

城市形态的分维估算与分形判定

陈彦光

(北京大学城市与环境学院,北京 100871)

摘要:城市形态的分形是城市发育到一定阶段涌现出的有序格局和复杂结构,其基本特征是空间分布的无尺度性质。当研究者基于某个显著性水平推断城市分维存在时,实际上就是基于相应的置信度判断分形特征。虽然分形城市研究已经多年,但大量有关维数测算的基础问题依然悬而未决。本文根据分形几何学的基本思想论证城市形态分维测算的若干问题。分维测量的准则是最佳覆盖——不多不少、恰到好处覆盖。盒子覆盖是最容易理解的测量方法。采用盒子覆盖法测量城市形态分维时,应考虑三个标准:一是快速逼近,二是简便操作,三是稳定拟合。直观估计分维的办法是利用双对数坐标图。由于城市形态不是严格意义的分形,而是类似于文献中的“前分形”,测量尺度与相应测度的幂律关系通常仅在一定尺度范围内有效,从而形成所谓标度区。本文围绕城市形态的分维测量和分形判断开展一系列讨论,包括尺度选取、标度区识别和统计标准等问题,对今后城市分形研究具有理论启示和方法论的参考价值。

关键词:城市形态;分形;分维;自仿射;盒子计数法;多分形;标度区

1 引言

分维的本质是一个特征指数,是在无特征尺度定量描述的研究对象中找到的一个有特征的值。利用分维可将众多的地理空间数据浓缩为一个简单的数字,据此揭示城市背后隐含的时空信息。自1985年开始,分形城市研究至今已有30多年,理论、方法和实证分析都有很大进步(Batty et al, 1994; Frankhauser, 1994, 1998; White et al, 1994; 刘继生等, 1998, 1999a, 1999b, 2003; De Keersmaecker et al, 2003; 姜世国等, 2006; Thomas et al, 2007, 2008; 陈彦光, 2008; Feng et al, 2010; Ariza-Villaverde et al, 2013; Chen et al, 2013; Murcio et al, 2015; 秦静等, 2015)。然而,许多基本问题尚未解决,一些问题在理论界不成问题,但就经验应用而言知者甚少。这些问题包括:测量尺度问题,标度区的识别问题,城市分形的判定问题,统计分析标准的确定问题等。分维的测算通常采用几何尺度,即测量尺度以指数的方式衰减,相应的观测数目以

指数的方式递增。于是,在双对数坐标图中,数据点不会太多,很少有超过15个样本点。因此有人怀疑,样本是不是太小?是否可以改用算术尺度增加样本点?对于分形机理不太熟悉而对统计学分析又一知半解的初学者大多难免产生诸如此类的猜想。由于城市空间格局不是数学意义的分形,双对数坐标图中的数据点通常并非形成一条直线,而是局部出现直线段即所谓标度区,于是又产生了标度区上、下界限的识别问题。城市分维测量是一种统计分析,统计推断没有绝对而截然的结论,那么怎样判断一个城市的形态具有分形性质呢?虽然有学者提出初步的判断标准(Benguigui et al, 2000; Feng et al, 2010),但至今尚未形成定论。这类问题如不澄清,会对今后的城市分形研究带来很多困惑、误会乃至混乱。

近10多年来,不断有学生、同行咨询有关城市分维测量方法,上述问题是众多研究者共同反映的基本问题。但今天看来,这些问题已不再是难题了,因为有条件给出令人信服的答案。本文的写作

收稿日期:2016-05;修订日期:2017-03。

基金项目:国家自然科学基金项目(41671167)[Foundation: National Natural Science Foundation of China, No. 41671167]。

作者简介:陈彦光(1965-),男,河南罗山人,教授,博士,从事城市理论和理论地理学研究,E-mail: chenyg@pku.edu.cn。

引用格式:陈彦光. 2017. 城市形态的分维估算与分形判定[J]. 地理科学进展, 36(5): 529-539. [Chen Y G. 2017. Approaches to estimating fractal dimension and identifying fractals of urban form[J]. Progress in Geography, 36(5): 529-539.]. DOI: 10.18306/dlkxjz.2017.05.001

目的,就是根据作者多年的理论和实证研究经验提供自己的回答。不过,要理解这些答案,可能需要相应的知识结构——只有知识结构相似的学者之间才能进行有效的交流。任何一个学术概念都不是孤立的,其背后涉及一整套概念体系,如果不了解这套概念体系,就容易因为断章取义而望文生义,进而形成误解。城市形态是一种空间现象,空间上的分维测量的基本方法之一是盒子计数法(Benguigui et al, 2000; Shen, 2002)。本文将盒子测量法为基础,将有关理论、方法和技术问题逐步展开,依次说明。

2 分形的特性和分维测算的思想

2.1 分形的结构性特征

提到分形的时候,大多学者会想到Feder与Mandelbrot通信中的那个定义:“分形是由以某种方式与整体相似的部分构成的一种形体”(Feder, 1988)规则分形的显著特征是自相似性,随机分形的基本特征是破碎和不规则性。支离破碎、不规则仅仅是表象,随机分形的背后则是统计自相似性。换言之,随机分形表面的无序后面隐含着有序的深层结构。自相似性在数学上的表现则是无尺度性,或者叫做尺度不变性(scale invariance)、标度对称性(scaling symmetry)。上述概念容易理解,但仅掌握上述概念,不足以深刻认识分维的测量方法。要想真正理解分维的测量思路,至少要了解分形的如下性质。

其一,分形是一个具有递阶结构(cascade structure)的等级体系(hierarchy)。“飞流直下三千尺”的庐山瀑布是没有递阶结构的。不妨想象一条比庐山三叠泉更多层次的瀑布,这条瀑布从悬崖上跌落一定距离,遇到岩石阻挡分为两条较小的瀑布;这两条较小的瀑布下落一定距离,又遇到岩石妨碍各自分解为两条更小的瀑布;这四条更小的瀑布下降一定尺度,遇到岩石阻碍再次分开成为八条小瀑布……。如此层层分解,表现为递阶结构;分解的最终结果,则形成一个瀑布的等级体系。分形体正是这样的等级体系。一个分形体由 N 个下级分形单元构成,每个分形元的尺度是分形体的 $1/k$;这 N 个下级分形元又由 N^2 个更小的下级分形元构成,每个小分形元的尺度是分形体的 $1/k^2$ 。到第 m 级,分形元的数目变成 $N_m=N^{m-1}$ 个,尺度则下降为 $r_m=1/k^{m-1}$ 。

