

文章编号: 1007-4619 (2000) 04-0316-06

# 矢量 GIS 图上地理曲线的定位误差模型

刘文宝<sup>1,2</sup>, 邓 敏<sup>1</sup>, 夏宗国<sup>2,3</sup>

(1. 山东科技大学 地球科学系, 山东 泰安 271019; 2. 美国麻省大学 地球地理科学系, 美国 波士顿 02125;

3. 武汉测绘科技大学, 湖北 武汉 430070)

**摘 要:** 为了分析 GIS 图上地理曲线的定位精度, 首先探讨地理曲线的表达与定位误差, 区分“数字曲线”和“连续曲线”两个概念; 然后结合由数字曲线生成连续曲线的 GIS 算法, 建立连续曲线的误差带模型, 并导出地理曲线长度的误差估计公式; 最后通过算例说明地理曲线误差带的可视化方法和曲线长度误差估计公式的应用。

**关键词:** 地理曲线; 位置误差; 可视化

**中图分类号:** TP79/P23 **文献标识码:** A

## 1 引 言

GIS 空间数据质量是当前国际 GIS 和制图界的重要研究专题之一<sup>[1]</sup>。现已公认的空间数据质量评价指标为 7 项: 渊源信息 (Lineage), 位置精度 (Positional Accuracy), 属性精度 (Attribute Accuracy), 完整性 (Completeness), 逻辑一致性 (Logical Consistency), 语义精度 (Semantic Accuracy) 和时态信息 (Temporal Information)<sup>[2]</sup>。简言之, 空间数据质量报告将描述最终的数据文件与自然界原型客观实体的一致性程度。

在 GIS 中, 描述自然界客观实体的地理信息是采用目标或场的方法表达的<sup>[3]</sup>。前者利用矢量数据结构, 将自然界视为一种几何原型, 经过抽象和综合, 由点、线和面表达, 涉及几何位置、拓扑关系和非空间的属性值。后者利用栅格数据结构, 仅提供任意位置处的属性值, 不直接涉及单个空间目标和属性值间的关系。因此, 矢量数据质量评价远复杂于栅格数据质量评价问题<sup>[4]</sup>。

矢量数据位置精度的研究对象主要是点、线和面的几何定位精度<sup>[5]</sup>。点位精度问题在测绘界已进行了长期研究, 点位协方差阵已被公认为度量点位

精度的完备指标。为了实用上的方便, 由点位协方差阵简化导出了一些数值指标, 例如点位均方差、坐标分量误差, 以及点位协方差阵的几何图像——误差椭圆和误差曲线等可视化指标<sup>[6]</sup>。其中, 点位均方差和误差圆已在 GIS 中广泛应用<sup>[5]</sup>。对于一维高程点的数据, 文献<sup>[6]</sup>提出了一种可视化精度指标——误差区间。而线和面的定位精度是伴随 GIS 的诞生而提出的新问题。文献<sup>[7]</sup>对 Perkal 提出的实验曲线的“ $\epsilon$ -带”概念<sup>[8]</sup>进行了扩展, 用于度量一段制图线段的不确定性。“ $\epsilon$ -带”模型虽然绘制简单, 但缺少统计学基础。为此, 文献<sup>[6]</sup>提出了平面随机线段的统计表现为二元随机过程的思想, 并导出了能完备描述随机线段线位精度的线位相关函数表达式, 以及相应的数值简化指标, 从而将点位协方差阵扩展到线位相关函数。文献<sup>[9]</sup>进一步研究了线位相关函数的几何图像——线位误差带, 利用包络线原理导出了线位误差带的边界线方程, 从而将点位误差椭圆可视化指标扩展到线位误差带可视化指标。类似地, 根据文献<sup>[6]</sup>提出的平面随机多边形的统计表现为二元随机场的思想, 文献<sup>[10]</sup>导出了能完备描述随机多边形精度的形位相关函数及其几何图像——误差环。由于折线状目标本身在矢量 GIS 空间数据库中是以折线方式存贮位置数据的, 例如

