A meta-analysis of best management practice effects on nonpoint source pollutants.

Michael Schramm 1, Duncan Kikoyo 2, Janelle Wright 3, Shubham Jain 4

1 Research Specialist IV, Texas Water Resources Institute, Texas A&M AgriLife Research

2 Research Specialist III, Texas Water Resources Institute, Texas A&M AgriLife Research

3 Graduate Research Assistant, Texas Water Resources Institute, Texas A&M AgriLife Research

4 Graduate Research Assistant, Texas Water Resources Institute, Texas A&M AgriLife Research

April 1, 2021

Lorem ipsum

# Introduction

The impact of land use change on hydrology and water quality is well established. Best management practices (BMPs) are control measures used to mitigate increased runoff volume or pollutant concentrations resulting from diffuse nonpoint sources. A suite of empirical approaches and mechanistic watershed models have been developed for practitioners to develop planning scenarios and forecast the hydrologic and water quality impacts of BMP implementation from site to basin scales. Resource managers depend on these models and global databases, such as the International Stormwater Database, to understand the performance of BMPs and plan their design and implementation within impaired watersheds. Despite progress in the development and implementation of BMPs, nonpoint source driven impairments remain a substantial challenge in the United States. Fecal indicator bacteria (FIB) and nutrients, largely driven by nonpoint sources, remain the major cause of water quality impairments in the nation’s rivers and streams despite decades of work. In 2017, the Environmental Protection Agency (EPA) estimated 41% or more of the nation’s rivers and streams rated poorly for biological condition due to excess nitrogen or phosphorus (EPA 2017).

# Methods

## Review and data extraction

We conducted a systematic review of recent (2000-2022) literature to compile field studies documenting the effectiveness of best management practices on fecal indicator bacteria and nutrients concentrations. The systematic review followed guidance provided in the Collaboration for Environmental Evidence systematic review guidelines (Collaboration for Environmental Evidence 2018). In order to maximize the number of studies included in the review, we included both peer-reviewed studies and unpublished white papers to reduce potential bias against negative results. The inclusion criteria filtered out (1) non-field studies, (2) modelling results, (3) studies that did not evaluate specific BMPs, (4) studies conducted outside of the U.S. or published in a language other than English. We ran search queries in Texas A&M Library Catalog, Web of Science, and Google Scholar. Although results from Google Scholar are not always replicable, we utilized the service to maximize search results for studies not published in academic journals.

Results from each database were first filtered to remove duplicates. After removal of duplicates, each member of the research team (*n* = 4) was assigned a subset of studies to evaluate if they should be included. **include supplementary tables with inclusion criteria and data variables** Each study was reviewed by two team members and differences in opinion were collectively discussed and agreed upon before progressing. The remaining studies were split among team members for data extraction, again with at least two team members reviewing each study. If data was provided in figures, the data was extracted with the WebPlotDigitizer tool (Rohatgi 2022).

## Meta-analysis

Robust meta-analysis produces estimates of overall effect size of interventions by weighting individual studies using samples sizes or uncertainty estimates in the calculation of overall effect size (Hedges, Gurevitch, and Curtis 1999). Prior BMP reviews routinely report measures of central tendency (mean or median) of BMP efficiency values (see Agouridis et al. 2005; Kroger et al. 2012; Liu et al. 2015; Simpson and Weammert 2009). Efficiency values are calculated as:

where *xcontrol* is the pre-treatment or control pollutant concentration and *xexperiment* is the pollutant concentration measured after the BMP intervention. The resulting efficiency or percent change values are asymmetric, skewed, and are not additive (Nuzzo 2018; Cole and Altman 2017). Log response ratios (*R*), commonly used in ecological meta-analysis, quantifies the difference in means between the control and experimental group (Hedges, Gurevitch, and Curtis 1999):

where *xi,control* and *xi,experiment* are the mean pollutant concentrations for experiment *i*. The statistical properties of the log response ratio (normal distribution around zero and additive properties) are preferable to the more commonly reported efficiency metric (Osenberg, Sarnelle, and Cooper 1997; Hedges, Gurevitch, and Curtis 1999). This required the exclusion of studies that only provided measures of BMP efficiency and not the underlying data used to derive the metric.

