

地理数据尺度转换方法研究进展

孟 斌^{1,2,3}, 王劲峰¹

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 北京联合大学应用文理学院, 北京 100083;
3. 中国科学院研究生院, 北京 100039)

摘要: 尺度问题在地理学、生态学和人文科学等众多领域都具有非常重要的地位。近年来, 随着对地观测技术和地理信息科学的飞速发展, 解决地理数据的尺度转换问题成为目前地理信息科学及相关研究中的热点和难点问题之一。在地理信息科学相关领域中, 地图学和遥感科学研究人员从尺度概念的理解到尺度转换的理论和方法都做了大量的研究, 对解决地理数据空间特征的尺度转换做出了重要的贡献。在地理数据属性特征的尺度转换研究领域, 地理信息科学研究者提出的面域插值方法是解决此问题的主要方法之一。同时, 在社会经济科学领域, “小区域统计学”也发展了一套相关的理论和方法, 试图解决统计单元的变更问题。文章在全面回顾和比较不同研究领域解决“尺度转换”问题方法的基础上, 重点介绍面域插值方法和小区域统计学的基本原理及其典型应用。

关键词: 尺度; 尺度转换; 属性数据; 空间插值; 小区域统计

1 引言

尺度间的相互依赖是地理学家观察世界的三个“透镜”之一, 对尺度的关注, 可以至少避免两种错误, 即用错误的空间尺度观察问题和对因果关系的曲解^[1]。因此尺度问题是地理学研究的核心问题之一^[2]。在气候、地貌、水文、生物多样性等广泛领域中, 众多学者从不同侧面对尺度问题进行了大量的研究, 如路云阁等^[3]结合平均气温和降水数据研究了气候变化的尺度效应, 朱蓉等^[4]讨论了气象数值模拟中参数选择的多尺度问题; 吴凡等^[5]讨论了地貌形态的多尺度表达的方法, 刘纪根等^[6]研究了土壤侵蚀模型中的尺度转换问题; 夏军^[7]、丁晶等^[8]回顾了尺度研究在水文学研究中的重要地位和目前水文研究中尺度相关热点问题; 周红章等^[9]研究了生物多样性的变化格局与时空尺度的关系。在经济地理学中, Openshaw^[10]提出的可变面元问题 (Modifiable Areal Unit Problem, MAUP) 成为尺度研究的代表问题。在资源开发和利用研究中, 也有学者关注到旅游资源^[11]、渔业资源^[12]、土地利用^[13, 14]等研究中存在的尺度效应和尺度选择问题。在地质学相关学科中, 生态学家也很早就认识到了尺度问题的重要性, Robinson 在 1950 年就提出生态学中著名的“尺度谬论”问题^[15]; 邬建国^[16]、吕一河等^[17]回顾和总结了尺度研究在生态学研究中的最新进展及其重要地位, 赵文武等^[18]探讨了生态学研究尺度转换的理论和方法, 程序等^[19]研究了流域生态—经济重建中的尺度选择问题。

在地理信息科学领域, “尺度转换”和“最适合尺度选择”问题一直是研究人员关注的热点问题^[20-22]; 由于地理信息科学对计算单元的严格要求, Goodchild 认为尺度是地理信息科学中最重要的问题^[23], 甚至提出“尺度科学”的概念^[24]。陈军也提出在中国空间

收稿日期: 2004-01-20; 修订日期: 2004-12-14

基金项目: 国家 973 项目 (2001CB5103); 国家 863 计划项目 (2002AA135230-1) [Foundation: National "973" Program, No.2001CB5103; National High Technology Research and Development Program, No.2002AA135230-1]

作者简介: 孟斌 (1971-), 男, 安徽肥东人, 讲师, 博士研究生, 中国地理学会会员, 研究方向为地理信息技术与空间分析方法、区域可持续发展。Email: mengb@lreis.ac.cn

数据基础设施建设中要加强多尺度数字地理空间数据框架建设^[25]。传统上,地理信息系统的广泛应用的一大瓶颈就是数据问题。地理数据是各种地理特征和现象间关系的符号化表示,包括空间位置、属性特征和时态特征三个部分^[26]。GIS 发展之初主要问题是数据来源渠道少。但随着技术的进步,获取数据的能力有了极大提高^[27]。但实际应用和研究过程中,合适的数据来源还是困扰研究人员的一个难题。在自然灾害灾情空间预报^[28]环境流行病学^[29]以及土地利用与土地覆被变化^[30]等众多人文与自然综合研究的领域,由于数据来源的不同,尤其是这些数据依附的空间单元尺度的不同,使得数据融合成为难题。因此,随着自然和人文综合研究的趋势日渐加强和地理信息技术的广泛应用,解决“尺度鸿沟”的地理数据的空间尺度融合的要求越来越紧迫。NCGIA 将空间化(非空间信息的空间表达)作为研究重点之一,UCGIS1998 年提出的优先研究领域也包括“尺度”问题的研究。在尺度转换研究中,不同领域对地理数据三个方面关注的重点有所差异,本文重点讨论其中属性数据的尺度转换问题。

2 尺度转换的理解

尺度指的是经历时间的长短或在空间上涵盖范围的大小,即通常所指尺度有时间和空间两方面的含义。Thompson 早在 1917 年就指出,“尺度的影响,不仅仅是尺度本身,而是与其周围的环境紧密联系”。在不同学科领域,同一术语如 Scale 表达了不同的含义,有时不同用词又表达同样的含义,因此有必要对尺度及相关的概念进行比较(表 1)。

表 1 尺度及其相关概念^[31]

Tab. 1 Scale and its related concepts

术语	含义	使用领域
尺度	指的是一个事物或过程经历时间的长短或在空间上涵盖范围的大小	广泛使用
绝对尺度	实际的距离、方向、形状和几何特性等	
相对尺度	利用相对距离、方向、形状和几何特性以及特定的函数关系表达绝对尺度。	
制图比例尺	地图距离和地球表面实际距离的比率	地图学
分辨率	测量的精确程度,空间采样单元的大小	遥感
粒径	给定数据的最大分辨率	景观生态学
范围	研究区域的大小或考虑的时间范围	
支集	度量或定义(属性)值的空间	地统计学
步长	相邻现象、采样或分析单元间隔的度量	空间分析生态学