收稿日期: 1999-12-02

基金项目: 国家自然科学基金资助项目, 编号 49801016 和 49671063。

作者简介: 刘文宝(1965—), 男, 教授, 1987, 1990 年于山东科技大学获工学学士、硕士学位; 1996 年于武汉测绘科技大学获工学博士学位; 1996 年 7 月至 1998 年 6 月为东南大学博士后。主要研究领域为遥感与地理信息系统中的数据质量、误差与不确定性, 以及信息可视化, 已发表论文 40 余篇, 合著 1 部。

ARC/INFO, 因而上述直线段的线位误差带可以直接扩展用于度量折线精度<sup>[11]</sup>。

尽管自然界中的大量曲线状目标, 例如海岸线、生态环境分界线与人文景观分界线等, 在矢量 GIS 空间数据库中也是以折线方式存贮的, 但在 GIS 输出图上却表现为连续的地理曲线。本文探讨这种地理曲线的平面定位精度和曲线长度的误差估计。

## 2 地理曲线的误差模型

### 2.1 地理曲线的表达与定位误差

自然界中的地理曲线可区分为两种: 一种是定义比较明确的, 例如道路、河流、管线和行政区划分界线等; 另一种是定义不明确的, 例如沙漠草场分界线和土壤特性分界线等。一旦确定地理曲线的特征点, 并获取采样值后, 则两种曲线的定义都变为明确的。无论是从自然界中直接采样, 还是从已有纸质地图或专题图上采样, 地理曲线在矢量 GIS 空间数据库中都将以数字文件方式存储。故这种曲线可称为“数字曲线”。针对用户的不同要求, 由“数字曲线”经 GIS 算法可生成各种专题图上的“连续曲线”, 例如土地利用图上的边界线<sup>[12]</sup>。显然, 这种“连续曲线”与自然界中的地理曲线原型所存在的差别反映了“连续曲线”(以下简称地理曲线)的定位精度。由于这种精度取决于多种因素的影响, 例如所选取的特征点的代表性、数目、采样精度、可靠性和尺度以及绘图误差等。因此, 地理曲线的精度估计是一个非常复杂的问题。

在对地理曲线进行实际采样时, 原始采样点数目一般较多, 存在大量信息冗余。在存储于 GIS 空间数据库之前, 需进行“数据压缩”处理, 仅将最关键的特征点数据存于空间数据库中, 形成“数字曲线”。有关数据压缩的准则和方法在制图综合、计算机科学和计算数学领域均有大量研究<sup>[13]</sup>。这种“数字曲线”的定位误差, 可直接由折线误差带度量<sup>[11]</sup>。

由“数字曲线”所生成的 GIS 图上的连续地理曲线, 根据地理曲线本身固有的特性不同, 分别表达为经典几何曲线、分形曲线和模糊曲线<sup>[13]</sup>。例如, 土地利用图上的边界线为经典几何曲线, 海岸线和土壤特性分界线等为分形曲线, 沙漠草场的分界线则为模糊曲线。不同的曲线有不同的生成方法<sup>[13]</sup>。由于大多数 GIS 专题图上的曲线属于经典几何曲线类, 本文讨论这类曲线。

### 2.2 地理曲线的函数模型

如上所述, 经过数据压缩后所得的数字曲线, 其中既含有地理曲线原型的消息量最大, 又使折点数量少。在将 GIS 空间数据库中的一条数字曲线转换为矢量地图上的一条连续曲线时, 根据矢量图的尺度确定是否利用数字曲线的所有点。转换过程是借助于拟合、插值或光滑函数分段生成的, 其中包括模型选择和参数识别<sup>[13]</sup>。首先根据数字曲线折点的位置坐标和误差信息, 确定函数模型; 然后采用解线性(或非线性)方程组或者迭代算法确定拟合或插值函数中的未知参数。在 GIS 中, 通常采用的函数模型为回归模型、特征多项式模型和插值函数模型<sup>[13]</sup>。