The application of overall BMP efficiency values have been shown to result in inaccurate characterization of treated runoff quality due primarily to the fact that efficiency is a largely a function of influent pollutant concentration (Barrett 2005). In general, higher influent concentrations yield higher efficiencies while relatively clean water yield lower efficiencies. We explored relationships between *R* and influent concentration using meta-analytic multi-level random effects models using the “metafor” package (Viechtbauer 2010) in R (R Core Team 2023).

A key feature of meta-analysis is the weighting of effects based on sampling variance. 54.24% of effect sizes for fecal indicator bacteria were missing standard deviation required to estimate sampling variance. Sampling variance for missing cases was imputed via the weighted average coefficient of variation within groups to avoid bias associated with excluding incomplete cases (Nakagawa, Noble, et al. 2023; Kambach et al. 2020). Log transformed pre-treatment pollutant concentration, BMP category, and interactions were included as fixed effects in the regression model. We also specified a between study random effects term and within study random effects term to account for non-independence of multiple effect sizes report within individual studies. The design of many BMP experiments (pre-post or upstream-downstream studies) are prone to statistical non-independence that inflates Type I error. To account for statistical non-independence among sampling errors we applied robust variance estimation (RVE) methods to obtain adjusted standard errors and confidence intervals (Hedges, Tipton, and Johnson 2010). Relative heterogeneities between and within studies were calculated using the *I2* metric described in Nakagawa and Santos (2012). Marginal *R2* was used to describe the amount of variance explained by the independent variables (Nakagawa and Schielzeth 2013). We tested for evidence of publication bias, in the form of small study effect, by using the extension of Egger’s regression applied to the multilevel model framework that included adjusted sampling error as a moderator (Nakagawa, Yang, et al. 2023). We did not find evidence of publication bias in the surveyed studies for fecal indicator bacteria (include z stat and p value), total nitrogen, total phosphorus, or total suspended solids. Adjustments for publication bias were not included in the final models. We conducted a sensitivity analysis of the overall effect sizes to individual studies using leave-one-out analysis (Nakagawa, Yang, et al. 2023). Outlier studies identified in the sensitivity analysis are highlighted in Table X and were removed prior to fitting the final models.

