Agriculture, Ecosystems and Environment 148 (2012) 72–82

Contents lists available at SciVerse ScienceDirect

Agriculture, Ecosystems and Environment

journal homepage: www.elsevier.com/locate/agee

Assessment of the quality of meta-analysis in agronomy

Aurore Philibert[a,*], Chantal Loyce[b], David Makowski[a]

[a] INRA, UMR 211 Agronomie, F-78000 Thiverval Grignon, France
[b] AgroParisTech, UMR 211 Agronomie, F-78000 Thiverval Grignon, France

ARTICLE INFO

Article history:
Received 3 June 2011
Received in revised form
28 November 2011
Accepted 1 December 2011
Available online 30 December 2011

Keywords:
Meta-analysis
Nitrous oxide
Mixed model
Bayesian statistics
Sensitivity analysis
Legume

ABSTRACT

A meta-analysis is a statistical treatment of a dataset derived from a literature review. Meta-analysis appears to be a promising approach in agricultural and environmental sciences, but its implementation requires special care. We assessed the quality of the meta-analyses carried out in agronomy, with the intent to formulate recommendations, and we illustrate these recommendations with a case study relative to the estimation of nitrous oxide emission in legume crops. Eight criteria were defined for evaluating the quality of 73 meta-analyses from major scientific journals in the domain of agronomy. Most of these meta-analyses focused on production aspects and the impact of agriculture activities on the environment or biodiversity. None of the 73 meta-analyses reviewed satisfied all eight quality criteria and only three satisfied six criteria. Based on this quality assessment, we formulated the following recommendations: (i) the procedure used to select papers from scientific databases should be explained, (ii) individual data should be weighted according to their level of precision when possible, (iii) the heterogeneity of data should be analyzed with random-effect models, (iv) sensitivity analysis should be carried out and (v) the possibility of publication bias should be investigated. Our case study showed that meta-analysis techniques would be beneficial to the assessment of environmental impacts because they make it possible to study between site-year variability, to assess uncertainty and to identify the factors with a potential environmental impact. The quality criteria and recommendations presented in this paper could serve as a guide to improve future meta-analyses made in this area.

1. Introduction

Systematic reviews are frequently carried out to compile research studies on a specific subject (Evans and Foster, 2011). They involve a rigorous scientific approach comprising the collection, evaluation and synthesis of all studies on a given topic, sometimes contradictory ones, while limiting the introduction of bias (Bland et al., 1995). Systematic reviews may be qualitative if they provide a synthesis of research studies (*e.g.*, Robson et al., 2002), or quantitative, if they involve the processing of a set of data gathered from previous publications. These two kinds of approaches are useful for summarizing large numbers of papers and for the objective establishment (Gurevitch and Hedges, 1993) of what is known and unknown in a specific field (Yuan and Hunt, 2009).

Quantitative systematic reviews are generally referred to as "meta-analyses" when a statistical treatment is applied to a dataset derived from a literature review. The term "meta-analysis" was first coined by Glass in 1976, in the field of educational science, and is defined as a "statistical analysis of a large collection of results from individual studies" (Glass, 1976). A meta-analysis includes typically the following steps (Borenstein et al., 2009; Doré et al., 2011): (i) definition of the objective of the meta-analysis and of the response variable to be estimated from the data. For example, in Miguez and Bollero (2005), the response variable is the ratio of maize yield after a winter cover crop to maize yield in the absence of a cover crop, (ii) systematic review of the literature and/or of the dataset reporting values of the response variable, (iii) analysis of data quality (*i.e.*, quality of experimental design and measurement techniques, precision of the response variable), (iv) assessment of between-study variability and heterogeneity, (v) assessment of publication bias, and (vi) presentation of the results and of the level of uncertainty.

To date, most of the meta-analyses carried out concerned medical science (Normand, 1999; Sutton et al., 2000). In this field, meta-analysis aims (i) to detect an overall treatment effect, (ii) to evaluate the variability between studies, or (iii) to identify study characteristics associated with really effective treatments (Normand, 1999).

Meta-analysis has become an essential technique in human health, and an international organization, the Cochrane Organization, was created in 1993 to prepare, update and promote meta-analyses in this domain (http://www.cochrane.org/). In human health, meta-analyses have long been considered as a field of research in their own right (Cucherat et al., 1997). Meta-analysis

* Corresponding author.
E-mail address: Aurore.Philibert@grignon.inra.fr (A. Philibert).

0167-8809/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.agee.2011.12.003

Table 1
Definition of the eight criteria used to assess the quality of meta-analyses.

Criterion	Definition
Repeatable procedure	A repeatable procedure for the selection of papers for the meta-analysis is presented
References	A list of the references used for the meta-analysis is provided
Heterogeneity	The origins of the variability of the results are analyzed
Sensitivity analysis	The sensitivity of the conclusion to observations or methods is analyzed
Investigation of publication bias	The publication bias is studied
Weighting	Observations are weighted according to their level of accuracy in the statistical model
Availability of the dataset	The dataset is available in an electronic format or published directly in the paper
Availability of the program	The program used for statistical analysis is made available

has been also applied to other areas of science (although less systematically than in human health), such as ecology (*e.g.*, Arnqvist and Wooster, 1995; Cardinale et al., 2006; Stewart, 2010), plant pathology (Rosenberg et al., 2004; Madden and Paul, 2011) and animal science (Sauvant et al., 2008).

Doré et al. (2011) recommended the more systematic use of meta-analysis in agronomy. A considerable amount of experimental data is available from papers published in agronomic journals, and such data could be reviewed, combined and analyzed with statistical techniques to rank cropping systems (within a given environment) according to their impact on crop production and on key environmental variables, such as water nitrate content, the emission of greenhouse gases (*e.g.*, N_2O) or the presence/absence of species of ecological interest (*e.g.*, earthworms, birds). According to Doré et al. (2011), the meta-analysis framework provides an interesting alternative to dynamic crop models (*e.g.*, Brisson et al., 2003; Jones et al., 2003; Keating et al., 2003; Stöckle et al., 2003; van Ittersum et al., 2003) because these models include several sources of uncertainty (Monod et al., 2006) and their predictions are not always reliable (*e.g.*, Barbottin et al., 2008; Makowski et al., 2009).

Meta-analysis appears to be a promising approach for assessing the agronomic and environmental performances of cropping systems, but its implementation requires special care and the value of a meta-analysis may be greatly decreased by the use of inappropriate techniques. Indeed, there is a risk of biased estimation, misinterpretation and incorrect conclusions in meta-analyses performed without sufficient quality control (Sutton et al., 2000). Several authors have proposed quality criteria that could be used to assess the quality of a meta-analysis (Borenstein et al., 2009; Gates, 2002; Roberts et al., 2006; Sutton et al., 2000), but these criteria have not yet been used to assess the quality of the meta-analyses carried out in agronomy.

We therefore assessed the quality of the meta-analyses carried out in agronomy, with the intent to formulate recommendations. We illustrate these recommendations with a case study on the estimation of the emission by legume crops of nitrous oxide, a very potent greenhouse gas with a global warming potential 296 times greater than that of CO_2 (IPCC, 2007).

2. Materials and methods

2.1. Criteria for quality assessment

We defined eight criteria (Table 1), based on the findings of previous studies (Borenstein et al., 2009; Roberts et al., 2006; Gates, 2002), for assessment of the various steps in meta-analyses carried out in agronomy:

(1) Correct description of the bibliographic search procedures used by the authors to select the individual studies (*i.e.*, papers) and the repeatability of these procedures.
(2) Listing of the references of the selected individual studies used in the meta-analysis.
(3) Analysis of the variability of the results of individual studies, including checking to see whether the results vary between the selected individual studies and, when relevant, investigation of the sources of between-study variability (*e.g.*, using random-effects model). Evaluation of the between-study variability of the response variable and of differences in the accuracy of individual estimates is an important step in a meta-analysis and several statistical methods have been proposed for the estimation of between- and within-study variances (Borenstein et al., 2009).
(4) Analysis of the sensitivity of the conclusions to any change in the dataset and/or in the statistical method used to analyze the data. Sensitivity analyses should be carried out to identify influential data and to assess the robustness of the main conclusions of a meta-analysis to the assumptions made in the statistical analysis.
(5) Assessment of the publication bias, which occurs when only studies with highly significant results are published. In this case, a meta-analysis can lead to a biased conclusion and an overestimation of the effect of a given factor. Publication bias is a predominant issue in meta-analysis and several methods such as funnel plots (*e.g.*, Borenstein et al., 2009; Light and Pillemer, 1984) have been developed to detect the presence of such bias in datasets including published results.
(6) *Data weighting*. When the results reported in the individual studies differ in their levels of accuracy, weighting of the data according to their levels of precision is recommended, based, for example, on the inverse of the variance of the measurements, as suggested by Hedges and Olkin (1985).
(7) Availability of the dataset.
(8) Availability of the program used for statistical analysis.

