

计算植被生态学及其解决方案^{*}

王本洋, 余世孝

(中山大学生命科学学院, 广州 510275)

摘 要: 初步提出了计算植被生态学这一新的生态学分支学科及其面向算法的计算机技术支持新方案, 认为数据结构 and 算法研究是其核心研究内容之一。提出了按数学方法和计算机技术在生态学中的作用对现代生态学分支学科进行分类的 MC 划分方法, 该方法能反映计算植被生态学的学科地位, 其分类效果优于景观生态学时空尺度划分方法和应用生态学应用性—基础性划分方法。按照数据采集、存贮、分析和显示的流程分析了计算植被生态学的研究方法、研究内容及其主要支撑技术, 认为计算植被生态学的建立有助于产生可被重复利用的生态学模型和算法, 实现植被研究的计算机化, 为区域可持续发展服务。

关键词: 数学方法; 计算机技术; 计算植被生态学; MC 划分方法

中图分类号: Q948.15 **文献标识码:** A **文章编号:** 0529-6579 (2004) 03-0075-05

1 问题的提出

自 Williams 和 Lambert^[1] 首次应用计算机成功地实现样方分类以来, 计算机技术在植被科研中的应用不断深入。近年来计算机技术尤其是相关植被信息系统^[2] 等赖以建立的 GIS (Geographic Information System) 技术等更是极大地推动了植被科研的标准化、可视化、集成化、新发展。另一方面, 各种数学方法也不断被引入到植被分析中来, 如地统计学、分形理论等, 并广泛应用于植被研究的各个方面。

植被研究的复杂化使得计算机成为实际工作中不可缺少的工具; 而基于计算机技术的各类植被研究最终要落实到计算机软件需求上来, 因此有必要对应用于植被分析的计算机应用程序开展相关研究。当前植被研究的计算机技术支持方案都具有面向结果的特点, 即为了解决某一问题而采用既有计算机软件包, 如各种统计软件等。但这种面向结果的解决方案不能满足植被科研的需要, 因为:

一是计算机软件的老化现象。现有植被数量分析程序包的设计语言各不相同, 如 GINKO 程序包采用了 C++ 语言^[3] 等, 但可视化的程序设计语言已成为当今软件设计的首选。因各自运行环境的差异会带来计算机兼容性方面的问题; 因用户对其内部实现知之甚少, 还会造成技术交流、算法改进和程序更新等方面的困难。这些都会导致既有生态学

程序变得老化, 其利用价值逐渐降低。

二是植被分析的集成化趋势。就生态学的复杂程度而言, 通用统计软件如 SPSS, SAS 等的功能远远不够^[4], 因其不具备数字化功能, 也不支持对多媒体数据的处理, 智能化分析功能较低。目前公认的标准数值分析软件 MATLAB 虽可用于进行生态学系统分析、生态学图像处理 and 生态系统仿真研究^[4], 但不具备空间分析、三维分析等功能。通用 GIS 软件具有强大的空间分析、网络分析等功能, 但其统计分析功能却不及上述统计软件。

植被科学的发展具有两个鲜明的特点: 一是数学化, 即其应用的数学方法渐趋复杂; 二是计算机化, 即其对计算机技术的依赖程度越来越强, 对计算机软件的要求越来越高。植被科研被动利用计算机技术是导致其与计算机技术应用不能协调发展的根本原因。本文提出计算植被生态学及其面向算法的计算机技术支持新方案, 对如何解决上述问题进行初步探讨。

2 计算生态学

科学计算的兴起是 20 世纪最重要的科学进步之一。计算已成为第三种方法, 与实验、理论共同成为科学方法论的基本环节; 大规模科学计算的发展水平被认为是衡量一个国家综合国力的重要标志, 已得到世界各国的重视^[5,6]。

