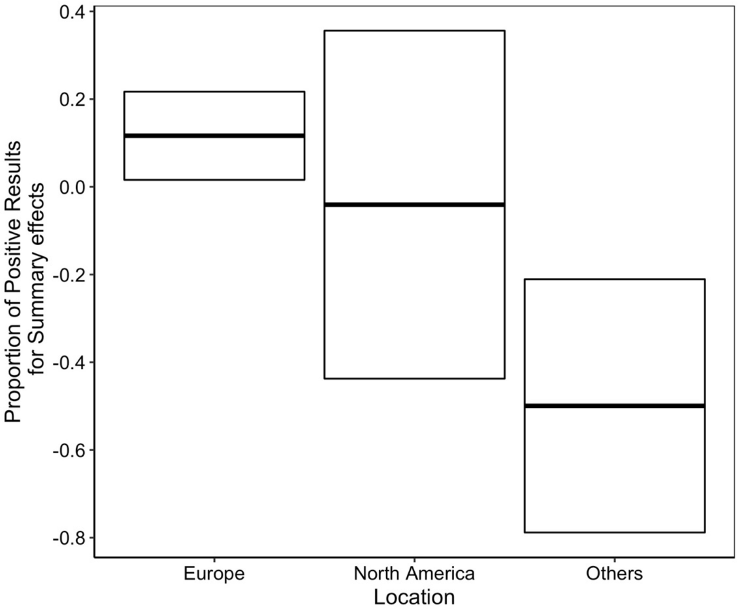


Figure 2.2. Proportion of positive results depending on the location where the article was published. Darker lines are the means and boxes represent the standard deviations. The PPR ds1 (proportion of positive results for summary effect size results) shown in the y axis are the residuals of the nonlinear least squares model.

The proportion of positive results recorded for the effects of moderators (ds2) was not associated with any of the response variables we proposed. The occurrence of positive results bias regarding individual effect size results (ds3), on the other hand, was predicted by the shares of doctorate degrees per country (Table 2. 3, Figure 2. 3). However, it explained only a small part of the variation.

Table 2. 3. The proportion of positive results of moderators (ds2) and of individual effect sizes (ds3) modelled separately in response to publication year, country, number of authors and shares of doctorate degrees per country as the explanatory variables.

		df	**SQ**	**MSQ**	**F**	***p***
DS2	Location	2	0.17	0.08	0.75	0.48
	Year	1	0.01	0.01	0.12	0.74
	Number of authors	1	0.02	0.01	0.13	0.72
	PhD graduates	1	0.15	0.15	1.37	0.25
	Residuals	42	4.67	0.11		
DS3	Location	3	0.41	0.14	1.89	0.14
	Year	1	0.00	0.00	0.05	0.82
	Number of authors	1	0.27	0.27	3.76	0.056
	PhD graduates	1	0.45	0.45	6.13	**0.02**
	Residuals	73	5.31	0.07		

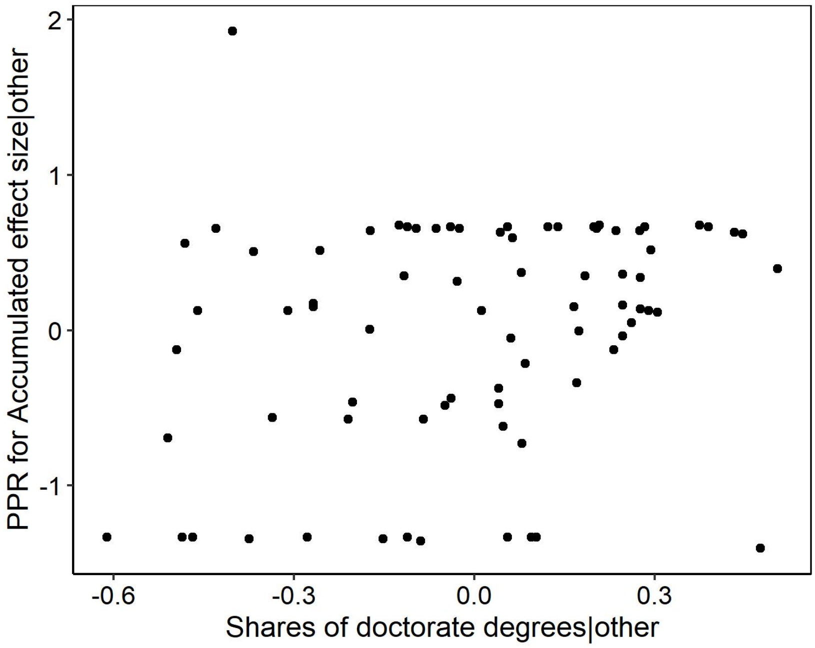


Figure 2. 3. Partial regression plot showing the influence of the shares of 25-64-year-olds with a doctorate degree per country on the proportion of positive results (PPR) for individual effect size results.

Discussion

Positive results were more frequent than negative results in the meta-analyses evaluated. In general, this result is in line with findings from studies focused on other disciplines (Csada et al., 1996; Jennions & Møller, 2002; Knight, 2003; Ioannidis, 2008; Fanelli, 2010b, 2012; Fanelli et al., 2017; Fraser et al., 2018) and may be a cause of concern (Ioannidis & Trikalinos, 2007a; Fanelli, 2012; Fanelli et al., 2017). A high incidence of positive results in meta-analyses illustrates how a high proportion of positive results in primary studies may be propagated to different research levels. Efforts are made in a meta-analysis to include data of different natures (e.g., grey literature) (Ioannidis, 2010). However, there is much room to increase the efforts to include grey literature data.

Excessive positive results in primary studies may indicate a difficulty in publishing negative results throughout the stages involved in publishing a paper (from the decision to submit to the acceptance for publication; Dickersin & Min, 1993; Ioannidis, 1998; Møller & Jennions, 2001; Knight, 2003), negatively affecting scientific progress (Alatalo et al., 1997; Browman, 1999; Fanelli, 2012). Concerns about the low occurrence of negative results have often been raised in the literature (Gould, 1993; Fanelli, 2010b, 2012; Graqvist, 2015). However, studies have indicated that negative results are in fact increasing (de Winter & Dodou, 2015) and only *seemingly* disappearing from papers (Fanelli, 2018). Negative results would be "increasingly embedded in longer publications that contain multiple results" and would "remain accessible to any researcher interested in finding them" (Fanelli, 2018). I included multiple results per study, in multiple levels, along with the negative results therein. Still, positive results prevailed in all three levels evaluated. In turn, the increased incidence of positive results may reflect more mature and well-developed theories (Fanelli, 2018).

A high prevalence of positive results in meta-analyses may reflect biases in two different levels, specifically, in primary studies and within the meta-analyses themselves. For instance, the file drawer problem (Rosenthal, 1979), selective reporting, hypothesizing after the results are known (Fraser et al., 2018), P-hacking (e.g., by changing statistical methods and choosing how much data to include or remove based on significance; Head et al., 2015), and selective analysis (using analyses that provide positive results) are all biases related to 'seeking positive results' that can occur in primary studies. A file drawer problem may occur if a paper, in this case with negative results, is never submitted, or is submitted but rejected during the review process (Csada et al., 1996; Pautasso, 2010; Fraser et al., 2018). Prestigious journals tend to publish results that are significant (Csada et al., 1996; Fraser et al., 2018), with high effect sizes, and that support a new (usually counter-intuitive) hypothesis, based on a new theoretical framework. This tendency of prestigious journals may discourage authors from submitting their negative findings. However, the existence of journals or journal sections dedicated to publishing negative results shows that the scientific community is aware of the bias against negative results, and consequently may reduce the reluctance of authors to write up and/or submit a paper with negative findings (Graqvist, 2015; Ekmekci, 2017). Also, steps have been taken towards a more transparent (Parker et al., 2018) and less P-value oriented research (Nakagawa, 2004). Still, other issues associated with publishing negative findings, such as time and interest, go beyond the presence of journals dedicated to negative results (DeVito & Goldacre, 2019) and along with other factors (see Fraser et al., 2018) may still favor the file drawer problem.

Selective reporting was recorded by researchers in psychology who showed that 70% of the results in a study were not reported (Franco et al., 2016) based on statistical significance (Fraser et al., 2018; Van Aert et al., 2019). High rates (> 50%) of cherry picking and hypothesizing after the results have also been reported among ecologist and evolutionary biologists (Fraser et al., 2018). Some go as far as to suggest that conducting questionable research practices (QRP) is a norm among scientists (John et al., 2012).

Meta-analyses are also affected by QRP, however differently. Meta-analyses have been criticized based on the "garbage in garbage out" argument. The structure of meta-analyses itself calls for an objective methodological assessment that incorporates different results, regardless of the direction (Ioannidis & Lau, 1998). A prevailing incidence of positive results in meta-analyses of the same field, indicates an issue probably with biases in primary studies. For instance, no excess in positive findings is seen in a meta-analysis of individual-level data (prospective meta-analysis), where the study is designed in a more controlled manner (Ioannidis & Trikalinos, 2007a). Similarly, results of studies with an elevated contribution of non-published results and analyzed in homogenous subsets, yielded a low incidence of positive results (Van Aert et al., 2019).

Truly a meta-analysis composed of mostly published (and positive) results will probably yield a positive result itself, but other factors beyond primary studies also play a role in determining the results, for instance, P-hacking in more heterogeneous scenarios (van Aert et al., 2016). For instance, the frequent updating of a meta-analysis may "increase the chances of making a false positive conclusion" (AlBalawi et al., 2013). In addition, cherry picking and/or HARKing based on the analyses of moderator results can also increase the proportion of significant results. The prevalence of positive results in summary effects may also indicate that our theories are going in the right direction (Fanelli, 2018). On the other hand, excessive positive results in moderators may indicate cherry picking or a 'fishing expedition'.

Meta-analyses do have several tools to manage publication bias (Ioannidis & Trikalinos, 2007a; Marks-Anglin & Chen, 2020). The meta-analytical method itself controls to some extent primary results derived from P-hacking (Head et al., 2015). However, we may be using such tools beyond their real capacity (Ioannidis & Trikalinos, 2005, 2007b; Lau et al., 2006; Ioannidis, 2008) because most publication bias tests are not intended to test significance-chasing biases (Ioannidis, 2010), and may also mistake heterogeneity for bias when the adequate design is not used (Ioannidis & Trikalinos, 2007a). Still, conducting publication bias tests is important, and unfortunately often forgotten (Aytug et al., 2012; Koricheva & Gurevitch, 2014; Lodi et al., 2021). We previously reported that at least 50% of the meta-analysis did not conduct publication bias analysis (Lodi et al., 2021). A posterior exploratory analysis showed that those that did evaluate publication bias, conducted multiple tests, specifically funnel plot, followed by rank correlation and fail-safe numbers (Orwin's, Rosenthal's, Rosemberg's, and Egger's; see Figure S2.1), as recommended (Coburn & Vevea, 2015). The use of multiple publication bias evaluation methods is recommended because "some methods do not perform well in some conditions and none of the publication bias methods outperform all the other methods under each and every condition" (Van Aert et al., 2019).

Statistical methods have been continuously developed to address publication bias (Marks-Anglin & Chen, 2020). Thus, one would expect a reduction in the prevalence of positive results in time (de Winter & Dodou, 2015). However, only positive results of summary effect sizes increased time and non-linearly. I speculate that non-linearity may be a consequence the increasing number meta-analyses in freshwater ecology.

Researchers have reported a lack of cultural differences on the incidence of positive results along with a slow increase in negative results in primary studies (de Winter & Dodou, 2015). Similarly, my findings indicate that cultural differences (i.e., country) were irrelevant for the incidence of positive results in individual effect sizes and moderator results. Still, I have found a higher frequency of positive results (considering summary effect sizes) in Europe, followed by North America. Not only "location", but locations more prone to becoming involved in scientific endeavors, as represented by shares of doctorate degree holders, would have reduced incidence of positive results. However, I observed quite the opposite for individual effect sizes with increased positive results in locations with higher shares of doctorate degree holders. However, the explanatory power was low, and these results should be observed with caution.

I had also expected that possible bias towards positive results would be reduced due to mutual control of team members (Fanelli et al., 2017), but the number of authors involved in the publication did not explain the proportion of positive results in any of the scales evaluated. The effect of collaboration in reducing bias has been reported as dependent on the geographical distance between collaborators such that the longer the distance the higher the probability of QRPs (Fanelli et al., 2017). However, we did not adjust for the distance factor. My results added to not having accounted for distance may have two implications. First, there is, in fact no relationship between the number of authors and bias towards positive results. Second, the effect of the distance between authors would buffer such relationship.

