

指示值方法及其在昆虫中的应用*

李 巧**

(西南林业大学云南省森林灾害预警与控制重点实验室 昆明 650224)

摘 要 指示值(IndVal)方法于1997年由Dufrêne和Legendre提出,是生物指示方法论研究中最有意义的进步,在国外被广泛运用于生态指示物研究,然而在国内却没有得到应有的重视。本文介绍了如何运用R语言软件进行IndVal值的计算,IndVal方法在指示昆虫研究中的应用,以及运用该方法进行昆虫指示物种的判断和检验。旨在普及该方法的运用,推动我国的生物指示研究。

关键词 方法论,指示生物,生态指示,生物指示,R语言软件

Indicator Value (IndVal) method and its application

LI Qiao**

(Key Laboratory of Forest Disaster Warning and Control in Yunnan Province, Southwest Forestry University, Kunming 650224, China)

Abstract The Indicator Value (IndVal) method, developed by Dufrêne and Legendre in 1997, is the most significant methodological advance in bioindication but is not commonly used in China. This paper describes how to calculate Indicator Values with R language software, the application of the IndVal method to insect indicators, and how to identify and test bioindicators with the IndVal method.

Key words methodology, bioindicators, ecological indication, bioindication, R language software

生物指示(bioindication or bio-indication)研究在本质上是保护生物学的分支学科,旨在建立用于从生态系统中提取生物信息并利用这些信息进行科学性管理决策的生物指示系统(bioindication system),将科学知识应用于生态关系的管理中(Caughley and Gunn, 1996; van Straalen and Krivolutsky, 1996; McGeoch, 2007)。指示值(IndVal)方法于1997年由Dufrêne和Legendre提出,它将一个物种对某一生态状态的专一性程度的测量与它对该状态的忠实性测量结合起来,进行指示生物价值的计算,在环境状态内具有高的专一性(specifity)和高的忠实性(fidelity)的物种将具有对该状态的高指示价值。该方法一经提出,就被广泛运用,是生物指示方法论研究中最具影响力的方法,是生态指示研究的一个特别高效的工具(McGeoch *et al.*, 2002; McGeoch, 2007)。然而在国内仅见于极少数研究(刘新民和杨劫, 2005),没有得到应有的重视。

本文介绍了如何运用R语言软件(R Development Core Team, 2009)对调查数据进行分析,计算物种的IndVal值,并通过IndVal值进行指示物种的选择。以期在国内普及该方法的运用,推动我国的昆虫指示研究。

1 IndVal值及其分析软件

IndVal值根据以下公式进行计算(Dufrêne and Legendre, 1997):

$$A_{ij} = N_{individuals_{ij}} / N_{individuals_i}$$

$$B_{ij} = N_{sites_{ij}} / N_{sites_j}$$

$$IndVal_{ij} = A_{ij} \times B_{ij} \times 100$$

式中: A_{ij} 反映了物种*i*对样地组*j*的专一性,是物种*i*在样地组*j*的平均多度(个体数)与该物种在所有样地组的平均多度之和的比值; $N_{individuals_{ij}}$ 表示物种*i*在样地组*j*的平均个体数,而 $N_{individuals_i}$ 是物种*i*在所有样地组的平均个体数之和。 B_{ij} 反映了物种*i*对样地组*j*的忠实性,是

* 资助项目:西南林业大学重点科研基金项目和动物学重点学科(XKX200903)。

** 通讯作者, E-mail: lqfcb@126.com

收稿日期:2010-04-06, 接受日期:2010-09-16

样地组 j 中存在物种 i 的样地数与该样地组的所有样地数的比值; $N_{sites_{ij}}$ 表示样地组 j 中存在物种 i 的样地数, N_{sites_j} 表示样地组 j 中的所有样地数。

在使用中,该公式常被简化为:

$$IndVal_{ij} = A_{ij} \times B_{ij}$$

IndVal 值的计算可以通过 IndVal 2.0、PC-ORD、R 语言等多个统计软件实现。其中,在免费统计软件 R 语言环境中运用“labdsv”程序包的“duleg”函数,能够迅速便捷的完成物种的 IndVal 值计算。R 语言及“labdsv”程序包均可以从相关网站上免费下载 (Roberts, 2007; R Development Core Team, 2009)。下载并安装 R 语言及相关程序包,方法见文献 (Legendre, 2009a; 2009b) 及丁国徽译的《R 导论》(2005)。

