

Meta-Analysis 及其在农业水土工程中的应用前景

谢 岷^{1,2} 魏占民¹ 屈忠义¹ 张永平³

(1. 内蒙古农业大学水利与土木建筑工程学院, 内蒙古 呼和浩特 010018;

2. 加拿大农业食品部园艺发展与研究中心, J3B 3E6; 3. 内蒙古农业大学农学院, 内蒙古 呼和浩特 010019)

摘要 介绍 Meta-Analysis 方法的发展和现状, 重点介绍 Meta-Analysis 研究方法和特点。在总结 Meta-Analysis 方法成功应用实例的基础上, 探讨 Meta-Analysis 方法在农业水土工程学科中的应用前景。认为 Meta-Analysis 方法在农业水土工程学科中的深入应用将会显著提高试验数据的处理水平, 使得定量地对以往数据做出综述分析成为可能。

关键词 Meta-Analysis; 效应值; 农业水土工程

中图分类号 S27 **文献标识码** A **文章编号** 1006-764X(2011)S1-0095-04

当前, 由于科学技术手段的突飞猛进, 各科学研究领域每年均有大量的试验结果被报道, 除了一些新研究与新发现之外, 有相当部分的试验设计与方法都存在着不同程度的重复, 同时, 纵览各学科的研究, 多年来均有大量的文献资料积累, 所以怎样从海量的试验结果和资料文献中找出它们之间的相互联系、进行彼此综合、提取、比较和分析以往的研究结论, 得出一个理性且严谨的研究综述, 已成为各学科研究中必不可少的关键环节。传统的综述方法由于主观性较强、局限性较大, 易受综述者人为因素的影响和缺乏理论数据支持等, 往往很难做到真正的客观、全面和有理有据。因此, Meta-Analysis 作为一种可以对同一主题下不同研究成果进行整理、综合的统计推断方法被广泛应用于各个学科领域研究综述中。如在农学中, 土壤条件、气候特性等影响因子对作物产量影响效应的研究; 在生态学中, 用于各个生理生态指标对大气 CO₂ 浓度倍增的相应研究; 在医药学中, 医药企业通过对药品药理效果的 Meta-Analysis 而获得药效的评定, 等等。

Meta-Analysis 顾名思义即是对诸多独立研究的结果进行综合、分析的统计方法。1993 年首次被赵宁等^[1]介绍到国内的医学界, 1994 年徐勇^[2]介绍了 Meta-Analysis 所需的资料类型及其统计方法, 之后彭少麟等^[3]把 Meta-Analysis 引入到生态学界并翻译为整合分析。近年来 Meta-Analysis 因其客观、严谨、

综合性强等突出优点而备受各方关注, 成为目前各学科用于综述性研究最热门的定量统计方法。但与此同时, 亦有相当一些研究表明在收集 Meta-Analysis 原数据的过程中, 在出现数据缺失、发表偏见、索引论文数据的质量不高、观测值缺乏独立性和整合不同类型的数据集合等情况下, 运用 Meta-Analysis 将会影响到最终结果的可靠性并产生一定程度的误差^[4]。本文将对 Meta-Analysis 做一个简要的介绍, 以此来展开 Meta-Analysis 在农业水土工程学科应用的探讨与展望。

1 Meta-Analysis 简介

Meta-Analysis 是一种在同一主题下从多个研究成果(包括未发表的成果)中整理、合并研究数据的统计分析方法, 当各项研究在时间或范围上有一定的连续性和相关性时, 它可以用来识别诸多研究的共同成果, 从这些研究成果中发现一些变化的原因和潜在的规律, 从而对以往研究成果做一个定量综述, 也可以为今后的研究提供一定程度的启示和指导。

1.1 Meta-Analysis 的发展历程与现状

1.1.1 萌芽阶段

现代 Meta-Analysis 思想的起源最早可追溯到 20 世纪 30 年代, 由 Cochran^[5], Fisher^[6], Pearson^[7]和 Tippet^[8]的统计显著性联合检验理论发展而来^[4], 也有学者认为现代 Meta-Analysis 在 1904 年 Pearson

的《Report on certain enteric fever inoculation statistics》一文中已见雏形^[9]。在 1952 年,Eysenck^[10]断定 Meta-Analysis 的思想在心理治疗方法研究上没有任何的应用价值。由此,对于 Meta-Analysis 是否具有应用价值的争论进入了一个白热化的阶段。在此期间,比较有影响的研究学者有^[11-13]:Brozek,Tiede,Sakoda,Cohen,Beall,Sterling 等。