The ability to refute a very prestigious hypotheses may also be very attractive (for example, in molecular epidemiology high effects are published and quickly refuted; Ioannidis, 2010), and is known as the proteus phenomenon (Ioannidis & Trikalinos, 2005). A similar phenomenon has been reported for behavioral ecology (Chuard et al., 2019). However, this phenomenon is very discipline specific, and not likely to be occurring in freshwater ecology. Gould (1993) argues that the unwillingness of authors to publish negative findings is because positive results provide more interesting stories, while negative results "*often appear only in highly technical journals read by more restricted audiences and, as nonstories*". Following his reasoning, a commitment to reporting negative findings under a more interesting narrative, as more of a "good story", could correct publication bias to a large extent (Gould, 1993). Publication bias and the incidence of QRPs must continue to be studied empirically and using rigorous methods of investigation (Krosnick, 2019).

While reproducibility has been widely debated as a crisis (Baker & Penny, 2016) or a non-crisis (Fanelli, 2018; França & Monserrat, 2018), Fanelli (2018) argues that efforts to promote transparency and reproducibility would be better fostered under a narrative of transformation and empowerment. Following this reasoning, I add that more important than efforts to increase reproducibility would be to focus on encouraging researchers to turn their attention to writing up well-developed, well-founded manuscripts that answer *interesting/relevant* questions, regardless of the answer itself. Policies that encourage sharing data and data-analysis protocols are undoubtedly important for reasons that go beyond avoiding bias, for example, accelerating scientific advances by increasing exchange and efficiency (Woelfle et al., 2011; Fecher & Friesike, 2014). Still, I fail to see how science could truly advance without the continued increase in tolerance to unexpected/negative results. Especially as biological sciences are concerned, I believe 'unexpected' may be more common than not, and only by allowing the variability inherent to nature to surface, may ever more generalizable theories come forward.

Conclusion

Positive results were more frequent than negative results in the meta-analyses evaluated indicating that the high proportion of positive results in primary studies may be propagated to different research levels. I analyzed results in multiple levels per study; still, positive results prevailed in all levels evaluated. Such high prevalence of positive results in meta-analyses may reflect biases in primary studies as well as within the meta-analyses themselves. A meta-analysis composed of mostly published (and positive) results will probably yield a positive result, but other factors also play a role in determining the results. The prevalence of positive results in summary effects may also indicate that our theories are going in the right direction. Meta-analyses do have several tools to manage publication bias and the use of multiple publication bias evaluation methods is recommended. More important than efforts to increase transparency and systematization would be to focus on encouraging researchers to return their attention to writing up manuscripts focused on interesting/relevant questions, and not on the answer itself.

References

Al-Marzouki, S., S. Evans, T. Marshall, & I. Roberts, 2005. Are these data real? Statistical methods for the detection of data fabrication in clinical trials. BMJ 331: 267–270, http://www.bmj.com/lookup/doi/10.1136/bmj.331.7511.267.

Alatalo, R. V., J. Mappes, & M. A. Elgar, 1997. Heritabilities and paradigm shifts [11]. Nature. .

AlBalawi, Z., F. A. McAlister, K. Thorlund, M. Wong, & J. Wetterslev, 2013. Random error in cardiovascular meta-analyses: How common are false positive and false negative results?. International Journal of Cardiology 168: 1102–1107, https://linkinghub.elsevier.com/retrieve/pii/S016752731201529X.

Aria, M., & C. Cuccurullo, 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics 11: 959–975.

Aytug, Z. G., H. R. Rothstein, W. Zhou, & M. C. Kern, 2012. Revealed or concealed? transparency of procedures, decisions, and judgment calls in meta-analyses. Organizational Research Methods.

Bakan, D., 1966. The test of significance in psychological research. Psychological Bulletin .

Baker, M., & D. Penny, 2016. Is there a reproducibility crisis?. Nature. , 452–454.

Berlin, J. A., & R. M. Golub, 2014. Meta-analysis as Evidence. JAMA 312: 603, http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2014.8167.

Borenstein, M., L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein, 2009a. Introduction to Meta-Analysis. Introduction to Meta-Analysis. .

Borenstein, M., L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein, 2009b. How a Meta-Analysis Works Introduction to Meta-Analysis. .

Borenstein, M., L. V Hedges, J. P. T. Higgins, & H. R. Rothstein, 2009c. Criticisms of Meta-Analysis Introduction to Meta-Analysis. : 377–387, http://dx.doi.org/10.1002/9780470743386.ch43.

Browman, H. I., 1999. The uncertain position, status and impact of negative results in marine ecology: Philosophical and practical considerations. Marine Ecology Progress Series .

Chuard, P. J. C., M. Vrtílek, M. L. Head, & M. D. Jennions, 2019. Evidence that nonsignificant results are sometimes preferred: Reverse P-hacking or selective reporting?. PLOS Biology 17: e3000127, https://dx.plos.org/10.1371/journal.pbio.3000127.

Coburn, K. M., & J. L. Vevea, 2015. Publication bias as a function of study characteristics. Psychological Methods .

Cohen, J., 1994. The earth is round (p LT .05). American psychologist .

Cote, I. M., 2013. The Procedure of Meta-analysis in a Nutshell. Handbook of Meta-analysis in Ecology and Evolution .

Csada, R. D., P. C. James, & R. H. M. Espie, 1996. The “File Drawer Problem” of Non-Significant Results: Does It Apply to Biological Research? Oikos 76: 591, https://www.jstor.org/stable/3546355?origin=crossref.

de Winter, J. C., & D. Dodou, 2015. A surge of p -values between 0.041 and 0.049 in recent decades (but negative results are increasing rapidly too). PeerJ 3: e733, https://peerj.com/articles/733.

DeVito, N. J., & B. Goldacre, 2019. Catalogue of bias: publication bias. BMJ Evidence-Based Medicine 24: 53–54, https://ebm.bmj.com/lookup/doi/10.1136/bmjebm-2018-111107.

Dickersin, K., 1990. The Existence of Publication Bias and Risk Factors for Its Occurrence. JAMA: The Journal of the American Medical Association .

Dickersin, K., & Y.-I. Min, 1993. Publication Bias: The Problem That Won’t Go Away. Annals of the New York Academy of Sciences 703: 135–148, http://doi.wiley.com/10.1111/j.1749-6632.1993.tb26343.x.

Ekmekci, P. E., 2017. An increasing problem in publication ethics: Publication bias and editors' role in avoiding it. Medicine, Health Care and Philosophy 20: 171–178, http://link.springer.com/10.1007/s11019-017-9767-0.

Fanelli, D., 2010a. "Positive" results increase down the hierarchy of the sciences. PLoS ONE 5:.

Fanelli, D., 2010b. Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. PLoS ONE 5:, https://dx.plos.org/10.1371/journal.pone.0010271.

Fanelli, D., 2012. Negative results are disappearing from most disciplines and countries. Scientometrics 90: 891–904.

Fanelli, D., 2018. Opinion: Is science really facing a reproducibility crisis, and do we need it to?. Proceedings of the National Academy of Sciences 115: 2628–2631, http://www.pnas.org/lookup/doi/10.1073/pnas.1708272114.

Fanelli, D., R. Costas, & J. P. A. Ioannidis, 2017. Meta-assessment of bias in science. Proceedings of the National Academy of Sciences 114: 3714–3719, http://www.pnas.org/lookup/doi/10.1073/pnas.1618569114.

Fanelli, D., R. Costas, & V. Larivière, 2015. Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. PLOS ONE 10: e0127556, https://dx.plos.org/10.1371/journal.pone.0127556.

Fecher, B., & S. Friesike, 2014. Open Science: One Term, Five Schools of Thought In Bartling, S., & S. Friesike (eds), Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing. Springer International Publishing, Cham: 17–47, https://doi.org/10.1007/978-3-319-00026-8_2.

Forstmeier, W., E.-J. Wagenmakers, & T. H. Parker, 2017. Detecting and avoiding likely false-positive findings - a practical guide. Biological Reviews 92: 1941–1968, http://doi.wiley.com/10.1111/brv.12315.

França, T. F., & J. M. Monserrat, 2018. Reproducibility crisis in science or unrealistic expectations?. EMBO reports 19:, https://onlinelibrary.wiley.com/doi/abs/10.15252/embr.201846008.

Franco, A., N. Malhotra, & G. Simonovits, 2016. Underreporting in Psychology Experiments: Evidence From a Study Registry. Social Psychological and Personality Science .

Fraser, H., T. Parker, S. Nakagawa, A. Barnett, & F. Fidler, 2018. Questionable research practices in ecology and evolution. PLoS ONE .

Gardner, W., C. W. Lidz, & K. C. Hartwig, 2005. Authors' reports about research integrity problems in clinical trials. Contemporary Clinical Trials 26: 244–251, https://linkinghub.elsevier.com/retrieve/pii/S155171440500008X.

Gould, S. J., 1993. Cordelia's dilemma: silence, though usually undervalued, can be golden. Natural History 2: 10–18.

Graqvist, E., 2015. Why science needs to publish negative results. Innovation in Publishing .

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

Head, M. L., L. Holman, R. Lanfear, A. T. Kahn, & M. D. Jennions, 2015. The Extent and Consequences of P-Hacking in Science. PLOS Biology 13: e1002106, https://dx.plos.org/10.1371/journal.pbio.1002106.

Hopewell, S., M. Clarke, & S. Mallett, 2006. Grey Literature and Systematic Reviews Publication Bias in Meta-Analysis. John Wiley & Sons, Ltd, Chichester, UK: 49–72, http://doi.wiley.com/10.1002/0470870168.ch4.

Ioannidis, J. P. A., 1998. Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials. JAMA 279: 281, http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.279.4.281.

Ioannidis, J. P. A., 2005. Why most published research findings are false. PLoS Medicine. , 0696–0701.

Ioannidis, J. P. A., 2008. Interpretation of tests of heterogeneity and bias in meta-analysis. Journal of Evaluation in Clinical Practice 14: 951–957, http://doi.wiley.com/10.1111/j.1365-2753.2008.00986.x.

Ioannidis, J. P. A., 2010. Meta-research: The art of getting it wrong. Research Synthesis Methods 1: 169–184, http://doi.wiley.com/10.1002/jrsm.19.

Ioannidis, J. P. A., & T. A. Trikalinos, 2005. Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. Journal of Clinical Epidemiology .

Ioannidis, J. P. A., & T. A. Trikalinos, 2007a. An exploratory test for an excess of significant findings. Clinical Trials 4: 245–253, http://journals.sagepub.com/doi/10.1177/1740774507079441.

Ioannidis, J. P. A., & T. A. Trikalinos, 2007b. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne 176: 1091–1096, http://www.ncbi.nlm.nih.gov/pubmed/17420491.

Ioannidis, J. P., & J. Lau, 1998. Can quality of clinical trials and meta-analyses be quantified?. The Lancet 352: 590, https://linkinghub.elsevier.com/retrieve/pii/S0140673698220344.

Jennions, M. D., & A. P. Møller, 2002. Publication bias in ecology and evolution: An empirical assessment using the "trim and fill" method. Biological Reviews of the Cambridge Philosophical Society. .

John, L. K., G. Loewenstein, & D. Prelec, 2012. Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. Psychological Science .

Knight, J., 2003. Negative results: Null and void. Nature .

Koricheva, J., & J. Gurevitch, 2013. Place of Meta-Analysis among Other Methods of Research Synthesis. Handbook of Meta-analysis in Ecology and Evolution 498.

Koricheva, J., & J. Gurevitch, 2014. Uses and misuses of meta-analysis in plant ecology. Journal of Ecology 102: 828–844, http://doi.wiley.com/10.1111/1365-2745.12224.

Krosnick, J. A., 2019. Chapter 6. Comments on the Present and Future of Science, Inspired by Naomi Oreskes Why Trust Science?. Princeton University Press: 202–212, https://www.degruyter.com/document/doi/10.1515/9780691189932-009/html.

Kuhn, T. S., 1996. The Structure of Scientific Revolutions. The Structure of Scientific Revolutions. University of Chicago Press, http://www.bibliovault.org/BV.landing.epl?ISBN=9780226458083.

Lakatos, I., 2014. Falsification and the methodology of scientific research programmes Philosophy, Science, and History: A Guide and Reader. .

Lau, J., J. P. A. Ioannidis, N. Terrin, C. H. Schmid, & I. Olkin, 2006. The case of the misleading funnel plot. British Medical Journal. .

Lawrence, P. a, 2003. The politics of publication - Authors, reviewers and editors must act to protect the quality of research. Nature .

Lodi, S., B. S. Godoy, J. C. gonçalves Ortega, & L. M. Bini, 2021. Quality of meta-analyses in freshwater ecology: A systematic review. Freshwater Biology 66: 803–814, https://onlinelibrary.wiley.com/doi/10.1111/fwb.13695.

Marks-Anglin, A., & Y. Chen, 2020. A historical review of publication bias. Research Synthesis Methods 11: 725–742, https://onlinelibrary.wiley.com/doi/10.1002/jrsm.1452.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J. A. Berlin, J. Clark, M. Clarke, D. Cook, R. D'Amico, J. J. Deeks, P. J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P. C. Gøtzsche, J. Grimshaw, G. Guyatt, J. Higgins, J. P. A. Ioannidis, J. Kleijnen, T. Lang, N. Magrini, D. McNamee, L. Moja, C. Mulrow, M. Napoli, A. Oxman, B. Pham, D. Rennie, M. Sampson, K. F. Schulz, P. G. Shekelle, D. Tovey, & P. Tugwell, 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine. .