# Results

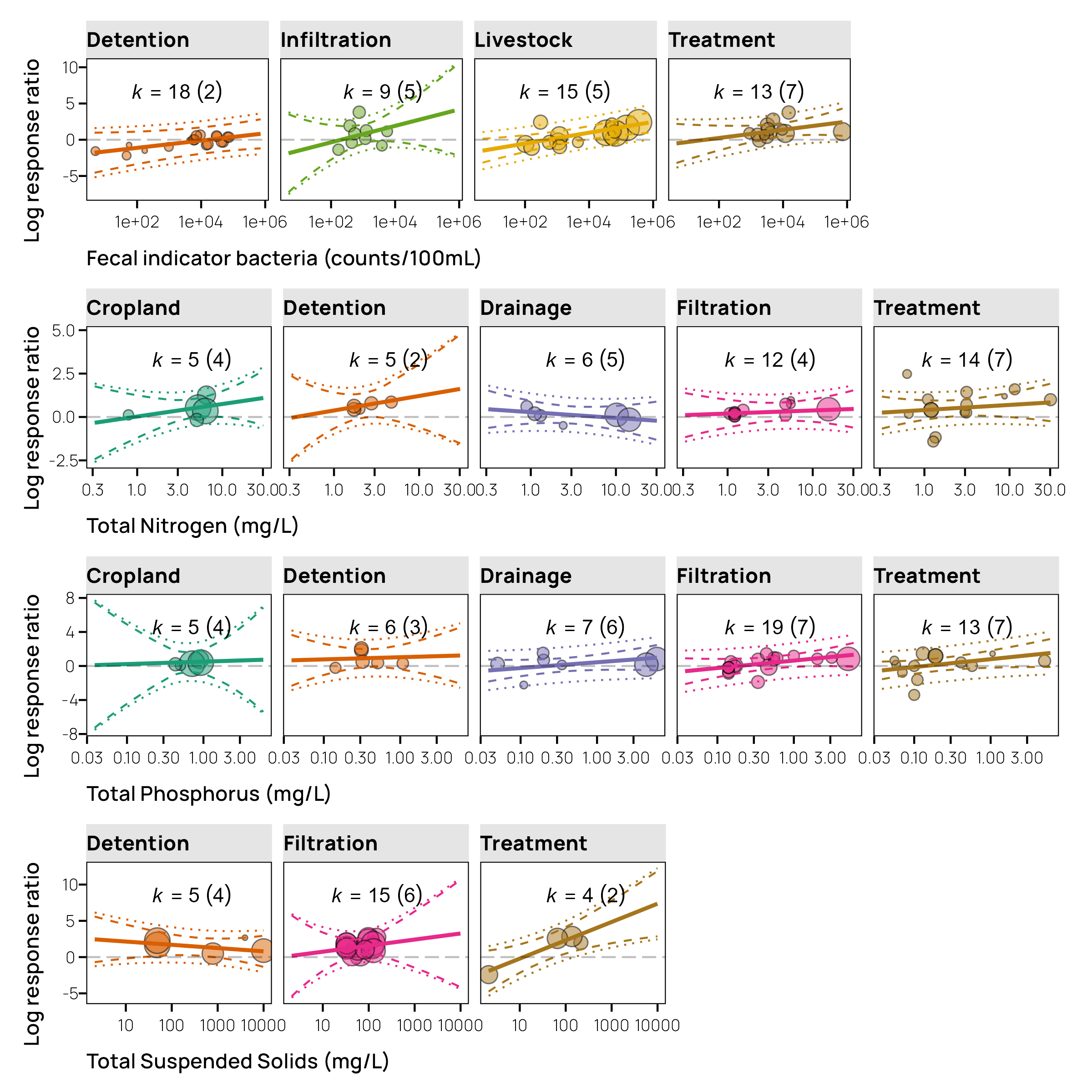


Figure . Marginal means….

# Discussion

## Study limitations

**probably goes in discussion** Our analysis was constrained by the lack of reported sampling variation and potential confounding variables such as watershed conditions or climatic variables that can serve as independent variables in a meta-analysis context. Eagle et al. (2017) reported that five independent teams of researchers identified similar constraints when conducting a meta-analysis of nutrient management BMP studies. The lack of consistent study design and reporting standards presents a significant challenge for pooling results from different studies. Reporting requirements that facilitate meta-analysis are much more common in epidemiology and human health studies where meta-analysis is routinely conducted. The adoption of meta-analysis in fields such as ecological science is more recent. As meta-analysis becomes increasingly utilized in fields such as hydrology and watershed science, we hope that field coalesces around a standard requirement of data required for publication.**probably goes in discussion**

# Conclusion

# Bibliography

Agouridis, Carmen T., Stephen R. Workman, Richard C. Warner, and Gregory D. Jennings. 2005. “Livestock Grazing Management Impacts on Stream Water Quality: A Review.” *Journal of the American Water Resources Association* 41 (3): 591–606. <https://doi.org/10.1111/j.1752-1688.2005.tb03757.x>.

Barrett, Michael E. 2005. “Performance Comparison of Structural Stormwater Best Management Practices.” *Water Environment Research* 77 (1): 78–86. <https://doi.org/10.2175/106143005X41654>.

Cole, Tim J, and Douglas G Altman. 2017. “Statistics Notes: What Is a Percentage Difference?” *BMJ*, j3663. <https://doi.org/10.1136/bmj.j3663>.

Collaboration for Environmental Evidence. 2018. “Guidelines and Standards for Evidence Synthesis in Environmental Management.” Version 5.0. [www.environmentalevidence.org/information-for-authors.](https://www.environmentalevidence.org/information-for-authors.)

Eagle, Alison J., Laura E. Christianson, Rachel L. Cook, R. Daren Harmel, Fernando E. Miguez, Song S. Qian, and Dorivar A. Ruiz Diaz. 2017. “Meta‐analysis Constrained by Data: Recommendations to Improve Relevance of Nutrient Management Research.” *Agronomy Journal* 109 (6): 2441–49. <https://doi.org/10.2134/agronj2017.04.0215>.

EPA. 2017. “National Water Quality Inventory: Report to Congress.” EPA 841-R-16-011. Washington, D.C.: U.S. Environmental Protection Agency. <https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100TG0C.txt>.

