

基于 GIS 的中药材产地适宜性 分析系统的设计与实现*

□孙成忠 刘召芹 (中国测绘科学研究院 北京 100039; 山东科技大学 青岛 266510)

陈士林** 魏建和 (中国医学科学院 药用植物研究所 北京 100094)
(中国协和医科大学)

赵润怀 王继永 周应群 (中国药材集团公司科技研发中心, 北京, 100094)

摘 要: 根据物种引种原则气候相似性, 设计了基于数据库的 GIS 中药材产地适宜性分析系统, 国家基础地理信息数据库、气候因子数据库、土壤数据库和第三次全国中药普查数据库为分析的数据基础。初步研究结果表明, 将 GIS 的空间聚类分析与空间分析应用于中药材产地适宜性分析中, 能够科学、快速、准确地分析出与药材道地产区最为相近的地区, 结束了依靠传统经验和单个药材、单个气候因子、单个产地分析的低效、准确性差的做法。

关键词: TCMGIS-I 产地适宜性 聚类分析 空间分析

中药材的野生分布区(原生起源中心)是保证药材质量和生长最适宜区域, 但由于原产区种植条件不合适或各种社会因素限制等难以保证药材的规模化生产, 因此, 很多药材需要进行引种扩种。为保证药材质量和规模化生产, 需要采用一定的科学理论和方法指导中药材产地的选择, 即选择与野生分布区生长环境相似的区域, 发展道地药材生产。

生物引种的基本原则是引种地气候必须与原产地气候相似保证引种中药材的正常生长。对引种地中药材不仅要求一定的产量而且要求生产出来的药材与原产区的药材具有相同或相似的品质以保证药材

的药性。而生长环境中的诸多因子如土壤条件、温差、海拔、降水量等与中药材品质有密切关系。本文采用地理信息系统(GIS)作为分析平台, 评价药材道地产区气候和土壤因子与全国其他地区的相似性差异, 根据相似程度等级来确定药材的适宜产地, 为药材引种地的选择提供科学地分析依据。这一系统称为《中药材产地适宜性分析地理信息系统》(TCMGIS-I)

一、地理信息系统的涵义

地理信息系统 GIS (Geographic Information System)是集计算机科学、地理学、测绘遥感学、环境科学、城市科学、空间科学、信息科学为一体的新兴学

组稿日期: 2006-06-06

* 北京市科技项目(Y0804001000091): 北京地区中药材种植的区划研究, 负责人: 陈士林; 国家发改委中药材扶持项目(2005-01): 道地药材产地适宜性数值区划, 负责人: 魏建和。

** 联系人: 陈士林, 本刊编委, 研究员, 博导, 主要研究方向: 中药资源可持续利用, Tel: 010-62899700, Email: slchen@implad.ac.cn。

112 [World Science and Technology/Modernization of Traditional Chinese Medicine and Materia Medica]

科,是在计算机硬件和软件支持下,运用系统工程和信息科学的理论,科学管理和综合分析具有空间内涵的地理数据,以提供对规划、管理、决策和研究所需信息的空间信息系统。

地理信息系统可以利用各种空间分析方法管理空间数据信息,在对各种信息综合分析基础上以寻求不同空间实体间的相互关系,并分析和处理一定区域内的分布现象和过程。空间分析对空间信息(特别是隐含信息)所具有的分析 and 提取已成为地理信息系统区别于一般信息系统的主要特征,亦是评价地理信息系统功能的主要指标。

二、空间分析-TCMGIS- I 开发原理

空间分析是基于地理对象的位置和形态特征的数据分析技术,从宏观上可以归纳为以下几种:

(1)空间图形数据的分析运算;

(2)非空间属性数据的运算;

(3)空间和非空间数据的联合运算。运算手段包括各种几何逻辑运算、数理统计分析、代数运算等以解决有关地理空间的实际问题,提取和传输地理空间信息尤其是隐含信息。空间分析的目的主要有以下 4 方面:

(1)认知——通过有效地获取、科学地描述空间数据,利用数据来再现事物本身;

(2)解释——理解和解释生成观察地理图案和数据的背景过程,认识事件的本质;

(3)预报——了解、掌握事件发生的现状,运用推测规律对未来的状况作出预测;

(4)调控——调控在地理空间上发生的事件。

GIS 的空间分析分为栅格数据空间分析和矢量数据空间分析两大类。本系统中栅格数据空间分析用于对中药材生长的气候环境进行分析后聚类以寻找中药材的最适宜生长区域,矢量数据空间分析用于对分析出的生长区域和行政区划进行叠加,进一步分析适宜区的行政范围及面积等。