于是分维为 $D=\ln(N_{m+1}/N_m)/\ln(r_m/r_{m+1})=\ln(N)/\ln(k)$,这就是所谓的相似维数。现实的城市结构是递阶的等级体系(Kaye, 1989; Batty et al, 1994; Frankhauser, 1998; 陈彦光, 2008; Chen et al, 2013),只不过是这种自相似的等级体系被随机性和复杂性掩盖。通过分维测算,可将隐藏在无序表象背后的标度秩序揭示出来。

其二,分形体与欧氏几何体存在“对偶”关系。对偶原本是一个线性规划建模的概念,指代调换目标函数和约束条件的模型表达。实际上,所有转换目的与手段的计算或建模关系都是对偶关系。欧氏几何学中基于维数测量面积和分形几何学中基于面积测量维数的关系属于对偶转换关系。一个欧氏几何体的维数不测可知(点为0维,线为1维,面为2维,体为3维),但其大小(长度、面积、体积)需要测量才能知道;一个分形体的大小理论上不测可知(长度、面积、体积为0或者无穷大),但其维数需要测量才能明确。欧氏几何体的测度(长短、大小)容易理解,分形几何体的测度举例如下:Cantor集的长度为0,Koch曲线的长度为无穷大,Sierpinski垫片的面积为0,但内边界线为无穷大,Menger海绵的体积为0。上述对偶性质对理解分形和分维非常重要:我们可将分形体视为一个欧氏几何体,测量它的大小即长度、面积或体积。这种测量没有直接意义——得不出长度、面积或者体积,但有间接结果:基于测量过程建立标度关系,利用标度指数估计分维数值。换言之,分维是对分形体长度、大小测量的间接成果。

其三,分维是分形体的特征参数。定量描述的关键在于找到特征尺度,或称特征长度(characteristic length)。只要量出特征长度,一个研究对象的数字特征就清楚了。欧氏几何体都是有特征长度的,圆有半径,正方形、矩形有边长,三角形、梯形、平行四边形有底边和高……。有了特征长度,就可以计算周长、面积、体积,从而研究对象数字信息就揭示出来。至于它们的维数,如前所述,无非0、1、2、3,不测可知,没有信息可言。而分形体就不同,一方面,它们没有特征尺度,故不能采用长度、大小来描述。改变测量尺度,测出的长度、面积、体积将会随之改变,亦即测量结果存在尺度依赖性;另一方面,分形体的维数却具有两方面的特征:一是具有空间信息,不测量不能知道维数的数值;二是具有特征尺度,分维特别高的城市和特别低的城市都很少,

绝大多数城市形态的分维围绕某个期望值上下波动。以城市形态为例,在二维空间中,分维通常围绕1.7变化(Batty et al, 1994),大于1.9和小于1.4的城市形态比较罕见。正因为海岸线没有一个确定的长度,无法找到一个特征长度来刻画它。但是,海岸线的维数既具有信息又有特征尺度,Mandelbrot (1967)因此建议采用分维这个参数来描述海岸线。对于一段海岸线,无法有效确定它的绝对长度(不依赖于尺度的长度),但可以肯定的是,分维越高的海岸线越是曲折,从而相对长度(相应于给定尺度的长度)也就越长。

这里涉及两个基本知识:一是信息论的知识。信息是对不确定性的消除。如果一个系统的某个测度或者参数不经测量就事先知道,则这个测度或者参数没有信息。信息出现的前提是不确定性——未经观测不可能知道确切的结果。二是复杂性的知识。这涉及有特征尺度分布和无特征尺度分布两个概念。有特征尺度分布通常是中庸型分布:中间高、两头低,高者多而低者少;无特征尺度分布通常是极端型分布:一端高、一端低,高者少而低者特多(陈彦光, 2015)。

2.2 分维测算的数理思想

认识分形的结构性特征之后,还必须了解分维测算的基本数理思想。只有理解分维测算的本质,才能把握测算过程的处理技巧。

其一,分维概念来自几何测量。如前所述,一个分形体的长度、面积、体积理论上是不测可知的,但还是可以像对待欧氏几何体那样测量它们的大小。这种测量对于长度、面积、体积而言当然不会产生有效结果,但我们需要的不是直接的结果,而是测量过程。利用测量过程建立幂指数关系,据此估计分维。举例说来,如果想要知道一个矩形桌面的大小,却没有测量的尺子,就可以采用已知大小的扑克牌来覆盖它。不言而喻,这个覆盖过程要满足如下条件才能尽可能得到接近真实的面积:①扑克牌没有重叠,也没有空隙;②桌子的边沿没有裸露,扑克牌也没有超越桌子的边沿。简而言之,采用不多不少、不大不小的覆盖方式,这个覆盖叫做“最佳覆盖”。如果满足上述条件,则扑克牌的大小乘以扑克牌的数量,就是桌面的面积。假定扑克牌的长、宽为 a 、 b ,最佳覆盖需要的扑克牌数量为 N ,则桌子的面积为 $A=Nab$ 。由于扑克牌的长宽之比是固定的,不妨设 $a/b=1/h$,即有 $b=ha$,这里 h 为比例

系数,于是 $A=Nb^2$ 。这个式子给出了3个量的关系:桌面的面积 A ,桌面的维数 $d=2$,以及扑克牌的线性尺度 b (特征长度)。在现实中,不可能一次性地找到一种扑克牌,其大小刚好满足上述关系,从而 N 不可能是整数。换言之,采用整数的扑克牌不大可能一次性地完美无缺地覆盖整个桌面,从而得到准确的桌面面积。因此,上述关系应取近似表达,即 $A \approx Nb^2$ 。一种解决的办法是,将扑克牌均等地一分为四,于是各个扑克牌的线性尺度变成原来的 $1/2$ 。采用这种小扑克牌来覆盖,得到的结果 Nb^2 更为接近桌面面积 A 。进一步地,将小扑克牌再次一分为四,则其线性尺寸变成原来的 $1/4$ 。随着扑克牌的不断缩小,覆盖结果越来越接近完美覆盖。也就是说,当 b 趋近0的时候,可以得到 $A=Nb^2$ 。显然,在上述测量过程中,桌面的面积 A 其实是一个常数,而扑克牌的数量 N 却是一个变量。于是上面的式子可以写作 $N=Ab^{-2}$ 。更一般地,这个关系表示如下:

$$N(b)=Ab^{-d} \quad (1)$$

式中: $d=2$ 为桌面的维数。

常见欧氏几何体如桌面的测量,有助于我们从“对偶”的视角理解分形测量过程和目的。进一步设想,测量的对象不是欧氏几何体的桌面,而是图1所示的Vicsek分形体(Jullien et al, 1987; Vicsek, 1989; Batty et al, 1994; Chen et al, 2013)——该分形常被用于隐喻城市生长——那又如何?理论上,分形体的面积为0(图中的分形无限细化之后,基本单元最终为无数个点)。但是,还是可采用正方块覆盖来测量它的“面积”。第一级,需要一个边长为1的正方块;第二级,需要5个边长为 $1/3$ 单位的正方块;第三级,需要25个边长为 $1/9$ 单位的正方块;第 m 级,需要 $N_m=5^{m-1}$ 个边长为 $r_m=1/3^{m-1}$ 单位的正方块。由于分形体的面积为0,这个过程可以无限度地进行下去。于是得到如下负幂律关系

$$N_m=N_1r_m^{-D} \quad (2)$$

式(2)可以等价地表作如下标度关系:

$$N(r)=N_1r^{-D} \quad (3)$$

式(3)中:比例系数 $N_1=1$ 为第一级非空盒子个数, D 表示分维,且有

$$D=-\frac{\ln(N(r))}{\ln(r)}=-\frac{\ln(N_{m+1}/N_m)}{\ln(r_{m+1}/r_m)}=-\frac{\ln(5^{m-1})}{\ln(1/3^{m-1})}=\frac{\ln(5)}{\ln(3)} \approx 1.465 \quad (4)$$

最终测量的“面积”为 $A=N(r)r^D=N_1=1$ 单位,而真实

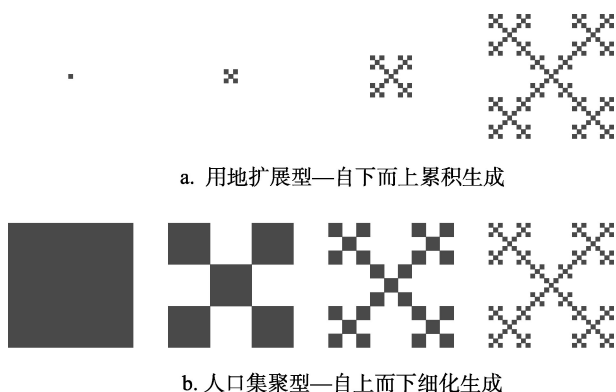


图1 一种规则生长分形的两种生成方法(前四步)
(根据 Jullien et al, 1987 修改)

Fig.1 Two approaches to generating a regular growing fractal that is often employed to model urban growth (the first four steps)
(modified from Jullien et al, 1987)

的分形面积为0,这是一个矛盾的结果。矛盾结果反证分形体的面积是不可测的,但我们要的不是结果,而是过程,因为通过这个过程,可得到一个具有特征尺度的参数,那就是分维 $D \approx 1.465$ 。

其二,分维的测量效果与覆盖效果正相关。如上所述,可以将一个分形体当作欧氏几何体来测量,通过测量长度、面积、体积之类的过程寻找分维数值。测量方式有多种,多快好省的测量过程以最佳覆盖为前提。前述 Vicsek 分形的维数测算,一个方块相当于一个二维盒子,每一个级别的盒子覆盖都是不多不少、不大不小恰到好处(图2)。另外,也

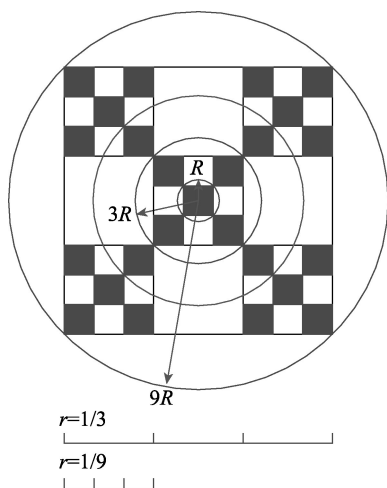


图2 一种规则生长分形的两种分维测量方法(前三步)
(根据 Batty et al, 1994 修改)

Fig.2 Two approaches to estimating fractal dimension of a regular fractal (the first three steps)
(modified from Batty et al, 1994)

可以采用面积—半径标度即半径法(radial method)来测量分维(Batty et al, 1994; Frankhauser, 1994)。第一个半径 $R=1$ 单位,将核心的1个分形单元盖住,面积为1单位;第二个半径 $R=3$ 单位,将中央的5个分形单元盖住,面积为5单位;第三个半径 $R=9$ 单位,将中央的25个分形单元盖住,面积为9单位;第 m 个半径 $R_m=3^{m-1}$ 单位,将围绕中心的 $N_m=5^{m-1}$ 个分形单元盖住,面积为 $A_m=N_m=5^{m-1}$ 单位。这个过程可以一直进行下去。于是得到如下正幂律关系:

$$A_m = A_1 R_m^D \quad (5)$$

式(2)可以等价地表作如下标度关系:

$$A(R) = A_1 R^D \quad (6)$$

式中:比例系数 $A_1=1$ 为第一分形单元的面积, D 表示维数,这个维数叫做径向维数(radial dimension)(Frankhauser, 1994),且有

$$D = \frac{\ln(A_{m+1}/A_m)}{\ln(R_{m+1}/R_m)} = \frac{\ln(5)}{\ln(3)} \approx 1.465 \quad (7)$$

可见,对于规则分形,如果方法得当,辐射维数与盒子维数相等。不过,最终测量的“面积(实际上是分形单元数目)”为无穷大。这个面积计量结果与分形的真实面积矛盾,也与前面的盒子法测量结果矛盾:不同的方法给出的面积是不一样的。这再次反证了分形体大小的不可测性质。