#### 2.2.1 回归函数

一元回归函数模型为:

$$y(x) = c_0 + c_1x + c_2x^2 + \dots + c_kx^k \quad (1)$$

式中,  $x, y$  分别为描述地理曲线点坐标关系的自变量和参变量,  $c_i (0 \leq i \leq k)$  为回归多项式系数;  $k$  为多项式阶数, 根据特征点个数和模型辨识确定。特别地, 对于直线状地理边界,  $k$  取 1。

#### 2.2.2 特征多项式拟合函数

根据用户对图纸所要求的尺度, 不妨设需选取数字曲线上的  $n$  个特征点, 则拟合曲线方程与特征点坐标之间的关系式为:

$$\begin{cases} x = f_x(x_1, y_1, x_2, y_2, \dots, x_n, y_n, l) \\ y = f_y(x_1, y_1, x_2, y_2, \dots, x_n, y_n, l) \end{cases} \quad (2)$$

式中,  $z_i = (x_i y_i)^T$  为特征点坐标;  $z = (xy)^T$  为地理曲线上离起始点弧长为  $l$  处的点的坐标。若  $n$  等于 2, 则地理曲线退化为直线, (2) 式简化为:

$$\begin{cases} x = f_x(x_1, y_1, x_2, y_2, l) = (1-r)x_1 + rx_2 \\ y = f_y(x_1, y_1, x_2, y_2, l) = (1-r)y_1 + ry_2 \end{cases} \quad (3)$$

式中, 参数  $r = z_1 z_i / z_1 z_2$ , 并满足  $0 \leq r \leq 1$ ,  $z_i$  为线段上任意点<sup>[6]</sup>。

#### 2.2.3 插值多项式函数

插值多项式函数的一般形式为<sup>[13]</sup>:

$$\begin{cases} x_i = f_\alpha(\alpha^T, t) \\ y_i = f_\beta(\beta^T, t) \end{cases} \quad (4)$$

式中  $x_i, y_i$  为插值点坐标,  $\alpha, \beta$  为插值函数的系数向量,  $t$  为插值参数。插值函数不同, 所利用的数字特征点数也不同。常用的插值函数有多项式函数、三点抛物线函数、两点抛物线求导法和张力样条函数等<sup>[13]</sup>。

### 2.3 地理曲线的误差模型

#### 2.3.1 回归函数误差模型

模型(1)所伴随的误差模型为:

$$y(x) = \bar{y}(x) \pm \sigma_\gamma \quad (5)$$

式中:  $\bar{y}(x) = \bar{c}_0 + \bar{c}_1x + \bar{c}_2x^2 + \dots + \bar{c}_kx^k$ , 而系数  $\bar{c}_i$  ( $0 \leq i \leq k$ ) 为最小二乘估计值; 参数  $\sigma_\gamma$  为在  $\gamma$  置信水平下误差带的宽度<sup>[12]</sup>。

#### 2.3.2 特征多项式拟合函数误差模型

根据最小二乘拟合公式<sup>[13]</sup>, 可导出地理曲线上任意一点  $z = (xy)^T$  的点协方差矩阵为:

$$\Sigma_z = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix} = A^T \Sigma_{00} A \quad (6)$$

式中:  $\Sigma_z$  为任意点  $z = (xy)^T$  的协方差阵;  $A^T = \begin{pmatrix} p_1 q_1 p_2 q_2 \dots p_n q_n \\ h_1 j_1 h_2 j_2 \dots h_n j_n \end{pmatrix}_{2 \times 2n}$ , 且  $p_i = \partial f_x / \partial x_i$ ,  $q_i = \partial f_x / \partial y_i$ ,  $h_i = \partial f_y / \partial x_i$ ,  $j_i = \partial f_y / \partial y_i$ ;  $\Sigma_{00}$  为  $2n \times 2n$  维的特征点坐标的协方差矩阵。显然,  $\Sigma_z$  表达了地理曲线上任意点处的位置误差。