1998 年在意大利召开的统计生态学国际大会

^{*} 收稿日期: 2003-09-08

基金项目: 教育部骨干教师基金资助项目; 广东省千百十工程优秀人才基金资助项目

作者简介: 王本洋 (1973 年生), 男, 博士生; 通讯联系人: 余世孝; E-mail: lssx@zsu.edu.cn

上提出了计算生态学 (Computational Ecology) 的概念^[4]。该概念一经提出就倍受重视,目前已发展成为生态学的前沿研究领域。它的产生具有特殊的历史背景^[4],一方面 是受到现代技术理论和计算方法的带动和影响,而且生态学自身也具备相应的计算技术储备;另一方面也是完善生态学研究模式和解决现代生态学问题的实际需要。

3 计算植被生态学的提出

计算植被生态学 (Computational Vegetation Ecology) 是以数学生态学、统计生态学等生态学理论和现代计算理论、计算方法为理论基础,以数据结构和算法研究为核心研究内容,以多种高新技术特别是计算机技术为支撑,实现植被研究计算机化的科学,是计算生态学 (Computational Ecology) 在植被水平上的一门生态学分支学科。

提出计算植被生态学的学科概念,有利于强调数学方法与计算机技术应用在解决现代植被科学问题中的重要作用,反映现代植被科学发展所具有的数学化、计算机化特点,有利于系统总结植被数量分析方法和数学生态学、统计生态学等学科的研究成果,对明确植被研究进一步发展方向具有重要的理论和实践意义。从问题解决的路线来看,计算植被生态学的研究路线具有图 1 所示的一般形式。

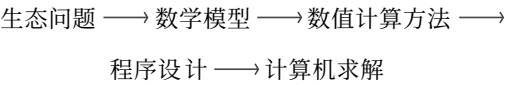


图 1 计算植被生态学的研究路线

Fig 1 Research Route of Computational Vegetation Ecology

数学生态学、统计生态学等虽然有借助相关专业知识和数学理论建立数学模型这一过程,但其以理论研究为主,基本不考虑模型的计算机实现等问题;虽然也利用计算机技术解决实际问题,但以利用现有计算机软件为主,并不注重软件的算法研究及其编制等问题,或者有少数涉及到软件编制过程(如植被数量分析方法)。信息生态学或生态信息学则强调对现代计算机技术的利用,并不注重数据结构和算法研究等问题。

计算植被生态学系统地研究从生态学问题的提出到计算机求解的整个过程(见图 1)涉及的问题,它一方面大量利用数学生态学、统计生态学及植被数量分析方法等理论成果,另一方面同样重视数学模型及其它生态学理论的计算机实现技术。它广泛利用现代计算理论、科学计算方法和以计算机技术为核心的现代高新技术如 GIS 技术、RS (Remote Sensing) 技术、GPS (Global Positioning

System) 技术、VR (Virtual Reality) 技术、Web 技术等,是数学生态学、统计生态学以及信息生态学等生态学分支学科进一步向纵深发展的必然趋向。

4 数据结构与算法的核心地位

数学是自然科学各分支的基础,其力量在于提供一种能表达思想特别是极复杂关系符号逻辑的能力,同时又保持表达式的简明性^[3]。但是,在数学理论上非常漂亮的结果应用于实际计算时却经常出现行不通的现象,如解线性方程组的 Cramer 法则因其运算量太大而根本不能用于实际计算。数量分析方法和统计分析方法是植被生态学研究两种重要的分析方法。根据 Allaby^[7]的定义,这两种分析方法都不考虑其计算机实现问题,或者不把计算机实现作为其研究的重点。而计算植被生态学把植被研究方法的计算机实现作为重点研究的问题,其核心工作之一就是数据结构和算法研究。

数据结构和算法被认为是计算机科学的两个重要支柱,两者构成一个不可分割的整体。数据结构的 不同会直接影响到算法的选择和效率。数据结构也称数据的逻辑结构,指数据元素之间抽象化的相互关系,这种相互关系可用一组运算及相应的运算规则来描述。算法 (algorithm) 一词源于算术 (algorism),它是对特定问题求解步骤的一种描述,是计算机指令的有限序列。计算机程序实际上就是用计算机语言描述的算法,而一种算法可用多种计算机语言来描述,即算法描述可独立于具体的语言工具而存在。以某种计算机语言为工具,采用合适的数据结构,根据既有算法,就可以方便地编制或重新生成某种计算机语言的程序,解决植被生态学的实际问题。