These last two criteria are used to determine whether the meta-analysis could easily be re-run.

2.2. Assessment of the quality of the meta-analyses carried out in agronomy

The quality of the meta-analyses carried out in agronomy was assessed with the eight criteria listed above. One hundred and thirty-six scientific journals publishing papers in agronomy were selected for this purpose. These journals were referred to in the *Journal of Citation Report* (JCR) as journals publishing papers in *Agronomy*, *Agriculture Multidisciplinary*, *Agricultural Engineering*, or *Environmental Sciences*. Journals belonging to the first three categories are further referred to as *Agronomy* and *Agricultural* journals. The scopes of these journals were analyzed and found to be consistent with either the American or European definition of agronomy. The American Society of Agronomy (ASA) definition is "the application of soil and plant sciences to crop production that incorporates the wise use of natural resources and conservation practices to produce food, feed, fuel, fiber, and pharmaceutical crops while maintaining and improving the environment". The definition of the European Society of Agronomy (ESA) is" the relationships between crops, soils, climates and agricultural practices, and between agriculture and the environment".

The papers published in these journals were screened with a systematic literature search (until August 16, 2011) using the key-word "meta-analysis" (or the equivalent in other languages, according to the language of the journal) except for *Environmental Sciences* journals where the key-words "meta-analysis" and crop* were used. The ScienceDirect, WileyInterScience and SpringerLink databases and the websites of the 136 journals were analyzed and 1729 articles were identified in 94 of the 136 selected journals.

The 1729 papers were studied, and 1645 were excluded for the following reasons:

- The word "meta-analysis" appeared only in the references of 1444 papers.
- The topic concerned was related to a discipline other than agronomy (*e.g.*, social science, animal science, focus on forest crop only, plant physiology) in 125 papers.
- Meta-analysis was mentioned but not carried out in 44 papers.
- 32 papers were not research articles (index, table of contents).

Thus, 84 papers in total were found to report the results of a meta-analysis. Thirteen of these papers reported a qualitative analysis of the dataset, 16 reported the results of a quantitative analysis of unpublished data, and 55 reported the results of a quantitative analysis of published data (Fig. 1).

We focused on the 55 papers reporting the results of quantitative analyses of published data. Thirty-one were published in *Agronomy and Agricultural* journals (A&A journals) and 24 in *Environmental Sciences* journals (ES journals). As several meta-analyses were conducted in some of these papers, there were 73 quantitative meta-analyses in total; 45 published in A&A journals and 28 in ES journals. Each of these meta-analyses was assessed according to the eight criteria listed in Table 1. Recommendations concerning statistical procedures in meta-analysis were then formulated, based on the results of the quality assessment.

2.3. Case study on the estimation of nitrous oxide emission by legume crops

The recommendations formulated after the quality assessment were illustrated by a case study of nitrous oxide emission by legume crops during the growing season.

2.3.1. Data

The dataset was extracted from the paper by Rochette and Janzen (2005). It included seventeen values for nitrous oxide emission measured in "pure legume forage stands" – alfalfa (*Medicago sativa*) and clover (*Trifolium pretense* and *Trifolium repens*) – published in nine papers. These papers, published from 1982 to 2004, reported the results of experiments carried out at nine experimental sites. Alfalfa experiments were located in North America and clover experiments in Europe and New Zealand.

The response variable of the meta-analysis was nitrous oxide emission, expressed in $\text{kg N ha}^{-1}\,\text{year}^{-1}$. We used the following explanatory variables: (i) location of the experiment, (ii) type of crop (alfalfa or clover), and (iii) number of days of the experiment.

2.3.2. Statistical models

Sixteen models were fitted to data, to determine the relationship between nitrous oxide emission and the explanatory variables. These models differed in terms of their explanatory variables (with/without crop effect, with/without location effect), the probability distribution of the residual error of the model (with/without weighting by the number of days of the experiment) and the estimation method used (frequentist/Bayesian) (Table 2).

The type of crop was defined as a fixed effect and the location of the experiment was defined as a random effect describing between-site variability.

Specifically, model 5 is defined by: $\log(Y_{ij}) = \mu + \alpha X_{ij} + b_i + \varepsilon_{ij}$ with $b_i \sim N(0,\sigma^2)$ and $\varepsilon_{ij} \sim N(0,\tau^2)$, where Y_{ij} is the nitrous oxide emission in $\text{kg N ha}^{-1}\,\text{year}^{-1}$, μ is emission from alfalfa and $\mu + \alpha$ is emission from clover (X_{ij} is 1 for clover observations and zero for alfalfa observations), $i = 1, \ldots, 9$ is the number of experimental sites (equal to the number of papers) and $j = 1, \ldots, n_i$ with n_i the number of observations per location; $n_i \in [1;5]$. b_i and ε_{ij} are two random terms corresponding to the site effect and to the residual error of the model, respectively, both of which are assumed to be independent and normally distributed. The variability of the measurements within a given location is described by the residual error ε_{ij}. We carried out a log transformation of the response variable to normalize the residuals.

Models 6, 7, and 8 are simplified versions of model 5. Model 6 includes a random location effect but no crop effect, and is defined by $\log(Y_{ij}) = \mu + b_i + \varepsilon_{ij}$. Model 7 includes a crop effect but no location effect, and is defined by $\log(Y_{ij}) = \mu + \alpha X_{ij} + \varepsilon_{ij}$. Model 8 is simply defined by $\log(Y_{ij}) = \mu + \varepsilon_{ij}$.

Models 13–16 are based on the same equations as models 5–8 except that the residual error variance of the model is defined by $\text{Var}(\varepsilon_{ij}) = \tau^2/Z_{ij}$, with Z_{ij} the number of days of each experiment at each location. Thus, according to models 13–16, the higher the number of days of experiment, the lower the variance is.

The parameters of models 5–8 and models 13–16 were estimated from data using a frequentist method (maximum likelihood).

Models 1–4 and models 9–12 are based on the equations of models 5–8 and of models 13–16, respectively, but their parameters were estimated using a Bayesian method (Markov chain Monte Carlo, MCMC). Prior parameter probability distributions were defined with Gamma and Normal distributions, as σ^{-2}, $\tau^{-2} \sim \Gamma(0.001,0.001)$ and $\mu,\alpha \sim N(0,1000)$. μ and α had a mean of zero and a standard deviation of 32, which is large given the range of response variable values, which extended from 0.5 to 5. Along the same lines, σ^{-2} and τ^{-2} had a mean of 1 and a standard deviation of 1000. These distributions represent broad *a priori* distribution when compared with the data.

Frequentist models were assessed by calculating the Akaike Information Criterion (AIC) and Schwartz Criterion (BIC) (Akaike, 1974; Burnham and Anderson, 2002). Bayesian models were assessed by calculating the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). The best models are those with the lowest AIC, BIC and DIC.

The Jackknife technique (Tudoreanu and Phillips, 2004) was used to study the sensitivity of estimates of N_2O emission to each observation of the dataset. Each observation was removed from the dataset in turn, and the models were fitted to the remaining data. The relationship between the values of the measured emission and their influence on the estimates was analyzed.

We used R software for frequentist models and Winbugs software (implementing MCMC algorithms) for Bayesian models. For MCMC simulation, three chains were run, with 20,000–200,000 iterations, depending on the model.

2.3.3. Quality assessment

The dataset used in our case study was published in the paper by Rochette and Janzen (2005) (criterion 1); the list of references and the dataset (quality criteria 2 and 7) were presented in this paper.

