Published online November 7, 2017

FORUM

Meta-Analysis Constrained by Data: Recommendations to Improve Relevance of Nutrient Management Research

Alison J. Eagle,* Laura E. Christianson, Rachel L. Cook, R. Daren Harmel, Fernando E. Miguez, Song S. Qian, and Dorivar A. Ruiz Diaz

ABSTRACT

Five research teams identified parallel obstacles when concurrently attempting to conduct meta-analyses on the air and water quality impacts of on-farm 4R nutrient management practices. Across projects, system complexity and the lack of relevant data from cultivated and grassland agriculture field trials impeded the application of standard meta-analytical procedures. Because challenges were comparable across projects, the 4R Research Fund technical leadership tasked the researchers with recommending improvements in field research design, data collection, and reporting to enhance future agri-environmental data syntheses and meta-analyses. Here we outline statistical and analytical issues unique to meta-analysis and data synthesis in agriculture, discuss critical data and reporting gaps in the existing literature, and provide specific recommendations for researchers, funders, and journals. Key obstacles developed when field studies did not include complete descriptive or response data (per treatment and experiment year), measurement uncertainty, estimation error in treatment effects, or simultaneously measured nutrient losses and crop yield. Others did not report crop nutrient uptake or their apparent recovery efficiencies. To alleviate such challenges for subsequent research, we make the following recommendations: (i) use common meta-data protocols for consistent units and terminology; (ii) clearly define treatments and controls; (iii) provide complete, tabular, full-factorial response data for each year and location; (iv) collect and report a minimum set of auxiliary data; and (v) establish requirements for data curation and repositories in funding and publication cycles. Implementing these in future nutrient management research will facilitate more robust meta-analyses and other data synthesis efforts.

Core Ideas

- Data and reporting deficiencies reduce the effectiveness of agri-environmental meta-analysis.
- Standardization and consistency across studies will enhance data synthesis and meta-analysis.
- Reporting standard sets of data and meta-data will extend the value of agricultural field research.
- Journals and funders have a unique opportunity to support data management and curation.

Published in Agron. J. 109:2441–2449 (2017)
doi:10.2134/agronj2017.04.0215
Available freely online through the author-supported open access option

Population growth and consumer demand require that agriculture continue to increase cropping system productivity while managing environmental impacts. Policymakers, environmental interests, and producer groups agree on the need for broader application of practices that benefit water or air quality while maintaining or enhancing production. Efficient farm production and environmental management need well-informed and scientifically based strategies to support decision making on practice selection, practice implementation, site selection, and cost-effectiveness. To provide this critical information, the ever-increasing volume of data from agricultural field research must be better summarized, assessed, and interpreted.

Although reviews of the literature summarize available research, systematic quantitative reviews with synthesis go a step further to collate and summarize studies and capture all available information in a repeatable manner (Higgins and Green, 2011). With sufficient available data, meta-analysis uses a variety of statistical methods to quantitatively analyze a treatment effect across multiple studies (Arnqvist and Wooster, 1995; Hedges et al., 1999), estimating overall treatment or factor effects and their precision. By definition, meta-analysis does more than systematic review and synthesis, allowing for quantification of the impact of an experimental treatment relative to a control that is consistently defined across all studies. Meta-analysis broadens the potential impact of primary research studies by placing them as substantial contributions within the larger picture of a research topic (Gerstner et al., 2017).

When results from individual experiments disagree, or when results are inconclusive, the use of meta-analysis can increase “replication” by combining multiple studies, thereby giving a smaller estimation error in overall treatment comparison results. By combining multi-study data with robust statistical methods, meta-analysis of agricultural experimental data can be used to find overall benefits of management practices that

A.J. Eagle, Environmental Defense Fund, 4000 Westchase Blvd., Suite 510, Raleigh, NC 27607; L.E. Christianson, Dep. of Crop Sciences, Univ. of Illinois, 1102 S. Goodwin Ave., Urbana, IL 61801; R.L. Cook, Dep. of Forestry and Environmental Resources, NC State Univ., 2820 Faucette Dr., Campus Box 8001, Raleigh, NC 27695; R.D. Harmel, Center for Agricultural Resources Research, USDA Agricultural Research Service, 2150 Centre Ave., Fort Collins, CO 80526; F.E. Miguez, Dep. of Agronomy, Iowa State Univ., 1206 Agronomy Hall, Ames, IA 50011; S.S. Qian, Dep. of Environmental Sciences, The Univ. of Toledo, 2801 W. Bancroft St., Mail Stop #604, Toledo, OH 43606; D.A. Ruiz Diaz, Dep. of Agronomy, Kansas State Univ., 2711 Throckmorton PSC, Manhattan, KS 66506. Received 12 Apr. 2017. Accepted 18 July 2017. *Corresponding author (aeagle@edf.org).

may otherwise be difficult to fully understand with individual, often short-term, research projects, most of which are limited to particular climatic and soil conditions. For example, recent meta-analyses have improved the scientific certainty as well as the understanding of the implications of tillage on nitrous oxide (N_2O) emissions (van Kessel et al., 2013), the impact of cover crops on N_2O emissions (Basche et al., 2014), and the effect of nitrogen (N) fertilizer on N_2O emissions and the interaction effect between fertilizer rate, fertilizer type, and soil organic carbon (Qian et al., 2010).

In the spring of 2014, the 4R Research Fund supported five research teams to conduct a posteriori (as opposed to predetermined) meta-analyses of the nutrient loss and crop yield impacts of on-farm 4R nutrient management (i.e., application of the Right fertilizer source at the Right rate, at the Right time, and in the Right place) (Bruulsema et al., 2009; IPNI, 2015). The databases compiled or expanded in this work contain research results from a large number of field trials in North American cultivated and grassland agricultural systems. These five evaluations covered a comprehensive range of N and phosphorus (P) loss pathways and yield impacts, with a strong focus on corn-based cropping systems in the midwestern United States (Table 1). All project teams performed a systematic review and synthesis, and, when data were sufficient, a meta-analysis. Researchers also identified key data gaps in peer-reviewed research articles, with the goal of targeting future research into these practices, regions, and cropping systems.