单靠提高计算机的运算速度和增加计算机的容量来解决计算问题极为困难;另外,由于发达国家对我国实施的最先进计算机技术和先进软件(包括系统软件和应用软件)的封锁^[9],重视科学计算的作用,研究如何节省存储和计算时间的数据结构和新算法很有必要。数据结构和算法研究应该是计算植被生态学的研究重点和核心内容之一。

由于数据结构和算法研究对外界提供了计算机实现的理论依据和关键技术,因此有助于实现技术的交流、共享和更新,加快植被研究计算机化的进程。与面向结果的方案不同,这种面向算法的植被研究计算机技术支持方案不存在对语言的依赖性,也不存在软件老化等问题,能够使植被研究与计算机技术协调发展,是解决植被研究计算机技术支持的较好方案。

5 生态学学科划分方法

5.1 传统的生态学学科划分方法

当前生态学学科的划分方法大致有两条途径：一是景观生态学的时空尺度途径；二是应用生态学的基础性—专业性的途径。

5.1.1 景观生态学的时空尺度划分方法 根据景观生态学的时空尺度观点^[8]划分的结果，每一个时空等级都可对应于一个生态学分支学科，如种群生态学、群落生态学等，但这种划分方法没有反映数学对生态学各分支学科的深刻影响，也没有体现生态学向宏观—微观两极化发展^[9]的特点。实际上，数学方法被应用于每个等级的生态学研究之中，不同时空等级尺度都对应着数学生态学的一个研究领域。如在个体水平上的种群动态模拟、群落水平上的植被分类与排序方法等。这种划分方法也没有体现计算机技术对生态学学科发展的巨大推动作用。因此，景观生态学的时空尺度划分法不能包括计算生态学这一新的生态学分支学科。

5.1.2 应用生态学的基础性—专业性划分方法 根据周启星等^[9]的观点，应用生态学以解决资源可持续利用问题、生态问题和环境污染问题等三大问题为其重点研究的内容，并通过生态过程的研究与基础生态学紧密联系，并行发展。根据应用生态学的基础性—专业性划分的结果^[9]，数学生态学是基础生态学的一个分支学科。这种划分方法反映了数学在生态科学中的重要地位和作用，但也没有体现计算机技术对现代生态学发展的巨大作用。因此，计算生态学不属于基础生态学的范畴，也不属于应用生态学的范畴。

5.2 生态学学科的 MC 划分方法

由于景观生态学的时空尺度划分方法或应用生态学的基础性—专业性划分方法没有反映数学方法对生态学发展的基础地位与作用，或者计算机技术对生态学发展的巨大推动作用；也不能很好地包括计算生态学这一新的分支学科、反映生态学学科发展的新动向，因而都是不够全面的划分方法。

采用 MC 划分方法可克服上述两种划分方法的不足。其思路是，以互相垂直的数学方法轴 (Mathematical Methods) 即 M 轴和计算机技术轴 (Computer Technologies) 即 C 轴构成 MC 分类空间 (MC Classification Space, 即 MCCS)，然后把生态学各分支学科展布在 MCCS 上 (图 2)。