Møller, A. P., & M. D. Jennions, 2001. Testing and adjusting for publication bias. Trends in Ecology and Evolution. .

Nakagawa, S., 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. Behavioral Ecology 15: 1044–1045, https://academic.oup.com/beheco/article-lookup/doi/10.1093/beheco/arh107.

Oreskes, N., 2019. Chapter 1. Why Trust Science? Perspectives from the History and Philosophy of Science Why Trust Science?. Princeton University Press: 15–68, https://www.degruyter.com/document/doi/10.1515/9780691189932-003/html.

Parker, T. H., S. C. Griffith, J. L. Bronstein, F. Fidler, S. Foster, H. Fraser, W. Forstmeier, J. Gurevitch, J. Koricheva, R. Seppelt, M. W. Tingley, & S. Nakagawa, 2018. Empowering peer reviewers with a checklist to improve transparency. Nature Ecology and Evolution .

Parker, T. H., & S. Nakagawa, 2014. Mitigating the epidemic of type I error: ecology and evolution can learn from other disciplines. Frontiers in Ecology and Evolution .

Pautasso, M., 2010. Worsening file-drawer problem in the abstracts of natural, medical and social science databases. Scientometrics 85: 193–202, http://link.springer.com/10.1007/s11192-010-0233-5.

Popper, K., 2005. The Logic of Scientific Discovery. The Logic of Scientific Discovery. Routledge, https://www.taylorfrancis.com/books/9781134470020.

Ranstam, J., M. Buyse, S. L. George, S. Evans, N. L. Geller, B. Scherrer, E. Lesaffre, G. Murray, L. Edler, J. L. Hutton, T. Colton, & P. Lachenbruch, 2000. Fraud in Medical Research. Controlled Clinical Trials 21: 415–427, https://linkinghub.elsevier.com/retrieve/pii/S0197245600000696.

Rosenthal, R., 1979. The file drawer problem and tolerance for null results. Psychological Bulletin 86: 638–641, http://content.apa.org/journals/bul/86/3/638.

Rosenthal, R., & K. L. Fode, 1963. Psychology of the Scientist: V. Three Experiments in Experimenter Bias. Psychological Reports 12: 491–511, http://journals.sagepub.com/doi/10.2466/pr0.1963.12.2.491.

Simmons, J. P., L. D. Nelson, & U. Simonsohn, 2011. False-Positive Psychology. Psychological Science 22: 1359–1366, http://journals.sagepub.com/doi/10.1177/0956797611417632.

Stewart, G. B., I. M. Côté, H. R. Rothstein, & P. S. Curtis, 2013. First steps in beginning a meta-analysis Handbook of Meta-analysis in Ecology and Evolution. .

van Aert R. C. M., J. M. Wicherts, & M. A. L. M. van Assen, 2016. Conducting Meta-Analyses Based on p Values: Reservations and Recommendations for Applying p-Uniform and p-Curve. Perspectives on Psychological Science .

Van Aert, R. C. M., J. M. Wicherts, & M. A. L. M. Van Assen, 2019. Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. PLoS ONE .

Whittaker, R. J., 2010. Meta-analyses and mega-mistakes: Calling time on meta-analysis of the species richness-productivity relationship. Ecology 91: 2522–2533.

Woelfle, M., P. Olliaro, & M. H. Todd, 2011. Open science is a research accelerator. Nature Chemistry 3: 745–748, http://www.nature.com/articles/nchem.1149.

Yu, L., Y. Huang, F. Sun, & W. Sun, 2017. A synthesis of soil carbon and nitrogen recovery after wetland restoration and creation in the United States. Scientific Reports .

Supplementary Information

Table S2.1. Three datasets (ds1, ds2 and ds3) used in this paper. ds1 corresponds to the PPR of summary effect sizes, ds2 of effects of moderators, and ds3 of individual effect size results. The column PPR refers to ratio of positive to negative and non-significant results (positive/(negative + non-significant)). Neg = Negative; Pos = Positive, ns = Non-significant, Neg+ns = the total amount of negative results encountered. The column 'Reference number' indicates the reference evaluated, following (Lodi et al., 2021).

Reference Number	**ds1**					**ds2**					**ds3**				
	Neg	**Pos**	**ns**	**Neg+ns**	**PPR**	**Neg**	**Pos**	**ns**	**Neg+ns**	**PPR**	**Neg**	**Pos**	**ns**	**Neg+ns**	**PPR**
1	0	1	0	0	1	3	2	1	4	0.33	3	10	7	10	0.50
2	0	1	0	0	1	1	3	0	1	0.75	2	5	0	2	0.71
3	1	1	0	1	0.50	1	1	6	7	0.125	1	12	3	4	0.75
4	1	0	0	1	0	4	4	0	4	0.5	0	0	0	0	NA
5	0	1	0	0	1	0	1	0	0	1	0	7	0	0	1
6	1	1	0	1	0.50	0	0	0	0	NA	2	2	10	12	0.14
7	0	1	0	0	1	0	0	0	0	NA	0	0	0	0	NA
8	0	0	0	0	NA	0	3	0	0	1	1	5	1	2	0.71
9	0	1	2	2	0.33	0	0	0	0	NA	0	10	14	14	0.42
10	0	2	0	0	1	0	1	5	5	0.17	1	17	3	4	0.81
11	1	0	0	1	0	1	0	0	1	0	2	0	0	2	0
12	0	1	0	0	1	0	1	1	1	0.5	0	3	1	1	0.75
13	0	0	0	0	NA	0	1	1	1	0.5	0	3	5	5	0.38
14	0	0	0	0	NA	0	1	0	0	1	0	1	1	1	0.50
15	0	1	0	0	1	0	1	0	0	1	0	3	0	0	1
16	0	0	0	0	NA	2	4	0	2	0.67	1	7	10	11	0.39
17	0	0	0	0	NA	1	6	0	1	0.86	0	0	0	0	NA
18	0	0	0	0	NA	2	4	0	2	0.67	0	6	0	0	1
19	0	0	0	0	NA	0	1	0	0	1	0	5	0	0	1
20	0	1	0	0	1	1	3	0	1	0.75	5	11	1	6	0.65
21	0	3	0	0	1	0	0	0	0	NA	0	22	14	14	0.61
22	0	1	0	0	1	2	2	0	2	0.5	0	11	0	0	1
23	0	1	0	0	1	4	6	1	5	0.55	1	13	4	5	0.72
24	0	0	0	0	NA	0	8	1	1	0.89	0	22	0	0	1
25	0	0	0	0	NA	0	2	0	0	1	0	0	0	0	NA
26	0	0	0	0	NA	0	2	3	3	0.4	1	4	4	5	0.44
27	0	0	0	0	NA	0	0	0	0	NA	0	24	24	24	0.50
28	0	1	0	0	1	0	2	0	0	1	0	5	3	3	0.63
29	0	0	0	0	NA	0	3	2	2	0.6	0	11	4	4	0.73
30	0	0	2	2	0	4	1	0	4	0.2	4	6	2	6	0.5
31	0	0	0	0	NA	0	0	0	0	NA	0	6	0	0	1
32	0	0	0	0	NA	0	2	0	0	1	10	2	0	10	0.17
33	0	0	0	0	NA	0	0	0	0	NA	0	6	0	0	1
34	0	0	0	0	NA	0	0	0	0	NA	0	0	8	8	0
35	0	0	0	0	NA	0	4	1	1	0.8	3	18	1	4	0.82
36	0	0	0	0	NA	0	7	2	2	0.78	3	24	18	21	0.53
37	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA

38	0	0	0	0	NA	0	0	0	0	NA	0	2	1	1	0.67
39	0	0	0	0	NA	0	0	0	0	NA	0	2	3	3	0.40
40	0	0	0	0	NA	0	3	1	1	0.75	0	9	2	2	0.82
41	0	1	0	0	1	0	0	0	0	NA	0	0	0	0	NA
42	0	0	0	0	NA	0	0	0	0	NA	0	0	4	4	0
43	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
44	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
45	0	0	0	0	NA	0	0	0	0	NA	0	1	0	0	1
46	0	2	0	0	1	0	2	4	4	0.33	0	2	12	12	0.14
47	0	1	0	0	1	1	3	0	1	0.75	0	8	6	6	0.57
48	0	0	0	0	NA	0	0	0	0	NA	1	2	0	1	0.67
49	0	2	0	0	1	0	0	0	0	NA	0	0	0	0	NA
50	0	4	0	0	1	3	0	0	3	0	0	17	2	2	0.89
51	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
52	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
53	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
54	0	0	0	0	NA	0	0	0	0	NA	0	16	4	4	0.8
55	0	0	0	0	NA	0	7	1	1	0.88	0	0	0	0	NA
56	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
57	0	0	0	0	NA	0	0	0	0	NA	2	6	0	2	0.75
58	0	0	0	0	NA	1	2	0	1	0.67	0	0	0	0	NA
59	1	1	0	1	0.50	0	4	0	0	1	0	6	5	5	0.55
60	0	0	0	0	NA	0	0	0	0	NA	0	2	0	0	1
61	0	0	0	0	NA	0	3	11	11	0.21	0	1	5	5	0.17
62	0	0	0	0	NA	0	0	0	0	NA	0	1	1	1	0.50
63	0	0	0	0	NA	0	0	0	0	NA	1	1	2	3	0.25
64	0	0	0	0	NA	0	0	0	0	NA	1	2	0	1	0.67
65	0	0	0	0	NA	0	0	0	0	NA	1	4	7	8	0.33
66	0	0	0	0	NA	0	6	0	0	1	1	12	0	1	0.92
67	0	0	0	0	NA	0	0	0	0	NA	1	1	0	1	0.50
68	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
69	0	0	0	0	NA	1	1	0	1	0.5	9	37	29	38	0.49
70	0	0	0	0	NA	0	3	0	0	1	0	12	1	1	0.92
71	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
72	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
73	0	0	0	0	NA	0	0	0	0	NA	1	12	4	5	0.71
74	0	0	0	0	NA	0	0	0	0	NA	7	3	14	21	0.13
75	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
76	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
77	0	1	0	0	1	0	3	3	3	0.5	3	12	0	3	0.8
78	0	0	0	0	NA	0	0	0	0	NA	0	3	0	0	1
79	0	0	0	0	NA	0	0	0	0	NA	0	2	0	0	1
80	0	0	0	0	NA	1	1	3	4	0.2	0	23	14	14	0.62
81	0	0	0	0	NA	0	0	0	0	NA	0	9	13	13	0.41
82	0	0	0	0	NA	0	3	2	2	0.6	1	1	0	1	0.5
83	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
84	0	0	0	0	NA	0	0	0	0	NA	7	20	7	14	0.59
85	2	7	1	3	0.70	0	0	0	0	NA	0	0	0	0	NA

86	0	1	1	1	0.50	0	1	0	0	1	2	19	0	2	0.90
87	0	0	0	0	NA	0	0	0	0	NA	0	11	2	2	0.85
88	0	0	0	0	NA	0	1	3	3	0.25	0	12	2	2	0.86
89	0	0	0	0	NA	0	0	0	0	NA	0	3	0	0	1
90	0	0	0	0	NA	0	0	1	1	0	0	4	9	9	0.31
91	0	0	0	0	NA	0	0	0	0	NA	0	1	2	2	0.33
92	0	0	1	1	0	0	0	0	0	NA	0	3	8	8	0.27
93	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
94	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
95	0	0	0	0	NA	0	0	0	0	NA	1	3	2	3	0.5
96	0	0	0	0	NA	0	0	0	0	NA	0	1	0	0	1
97	0	0	0	0	NA	0	0	0	0	NA	0	2	1	1	0.67
98	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
99	0	0	0	0	NA	0	2	0	0	1	0	5	1	1	0.83
100	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
101	0	1	0	0	1	0	1	0	0	1	0	12	11	11	0.52
102	0	1	0	0	1	0	2	0	0	1	0	3	0	0	1
103	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
104	0	1	0	0	1	0	0	0	0	NA	0	3	0	0	1
105	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
106	0	1	0	0	1	0	0	0	0	NA	0	0	0	0	NA
107	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
108	0	0	0	0	NA	0	0	0	0	NA	1	14	5	6	0.7
109	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
110	0	2	1	1	0.67	0	0	0	0	NA	0	16	18	18	0.47
111	0	0	0	0	NA	0	0	0	0	NA	0	2	7	7	0.22
112	0	0	0	0	NA	0	0	0	0	NA	0	0	0	0	NA
113	0	0	0	0	NA	0	0	0	0	NA	0	4	3	3	0.57

Table S2.2. Ratios obtained considering the direction of the hypothesis partitioned among three explanatory variables, namely number of authors, year of publication and location. Effects of each category of location are shown.