Hedges, Larry V., Jessica Gurevitch, and Peter S. Curtis. 1999. “The Meta-Analysis of Response Ratios in Experimental Ecology.” *Ecology* 80 (4): 1150–56. <https://doi.org/10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2>.

Hedges, Larry V., Elizabeth Tipton, and Matthew C. Johnson. 2010. “Robust Variance Estimation in Meta‐regression with Dependent Effect Size Estimates.” *Research Synthesis Methods* 1 (1): 39–65. <https://doi.org/10.1002/jrsm.5>.

Kambach, Stephan, Helge Bruelheide, Katharina Gerstner, Jessica Gurevitch, Michael Beckmann, and Ralf Seppelt. 2020. “Consequences of Multiple Imputation of Missing Standard Deviations and Sample Sizes in Meta‐analysis.” *Ecology and Evolution* 10 (20): 11699–712. <https://doi.org/10.1002/ece3.6806>.

Kroger, R., M. Perez, S. Walker, and A. Sharpley. 2012. “Review of Best Management Practice Reduction Efficiencies in the Lower Mississippi Alluvial Valley.” *Journal of Soil and Water Conservation* 67 (6): 556–63. <https://doi.org/10.2489/jswc.67.6.556>.

Liu, Yaoze, Laurent M. Ahiablame, Vincent F. Bralts, and Bernard A. Engel. 2015. “Enhancing a Rainfall-Runoff Model to Assess the Impacts of BMPs and LID Practices on Storm Runoff.” *Journal of Environmental Management* 147 (January): 12–23. <https://doi.org/10.1016/j.jenvman.2014.09.005>.

Nakagawa, Shinichi, Daniel W. A. Noble, Malgorzata Lagisz, Rebecca Spake, Wolfgang Viechtbauer, and Alistair M. Senior. 2023. “A Robust and Readily Implementable Method for the Meta‐analysis of Response Ratios with and Without Missing Standard Deviations.” *Ecology Letters* 26 (2): 232–44. <https://doi.org/10.1111/ele.14144>.

Nakagawa, Shinichi, and Eduardo S. A. Santos. 2012. “Methodological Issues and Advances in Biological Meta-Analysis.” *Evolutionary Ecology* 26 (5): 1253–74. <https://doi.org/10.1007/s10682-012-9555-5>.

Nakagawa, Shinichi, and Holger Schielzeth. 2013. “A General and Simple Method for Obtaining *R* 2 from Generalized Linear Mixed‐effects Models.” Edited by Robert B. O’Hara. *Methods in Ecology and Evolution* 4 (2): 133–42. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>.

Nakagawa, Shinichi, Yefeng Yang, Erin L. Macartney, Rebecca Spake, and Malgorzata Lagisz. 2023. “Quantitative Evidence Synthesis: A Practical Guide on Meta-Analysis, Meta-Regression, and Publication Bias Tests for Environmental Sciences.” *Environmental Evidence* 12 (1): 8. <https://doi.org/10.1186/s13750-023-00301-6>.

Nuzzo, Regina. 2018. “Percent Differences: Another Look.” *PM&R* 10 (6): 661–64. <https://doi.org/10.1016/j.pmrj.2018.05.003>.

Osenberg, Craig W., Orlando Sarnelle, and Scott D. Cooper. 1997. “Effect Size in Ecological Experiments: The Application of Biological Models in Meta‐analysis.” *The American Naturalist* 150 (6): 798–812. <https://doi.org/10.1086/286095>.

R Core Team. 2023. “R: A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Rohatgi, Ankit. 2022. “WebPlotDigitizer.” Pacifica, California. <https://automeris.io/WebPlotDigitizer>.

Simpson, Dr Thomas, and Sarah Weammert. 2009. “Developing Best Management Practice Definitions and Effectiveness Estimates for Nitrogen, Phosphorus and Sediment in the Chesapeake Bay Watershed.” University of Maryland Mid-Atlantic Water Program. <https://owl.cwp.org/mdocs-posts/developing-best-management-practice-definitions-and-effectiveness-estimates-for-nitrogen-phosphorus-and-sediment-in-the-chesapeake-bay-watershed/>.

Viechtbauer, Wolfgang. 2010. “Conducting Meta-Analyses in R with the Metafor Package.” *Journal of Statistical Software* 36 (3). <https://doi.org/10.18637/jss.v036.i03>.