In our case study, the heterogeneity (criterion 3) of nitrous oxide emissions between individual studies was analyzed with the statistical models presented above. A sensitivity analysis (criterion 4) was performed to identify influential observations and to assess the effect of the statistical models on the estimated values (by testing sixteen different models). Publication bias (criteria 5) was investigated by generating a funnel plot (Fig. 2). A funnel plot usually presents the precision of the effect size (inverse of its standard

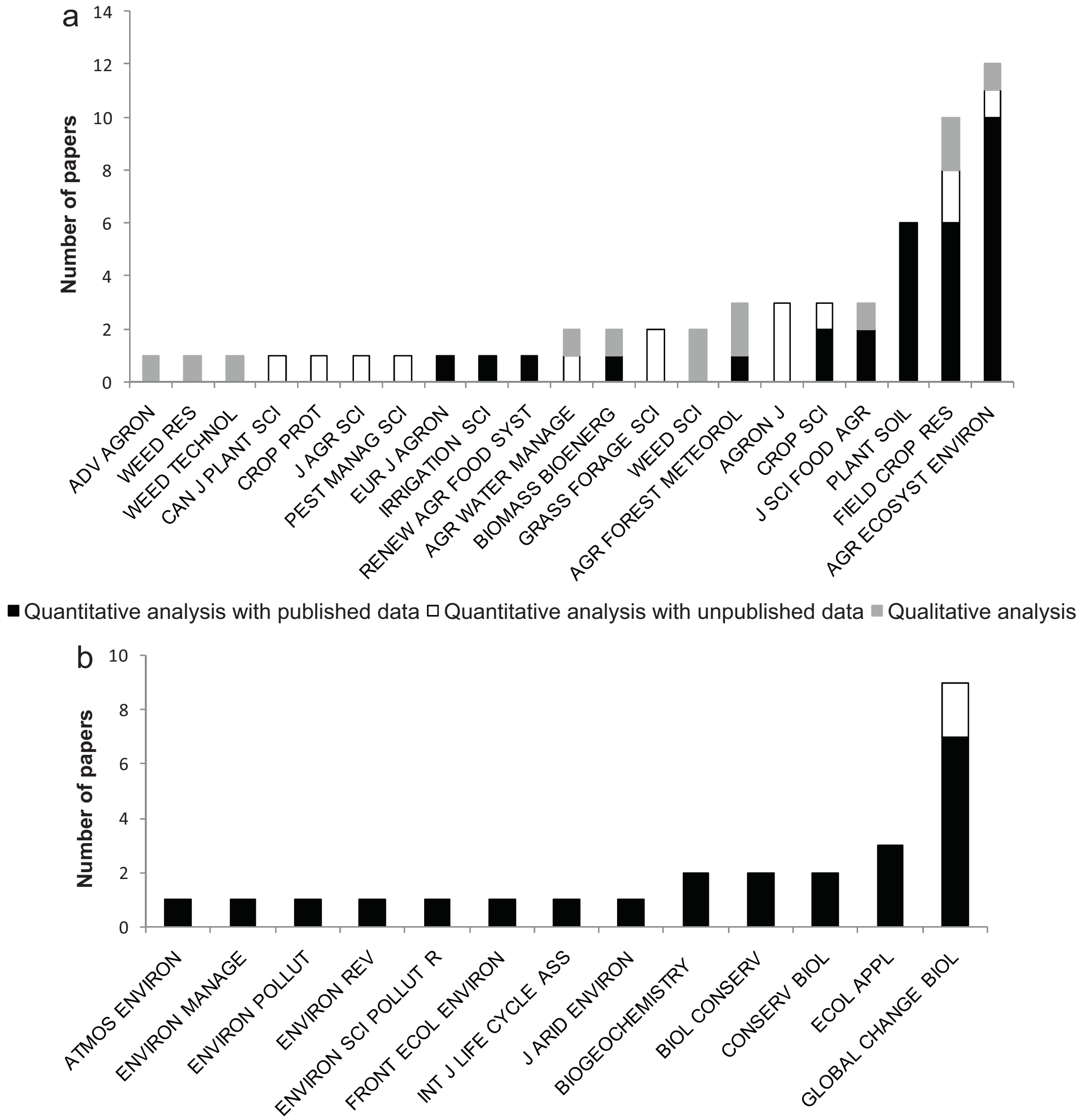

Fig. 1. Number of papers reporting results of meta-analyses in agronomy published in (a) *Agronomy*, *Agriculture Multidisciplinary* and *Agricultural Engineering* journals, and in (b) *Environmental Sciences* journals.

Table 2
Description and assessment of the 16 statistical models for the analysis of N_2O emission data: use of a crop effect, use of a random site effect, use of observation weights, frequentist or Bayesian estimation method, values of the Akaïke, Schwartz, and Deviance Information criteria (AIC, BIC, and DIC).

Model number	Crop effect	Random site effect	Weight	Frequentist/Bayesian	AIC	BIC	DIC
1	Yes	Yes	No	Bayesian	–	–	24.23
2	No	Yes	No	Bayesian	–	–	25.96
3	Yes	No	No	Bayesian	–	–	35.16
4	No	No	No	Bayesian	–	–	37.1
5	Yes	Yes	No	Frequentist	32.67	35.51	–
6	No	Yes	No	Frequentist	34.6	36.92	–
7	Yes	No	No	Frequentist	36.94	39.06	–
8	No	No	No	Frequentist	38.78	40.33	–
9	Yes	Yes	Yes	Bayesian	–	–	23.65
10	No	Yes	Yes	Bayesian	–	–	24.62
11	Yes	No	Yes	Bayesian	–	–	36.43
12	No	No	Yes	Bayesian	–	–	39.01
13	Yes	Yes	Yes	Frequentist	33.84	36.67	–
14	No	Yes	Yes	Frequentist	35.28	37.6	–
15	Yes	No	Yes	Frequentist	38.8	40.92	–
16	No	No	Yes	Frequentist	40.77	42.32	–

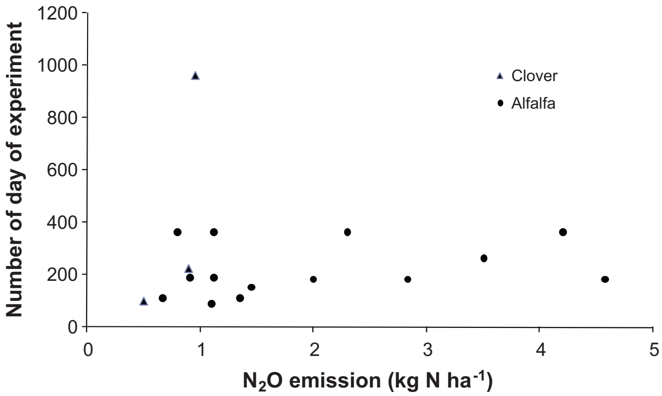

Fig. 2. Funnel plot of nitrous oxide emission and number of days of experiment.

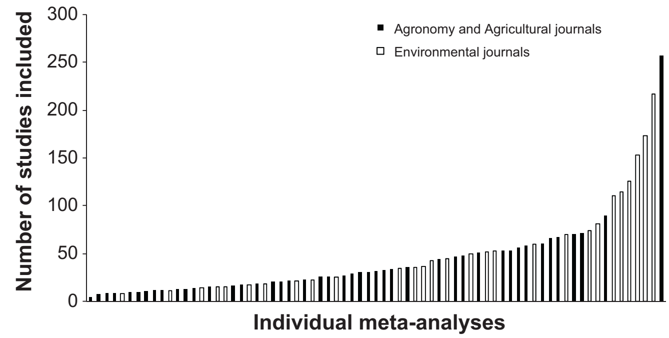

Fig. 3. Number of individual studies used in each of the 73 quantitative meta-analyses. The median number was 32. Results for meta-analyses published in A&A journals were presented in black whereas results for meta-analyses published in ES journals were presented in white.

error) *versus* the effect size (Borenstein et al., 2009). As the variances of the N_2O measures were not available in our dataset, the lengths of the experiments (numbers of days) were plotted in function of the measured N_2O emissions, and no publication bias was found based on this figure (Fig. 2). Individual data were weighted (criterion 6) by the number of days of the experiment (Models 9–16), as the variances for nitrous oxide emissions were not available. Finally, the code used for statistical analysis is available, on request, from the corresponding author (criterion 8).

3. Results

3.1. Description of the 73 selected meta-analyses

The 55 papers (Appendix A) reporting meta-analyses of published data were found in 23 journals (Fig. 1): *Agriculture, Ecosystems & Environment* (10 papers), *Global Change Biology* (7 papers), *Field Crops Research* (6 papers), *Plant and Soil* (6 papers), *Ecological Applications* (3 papers), *Biogeochemistry* (2 papers), *Biological Conservation* (2 papers), *Conservation Biology* (2 papers), *Crop Science* (2 papers), *Journal of the Science of Food and Agriculture* (2 papers), *Agricultural and Forest Meteorology* (1 paper), *Atmospheric Environment* (1 paper), *Biomass and Bioenergy* (1 paper), *Environmental Management* (1 paper), *Environmental Pollution* (1 paper), *Environmental Reviews* (1 paper), *Environmental Science and Pollution Research* (1 paper), *European Journal of Agronomy* (1 paper), *Frontiers in Ecology and the Environment* (1 paper), *International Journal of Life Cycle Assessment* (1 paper), *Irrigation Science* (1 paper), *Journal of Arid Environments* (1 paper) and *Renewable Agriculture and Food Systems* (1 paper). These papers were published from 2001 to 2011. Eleven papers were published between 2001 and 2007, the others being published after 2007.