尺度转换是将数据或信息从一个尺度转换到另一个尺度的过程,可以是向上尺度转换,也可以是向下尺度转换。不同领域的学者往往采用不同的术语来表达同样的概念,这对中文的准确表达造成更大的困难,如有学者用“尺度上推”和“尺度下推”来区别向上和向下的尺度转换过程^[17],还有国内学者称之为“尺度推绎”^[18],也有学者建议采用“尺度压缩”表达向下尺度转换过程。本文采用直译,统一将从较小尺度观测结果获得较大尺度信息的向上尺度转换过程称为尺度扩展;而把大尺度上的信息分解到更小的尺度向下尺度转换的过程称为尺度收缩(表 2)。Wu 认为尺度扩展是将信息从精确的尺度向模糊的尺度转换的过程,尺度收缩则是相反的过程^[32]。Nordhaus 认为尺度收缩是试图融合数据的主要过程或方法,其可能的的方法包括回归(regression)、三角内插(triangular interpolation)和“择多原则”(majority rule)^[33]。

不同领域对尺度转换的关注有所不同,与社会经济数据空间尺度转换问题相关的研究一般称为“不同分区系统间数据转换”^[34]或“空间(表面)转换”^[35],或直接以“面插值”(areal interpolation)代替^[36,37]。Anselin 将此问题称为“数据综合”或“数据转换”,并认为此问题是空间分析在社会经济领域应用的三个最重要的工具之首^[38]。

表 2 常用关于空间尺度转换的表述

Tab. 2 Different expressions of scaling

	英文表述	其它中文翻译	具体含义
尺度 扩展	Upscaling Scaling-up Top-down	向上尺度转换 尺度上推	从较小尺度(空间)观测中获得较大尺度(空间)上信息的过程; 是将信息从精确的尺度(高分辨率)向模糊的尺度转换的过程
尺度 收缩	Downscaling Scaling-down Bottom-up	向下尺度转换 尺度下推	将大尺度上(空间)的信息分解到更小的尺度(空间)上的过程; 是将信息从模糊的尺度向精确的尺度转换的过程

3 尺度转换主要方法

尺度转换在自然地理及其相关领域研究中, 常常与层次理论(hierarchy theory)紧密联系。Wilby等认为尺度转换就是信息在不同层次水平的变化^[39]。层次理论最初的产生是因为研究“复杂性”问题的需要, 在管理科学、经济学、心理学、生物科学、生态学和系统科学等领域得到极大发展^[32, 40-43]。如O'Neill等将层次理论(hierarchy theory)作为生态学中联系空间尺度和信息的理论框架^[42]。在景观生态学中, 结合斑块动力学理论和层次理论, 发展了层次斑块动力学理论(Hierarchical patch dynamics paradigm, HPDP)。层次斑块动力学理论通过结合层次理论和斑块动力学理论, 提供了一个研究时空异质性、尺度和层次组织如何影响生态系统结构和动态的理论框架^[44]。鲁学军等^[45]探讨应用层次理论建立地理空间的尺度—结构分析模式问题。

尺度问题也是目前地图学研究中难点和热点问题之一。制图综合是其中非常重要的一个焦点问题^[46, 47]。毋河海^[47]详细论述了地图综合的基础理论问题和技术方法问题, 认为地图综合过程的综合算子可以归为两类: 信息变换类和图形再现类。高文秀等^[48]讨论了专题属性数据综合的若干问题。陈军等^[49]讨论了应用Voronoi图理论改善空间地图数据概括问题, 认为Voronoi图在制图综合中应用大致可归纳为: 基于凸壳的综合、基于骨架线的综合、基于邻域的综合等三个方面。吴凡^[50]讨论了基于小波分析的线状特征数据无级表达问题。由于数字地图综合问题未能取得突破性进展^[46, 47], 有学者指出解决多重表达和协同更新问题是目前解决地理空间框架数据的多尺度问题的重要手段^[51]。

多比例尺问题是地图学研究和尺度有关的另外一个重要领域。多比例尺问题即无级比例尺空间数据压缩与复现。对于矢量数据多比例尺的数据组织与表示方法主要有三种: 即建立同一数据库的多个比例尺版本; 运用面向对象方法、语义数据模型等技术来支持空间数据的多级表达; 通过地图自动综合来自动生成同一数据库的多个版本^[52]。尹连旺等根据地理要素的空间属性特征和重要程度, 提出了复合指数开方根模型, 试图解决不同地理环境条件、不同地物要素信息综合的问题^[53]。李德仁等提出以空间信息多级网格为基础的解决全球地理信息多尺度表示和数据组织问题的新思想^[52]。

遥感图像处理研究中, 尺度问题同样非常重要^[54, 55]。遥感中的尺度问题主要体现在三个方面^[56]: 遥感信息和遥感模型的尺度效应问题; 选取合适空间分辨率的遥感数据; 遥感信息的尺度转换问题。Woodcock和Strahler^[57]研究了尺度因子与遥感分类精度的关系, 推动了遥感信息的尺度效应问题研究; 苏理宏等^[58]讨论了遥感研究中的尺度问题最新进展。在遥感信息的尺度转换方法研究中, Xia认为傅立叶变换(fourier transform)、变异函数(variograms)、小波理论和分形理论是主要的描述和处理尺度变换的方法, 并且在该文中详细介绍了分形理论在尺度转换的应用^[59]。其后, 分形理论在遥感图像尺度转换中的研究不断深入^[60]。20世纪80年代发展起来的小波理论和方法^[61, 62]和基于粗神经网络的多传感器数据融合^[63]也在遥感数据的尺度研究领域也表现出一定的优势。张颢等提出利用直方变

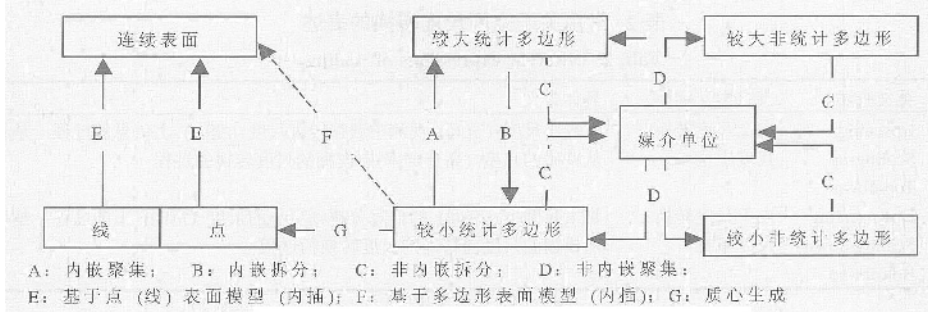


图 1 属性数据空间尺度转换系统框架 引自 Briggs^[35]

Fig. 1 Spatial transformation methods for the analysis of geographic data (After Briggs^[35])