2.3 从规则分形到随机前分形

前面给出的分形例子是规则分形(regular fractal)。一个分形的三要素包括形态、机遇和维数(Mandelbrot, 1977, 1983)。如果在分形的生成过程中引入机遇因子,就会形成随机分形。以前述生长分形为例,每一步都是将一个分形单元一分为九,然后去四留五。对于规则分形,总是去掉上下左右四个部分,留下中央和左上、右上、左下、右下四个部分(图1)。如果去掉和保留的部分由随机决定,则最终得到随机分形(Vicsek, 1989),而随机分形可以更好地模拟城市系统的结构和形态(陈彦光, 1998)。然而,数学意义的随机分形与城市形态还是有区别的。数学世界的分形,无论规则分形还是随机分形,都是无穷层次的。城市形态的层次不可能达到无穷。在一幅电子地图上,城市形态的分辨率是有限度的,从而层次存在限度。实际上,即便是计算机生成的理论意义分形图形,也不可能是无穷层次的,无穷层次的分形只能存在于想象中的数学世界,而不会出现在现实世界。因此,人们所见的分形,不论是现实的,抑或是模拟的,都是所谓前

分形(prefractal)——具有有限层次的自相似形体(Addison, 1997)。在这个意义上,城市形态可以看作一种随机前分形。

对于随机前分形来说,在分维测量过程中不可能找到所谓最佳覆盖,只能采用某种方法逐步逼近。因此,一方面,测量的前几步由于尺度太大,测量体与分形体吻合不好,测量尺度与相应测度的标度关系不能很好地反映分维数值;另一方面,当测量的尺度足够小时,就会达到分辨率的极限,或者超越分形最小单元的界限,从而尺度与相应测度的标度关系失效。这样,只有尺度不是太大也不是太小的时候,测量尺度与相应测度的标度表现才会给出有效的分维估计结果,这就涉及后面讨论的标度区问题。

3 分维测量中的尺度与标度

3.1 算术尺度与几何尺度

前述两种测量方法即盒子法和半径法是显著不同的。盒子法是自上而下的测算,标度关系表现为负幂律;半径法是自下而上的测量,标度关系表现为正幂律(图2)。尽管如此,分维测算结果是一致的。两种测量方法的共性为:①采用不多不少的最佳覆盖。②盒子或半径尺度取几何递减或几何递增数列,相应的非空盒子数或者面积计量结果也表现为几何递增数列。简而言之,测量尺度与递阶结构的等级体系形成对应的标度关系。几何数列连续化的结果是指数函数:尺度以指数的方式衰减或者增长,数目或者面积以指数的方式上升。在城市形态分维测量过程中,采用几何数列会很快达到分辨率的极限,给出的数据点通常是10个左右,很难达到15个数据点。级别过多,目前的计算机无法处理。于是有人提出疑问,10个左右的数据点,样本是不是太小了?是否有必要将测量尺度由几何数列改为算术数列?

这类问题是很多初步涉及分形理论、并且懂得统计学知识的人都会想到的。实际上,在分维测量过程中,既可以采用几何测度,也可以采用算术测度。但是,在绝大多数情况下,算术尺度并不优于几何尺度。如果为了扩大双对数坐标图中的数据点,则没有必要施行这种转换。分维测量的基准是最佳覆盖,因此需要满足以下三方面条件:一是测量尺度要与分形结构对应,二是测量过程要方便操

作,三是数据点不能太少。上述三方面条件相互制约,可进一步阐述如下:①分形结构。几何尺度与等级体系的递阶结构一致。如前所述,分形体系是一个递阶结构的等级体系。一个分形体的结构描述可以采用一个幂函数描述,也可采用两个指数函数描述,而指数函数对应的就是几何数列(等比数列)。只有等比数列才能与分形体的真实结构吻合,从而接近所谓的最佳覆盖。②逼近效果。对于分维估计,盒子尺度的迅速变小及其对测量对象单元的快速逼近要比数据点数更重要。只有当盒子迅速变小的时候,盒子数目与其大小的乘积才接近于测量对象的某个测度(长度、面积或者体积);只有当盒子数目与其大小的乘积接近于几何体的测度时,分维估计值才接近真实值。几何数列不仅逼近速度快,而且容易捕捉等级体系的递阶结构特征。③去噪功能。几何数列可以过滤测量过程中数据提取的噪声。采用等比数列估计分维,虽然降低了数据点数,同时也将大量随机变化引起的随机干扰滤除。采用等差数列虽然增加了数据点数,但同时也受到随机噪声更大干扰。④操作方便。不同的处理方式具有不同的操作效果,越是简单、便捷的测量方法越是可取。最方便的操作是将盒子的边长一分为二、二分为四、四分为八,即以二倍数的方式划分。一分为四、一分为八与一分为二等价,一分为三、一分为五、一分为七就不方便操作了。

以二维空间中的盒子法为例,分维的测量效果与盒子覆盖对分形体的“面积”逼近效果正相关。逼近效果越好,分维测量效果越好。对于规则分形,容易找到最佳覆盖方式,从而可以通过有限步骤的测量准确计算分维。但是对于随机分形,情况就不同了,无法找到最佳覆盖,只能通过不断细分逐步逼近。理论上,当盒子尺度趋于无穷小时,可望实现最佳逼近,从而最好地估计分维。正因为如此,分维测量公式中通常有一个极限符号 \lim ,表示尺度 r 趋于无穷小。例如式(4)可以规范地表示为:

$$D = -\lim_{r \rightarrow 0} \frac{\ln(N(r))}{\ln(r)} \quad (8)$$

那么,既然追求快速逼近,是否可干脆将盒子边长一分为十、十分为百,那很快盒子就趋于无穷小了。问题在于:快速逼近仅仅分维估计效果的条件之一,其他条件包括分形结构、样本大小和操作方便。其实,log-log坐标图中的数据点不是严格意义的样本大小,真正的样本是研究对象中包括的像

元数目(对于城市形态)或者城市数目(对于城市体系)。当然,数据点数与统计分析的自由度有关。尽管如此,一个模型的参数估计效果,不在于测量多么精确,而在于测量结果的无偏性;一个参数的估计值,不在于样本大小,而在于估计结果的置信度高低和误差范围的大小。