1.1.2 发展阶段

尽管 Meta-Analysis 的思想在 20 世纪初就已经被有关学者发现,但其发展却是在 70 年代左右。1976 年,Glass 首次正式提出把这种新的统计方法命名为 Meta-Analysis。此后,Glass 和 Smith 有大量关于 Meta-Analysis 的文章问世,并均取得了满意的效果^[14],自此也结束了学术界长达 20 多年对 Meta-Analysis 是否具有应用价值的大争论。随后的十几年里,Meta-Analysis 方法在社会学^[15]、心理学^[14]和医学领域^[16]得到了快速的发展,并有大量的文章发表,这段时间 Meta-Analysis 得到了快速的发展,主要体现在以下 2 个方面:①Glass 提出用效应值替代标准均值差,并将 p 值和投票表法转化为参数和模型以便于研究和分析;②Hedges 等^[17]在 1985 年提出把固定效应模型与随机效应模型运用到 Meta-Analysis 中,2 种模型对平均效应值做显著性测试和置信度区间评估,之后又相继把方差权重引入到 Meta-Analysis 中,使 Meta-Analysis 的结果更加准确,符合实际。

1.1.3 完善与应用阶段

从 20 世纪 90 年代初开始,Meta-Analysis 的思想逐渐地被各学科领域关注,进入了高速发展和全面应用的阶段。除了之前被广泛运用于医学和社会科学研究中,1991 年生态学家 Järvinen 成功地把 Meta-Analysis 引入到生态学中,自此拉开了 Meta-Analysis 向农业相关学科渗入的序幕。截至 2009 年 6 月 18 日,从 Scopus 上键入 Meta-Analysis 为关键词检索到的相关文章共计 50 487 篇,其中关于医学和药理学的最多,分别达到 40 188 篇和 5 463 篇;关于环境学、农学与生物学、工程学的分别是 1 356 篇、1 217 篇和 211 篇。从图 1 中可以看出,从 20 世纪 90 年代初期开始,Meta-Analysis 进入了一个全新的发展时期,尤其在近 10 年,得到了各学科的广泛应用。从图 2 也可以看出,自 20 世纪 90 年代中期开始,Meta-Analysis 在农学和生物学中的应用迅速增长,仅 2008 年,从 Scopus 可以检索到已发表的文章已达 166 篇。尽管在 20 世纪 90 年代初 Meta-Analysis 已被引入到国内,但其应用和发展只限于医药学、心理学、社会学和生态学几个领域,遗憾的是目前国内还没有发现 Meta-

Analysis 在农业工程学科方面的应用。笔者认为 Meta-Analysis 在农业水土工程学科中应用前景广泛,应该尽早地把它运用到农业水土工程的研究中来。

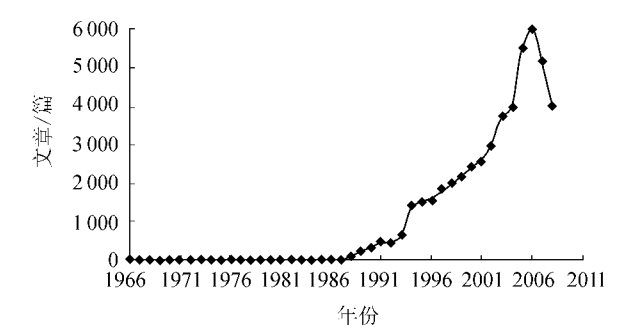


图 1 以 Meta-Analysis 为检索词从 Scopus 搜索到的所有文章数

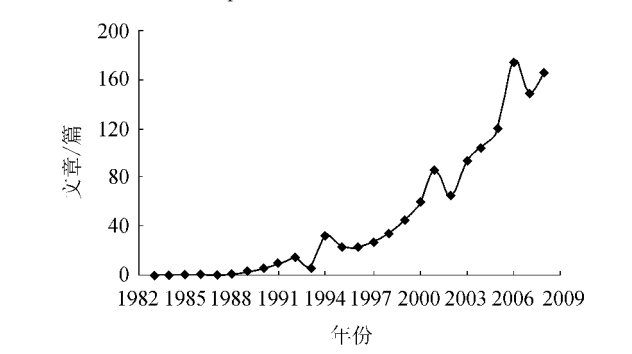


图 2 在 Scopus 中农学和生物学目录下搜索到的有关 Meta-Analysis 的文章数

1.2 Meta-Analysis 研究方法

实际上,Meta-Analysis 本身即是一项统计学研究,它经过几十年的发展已有了固定的分析过程与方法。现以文献[18]为例,将 Meta-Analysis 的具体方法和步骤介绍如下。