1. 矢量数据的空间分析

矢量数据空间分析通常包括:空间数据查询和属性分析,多边形的重新分类、边界消除与合并,点

线、点与多边形、线与多边形、多边形与多边形的叠加,缓冲区分析,网络分析,面运算,目标集统计分析等。矢量数据的叠加分析主要分为擦除、相交、合并等,其算法模型(见图 1)。

2. 栅格数据的空间分析

栅格数据空间分析功能通常包括:记录分析、叠加分析、滤波分析、扩展领域操作、区域操作、统计分析等。栅格数据的叠加分析要优于矢量数据的叠加分析,其基本的算法模型(见图 2)。

3. 空间聚类分析

聚类分析亦称为群分析,是数值分类学的一个分支。空间聚类是采用聚类分析处理带有空间位置的数据。传统的聚类方法在进行数据分析时一般针对数据的属性要素而很少考虑数据的空间特征;而

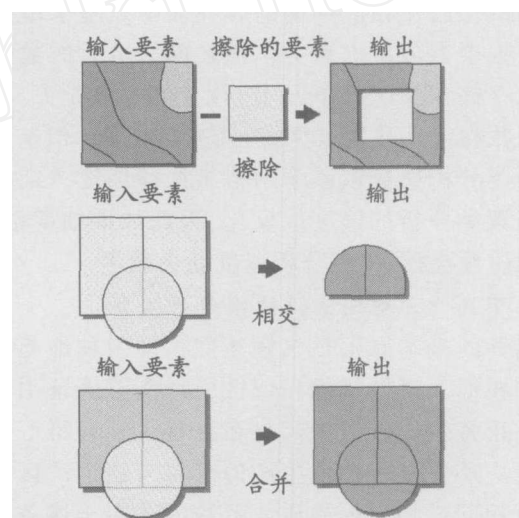


图 1 矢量数据算法模型

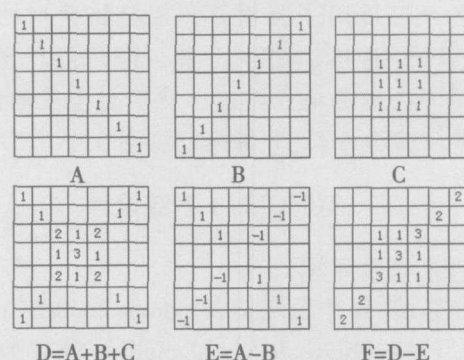


图 2 栅格数据算法模型

空间聚类方法不仅考虑属性数据的聚类分析,而且要考虑其空间特征。

空间聚类是指将数据对象集按照相似度组成不同的簇,这样同一簇中的对象之间具有较高的相似度,而不同簇中的对象差别较大,即相异度较大。空间聚类方法通常可以分为四大类:划分法、层次法、基于密度的方法和基于网格的方法。根据应用目的选择不同算法,如对于栅格数据分析,基于网格和密度的算法更为合适。此外,算法的速度、聚类质量以及数据的特征,包括数据的维数、噪声、数量等因素都影响到算法的选择。

基于格网的空间聚类分析是把将空间划分为有限数目的正方形网格,以构成可以进行聚类分析的网格结构。这样,每一个网格相当于一个数据对象,其观测值存放在相应网格的属性表中。基于单元网格的数据聚类分析优点为:描述网格单元的数据统计信息存储在相应的单元中,与查询要求无关;有助于实现并行计算和增量更新;聚类速度快。但基于网格的聚类分析质量依赖于网格结构的细度,细度的高低与聚类分析的速度成反比,因此如果细度很大,分析的质量会很好,但分析速度就会较慢。

4. 中药材产地适宜性栅格分析模型

物种引种的基本原则是气候相似,即引种地与原产地气候相似。根据气候相似性引种能够保证引种中药材的正常生长。但中药材在正常生长获得产量的基础上,还要求引种地生产的药材与道地产区的药材具有相同或相似的品质。环境因子中土壤条件与中药材品质形成有密切关系。把与中药材生长环境相关的各个环境因子分别量化成栅格数据结构,即将海拔、降水、温度、湿度、土壤、日照等因子数据转换成1平方公里的单元网格,然后对生成的网格数据进行空间聚类分析,分析模型见图3。

三、TCMGIS- I 系统设计

1. 开发运行环境

操作系统:操作系统选 Windows 2000/XP;