3.2 双对数坐标图与标度区

分形是一种分布广泛的具有标度性质的现象。所谓标度,就是具有时空伸缩变换下的尺度不变性,而这种不变性又导致测量上的尺度依赖性。标度是目前科学界的前沿概念之一,代表数学建模的新理念。地理学家最应认识和理解标度,因为地理绘图的过程就是典型的建模过程,而一幅地图按照一定的比例尺连续缩放而形态不变的过程就是典型的标度过程。标度关系在经验上是有尺度范围的:尺度太大或者太小,标度关系就会破坏(Bak, 1996; 陈彦光等, 2007)。在log-log坐标图上,测量尺度(线性尺寸)与相应测度(长度、面积、体积、数量等)之间的关系很少形成一条直线,通常在中间一段表现为直线段,这个直线段叫做标度区(scaling range)或者无尺度区(scale-free range),过去常翻译为无标度区。标度区的成因有以下几方面:一是测量方法导致。由于无法找到随机分形的最佳覆盖,当盒子尺度太大时,盒子与分形体的匹配关系不好,从而分维逼近效果不佳;当盒子尺度小到一定程度,盒子与分形体形成更好的匹配。这样,双对数坐标上不是一条直线,而是两条以上的直线段。二是分形体自身特性。严格意义的分形仅存于数学世界,不属于现实世界。即便采用计算机生成分形图形,得到的结果也不是真正意义的分形,而是一种前分形。真正的分形是无穷层次、基本单元无穷小的,计算机生成的分形和现实中的分形都是在有限尺度范围内像分形。因此,当尺度小到一定程度之后,就会达到分辨率的极限。如果继续测量下去,尺度与相应测度的关系就不会一如既往地保持先前的比例。正是由于第一个原因,形成标度区的上限(对应于大尺度);由于第二个原因,形成标度区的下限(对应于小尺度)。在动力学上,标度区与广义的空间关联过程有关。任何一个分形模型,理论上都是一种特定的关联函数。如果尺度太小,则没有关联对象;如果尺度太大,则失去显著的关联效果。因此,只有中间尺度形成有效的空间关联。对于理论上的规则分形,标度区没有局限,整个双对

数图表现为一条直线。然而,对于现实中的、被人们视为分形的东西,如城市、海岸线,一定存在有限的标度区的问题。在研究城市形态时,标度区更是不可避免的概念。问题在于,由于测量困难,人们处理的尺度往往没有达到极限,从而看不到显著的标度区边界。

如果采用盒子法测量城市形态的分维,标度区相对容易确定。当盒子尺度较大时,不存在非空盒子,对应于整数维,这是嵌入空间的欧氏维数;当盒子尺度逐步变小时,非空盒子出现了,第一次出现非空盒子时,可作为标度区的起点和上限。由于现实研究对象如城市通常是前分形,而不是数学意义的分形,当盒子尺度小到一定程度之后,非空盒子数增加速度会突然变小,从而在log-log坐标图中会出现突然的转折,转折后线段的斜率理论上对应于拓扑维数。一个分形的盒子维数值介于分形体的拓扑维数和其嵌入空间的欧氏维数之间。对于有经验的研究者而言,这第二个转折点凭目测即可判断。即便不能准确地断定分界点也没有关系。确定标度区的范围不是目的,目的是估计分维。将标度区少算几个观测点原则上不会显著影响分维估计结果,因而也就没有必要准确无误地找到分界点。

然而,对于面积一半径标度法,情况就很不相同。基于半径法的标度区的上、下限通常都是比较模糊的,难以找到客观的分界点,而且改变标度区的范围会对分维估计值形成明显影响。不少人曾经提出一些客观确定标度区的方法,但在实测中都被证明不具备普适性。因为现实中的研究对象非常复杂,现实世界没有真正的简单分形,大都是自仿射分形或者多分形性质的对象。如果分形是自仿射的,盒子法可以给出比较明确的标度区范围,但面积一半径标度法无法给出明显的标度区界限。其原因在于:自仿射分形是各向异性的,采用各向同性的面积一半径标度法测量,结果往往似是而非。如果分形是多标度的,情况就更加复杂了。一个有效的标度区的判断方法,必须保证基于这个标度区的多分维谱正常。可是,从简单分形角度来看非常合理的标度区范围,给出的多分维谱有时却是混乱的;而看起来不够合理的标度区范围,却能给出正常的多分维谱。举例说来,尺度较大时,测量的前几步没有非空盒子,这几步应该视为溢出标度区外。然而,在测量多分维谱时,将这前几步剔除之后,却往往不能给出正常的多分维谱,倒是考

虑这几步反而给出合理的多分维谱系。由此可见,标度区范围的确定是非常困难的。

今天看来,客观的标度区范围的判据也许可以找到,也许永远找不到。特别是,当自仿射性与多标度分形性质杂合在一起的时候,很难借助半径法给出明确的标度范围。这就引发如下问题:分维值还是客观的吗?当然不够客观,但对于有经验的研究者而言,标度区的判断结果大同小异。估计的分维值可能有差别,但只要不存在显著性的差异,问题就不大;同时,分维作为一种测度的替代参数,其数值意义在比较中才能显现出来。单独一个分维数值没有太多的信息,多个分维值比较才有意义。对于分维测量而言,分维数值的可比较性要比标度区的客观性或者数值的精确性更为重要。

4 城市形态的分形判断

4.1 标准误差与分形判定

地理规律是演化的规律,而不是存在的规律。城市分形不是一开始就有的,而是通过自组织过程逐步演化出来的(Benguigui et al, 2000; 陈彦光, 2008)。分形、位序—规模分布之类都是复杂系统进入自组织临界状态出现的有序图式(Bak, 1996)。而且,如果城市管理措施不当,进化后的分形结构和自相似分布也可能受到干扰和破坏。如何判断一个城市的形态是分形的?具体操作的办法为:建立测量尺度(盒子线性尺寸,圆环的半径)与相应测度(非空盒子数目或信息熵,圆环内的面积、密度等)的双对数关系。如果log-log坐标图上散点呈现直线趋势,或者出现一个直线段,则可以判断城市形态具有分形性质,直线的斜率或其绝对值给出分维的估计值或标度指数。如果开展统计分析,可以借助双对数线性回归估计分维值,并用相关系数平方即 R^2 值对拟合效果进行评估。然而,并非所有学者都赞同这种定性的标准,以色列学者Benguigui研究团队提出了一个分维的标准误差 δ 作为判据。他们认为,只要分维的标准误差小于 $\delta < 0.04$,这个分维数就可以接受,从而判断城市形态是自相似结构(Benguigui et al, 2000)。问题在于,拟合优度 R^2 也好,参数标准误差 δ 也好,它们都是描述性的统计量,不是推断性的统计量(Shelberg et al, 1982),不能据之判断分形存在与否,只能借助它们对分维计算结果的可信程度做出估价。