Source of variation	Estimate	Std.Error	t	*P*
Intercept	-49.165	26.162	-1.879	0.066
Number of authors	-0.011	0.031	-0.355	0.724
Year	0.025	0.013	1.917	0.061
Location - CN	0.221	0.485	0.456	0.650
Location - EU-15	0.368	0.480	0.768	0.446
Location - NZ-AM	0.250	0.575	0.435	0.666
Location - US	0.436	0.481	0.906	0.369

Table S2.3. Ratios obtained when I considered significant values as positive results (inclusive ratio) partitioned among three explanatory variables, namely number of authors, year of publication and location. Effects of each category of location are shown

Source of variation	Estimate	Std.Error	t	*P*
Intercept	-52.551	19.004	-2.765	0.007
Number of authors	-0.017	0.013	-1.259	0.211
Year	0.027	0.009	2.816	0.006
Location - CN	0.235	0.267	0.882	0.380
Location - EU-15	0.276	0.256	1.079	0.284
Location - NZ-AM	0.025	0.327	0.075	0.940
Location - US	0.432	0.260	1.662	0.100

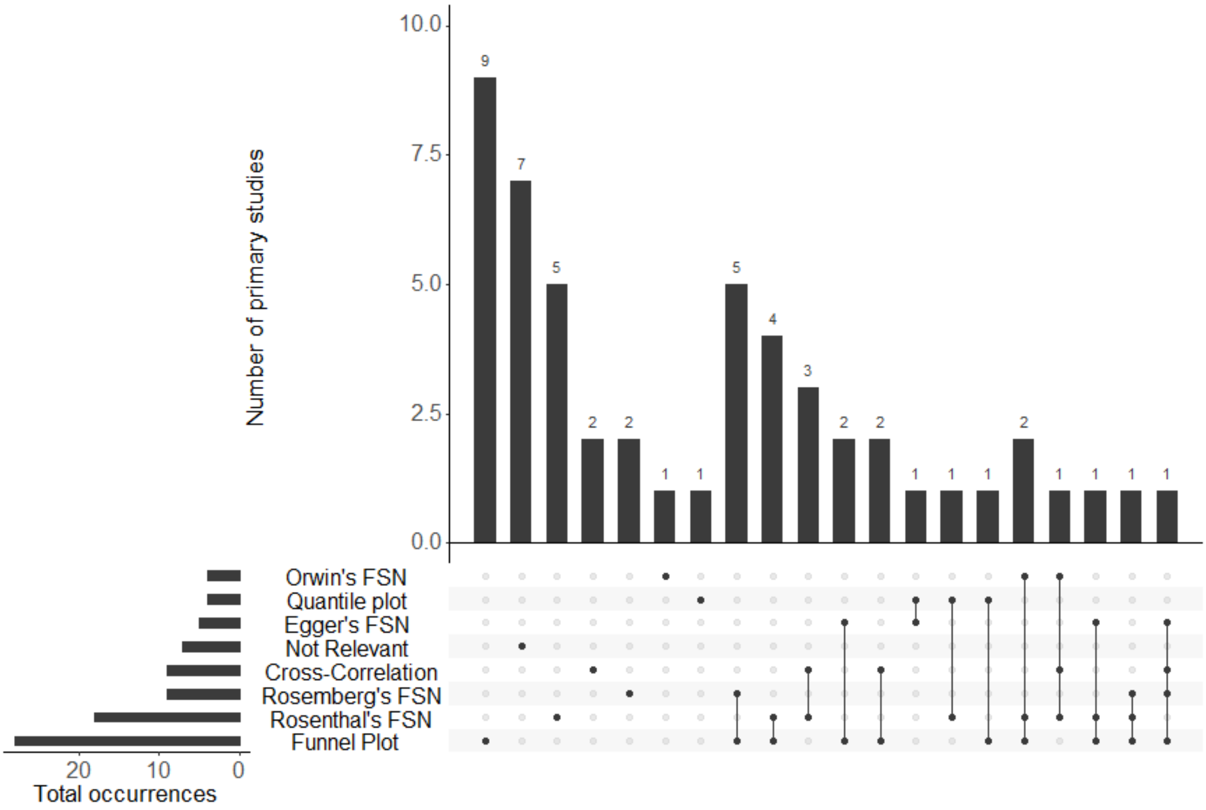


Figure S2.1 Strategies used to check for de occurrence of publication bias. The histogram with vertical bars shows the number of studies that used each combination of assessment strategy shown below the graph. The dots indicate the usage of a strategy, and connected dots indicate the multiple strategies used. The histogram with horizontal bars (to the left) shows the occurrence of each combination used in the studies evaluated. Note that 62 papers didn't use any strategy to check for publication bias, despite its relevance, and were therefore not included in this graph. FSN = Fail Safe Number

Chapter III

Relative importance of land use and local scale predictors of freshwater communities

Resumo

Ecólogos têm utilizado variáveis ambientais obtidas em diferentes escalas para tentar predizer a estrutura das comunidades aquáticas. Para alcançar este objetivo, as variáveis de uso do solo podem complementar as variáveis obtidas em escala local. Neste estudo, utilizei resultados de análises canônicas para avaliar a importância relativa das variáveis locais e de uso do solo na estruturação de comunidades aquáticas, utilizando uma abordagem meta-analítica. Os resultados não apoiaram a noção de que as variáveis de uso do solo são melhores para explicar a estrutura das comunidades do que as variáveis locais. Verifiquei que os dois tipos de variáveis são importantes no estudo de correlatos da estrutura de comunidades aquáticas. Para um tipo de tamanho de efeito que analisei (correlação máxima com o primeiro eixo de ordenação), a importância relativa dos tipos de variáveis dependeu da diferença entre o número de variáveis representando cada grupo. Aproximadamente metade das variáveis incluídas nos estudos foram utilizadas apenas uma vez, indicando uma falta de consenso sobre variáveis devem ser utilizadas. Verifiquei a necessidade de melhorar a apresentação dos resultados em estudos que utilizam métodos multivariados para facilitar futuros estudos de meta-análise.

Abstract

Ecologists have become interested in integrating terrestrial and limnological knowledge to better understand the structure of aquatic communities. Land use variables may complement local scale variables in explaining the effects of the environment on aquatic communities. In this study, I used results from canonical analyses to evaluate the relative importance of local and land use variables in determining the composition of freshwater communities using a meta-analytical approach. The results did not support the notion that land use variables are better at explaining community structure than local variables. I found that both variable types are important when studying correlates of aquatic community structure. For one type of effect size (maximum correlation with the first ordination axis), the relative importance of the types of variables was dependent on the difference between the number of variables used to represent each group. Approximately half of the variables included in the studies were used only once, indicating a lack of consensus on which variables to use among studies. I found the need to improve the reporting of multivariate studies to facilitate future meta-analysis studies.

Introduction

Limnology has traditionally considered numerous environmental variables to describe aquatic ecosystems in different spatial and temporal scales (see Jenkins & Buikema, 1998; Beisner et al., 2006). However, despite the large number of variables that are often used, multivariate models aiming to unveil the processes that shape the communities of both lotic and lentic environments have, in general, very low explanatory powers (e.g., Soininen, 2014; 2016). An explanation for this imbalance between describing and understanding community patterns may occur due to the historical assumption that aquatic environments are closed systems, with little influence form the surrounding environment (Forbes, 1925; Hansson *et al.*, 2013).

More recently ecologists became interested in integrating terrestrial and limnological knowledge to better understand the structure of aquatic communities (Wiley, Kohler & Seelbach, 1997; Allan & Castillo, 2007; Nessimian *et al.*, 2008) following the principle that aquatic and terrestrial ecosystems would be linked through the movement of animals, energy and matter (Soininen *et al.*, 2015). There would be four dimensions regulating stream dynamics: lateral (exchange between terrestrial and aquatic environments), longitudinal (movement of aquatic organisms and material downstream), vertical (movement within the water column and the hyporheic zone) and temporal dimensions (Ward, 1989; Mello *et al.*, 2020). Lake dynamics is also affected by anthropogenic and catchment scale factors (Liu *et al.*, 2012). Land use variables would add temporal information that the "in-water body" variables would disregard (Carneiro *et al.*, 2014). Local variables would be a snapshot representation of the environmental conditions whilst land use variables would show the events that the water body was, across time, subject to. Considering this framework, land use variables would complement (Mikulyuk *et al.*, 2011; Carneiro *et al.*, 2014; Soininen *et al.*, 2015; Mello *et al.*, 2020) or even be a better predictor of community structure than local variables (Miserendino *et al.*, 2011). In addition to the theoretical considerations, variables obtained from the environment surrounding water bodies (e.g., land use), often obtained by geoprocessing, enable large-scale analyses (see Wiley *et al.*, 1997 for an example). Ecologists do agree that large-scale processes are important (Ricklefs, 1987; Levin, 1992; Root & Schneider, 1995; Wiley *et al.*, 1997). Effects of biotic interactions and more local factors driving species distributions are averaged out at large scales becoming Eltonian noise (Eltonian factors usually associated with interactions may not affect distributions at large extents and low resolutions; Soberon & Nakamura, 2009). In addition, at least one checklist with standardized variables associated with land-use is available and widely used (e.g., Nessimian *et al.*, 2008).

Despite many calls for the inclusion of land use variables as potential predictors of aquatic community structure, evidences show contrasting results on how these variables affect different biological groups (Vannote *et al.*, 1980; Allan, Erickson & Fay, 1997; Sponseller, Benfield & Valett, 2001; Hoffmann & Dodson, 2005; Carneiro *et al.*, 2014; Rocha *et al.*, 2020). For instance, land use has been shown to structure macroinvertebrate communities (Sponseller *et al.*, 2001; Fierro *et al.*, 2017). On the other hand, Ometo *et al.* (2000) are of the opinion that these communities are mainly structured by local factors. Zooplankton community has been reported to be more affected by lake development (indicating land-use intensity; Hoffmann & Dodson, 2005), by land-use (Sługocki *et al.*, 2018) and by spatially structured environmental factors (Rocha *et al.*, 2020). Studies report that the fish community diversity is affected by land use (Stanfield & Kilgour, 2013; Dala-Corte *et al.*, 2016) depending on the scale (Allan *et al.*, 1997). Also, the influence of land use may depend on the dispersal ability of the taxa being studied (Firmiano *et al.*, 2021).

There is also considerable redundancy in the variables usually considered, leaving room to improve parsimony in data gathering and, consequently, increase sample sizes with the resources saved. The number/redundancy of the variables used to study a relationship between community structure and environmental factors directly relates to the reliability of the results obtained. When the number of explanatory variables increases, so does the odds of multicollinearity (especially if variables are describing similar/linked phenomena). Multicollinearity inflates the fit of models and, according to Graham (2003), "multicollinear explanatory variables are difficult to analyze because their effects on the response can be due to either true synergistic relationships among the variables or spurious correlations", yielding possibly confusing results.

The degree of relationship between the environment and biological communities is typically examined using direct gradient analyses or canonical analyses (Canonical Correspondence Analysis or Redundancy Analysis) (Cottenie, 2005; Soininen, 2014, 2016). In this study I use results from canonical analyses to evaluate the relative importance of local and land use variables in determining the composition of plankton, macroinvertebrate, and fish communities using a meta-analytical approach. I aim to evaluate the hypotheses that land use variables will be important determinants of aquatic communities. Different components of the benthic macroinvertebrates and fish communities have been reported sensitive to land-use changes (see Harding *et al.*, 1998; Allen *et al.*, 1999; Cortes *et al.*, 2011; Cunha & Juen, 2017), while the effects of land use on planktonic communities tend to be smaller (Rocha *et al.*, 2020). The relative importance of land use variables will vary depending on the scale studied, reducing at a local scale, and increasing as we focus on larger scales.

Methods

Search strategy and criteria

I selected studies from the Web of Science and Scopus platforms. The terms "(assemblage* OR communit*) AND (fish* OR macroinvertebrate* OR *plankton) AND ("land us*" OR disturbance* OR "soil us*" OR "land cov*" OR regional)" (Moher *et al.*, 2009b, 2015) were searched in the title of the articles and filtered by Web of Science categories excluding irrelevant categories (e.g., psychology experimental, hematology, psychology applied, food science technology, immunology, or infectious diseases). I included articles conducted in freshwater environments, that studied the simultaneous effects of land-use and of local scale variables on aquatic community structure (specifically for zooplankton, phytoplankton, fish, and macroinvertebrate data; see results). Also, I only included studies that used canonical analyses and reported the correlation between the environmental variables (local or land use) and the ordination axes (following Siefert *et al.*, 2012). These results were retrieved from the ordination diagrams (see below). Studies that did not meet the inclusion criteria were excluded. Specifically, our exclusion criteria were not being i) ecology, ii) freshwater, iii) land use or vi) fish, zooplankton, phytoplankton, or macroinvertebrates, v) a primary study. I also excluded studies that vi) studied phytotelmas, vii) did not use canonical methods, viii) studied functional or beta diversity, ix) conducted a partial canonical analysis (because partial correlations depend also on correlations among predictor variables and not only on the relationship between the predictor and the response variable following Siefert *et al.*, 2012), x) used only land use variables, xi) was in Chinese or French, xii) was unavailable for download and/or not sent by authors, and that xiii) reported only local environmental variables. I collected and screened the articles following the PRISMA protocol. Study selection is summarized in the PRISMA flow diagram (Figure 3.1).