#### 2.3.3 插值多项式函数误差模型

参数方程(4)式的一种常用形式为:

$$\begin{cases} x_t = f_\alpha(\alpha^T, t) = a_0 + a_1t + a_2t^2 + \dots + a_mt^m \\ y_t = f_\beta(\beta^T, t) = b_0 + b_1t + b_2t^2 + \dots + b_nt^n \end{cases} \quad (7)$$

其中,  $\alpha^T = (a_0 a_1 a_2 \dots a_m)$ ,  $\beta^T = (b_0 b_1 b_2 \dots b_n)$  为参数方程的系数向量,  $m, n$  分别为阶数,  $t$  为参数<sup>[13]</sup>。记  $U^T = (\alpha^T \beta^T)$ ,  $B(t) = \text{diag}(B_\alpha^T B_\beta^T)$ ,  $B_\alpha^T = (1 t^2 \dots t^m)$ ,  $B_\beta^T = (1 t^2 \dots t^n)$ , 则地理曲线  $\{z_t = (x_t y_t)^T; t = t_1, t_2, \dots, t_n\}$  上任意内插点的协方差阵为:

$$\Sigma_{z_t z_t} = B(t) \Sigma_{UU} B^T(t) \quad (8)$$

即:

$$\begin{pmatrix} \sigma_{x_t}^2 & \sigma_{x_t y_t} \\ \sigma_{y_t x_t} & \sigma_{y_t}^2 \end{pmatrix} = \begin{pmatrix} B_\alpha^T \sigma_\alpha^2 B_\alpha & B_\alpha^T \sigma_{\alpha\beta} B_\beta \\ B_\beta^T \sigma_{\beta\alpha} B_\alpha & B_\beta^T \sigma_\beta^2 B_\beta \end{pmatrix} \quad (9)$$

其中,

$$\Sigma_{UU} = \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\beta\alpha} & \sigma_\beta^2 \end{pmatrix}, \quad \sigma_\alpha^2 = \begin{pmatrix} \sigma_{a_0}^2 & \sigma_{a_0 a_1} & \dots & \sigma_{a_0 a_m} \\ & \sigma_{a_1}^2 & \dots & \sigma_{a_1 a_m} \\ & & \dots & \sigma_{a_m}^2 \end{pmatrix},$$

$$\sigma_{\alpha\beta} = \begin{pmatrix} \sigma_{a_0 b_0} & \sigma_{a_0 b_1} & \dots & \sigma_{a_0 b_n} \\ \sigma_{a_1 b_0} & \sigma_{a_1 b_1} & \dots & \sigma_{a_1 b_n} \\ \dots & \dots & \dots & \dots \\ \sigma_{a_m b_0} & \sigma_{a_m b_1} & \dots & \sigma_{a_m b_n} \end{pmatrix}, \quad \sigma_\beta^2 = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} & \dots & \sigma_{b_0 b_n} \\ & \sigma_{b_1}^2 & \dots & \sigma_{b_1 b_n} \\ & & \dots & \sigma_{b_n}^2 \end{pmatrix}.$$

其中,  $\Sigma_{UU}$  为插值函数参数的协方差阵, 根据数字曲线特征点的误差信息求得<sup>[13]</sup>。

#### 2.3.4 地理曲线误差带概念

根据(9)式, 可导出地理曲线  $\{z_t = (x_t y_t)^T; t = t_1, t_2, \dots, t_n\}$  上任意内插点在  $\varphi_t$  方向上的位差为:

$$\sigma_{\varphi_t}^2 = (B_\alpha^T \sigma_\alpha^2 B_\alpha) \cos^2 \varphi_t + (B_\alpha^T \sigma_{\alpha\beta} B_\beta) \sin^2 \varphi_t + (B_\beta^T \sigma_\beta^2 B_\beta) \sin^2 \varphi_t \quad (10)$$