差图研究图像融合问题^[64]。柏延臣等研究了尺度与遥感图像分类不确定性的关系^[56]。

4 GIS 空间插值

地理信息科学中, 社会经济数据和自然环境数据常常具有不同的地理单元, 即“可变更面元”问题 (MAUP)^[10]。处理此问题的根本即空间单元转换问题可以描述为“地理空间单元间属性数据的转换过程”^[36]。MAUP 可以理解为“由于对连续的地理现象的空间报告单元的人为划分, 并在此基础上产生的人为的空间格局”。MAUP 的研究最早可以追溯到 1934 年^[65], Openshaw 和 Taylor 于 1979 年首次提出此概念^[66], 并在 1983 年详细阐述了这一概念^[10]。Flowerdew 和 Openshaw 明确提出 MAUP 的概念及其研究的迫切性^[67]。在其后, 有关 MAUP 问题的一系列研究成果见诸报端。MAUP 主要研究聚集 (aggregation), 而对于拆分 (disaggregation) 研究较少。聚集主要解决属性数据从小区域向大区域转换问题, 与前述尺度扩展 (upscaling) 含义相同。拆分则指从较大区域向较小区域的数据转换, 与前述尺度收缩 (downscaling) 对应, 或是不同边界区域间数据转换。

在地理分析中, Kolaczyk 和 Huang 提出所谓“多尺度统计模型” (multiscale statistical models) 来解决尺度转换问题^[68]。Mugglin 和 Carlin 也利用层次模型和贝叶斯方法结合, 实现人口数据的尺度转换^[69]。Briggs 提出 GIS 与空间数据地理单元有关的分析技术和方法可以分为聚集, 拆分和空间建模, 并给出了系统框架 (图 1)^[35], 是目前为止关于尺度转换的比较全面的描述。

当前处理空间尺度拆分问题方法有点与多边形叠加^[34, 70]、面域加权^[71, 73]和修正的面域加权^[36]等 (表 3)。与地理数据空间尺度转换有关的空间模型均以面插值为核心, 对“聚集”和“拆分”两方面的内容均有考虑^[37, 73]。Turner 和 Openshaw 特别研究了处理拆分的空间插值方法, 比较了面域加权、最大化保留插值方法 (pynophilactic) 和利用辅助数据的“巧妙的空间插值方法” (smart SIM), 并提出利用神经网络作为“聪明的空间插值方法” (clever SIM)^[73]。陈军讨论了应用 Voronoi 图理论改善空间插值方法的问题^[49]。朱会义^[74]也特别讨论了自然地理要素空间插值的问题。

表 3 GIS 中属性数据空间尺度转换方法

Tab. 3 Methods of scaling in GIS

数据处理过程	可能的方法	应用实例
内嵌聚集	模拟退火 优化	
内嵌拆分	面域加权 修正的面域加权 点与多边形叠加	[71]; [72]
非内嵌拆分	修正的面域加权	[36], [83]
非内嵌聚集	最大化保留插值 加权形心平滑 多变形过滤 “聪明的空间插值方法”和其他点模型	[73], [80], [81]

点与多边形叠加本质上是点与多边形的空间叠加过程。面域加权处理的实际上是多边形叠加。Pycnophilactic, 其含义为最大化保留 (mass preserving) 是简单面域加权的扩展, 其原理就是对简单面域加权计算出的目标区域的值采用邻近区域的值进行修正。修正的面域加权与简单的面域加权方法最大的差异在于采用了面积以外辅助的信息, 即“借力” (borrow strength) 问题。

4.1 点与多边形叠加

点与多边形叠加本质上是点与多边形的空间叠加过程。在 ARC/GIS 中也有相关模块“PointGrid”。就是将属性信息 (依附于点状实体) 转化为 GRID 单元, GRID 的单元根据其覆盖的点 (或多个点) 被编码和赋值, 多个点时, 可以以出现次数最多的点来赋值; 如果没有点, 则赋值为“NODATA”。该方法的一般算法如下:

$$\hat{Y}_t = \sum_s Y_s \quad s \in t \quad (1)$$

式中: \hat{Y}_t 为待求的属性值; Y_s 为目标在源区域的属性值。

该方法算法相对简单, 结果的准确性与目标区域的大小和点的分布有关, Sadahiro 讨论了有代表性的最佳点分布问题^[70]。Okabe 和 Sadahiro 详细分析了该方法的误差问题^[34]。

虽然点与多边形叠加处理的是“拆分”问题, 但从此方法本质上, 可以看出其更适用于处理“聚集”过程。

4.2 面域加权 (Areal Weighting)

面域加权处理的实际上是多边形叠加 (polygon overlay), 就是将目标区和源区叠加, 分别计算各交叉区域的属性值, 再按目标区计算。也有学者将其称为“比例分配”^[75]。

根据属性变量性质不同, 可以将属性分为广延量 (extensive) 和强度量 (intensive), 广延量是一类可以累加的变量, 如人口总数, 农作物产量等。在 Areal Weighting 中, 目标区中广延量的值, 一般等于源区域与目标区交叉区域内变量值之和。强度量是表示变量之间比率关系的一类变量, 不可以直接累加, 如出生率, 百分比数据等。在面域加权中, 目标区中强度量的值, 一般等于源区域与目标区交叉区域内强度量值的加权平均。其基本算法分别如下^[76, 77]:

$$V_t = \sum (V_s \times A_{st} / A_s) \quad \text{属性为广延量} \quad (2)$$

$$\text{或 } V_t = \sum (V_s \times A_{st} / A_t) \quad \text{属性为强度量} \quad (3)$$

式中: V_t 为目标区域的属性值, V_s 为源区域的属性值, A_t 、 A_s 和 A_{st} 分别为目标区域、源区域和交叉区域的面积。

Sadahiro 分析了面域加权方法的误差^[78]。Gregory 以英格兰和威尔士长期人口变迁为例比较了主要的“面插值”方法, 尤其是面域加权方法在人口分布研究中的适应性^[79]。

4.3 最大化保留 (Pycnophilactic)

最大化保留是简单面域加权的扩展, 其基本原理就是考虑到地理学第一定律, 对面域加权方法计算出的目标区域属性值采用邻近区域的计算结果进行修正^[80], 步骤如下:

- 第一步, 对研究区域生成标准大小的单元格网 (GRID);
- 第二步, 对每个单元格利用前述面源比重法进行赋值;
- 第三步, 对每个单元格利用其邻域单元计算结果进行平滑;
- 第四步, 汇总整个源区域单元格网得到源区域属性值;
- 第五步, 对在同一源区域的目标区的单元格赋相同权重, 保证源区域属性值的稳定;