实际上,借助最小二乘法估计分维,相当于基于双对数线性关系开展一元线性回归。对于一元线性回归来说, R^2 与 F 统计量以及分维的 T 统计量彼此等价。分维的标准误差与 R 值虽不是等价关系,但却存在严格的函数关系。根据统计学的有关知识,可以推导出分维 D 、分维的标准误差 δ 和相关系数 R 的关系如下

$$\frac{D}{\delta} = \sqrt{F} = T = \frac{R}{\sqrt{\frac{1}{n-2}(1-R^2)}} \quad (9)$$

式中: n 为参与回归的数据点数, F 为回归分析的 F 统计量, T 为分维的 T 统计量。上式变形可得:

$$\delta = D \sqrt{\frac{1/R^2 - 1}{n-2}} \quad (10)$$

由此可知:①分维的标准误差与分维自身大小正相关,与拟合优度成反比例关系。在 R^2 一定的情况下,分维越大,则标准误差越大。②在分维一定的情况下,拟合优度越高,则标准误差越小。③标准误差与自由度成反比关系,增加数据点可以减小分维值的标准误差,从而缩小误差范围。④对于分维估算, F 检验和 T 检验均无必要,只要给出拟合优度检验即可。

城市形态的分形研究过程是一种数学建模,模型参数估计需要借助于算法,估计效果需要根据相关统计量进行评估。统计检验是对模型参数估计效果的一种评价,任何评价方法都有局限性。统计检验通过与否取决于显著性水平的门槛,置信度低不一定在于研究对象是否分形结构,而可能是算法本身引起的。因此,无论根据 R^2 、参数标准误差 δ 抑或参数分布特征,都不能确定一个系统的分形性质。换言之,描述性统计量不是分形结构的决定性判据,仅仅是分维估计效果的评价指数。那么,怎样判断城市形态的分形结构存在呢?其实,统计分析永远不可能得到100%确定的结论,关键在于利用置信陈述——给出分形判断结论的置信度以及分维估计的误差范围。

4.2 分维测量与置信陈述

模型参数的显著性检验从来没有一个客观的标准,其置信度都是人为规定的。一般取0.05的显著性水平(对应于95%置信度),这个标准的好处是置信门槛和误差范围比较合适,并且数值估计通常较为方便。没有任何证据表明必须通过95%的置信度的检验。对于分维测量结果,关键在于给出置信陈述:结论的显著性水平的高低,分维的误差范

围大小。不少文献采用分维 D 加减标准误差 δ 表示误差范围($D \pm \delta$),这是误会。只有简单的比率统计可以采用标准误差表示计算结果的误差范围。分维的误差范围应该是分维 D 加减标准误差 δ 乘以 T 统计量的门槛值 $T_{\alpha,n-k-1}$,计算公式如下:

$$D^* = D \pm \delta \cdot T_{\alpha,n-2} \quad (11)$$

式中: D^* 表示分维的上下限(D_u 、 D_l), $T_{\alpha,n-2}$ 表示 T 统计量的门槛值, α 为显著性水平(通常取0.05或者0.01,相应的置信度为95%或者99%), n 为用于估计分维的数据点数。假定 $n=10$,则自由度 $df=n-2=8$,取 $\alpha=0.05$,在Excel的任意单元格输入函数“=tinv(0.05,8)”立即得到 $T_{\alpha,n-2}=T_{0.05,8}=2.306$ 。假定分维估计值为 $D=1.7$,标准误差为0.01,则误差范围约为 $D_l=1.677$ 到 $D_u=1.723$ 。结论是:我们有95%的把握相信,分维值约为1.7,误差范围大约为1.677~1.723。如果显著性水平取 $\alpha=0.01$,则误差范围为 $D_l=1.666$ 到 $D_u=1.734$ 。结论是:我们有99%的把握相信,分维值约为1.7,误差范围大约为1.666~1.734。显然,置信度越高,误差范围也就越大,统计检验并非置信度越高越好。如果采用标准误差直接估计误差范围,则与置信度无关,这不符合统计学的置信表述规则。将式(10)代入式(11)可得

$$D^* = D(1 \pm \sqrt{\frac{1/R^2 - 1}{n-2}} \cdot T_{\alpha,n-2}) \quad (12)$$

由式(12)可以看出,给定分维 D 值,误差范围与拟合优度 R^2 、自由度 $n-2$ 和显著性水平 α 有关。拟合优度越高,误差范围越小;自由度越高,误差范围越小;显著性水平越高(置信度越低),误差范围也越小。增加双对数坐标图中的观测数据点可以改变分维估计值的误差范围,但对分维估计值本身影响并不明显。

至此可知,在分维估计过程中,样本大小或者数据点的多少不是主要问题,问题在于置信水平和误差范围。如果样本大、数据点多,则误差范围小、分维估计值稳定,或者置信度高;相反,如果样本小、数据点少,则误差范围大、分维值不太稳定,或者置信度低。统计检验的门槛总是由客观的自由度和主观定义的置信度共同决定(陈彦光等,2014)。只要采样或者数据处理过程没有偏性,则分维估计值不会对实际值有太大偏离,只是误差范围大小不同而已。

4.3 城市分形的置信判断

城市分形判定的简单途径就是测量,测算城区面积、边界长度、分布密度,如此等等。如果测量结

果很快收敛到确定的数值,则研究对象是非分形的;反之,如果无法测出稳定的结果,则研究对象具有标度性质,可将其作为分形对待。在技术层面,城市分形的识别和判断步骤如下:

第一步,直观判断。借助双对数坐标图,可对城市形态的分形进行直观的初步识别。分形的基本性质是自相似,统计自相似表现为统计分析中的标度关系。如果测量尺度如盒子的线性尺寸 r 与相应测度如城市形态的面积 $A(r)$ 或非空盒子数目 $N(r)$ 满足幂律关系,则可初步认定城市形态是分形的。幂律是标度关系的数学反映,它是分形结构存在的必要条件,而不是充分条件。只有当一个城市形态存在分维的时候,才可以进一步确定其分形性质。

第二步,定量判断。借助分维测量,可以定量地判断城市形态的分形形状及其空间结构特征。如果一个系统是分形的,则它一定存在一个分数的维数。基于空间测量得到的分维值 D 一定大于拓扑维数0而小于嵌入空间的欧氏维数 d 。因此,可接受的分维值是城市分形的定量判据。有分形不一定可以估计分维——不是没有分维,而可能是技术障碍。但是,如果有办法估计出分维,并且估计的分维值 D 落入合理范围之内($0 < D < d$),则可以判断相应的系统具有分形结构。欧氏维数、分维和拓扑维数分别对应于log-log坐标图上的三个线段。但是,由于图像资料分辨率的问题,经验上拓扑维数往往被伪欧氏维数所替代。