应用 MC 划分方法可获得与景观生态学的时空尺度划分法相似的结果 (图 2)。按生态学研究的组织水平和时空尺度来划分，有宏观生态学

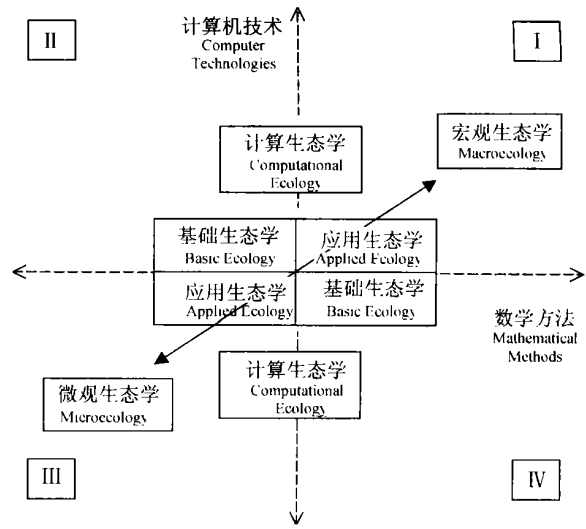


图 2 MC 分类空间中的生态分支
Fig. 2 Branches of ecology in MC classification space

(Macroecology) 和微观生态学 (Microecology) 两极。在个体、种群以至全球水平上的生态学研究，其时间尺度为年到千年、空间尺度为米到万米^[8,9]，属于生态研究中宏观的一极^[9]，共同构成宏观生态学，分布在 MCCS 的第 I 象限；而在种、个体、部件以至分子或基因水平的应用生态学研究，是微观的一极^[9]，共同构成微观生态学，分布在 MCCS 的第 III 象限。把宏观生态学和微观生态学展布于 I，III 象限中，可反映数学方法和计算机技术在宏观和微观各个层次上的作用和影响。

应用 MC 划分法也可反映基础生态学与应用生态学的关系 (图 2)。按应用生态学的基础性—专业性来划分，有基础生态学和应用生态学两大分支，两者并行发展^[9]，可对称地展布在 MCCS 的四个象限中。

应用 MC 划分方法，可把计算生态学展布于 MCCS 的 C 轴且位于应用生态学 and 基础生态学之上 (图 2)。这样既反映了计算生态学、统计生态学等学科之间存在的继承与发展的关系和数学方法的基础地位与作用，也反映了计算生态学对计算机技术的强烈依赖特点，且与数学生态学可看作基础生态学的一个分支^[9]的观点相一致。

MC 划分方法充分反映了数学方法和计算机技术对现代生态学发展的重要作用，具有一定的可扩展性，可包括计算生态学等新的生态学分支学科。MC 划分方法具有简洁明了、重点突出、前瞻性好等优点，其分类效果优于景观生态学时空尺度划分方法和应用生态学应用性—基础性划分方法，是一种比较好的生态学学科分类方法。

6 基于GIS 的计算植被生态学研究流程

GIS 系统为计算植被生态学研究提供了理想的技术平台。许多先进的 GIS 系统都具有数字化的数据采集能力以及整合 GPS、RS 数据等多源数据的能力, 集成了数据库技术、可视化技术、网络技术、多媒体技术等先进技术; 还提供完善的二次开发语言以支持二次开发、增强系统功能, 或者提供系统与外部程序进行信息交换的机制以弥补二次开发语言不足。

基于 GIS 的计算植被生态学的研究流程可大致划分为数据的采集、存贮、分析、显示等四个阶段, 数据结构和算法研究贯穿始终 (图 3)。

6.1 数据采集

新技术的引入扩展了植被研究的数据来源, 如 GPS 数据、RS 数据、各种专业性的图件及群落调查数据等。数据采集的标准化是实现计算植被生态学研究的基础。为了实现数据资源的共享, 增强研究结果的可比性, 保证基于计算机的计算植被生态学研究正常进行, 必须建立完整的可操作性强的相关标准。其中, 标准化的植被分类系统和数据采集方法是最重要的两个标准。如美国国家地理数据委员会于 1996 年制订了植被分类系统和植被信息标准, 并建立了通用植被数据库。该系统已在美国开始使用^[10]。我国目前还没有类似的标准。