Twelve of the 55 papers reported results from more than one meta-analysis, bringing the total number of meta-analyses published in these papers to 73, as explained above. The number of individual studies analyzed in these meta-analyses ranged from 5 to 257, with a median of 32 for the 73 meta-analyses, 27 for meta-analyses published in A&A journals and 44 for meta-analyses published in ES journals (Fig. 3).

Crop yield and production quality were used as response variables respectively in 40 and 3 meta-analyses. In fifteen meta-analyses, environmental characteristics: soil carbon (10 meta-analyses), gas emissions (4) and nitrate leaching (1) were considered as response variables. Pests or natural enemies of pests were used as response variables respectively in 5 and 4 meta-analyses. The response variable was a biodiversity indicator in 8 meta-analyses. In the remaining two meta-analyses, the response variable was related to cropping system (crop rotation, Thenail et al., 2009) and plant response (Kaschuk et al., 2011). In four meta-analyses, several types of response variable were considered.

Explanatory variables identified to explain the variability of the response variables were management practices (46 meta-analyses), environmental characteristics (climatic variables in 11 meta-analyses and soil characteristics in 5 meta-analyses), land use (land use characteristics in 9 meta-analyses, land use change in 6 meta-analyses), plant characteristics (3 meta-analyses), and pest natural enemy (1 meta-analysis). Several types of explanatory variables were considered in 5 meta-analyses.

Lastly, in 30 of the 73 meta-analyses, the response variables dealing with crop production were related to explanatory variables dealing with management practices.

The statistical methods used by the authors are summarized in Table 3. Mean effect size method (*i.e.*, mean individual estimated effect size weighted by the inverse of the variance of the effect size) was used in 40 of the 73 meta-analyses. In the other meta-analyses, statistical analysis was based on a regression model with random coefficients (21) or fixed coefficients (8) or statistical tests (4). MetaWin, a specialized software for meta-analyses, was used in 30 meta-analyses, R in 10 meta-analyses and SAS® in 9 meta-analyses. Software and statistical methods were found to be significantly related (Chi-squared test, p value = 2.2×10^{-10}): random coefficients regression models were generally implemented in SAS® and R, and mean effect size was systematically computed with MetaWin. Bayesian statistical techniques were not used in any of these meta-analyses.

3.2. Assessment of the quality of the meta-analyses reviewed

The “References” criterion (criterion 2) was satisfied by 92% of the meta-analyses (87% of the meta-analyses published in A&A journals and 100% of the meta-analyses published in ES journals). The “Heterogeneity” criterion (criterion 3) was satisfied by 88% of the meta-analyses; 84% of the meta-analyses published in A&A journals and 93% of the meta-analyses published in ES journals (Fig. 4). The “Repeatable procedure” criterion (criterion 1) was satisfied by 22% of the meta-analyses (27% published in A&A journals and 14% in ES journals) and publication bias in the selection of papers (criterion 5) was studied in 16% of the meta-analyses (16% published in A&A journals and 18% in ES journals). Sensitivity analysis (criterion 4) was performed in 8% of the meta-analyses (7% published in A&A journals and 11% in ES journals) and the program was made available (criterion 8) in 4% of the meta-analyses (4% published in A&A journals and 4% in ES journals). Observations were weighted (criterion 6) in 37% of the meta-analyses (27% of the meta-analysis published in A&A journals and 54% of the meta-analyses published in ES journals, p value < 0.05). The dataset was

Table 3
Statistical methods and software used in the 73 meta-analyses reporting the results of quantitative analysis of published data.

Statistical methods	Software											Total
	MetaWin	R	SAS	Blossom	Genstat	Irene	MIX	SPLUS	SPSS	Statistica	Unknown	
Mean effect size	30		2								8	40
Random coefficients regression model		6	5		1		1	1	1		6	21
Regression model			2	2		1				1	2	8
Statistical tests		4										4
Total	30	10	9	2	1	1	1	1	1	1	16	73

made available in 18% of the meta-analyses (9% and 32% of the meta-analyses for A&A and ES journals respectively, p value < 0.05).

None of the meta-analyses fulfilled all eight criteria, but three meta-analyses satisfied at least six criteria (Philpott et al., 2008; Kiær et al., 2009; Akiyama et al., 2010).

Publication bias, sensitivity analysis, and availability of the program were the criteria with the lowest scores for both A&A and ES journals. The fourth criterion with the lowest score was the repeatable procedure for meta-analyses published in ES journals, and the availability of the dataset for meta-analyses published in A&A journals.

The use of sensitivity analysis, weighting and Bayesian techniques is illustrated in the case study below.

3.3. Case study of the estimation of nitrous oxide emission by legume crops

3.3.1. Selection of statistical models

AIC and BIC values showed that models including random site effects gave better results; significant between-site variability of N_2O emission was observed. AIC, BIC, and DIC values also showed that the crop effect was significant in both frequentist and Bayesian models. Finally, the use of a model with a weighted residual variance decreased DIC, but not AIC and BIC.

Model 9 (*i.e.*, the model with a weighted residual variance) was the Bayesian model with the lowest DIC. The frequentist model giving the lowest AIC and BIC values was model 5. However, the DIC of model 1 was close to the value obtained for model 9, and the AIC and BIC values of model 13 were close to the values obtained for model 5. Hence, only the results obtained with models 1, 5, 9, and 13 were considered further for the estimation of nitrous oxide emission, since these models showed the best performances (*i.e.*, the lowest AIC, BIC or DIC).

3.3.2. Estimation of nitrous oxide emission

Results were initially obtained in log units and were back-transformed for expression of the model estimates in kg N ha^{-1} $year^{-1}$ units. All the selected models included a random site effect, so their output was the probability distribution describing the between-site variability of N_2O emissions. These distributions are presented for alfalfa and clover separately, in Fig. 5. The means, standard deviations and percentiles of these distributions were similar (Table 4).

All models showed that emissions were much higher in alfalfa than in clover. Estimated mean nitrous oxide emission ranged from 2 to 2.04 kg N ha^{-1} $year^{-1}$ in alfalfa, depending on the model used, and from 0.86 to 0.97 kg N ha^{-1} $year^{-1}$ in clover. The 95% confidence intervals obtained were slightly larger with frequentist models than with Bayesian models for alfalfa N_2O emission, and larger for Bayesian models than for frequentist models for clover N_2O emission. However, overall, the two types of model gave similar results. The differences between the 2.5 and 97.5 percentiles were large, with all models revealing considerable between-site variability for N_2O emissions (Table 4 and Fig. 5).

We compared the mean values estimated with our models with the estimate used by Rochette and Janzen (2005) (1.8 kg N ha^{-1} $year^{-1}$ for clover and alfafa), by calculating the probability of exceeding this value based on the probability distributions shown in Fig. 5. The value of 1.8 kg N ha^{-1} $year^{-1}$ was exceeded with a probability close to 0.5 for alfalfa and a probability of 0.04–0.08 for clover, depending on the model (Table 4). As the between-site variability of N_2O emissions is strong, it is more relevant to present

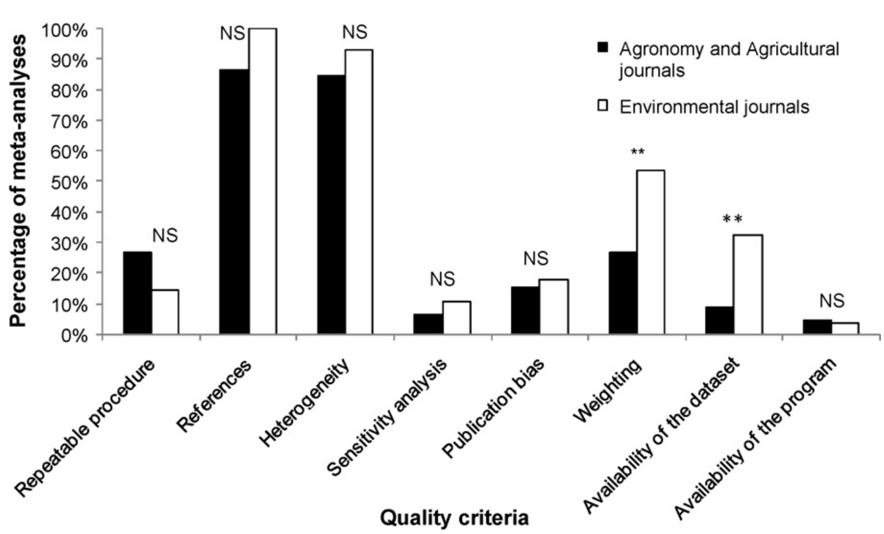

Fig. 4. Number of meta-analyses (out of 73 studied) satisfying each of the 8 quality criteria. NS for a non-significant difference for this criterion, between meta-analyses published in A&A and ES journals and ** for a significant difference with a p value < 0.05. Results for meta-analyses published in A&A journals were presented in black whereas results for meta-analyses published in ES journals were presented in white.