1.2.1 确定研究方向和目标

由于 Meta-Analysis 本身就是一项独立的研究,所以在对某项科学研究进行 Meta-Analysis 之前设计一个合理、可行的 Meta-Analysis 方案和步骤至关重要,它直接关系到最终的 Meta-Analysis 结果以及对研究的综述和评价。所以,要进行一个 Meta-Analysis,分析者要对所要整合的研究有一定程度的了解,知道此研究当前的发展情况和研究历史,并以此来确定研究的具体方向和 Meta-Analysis 的目标。文献[18]的作者以整合磷肥施用对作物的影响为研究方向,从芬兰 80 年间在作物上施用磷肥的试验中找出磷肥、产量和土壤结构与质地之间的关系。

1.2.2 制定一个合适的标准和研究范围

在明确 Meta-Analysis 要研究的方向后,对所要进行的 Meta-Analysis 制定一个合适的研究范围和标准是 Meta-Analysis 成功地必要基础。合适的标准包括文献搜索的方向和标准、搜索数据库时遇到问题

的处理标准、所搜集到文献的偏见的处理标准与对以前地研究质量的评判标准等等。研究范围的选择和制定通常包括所研究试验设计的选择、所搜集文献发表年份的选择、文献的语言和出处的选择和样本数量的选择等等。

1.2.3 数据的收集和提炼

在确定研究方向和制定合适的标准之后,确定所要搜集文献的数据库,键入检索词,开始搜索文献。在数据的收集和提炼过程中需要注意几个方面:①收集地文献资料数据库是否权威、全面、便于搜索;②区别相互联系的文献(比如同样的原始数据从不同角度来分析的情况),选出更适合自己的研究的数据源;③如果部分文献并没有提供全部相关数据,可通过联系作者索取所需要的数据;④对于一些结论模糊和存在偏见的文献(比如有显著差异或显著区别的文献更易被发表和收录)要慎重对待,从科学的角度来决定取舍;⑤建议对收集到的所有文献做一个质量评判,以保证文献数据库的质量(比如可以建立一个效应值散点图,以效应值为 x 轴、以样本单位为 y 轴作一个分布图,如果图形显示是正态分布,则说明文献发表偏见最小)。

1.2.4 选定分析方法和模型,作统计学分析

在准备好上述几步之后,即可开始进行 Meta-Analysis。具体步骤如下:①确定效应值的估计值;②为研究选定一个合适的效应值;③选定一个分析计算的模型;④根据资料数据库所提供的对照组和研究组的均值、标准差、样本重复数等信息计算效应值,然后进行均质性分析;⑤分析 Meta-Analysis 结果,总结报告。现今,由于 Meta-Analysis 的迅速发展,多种针对 Meta-Analysis 计算的应用软件相应产生。如 Comprehensive Meta-Analysis, Meta Analyst, Meta Win 2.0 等等,一定程度上简化精确了 Meta-Analysis 的计算步骤与过程。

2 Meta-Analysis 的特点

Meta-Analysis 与其他传统的综合方法相比有其固有的特点,且这些特点突出,这也是 Meta-Analysis 近年来成为各学科和各研究领域综述分析热点和焦点的根本原因。

2.1 效应值

效应值的引入标志着现代 Meta-Analysis 的诞生,它的出现使得把各个单独研究的结果放在同一水平比较成为可能。现介绍 2 种在农学和生态学 Meta-Analysis 中最常用的效应值,见表 1。

在选取效应值的过程中,有 3 个标准应该作为选用效应值的主要依据。第一,来源于不同研究中

表 1 2 种常用的效应值

名称	符号	公式	方差
Hedges'd	D	$d = \frac{(\bar{X}_E - \bar{X}_C)J}{S}$	$V_d = \frac{N_C + N_E}{N_C N_E} + \frac{d^2}{2(N_C + N_E)}$
Response ratio	R	$R = \frac{\bar{X}_E}{\bar{X}_C} \ln R = \ln\left(\frac{\bar{X}_E}{\bar{X}_C}\right)$	$V_{\ln R} = S_{\text{pooled}}^2 \left(\frac{1}{N_E \bar{X}_E^2} + \frac{1}{N_C \bar{X}_C^2} \right)$