遥感影像数据是计算植被生态学研究的重要数据源之一。信息提取的标准化是提高分析结果可比性的基础, 如何从遥感影像中提取其中所包含的大

量生物信息以及所获取信息的保真度对后续的分析、模拟、制图等具有重大影响。因此, 也是研究的重要内容之一。

6.2 数据存贮

群落调查数据等生态学专业数据是计算植被生态学数据存贮方案研究的重点, 因为其它如遥感影像数据等均可由成熟的商业软件来处理。数据库是后续数据分析和算法设计的基础, 因而数据库技术决定了整个计算植被生态学研究的水平和质量的高低, 必须给予足够的重视。

群落调查所得数据一般是二维数据矩阵, 如样方一种矩阵等。在群落研究中, 常见特殊矩阵主要为实对称矩阵和稀疏矩阵, 后者系指零元素占优势 (比如占 80%以上) 的矩阵。为了提高后续分析与模拟等操作的效率, 实现对高阶矩阵的运算, 必须对不同类型特殊矩阵的特点、存贮方案及计算方法进行研究。

6.3 数据分析

计算植被生态学的数据分析方法, 除传统的分析方法之外, 还应尽可能利用其它多种数学方法, 如地统计学和分形理论等, 并充分利用集成于 GIS 平台的多种新技术进行时空分析、三维分析、网络分析等。可视化分析环境是基于 GIS 技术的计算植被生态学数据分析的重要特色。

时间序列分析是指对一定区域的植被沿时间序列作动态的分析, 可用于植被演替研究。空间分析包括空间结构分析、空间分布和动态模拟、点数据的插值、图层迭置等过程^[11]。它是地理信息系统

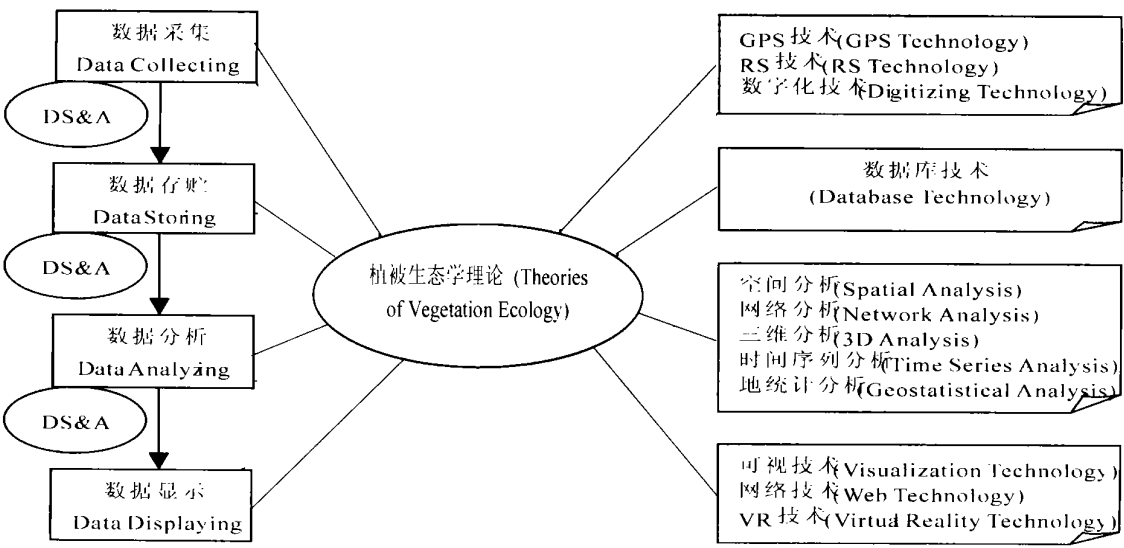


图 3 基于 GIS 的计算植被生态学的研究流程

Fig. 3 Research flowchart of GIS-based computational vegetation ecology

DS & A 表示数据结构和算法研究

的核心功能，可分为空间检索、空间拓扑迭加分析和空间模拟分析等三个不同的层次^[12]。利用地统计学的原理，把时间和空间结合起来，作时空复合分析，可用于定量评估区域化变量的变异程度及其随时间的变化^[11]，对植被进行时间序列分析和空间分析的复合分析。