开发环境:Microsoft Visual Studio. NET 2003 C#

GIS 基础软件:选用 ESRI 公司的 ArcEngine 作为

二次开发组件,ArcEngine 是基于 COM 的二次开发组件,提供了丰富的图形管理功能及大量的空间分析和多元统计分析接口,可以快速方便地用于开发自己的 GIS 应用系统。

数据库管理软件:采用关系数据库管理软件 Microsoft SQLServer2000 来存储和管理空间数据和属性数据。空间数据及相应的属性数据通过 ESRI 公司的 ArcSDE 进行存储和访问,其他的属性数据通过 ADO.net 进行访问。

2. 系统结构设计

系统为 C/S 结构设计,系统结构见图4。

3. 空间数据库的调用

为了进行基于气候的中药材产地适宜性分析,选择了以下4个数据库:

(1)基础地理信息数据库。

基础地理信息数据库选取了国家 1:1000000 的

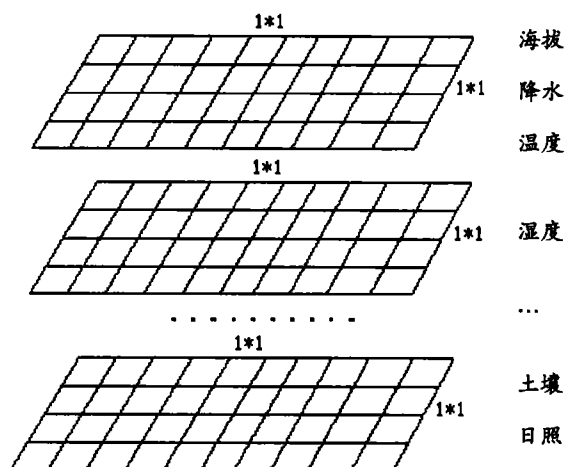


图3 中药材产地适宜性栅格分析模型

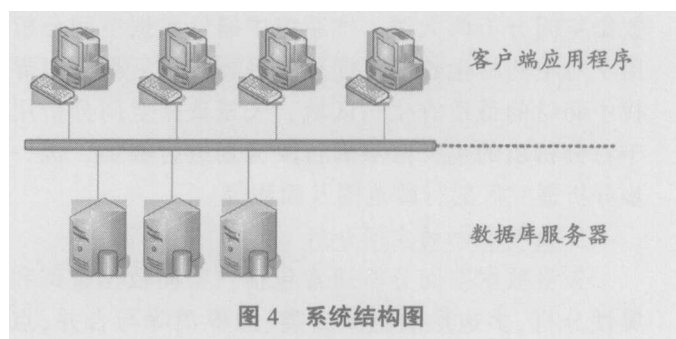


图4 系统结构图

基础地理信息数据,主要包括政区、居民地、交通与管网、水系及附属设施、地貌、地名、测量控制点等内容。基础地理信息数据以点、线、面的矢量形式存放于空间数据库中。

基础地理信息数据用于和分析后的栅格和矢量数据进行叠加分析和制图,以丰富的图形方式表达分析结果。

(2) 气候资源数据库。

气候资源数据库储存了基于全国气象站点 30 年(1971 年~2000 年)的逐月地面气象数据,按国家认可的方法计算所得的按月栅格数据(每平方公里)及衍生数据,包括 7 个指标:平均温度,相对湿度,降水量,日照时数,以及极端最低温度,极端最高温度。

(3) 土壤数据库。

土壤数据库采用的是 1:4000000 第 3 次全国土壤普查至亚类的土壤数据。

(4) 中药材普查数据库。

中药材普查数据库为第 3 次全国中药资源普查数据,用于中药材道地环境的查询及分析对比。

4. 系统主要实现的功能

根据中药材产地适宜性分析的需要,系统主要功能包括:用户管理、图层管理、图形显示、图形编辑、数据导入、数据导出、空间查询、空间分析、多元统计、输出功能等,其主界面见图 5。

(1) 用户管理。

用户管理是为了保证系统安全权限管理,分为

角色管理和用户管理,可以根据用户群的需要,设定不同的角色,每个角色的权限又分为系统权限和操作权限,系统权限控制系统菜单权限,主要控制系统主菜单下的各权限,如一般的用户只能具有数据查询权限,而只有高级用户才具有系统维护、图层管理、数据编辑等权限。操作权限控制用户对不同数据操作的权限。