注: $\bar{X}_E$ 和 $\bar{X}_C$ 分别是各独立研究试验组和对照组的平均值, S_E 和 S_C 分别是试验组和对照组的标准差, N_E 和 N_C 分别为试验组和对照组的样本数, S 为混合方差, $S_{\text{pooled}} = \sqrt{\frac{(N_E - 1)S_E^2 + (N_C - 1)S_C^2}{N_E + N_C - 2}}$, J 为取样误差的修正系数, $J = 1 - \frac{3}{4(N_C + N_E - 2) - 1}$ 。

的效应值必须是指针对同一参数和指标的研究,而不是依据试验设计中的样本数、协方差等因素来确定效应值,否则效应值的比较就没有科学意义。比如:在文献 [18] 中,作者选取 R_R (Response Ratio) 为效应值,选定无磷处理为对照处理,所有的施磷处理为试验处理,对照处理的作物产量即为 $\bar{X}_E$, 试验处理的产量即为 $\bar{X}_C$; 第二,计算效应值所需的各数值应该在相关文献中被直接报道,一般不需要重复计算原始试验数据推导计算效应值所需的各个数值,除非有特殊需要;第三,要根据原始资料库中的数据情况选取那些更容易被计算的效应值,比如原始资料中给出数据取样分布情况,效应值中的方差和置信区间可以被直接计算 [19]。

2.2 模型的选择

固定效应模型和随机效应模型在大多数 Meta-Analysis 中被用来计算合并效应值。其中,当所有的研究都共享一个相同的真实效应值,且影响效应值的所有因素都相同时,固定效应模型将被选择来计算合并效应值。例如:在对照和处理试验中,如果所有的气候、土壤、作物、田间管理因素相同,即可认为所有的试验共享一个相同的真实效应值。与此相反,当所有的研究不能共享同一个效应值,而只能共享一个效应值的平均值时,随机效应模型将被选择来计算合并效应值。也就是说,随机效应模型比固定效应模型更加平衡和具有更大的标准差和置信区间。

2.3 均质性分析

均质性分析是对所有效应值的一种假设。在 Meta-Analysis 中,如果对各效应值的均质性假设不成立,则说明各效应值具有异质性,异质性是描述效应值变异程度的一个重要指标。在 Meta-Analysis 中,如果是均质性研究,随机效应模型和固定效应模型没有区别;如果研究是异质性的,则随机效应模型在计算上要比固定效应模型更严谨。在大多数 Meta-Analysis 研究中,异质性又被分解为总异质性、组内异质性和组间异质性。

3 Meta-Analysis 在农业水土工程中的应用前景

3.1 在农业水土环境的监测和水土资源可持续利用方面

农田水利工程与措施对农业区域环境效应的影响一直是本学科研究的重点。但大多数综述性研究由于研究方法的局限,基本都集中在单个因素或个别几个因素对农业水土环境的影响和监测上,很难把长时间跨度内的诸多因素放到同一水平下进行评估。利用 Meta-Analysis 的特点,便可以对各种农田水利工程与措施对水土环境的影响做出更为合理地预测、分析和评估。比如,可以利用 Meta-Analysis 研究节水灌溉对农田生态环境各因素的影响、对地表水与地下水水文循环的影响、对土壤盐渍化以及地下水三氮污染的影响等等。

3.2 在农业水的高效利用方面

近年来,非充分灌溉和节水灌溉一直是农业各相关专业的热点研究领域,各学科均投入了大量的人力、物力和财力,并得到大量地试验数据和统计结果。但现实情况是各试验数据和结论均相互独立,缺乏有效的整合和系统的综述。利用 Meta-Analysis,可以在灌溉量、施肥量、产量等几个关键词下进行集中,对各种非充分灌溉与节水灌溉的措施和制度进行综述、分析和评估,在一定程度上为提高农业水的利用效率给予指导。