6.4 数据显示

数据显示主要指分析结果的显示，其表示方法主要分为三大类：一类是各种数据表，以表格的形式存储统计分析的结果，如群落分析表、植被分类系统中各分类单元的面积数据等；第二类是各种曲线图，如种—面积曲线图、重要值—面积曲线图等；第三是各种图件，如植被分布图、群落剖面图、表示种间联结的半矩阵图和星座图等。

7 结 论

计算植被生态学的建立有助于消除生态学计算研究的“破碎化”状况^[4]，产生可重复使用的生态学模型和计算方法。本文提出的面向算法的计算机技术支持方案是计算植被生态学的较好解决方案。

计算植被生态学的本质是植被研究的计算机化。植被研究数据的结构和相关算法研究贯穿始终，是其核心研究内容之一。这种面向算法的计算机技术支持方案是计算植被生态学的重要特征之一。该方案不存在对计算机语言的依赖性，能够适应语言多样的形势、克服软件老化等问题，有助于实现植被分析技术的交流、共享和更新。

计算植被生态学以数学生态学、统计生态学等生态学理论为基础，借助现代高新技术特别是计算机技术，实现植被研究的计算机化，可在植被分

类、生态学过程的监测、生物多样性、全球变化等多方面的研究中发挥巨大作用，为区域可持续发展服务。

参考文献:

[1] Williams W T, Lambert J M. Multivariate methods in plant ecology. I. Association—analysis in plant communities[J]. Journal of Ecology, 1959, 47: 83—101.

[2] 余世孝, 罗睿, 陈里娥, 等. 广东生物地理信息系统的研究Ⅱ[J]. 中山大学学报(自然科学版), 1996, 35(6): 85—89.

[3] 余世孝. 数学生态学导论[M]. 北京: 科学技术出版社, 1995.

[4] 丛沛桐, 阎秀峰, 赵则海, 等. 计算生态学的形成与发展[J]. 生态学杂志, 2002, 21(1): 58—61.

[5] 石钟慈, 袁亚湘. 奇效的计算—大规模科学与工程计算的理论和方法[M]. 湖南: 湖南科学技术出版社, 1998.

[6] 石钟慈. 第三种科学方法: 计算机时代的科学计算. 北京: 清华大学出版社; 广州: 暨南大学出版社, 2000.

[7] Allaby M. Oxford Dictionary of Ecology[M]. Shanghai: Shanghai Foreign Language Education Press, 2001.

[8] 陈文波, 肖笃宁, 李秀珍. 景观空间分析的特征和主要内容[J]. 生态学报, 2002, 22(7): 1135—1142.

[9] 周启星, 孙顺江. 应用生态学的研究与发展趋势[J]. 应用生态学报, 2002, 13(7): 879—884.

[10] 李博, 杨持, 林鹏, 等编. 生态学[M]. 北京: 高等教育出版社, 2000.

[11] 王正军, 张爱兵, 程家安, 等. 基于GIS的种群动态的时空分析与模拟研究的方法进展[J]. 生态学报, 2002, 22(1): 104—110.

[12] 陈述彭, 鲁学军, 周成虎, 编著. 地理信息系统导论[M]. 北京: 科学出版社, 2000.

Computational Vegetation Ecology And Its Solutions

WANG Ben yang, YU Shi xiao

(School of Life Sciences, Sun Yat sen University, Guangzhou 510275, China)

Abstract: A new branch of ecology called computational vegetation ecology (CVE)) is proposed. Data structure and algorithm are considered as the core of its researches. The methodology, contents and supporting technologies are discussed according to the research flowchart of data collecting, storing, analyzing and displaying. The branches of modern ecology are also reclassified based on the degree of utilizing mathematical methods and computer technologies (MC classification method). Computational vegetation ecology will help producing replicable ecological models and algorithms, realize computerization of vegetation study, help serving regional sustainable development.

Key words: mathematical methods; computer technology; computational vegetation ecology (CVE); MC classification method