第六步, 重复 3 到 5 步, 直至达到预定目标。

该方法的著名的应用是 NCGIA 关于全球人口分布研究^[81]。

4.4 修正的面域加权

修正的面域加权与简单的面域加权方法最大的差异在于采用了面积以外辅助的信息, 即“借力”问题。该方法又可以分为。

1) 使用控制区的面域加权法 当源区和目标区内的属性都不是均匀分布时, 可以引入控制区的概念, 即假设存在一个中间单元区称为控制区, 它的属性分布是均质的, 可以通过控制区作为中间步骤获得目标区属性信息的估计^[36, 82]。其计算过程是: 先将源区与控制区叠加, 在已知源区的属性信息和叠加后的交叉区域面积的情况下, 控制区的属性信息的密度可以通过 (4) 式得出。再将控制区与目标区叠加, 通过 (5) 式求得目标区人口。

$$P_s = \sum_c d_c a_{cs} \quad (4)$$

$$P_t = \sum_c d_c b_{ct} \quad (5)$$

式中: d_c 是控制区的属性值的密度, a_{cs} 和 b_{ct} 分别是控制区与源区和目标区的单元叠加面积, c 为控制区的个数, P_s 和 P_t 分别是源区和目标区的属性值。

2) 使用回归关系的面域加权 当源区和目标区内的属性都不是均匀分布时, 假设所求目标区的属性值与源区若干要素相关, 利用要素间的回归关系, 可以计算目标区的属性值。Gallego 将土地利用作为辅助信息, 对欧洲人口分布进行研究^[83]。其基本模型如下:

$$Y_{cm} = U_c \frac{X_m}{\sum_c S_{cm} U_c} \quad (6)$$

$$X_m = \sum_c S_{cm} U_c W + \varepsilon_m \quad (7)$$

式中: Y_{cm} 为区域 m 内 c 种土地利用类型的人口密度, X_m 为 m 区域已知人口总数, U_c 为 c 种土地利用类型与人口数相关系数, S_{cm} 为 m 区域内 c 种土地利用类型的总面积, W 为修正系数, 保证人口数的一致。由于土地利用类型的划分单元与人口统计单元尺度的差异, 该方法事实上实现了属性数据的尺度融合。

刘纪远等利用“格点生成法”模拟中国人口密度^[84], 该方法本质上与使用回归关系的面域加权方法是一致的。

4.5 GIS 空间插值方法比较

空间插值思想是解决尺度融合的主要理论和方法基础, 这些方法的基本假设有所不同, 导致它们具有不同的优点和不足 (表 4)。

5 小区域统计学 (Small Area Statistics)

所谓“小区域”, 表面含义是较小的地理区域, 但本质上是指区域内样本点较少, 因此在统计分析过程中, 需要从相关区域“借力”来获得可靠的分析结果, 其产生和发展与邻居统计学 (neighbourhood statistics) 密切相关^[85]。小区域统计学本质上是一种间接估计, 其核心问题建立相关区域 (数据) 的联系模型。由于小区域统计“借力”的区域往往具有不同大小的空间尺度, 所以小区域统计方法也是解决属性数据尺度融合的途径之一。Anselin 就指出解决“数据综合”问题的方法框架包括统计学领域的“小区域估计”和 GIS 中的“空间插值”或“面插值”^[38]。Simpson 也指出从统计学的观点来看, 地理学中

表 4 GIS 空间插值方法比较
Tab. 4 Comparison of the method of areal interpolation in GIS

方法名称	基本假设	主要优点	主要缺点
点与多边形叠加	点代表区域的所有属性	简单, 方便计算	选择质点位置具有很大的不确定性; 无法保证源区与目标区的属性值相等
面域加权	假定属性在源区内的分布是均匀	能够保证源区与目标区的属性总量相等	区域边界处变量的差异可能很大
最大化保留	相邻的单元相互影响, 且较近区域影响较大	① 保证了平滑前后的统计单元的属性总量一致性; ② 不要求统计单元中变量分布的均质性	单元边界处变量的较大变化会影响插值的效果
修正的面域加权	辅助变量在源区与目标区都可获得, 且和目标变量的关系相同	能够保证源区与目标区的属性总量相等	单元边界处变量的差异可能很大

的尺度转换可以被认为小区域统计中“综合估计”(synthetic estimators)的特例^[86]。

小区域统计主要包括“综合估计”、“复合估计”(composite estimators)和“基于模型的估计方法”(model-based estimators)几大类方法^[87]。综合估计假定较大区域中属性值可以通过直接估计获得, 在此基础上假设不同大小区域的某些属性具有相同特征, 实现对小区域的估计^[88]。复合估计是一种平衡综合估计和直接估计的方法, 通过确定两者权重实现。基于模型的方法又可以分为“区域模型”和“单元模型”。

小区域统计学的核心算法是贝叶斯统计的思想, 广泛应用于人口及和人口相关的统计研究中, 比如世界银行组织对世界人口贫困状况问题的研究^[89]。同时, 在作物估产、疾病、房地产等研究中也有广泛应用^[90, 91]。

以下主要参考 Rao 的相关文章^[85, 92], 介绍小区域统计的主要方法及其应用。

5.1 综合估计

假设在较大区域里, 可以将样本 Y 划分为 g 层, 将区域划分为 i 个小区域。对于每一层中的 $Y_{\cdot g}$ 可以直接利用观测数据计算, 并且 $Y_{\cdot g} = \sum Y_{ig}$, 即 Y 等于所有区域中所有层的 y 之和。小区域统计就是要计算 Y_i , 即第 i 个小区域中变量 Y 的值。问题的关键依然是找到 Y 有关的辅助变量 X 。综合估计的基本算法如下:

$$\hat{Y}_i^s = \sum X_{ig} (\hat{Y}_{\cdot g} / \hat{X}_{\cdot g}) \quad (8)$$

式中: \hat{Y}_i^s 是 Y_i 的综合估计, 且

$$Y_{\cdot g} = \sum_i Y_{ig}, \quad X_{\cdot g} = \sum_i X_{ig}$$

式中: Y_{ig} 是第 i 个小区域 g 层的变量的值, X_{ig} 是已知的辅助变量 X 在 i 个小区域 g 层中的值。

衡量综合估计的结果好坏, 一般考虑其均方差 (Mean squared error, MSE)。对 \hat{Y}_i^s 的均方差可以采用下式计算。

$$\text{MSE}(\hat{Y}_i^s) = (\hat{Y}_i^s - \hat{Y}_i)^2 - v(\hat{Y}_i) \quad (9)$$

式中: \hat{Y}_i^s 是 Y_i 的综合估计, \hat{Y}_i 是 Y_i 的直接估计, $v(\hat{Y}_i)$ 是 \hat{Y}_i 的方差。

Gonzalez 等将美国全国出生调查数据根据母亲的种族、年龄和生养孩子数目进行分层, 利用综合估计的方法对各州的一些卫生与健康变量如婴儿的低出生体重比率等进行

估计, 并比较了综合估计与直接估计的结果^[93]。

5.2 复合估计

复合估计是为了综合属性值的间接估计值和直接估计值 \hat{Y}_i^s , 通过权重 \hat{Y}_i 来计算:

$$\hat{Y}_i^c = \alpha_i \hat{Y}_i + (1 - \alpha_i) \hat{Y}_i^s, \quad \alpha_i \in [0, 1] \quad (10)$$