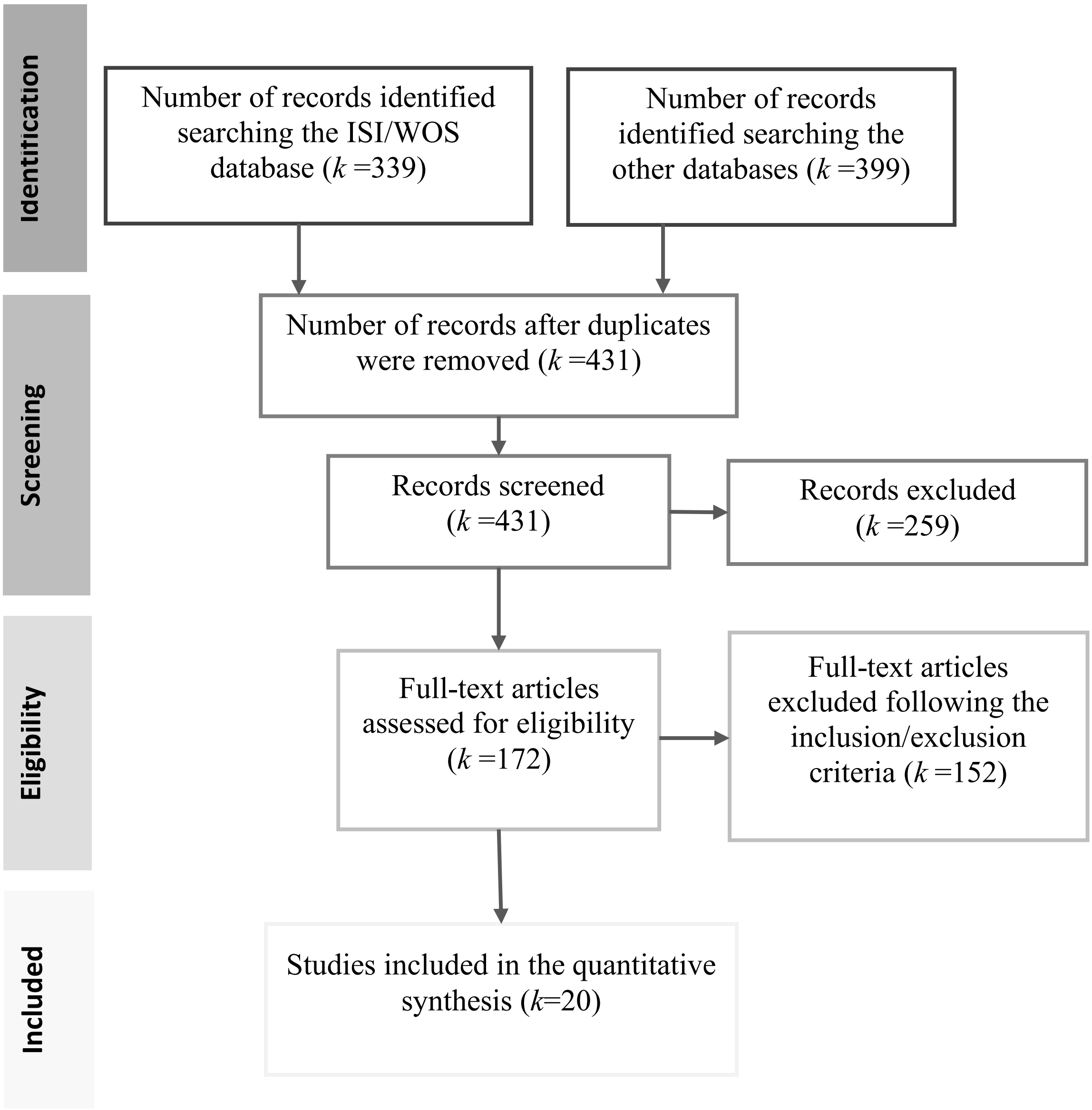


Figure 3.1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram summarizing study inclusion and exclusion phases (adapted from Moher *et al.*, 2009a)

Data and Statistical analysis

Considering an ordination plot from a study, we first retrieved the correlations (i.e., loadings) between the first axis and a given type of environmental variable (local and land use). Then, we averaged the absolute values of the correlations for each type of variable. After repeating this procedure with our set of 20 studies (totaling 24 ordination plots), we created two vectors: one with the average correlations between the local variables and the first axis ($r_{Ax1,local}$) and other with the average correlations between the land use variables and the first axis ($r_{Ax1,land\ use}$). These procedures were repeated for the second axis ($r_{Ax2,local}$ and $r_{Ax2,land\ use}$) and for the maximum correlation values for each ordination axis (instead of averages; generating $rmax_{Ax1,local}$, $rmax_{Ax2,local}$, $rmax_{Ax1,land\ use}$ and $rmax_{Ax2,land\ use}$). Figure 3.2 summarizes these procedures.

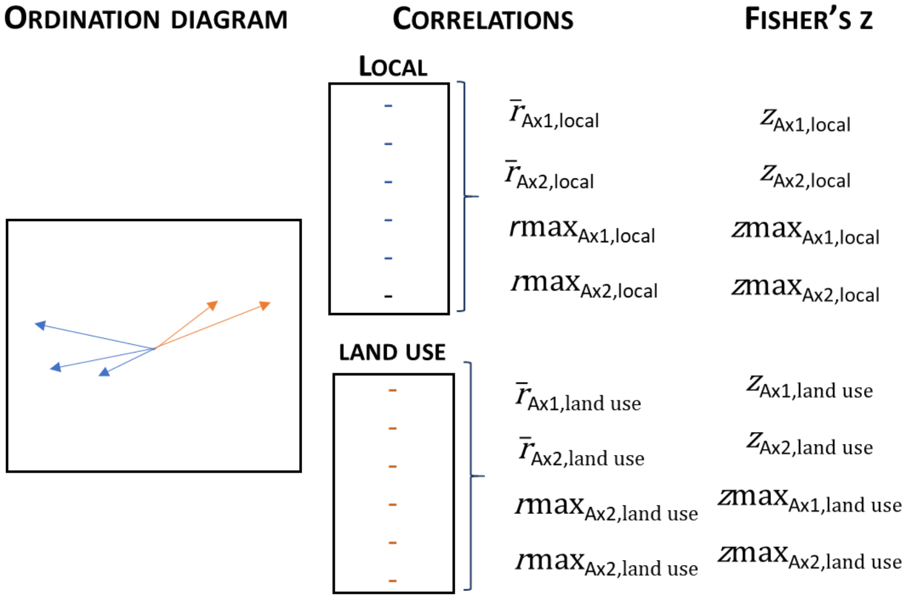


Figure 3.2. Schematic representation of the analyses conducted in this study. max=maximum value, Ax= Axis of an ordination diagram.

Each type of correlation (mean of the correlations between an axis and the local and land use variables and their respective maximum values) was transformed into Fisher's *z* (see equation 6.2 from Borenstein *et al.*, 2009). We then estimated a random-effects weighted effect size (following procedures described in Borenstein *et al.*, 2009) for each type of environmental variable (local and land use; $z\text{-}local_{++}$ and $z\text{-}land\ use_{++}$, respectively) and considering the different types of correlation. We used the overlap between the confidence intervals to ascertain whether the weighted effect sizes differed between local and land use variables. All estimates were made using the metafor package (Viechtbauer, 2010) for R (R Development Core Team, 2020).

The comparison between the two correlation coefficients (e.g., between $r_{Ax1,local}$ and $r_{Ax1,land\ use}$), in our study, should consider that the correlations are overlapping. In other words, the two variables being compared (i.e., local and land use) have at least one variable in common, an aquatic community (Diedenhofen & Musch, 2015). Thus, we used the cocor package (Diedenhofen & Musch, 2015) to compare the strengths of the relationships between each type of environmental variable (local and land use) and an aquatic community for each study (for the different statistical tests available in this package, see Table 2 in Diedenhofen & Musch, 2015). We assumed that the correlation between local and land use variables was zero (a conservative criterion); however, other values (e.g., 0.25, 0.50 and 0.75) produced similar results (not shown).

Moderators

For each study, I also collected the following moderators (along with the effect sizes described above): biological group (zooplankton, phytoplankton, fish, or benthic macroinvertebrates), type of environment studied (lotic or lentic), number of variables included for each type of environmental variable (local and land use) and study spatial extent (defined in this study as the largest distance between two sites in a straight line). The effects of moderators were evaluated using the *Z*-tests.

Results

Most of the articles were excluded for not studying land use variables (125 out of 411 excluded), for not having analyzed the data using a canonical multivariate analysis (119 out of 411), and for not being conducted in freshwater ecosystems (116 out of 411; Figure 3.3). After applying these criteria, I included 24 analyses from 20 articles.

The number of local and of land use variables used in the canonical analyses varied greatly among studies (from 2 to 18 and from 1 to 14, respectively). However, the number of variables considered in the primary studies were larger than those used in the canonical analyses (from 6 to 33 and 1 to 20, respectively).

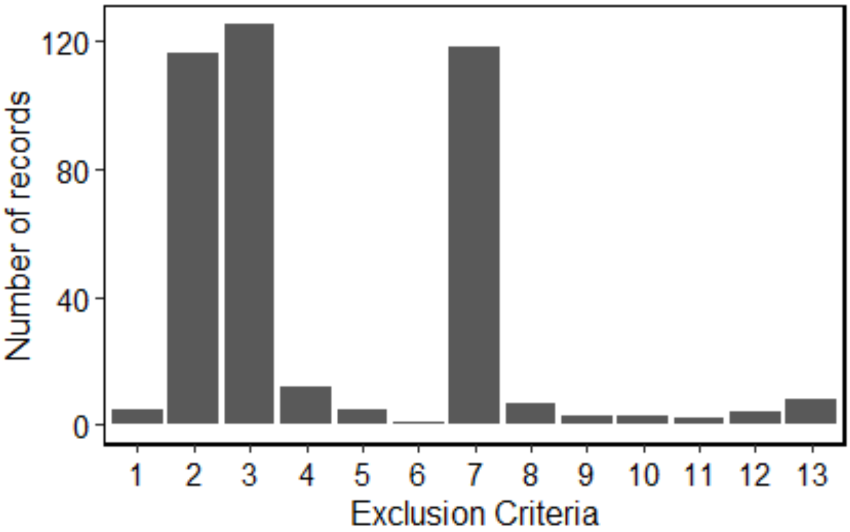


Figure 3.3. Number of records excluded per exclusion criterion. 1- Not ecology; 2- Not freshwater; 3- Not land use 4 Not fish, zooplankton, phytoplankton, or macroinvertebrates; 5- not a primary study; 6- Phytotelmas; 7- Did not use canonical methods; 8- used functional or beta diversity; 9- conducted a partial canonical analysis, 10- reported only land use variables; 11 – was in Chinese or French; 12 – data not available; 13 – reported only local environmental variables.

A total of 64 local scale variables and 65 land use variables were reported by the primary studies. Many of the variables were reported in only one or two studies, especially concerning land use variables (see Figure S3.1). The most used local-scale variables were dissolved oxygen (9 studies), alkalinity (8 studies), width of the water body (7 studies), water flow (7 studies), depth (6 studies), transparency (5 studies), salinity (4 studies), total suspended solids (4 studies), total dissolved solids (4 studies), Nitrogen/Phosphorous ratio (4 studies), and total organic carbon (4 studies). The variables permanganate index, total phosphorous, conductivity, total nitrogen, NO_3, NH_3, NH_4, and pH were used by 3 studies. The most used land use-scale variables were forest (6 studies), crop (4 studies), grass (4 studies), urban (4 studies), grazing (3 studies), tree removal (3 studies) and catchment area (3 studies).

For the first ordination axis, the weighted effect size of the relationship between the local environmental variables and the aquatic communities ($z\text{-}local_{++} = 0.316$; 95% CI = 0.304-0.418) did not differ substantially from that estimated for the land use environmental variables ($z\text{-}land\ use_{++} = 0.445$; 95% CI = 0.302-0.587). However, the heterogeneity among the effect sizes estimated with the local environmental variables ($Q = 20.74$, $k = 23$, $P = 0.5971$) was clearly lower than the heterogeneity among the effect sizes estimated with the land use variables ($Q = 121.57$, $k = 23$, $P < 0.0001$). In general, the results of the comparisons between the local and land use variables, using the tests summarized by Diedenhofen & Musch (2015), also indicated that these groups of variables were similarly correlated with the aquatic communities. Indeed, we found that in only three studies the land use variables were more strongly correlated with the aquatic communities than the local variables (Figure 3.4a). We also did not find evidence that the differences between the types of variables (local and land use) were dependent on the group of organisms under study (Figure S3.2a), type of environment (lotic x lotic, not shown), on the difference in the number of variables of each type ($r = 0.177$; $P = 0.4083$) and on the spatial extent of the study ($r = -0.168$; $P = 0.468$).