第三步,置信结论。分维的统计检验是基于某个显著性水平的,一个显著性水平对应于一个置信度。所有的统计结论都是在一定置信度上得出的,没有100%可信的结论。城市分形的判断是以分维估计和检验为基础的。如果一个分维可以通过0.05的显著性水平的检验,就有95%的把握判断城市分形存在;如果一个分维可以通过0.01的显著性水平的检验,就有99%的把握判断存在城市分形……。其余依此类推。

由此可见,样本点的多少、回归分析的相关系数多高或者分维的标准误差多大,已经不重要了,重要的是置信陈述(陈彦光等,2014);研究人员究竟有多大把握判断城市形态具有分形性质。

5 结束语

分形城市研究的根本原因之一是城市形态没

有特征尺度,不能采用传统的测度如长度、面积、体积、规模等等对城市进行客观描述。分维是一个有特征尺度的参数,采用分维描述城市形态可解决城市格局空间分析的尺度依赖性问题。同时,分维作为一个空间指数,可将众多的测量数据浓缩为一个简单的数值,据此反映研究对象的空间特征。单纯一个分维说明不了太多的问题,但通过分维演变或者差异的纵横比较,就可揭示城市演化和分布的一些重要的时空信息。要想采用分形方法对城市开展卓有成效的研究,必须适当地估计分维,并据此对城市形态的分形特征进行有效地判断。本文的要点可以归结如下。

第一,分维估计的基本思想是最佳覆盖,就是采用测量体不多不少、不大不小地覆盖住被测量的分形对象。覆盖的效果即测量体与被测量对象之间的逼近效果越好,分维估计值越是接近于真值。对于现实中具有随机前分形特征的城市形态,无法找到最佳覆盖方式,这个问题在统计分析过程中可以通过以下三方面的测量准则来解决:一是快速逼近准则,即测量尺度与分形结构具有一致性。分形体是以几何数列组织起来的递阶结构,分维测量的尺度最好采取几何尺度,而不是算术尺度。采用几何尺度可以捕捉分形的结构特征,实现测量过程的快速逼近。二是操作方便准则,即测量过程不能太复杂。在各种测量方法中,尺度一分为二、二分为四的过程最为便捷,一分为三、一分为五的方法理论上可行,但不便于操作。三是结果稳定准则,即用于估计分维的数据点数不宜太少。采用尺度一分为二的方式,即能保证数据点较多,又能与分形的递阶结构一致,同时操作也最为方便。

第二,分形判断的统计方法是置信陈述,即基于某个给定的显著性水平得出结论。理论上的规则分形是容易识别和判断的,但现实中的随机分形往往似是而非,不太可能给出百分之百肯定的结论。当人们借助观测数据和统计分析技术做出判断的时候,最可取的方法就是置信判断和陈述。置信表述的要件是分维值、置信度和误差范围。只要估计出分维值,一定可以给出置信陈述。首先,借助尺度和相应测度的幂指数关系估计分维,并对测量结果进行统计检验。如果分维数值通过某个显著性水平如0.05的检验,就可以基于相应的置信度(如95%)得出结论,判断分形特征存在。其次,基于某个显著性水平给出分维的误差范围。显著性水

平越低,相应的置信度越高,但分维的误差范围也会越大。有了置信陈述,样本的大小或者数据点的多少就不是主要问题。数据点越少,自由度也就越小,从而统计检验的标准也就越是苛刻。只要通过基于某个显著性水平的检验,那就表明相应的分维值在某个置信度上是可信的;只要分维可信,城市形态分形性质的可信程度也就随之提高。

参考文献(References)

- 陈彦光. 1998. 中心地体系中的分形和分维[J]. 人文地理, 13(3): 19-24. [Chen Y G. 1998. Fractals and fractal dimensions in central place systems[J]. Human Geography, 13(3): 19-24, 18.]
- 陈彦光. 2008. 分形城市系统: 标度·对称·空间复杂性[M]. 北京: 科学出版社. [Chen Y G. 2008. Fractal urban systems: Scaling, symmetry, spatial complexity[M]. Beijing, China: Science Press.]
- 陈彦光. 2015. 简单、复杂与地理分布模型的选择[J]. 地理科学进展, 34(3): 321-329. [Chen Y G. 2015. Simplicity, complexity, and mathematical modeling of geographical distributions[J]. Progress in Geography, 34(3): 321-329.]
- 陈彦光, 刘继生. 2007. 城市形态分维测算和分析的若干问题[J]. 人文地理, 22(3): 98-103. [Chen Y G, Liu J S. 2007. On fractal dimension calculation and analysis of urban form[J]. Human Geography, 22(3): 98-103.]
- 陈彦光, 张莉. 2014. 信阳城市人口—城区用地异速生长分析[J]. 地理科学进展, 33(8): 1058-1067. [Chen Y G, Zhang L. 2014. An allometric analysis of the scaling relations between population and urban area of Xinyang[J]. Progress in Geography, 33(8): 1058-1067.]
- 姜世国, 周一星. 2006. 北京城市形态的分形集聚特征及其实践意义[J]. 地理研究, 25(2): 204-212. [Jiang S G, Zhou Y X. 2006. The fractal urban form of Beijing and its practical significance[J]. Geographical Research, 25(2): 204-212.]
- 刘继生, 陈彦光. 1998. 城镇体系等级结构的分形维数及其测算方法[J]. 地理研究, 17(1): 82-89. [Liu J S, Chen Y G. 1998. Fractal dimensions of hierarchical structure of urban systems and the methods of their determination[J]. Geographical Research, 17(1): 82-89.]
- 刘继生, 陈彦光. 1999a. 城镇体系空间结构的分形维数及其测算方法[J]. 地理研究, 18(2): 171-178. [Liu J S, Chen Y G. 1999a. Fractal dimensions of spatial structure of an urban system and the methods of their determination[J]. Geographical Research, 18(2): 171-178.]
- 刘继生, 陈彦光. 1999b. 交通网络空间结构的分形维数及其