2 IndVal 值计算

假设对 3 种类型的样地 (plot 1-12 分别属于 3 个样地组,其中 plot 1-5、6-8、9-12 分别属于样地组 1、2、3) 进行多样性调查,得到 95 个物种。通过以下两个步骤完成数据文本的准备。

第 1 步:通常以样本为行、物种为列在 Excel 工作表上进行数据统计,如图 1 (sheet 1),它表示 12 个样地中 95 个物种的多度 (个体数)。第 1 行为物种名或物种编号 (本例中为物种编号),第 1 列为样地编号。第 2 步:将 sheet 1 中除第 1 列外的所有内容复制到一张新工作表,将该工作表以 CSV (逗号分割) 文档形式保存于 R 语言的工作目录下。本例中 CSV 文档的文件名为“antlv3” (图 2),从第 1 列可以看出,第 1 个物种在样地 2 和 12 中的多度分别是 3 和 22,其余样地中均未发现该物种。

数据文本准备好后启动 R 语言程序,运用以下的 6 个指令进行 IndVal 值的计算:

指令 1:读取数据 (图 3);

```
> antlv3 <- read.csv("antlv3.csv")
```

指令 2:调用“labdsv”程序包 (程辑包) (图 3);

```
> library(labdsv)
```

指令 3:对生境类型进行分类,输入分类结果 (图 3);

```
> class <- c(1, 1, 1, 1, 1, 2, 2, 2, 3, 3, 3, 3)
```

指令 4:调用“labdsv”程序包中的“duleg”进行计算 (图 3);

```
> duleg(antlv3, class, numitr = 1000)
```

所有物种的 IndVal 值及相关结果将计算并显示出来 (图 4)。

指令 5:筛选具有统计学差异的物种 (图 5);

```
> result <- duleg(antlv3, class, numitr = 1000)
```

指令 6:显示结果 (图 5)。

```
> summary(result)
```

3 个样地组中具有统计学显著差异的 18 个物种被筛选出来,见图 5。

3 指示物种选择

通过以上指令得到的 18 个物种并不都是指示物种,一般选择指示值大于 0.7 的物种作为指示物种 (McGeoch *et al.*, 2002; Nakamura *et al.*, 2007)。本例中,根据数据计算结果,样地组 1 有 5 个指示物种,按指示值从大到小排列分别是: sp. 90、sp. 23、sp. 47、sp. 26 和 sp. 89; 样地组 2 有 4 个指示物种,分别是: sp. 10、sp. 16、sp. 44 和 sp. 68; 样地组 3 仅有 1 个指示物种,为 sp. 71。根据物种的鉴定结果可以明确物种的学名。

4 IndVal 方法的应用

Hiddink (2005) 认为指示生物的多度可能因季节及分布因素的影响而发生变化,与环境参数之间缺乏简单直接的关系,这种指示生物忠实性测量中的不确定因素可能导致在使用 IndVal 中出现问题。但该方法仍被认为优于其它用于生态指示的指示物测量方法 (McGeoch, 2007)。Nakamura 等 (2007) 在以蚂蚁为指示生物揭示澳大利亚亚热带雨林开垦和土地利用的影响研究中发现,由于大多数蚂蚁种类表现出块状分布,因而不能可靠用于生态指示;为克服这一问题,他们提出了复合指数,将同一类群的指示物种进行合并,在更高分类单位上进行指示值分析。他们的研究将 IndVal 方法的使用由种级分类单位扩展到属级、科级,由单个物种 IndVal 值的计算扩展到指示类群的指示值的计算。De Cáceres 和 Legendre (2009) 比较了相关性分析方法和指示值方法在物种-样地组关系分析上的差异,指出生态学家应根据研究目的选择最适合的关联强度指数,并且计算置信区间以确定估计精度。该研究强调了运用指示值方法时需注意的问题,进一步完善了指示值方法。

	sp.1	sp.2	sp.3	sp.4	sp.5	sp.6	sp.7	sp.8	sp.9	sp.10	sp.11	sp.12
plot 1	0	2	0	0	0	0	0	0	0	0	0	0
plot 2	3	1	0	1	0	0	0	0	0	0	0	0
plot 3	0	0	9	0	0	0	0	0	0	0	0	0
plot 4	0	0	1	0	0	0	0	0	0	0	0	0
plot 5	0	0	0	0	0	0	0	1	0	1	0	0
plot 6	0	0	0	2	0	1	0	0	0	1	0	0
plot 7	0	0	0	0	0	0	0	0	0	1	0	0
plot 8	0	0	66	0	0	0	0	4	0	1	0	0
plot 9	0	0	0	0	0	0	0	0	0	0	0	0
plot 10	0	0	0	0	0	0	1	0	0	0	0	0
plot 11	0	0	0	0	0	0	0	0	0	0	0	1
plot 12	22	0	0	1	1	0	0	0	1	0	0	0