3.3 在农业水土工程学中的应用实例

Sadras^[20]应用 Meta-Analysis 对局部根区控制灌溉对田间灌溉水生产率进行研究,对比了局部根区控制灌溉与非充分灌溉两者在产量、田间灌溉水生产率和水分利用率方面的区别,建议在实践中从经济效益和灌溉管理的可行性 2 方面来慎重选择 2 种灌溉方式,从而为灌区灌溉制度的制定与区域环境的规划和管理提供更有效的参考和决策依据。

4 结 语

Meta-Analysis 具有坚实的统计学基础,可以对同一主题下的大量试验数据进行综述分析。虽然该方法本身也有一些缺陷,而且不适合所有领域的科学研究,但把 Meta-Analysis 引入到农业水土工程学科的研究中,可以对过去的试验进行总结和提炼,找出诸多试验中的一些本质的规律,排除个别试验中产生的随机误差,为将来的研究指明方向,不失为一种实用而有效地试验统计与综合方法;在农业水土工程学科中应用 Meta-Analysis 亦可以在当前的试验数据分析和整理上做出一定地创新与尝试,尽早地与国际研究接轨,提高当前的学科研究水平。

参考文献:

- [1] 赵宁,俞顺章. Meta-analysis: 一种新的定量综合方法[J]. 中国慢性病预防与控制, 1993, 1(6): 277-281.
- [2] 徐勇. Meta 分析常见资料类型及统计分析方法[J]. 中华预防医学杂志, 1994, 28(3): 303-307.
- [3] 彭少麟,唐小焱. Meta 分析及其在生态学上的应用[J]. 生态学杂志, 1998, 17(5): 74-79.
- [4] GUREVITCH J, CURTIS P S, JONES M H. Meta-analysis in ecology[J]. Advances in Ecological Research, 2001, 32: 199-247.
- [5] COCHRAN W G. Problems arising in the analysis of a series of similar experiments[J]. Roy Stat Soc, 1937, 4(suppl): 102-118.
- [6] FISHER R A. Statistical methods for research workers[M]. London: Oliver and Boyd, 1932.
- [7] PEARSON K. On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random[J]. Obst. Gynecol, 1933, 89: 628-633.
- [8] TIPPETT L H C. The methods of statistics[M]. London: Williams & Norgate, 1931.
- [9] 彭少麟,郑凤英. Meta 分析: 综述中的一次大革命[J]. 生态学杂志, 1999, 18(6): 65-70.
- [10] EYSENCK H J. The effects of psychotherapy: An evaluation[J]. Journal of Consulting Psychology, 1952, 16: 319-324.
- [11] BROZEK J, TIEDE K. Reliable and questionable significance of combined results[J]. Psychol Bull, 1952, 49: 339-341.
- [12] SAKODA J M, COHEN B H, BEALL F. Tests of significance for a series of statistical tests[J]. Psychol Bull, 1954, 51: 172-175.
- [13] STERLING T C. Publication decisions and their possible effects of inference drawn from tests of significance or vice versa[J]. Am Stat Assoc J, 1959, 54: 30-34.
- [14] SMITH M L, GLASS G V. Meta-analysis of psychotherapy outcome studies[J]. American Psychologist, 1977, 32(9): 752-760.
- [15] DROR Y. Proposed policymaking scheme for the Knesset Committee for the Examination of the Structure of Elementary and Post-Elementary Education in Israel-an illustration of a policy analysis memorandum[J]. Socio-Economic Planning Sciences, 1969, 3(1): 13-24.
- [16] KAZRIN A, DURAC J, AGTEROS T. Meta-meta analysis: A new method for evaluating therapy outcome[J]. Behaviour Research and Therapy, 1979, 17(4): 397-399.
- [17] HEDGES L V, OLKIN I. Statistical methods for Meta-analysis[M]. Orlando, FL: Academic Press, 1985.
- [18] VALKAMA E. Phosphorus fertilization: a meta-analysis of 80 years of research in Finland[J]. Agriculture, Ecosystems and Environment, 2009, 130(3/4): 75-85.
- [19] BORENSTEIN M, HEDGES L V. Introduction to Meta-Analysis[M]. New York: John Wiley and Sons, 2009.
- [20] SADRAS V O. Does partial root-zone drying improve irrigation water productivity in the field? A meta-analysis[J]. Irrigation Science, 2009, 27(3): 183-190.

(收稿日期: 2011-03-25 编辑: 高建群)