The results for the second axis also indicated that the weighted effect sizes of the relationships between the environmental variables and the aquatic communities were similar ($z\text{-}local_{++} = 0.305$; 95% CI = 0.237-0.372; $z\text{-}land\ use_{++} = 0.241$; 95% CI = 0.177-0.587). However, the heterogeneity among the effect sizes was much lower ($Q = 21.20$, $P = 0.3852$ and $Q = 11.32$, $P = 0.9375$, for local and land use variables, respectively). Similar to the first axis, the differences between local and land use variables, in terms of their effect sizes on aquatic communities (Figure 3.4b), did not vary consistently with the group of organisms (Figure S3.2b), type of environment (not shown), the difference in the number of variables in each type ($r = -0.086$; $P = 0.712$) and the spatial extent of the studies ($r = 0.043$; $P = 0.854$).

The maximum correlations between the two types of environmental variables (local and land use) and the aquatic communities (along the first ordination axis) were also similar ($z\text{-}local_{++} = 0.562$; 95% CI = 0.408-0.715; $z\text{-}land\ use_{++} = 0.584$; 95% CI = 0.431-0.737). The amount of heterogeneity was high and similar for the two types of environmental variables ($Q = 142.52$, $P < 0.0001$ and $Q = 141.46$, $P < 0.0001$, for local and land use variables, respectively). The number of significant differences between the maximum correlations was more frequent as compared to mean correlations (Figure 3.4c). Specifically, the correlations between local environmental variables and aquatic communities were stronger than the correlations between land use environmental variables and aquatic communities in four datasets. In other four datasets, we found the opposite ($rmax_{Ax1,local} < rmax_{Ax1,land\ use}$). The effects of types of environments (not shown), spatial extent ($r = -0.287$, $P = 0.207$), or biological groups (Figure S3.2c) were also negligible; however, the difference in the number of variables used in the primary studies to represent local and land use variables was positively and significantly correlated with the difference in effect sizes ($r = 0.483$, $P = 0.0168$). Thus, local variables tended to have a variable with higher maximum correlation, as compared to land use variables, when the former type contained a larger number of variables, and vice-versa.

Considering the maximum correlations with the second axis, the difference between the weighted effect sizes of local and land use variables was the highest among the metrics we used ($z\text{-}local_{++} = 0.632$; 95% CI = 0.504-0.761; $z\text{-}land\ use_{++} = 0.422$; 95% CI = 0.310-0.534). Heterogeneity values among the effect sizes were significant for both types of environmental variables ($Q = 65.05$, $P < 0.0001$; $Q = 48.53$, $P = 0.0004$; for local and land use variables, respectively). All five datasets with significant differences between the correlation coefficients indicated that local variables were more strongly correlated with aquatic communities than land use variables (Figure 3.4d). None of the moderators [group of organisms (see Figure S3.2d), types of ecosystems, differences between the number of variables (r = 0. 108, P = 0.640) and spatial extent (r= -0.053, P= 0.821)] were associated with the difference between $rmax_{Ax2,local}$ and $rmax_{Ax2,land\ use}$.

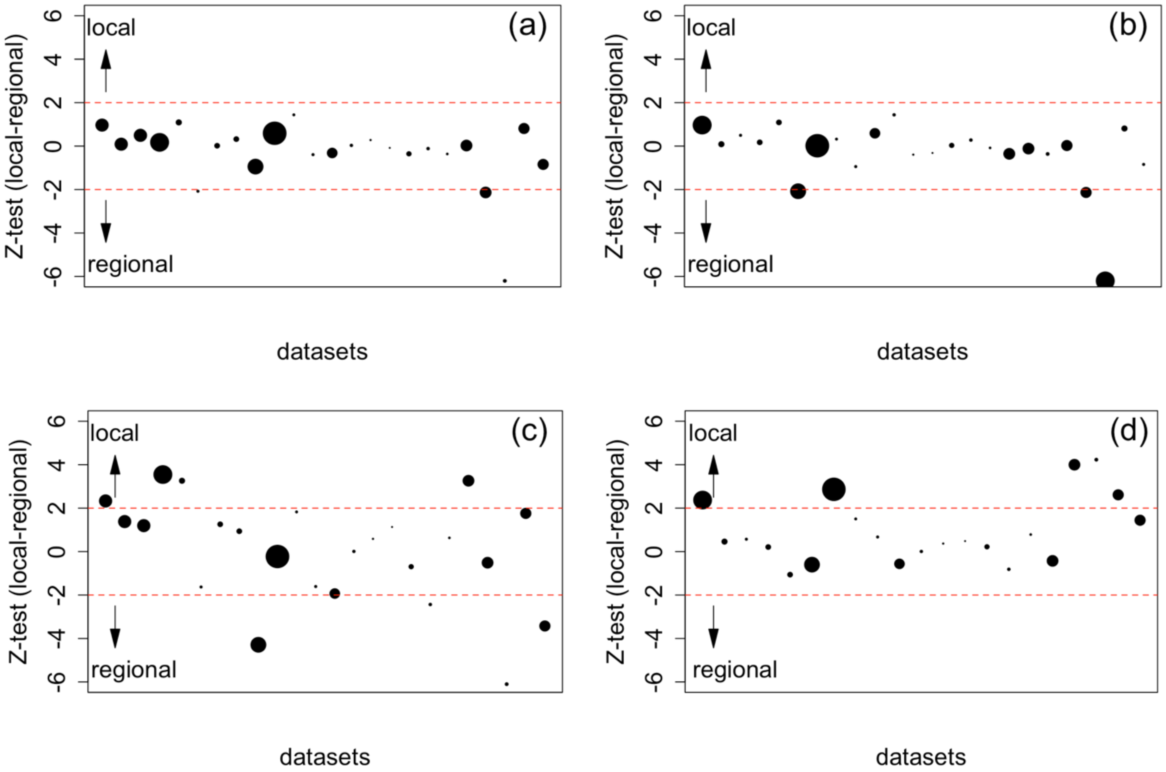


Figure 3.4. Z-tests evaluating the difference in the strengths of the correlations between environmental variables and (local - land use variables) and aquatic communities. Positive values indicate that the relationships between local variables and aquatic communities were stronger than the relationships between land use variables and aquatic communities. Negative values indicate the opposite. Z-values outside the dashed lines are statistically significant. Sizes of the circles are proportional to sample sizes. The analyses were based on the mean correlation for axis 1 (a), mean correlation for axis 2 (b), maximum correlation value for axis 1 (c), and maximum correlation value for axis 2 (d). The results were similar for the different tests discussed by Diedenhofen & Musch (2015).

Discussion

Several studies have emphasized the need to consider variables measured in different scales to better understand the factors underlying community structure (Ward, 1989; Allan & Castillo, 2007; Nessimian *et al.*, 2008; Chessman & Townsend, 2010; Soininen *et al.*, 2011, 2015; Mello *et al.*, 2020). Land use and other regional scale variables have been used to describe processes that structure aquatic communities. This idea relies on the assumption that land use variables can be a better proxy of, for example, nutrient loading (Allan et al., 1997) than local scale variables. Also, land use variables may better indicate the long-term impacts to which a water body is subjected (e.g., Miserendino *et al.*, 2011; Soininen *et al.*, 2015). If so, one would expect that land use variables would be more important predictors of aquatic community structure than local environmental variables (e.g., pH, conductivity, dissolved oxygen, etc.). However, our results did not support this view.

The maximum correlation values along the second axis suggested a greater importance of local rather than land use environmental variables. These results seem to be consistent for different biological groups and ecosystem types. The heterogeneity among the effect sizes estimated with the local environmental variables was lower than the heterogeneity among the effect sizes estimated with the land use variables. Therefore, local scale variables seem to be less variable in predicting the structure of biological communities. On the other hand, the small difference between local and land-use correlations does not allow us to argue that only local variables should be used. Therefore, we support the idea that using both variable types is important (Root & Schneider, 1995) when studying putative correlates of aquatic community structure (e.g., Lenat & Crawford, 1994; Brown & Swan, 2010; Cañedo-Argüelles *et al.*, 2015; Firmiano *et al.*, 2021). For example, the explanatory power of the analyses (e.g., RDA) would be reduced upon the use of only one group of variables.

The relative importance of different environmental factors may be dependent on the spatial scale (Sponseller *et al.*, 2001; Soininen, 2007; Siefert *et al.*, 2012; Zhang *et al.*, 2019). Spatial dynamics can affect the identification of important relationships since environment-community relationships have a strong spatial component (mainly due to spatial autocorrelation). A group of variables structured in space, for example, could emerge as more important (Rocha *et al.*, 2020), resulting in biased results if we seek to disentangle local and land use effects disregarding spatial structure. Spatially structured environmental data have been reported to have higher effect on diversity than pure environmental effects (Soininen, 2007; Zhang *et al.*, 2019), which shows the importance of also considering the spatial structure. Land use variables are to some extent spatially structured. As scale increases the importance of larger scaled variables should increase as well (Soberon & Nakamura, 2009), and the opposite is also expected. Still, we did not find a relationship between the difference between local and land use variables, in terms of their effect sizes on aquatic communities, and spatial extent.

Space must be regarded in variance partitioning to avoid that either the effects of local or regional environmental variables be overestimated due to autocorrelation (Lennon, 2000). If local variables are autocorrelated, then the difference in the correlation that I estimated would be underestimated (i.e., regional effects could be more important after spatial control, implying in more negative Z values in Figure 3.4). On the other hand, if regional variables included were more autocorrelated, then I would be overestimating the difference in correlations (i.e., local effects could be more important after spatial control, implying in more positive Z values in Figure 3.4). However, we were not able to assess the effect of taking the spatial component into account because of our exclusion criterion and also because few studies conducted variation partitioning analyses. Despite the problems of using the results of variation partitioning analyses (as discussed by Siefert *et al.*, 2012), I believe that analyzing this effect, within the context of this work, would be an interesting avenue for future studies.

Our results based on average correlations did not evidence that the strength of the association between environmental variables and aquatic communities are dependent on the type of variable (whether local or land use). The relative importance of the variables was dependent on the difference between the number of variables used to represent each group when the maximum correlations values along the first axis were considered. The higher the number of explanatory variables included (local or land-use) the higher the probability of a variable with larger maximum correlation being encountered. This is a double-edged sword once limnologists could firstly be inclined to increase (even more) the number of variables included in a study. However, as the number of explanatory variables increases, so does the incidence of spurious correlations (Garsd, 1984) and of possible issues with multicollinearity. In addition, the explanatory power increases while our ability to explain our findings is greatly reduced with increasing number of explanatory variables (Weiner, 1995). It has been shown that corelated variables tend to be artificially more important in a regression (Lennon, 2000). Therefore, we reinforce recommendations that we must be cautious when choosing which variables to include in a study.

There seems to be a context dependency concerning the association between communities and the environment (Soininen, 2014). This dependency would be due to differences in environmental variation among regions or differences in the variables chosen as predictors (number and/or specific variables). Notedly, over half of the variables included in the studies were used only once (see Figure S3.1). Few land use variables were used frequently, indicating a lack of consensus on which variables to use among studies. Moving our focus to a set of relevant environmental variables is paramount if we want to advance our knowledge on how communities are affected by land use. I observed in this study that dissolved oxygen, alkalinity, width of the water body, flow of the water body, depth and transparency were the most used local variables, and forest, crop, grass, urbanization, grazing, tree removal and catchment were the most used land use/large scale variables. Note that these were the most used variables in the ordination plots. This implies that they were more selected by significance and other procedures, for example, the forward selection was used to select variables in some of the studies.

A gradient analysis (canonical correspondence analysis, CCA; or redundancy analysis, RDA) is expected in studies of the effects of environmental variables on communities (Soininen, 2014, 2016). A good reporting of the results is necessary for a thorough understanding of the study. Regardless of the method used, reporting a single effect size statistics for each response variable, depicting the amount of variation explained by it, is essential in future studies (Siefert *et al.*, 2012). Reporting the loadings of each environmental variable included and the biplot, in addition to (and not only) the adjusted R^2 of the model is crucial. Knowing how much of the variance (adjusted R^2) was explained shows the strength of the model, while the relationship between predictor variables and response variables (loadings) shows how the variance was partitioned (Legendre & Legendre, 2012). However, methods used varied greatly among studies sampled (114 studies excluded for not conducting or accurately reporting a canonical analysis); this variation was the biggest challenge in this study.