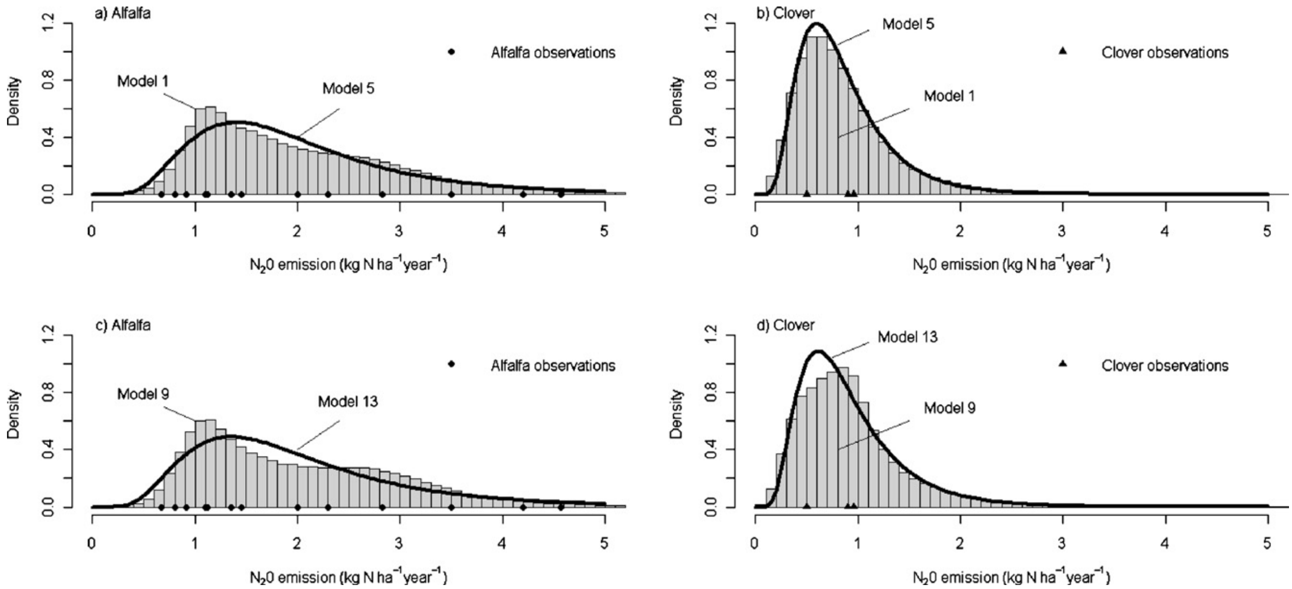

Fig. 5. Probability distribution of N_2O emission obtained with frequentist models 5 and 13 (black curves) and with Bayesian models 1 and 9 (gray histograms). The seventeen observations used for the statistical analysis are represented in black circles for alfalfa crops and in black triangles for clover crops.

confidence intervals for N_2O emissions (Table 4) than absolute-number estimates (*e.g.*, 1.8 kg N ha^{-1} year^{-1}).

3.3.3. Analysis of sensitivity to individual data

The results of the Jacknife procedure are presented in Fig. 6. The circles and triangles at some distance from the gray vertical line indicate influential data respectively for alfalfa and clover. The estimated mean level of N_2O emission in clover was sensitive to the data collected in clover fields, but not to the data collected in alfalfa fields (Fig. 6a and c). For N_2O emissions in clover, the most influential observation was #1; the removal of this observation increased the estimated mean emission from 0.86 to 1.06 kg ha^{-1} year^{-1} with model 5, and from 0.97 to 1.18 kg ha^{-1} year^{-1} with model 9.

Estimated mean N_2O emission in alfalfa was sensitive to data #9, #12, #13, #15, #16 and #17 (Fig. 6b and d). These data were collected in alfalfa fields. The strong influence of data #15, #16 and #17 may reflect the association of these data with high N_2O emissions. However, data #9, #12 and #13 were also influential but corresponded to only moderate levels of N_2O emission.

The sensitivity analysis identified several influential data, but mean emissions in clover crops never exceeded 1.18 kg N ha^{-1} year^{-1}, whereas alfalfa emissions were never lower than 1.8 kg N ha^{-1} year^{-1}; mean emissions were thus consistently higher in alfalfa crops located in North America than in clover crops located in Europe and New Zealand.

4. Discussion

It is useful to compare our quality assessment with similar quality assessments carried out in the past for meta-analyses in the fields of ecology and medical science. In our discussion of the quality of meta-analyses in agronomy we consider two previous studies for this purpose: Roberts et al. (2006), comparing the characteristics of 73 meta-analyses in ecology with those of 73 meta-analyses in medical science (*i.e.*, the same total number of meta-analyses as the number considered in our study) and Gates (2002), analyzing 29 meta-analyses in ecology.

The results of these previous studies show that criterion 1 (description of the procedure used for literature search) is more frequently satisfied in medical science (meta-analyses described in 2006: 100%) than in ecology (meta-analyses described in 2002: 66%, p value < 0.05; meta-analyses described in 2006: 0%, p value < 0.05) and in agronomy (meta-analysis described here: 22%, p value < 0.05). The references of individual studies were presented in details in all fields (medical science: 100%; ecology, 2002: 93%; ecology, 2006: 100% and agronomy: 92%). However, the existence of a possible publication bias was rarely investigated in any of the fields considered (medical science: 23%; ecology, 2002: 34%; ecology, 2006: 8% and agronomy: 16%) despite the development of several methods for dealing with this issue (*e.g.*, funnel plot). Sensitivity analyses were more frequently carried out in medical science (32%) than in ecology (2006: 0%, p value < 0.05) or agronomy (8%,

Table 4
Mean, standard deviation, 2.5% and 97.5% percentiles of nitrous oxide emission calculated with models 1, 5, 9 and 13 for alfalfa and clover crops.

Model number	Crop	N_2O emission (kg N ha^{-1} year^{-1})				P (N_2O emission ≥ 1.8)[b]
		Mean	SD[a]	Percentile 2.5%	Percentile 97.5%	
1	Alfalfa	2	1.10	0.75	4.36	0.49
5	Alfalfa	2.03	1.08	0.68	4.78	0.50
9	Alfalfa	2.02	1.10	0.71	4.32	0.49
13	Alfalfa	2.04	1.14	0.64	4.96	0.49
1	Clover	0.89	0.63	0.24	2.38	0.06
5	Clover	0.86	0.45	0.28	2.01	0.04
9	Clover	0.97	0.88	0.24	2.61	0.08
13	Clover	0.92	0.51	0.29	2.23	0.06

[a] SD: standard deviation.
[b] The probability of exceeding the value of 1.8 kg N ha^{-1} year^{-1} was calculated with each model for each crop.