Comparable challenges encountered in all projects arose from incomplete data collection, reporting, and availability within the source literature. Specific improvements in field research design, data collection, and data reporting could address these issues, further supporting well-informed future meta-analyses and syntheses of agricultural research as well as making data more accessible for validation and calibration of process models.

The overall goals in the present paper are to use insights gained in these analyses; to inform field scientists, journal editors, and research funders of the descriptive and response data gaps; to suggest realistic steps to improve reporting and availability of field data; and ultimately to better inform sound environmental nutrient policy, extension, and on-farm decision-making. This article addresses (i) statistical and analytical issues unique to meta-analyses and systematic reviews and syntheses of agricultural research data, with an emphasis on statistical causal inference; (ii) critical data gaps found by the five meta-analysis projects; and (iii) recommendations for funders, researchers, and journals publishing field research results.

STATISTICAL CONSIDERATIONS FOR META-ANALYSIS IN AGRICULTURE

The primary objective of a meta-analysis is to combine studies and accumulate evidence, allowing us to infer the causal effect of the treatment of interest. Statistical methods designed for experimental data evaluate the differences between units with and without the treatment of interest, but causal inference predicts the effect of that treatment as if it were applied to the same experimental units (Gelman and Hill, 2007). As a result, causal inference must be conducted under more

Table 1. Five meta-analysis projects evaluating the nutrient loss and crop yield impact of agricultural management practices, specifically the 4R practices.

Study and key topic	Data source and time frame (number of studies; number of observations)	N and P losses	Crop yield	4R and other practices	Geography; cropping system
Christianson and Harmel (2015a, 2015b; Christianson et al., 2016), N and P losses with artificial drainage	peer-reviewed literature, 1961–2012 (91 studies; 1279 obs)	total and dissolved N; total, dissolved, and particulate P; all as loads in surface and subsurface drainage; no lysimeter studies	yes	source, rate, timing, placement; tillage; crop rotation	North America; all crops, though corn-based systems were primary
Cook et al. (2015), corn yield, N losses, and enhanced efficiency fertilizers	peer-reviewed literature and other reports, 1994–2014 (Yield: 52 studies, 1015 obs; NO_3: 6 studies, 91 obs; N_2O: 11 studies, 127 obs)	N_2O emissions and NO_3 leaching	yes	source (enhanced efficiency products), timing, placement, rate	US Midwest; corn
Eagle et al. (2017), nitrous oxide and nitrate losses in corn-based systems	peer-reviewed literature, 1980–2012 (NO_3: 22 studies, 396 obs; N_2O: 27 studies, 408 obs)	N_2O emissions and NO_3 leaching	Yes (yield was a selection criterion)	source, rate, timing, placement; tillage, crop rotation, cover crops	North America; corn
Ruiz Diaz and Edwards (2016), P surface runoff losses and yield in corn–soybean rotations	peer-reviewed literature and other reports, 1980–2014 (corn yield: 29 studies, 2339 obs; soybean yield: 42 studies, 7666 obs; P loss: 5 studies; 169 obs)	Total P in surface runoff	yes	placement; tillage	US Midwest and Great Plains; corn/soybean
Nummer (2016), conservation practices and P losses	peer-reviewed literature, 1969–2014 (65 studies, 1980 obs)	dissolved, particulate, and total N and P loads and concentrations in surface runoff	yes (when reported)	source, rate, timing, placement; plus conservation practices	North America; all crops

stringent assumptions about the data. These more stringent assumptions are reflected in typical meta-analysis methods, which assume that studies used similar experimental design (most likely with randomization and replication). Furthermore, these individual studies included in the meta-analysis are all designed (or assume to be so) to investigate the causal effect of a specific treatment, so that they are statistically treated as "exchangeable"; they are different but with commonalities (Lindley and Novick, 1981). That is, although data from individual studies cannot be treated as exact replicates, they also are not unrelated separate entities. Combining these data can improve the estimate of the mean treatment effect. Statistical methods for meta-analysis are designed to handle such exchangeable events. Without a consistent design, meta-analysis of these observational data can lead to paradoxical results, a phenomenon widely discussed in statistical literature as Simpson's paradox, Lord's paradox, and the suppression effect (Tu et al., 2008).

Depending on the overall goals, treatment effect sizes for individual studies may be expressed as response ratios, mean differences, scaled differences, and, more rarely, correlations. A response ratio is the ratio of the experimental treatment mean to the control treatment mean (Hedges et al., 1999). Quantification of effect sizes for a treatment requires controls and treatments to be consistently interpreted and applied across relevant studies. However, studies of agricultural management practices rarely use a consistent assignment of treatment and control groups. Furthermore, studies performed in different locations or at different time periods are inevitably associated with different conditions that may also influence the outcome. The conditions may vary because of environment (e.g., weather or soil) or management (e.g., crop variety, planting date, or tillage). As a result, agricultural scientists usually cannot treat different studies as exchangeable. For example, using the MANAGE database (Harmel et al., 2006b, 2008) to evaluate the effect of water and soil conservation practices on nutrient loss from agricultural fields, Qian and Harmel (2016) found that fields with conservation practices tend to have higher fertilizer applications than fields without any conservation practices. As a result, nutrient losses between fields with and without conservation practices (the treatment) cannot be directly compared to evaluate their effectiveness.

Differences between studies in methodology or reporting of response data add other factors that could systematically affect results. In nutrient management studies, researchers may use different application rates for a nitrification inhibitor, measure N_2O loss at different frequencies or time of day, monitor losses for different total time periods, or use different methods for measuring N_2O or nitrate (NO_3) losses. This further increases the value of methods and models that include multiple explanatory factors to evaluate variability and treatment effects.