图 1 Excel 工作表数据示范 1

Fig. 1 Data example 1 in an Excel table

	sp.1	sp.2	sp.3	sp.4	sp.5	sp.6	sp.7	sp.8	sp.9	sp.10	sp.11	sp.12	sp.13
plot 1	0	2	0	0	0	0	0	0	0	0	0	0	0
plot 2	3	1	0	1	0	0	0	0	0	0	0	0	2
plot 3	0	0	9	0	0	0	0	0	0	0	0	0	0
plot 4	0	0	1	0	0	0	0	0	0	0	0	0	0
plot 5	0	0	0	0	0	0	0	0	0	1	0	1	0
plot 6	0	0	0	0	0	0	0	0	0	1	0	1	0
plot 7	0	0	0	2	0	1	0	0	0	1	0	0	0
plot 8	0	0	0	0	0	0	0	0	0	1	0	0	0
plot 9	0	0	66	0	0	0	0	4	0	1	0	0	0
plot 10	0	0	0	0	0	0	0	0	0	0	0	0	0
plot 11	0	0	0	0	0	0	1	0	0	0	0	0	0
plot 12	0	0	0	0	0	0	0	0	0	0	0	0	1
plot 13	22	0	0	1	1	0	0	0	0	1	0	0	0

图 2 Excel 工作表 (CSV 格式) 数据示范 2

Fig. 2 Data example 2 in an Excel table (CSV format)

```

RGui - [R Console]
文件 编辑 查看 其他 程序包 窗口 帮助

R version 2.9.2 (2009-08-24)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R是自由软件软件，不带任何担保。
在某些条件下你可以将其自由散布。
用'license()'或'licence()'来看散布的详细条件。

R是个合作计划，有许多人为之做出了贡献。
用'contributors()'来看合作者的详细情况
用'citation()'会告诉你如何在出版物中正确地引用R或R程序包。

用'demo()'来看一些示范程序，用'help()'来阅读在线帮助文件，或
用'help.start()'通过HTML浏览器来看帮助文件。
用'q()'退出R。

[原来保存的工作空间已还原]

> antlv3<-read.csv("antlv3.csv")
> library(labdsv)
载入需要的程辑包: mgcv
This is mgcv 1.5-5 . For overview type 'help("mgcv-package")'.
载入需要的程辑包: MASS

载入程辑包: 'labdsv'

The following object(s) are masked from package:stats :

density

> class<-c(1,1,1,1,1,2,2,2,3,3,3)
> duleg(antlv3,class,numitr=1000)
$relfrq
  1      2      3
sp.1 0.2 0.0000000 0.25
sp.2 0.4 0.0000000 0.00
sp.3 0.4 0.3333333 0.00