(2) 图层管理。

图层管理图层的增加、删除、风格设置等。系统提供了加载所有 ArcGIS 支持的数据格式,如:shp、Geodatabase、grid、sde 图层等。

(3) 图形显示。

图形显示的功能包括控制图形的显示与关闭、加载影像、放大缩小和平移及专题图制作等。

(4) 图形编辑。

图形编辑用于对数据的变更进行变更,保持数据的现势性。

(5) 空间查询。

空间查询包括从图形到属性和从属性到图形的双向查询,提供了点击查询、选择查询、模糊查询等多种查询方式。

(6) 空间分析。

空间分析是系统的主要功能,系统提供了矢量数据的空间分析、栅格数据的空间分析、及矢量和栅格的混合分析功能。

(7) 多元统计。

多元统计功能是系统的重要组成部分,多元统计包括监督分类、非监督分类、最大似然分类等分类方法。监督分类中提供了手动分类、等间隔分类、自定义间隔分类、标准差分类方法。

(8) 输出功能。

根据用户具体要求可以输出多种形式的数据、报表、图表。

四、TCMGIS- I 系统分析流程

1. 目标值的确定

要对中药材引种适宜性进行了分析,首先要确定中药材道地产地的生长因子,输入药材的最适宜

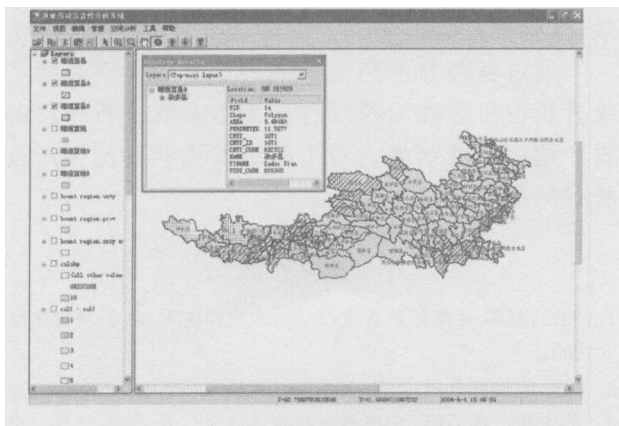


图 5 系统主界面图

的生长环境因子。生长环境因子的输入有两种方法,其一是手动输入各环境因子值;其二是输入药材的道地产区,系统根据输入的道地产区的环境因子进行因子分析,分析出药材的最适宜生长环境因子。

2. 数据标准化

气候、土壤数据在进行相似性的聚类分析前必须进行标准化处理。数据的标准化方法有多种:总和标准化、标准差标准化、极大值标准化、极差标准化及绝对偏差均值标准化。分析者可以根据分析的需要选择各种标准化方法。

3. 相似系数计算

经过标准化后的数据可以进行距离的计算。系统按照统计学的原理也提供了多种可供选择的方法,如:绝对值距离、欧氏距离、明科夫斯基(Minkowski)距离法。上述3种距离可以统一为明科夫斯基距离公式:

$$d_{ij} = \left[\sum_{k=1}^n |x_{ik} - x_{jk}|^q \right]^{1/q}$$

上式中当 $q=1$ 时为绝对值距离,当 $q=2$ 时为欧氏距离。

根据距离计算公式,计算各个栅格相对于目标值的距离,可以看出,距离越小的和目标值相似,反之差异越大。

4. 权重的确定

根据药材对不同环境因子的依赖性不同,确定各个环境因子的权重,依赖性越强的因子权重越大,反之越小。各个因子的权重之和为1。

5. 适宜区空间分析

空间分析分为栅格数据的空间聚类分析和矢量数据的空间叠加分析。栅格数据的空间聚类分析是根据前面生成的标准化后的数据和距离方法,进行加权空间聚类分析,将所有环境因子数据进行聚类,生成新的已分类的栅格数据,栅格数据中值最大的那一类即为适宜生长区域,次一级的为次适宜区,再次一级为一般区,其他的为不适宜区。

为进一步分析适宜区,将分类的栅格数据转换成面状的矢量数据文件,将生成的矢量数据和行政区划(到县)数据进行矢量数据的相交运算,然后利

用行政区划数据对运算后的适宜区数据进行空间查询,查询出各行政区划中的适宜区、次适宜区和一般区。

6. 成果输出

将转换后的适宜区数据和基础地理信息数据进行叠加显示,选择输出类型,即可输出预先设计的各种类型表格、图表。

五、结 语

利用本系统对5个中药材品种进行了分析,包括暗紫贝母、浙贝母、人参、川芎和附子,根据这些药材的原生地(道地产区)的气候因子数据(最低温度、最高温度、降雨量、日照等)、地形因子数据(海拔高度)和土壤类型数据,采用等权绝对值距离法分析后得出每种药材的适宜区、次适宜区和一般区。