Although many statistical analyses reveal a correlation, such correlation is not the same as causation. In studying the effect of agricultural management practices, scientists and practitioners are interested in the "causal effect." Consequently, additional steps are needed to ensure that the effect can be truly attributed to the treatment. In statistics, a causal analysis is performed either through careful experimental design (e.g., a randomized experiment) or through additional steps in data analysis when working with "observational" data (i.e., data collected without a standardized process of treatment assignment across studies and possibly with different objectives and study designs) (Rosenbaum, 2002). Because the researcher working with observational data has no influence on study design, factors other than the management practice of interest are not "controlled." With such potential confounding factors, additional effort is needed to try to account for their statistical outcome effect. It is very important to include as much of this information as possible in the analysis. This allows the variation due to confounding factors to be separated from the impacts of management practices (Cochran, 1965).

Although many statistical methods for analyzing observational data are available (Gelman and Hill, 2007; Rosenbaum, 2002), almost all are based on assumptions that cannot all be easily verified. For example, a commonly used method for causal inference using observational data, the propensity score method (Rubin, 2006), assumes that all confounding factors have been included in the data to produce the propensity score. Because all confounding factors are not likely known before a study, it is not possible to know whether this assumption is valid or not. Multiple statistical methods estimating the same effect (Nummer, 2016; Qian and Harmel, 2016), and alternatively specified models within one method (Eagle et al., 2017), address this issue as well as other statistical assumptions, like normality and homogeneity of variances. When these alternative methods result in similar effect estimates, more confidence can be placed in the statistical outcome and inferences.

Individual study effects can be combined across all the studies to generate an overall mean effect, which represents the magnitude of the summarized treatment effect (Fig. 1). The standard methods for meta-analysis necessitate weighting of individual response ratios or effect sizes based on the sample size and variance of each study variable. Thus, studies with more observations and less variability are given greater weight in the overall estimate. However, this sort of statistical information is often not reported for field-scale ecological or agricultural studies. Although resampling procedures such as bootstrapping have been developed to deal with such missing information (Adams et al., 1997; Gurevitch and Hedges, 1999), reporting this variance information in some way or another is preferable and would contribute significantly to improved understanding of treatment effects.

CHALLENGES AND ISSUES IN SYSTEMATIC REVIEWS, SYNTHESES, AND META-ANALYSES

One consistent challenge throughout the five independent data synthesis projects was the difficulty in applying standard meta-analytical approaches to evaluate response effects. This was largely due to system complexity and a lack of necessary information in the published studies. Put simply, it was difficult to locate sufficient levels of available research data that met the selection criteria and reported on yield or nutrient losses in direct side-by-side comparisons of specific management practices. In some cases, research had not been conducted on the topic of interest (i.e., source, time, and place of application, in addition to or interacting with rates), or relevant results had not been published. Therefore, several of the meta-analyses identified a

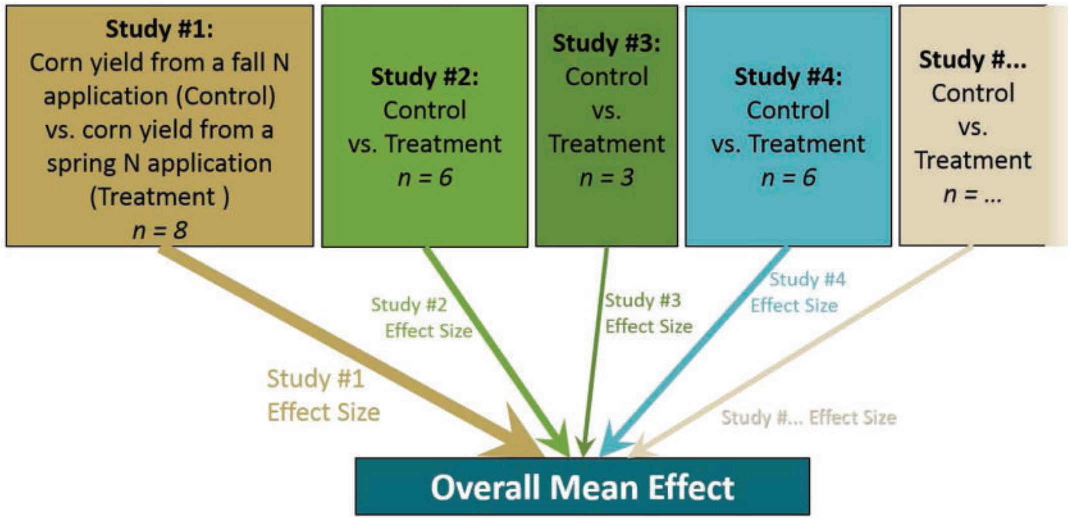

Fig. 1. Multiple studies investigating a treatment contribute to an overall effect size. Examples of weighting shown with box width and arrow size are illustrative only and are not intended to be to scale.

need for more published studies about nutrient loss and yield impacts of agricultural management practices. Published information was particularly lacking in drainage P studies (Christianson et al., 2016) and in Midwestern water quality studies testing enhanced-efficiency N fertilizers (Cook et al., 2015). Although the lack of direct side-by-side comparisons may be partially addressed with alternative modeling methods that can use data across multiple locations while correcting for other factors (Eagle et al., 2017; Qian and Harmel, 2016), this is only possible when data have been collected and reported on specific fertilizer management practices.

Results from some field research could not be used in the meta-analyses (or were difficult to include) because important and relevant information was not reported or control treatments were not consistently defined and applied. Regional differences in climate, soil, and agricultural practice make some variation in experimental methods and objectives inevitable, but meta-analysis techniques are well suited to deal with them as explanatory variables when data are generated by research that applies consistent definitions, units, and treatments. However, this was not always the case. For example, whereas the standard corn grain yield is reported at a moisture content of 15.5%, yield was reported in various studies at 0, 10, 14.5, and 15.0% moisture. Other researchers failed to specify moisture content. In addition, because the incentives and goals of field research do not necessarily align with those of synthesis and meta-analysis, published studies were often missing the additional detailed information on treatments and other management factors that would have been valuable for cross-study comparison.