```

图 3 R 语言操作示范

Fig. 3 Operation example in R console

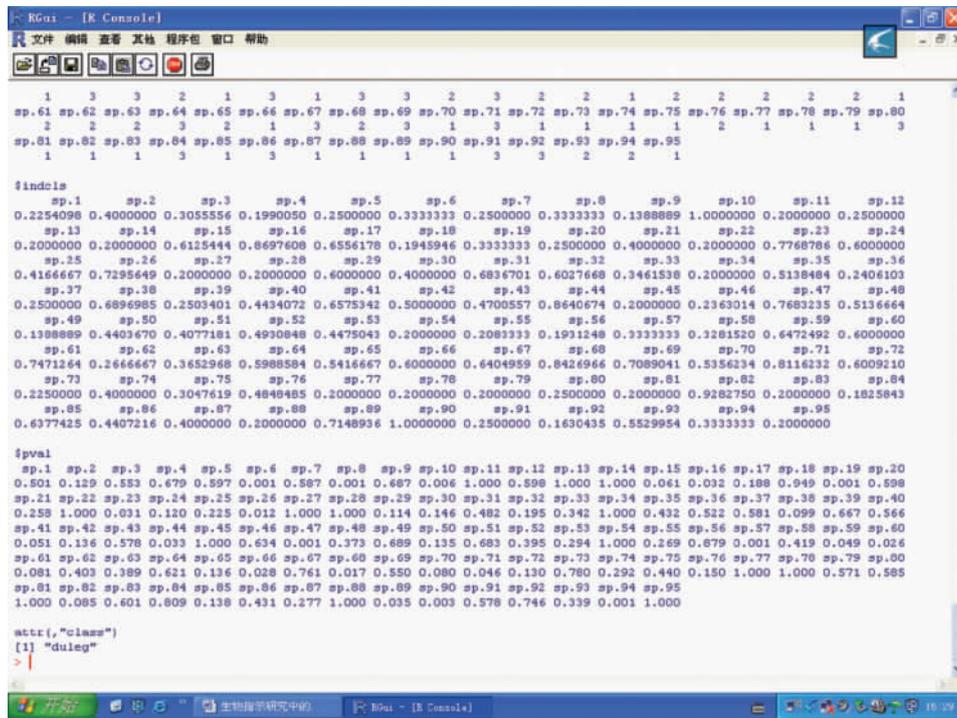


图 4 R 语言计算结果显示
Fig. 4 Calculation results in R console

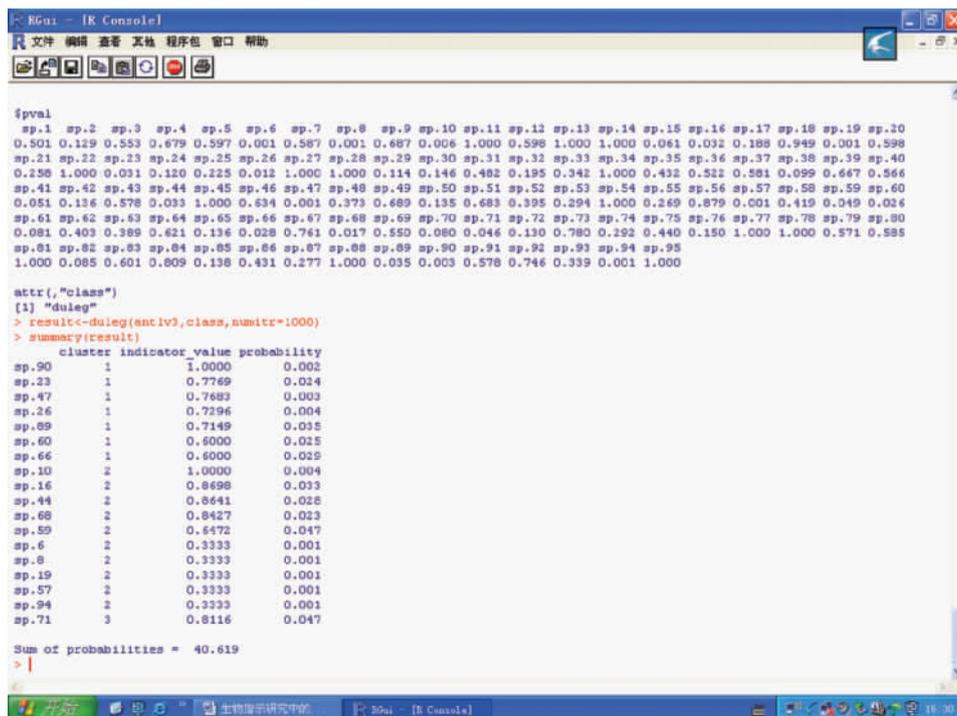


图 5 R 语言操作示范 2
Fig. 5 Operation example 2 in R console

5 IndVal 方法在指示昆虫研究中的应用

IndVal 方法的产生是生物指示研究方法论上最有意义的进步 (McGeoch, 2007)。而昆虫是颇受关注的指示生物类群,通过对不同生境、不同植被或是不同干扰类型昆虫 IndVal 值的计算,研究人员能够找到具有明显指示意义的昆虫物种。步甲是公认的指示昆虫, Elek 等 (2001) 在匈牙利 Bükk 国家公园的研究发现, *Carabus hortensis* 能指示郁闭度大的森林,而 *Carabus violaceus* 和 *Pterostichus melanarius* 则能指示山毛榉林; Magura 等 (2001) 发现,在匈牙利 Aggtelek 国家公园 *Pterostichus melanarius* 等 8 种步甲能指示草地,而 *Pterostichus oblongopunctatus* 等 7 种步甲则指示森林,其中 3 种指示森林内部,2 种能指示森林边缘; Larsen 等 (2003) 利用步甲来指示美国爱荷华州东北部的森林和牧场,其中 *Amphasia interstitialis* 等 13 种步甲能指示森林生境,而 *Amara cupreolata* 等 6 种步甲则指示牧场生境。蜣螂在生物指示研究中也受到关注,在南非滕贝大象公园, *Sisyphus sordidus* 等 6 种蜣螂能指示混交林,而只有一种蜣螂 (*Sisyphus* sp. *Y sensu* Paschalidis) 能指示沙林 (McGeoch *et al.*, 2002); Verdú 等 (2007) 在墨西哥中部的墨西哥生物圈保护区发现, *Dichotomius colonicus* 等 3 种蜣螂能指示地表裸露的亚高山灌丛,而 *Aphodius guatemalensis* 等 4 种蜣螂则指示有地被物的亚高山灌丛。