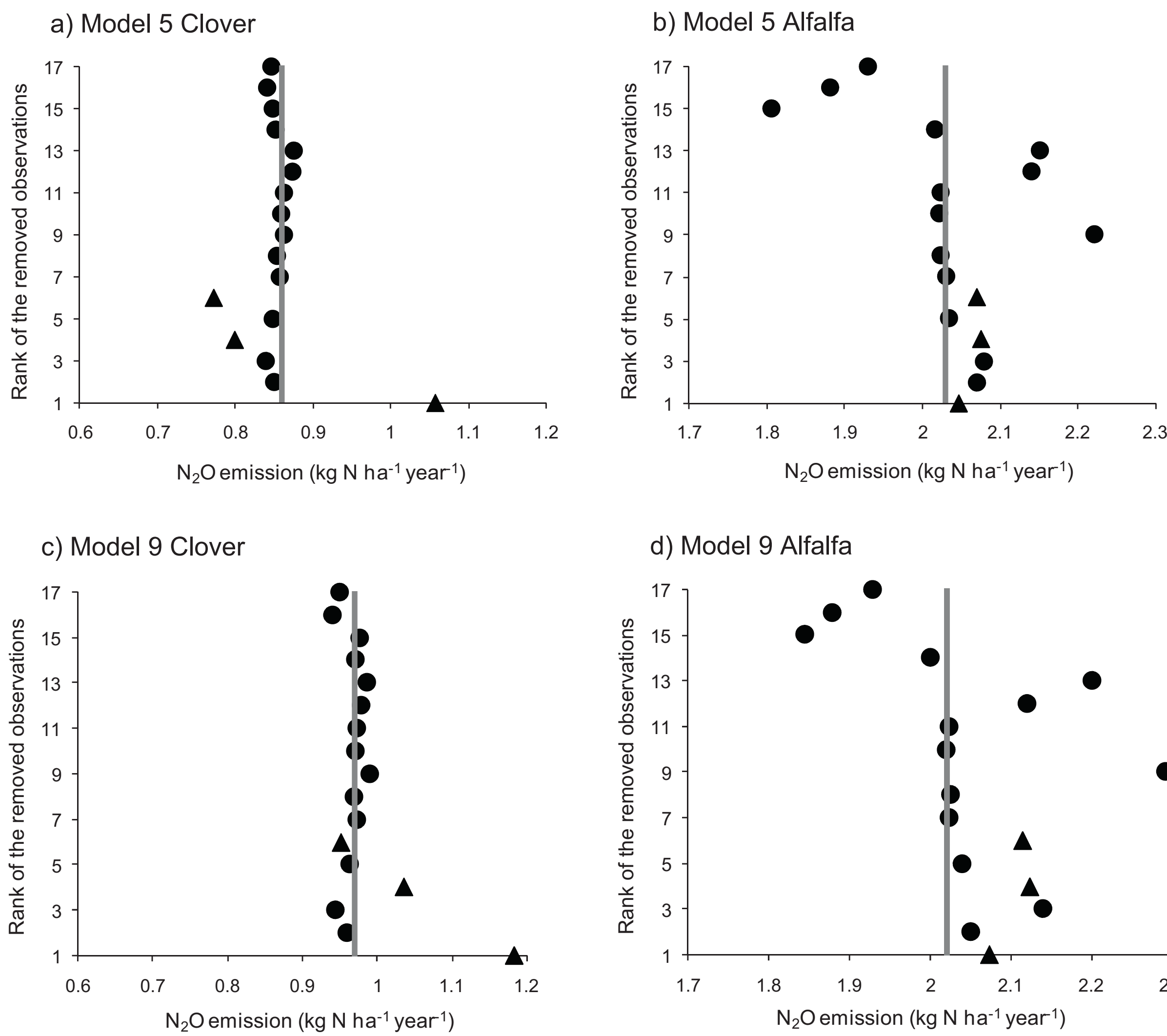

Fig. 6. Sensitivity analysis for models 5 and 9. Each point indicates N_2O emission estimated (on the *x* axis) with model 5 (a for clover and b for alfalfa) and model 9 (c for clover and d for alfalfa), when a given observation (on the *y* axis) is removed from the dataset. The *x* axis shows estimated nitrous oxide emissions. They *y* axis shows the rank of the information removed. Observations were ranked in decreasing order of N_2O emission (observation number 1 has the lowest value of nitrous oxide emission). Gray vertical lines represent the estimate obtained with the model for all observations. Black circles show estimates of nitrous oxide emission when one alfalfa observation is removed and black triangles show estimates when one clover observation is removed.

p value < 0.05). Sensitivity analysis was performed in less than one third of the reviewed meta-analyses in all fields, despite the work of Oxman (1994) stressing the importance of sensitivity analysis for systematic reviews. According to Oxman, sensitivity analysis is a useful method for "testing how robust the results of a review are relative to key decisions and assumptions that were made in the process of conducting the review".

The heterogeneity of individual results was more frequently studied, in all fields (medical science: 49%; ecology, 2002: 76%; ecology, 2006: 16%; agronomy: 88%). Sources of heterogeneity were analyzed almost systematically in agronomy, probably because it is important for agronomists to identify the crop management techniques and environmental characteristics yielding the best performances. As shown above, most of the analyses carried out in agronomy did in fact explain crop production variations according to crop management techniques.

Overall, more quality criteria were satisfied in past meta-analyses in medical science than in agronomy, with one exception: the criterion concerning the heterogeneity of the individual studies. The meta-analyses performed in ecology and in agronomy were more similar in terms of quality. None of the 73 meta-analyses in agronomy satisfied all eight criteria and only three satisfied six of these criteria.

The higher quality of meta-analyses in medical science is probably due to the intensive research conducted on meta-analysis in this area. In 1997, the Cochrane Collaboration (van Tulder et al., 1997 updated in 2003: van Tulder et al., 2003 and 2009: Furlan et al., 2009), introduced methodological guidelines for systematic reviews, suggesting a framework to help researchers with quantitative synthesis. Other groups, such as the QUORUM (Quality of reporting meta-analyses) group, have provided advice about ways of carrying out high-quality meta-analyses. This group, consisting of 30 epidemiologists, clinicians, statisticians, editors and researchers, produced a checklist of standard items to be included in meta-analyses (Moher et al., 1999).

An important difference between medical and agronomical meta-analyses is due to the fact that, in medicine, results are almost entirely constituted of Randomized Control Trials including a control treatment (Begg et al., 1996; Gates, 2002) whereas in agronomy experiment designs are more diverse and do not systematically include a control treatment. The between site-year variability of the agronomical response variable is usually high, and it is recommended to analyze this variability using a mixed-effect model as shown in our case study. However, results obtained with such models should be interpreted with care due to the possible effect of confounding variables. For example, in our case study, we found that emissions were significantly higher in alfalfa than in clover, but this difference may be as well due to a geographical effect because all the alfalfa experiments were located in North America and all the clover experiments were carried out in Europe and New Zealand.

Another interesting finding of the study was that Bayesian methods are entirely absent from the meta-analyses carried out in agronomy, despite the implementation of such methods in meta-analyses in medical science since the early 1990s (Smith et al., 1995; Sutton and Abrams, 2001) and more recently in ecology (Myers, 2001; Stewart, 2010). However, frequentist and Bayesian statistical methods were compared in our case study and the two types of methods were found to give similar results. The conclusions were thus insensitive to the type of statistical method used for parameter estimation. As this comparative study is the first one performed with agronomic data, further studies are required to determine the value of Bayesian statistics in meta-analysis.

Based on our quality assessment, we were able to formulate several recommendations for improving future meta-analyses in agronomy:

- The procedure used to search for papers in scientific databases should be presented.
- Individual data should be weighted according to the level of precision (like in ecology, as mentioned by Stewart (2010)). A typical approach involves weighting data by the inverse of the standard error of the measurements. In our case study, we used the number of days of the experiment, because standard error values were not available.
- The heterogeneity of data in individual studies should be analyzed with random-effect models (Madden and Paul, 2011). These models are useful for analyzing the between site-year variability of the response variables and for identifying relevant explanatory variables (*i.e.*, crop effect in our case study).
- The sensitivity of the estimated values to (i) the statistical method (frequentist *versus* Bayesian), explanatory variables and residual error distribution and (ii) individual data should be analyzed. Sensitivity analysis is useful for assessing the robustness of the main conclusions of a meta-analysis.
- Efforts should be made to check for the publication bias.
- The design generated by merging data from different studies should be checked for confounding effects (*e.g.*, continent effect in our case study).

5. Conclusions

Finally, the 73 meta-analyses reviewed in this paper focused largely on production aspects. The response variable was crop production in most cases. Our case study showed that agronomical studies would benefit from meta-analysis techniques, because these techniques would make it possible to study the between site-year variability of the response variable (N_2O emission in our case study), to assess uncertainty and to identify the factors potentially influencing environmental impacts. The quality criteria and recommendations presented here could serve as a guide to improve future meta-analyses made in this area.

Acknowledgements

This work was partly funded by the French Research Agency (ANR project ORACLE). We wish to thank P. Rochette for his help on describing the content of the database we used in the case study. We are grateful to J. Sappa for revising the English and the referees for their stimulating comments.

Appendix A. References of the 55 papers reported meta-analyses of published data in agronomy

Ainsworth, E.A., 2008. Rice production in a changing climate: a meta-analysis of responses to elevated carbon dioxide and elevated ozone concentration. Global Change Biol. 14, 1642–1650.

Ainsworth, E.A., Davey, P.A., Bernacchi, C.J., Dermody, O.C., Heaton, E.A., Moore, D.J., Morgan, P.B., Naidu, S.L., Yoo Ra, H., Zhu, X., Curtis, P.S., Long, S.P., 2002. A meta-analysis of elevated [CO_2] effects on soybean (*Glycine max*) physiology, growth and yield. Global Change Biol. 8, 695–709.

Akiyama, H., Yan, X., Yagi, K., 2010. Evaluation of effectiveness of enhanced-efficiency fertilizers as mitigation options for N_2O and NO emissions from agricultural soils: meta-analysis. Global Change Biol. 16, 1837–1846.