Other specific challenges included:

- Many studies reported ANOVA results but not the actual measured response data, which does not help in understanding the magnitude of effects or provide enough information for future meta-analysis. Although this reporting method shows significance or lack thereof, it does not allow comprehensive comparison across multiple studies.
- Actual measured results were reported in figures only (and not in tables), so that it was difficult to extract accurate values for modeling or meta-analysis.
- Results reported in figures or tables were often summarized across different treatments, locations, or years, making it impossible to separate data into specific observations associated with key management or environmental factors.
- The majority of field research did not report estimation error in treatment effects (i.e., variability among replicates) in the outcome data, making it difficult to determine the precision of reported values and to assign appropriate weights to studies in the meta-analysis models.
- Most studies did not report any estimate of measurement uncertainty (i.e., uncertainty due to an instrument or method), even though researchers are increasingly calling for and explaining the benefits of publishing such values corresponding to measured data (Beven, 2006; Harmel et al., 2006a, 2009, 2014).
- Few studies reported both N and P losses; nor did they report different forms or loss pathways for an individual nutrient. For example, Eagle et al. (2017) found only one study that reported both N_2O and NO_3 losses, making it difficult to identify management synergies and tradeoffs between these two environmentally sensitive loss pathways.
- Crop yield response to treatments was not always reported, making it difficult to assess the relationship of losses to productivity. Yield (at a specified moisture content) is needed to address losses in yield-scaled units, a concept that has several merits, especially in environmental markets (Johnson et al., 2011; van Groenigen et al., 2010; Venterea et al., 2011; Zhou and Butterbach-Bahl, 2014).
- The frequency or interval of measurement and data collection for NO_3 losses and N_2O emissions was not always reported. However, both timeframe for measuring losses and frequency of measurement can have an impact on losses measured.
- The lack of consistently defined controls and treatments was a significant barrier to the calculation of treatment effects. Greater standardization in experimental design would make it easier to compare effects across locations and cropping systems.

The consistency of the control is crucial for a meaningful meta-analysis. To test alternative fertilizer management practices, however, many studies use a locally standard rate (or source, placement, timing) as their "control." In fertilizer N source comparisons, some studies tested a new fertilizer source against locally standard sources or used different timing or rates with the new source than with the control (Eagle et al., 2017). Although this may make sense from an agronomic practice point of view, the variability in these standards complicates the comparison across studies and environmental conditions. For example, Christianson and Harmel (2015b) examined 85 site-years from 13 studies with N fertilizer applied "out-of-season" (i.e., >2 mo before planting). However, only four and five studies, testing at-plant timing and pre-plant timing, respectively, reported sufficient data for developing a response ratio to compare out-of-season applications with the alternative. Therefore, using "out-of-season" N application timing as the control meant that only studies that reported out-of-season and at least one additional application timing could be used.

In the end, field research studies that reported more detailed experimental and meta-data had a much higher chance of inclusion in the meta-analyses and data syntheses. Although some data limitations can be (and were) addressed by contacting study authors for original datasets, this can pose an undue burden on field researchers. Data and meta-data are also less likely to be available with the progression of time (Michener et al., 1997). As an example of the value of reported data for meta-analysis, Fig. 2 illustrates the total number of studies that measured and reported N_2O or NO_3 losses in North American corn-system fertilizer management studies, separated into different publication time periods (Eagle et al., 2017). Data from studies that measured only short-term losses or reported concentrations or fluxes without cumulative totals could not be used in the meta-analysis, and data that were reported as multi-treatment or multi-year averages posed other problems. The meta-analyses with yield-scaled losses could use only those observations with both loss and yield data reported by year and treatment. Because variability measures were only reported in limited studies, the typical meta-analysis procedure of weighting by variability could not be performed without losing valuable observations.

RECOMMENDATIONS FOR AGRICULTURAL RESEARCH

Specific changes in field research design, reporting, and data publication would ease many of the common challenges experienced and would facilitate future systematic review and *a posteriori* meta-analysis. Recommendations presented here focus on agricultural nutrient management research but could be extended to other topics that have policy implications, especially those affecting agri-environmental interactions. Similar guidelines are advocated in related fields such as ecology and evolutionary science (Gerstner et al., 2017).

Field Research with Common Protocols and Definable Treatments

Thorough, standardized meta-data and response data collection and reporting make cross-study comparisons more reliable. Some recently established large-scale, multi-site agricultural research projects have coordinated project-design efforts with a goal of producing end data that can be combined and reported across experimental studies. Such efforts are far from trivial, given varying study objectives, funding limitations, regional specificity, and other issues. For example, the Long Term Agricultural Research network (Walbridge and Shafer, 2011) spent considerable time and effort developing meta-data protocols that, if maintained for all primary research, could be extremely valuable for meta-analysis.