由此可得地理曲线上任意内插点误差椭圆的长半轴方向及长、短半轴计算式为:

$$\text{tg}^2 \varphi_\tau = \frac{2 B_\alpha^T \sigma_{\alpha\beta} B_\beta}{B_\alpha^T \sigma_\alpha^2 B_\alpha - B_\beta^T \sigma_\beta^2 B_\beta} \quad (11)$$

$$\lambda_{1t}^2 = (B_\alpha^T \sigma_\alpha^2 B_\alpha) \cos^2 \varphi_\tau + (B_\alpha^T \sigma_{\alpha\beta} B_\beta) \sin^2 \varphi_\tau + (B_\beta^T \sigma_\beta^2 B_\beta) \sin^2 \varphi_\tau \quad (12)$$

$$\lambda_{2t}^2 = (B_\alpha^T \sigma_\alpha^2 B_\alpha) \sin^2 \varphi_\tau - (B_\alpha^T \sigma_{\alpha\beta} B_\beta) \sin^2 \varphi_\tau + (B_\beta^T \sigma_\beta^2 B_\beta) \cos^2 \varphi_\tau \quad (13)$$

利用(11)、(12)和(13)式可绘出地理曲线  $\{z_t = (x_t y_t)^T; t = t_1, t_2, \dots, t_n\}$  上任意内插点的误差椭圆族相互叠加所形成的带状区域, 称为地理曲线的误差带。

### 2.4 地理曲线长度的方差估计

地理曲线  $\{z_t = (x_t y_t)^T, t = t_1, t_2, \dots, t_n\}$  上任意点处的切向量为:

$$\begin{cases} x'(t) = \frac{dx}{dt} = a_1 + 2a_2t + \dots + ma_mt^{m-1} \\ y'(t) = \frac{dy}{dt} = b_1 + 2b_2t + \dots + nb_nt^{n-1} \end{cases} \quad (14)$$

则曲线长度为:

$$l = \int_{t_1}^{t_2} \sqrt{x'^2(t) + y'^2(t)} dt \quad (15)$$

其方差为:

$$\sigma_l^2 = e^T \sigma_{\alpha\epsilon}^2 e + 2e^T \sigma_{\alpha\beta} f + f^T \sigma_\beta^2 f \quad (16)$$

其中,  $e^T = (e_0 e_1 \dots e_m)$ ,  $f^T = (f_0 f_1 \dots f_n)$ 。令  $\theta(t) = \frac{x'(t)}{\sqrt{x'^2(t) + y'^2(t)}$ ,  $\phi(t) = \frac{y'(t)}{\sqrt{x'^2(t) + y'^2(t)}$ , 则  $e, f$

中元素  $e_i, f_j$  的计算式为:

$$e_0 = 0, e_1 = \int_{t_1}^{t_n} \theta(t) dt, e_2 = 2 \int_{t_1}^{t_n} t \theta(t) dt, \dots,$$

$$e_m = m \int_{t_1}^{t_n} t^{m-1} \theta(t) dt$$

$$f_0 = 0, f_1 = \int_{t_1}^{t_n} \phi(t) dt, f_2 = 2 \int_{t_1}^{t_n} t \phi(t) dt, \dots,$$

$$f_n = n \int_{t_1}^{t_n} t^{n-1} \phi(t) dt$$

展开(16)式, 得:

$$\sigma_i^2 = \sum_{i=1}^m e_i^2 \sigma_{a_i}^2 + 2 \sum_{i=1}^m \sum_{j=i+1}^m e_i e_j \sigma_{a_i a_j} + \sum_{i=1}^n f_i^2 \sigma_{b_i}^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n f_i f_j \sigma_{b_i b_j} + 2 \sum_{i=1}^m \sum_{j=1}^n e_i f_j \sigma_{a_i b_j} \quad (17)$$