将每个药材品种的产地适宜区分析结果与第3次全国中药资源普查数据分布结果进行对比,可以得到适宜区分析结果的初步评价。

对比评价结果可以表明,产地适宜区分析结果与全部资源普查分布的区域基本吻合,同时可以看出在中药资源调查中没有的产地适宜区,这些区域将成为今后发展药材生产新的空间。

基于GIS的中药材产地适宜性分析系统(TCMGIS-I)的建立,将影响药材生长的环境因子进行了量化和综合分析,能科学、快速、准确地分析出与药材道地产区最为相近的地区,结束了依靠传统经验和单个药材、单个气候因子、单个产地分析的低效、准确性差的做法。

通过各药材中药生长适宜区的分析,表明了系统分析得可靠性与科学性,由本系统分析得出的药材产地适宜性结果,对于各地发展药材的生产具有较好的指导意义。

参考文献

- 1 邬伦,刘瑜.地理信息系统-原理、方法和应用,北京:科学出版社,2001.
- 2 郭仁忠.空间分析,武汉:武汉测绘科技大学出版社,1997.
- 3 陈士林,肖小河.中国药用植物的数值区划.资源开发与市场,1994,10(1),8~10.

- 4 陈士林,肖培根.中药资源可持续利用发展导论.北京:中国医药科技出版社,2006.
- 5 肖小河,陈士林,陈善墀.川产道地药材生产布局研究.中国中药杂志,1992,17(2),70~72.
- 6 肖小河,陈士林,陈善墀.四川乌头和附子气候生态适宜性研究.资源开发与保护,1990,6(3),51~153.
- 6 魏建和,陈士林,魏淑秋,等.北柴胡适生地分析及数值区划研究.世界科学技术—中医药现代化,2005,7(6):125.
- 7 何晓群.多元统计分析.北京:中国人民大学出版社,2004,54~92.
- 8 朱明.数据挖掘.北京:中国科学技术大学出版社,2002,129~164.

Design and Realization of traditional Chinese Medicine Adaptability Analyzing System Based on GIS

Sun Chengzhong, Liu Zhaoqin

(The Chinese Academy of Surveying and Mapping, Beijing 100039, China; College of Geo-information Science and Engineering, Shandong University of Science and Technology, Qingdao 266510, China;)

Chen Shilin **, Wei Jianhe

(Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Chinese Peking Union Medical College, Beijing 100094, China)

Zhao Runhui, Wang Jiyong, Zhou Yingqun

(Technology development center, China National Group Corp. of Traditional & Herbal Medicine, Beijing 100094, China)

Based on the species growth principle—climate comparability, the traditional Chinese medicine adaptability analyzing system (TCMGIS- I) based on GIS was designed, the state basic geographic information database, climate elements database, soil database and the third national Chinese medicine census database is the analytic base data. Preliminary study showed that utilizing the TCMGIS- I , in which the space clustering analyzing and space analyzing technology is the main technology, to analyze the traditional Chinese medicine adaptability to find out the most equivalent growth condition region to initial zone of medicine materials, the result is objective, rapid and the precision can be high, overcoming the disadvantages of traditional analyse based on experience ,single medicinal plant, single ecological factor and single producing area.

Keywords: TCMGIS- I area adaptability clustering analyzing space analyzing

(责任编辑:郭 屹,责任编审:柳 莎,张志华,责任译审:熊艳艳)

(Continued from page 124)

A geographic information system of suitable producing area evaluation of Traditional Chinese Medicine (TCMGIS- I) was applied to evaluate the producing area for *Fritillaria unibracteata*. The producing areas of different suitability rating in China were distinguished. The result was consistent with the characteristics of the biological and distribution of *Fritillaria unibracteata*. Thus, TCMGIS- I is a useful tool in TCM for the determination of suitable producing area, and will be instructive effectively for medicinal plants introduction, acclimatization and wild tending.

Keywords: TCMGIS- I ; suitability of producing area; *Fritillaria unibracteata*; *Bulbus Fritillariae Cirrhosae*

(责任编辑:郭 屹,责任编审:张志华,责任译审:熊艳艳)

[World Science and Technology/Modernization of Traditional Chinese Medicine and Materia Medica] 117