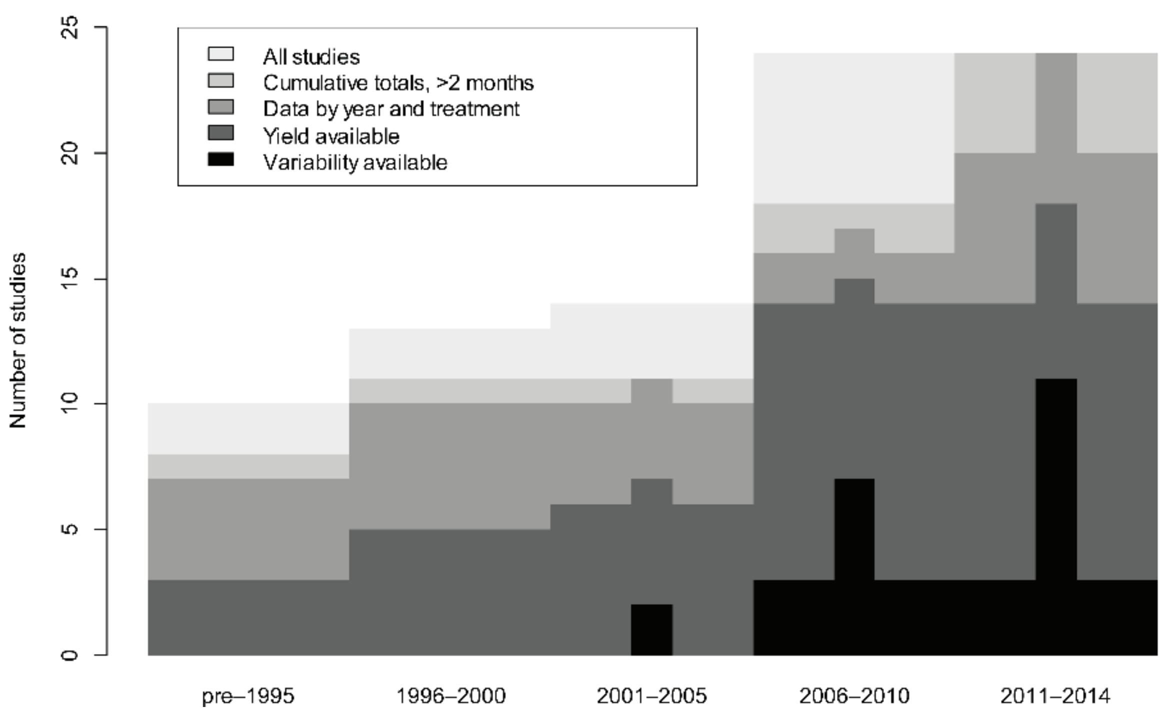

Fig. 2. Quality of data available from peer-reviewed studies of fertilizer management experiments that measured either N_2O or NO_3 losses in corn field trials in North America. Categories with darker shading have progressively more information available. The baseline for each timeframe and data quality category shows the number of papers for which data were available within the article or supplemental material. The mid-bar extensions upward show data gained after direct author contact and subsequent data provision.

Another USDA field research network of 35 experimental sites established certain standard treatments and variables, with room for treatment differences across sites. Standardized protocols were applied to soil and crop measurements, water measurements for drainage, and greenhouse gas flux measurements (Kladivko et al., 2014), and an online data entry interface enforced uniformity and structure (Herzmann et al., 2014). Extension of these protocols beyond these large, broadly coordinated research projects would increase the wider applicability of other field research data. Field studies designed and conducted with consistent measurement protocols facilitate clear comparisons between studies and permit effective comparisons between treatments—across multiple studies—in future meta-analyses.

Consistently defined controls and treatments across studies allow for meaningful comparisons and can reduce otherwise unexplained differences in effect sizes. Further, when creating either empirical or mechanistic models of biogeochemical processes, the implications of an individual practice change should be examined before the interactions between multiple practices (if applicable). For instance, do not change both fertilizer rate and tillage type without looking at the individual effects of each. Although complete factorial designs are a basic tenet of experimental design, short cuts take place because of inadequate time, plot space, or other resources (Casler, 2015). Field research that compares whole systems with one another (e.g., organic versus conventional) (Dobermann, 2012; Seufert et al., 2012) may have multiple explanatory factors for the measured outcome, making it difficult to identify the impact of an individual factor (e.g., reducing N fertilizer rate). Certainly, given the distinctiveness and individual challenges of each field experiment, it is not universally possible to apply a single experimental design. However, to aid meta-analysis and modeling, field research teams and collaborations can strategically design experiments as much as possible to standardize treatments and to change one management factor at a time and then follow up with combinations of different factors.

Recording and Reporting Data

Whereas standard criteria on data completeness and quality can determine the suitability of data for publication (White and van Evert, 2008), meta-analysis can extend the value of these data beyond the original experiment, and beyond replication, to application in policy and other decision-making. Experimental data are appropriately complete when the data and supporting information sufficiently allow the experiment to be reproduced, either in the field or with a simulation model (Hunt et al., 2001, 2006). The quality of a dataset involves precision as characterized by common statistical measures and the completeness of the data. Although it is not always easy for a researcher to anticipate the kind of data that will be important for future meta-analyses, storage in a database with a common protocol facilitates subsequent use.

For agricultural meta-analysis comparison across studies, management practices, and other conditions, the importance of potential controlling factors (whether or not they are statistically significant in the individual experiments) increases substantially. Therefore, noting applicable dates and crop stages, it is important to record all management activities that affect nutrient loss and plant utilization, in addition to general site characteristics such as weather and soil data (Table 2). The highly significant role of environmental controls in nutrient cycling makes it essential to report response data for each year (with different weather patterns) and each location (with different soils and other characteristics). Recording as many covariates as possible means that all confounding factors can be included as random or fixed effects, reducing unexplained statistical error and providing credibility to subsequent meta-analysis.

For numerical values, data should be reported in tables (either in the body of the paper or as supplemental data), not only in figures. Although software programs such as DataThief can help retrieve estimates of graphical data, this requires an extra step and may introduce additional uncertainty in the analyses. It is also essential that all data collected be clearly defined. For example, does the study report total P in runoff, or is it total dissolved? Is the value truly a total based on a digestion of liquid and sediment in the sample, or was sediment filtered?

The uncertainty due to an instrument or method as well as estimation error for treatment effects should also be reported for each type of measured data. For example, Harmel et al. (2006a, 2009) discuss methods to estimate uncertainty in discharge and nutrient flux measurements. This information is valuable for comparing datasets, for model evaluation, and for decision-making based on the measured data. Although meta-analyses in medicine, behavioral science, or even ecology tend to have many observations for each treatment (each of which may have other associated characteristics), the nature of agricultural field experiments with replicated treatments may require adaptation to typical synthesis methods. For example, the reported mean of three replicates for the control is compared with the mean of three treatment replicates, each with variability measures, so that a measure of variability is associated with the treatment effect of a "single" experimental subject. The variability between subjects within a study occurs at another level.