蚂蚁是指示生物多样性和环境变化的重要昆虫, Parr 等 (2004) 在非洲南部萨王纳地区发现 *Monomorium notulum* 等 6 种蚂蚁可以指示火烧过的榄仁树林; Botes 等 (2006) 在南非西开普省发现 *Tetramorium quadrispinosum* 等 6 种蚂蚁可以指示多汁卡罗植被, *Technomyrmex* sp. 指示 Strandveld 植被, *Camponotus mystaceus* 指示 Proteoid 和 Restioid 植被,而 *Pheidole* sp. 则指示 Fynbos 植被; 研究还发现,一些蚂蚁可以作为一定坡向不同海拔的指示物种,如 *Monomorium fridae* 是该地区西坡海拔 0 ~ 900 m 的指示物种,而 *Tetramorium quadrispinosum* 等 3 种蚂蚁是东坡海拔 500 m 的指示物种 (Botes *et al.*, 2006)。Nakamura 等 (2007) 比较了澳大利亚东部亚热带地区雨林和牧场的地表蚂蚁群落,发现 2 种大头蚁 (*Pheidole* spp.) 能指示雨林, *Rhytidoponera metallica* 可指示牧场; Sakchoowong 等 (2008) 对泰

国北部不同土地利用生境的蚂蚁群落研究显示, *Pseudolasius* sp. 能指示山地森林,而 *Centromyrmex feae* 则指示茶园生境。IndVal 方法还被用于其它指示昆虫类群的研究,如 Mata 等 (2008) 在巴西 Cerrado 生物带对果蝇的研究发现, *Drosophila nigricruria* 能指示未受干扰的萨王纳, *Drosophila paraguayensis* 等 5 种果蝇能指示未受干扰的长廊林,而 *Drosophila melanogaster* 和 *Drosophila cardinoides* 则指示城市环境。

上述研究在指示物种判断和指示值及 *P* 值有无上存在不同,一些研究以指示值 ≥ 0.7 为标准来确定指示物种,而有些研究则以指示值 ≥ 0.5 为标准;一些研究给出了明确的指示值和 *P* 值,而有些研究则缺乏明确的指示值或 *P* 值。建议在以昆虫作为指示生物进行生物指示研究时,统一以指示值 ≥ 0.7 作为确定指示物种的标准,并给出各物种的指示值和 *P* 值。此外,缺少对不同研究结果的比较分析,尤其是同一物种对不同地区不同生境的指示,如上述研究中, *Pterostichus melanarius* 既是山毛榉林 (匈牙利 Bükk 国家公园) 也是草地 (匈牙利 Aggtelek 国家公园) 的指示物种,显然这一结果是由于研究中样地设置的不同而造成的。针对同一类群的指示昆虫研究结果,有必要进行比较分析,明确其指示意义。