Breseghello, F., de Morais, O.P., Pinheiro, P.V., Silva, A.C.S., da Maia de Castro, E., Guimarães, É.P., de Castro, A.P., Pereira, J.A., de Matos Lopes, A., Utumi, M.M., de Oliveira, J.P., 2011. Results of 25 years of upland rice breeding in Brazil. Crop Sci. 51, 914.

Carrubba, A., 2009. Nitrogen fertilisation in coriander (*Coriandrum sativum* L.): a review and meta-analysis. J. Sci. Food Agric. 89, 921–926.

Chivenge, P., Vanlauwe, B., Six, J., 2010. Does the combined application of organic and mineral nutrient sources influence maize productivity? A meta-analysis. Plant Soil. 342, 1–30.

Conant, R.T., Paustian, K., Elliott, E.T., 2001. Grassland management and conversion into grassland: effects on soil carbon. Ecol. Appl. 11, 343–355.

Danielsen, F., Beukema, H., Burgess, N.D., Parish, F., Brühl, C.A., Donald, P.F., Murdiyarso, D., Phalan, B., Reijnders, L., Struebig, M., Fitzherbert, E.B., 2009. Biofuel plantations on forested lands: double jeopardy for biodiversity and climate. Conserv. Biol. 23, 348–358.

Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks—a meta-analysis. Global Change Biol. 17, 1658–1670.

Felton, A., Knight, E., Wood, J., Zammit, C., Lindenmayer, D., 2010. A meta-analysis of fauna and flora species richness and abundance in plantations and pasture lands. Biol. Conserv. 143, 545–554.

Feng, Z., Kobayashi, K., 2009. Assessing the impacts of current and future concentrations of surface ozone on crop yield with meta-analysis. Atmos. Environ. 43, 1510–1519.

Feng, Z., Kobayashi, K., Ainsworth, E.A., 2008. Impact of elevated ozone concentration on growth, physiology, and yield of wheat (*Triticum aestivum* L.): a meta-analysis. Global Change Biol. 14, 2696–2708.

Feng, Z., Wang, S., Szantoi, Z., Chen, S., Wang, X., 2010. Protection of plants from ambient ozone by applications of ethylenediurea (EDU): a meta-analytic review. Environ. Pollut. 158, 3236-3242.

Fletcher, R.J., Robertson, B.A., Evans, J., Doran, P.J., Alavalapati, J.R., Schemske, D.W., 2011. Biodiversity conservation in the era of biofuels: risks and opportunities. Front. Ecol. Environ. 9, 161–168.

Gardner, J.B., Drinkwater, L.E., 2009. The fate of nitrogen in grain cropping systems: a meta-analysis of 15 N field experiments. Ecol. Appl. 19, 2167–2184.

Garratt, M.P.D., Wright, D.J., Leather, S.R., 2011. The effects of farming system and fertilisers on pests and natural enemies: a synthesis of current research. Agric. Ecosyst. Environ. 141, 261–270.

Gobbi, M., Fontaneto, D., 2008. Biodiversity of ground beetles (Coleoptera: Carabidae) in different habitats of the Italian Po lowland. Agric. Ecosyst. Environ. 127, 273–276.

Grassini, P., Hall, A.J., Mercau, J.L., 2009. Benchmarking sunflower productivity in semiarid environments. Field Crop. Res. 110, 251–262.

Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. Global Change Biol. 8, 345–360.

Heaton, E., Voigt, T., Long, S.P., 2004. A quantitative review comparing the yields of two candidate C4 perennial biomass crops in relation to nitrogen, temperature and water. Biomass Bioenergy 27, 21–30.

Kaschuk, G., Alberton, O., Hungria, M., 2010. Quantifying effects of different agricultural land uses on soil microbial biomass and activity in Brazilian biomes: inferences to improve soil quality. Plant Soil. 338, 467–481.

Kiær, L.P., Skovgaard, I.M., Østergård, H., 2009. Grain yield increase in cereal variety mixtures: a meta-analysis of field trials. Field Crop Res. 114, 361–373.

Koellner, T., Scholz, R.W., 2006. Assessment of land use impacts on the natural environment. Int. J. Life Cycle Ass. 13, 32–48.

Letourneau, D.K., Armbrecht, I., Rivera, B.S., Lerma, J.M., Carmona, E.J., Daza, M.C., Escobar, S., Galindo, V., Gutiérrez, C., López, S.D., Mejía, J.L., Rangel, A.M.A., Rangel, J.H., Rivera, L., Saavedra, C.A., Torres, A.M., Trujillo, A.R., 2011. Does plant diversity benefit agroecosystems? A synthetic review. Ecol. Appl. 21, 9–21.

Linquist, B., Groenigen, K.J., Adviento-Borbe, M.A., Pittelkow, C., Kessel, C., 2011. An agronomic assessment of greenhouse gas emissions from major cereal crops. Global Change Biol. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2486.2011.02502.x/abstract (verified 09.09.11).

Lu, M., Zhou, X., Luo, Y., Yang, Y., Fang, C., Chen, J., Li, B., 2011. Minor stimulation of soil carbon storage by nitrogen addition: a meta-analysis. Agric. Ecosyst. Environ. 140, 234–244.

Luo, Z., Wang, E., Sun, O.J., 2010. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. Agric. Ecosyst. Environ. 139, 224–231.

Miguez, F., Bollero, G., 2005. Review of corn yield response under winter cover cropping systems using meta-analytic methods. Crop Sci. 45, 2318–2329.

Miguez, F.E., Villamil, M.B., Long, S.P., Bollero, G.A., 2008. Meta-analysis of the effects of management factors on Miscanthus x giganteus growth and biomass production. Agric. Forest Meteorol. 148, 1280–1292.

Ogle, S.M., Breidt, F.J., Paustian, K., 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry 72, 87–121.

Philpott, S.M., Arendt, W.J., Armbrecht, I., Bichier, P., Diestch, T.V., Gordon, C., Greenberg, R., Perfecto, I., Reynoso-santos, R., Soto-pinto, L., Tejeda-cruz, C., Williams-linera, G., Valenzuela, J., Zolotoff, J.M., 2008. Biodiversity loss in Latin American coffee landscapes: review of the evidence on ants, birds, and trees. Conserv. Biol. 22, 1093–1105.

Prieto-Benítez, S., Méndez, M., 2011. Effects of land management on the abundance and richness of spiders (Araneae): a meta-analysis. Biol. Conserv. 144, 683–691.

Roberts, P.D., Pullin, A.S., 2007. The effectiveness of management interventions used to control ragwort species. Environ. Manage. 39, 691–706.

Rotundo, J.L., Westgate, M.E., 2009. Meta-analysis of environmental effects on soybean seed composition. Field Crop. Res. 110, 147–156.

Rui, W., Zhang, W., 2010. Effect size and duration of recommended management practices on carbon sequestration in paddy field in Yangtze Delta Plain of China: a meta-analysis. Agric. Ecosyst. Environ. 135, 199–205.

Sadras, V.O., 2009. Does partial root-zone drying improve irrigation water productivity in the field? A meta-analysis. Irrigation Sci. 27, 183–190.

Salvagiotti, F., Cassman, K.G., Specht, J.E., Walters, D.T., Weiss, A., Dobermann, A., 2008. Nitrogen uptake, fixation and response to fertilizer N in soybeans: a review. Field Crop Res. 108, 1–13.

Sileshi, G., Akinnifesi, F.K., Ajayi, O.C., Place, F., 2008. Meta-analysis of maize yield response to woody and herbaceous legumes in sub-Saharan Africa. Plant Soil 307, 1–19.

Sileshi, G., Akinnifesi, F.K., Debusho, L.K., Beedy, T., Ajayi, O.C., Mong'omba, S., 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. Field Crop. Res. 116, 1–13.

Sullivan, T.P., Sullivan, D.S., 2003. Vegetation management and ecosystem disturbance: impact of glycophosphate herbicide on plant and animal diversity in terrestrial systems. Environ. Rev. 11, 37–59.

Taub, D.R., Miller, B., Allen, H., 2008. Effects of elevated CO_2 on the protein concentration of food crops: a meta-analysis. Global Change Biol. 14, 565–575.

Thenail, C., Joannon, A., Capitaine, M., Souchère, V., Mignolet, C., Schermann, N., Di Pietro, F., Pons, Y., Gaucherel, C., Viaud, V., Baudry, J., 2009. The contribution of crop-rotation organization in farms to crop-mosaic patterning at local landscape scales. Agric. Ecosyst. Environ. 131, 207–219.