Publishing Data

Publicly available original data, either linked to a publication or within a central data repository, are vital to creating more balanced datasets for meta-analyses and data syntheses. It may not always make sense to publish the full-factorial dataset of field research experiments, as some treatments may not have a significant impact, trends are consistent from year to year, or a treatment has essentially the same response across locations. However, for purposes of meta-analysis, and to understand the amount of variability resulting from weather and other factors, it is helpful to have outcome data available for each treatment comparison, for each year. These data could appear in supplemental materials online or be published with a unique DOI in a central database repository (Brouder and Gomez-Macpherson, 2014). Journal articles should have URLs or DOIs that point to where the primary raw data are located or the best way to obtain data from the authors.

However, with few incentives to publish datasets (and often several disincentives, such as lack of funding or established guidelines), mechanisms to facilitate this process are essential. Even researchers with the best of intentions will often move on to the next project without curating or publishing their datasets. Researchers with valuable datasets may also be

Table 2. Recommended reporting of descriptive and response data for agricultural nutrient management research (including field management, site characteristic, and environmental data) to facilitate meta-analysis and aid modeling efforts.

	Highly recommended data (minimum requirements)	Ideal additional data (preferred requirements)
Management data	planting (date, crop type, seeding rate) tillage (implement type) fertilizer (application date(s), rate, placement, formulation) harvest (date, crop yield, moisture content) pesticide, herbicide (rate, type, application method, timing) grazing (species, intensity, timing) management history, previous years (crop rotation, manure, fertilizer rate for 2 yr)	planting (variety, hybrid) tillage (depth of tillage, intensity/frequency) harvest (method) chemical (formulation) dates for other field activities (e.g., tillage) management history, longer term (10-yr manure application, 5-yr crop, 5-yr tillage)
General site data	location (state, county, latitude/longitude) plot size slope long-term conservation practice (type, NRCS code if applicable) artificial drainage (presence Y/N, tile pipe depth and spacing, drainage area)	long-term conservation practice (location) artificial drainage (tile history, design drainage coefficient) site history
Weather data	deviations from normal weather pattern annual and growing season precipitation (with location information, data can be obtained from nearest weather station)	daily precipitation daily max/min/avg air temperature growing degree days soil temperature wind speed and direction relative humidity radiation
Soil data	dominant soil series drainage class chemical properties (pH, soil organic C, inorganic and organic N, P) physical properties (soil texture) depth to restrictive layer significant changes over time in above properties	plant-available soil N (pre-plant, pre-sidedress/topdress, post-harvest) in the soil to at least 45–60 cm chemical properties (electrical conductivity, K, Mg, Ca) biological properties (respiration, microbial diversity) physical properties (infiltration rate, aggregate stability, bulk density, available water, water-filled pore-space)
Economic data	revenue (yield, commodity price)	costs of management activities, especially those that differ between treatments input costs (fuel, seed, fertilizer, chemical, labor)
Nutrient loss data (in water)	cumulative annual (or growing season) dissolved and particulate (or total) N and P losses differentiate between surface and subsurface losses	differentiate between surface runoff (storm event), baseflow, leaching (i.e., soil water), and artificial drainage measure or estimate volume (e.g., mm d^{-1} drainage discharge, total annual discharge) associated with each transport mechanism partition data into seasonal losses
Nutrient loss data (emissions)	cumulative annual (or growing season) losses	observed range(s) of gaseous fluxes partition data into seasonal losses
Experimental methods	methods used for determining soil properties frequency and number of loss measurements methods for measuring losses, yield, etc. time frame for data collected (seasonal, storm event, growing season, calendar year, water year)	dates for experimental measures (e.g., when soil or plant samples were collected) methods to estimate crop nutrient uptake and grain concentrations
Other	measured N and/or P uptake consistent units across studies for all measured outcomes, report either individual plot values or mean with standard deviation and number of replicates	

reluctant to provide data to multiple users because of time and resource requirements. Traditionally, there has also been a reluctance to make datasets available to others, thus maintaining control for future publications. Well-organized data repositories with effective technical support, an ability to embargo data for given time frames, and financial support within the research structure for data management are key to solving these issues.

Requirements by Journals or Funders

Guidelines or requirements can be put into place by leading agricultural and agri-environmental journals and funding agencies to help ensure that field research results survive beyond the initial publication and can be used for additional meta-analyses or other purposes, such as model validation or calibration. With growing support for open access publications and datasets, journals could require data to be accessible as part of the publication, with delayed access to protect authors' rights to use the data. This is already being practiced by *Nature*, whose policy requires authors to "make materials, data, code, and associated protocols" available on request, and *PLoS* journals, who will only consider manuscripts for which "data and related meta-data underlying the findings" will either be deposited in a public repository or provided as part of the article.

Accompanied by financial and other resources, data management standardization by funding entities and agencies may be more effective than requirements by journals. Such efforts can prioritize the creation and curation of both meta-data and primary datasets. For future meta-analyses, providing honorariums to field researchers that may need to spend a significant amount of time searching for archived data may enhance participation in the collection of unpublished data. Ideally, with proper standards to publish raw data, they can remain available and useful for many researchers rather than disappearing into a researcher's archives.

The increasing use of data repositories, and requirements by funders to use them, also expands available research data beyond that reported in published journal articles. Investment in such efforts promises high returns, allowing for much better synthesis and modeling of field research results and possibly reduction of unnecessarily repeated experiments by better retaining institutional memory. In medicine, data from clinical trials for FDA-approved drugs must be posted at a central website (ClinicalTrials.gov) within 1 yr of study completion, regardless of publication status. Because only about 50% of these results are ever published (Riveros et al., 2013), there is a significant amount of data being saved from possible loss simply by being placed in the repository. In the agricultural research setting, for example, this would mean that subsequent meta-analysis, data syntheses, and modeling would have access to possibly inconclusive experimental results (that when combined with others could tell a complete and compelling story). Similarly, the data from researchers who were unable to publish due to illness, death, or changes in career, and the data from PhD students who left academia before publishing would be accessible. Of key importance for this to be successful is inclusion of data management in research grant budgets and in institutional metrics for promotion and tenure.