式中: \hat{Y}_i^c 为复合估计值。 α_i 在 0-1 之间, 理想的 α_i 使得 $MSE(\hat{Y}_i^c)$ 达到最小, 一般认为最佳的 α_i 采用公式 (11) 计算^[87]。

$$\alpha_i = MSE(\hat{Y}_i^s) / [MSE(\hat{Y}_i^s) + v(\hat{Y}_i)] \quad (11)$$

式中: 相关符号含义同公式 (9)。

5.3 区域模型

基本的基于区域的模型由两部分组成: (1) 对第 i 个区域, 变量 y 的均值 \bar{Y}_i 的函数 $\hat{\theta}_i = g(\bar{y}_i)$, 等于 i 区域内变量的真实值 θ_i 和采样误差 e_i , 即 $\hat{\theta}_i = \theta_i + e_i$; 且 e_i 符合均值为零, 方差为 ψ_i 的正态分布。(2) 连接 θ_i 和区域变量 X 的线性回归模型 $\theta_i = X_i^T \beta + v_i$, v_i 为模型误差, 且 v_i 是均值为零, 方差为 σ_v^2 的正态分布。综合起来即可得到

$$\hat{\theta}_i = X_i^T \beta + v_i + e_i \quad (12)$$

式中: $\hat{\theta}_i$ 是基于区域模型的估计值, X_i 是区域 i 内的辅助变量, β 是回归系数, v_i 为模型误差, e_i 为采样误差。

基本的区域模型的发展主要表现在以下方面: 多元变量模型的研究^[94]、时间序列和截面 (cross-section) 数据模型^[95]、空间模型等^[96]。

5.4 单元模型

基于单元的模型是将 i 个区域划分为 j 个单元, 辅助变量在所有单元均可取得。最简单的基于单元的模型可以表示为:

$$y_{ij} = X_{ij}^T \beta + v_i + e_{ij} \quad j = 1, \dots, N_i, \quad i = 1, \dots, m, \quad (13)$$

式中: y_{ij} 是 i 区域 j 单元的估计值, 其他符号含义同 (12) 式。

6 结论与讨论

通过比较以上各类方法, 尤其是其中主要的基于模型的算法, 发现它们之间存在很多相似之处, 比如它们的前提和假设基本一致, 都假定变量的统计关系在不同尺度区域保持一致或可以定量表达。同时, 比较小区域统计学的模型和修正的面域加权模型, 可以发现其主要思想也基本一致。只是前者更注重统计, 而后者更着重机理。两者的结合, 将会是一种比较好的方法。但当小区域统计用于研究空间信息时, 也有传统统计学应用于空间信息分析所存在的同样问题, 即空间信息的空间相关性问题。

地理数据空间化研究在当前相关理论和方法的研究上都与实际需要存在距离。

属性数据尺度融合是地理研究中不可避免要面对的问题, 但由于尺度问题的时空复杂性, 虽然已经发展起来的处理尺度转换的方法十分丰富, 尺度问题依然是地理研究中难点问题之一。为了简化研究对象的复杂性, 以上方法都有其特定的假设前提, 这也是科学研究的常用方法。目前为止, 对上述方法的适用条件的研究依然欠缺。需要进一步研究的是各种方法假设前提对处理属性数据等具有空间特征的地理信息是否适用。另外,

不同的方法综合比较研究, 是一种选择最合理的解决问题途径的办法。同时, 对尺度转换过程中模型选择以及模型的不确定性分析等也是需要进一步加强的研究领域。

参考文献 (References)