When designing a meta-analysis, one must determine which effect sizes will be used. A highly heterogenous array of methods (as I encountered) hinders the inclusion of many papers. This heterogeneity in the methodological approach in studies that evaluate the effects of variables of different natures/scales on biological communities is not a trivial matter, and has hindered data obtention in this and other meta-analyses (Siefert *et al.*, 2012). We highlight the need to standardize basic reporting items of studies aiming to evaluate the effects of local scale and/or land use variables on biological communities. As such, I recommend that well-reported gradient analyses be called for in papers studying local/land use effects. Specifically, I recommend that studies clearly include (i) a table with the correlations between the environmental variables and sampling scores derived from the biological data (and not only the biplot), (ii) sample size, (iii) spatial extent and (iv) eigenvalues of the axes included. I believe more specific hypotheses and increased parsimony (using relevant variables) in choosing the explanatory variables will produce more reliable results. In addition, I recommend that future studies also partition variance into environmental and spatial components.

Conclusion

In general, the strengths of association between environmental variables and aquatic communities do not depend on the type of variable (local or land use). Thus, using both types of variables is important when studying aquatic communities to increase the explanatory power of the analyses. However, the relative importance of the variables depended on the number of variables used to represent each group for maximum correlation values. We recommend caution and parsimony when choosing to include many variables in a study, in addition to choosing variables that have been reported as relevant for biological communities. A more detailed reporting of multivariate results is also recommended.

References

Allan J.D. & Castillo M.M. (2007). *Stream ecology: structure and function of running waters*.

Allan J.D., Erickson D. & Fay J. (1997). The influence of catchment land use on stream integrity across multiple spatial scales. *Freshwater Biology* **37**, 149–161. https://doi.org/10.1046/j.1365-2427.1997.d01-546.x

Allen A.P., Whittier T.R., Kaufmann P.R., Larsen D.P., O'Connor R.J., Hughes R.M., *et al.* (1999). Concordance of taxonomic richness patterns across multiple assemblages in lakes of the northeastern United States. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 739–747. https://doi.org/10.1139/cjfas-56-5-739

Beisner B.E., Peres-Neto P.R., Lindström E.S., Barnett A. & Longhi M.L. (2006). The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* **87**, 2985–91. https://doi.org/10.1890/0012-9658(2006)87[2985:TROEAS]2.0.CO;2

Borenstein M., Hedges L. V., Higgins J.P.T. & Rothstein H.R. (2009). *Introduction to Meta-Analysis*.

Brown B.L. & Swan C.M. (2010). Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology* **79**, 571–580. https://doi.org/10.1111/j.1365-2656.2010.01668.x

Cañedo-Argüelles M., Boersma K.S., Bogan M.T., Olden J.D., Phillipsen I., Schriever T.A., *et al.* (2015). Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography* **42**, 778–790. https://doi.org/10.1111/jbi.12457

Carneiro F.M., Nabout J.C., Vieira L.C.G., Roland F. & Bini L.M. (2014). Determinants of chlorophyll-a concentration in tropical reservoirs. *Hydrobiologia* **740**, 89–99. https://doi.org/10.1007/s10750-014-1940-3

Chessman B.C. & Townsend S.A. (2010). Differing effects of catchment land use on water chemistry explain contrasting behaviour of a diatom index in tropical northern and temperate southern Australia. *Ecological Indicators* **10**, 620–626. https://doi.org/10.1016/j.ecolind.2009.10.006

Cortes R.M., Varandas S., Teixeira A., Hughes S.J., Magalhães M., Barquín J., *et al.* (2011). Effects of landscape metrics and land-use variables on macroinvertebrate communities and habitat characteristics. *Limnetica* **30**, 347–362

Cottenie K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* **8**, 1175–1182. https://doi.org/10.1111/j.1461-0248.2005.00820.x

Cunha E.J. & Juen L. (2017). Impacts of oil palm plantations on changes in environmental heterogeneity and Heteroptera (Gerromorpha and Nepomorpha) diversity. *Journal of Insect Conservation* **21**, 111–119. https://doi.org/10.1007/s10841-017-9959-1

Dala-Corte R.B., Giam X., Olden J.D., Becker F.G., Guimarães T. de F. & Melo A.S. (2016). Revealing the pathways by which agricultural land-use affects stream fish communities in South Brazilian grasslands. *Freshwater Biology* **61**, 1921–1934. https://doi.org/10.1111/fwb.12825

Dias J.D., Simões N.R., Meerhoff M., Lansac-Tôha F.A., Velho L.F.M. & Bonecker C.C. (2016). Hydrological dynamics drives zooplankton metacommunity structure in a Neotropical floodplain. *Hydrobiologia* **781**, 109–125. https://doi.org/10.1007/s10750-016-2827-2

Diedenhofen B. & Musch J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE* **10**, e0121945. https://doi.org/10.1371/journal.pone.0121945

Fierro P., Bertrán C., Tapia J., Hauenstein E., Peña-Cortés F., Vergara C., *et al.* (2017). Effects of local land-use on riparian vegetation, water quality, and the functional organization of macroinvertebrate assemblages. *Science of The Total Environment* **609**, 724–734. https://doi.org/10.1016/j.scitotenv.2017.07.197

Firmiano K.R., Cañedo-Argüelles M., Gutiérrez-Cánovas C., Macedo D.R., Linares M.S., Bonada N., *et al.* (2021). Land use and local environment affect macroinvertebrate metacommunity organization in Neotropical stream networks. *Journal of Biogeography* **48**, 479–491. https://doi.org/10.1111/jbi.14020

Forbes S. (1925). The Lake as a Microcosm. *Illinois Natural History Survey Bulletin* **15**, 537–550. https://doi.org/10.21900/j.inhs.v15.303

Garsd A. (1984). Spurious correlation in ecological modelling. *Ecological Modelling* **23**, 191–201. https://doi.org/10.1016/0304-3800(84)90100-5

Graham M.H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809–2815. https://doi.org/10.1890/02-3114

Hansson L.-A., Brodersen J., Chapman B.B., Ekvall M.K., Hargeby A., Hulthén K., *et al.* (2013). A lake as a microcosm: reflections on developments in aquatic ecology. *Aquatic Ecology* **47**, 125–135. https://doi.org/10.1007/s10452-013-9435-2

Harding J.S., Benfield E.F., Bolstad P. V., Helfman G.S. & Jones E.B.D. (1998). Stream biodiversity: The ghost of land use past. *Proceedings of the National Academy of Sciences* **95**, 14843–14847. https://doi.org/10.1073/pnas.95.25.14843

Hoffmann M.D. & Dodson S.I. (2005). Land use, primary productivity, and lake area as descriptors of zooplankton diversity. *Ecology* **86**, 255–261. https://doi.org/10.1890/03-0833

Jenkins D.G. & Buikema A.L. (1998). Do similar communities develop in similar sites? A test with zooplankton structure and function. *Ecological Monographs* **68**, 421–443. https://doi.org/10.1890/0012-9615(1998)068[0421:DSCDIS]2.0.CO;2

Legendre P. & Legendre L. (2012). *Numerical Ecology*, 24th edn. Elsevier, Amsterdam.

Lenat D.R. & Crawford J.K. (1994). Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. *Hydrobiologia* **294**, 185–199. https://doi.org/10.1007/BF00021291

Lennon J.J. (2000). Red-shifts and red herrings in geographical ecology. *Ecography* **23**, 101–113. https://doi.org/10.1111/j.1600-0587.2000.tb00265.x

Levin S.A. (1992). The problem of pattern and scale in ecology. *Ecology*. https://doi.org/10.2307/1941447

Liu W., Li S., Bu H., Zhang Q. & Liu G. (2012). Eutrophication in the Yunnan Plateau lakes: the influence of lake morphology, watershed land use, and socioeconomic factors. *Environmental Science and Pollution Research* **19**, 858–870. https://doi.org/10.1007/s11356-011-0616-z

Lodi S., Felipe L., Velho M., Carv P., Bini L.M., Velho L.F.M., *et al.* (2014). Patterns of zooplankton population synchrony in a tropical reservoir. *Journal of Plankton Research* **36**, 966–977. https://doi.org/10.1093/plankt/fbu028

Mello K. de, Taniwaki R.H., Paula F.R. de, Valente R.A., Randhir T.O., Macedo D.R., *et al.* (2020). Multiscale land use impacts on water quality: Assessment, planning, and future perspectives in Brazil. *Journal of Environmental Management* **270**, 110879. https://doi.org/10.1016/j.jenvman.2020.110879

Mikulyuk A., Sharma S., Van Egeren S., Erdmann E., Nault M.E. & Hauxwell J. (2011). The relative role of environmental, spatial, and land-use patterns in explaining aquatic macrophyte community composition. *Canadian Journal of Fisheries and Aquatic Sciences* **68**, 1778–1789. https://doi.org/10.1139/f2011-095

Miserendino M.L., Casaux R., Archangelsky M., Di Prinzio C.Y., Brand C. & Kutschker A.M. (2011). Assessing land-use effects on water quality, in-stream habitat, riparian ecosystems and biodiversity in Patagonian northwest streams. *Science of The Total Environment* **409**, 612–624. https://doi.org/10.1016/j.scitotenv.2010.10.034

Moher D., Liberati A., Tetzlaff J. & Altman D.G. (2009a). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, b2535–b2535. https://doi.org/10.1136/bmj.b2535

Moher D., Liberati A., Tetzlaff J., Altman D.G., Altman D., Antes G., *et al.* (2009b). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine* **6**

Moher D., Shamseer L., Clarke M., Ghersi D., Liberati A., Petticrew M., *et al.* (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* **4**, 1. https://doi.org/10.1186/2046-4053-4-1

Nessimian J.L., Venticinque E.M., Zuanon J., Marco P., Gordo M., Fidelis L., *et al.* (2008). Land use, habitat integrity, and aquatic insect assemblages in Central Amazonian streams. *Hydrobiologia* **614**, 117–131

Ometo J.P.H.B., Martinelli L.A., Ballester M.V., Gessner A., Krusche A. V., Victoria R.L., *et al.* (2000). Effects of land use on water chemistry and macroinvertebrates in two streams of the Piracicaba river basin, south-east Brazil. *Freshwater Biology* **44**, 327–337. https://doi.org/10.1046/j.1365-2427.2000.00557.x

R Development Core Team (2020). *R Development Core Team, R: a language and environment for statistical computing*.

Ricklefs R.E. (1987). Community diverstiy: relative roles of local and regional processes. *Science* **235**, 167–171. https://doi.org/10.1126/science.235.4785.167

Rocha B. da S., Souza C.A. de, Machado K.B., Vieira L.C.G. & Nabout J.C. (2020). The relative influence of the environment, land use, and space on the functional and taxonomic structures of phytoplankton and zooplankton metacommunities in tropical reservoirs. *Freshwater Science* **39**, 321–333. https://doi.org/10.1086/708949

Root T.L. & Schneider S.H. (1995). Ecology and Climate: Research Strategies and Implications. *Science* **269**, 334–341. https://doi.org/10.1126/science.269.5222.334

Siefert A., Ravenscroft C., Althoff D., Alvarez-Yépiz J.C., Carter B.E., Glennon K.L., *et al.* (2012). Scale dependence of vegetation-environment relationships: A meta-analysis of multivariate data. *Journal of Vegetation Science* **23**, 942–951. https://doi.org/10.1111/j.1654-1103.2012.01401.x

Sługocki Ł., Czerniawski R., Kowalska-Góralska M., Senze M., Reis A., Carrola J., *et al.* (2018). The Impact of Land Use Transformations on Zooplankton Communities in a Small Mountain River (The Corgo River, Northern Portugal). *International Journal of Environmental Research and Public Health* **16**, 20. https://doi.org/10.3390/ijerph16010020

Soberon J. & Nakamura M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences* **106**, 19644–19650. https://doi.org/10.1073/pnas.0901637106

Soininen J. (2014). A quantitative analysis of species sorting across organisms and ecosystems. *Ecology* **95**, 3284–3292

Soininen J. (2007). Environmental and spatial control of freshwater diatoms—a review. *Diatom Research* **22**, 473–490. https://doi.org/10.1080/0269249X.2007.9705724

Soininen J. (2016). Spatial structure in ecological communities - a quantitative analysis. *Oikos* **125**, 160–166. https://doi.org/10.1111/oik.02241

Soininen J., Bartels P., Heino J., Luoto M. & Hillebrand H. (2015). Toward More Integrated Ecosystem Research in Aquatic and Terrestrial Environments. *BioScience* **65**, 174–182. https://doi.org/10.1093/biosci/biu216

Soininen J., Korhonen J.J., Karhu J. & Vetterli A. (2011). Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. *Limnology and Oceanography* **56**, 508–520. https://doi.org/10.4319/lo.2011.56.2.0508

Sponseller R., Benfield E. & Valett M. (2001). Relationships between land use , spatial scale and stream macroinvertebrate communities. *Freshwater Biology* **46**, 1409–1424. https://doi.org/10.1046/j.1365-2427.2001.00758.x

Stanfield L.W. & Kilgour B.W. (2013). How proximity of land use affects stream fish and habitat. *River Research and Applications* **29**, 891–905. https://doi.org/10.1002/rra.2585