CONCLUSIONS

Many barriers exist to full data reporting in agriculture, but as the culture of data sharing has changed to a more "open-source" mentality, opportunities have arisen to learn from the experience of many investigators and studies. It is important to emphasize that we do not advocate for uniformity in primary study objectives and design. Each study is unique and should be designed to test specific hypotheses. Currently lacking, however, are both the availability of meta-data that aid subsequent analysis and better access to primary data that would allow for extraction of relevant meta-analysis statistics. With a few adjustments to how agricultural researchers design experiments and report data, how journals set standards for publication, and how funding agencies direct resources for database sharing and management, agricultural and environmental scientists have an opportunity to apply their collective wisdom and adapt their research methodology to ensure the results are better suited to contribute to the development of a more productive and environmentally sustainable agricultural system.

ACKNOWLEDGMENTS

This work was supported by a grant from the 4R Research Fund (IPNI-2015-USA-4RM11), which is supported by members of and operated by the Fertilizer Institute, the International Plant Nutrition Institute, Fertilizer Canada, and the Foundation for Agronomic Research. We also appreciate the research assistance provided by Cristie Edwards, Kansas State University. USDA is an Equal Opportunity employer and provider.

REFERENCES

Adams, D.C., J. Gurevitch, and M.S. Rosenberg. 1997. Resampling tests for meta-analysis of ecological data. Ecology 78:1277–1283. doi:10.1890/0012-9658(1997)078[1277:RTFMAO]2.0.CO;2

Arnqvist, G., and D. Wooster. 1995. Meta-analysis: Synthesizing research findings in ecology and evolution. Trends Ecol. Evol. 10:236–240. doi:10.1016/S0169-5347(00)89073-4

Basche, A.D., F.E. Miguez, T.C. Kaspar, and M.J. Castellano. 2014. Do cover crops increase or decrease nitrous oxide emissions? A meta-analysis. J. Soil Water Conserv. 69:471–482. doi:10.2489/jswc.69.6.471

Beven, K. 2006. On undermining the science? Hydrol. Processes 20:3141–3146. doi:10.1002/hyp.6396

Brouder, S.M., and H. Gomez-Macpherson. 2014. The impact of conservation agriculture on smallholder agricultural yields: A scoping review of the evidence. Agric. Ecosyst. Environ. 187:11–32. doi:10.1016/j.agee.2013.08.010

Bruulsema, T., J. Lemunyon, and B. Herz. 2009. Know your fertilizer rights. Crops Soils 42:13–18.

Casler, M.D. 2015. Fundamentals of experimental design: Guidelines for designing successful experiments. Agron. J. 107:692–705. doi:10.2134/agronj2013.0114

Christianson, L.E., and R.D. Harmel. 2015a. The MANAGE Drain Load database: Review and compilation of more than fifty years of North American drainage nutrient studies. Agric. Water Manage. 159:277–289. doi:10.1016/j.agwat.2015.06.021

Christianson, L.E., and R.D. Harmel. 2015b. 4R water quality impacts: An assessment and synthesis of forty years of drainage nitrogen losses. J. Environ. Qual. 44:1852–1860. doi:10.2134/jeq2015.03.0170

Christianson, L.E., R.D. Harmel, D.R. Smith, M.R. Williams, and K.W. King. 2016. Assessment and synthesis of 50 years of published drainage phosphorus losses. J. Environ. Qual. 45:1467–1477. doi:10.2134/jeq2015.12.0593

Cochran, W.G. 1965. The planning of observational studies of human populations. J. Royal Stat. Soc., Ser. A-Gen. 128:234–266.

Cook, R., A. Nail, A. Vigardt, A. Trlica, B. Hagarty, T. Williams, et al. 2015. Meta-analysis of enhanced efficiency fertilizers in corn systems in the Midwest. Report to the 4R Research Fund, Project IPNI-2014-USA-4RM06. http://research.ipni.net/page/RNAP-6510 (accessed 10 Apr. 2017).

Dobermann, A. 2012. Getting back to the field. Nature 485:176–177.

Eagle, A.J., J.B. Heffernan, L.P. Olander, K.L. Locklier, and E.S. Bernhardt. 2017. Fertilizer management and environmental factors drive N_2O and NO_3 losses in corn: A meta-analysis. Soil Sci. Soc. Am. J. 81:(in press). doi:10.2136/sssaj2016.09.0281

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge Univ. Press, New York.

Gerstner, K., D. Moreno-Mateos, J. Gurevitch, M. Beckmann, S. Kambach, H.P. Jones, et al. 2017. Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. Methods Ecol. Evol. 8:777–784. doi:10.1111/2041-210X.12758

Gurevitch, J., and L.V. Hedges. 1999. Statistical issues in ecological meta-analyses. Ecology 80:1142–1149. doi:10.1890/0012-9658(1999)080[1142:SIIEMA]2.0.CO;2

Harmel, R.D., R.J. Cooper, R.M. Slade, R.L. Haney, and J.G. Arnold. 2006a. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. Trans. ASABE 49:689–701. doi:10.13031/2013.20488

Harmel, R.D., S. Potter, P. Casebolt, K.H. Reckhow, C. Green, and R.L. Haney. 2006b. Compilation of measured nutrient load data for agricultural land uses in the United States. J. Am. Water Resour. Assoc. 42:1163–1178. doi:10.1111/j.1752-1688.2006.tb05604.x

Harmel, R.D., S.S. Qian, K.H. Reckhow, and P. Casebolt. 2008. The MANAGE database: Nutrient load and site characteristic updates and runoff concentration data. J. Environ. Qual. 37:2403–2406. doi:10.2134/jeq2008.0079

Harmel, R.D., D.R. Smith, K.W. King, and R.M. Slade. 2009. Estimating storm discharge and water quality data uncertainty: A software tool for monitoring and modeling applications. Environ. Model. Softw. 24:832–842. doi:10.1016/j.envsoft.2008.12.006

Harmel, R.D., P.K. Smith, K.W. Migliaccio, I. Chaubey, K.R. Douglas-Mankin, B. Benham, et al. 2014. Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. Environ. Model. Softw. 57:40–51. doi:10.1016/j.envsoft.2014.02.013

Hedges, L.V., J. Gurevitch, and P.S. Curtis. 1999. The meta-analysis of response ratios in experimental ecology. Ecology 80:1150–1156. doi:10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2

Herzmann, D.E., L.J. Abendroth, and L.D. Bunderson. 2014. Data management approach to multidisciplinary agricultural research and syntheses. J. Soil Water Conserv. 69:180A–185A. doi:10.2489/jswc.69.6.180A

Higgins, J.P.T., and S. Green, editors. 2011. Cochrane handbook for systematic reviews of interventions, Version 5.1.0. The Cochrane Collaboration. http://handbook.cochrane.org (accessed 10 Apr. 2017).