6 IndVal 方法在生物指示研究中的应用前景

当前生物指示研究中大多数缺乏对指示效果的检验,没有产生与环境参数关系牢固的指示生物 (McGeoch, 2007)。而 IndVal 方法不仅能够用于指示物种判断,还能够用于指示效果的检验 (McGeoch *et al.*, 2002): 运用独立数据对已经筛选出来的指示物种进行指示值分析,用于检验指示物种的指示效果,指示值 ≥ 0.7 (或 70%), 表明指示效果良好,反之则效果不佳。随着生物指示研究在我国的兴盛, IndVal 方法在判断和检验指示物种方面将起到重要作用。

参考文献 (References)

- Botes A, McGeoch MA, Robertson HG, Van Niekerk A, Davids HP, Chown SL, 2006. Ants, altitude and change in the northern Cape Floristic Region. *J. Biogeogr.*, 33: 71—90.

- Caughley G , Gunn A , 1996. Conservation Biology in Theory and Practice. Blackwell Science , Cambridge. 1—459.
- De Cúceres M , Legendre P , 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* , 90(12) : 3566—3574.
- Dufrêne M , Legendre P , 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* , 67(3) : 345—366.
- Elek Z , Magura T , Tóthmérész B , 2001. Impacts of non-native Norway spruce plantation on abundance and species richness of ground beetles (Coleoptera: Carabidae). *Web Ecology* , 2: 32—37.
- 丁国微译 2005. <http://www.biosino.org/R/R-doc/>
- Hiddink JG , 2005. Implications of Liebig' s law of the minimum for the use of ecological indicators based on abundance. *Ecography* , 28(2) : 264—271.
- Larsen KJ , Work TT , Purrington FF , 2003. Habitat use patterns by ground beetles (Coleoptera: Carabidae) of northeastern Iowa. *Pedobiologia* , 47: 288—299.
- Legendre P , 2009a. Introduction to the R statistical language. http://biol09.biol.umontreal.ca/BIO6077/Introduction_to_R.pdf.
- Legendre P , 2009b. Practicals using the R statistical language. http://biol09.biol.umontreal.ca/BIO6077/Practicals_in_R.pdf.
- 刘新民, 杨劼, 2005. 沙坡头地区人工固沙植被演替中大型土壤动物生物指示作用研究. *中国沙漠* , 25(1) : 40—44.
- Magura T , Tóthmérész B , Molnár T , 2001. Edge effect on carabid assemblages along forest-grass transects. *Web Ecology* , 2: 7—13.
- Mata RA , McGeoch M , Tidon R , 2008. Drosophilid assemblages as a bioindicator system of human disturbance in the Brazilian Savanna. *Biodivers. Conserv.* , 17: 2899—2916.
- McGeoch MA , 2007. Insects and bioindication: theory and progress//Stewart AJA , Lewis OT , New TR (eds.). *Insect Conservation Biology*. London: CABI Publishing. 144—174.
- McGeoch MA , van Rensburg BJ , Botes A , 2002. The verification and application of bioindicators: a case study of dung beetles in a savanna ecosystem. *J. Appl. Ecol.* , 39(4) : 661—672.
- Nakamura A , Catterall C , House A , Kitching R , Burwell CJ , 2007. The use of ants and other soil and litter arthropods as bio-indicators of the impacts of rainforest clearing and subsequent land use. *J. Insect Conserv.* , 11(2) : 177—186.
- Parr CL , Robertson HG , Biggs HC , Chown SL , 2004. Response of African savanna ants to long-term fire regimes. *J. Appl. Ecol.* , 41(4) : 630—642.
- R Development Core Team , 2009. R: A language and environment for statistical computing. R foundation for statistical computing , Vienna , Austria. ISBN 3-900051-07-0 , URL <http://www.R-project.org>
- Roberts DW , 2007. labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.3-1. <http://ecology.msu.montana.edu/labdsv/R>.
- Sakchoowong W , Jaitrong W , Ogata K , 2008. Ant diversity in forest and traditional hill-tribe agricultural types in Northern Thailand. *Kasetsart J. (Nat. Sci.)* , 42: 617—626.
- van Straalen NM , Krivolutsky DA , 1996. Bioindicator Systems for Soil Pollution. Dordrecht: Kluwer Academic. 1—261.
- Verdú J R , Moreno CE , Sánchez-Rojas G , Numa C , Galante E , Halfpeter G , 2007. Grazing promotes dung beetle diversity in the xeric landscape of a Mexican Biosphere Reserve. *Biol. Conserv.* , 140: 308—317.