- [1] Cai Yunlong, Lu Dadao, Zhou Yixing et al. Chinese progress and international trends of geography. *Acta Geographica Sinica*, 2004, 59(6): 803-810. [蔡运龙, 陆大道, 周一星等. 地理科学的中国进展与国际趋势. *地理学报*, 2004, 59(6): 803-810.]
- [2] Abler R F, Adams J, Gould P. *Spatial Organization*. New Jersey: Prentice-Hall, Inc., Englewood Cliffs, 1971.
- [3] Lu Yunge, Li Shuangcheng, Cai Yunlong. Multiscale studies on climatic change and its spatial heterogeneity in recent 40 years. *Scientia Geographica Sinica*. 2004, 24(4): 432-438. [路云阁, 李双成, 蔡运龙. 近 40 年气候变化及其空间分异的多尺度研究. *地理科学*, 2004, 24(4): 432-438.]
- [4] Zhu Rong, Xu Dahai. Multi-scale turbulent planetary boundary layer parameterization in mesoscale numerical simulation. *Quarterly Journal of Applied Meteorology*, 2004, 15(5): 543-555. [朱蓉, 徐大海. 中尺度数值模拟中的边界层多尺度湍流参数化方案. *应用气象学报*, 2004, 15(5): 543-555.]
- [5] Wu Fan, Zhu Guorui. Multi-scale representation and automatic generalization of relief based on wavelet analysis. *Geomatics and Information Science of Wuhan University*, 2001, 26(2): 170-176. [吴凡, 祝国瑞. 基于小波分析的地貌多尺度表达与自动综合. *武汉大学学报·信息科学版*, 2001, 26(2): 170-176.]
- [6] Liu Jigen, Cai Qiangguo, Fan Liangxin et al. Methods of scale transfer in modeling of soil erosion and sediment yield in catchments. *Journal of Sediment Research*, 2004, (3): 69-74. [刘纪根, 蔡强国, 樊良新等. 流域侵蚀产沙模拟研究中的尺度转换方法. *泥沙研究*, 2004, (3): 69-74.]
- [7] Xia Jun. Discussion of hydrology scale. *Journal of Hydraulic Engineering*, 1993, (5): 32-37. [夏军. 水文尺度问题. *水利学报*, 1993, (5): 32-37.]
- [8] Ding Jing, Wang Wensheng. Scale analysis in hydrology. *Journal of Sichuan University (Engineering Science Edition)*, 2003, 35(3): 9-13. [丁晶, 王文圣. 论水文学中的尺度分析. *四川大学学报 (工程科学版)*, 2003, 35(3): 9-13.]
- [9] Zhou Hongzhang, Yu Xiadong, Luo Tianhong et al. How does species diversity change? spatio-temporal patterns and scales. *Chinese Biodiversity*, 2000, 8(3): 325-336. [周红章, 于晓东, 罗天宏等. 物种多样性变化格局与时空尺度. *生物多样性*, 2000, 8(3): 325-336.]
- [10] Openshaw S. *The modifiable areal unit problem*, CATMOG 38. Norwich, UK: Geo Books, 1983.
- [11] Qi Deli. Indexes and scales-related fields in ecotourism resources evaluation. *Journal of Natural Resources*, 2004, 19(4): 508-518. [齐德利. 沿海生态旅游资源评价指标及尺度研究. *自然资源学报*, 2004, 19(4): 508-518.]
- [12] Ji Min. Multi-scale generalization and expression of ocean fishery thematic attribute data. *Bulletin of Surveying and Mapping*, 2004, (6): 28-31. [季民. 海洋渔业专题属性数据多尺度综合与表达. *测绘通报*, 2004, (6): 28-31.]
- [13] Chen Youqi. Multi-scale spatial characterization of land use/land cover in China. *Scientia Geographica Sinica*, 2000, 20(3): 198-202. [陈佑启. 中国土地利用土地覆盖的多尺度空间分布特征分析. *地理科学*, 2000, 20(3): 198-202.]
- [14] Liu Yanfang. Forewarning system of dynamic balance between cultivated land demands and supplies at multi-measures. *Geomatics and Information Science of Wuhan University*, 2004, 29(5): 420-425. [刘艳芳. 多尺度耕地供需动态平衡预警体系研究. *武汉大学学报·信息科学版*, 2004, 29(5): 420-425.]
- [15] Robinson W S. Ecological correlations and the behavior of individuals. *American Sociological Review*, 1950, 15: 351-357.
- [16] Wu Jianguo. *Landscape Ecology: Pattern, Process, Scale and Hierarchy*. Beijing: Higher Education Press, 2000. [邬建国. *景观生态学——格局、过程、尺度与等级*. 北京: 高等教育出版社, 2000.]
- [17] Lu Yihe, Fu Bojie. Ecological scale and scaling. *Acta Ecologica Sinica*, 2001, 21(12): 2096-2105. [吕一河, 傅伯杰. 生态学中的尺度及尺度转换. *生态学报*, 2001, 21(12): 2096-2105.]
- [18] Zhao Wenwu, Fu Bojie, Chen Liding. Some fundamental issues in scaling. *Advance in Earth Sciences*, 2002, 17(6): 905-911. [赵文武, 傅伯杰, 陈利顶. 尺度推绎研究中的几点基本问题. *地球科学进展*, 2002, 17(6): 905-911.]
- [19] Cheng Xu, Liu Guobin. Concept and methodology on upscaling issue of small watershed prototypes for reconstructing ecological economies in Loess Plateau. *Chinese Journal of Applied Ecology*, 2004, 15(6): 1051-1055. [程序, 刘国彬. 黄土高原小流域生态-经济重建模式的尺度概念和方法. *应用生态学报*, 2004, 15(6): 1051-1055.]
- [20] Atkinson P M, Tate N J. Spatial scale problems and geostatistical solutions: a review. *Professional Geographer*, 2000, 52(4): 607-623.
- [21] Yue Tianxiang, Liu Jiyuan. The digital model for transforming information at various scales. *Journal of Image and Graphics*, 2001, 6(9): 907-911. [岳天祥, 刘纪远. 第四代地理信息系统研究中的尺度转换数字模型. *中国图像图形学报*, 2001, 6(9): 907-911.]

- [22] Li Jun, Zhuang Dafang. Study on the appropriate scale of geo-spatial data. *Acta Geographica Sinica*, 2002, 57(suppl.): 52-59. [李军, 庄大方. 地理空间数据的适宜尺度分析. *地理学报*, 2002, 57(增刊): 52-59.]
- [23] Goodchild M F. Models of scale and scales of modeling. In: Tate Nicholas J, Atkinson Peter M (eds.), *Modelling Scale in Geographical Information Sciences*. Chichester: John Wiley & Sons, Ltd., 2001. 3-10.
- [24] Goodchild M F, Quattrochi D A. Scale, multiscale, remote sensing and GIS. In: Quattrochi D A, Goodchild M F (eds.), *Scales in Remote Sensing and GIS*. Boca Raton, FL: CRC/Lewis Publisher, 1997. 1-11.
- [25] Chen Jun. Developing multi-scale NSDI. *Surveying and Mapping*, 1999, (3): 18-22. [陈军. 多尺度空间数据基础设施的建设与发展. *中国测绘*, 1999, (3): 18-22.]
- [26] Chen Shupeng, Zhou Chenghu, Lu Xuejun (eds.). *Geographical Information System*. Beijing: Science Press, 2001. [陈述彭, 周成虎, 鲁学军 编著. *地理信息系统*. 北京: 科学出版社, 2001.]
- [27] Dungan J L, Dale M R T et al. A balanced view of scale in spatial statistical analysis. *Ecography*, 2002, 25: 626-640.
- [28] Wang Jinfeng (eds.). *Regionalization of Hazards in China*. Beijing: China Science & Technology Press, 1993. [王劲峰 编著. *中国自然灾害区划*. 北京: 中国科学技术出版社, 1993.]
- [29] Veugelers P J, Hornibrook S. Small area comparisons of health: applications for policy makers and challenges for researchers. *Chronic Diseases in Canada*, 2002, 23(3).
- [30] Cai Yunlong. Perspectives on innovation in physical geography. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2000, 36(4): 576-582. [蔡运龙. 自然地理学的创新视角. *北京大学学报(自然科学版)*, 2000, 36(4): 576-582.]
- [31] Turner M G, O'Neill R V, Gardner R H et al. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 1989, 3(3): 153-162.
- [32] Wu J. Hierarchy and scaling: extrapolating information along a scaling ladder. *Canadian Journal of Remote Sensing*, 1999, 25(4): 367-380.
- [33] Nordhaus W. Creation and Analysis of a Geophysically Scaled Economic Data Set, 2001.
<http://www.econ.yale.edu/~nordhaus/homepage/>
- [34] Okabe A, Sadahiro Y. Variation in count data transferred from a set of irregular zones to a set of regular zones through the point-in-polygon method. *Int. J. Geographical Information Science*, 1997, 11(1): 93-106.
- [35] Briggs. *Spatial Transformation Methods for the Analysis of Geographic Data*. U. K. Imperial College of Science (ed.), UN/ECE Work Session on Methodological Issues Involving the Integration of Statistics and Geography. Neuchatel, Switzerland: Statistical Commission and Economic Commission for Europe, 2000.
- [36] Goodchild M F, Lanselin Deichmann U. A framwork for the areal interpolation of socioeconomic data. *Environment and Planning A*, 1993, 25: 383-397.
- [37] Pan Zhiqiang, Liu Gaoquan. The research progress of areal interpolation. *Progress in Geography*, 2002, 21(2): 146-152. [潘志强, 刘高焕. 面插值的研究进展. *地理科学进展*, 2002, 21(2): 146-152.]
- [38] Anselin L. The future of spatial analysis in the social sciences. *Geographic Information Sciences*, 1999, 5(2): 67-76.
- [39] Wilby R L. Non-stationarity in daily precipitation series: implications for GCM down-scaling using atmospheric circulation indices. *International Journal of Climatology*, 1997, 17: 439-454.
- [40] Simon H A. The architecture of complexity. *Proceedings of the American Philosophical Society*, 1962, 106: 467-82.
- [41] Young G L. Hierarchy and central place. *Geographical Analysis*, 1978, 60(2): 71-78.
- [42] O'Neill R V, D DeAngelis. *A Hierarchical Concept of Ecosystems*. Princeton: Princeton University Press, 1986.
- [43] Marceau D J, Hay G J. Scaling and modelling in forestry: applications in remote sensing and GIS. *Canadian Journal of Remote Sensing*, 1999, 25(4): 342-346.
- [44] Wu J. From balance-of-nature to hierarchical patch dynamics. *Quarterly Review of Biology*, 1995, 70: 439-466.
- [45] Lu Xuejun, Zhou Chenghu. Analytical scheme on scale-structure of geographical space. *Progress in Geography*, 2004, 23(2): 107-114. [鲁学军, 周成虎. 地理空间的尺度, *地理科学进展*, 2004, 23(2): 107-114.]
- [46] Wang Jiayao, Cheng Yi. The multiple character of spatial data and automated generalization. *Hydrographic Surveying and Charting*, 2004, 24(4): 1-3. [王家耀, 成毅. 空间数据的多尺度特征与自动综合. *海洋测绘*, 2004, 24(4): 1-3.]
- [47] Wu Hehai. Research on fundamental theory and technical approaches of automated map generalization. *Journal of Wuhan Technical University of Surveying and Mapping*, 2000, 25(5): 377-386. [毋河海. 地图信息自动综合基本问题研究. *武汉测绘科技大学学报*, 2000, 25(5): 377-386.]
- [48] Gao Wenxiu. Thematic data generalization in GIS. *Geomatics and Information Science of Wuhan University*, 2002, 27(5): 505-510. [高文秀. GIS 中专题属性数据综合的若干问题. *武汉大学学报·信息科学版*, 2002, 27(5): 505-510.]
- [49] Chen Jun, Zhao Renliang, Qiao Chaofei. Voronoi diagram-based GIS spatial analysis. *Geomatics and Information Science of Wuhan University*, 2003, 28(Special Issue): 32-37. [陈军, 赵仁亮, 乔朝飞. 基于 Voronoi 图的 GIS 空间分析研究. *武汉大学学报·信息科学版*, 2003, 28 (特刊): 32-37.]
- [50] Wu Fan. Scaleless representations for polyline spatial data based on wavelet analysis. *Geomatics and Information*