- 测算方法探讨[J]. 地理学报, 54(5): 471-478. [Liu J S, Chen Y G. 1999b. A study on fractal dimensions of spatial structure of transport networks and the methods of their determination[J]. *Acta Geographica Sinica*, 54(5): 471-478.]
- 刘继生, 陈彦光. 2003. 河南省城镇体系空间结构的多分形特征及其与水系分布的关系探讨[J]. 地理科学, 23(6): 713-720. [Liu J S, Chen Y G. 2003. Multifractal measures based on man-land relationships of the spatial structure of the urban system in Henan[J]. *Scientia Geographica Sinica*, 23(6): 713-720.]
- 秦静, 方创琳, 王洋, 等. 2015. 基于三维计盒法的城市空间形态分维计算和分析[J]. 地理研究, 34(1): 85-96. [Qin J, Fang C L, Wang Y, et al. 2015. A three dimensional box-counting method for estimating fractal dimension of urban form[J]. *Geographical Research*, 34(1): 85-96.]
- Addison P S. 1997. *Fractals and chaos: An illustrated course* [M]. Bristol, Philadelphia: Institute of Physics Publishing.
- Ariza- Villaverde A B, Jiménez- Hornero F J, De Ravé E G. 2013. Multifractal analysis of axial maps applied to the study of urban morphology[J]. *Computers, Environment and Urban Systems*, 38: 1-10.
- Bak P. 1996. *How nature works: The science of self-organized criticality*[M]. New York, NY: Springer-Verlag.
- Batty M, Longley P. 1994. *Fractal cities: A geometry of form and function*[M]. London, UK: Academic Press.
- Benguigui L, Czamanski D, Marinov M, et al. 2000. When and where is a city fractal[J]. *Environment and Planning B: Planning and Design*, 27(4): 507-519.
- Chen Y G, Wang J J. 2013. Multifractal characterization of urban form and growth: The case of Beijing[J]. *Environment and Planning B: Planning and Design*, 40(5): 884-904.
- De Keersmaecker M-L, Frankhauser P, Thomas I. 2003. Using fractal dimensions for characterizing intra-urban diversity: The example of Brussels[J]. *Geographical Analysis*, 35(4): 310-328.
- Feder J. 1988. *Fractals*[M]. New York, NY: Plenum Press: 11.
- Feng J, Chen Y G. 2010. Spatiotemporal evolution of urban form and land-use structure in Hangzhou, China: Evidence from fractals[J]. *Environment and Planning B: Planning and Design*, 37(5): 838-856.
- Frankhauser P. 1994. *La Fractalité des structures urbaines (the fractal aspects of urban structures)*[M]. Paris, France: Economica.
- Frankhauser P. 1998. The fractal approach. A new tool for the spatial analysis of urban agglomerations[J]. *Population: An English Selection*, 10(1): 205-240.
- Jullien R, Botet R. 1987. *Aggregation and fractal aggregates* [M]. Singapore: World Scientific Publishing.
- Kaye B H. 1989. *A random walk through fractal dimensions* [M]. New York, NY: VCH Publishers.
- Mandelbrot B B. 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension[J]. *Science*, 156: 636-638.
- Mandelbrot B B. 1977. *Fractals: Form, chance, and dimension* [M]. San Francisco, CA: W. H. Freeman and Company.
- Mandelbrot B B. 1983. *The fractal geometry of nature*[M]. New York, NY: W. H. Freeman and Company.
- Murcio R, Masucci A P, Arcaute E, et al. 2015. Multifractal to monofractal evolution of the London street network[J]. *Physical Review E*, 92(6): 062130.
- Shelberg M C, Moellering H, Lam N. 1982. Measuring the fractal dimensions of empirical cartographic curves[J]. *Auto Carto*, 5: 481-490.
- Shen G Q. 2002. Fractal dimension and fractal growth of urbanized areas[J]. *International Journal of Geographical Information Science*, 16(5): 419-437.
- Thomas I, Frankhauser P, Biernacki C. 2008. The morphology of built-up landscapes in Wallonia (Belgium): A classification using fractal indices[J]. *Landscape and Urban Planning*, 84(2): 99-115.
- Thomas I, Frankhauser P, De Keersmaecker M-L. 2007. Fractal dimension versus density of built-up surfaces in the periphery of Brussels[J]. *Papers in Regional Science*, 86(2): 287-308.
- Vicsek T. 1989. *Fractal growth phenomena*[M]. Singapore: World Scientific Publishing.
- White R, Engelen G. 1994. Urban systems dynamics and cellular automata: Fractal structures between order and chaos [J]. *Chaos, Solitons & Fractals*, 4(4): 563-583.

Approaches to estimating fractal dimension and identifying fractals of urban form

CHEN Yanguang

(Department of Urban and Economic Geography, College of Urban and Environmental Sciences,
Peking University, Beijing 100871, China)

Abstract: Fractal cities and city fractals have been studied for about 30 years, but many basic problems have not yet been fully solved. Among the various basic fractal problems about cities, the most important are how to estimate fractal dimension and how to identify fractal nature of a city as a system or a system of cities in an effective way. Based on urban form and the box-counting method, this article discusses the approaches to calculating fractal dimension and determining fractal property of cities. The theoretical foundation of fractal dimension computation is the concept of perfect covering. In other words, in the process of fractal dimension measurement, a fractal object should be covered with boxes in the best way: nothing more, nothing less. In practice, it is hard to find the optimal way of box covering. Therefore, three rules should be followed. First, the sequence of measuring scales should be consistent with the cascade structure of a fractal city so that the fractal details can be captured in a reasonable way. Second, the operation of obtaining observational data should be simple and convenient to use so that the method can be applied by beginners. Third, the size of the dataset for fractal dimension estimation should be large enough so that the result of fractal parameters is stable. A conclusion can be reached that the geometric scale rather than the arithmetic scale should be employed to make a measurement because the fractal structure can be abstracted as geometric sequences instead of arithmetic sequences, and the measuring scale sequence should comply with the $1/2^n$ rule (that is, 1, $1/2$, $1/4$, ...). After estimating the fractal dimension of a city, the fractal property can be identified by the statistics from fractal dimension calculation. The good way of identifying the fractal nature of urban form is to use confidence statement, which consists of fractal dimension values, margin of error, and level of confidence. Given a level of significance (for example, $\alpha=0.05$), we can draw an inference that a city's form is of fractal structure because it has a fractal dimension with a determinate level of confidence (for example, $(1-\alpha)\times 100\%=95\%$). Using statistical analysis, however, one can never have full (100%) assurance that a city has a fractal form.

Key words: urban form; fractals; fractal dimension; self-affinity; box-counting method; multifractals; scaling range