Hunt, L.A., G.H. Hoogenboom, J.W. Jones, and J.W. White. 2006. Data standards for agricultural research and decision support. International Consortium for Agricultural System Applications, Honolulu, HI.

Hunt, L.A., J.W. White, and G. Hoogenboom. 2001. Agronomic data: Advances in documentation and protocols for exchange and use. Agric. Syst. 70:477–492. doi:10.1016/S0308-521X(01)00056-7

IPNI. 2015. 4R plant nutrition manual: A manual for improving the management of plant nutrition. T.W. Bruulsema, P.E. Fixen, and G.D. Sulewski, editors, International Plant Nutrition Institute, Peachtree Corners, GA.

Johnson, J.M.F., D.W. Archer, S.L. Weyers, and N.W. Barbour. 2011. Do mitigation strategies reduce global warming potential in the northern U.S. Corn Belt? J. Environ. Qual. 40:1551–1559. doi:10.2134/jeq2011.0105

Kladivko, E.J., M.J. Helmers, L.J. Abendroth, D. Herzmann, R. Lal, M.J. Castellano, et al. 2014. Standardized research protocols enable transdisciplinary research of climate variation impacts in corn production systems. J. Soil Water Conserv. 69:532–542. doi:10.2489/jswc.69.6.532

Lindley, D.V., and M.R. Novick. 1981. The role of exchangeability in inference. Ann. Stat. 9:45–58. doi:10.1214/aos/1176345331

Michener, W.K., J.W. Brunt, J.J. Helly, T.B. Kirchner, and S.G. Stafford. 1997. Nongeospatial metadata for the ecological sciences. Ecol. Appl. 7:330–342. doi:10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2

Nummer, S.A. 2016. Assessing the effects of conservation practices and fertilizer application methods on nitrogen and phosphorus loss from farm fields: A meta-analysis. MS thesis, Univ. of Toledo, OH.

Qian, S.S., T.F. Cuffney, I. Alameddine, G. McMahon, and K.H. Reckhow. 2010. On the application of multilevel modeling in environmental and ecological studies. Ecology 91:355–361. doi:10.1890/09-1043.1

Qian, S.S., and R.D. Harmel. 2016. Applying statistical causal analyses to agricultural conservation: A case study examining P loss impacts. J. Am. Water Resour. Assoc. 52:198–208. doi:10.1111/1752-1688.12377

Riveros, C., A. Dechartres, E. Perrodeau, R. Haneef, I. Boutron, and P. Ravaud. 2013. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. PLoS Med. 10:e1001566. doi:10.1371/journal.pmed.1001566

Rosenbaum, P.R. 2002. Observational studies, 2nd edition. Springer-Verlag, New York. doi:10.1007/978-1-4757-3692-2

Rubin, D.B. 2006. Matched sampling for causal effects. Cambridge Univ. Press, Cambridge, UK. doi:10.1017/CBO9780511810725

Ruiz Diaz, D., and C. Edwards. 2016. Experiences in looking at phosphorus management data for meta-analysis: Challenges and data gaps. In: Proceedings of the Great Plains Soil Fertility Conference, vol. 11. 1–2 Mar. 2016. Kansas State Univ., Manhattan. p. 18–21.

Seufert, V., N. Ramankutty, and J.A. Foley. 2012. Comparing the yields of organic and conventional agriculture. Nature 485:229–232. doi:10.1038/nature11069

Tu, Y.-K., D. Gunnell, and M.S. Gilthorpe. 2008. Simpson's paradox, Lord's paradox, and suppression effects are the same phenomenon: The reversal paradox. Emerg. Themes Epidemiol. 5:2. doi:10.1186/1742-7622-5-2

van Groenigen, J.W., G.L. Velthof, O. Oenema, K.J. van Groenigen, and C. van Kessel. 2010. Towards an agronomic assessment of N_2O emissions: A case study for arable crops. Eur. J. Soil Sci. 61:903–913. doi:10.1111/j.1365-2389.2009.01217.x

van Kessel, C., R.T. Venterea, J. Six, M.A. Adviento-Borbe, B. Linquist, and K.J. van Groenigen. 2013. Climate, duration, and N placement determine N_2O emissions in reduced tillage systems: A meta-analysis. Glob. Change Biol. 19:33–44. doi:10.1111/j.1365-2486.2012.02779.x

Venterea, R.T., B. Maharjan, and M.S. Dolan. 2011. Fertilizer source and tillage effects on yield-scaled N_2O emissions in a corn-cropping system. J. Environ. Qual. 40:1521–1531. doi:10.2134/jeq2011.0039

Walbridge, M.R., and S.R. Shafer. 2011. A long-term agro-ecosystem research (LTAR) network for agriculture. In: Proceedings of the Fourth Interagency Conference on Research in the Watersheds, Fairbanks, AK. 26–30 Sept. 2011. USGS, Reston, VA.

White, J.W., and F.K. van Evert. 2008. Publishing agronomic data. Agron. J. 100:1396–1400. doi:10.2134/agronj2008.0080F

Zhou, M.H., and K. Butterbach-Bahl. 2014. Assessment of nitrate leaching loss on a yield-scaled basis from maize and wheat cropping systems. Plant Soil 374:977–991. doi:10.1007/s11104-013-1876-9