- Science of Wuhan University, 2004, 29(6): 488-491. [吴凡. 基于小波分析的线状特征数据无级表达. 武汉大学学报·信息科学版, 2004, 29(6): 488-491.]
- [51] Chen Jun. Developing dynamic and multi-dimensional geo-spatial data framework. *Geo-Information Science*, 2002, 1: 7-13. [陈军. 多维动态地理空间框架数据的构建. 地球信息科学, 2002, 1: 7-13.]
- [52] Li Deren. From digital map to spatial information multi-grid. *Geomatics and Information Science of Wuhan University*, 2003, 28(6): 642-649. [李德仁. 从数字地图到空间信息网格. 武汉大学学报·信息科学版, 2003, 28(6): 642-649.]
- [53] Yin Lianwang, Li Jing. A technique on automatic compression and recovery of fundamental GIS spatial data on varying map scale. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 1999, 35(6): 842-849. [尹连旺, 李京. GIS 中基本要素的无级比例尺数据处理技术研究. 北京大学学报(自然科学版), 1999, 35(6): 842-849.]
- [54] Quattrochi D A, Goodchild M F. *Scales in Remote Sensing and GIS*. Boca Raton, FL: CRC/Lewis Publisher, 1997.
- [55] Tate Nicholas J. *Modelling Scale in Geographical Information Sciences*. Chichester: John Wiley & Sons, Ltd., 2001.
- [56] Bo Yanchen, Wang Jinfeng. *Uncertainty in Remote Sensing: Classification and Scale Effect Modeling*. Beijing: Geological Press, 2003. [柏延臣, 王劲峰. 遥感信息的不确定性研究: 分类与尺度效应模型. 北京: 地质出版社, 2003.]
- [57] Woodcock C E. The factor of scale in remote sensing. *Remote Sensing of Environment*, 1987, 21: 311-332.
- [58] Su Lihong, Li Xiaowen, Huang Yuxia. A review on scale in remote sensing. *Advance in Earth Sciences*, 2001, 16(4): 544-548. [苏理宏, 李小文, 黄裕霞. 遥感尺度问题研究进展. 地球科学进展, 2001, 16(4): 544-548.]
- [59] Xia Zongguo, Clarke K C. Approaches to scaling of geo-spatial data. In: Quattrochi D A, Goodchild M F, (eds.), *Scales in Remote Sensing and GIS*. Boca Raton, FL: CRC/Lewis Publisher, 1997. 309-360.
- [60] Quattrochi D A, Emerson C W. Fractal characterization of multitemporal remote sensing data. In: Tate Nicholas J, *Modelling Scale in Geographical Information Sciences*. Chichester: John Wiley & Sons, Ltd., 2001. 13-34.
- [61] Wu Fan. Scaleless representations for polyline spatial data based on wavelet analysis. *Geomatics and Information Science of Wuhan University*, 2004, 29(6): 488-491. [吴凡. 基于小波分析的线状特征数据无级表达. 武汉大学学报·信息科学版, 2004, 29(6): 488-491.]
- [62] Zhao Xi'an, Li Deren. Edge focusing based on multiscales wavelet transform. *Journal of Remote Sensing*, 2003, 7(4): 299-303. [赵西安, 李德仁. 小波多尺度影像边缘聚焦算法. 遥感学报, 2003, 7(4): 299-303.]
- [63] Bi Weiqiang. Multisensor data fusion and its application to global change. *Earth Science Frontiers*, 2002, 9(2): 387-391. [毕巍强. 全球变化研究中的数据融合. 地学前缘, 2002, 9(2): 387-391.]
- [64] Zhang Hao, Jiao Ziti, Yang Hua et al. Research on scale effect of histogram. *Science in China (Series D)*, 2002, 45(10): 949-960. [张颢, 焦子铤, 杨华等. 直方图尺度效应研究. 中国科学(D), 2002, 32(4): 307-315.]
- [65] Gehlke C E, Biehl K. Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 1934, 29(suppl.): 169-170.
- [66] Openshaw S, Taylor P. A million or so correlation coefficients: three experiments on the modifiable area unit problem. In: Wrigley N (ed.), *Statistical Applications in the Spatial Sciences*, London: Pion, 1979. 127-144.
- [67] Flowerdew R, Openshaw S. A review of the problems of transferring data from one set of areal units to another incompatible set. RR4, Northern Regional Research Laboratory, University of Newcastle Upon Tyne, England, 1987.
- [68] Kolaczyk E D, Huang H. Multiscale statistical models for the hierarchical spatial aggregation. *Geographical Analysis*, 2001, 33(2): 95-118.
- [69] Mugglin A S, Carlin B P. Hierarchical modelling in geographic information systems: population interpolation over incompatible zones. *Journal of Agricultural Biological and Environmental Statistics*, 1998, 3: 111-130.
- [70] Sadahiro Y. Accuracy of count data estimated by the point-in-polygon method. *Geographical Analysis*, 2000, 32(1): 64-89.
- [71] Flowerdew R, Green M. Statistical methods for inference between incompatible zonal systems. In: Goodchild M F, Gopal S (eds.), *Accuracy of Spatial Databases*. London: Taylor and Francis, 1989. 239-248.
- [72] Gallego J. Using land cover information to map population density, 2001.
<http://www.unece.org/stats/documents/2001/09/gis/21.e.pdf>
- [73] Turner A, Openshaw S. Disaggregative spatial interpolation.
<http://www.geog.leeds.ac.uk/staff/a.turner/papers/gisruk01/gisruk01.htm>
- [74] Zhu Huiyi. Problems of the spatial interpolation of physical geographical elements. *Geographical Research*, 2004, 23(4): 425-432. [朱会义. 自然地理要素空间插值的几个问题. 地理研究, 2004, 23(4): 425-432.]
- [75] Deichmann U, Balk D, Yetman G. Transforming population data for interdisciplinary usages: from census to grid, 2001. <http://sedac.ciesin.org/plue/gpw/GPWdocumentation.pdf>
- [76] Goodchild M F. Areal interpolation: a variant of the traditional spatial problem. *Geo-Processing*, 1980, (1): 297-312.
- [77] Flowerdew R, Green M. Areal interpolation and types of data. In: Fotheringham A S, Rogerson P A (eds.), *Spatial*