Vannote R.L., Minshall G.W., Cummins K.W., Sedell J.R. & Cushing C.E. (1980). The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences* **37**, 130–137. https://doi.org/10.1139/f80-017

Viechtbauer W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software* **36**. https://doi.org/10.18637/jss.v036.i03

Ward J. V. (1989). The Four-Dimensional Nature of Lotic Ecosystems. *Journal of the North American Benthological Society* **8**, 2–8. https://doi.org/10.2307/1467397

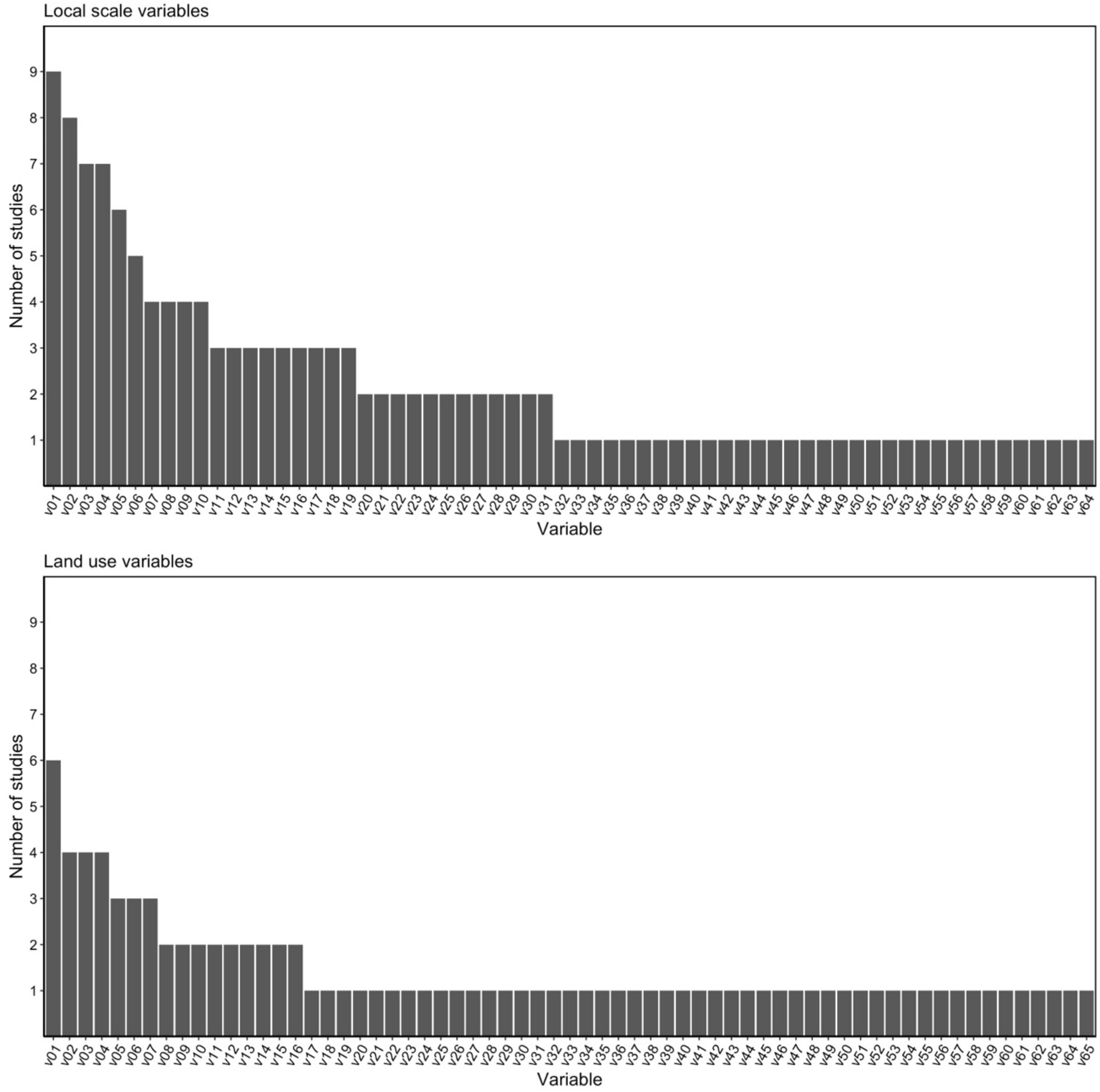
Weiner J. (1995). On the Practice of Ecology. *The Journal of Ecology* **83**, 153. https://doi.org/10.2307/2261159

Wiley Mi., Kohler S. & Seelbach P. (1997). Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. *Freshwater Biology* **37**, 133–148. https://doi.org/10.1046/j.1365-2427.1997.00152.x

Zhang M., García Molinos J., Su G., Zhang H. & Xu J. (2019). Spatially Structured Environmental Variation Plays a Prominent Role on the Biodiversity of Freshwater Macrophytes Across China. *Frontiers in Plant Science* **10**. https://doi.org/10.3389/fpls.2019.00161

Supplementary information

association between communities and the environment


Local scale variables

Land use variables

Figure S3.1. Number of local scale and land use variables included per study. Name of the variables are listed in Table S3.1following the respective numbers.

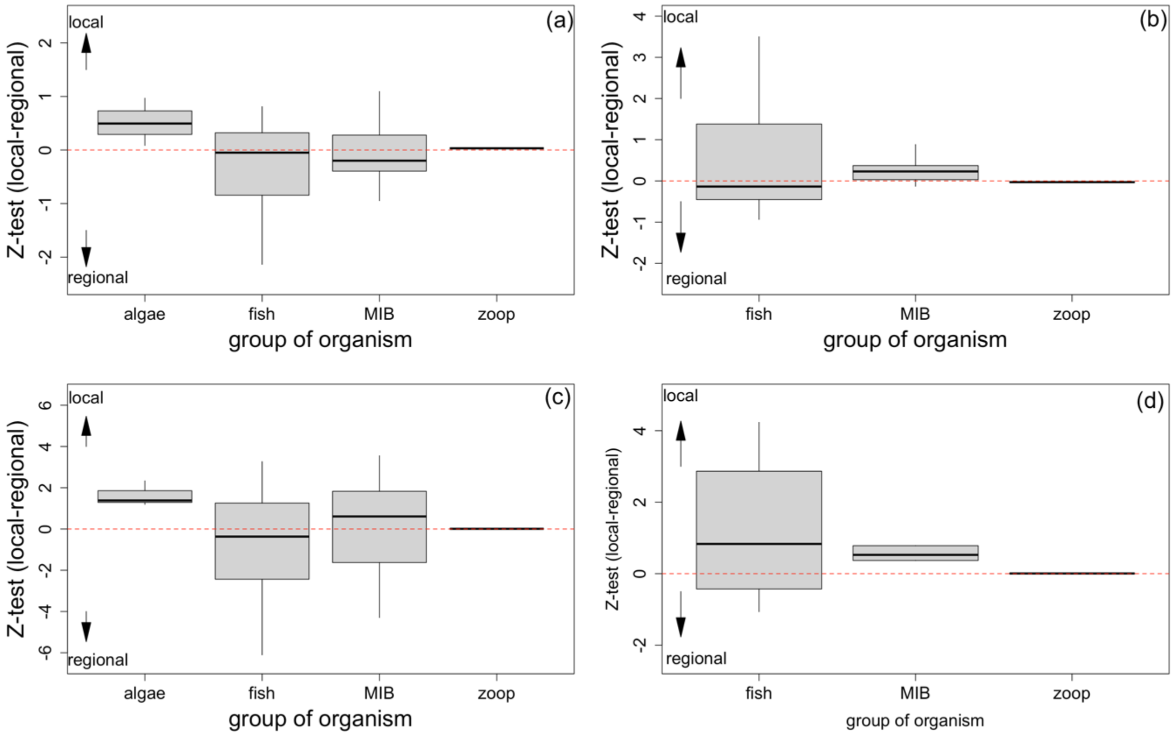


Figure S3.2. Z-tests evaluating the difference in the strengths of the correlations between environmental variables and (local - land use variables) and aquatic communities. Results were aggregated by type of organisms (individual results are shown in Figure 3.4). The analyses were based on the (a) mean correlation for axis 1 (a), mean correlation for axis 2 (b), maximum correlation value for axis 1 (c), and maximum correlation value for axis 2 (d)

Table S3.1. Local and land use scale variables included per study and the number of studies that used the respective variables

	Land use		**Local**	
v1	Forest	6	Dissolved oxygen	9
v2	Crop	4	Alkalinity	8
v3	Grass	4	Width	7
v4	Urban	4	Flow	7
v5	Grazing	3	Depth	6
v6	Tree removal	3	Transparency	5
v7	Catchment	3	Salinity	4
v8	Wetland	2	Total suspended solids	4
v9	Direct human activity	2	Total dissolved solids	4
v10	Preserved riparian	2	N/p	4
v11	Pca1	2	Total organic carbon	3

v12	Pca2	2	Permanganate index	3
v13	Agriculture	2	Total phosphorus	3
v14	Woody debris	2	Conductivity	3
v15	Impervious surface	2	Total nitrogen	3
v16	Altitude	2	Nitrate	3
v17	P fertilization	1	Ammonia	3
v18	K fertilization	1	Ammonium	3
v19	N fertilization	1	Ph	3
v20	C fertilization	1	Biological oxygen demand	2
v21	Pesticides	1	Temperature	2
v22	Total fertilizer	1	Chlorophyll	2
v23	Gross product	1	Phosphate	2
v24	Gross product agriculture	1	Discharge	2
v25	Gross product industry	1	Turbidity	2
v26	Gross other industry	1	Velocity	2
v27	Landfill	1	Altitude	2
v28	Irrigation	1	Wood in stream	2
v29	Tillage	1	Canopy	2
v30	Land slide	1	Coarse substrate	2
v31	Siltation	1	Boulder	2
v32	Area frag	1	Drainage to lake area	1
v33	Northing	1	Rifles	1
v34	Distance from source	1	Sand	1
v35	Regional anthropization index	1	Mud	1
v36	Order	1	Clay	1
v37	Shrub	1	Pca depth	1
v38	3rd order	1	Pca morphometry	1
v39	4th order	1	Pca chemical	1
v40	January temperature	1	Submerged vegetation	1
v41	July temperature	1	Macrophytes	1
v42	Precipitation	1	Litter	1
v43	Turf	1	Shrub	1
v44	Field	1	Debris	1
v45	Debris	1	Tree in stream	1
v46	Litter	1	Manganese	1
v47	Ph	1	Dissolved solids	1
v48	Soil OM	1	Sas5	1

v49	Open water	1	Aspt	1
v50	Other	1	Artificial substrate	1
v51	Sem	1	Hanging vegetation	1
v52	Nd2k (local index)	1	Hag	1
v53	Nd (local index)	1	Mmd	1
v54	Thalweg	1	Silt	1
v55	Riparian width	1	Coarse particulate organic material	1
v56	Bank stability	1	Wet channel width	1
v57	IHAS (local index)	1	Concrete	1
v58	Rain 15 days	1	No plant	1
v59	Rain 30 days	1	PC	1
v60	Rain 7 days	1	Width rapids	1
v61	Latitude	1	Under bank	1
v62	Longitude	1	Sd depth	1
v63	Exotic forest	1	Sd inc	1
v64	Qbr (local index)	1	Sd bank	1
v65	Habitat	1	-	-

Conclusion

Meta-analysis is a method increasingly used in freshwater ecology. There are issues that need to be addressed as quality is regarded. My primary concern is the use of the term ‘meta-analysis’. Authors should be more careful when labeling their study a meta-analysis when, in fact, different methods were used. I find that compliance with checklists should be widely fostered especially due to the implications of meta-analysis results as a robust summary evidence. I reiterate that authors, reviewers and editors should comprehensively use checklists to improve the quality of meta-analyses in freshwater ecology. The checklists and guidelines available provide the means for a good quality meta-analysis to be conducted.

Bias towards positive results is still a matter for concern, despite being historically discussed. I found a prevalence of positive results in freshwater ecology meta-analyses. This prevalence indicates that the high proportion of positive results in primary studies may be propagated to different research levels once a meta-analysis composed of mostly published positive results will probably yield a positive result as well. Meta-analyses have several tools to manage publication bias and using multiple publication bias evaluation methods is recommended. However, a thorough search of the grey literature also helps the inclusion of negative results, reducing possible biases towards positive results. A a long-term approach is also recommended. It’s important that we, as researchers, encourage colleagues and students to focus always more on the questions and proper development of a paper than on the results obtained.

Although some of the items of the quality checklist were not met, interesting results were obtained in our meta-analysis. I found no evidence of differences between land use and local scale variables, although local scale variables seemed to be more precise. In addition, I detected a need to narrow down which variables are included in land use studies as a step towards generalization and stronger analyses. I also recommended more detailed reporting of multivariate results to facilitate future meta-analyses.