Tirol-Padre A., Ladha, J.K., 2006. Integrating rice and wheat productivity trends using the SAS mixed-procedure and meta-analysis. Field Crop Res. 95, 75–88.

Tonitto, C., David, M., Drinkwater, L., 2006. Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: a meta-analysis of crop yield and N dynamics. Agric. Ecosyst. Environ. 112, 58–72.

Tudoreanu, L., Phillips, C.J., 2004. Empirical models of cadmium accumulation in maize, rye grass and soya bean plants. J. Sci. Food Agric. 84, 845–852.

Turmel, M.-S., Turner, B.L., Whalen, J.K., 2011. Soil fertility and the yield response to the system of rice intensification. Renew. Agric. Food Syst. 26, 185–192.

Unkovich, M.J., Baldock, J., Peoples, M.B., 2010. Prospects and problems of simple linear models for estimating symbiotic N_2 fixation by crop and pasture legumes. Plant Soil 329, 75–89.

Valkama, E., Uusitalo, R., Ylivainio, K., Virkajärvi, P., Turtola, E., 2009. Phosphorus fertilization: a meta-analysis of 80 years of research in Finland. Agric. Ecosyst. Environ. 130, 75–85.

Van den Putte, A., Govers, G., Diels, J., Gillijns, K., Demuzere, M., 2010. Assessing the effect of soil tillage on crop growth: a meta-regression analysis on European crop yields under conservation agriculture. Eur. J. Agron. 33, 231–241.

Vanlauwe, B., Kihara, J., Chivenge, P., Pypers, P., Coe, R., Six, J., 2010. Agronomic use efficiency of N fertilizer in maize-based systems in sub-Saharan Africa within the context of integrated soil fertility management. Plant Soil 339, 35–50.

Veres, A., Petit, S., Conord, C., Lavigne, C., 2011. Does landscape composition affect pest abundance and their control by natural enemies? A review. Agric. Ecosyst. Environ. doi:10.1016/j.agee.2011.05.027.

Veresoglou, S.D., Menexes, G., 2010. Impact of inoculation with Azospirillum spp. on growth properties and seed yield of wheat: a meta-analysis of studies in the ISI Web of Science from 1981 to 2008. Plant Soil 337, 469–480.

Virto, I., Barré, P., Burlot, A., Chenu, C., 2011. Carbon input differences as the main factor explaining the variability in soil organic C storage in no-tilled compared to inversion tilled agrosystems. Biogeochemistry. Available at http://www.springerlink.com/content/tmh62h277565022r/ (verified 09.09.2011).

Wang, S., 2011. Management and land use change effects on soil carbon in northern China's grasslands: a synthesis. Agric. Ecosyst. Environ. 142, 329–340.

Yahdjian, L., Gherardi, L., Sala, O.E., 2011. Nitrogen limitation in arid-subhumid ecosystems: a meta-analysis of fertilization studies. J. Arid Environ. 75, 675–680.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–723.
Arnqvist, G., Wooster, D., 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends Ecol. Evol. 10, 236–240.
Barbottin, A., Makowski, D., Le Bail, M., Jeuffroy, M., Bouchard, C., Barrier, C., 2008. Comparison of models and indicators for categorizing soft wheat fields according to their grain protein contents. Eur. J. Agron. 29, 175–183.
Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K.F., Simel, D., Stroup, D.F., 1996. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. J. Am. Med. Assoc. 276, 637–639.
Bland, C.J., Meurer, L.N., Maldonado, G., 1995. A systematic approach to conducting a non-statistical meta-analysis of research literature. Acad. Med. 70, 642–653.
Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. Introduction to Meta-Analysis, 1st ed. Wiley, Hoboken, NJ.
Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y., Cellier, P., Debaeke, P., Gaudillère, J., Hénault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309–332.
Burnham, K., Anderson, D., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoric Approach, 2nd ed. Springer-Verlag, New York.
Cardinale, B.J., Srivastava, D.S., Duffy, J.E., Wright, J.P., Downing, A.L., Sankaran, M., Jouseau, C., 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. Nature 443, 989–992.
Cucherat, M., Boissel, J.P., Leizorovicz, A., 1997. Manuel pratique de méta-analyse des essais thérapeutiques, Available at http://www.spc.univ-lyon1.fr/livreMA/frame.htm.
Doré, T., Makowski, D., Malézieux, E., Munier-Jolain, N., Tchamitchian, M., Tittonell, P., 2011. Facing up to the paradigm of ecological intensification in agronomy: revisiting methods, concepts and knowledge. Eur. J. Agron. 34, 197–210.
Evans, J.A., Foster, J.G., 2011. Metaknowledge. Science 331, 721–725.
Furlan, A.D., Pennick, V., Bombardier, C., van Tulder, M., 2009. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine 34, 1929–1941.
Gates, S., 2002. Review of methodology of quantitative reviews using meta-analysis in ecology. J. Anim. Ecol. 71, 547–557.
Glass, G.V., 1976. Primary, secondary, and meta-analysis of research. Educ. Res. 5, 3–8.
Gurevitch, J., Hedges, L., 1993. Meta-analysis: combining the results of independent experiments. In: Scheiner, S.M., Gurevitch, J. (Eds.), Design and Analysis of Ecological Experiments. Chapman and Hall, pp. 378–398.
Hedges, L., Olkin, I., 1985. Statistical Methods for Meta-analysis. Academic Press, New York.
IPCC, 2007. In: Solomon, S., et al. (Eds.), Climate Change 2007: The Physical Science Basis. Cambridge University Press, Cambridge.
Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265.
Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288.
Light, R., Pillemer, D., 1984. Summing Up: The Science of Reviewing Research. Harvard University Press, Cambridge.
Madden, L.V., Paul, P.A., 2011. Meta-analysis for evidence synthesis in plant pathology: an overview. Phytopathology 101, 16–30.
Makowski, D., Tichit, M., Guichard, L., Van Keulen, H., Beaudoin, N., 2009. Measuring the accuracy of agro-environmental indicators. J. Environ. Manage. 90, S139–S146.
Moher, D., Cook, D.J., Jadad, A.R., Tugwell, P., Moher, M., Jones, A., Pham, B., Klassen, T.P., 1999. Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. Health Technol. Asses. 3, 1–98.
Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. In: Wallach, D., Makowski, D., Jones, J. (Eds.), Working with Dynamic Crop Models. Elsevier, pp. 55–100.
Myers, R.A., 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES J. Mar. Sci. 58, 937–951.
Normand, S.L., 1999. Meta-analysis: formulating, evaluating, combining, and reporting. Stat. Med. 18, 321–359.
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. Biol. Conserv. 132, 409–423.
Robson, M., Fowler, S., Lampkin, N., Leifert, C., Leitch, M., Robinson, D., Watson, C., Litterick, A., 2002. The agronomic and economic potential of break crops for ley/arable rotations in temperate organic agriculture. Adv. Agron. 77, 369–427.
Rochette, P., Janzen, H.H., 2005. Towards a revised coefficient for estimating N_2O emissions from legumes. Nutr. Cycl. Agroecosyst. 73, 171–179.
Rosenberg, M.S., Garrett, K.A., Su, Z., Bowden, R.L., 2004. Meta-analysis in plant pathology: synthesizing research results. Phytopathology 94, 1013–1017.
Sauvant, D., Schmidely, P., Daudin, J.J., St-Pierre, N.R., 2008. Meta-analyses of experimental data in animal nutrition. Animal 2, 1203–1214.
Smith, T., Spiegelhalter, D., Thomas, A., 1995. Bayesian approaches to random-effects meta-analysis: a comparative study. Stat. Med. 14, 2685–2699.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. B 64, 583–639.
Stewart, G., 2010. Meta-analysis in applied ecology. Biol. Lett. 6, 78–81.
Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. Eur. J. Agron. 18, 289–307.
Sutton, A., Abrams, K., Jones, D., Sheldon, T., Song, F., 2000. Methods for Meta-analysis in Medical Research. Wiley, Chichester.
Sutton, A.J., Abrams, K.R., 2001. Bayesian methods in meta-analysis and evidence synthesis. Stat. Methods Med. Res. 10, 277–303.
van Ittersum, M., Leffelaar, P., van Keulen, H., Kropff, M., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. Eur. J. Agron. 18, 201–234.
van Tulder, M.W., Assendelft, W.J., Koes, B.W., Bouter, L.M., 1997. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. Spine 22, 2323–2330.
van Tulder, M., Furlan, A., Bombardier, C., Bouter, L., 2003. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine 28, 1290–1299.
Yuan, Y., Hunt, R.H., 2009. Systematic reviews: the good, the bad, and the ugly. Am. J. Gastroenterol. 104, 1086–1092.