- Analysis and GIS. London: Taylor and Francis, 1994. 121-145.
- [78] Sadahiro Y. Accuracy of count data transferred through the areal weighting interpolation method. *International Journal of Geographical Information Science*, 2000, 14: 25-50.
- [79] Gregory I N. An evaluation of the accuracy of the areal interpolation of data for the analysis of long-term change in England and Wales, 2000. <http://www.geocomputation.org/2000/GC045/Gc045.htm>
- [80] Tobler W A. Smooth pycnophylactic interpolation for geographical regions. *J. of American Statistical Association*, 1979, 74: 519-529.
- [81] Tobler W A, Deichmann U, Gottsegen J et al. The global demography project. In: NCGIA, Technical Report 95-6, Santa Barbara, 1995.
- [82] Flowerdew R, Green M. Statistical methods for inference between incompatible zonal systems. In: Goodchild M F, Gopal S (eds.), *Accuracy of Spatial Databases*. London: Taylor and Francis, 1989. 239-248.
- [83] Gallego J. Using land cover information to map population density, 2001. <http://www.unece.org/stats/documents/2001/09/gis/21.e.pdf>
- [84] Liu Jiyuan, Yue Tianxiang, Wang Ying'an et al. Digital simulation of population density in China. *Acta Geographica Sinica*, 2003, 58(1): 17-24. [刘纪远, 岳天祥, 王英安 等. 中国人口密度数字模拟. *地理学报*, 2003, 58(1): 17-24.]
- [85] Rao J N K. Some recent advances in model-based small area estimation. *Survey Methodology*, 1999, 24: 175-186.
- [86] Simpson L. Geography conversion tables, a framework for conversion of data between geographical units. *International Journal of Population Geography*, 2002, 8(1): 69-82.
- [87] Ghosh M, Rao J N K. Small area estimation: an appraisal (with discussion). *Statistical Science*, 1994, 9: 55-93.
- [88] Gonzalez M E. Use and evaluation of synthetic estimates. In: American Statistical Association. *Proceedings of the Social Statistics Section*, 1973. 33-36.
- [89] World Bank. Sources of small area data: indirect estimation of poverty indicators. www.worldbank.org/poverty/inequal/povmap/data4.htm. 2003.
- [90] Söstra K. Small area estimation methods. <http://www.stat.vil.ee/pks/analyys/jurmala.doc>
- [91] Heimsath C H. Small-area population estimation in absorption analysis. *The Journal of Real Estate Research*, 1991, 6 (3): 315-326.
- [92] Rao J N K. *Small Area Estimation*. Hoboken: John Wiley & Sons, Inc., 2003.
- [93] Gonzalez M E, Placek P J, Scott C. Synthetic estimation of followback surveys at the national center for health statistics. In: Schaible W L, *Indirect Estimators in U.S. Federal Programs*. New York: Springer Verlag, 1996. 16-27.
- [94] Datta G S, Ghosh M, Nangia N et al. Estimation of median income of four-person families: a Bayesian approach. In: Berry D A et al. (eds.), *Bayesian Analysis in Statistics and Econometrics*. Wiley, 1996. 129-140.
- [95] Rao J N K, Yu M. Small area estimation by combining time series and cross-sectional data. In: American Statistical Association, *Proceedings of the Survey Statistics Section*, 1992. 1-19.
- [96] Cressie N. Small-area prediction of undercount using the general linear model. *Proceedings of Statistics Symposium 90: Measurement and Improvement of Data Quality*. Statistics Canada, 1991. 93-105.

A Review on the Methodology of Scaling with Geo-Data

MENG Bin^{1,2,3}, WANG Jinfeng¹

(1. *Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;*

2. *College of Arts and Science of Beijing Union University, Beijing 100083, China;*

3. *Graduate School of Chinese Academy of Sciences, Beijing 100039, China)*

Abstract: The ability to change scale is one of the main development topics in GIS and RS. GIS and image processing systems have offered some tools for multi-scale approaches. But the capabilities are limited. Scaling of multi-scale Geo-data is one of the suitable methods to deal with this problem. In this paper, the detailed review was done by comparing different methods in different research areas. Areal interpolation is the most important way to solve the scaling problems in GISci. And sociologists and economists have developed the small area statistics to solve the problem of the changing of the statistical units. And the review shows that there are some similar principles in different methodologies, and each one has its special pivots.

Key words: scale; scaling; attribute data; areal interpolation; small area statistics