### 3 算例分析

下面通过算例说明地理曲线定位误差的可视化方法和曲线长度的误差估计。为便于分析地理曲线位置误差的内在规律以及误差模型比较, 考虑由6个折点构成的4种误差状态的数字曲线所生成的地理曲线, 即数字曲线折点坐标为: (I) 独立等精度的数字化值; (II) 等精度相关数字化值; (III) 非等精度独立数字化值; (IV) 非等精度相关数字化值。绘制地理曲线误差带的步骤为: (1) 首先根据文献[14]提出的方法估计数字曲线6个折点的方差/协方差, 结果列于表1; (2) 计算地理曲线插值参数的协方差阵  $\Sigma_{UV}$ ; (3) 按  $t_{i+1} - t_i = 0.1$  的等间隔参数值, 利用(11)、(12)和(13)式计算地理曲线上一系列内插点的误差椭圆参数, 其中6个折点处的结果列于表2;

表1 特征点坐标的方差-协方差

Table 1 Variance and covariance of coordinates of characteristic points

	I		II			III		IV		
	$\sigma_{x_i}^2$	$\sigma_{y_i}^2$	$\sigma_{x_i}^2$	$\sigma_{y_i}^2$	$\sigma_{x_i y_i}$	$\sigma_{x_i}^2$	$\sigma_{y_i}^2$	$\sigma_{x_i}^2$	$\sigma_{y_i}^2$	$\sigma_{x_i y_i}$
1	1.000	1.000	1.236	1.192	-0.114	1.236	1.192	1.236	1.192	-0.114
2	1.000	1.000	1.236	1.192	-0.114	1.230	1.179	1.230	1.179	-0.110
3	1.000	1.000	1.236	1.192	-0.114	1.226	1.196	1.226	1.196	-0.116
4	1.000	1.000	1.236	1.192	-0.114	1.222	1.184	1.222	1.184	-0.120
5	1.000	1.000	1.236	1.192	-0.114	1.226	1.196	1.226	1.196	-0.116
6	1.000	1.000	1.236	1.192	-0.114	1.242	1.192	1.242	1.192	-0.124

表2 内插点的误差椭圆参数

Table 2 Parameters of error ellipses of interpolated points

	$t$	I			II			III			IV		
		$A_t/cm$	$B_t/cm$	$\varphi_t/(^\circ)$	$A_t/cm$	$B_t/cm$	$\varphi_t/(^\circ)$	$A_t/cm$	$B_t/cm$	$\varphi_t/(^\circ)$	$A_t/cm$	$B_t/cm$	$\varphi_t/(^\circ)$
1	0	0.980	0.980	0	1.356	1.232	-39.55	1.200	1.177	0	1.130	1.027	-39.55
2	0.5	0.703	0.703	0	0.973	0.884	-39.56	0.859	0.844	0	0.810	0.736	-39.30
3	1.2	0.773	0.773	0	1.070	0.972	-39.46	0.944	0.927	0	0.889	0.810	-39.49
4	2.5	0.625	0.625	0	0.881	0.785	-34.62	0.762	0.759	0	0.736	0.658	-43.97
5	3.7	0.745	0.745	0	1.172	0.917	-28.60	0.984	0.924	0	0.986	0.772	37.98
6	4.6	0.574	0.574	0	1.352	0.634	-27.01	1.077	0.820	0	1.499	0.593	33.20

(4) 绘制各点的误差椭圆, 相互重叠形成地理曲线的可视化误差带。图1为利用表1中数据计算绘出的由上述4种误差状态的数字曲线所生成的地理曲线

的误差带。

从图1中可见, 地理曲线上靠近原数字曲线折点处的点的定位误差大于折点间的内插点的误差,

这与折线情况的结论一致<sup>[9,11]</sup>。图 1(I)中的结果表现为一系列半径不等的误差圆,所形成的误差带形状粗细相间。另外,比较图 1(II),(III),(IV)可知,折点的相关性对误差带的形状有较大影响。

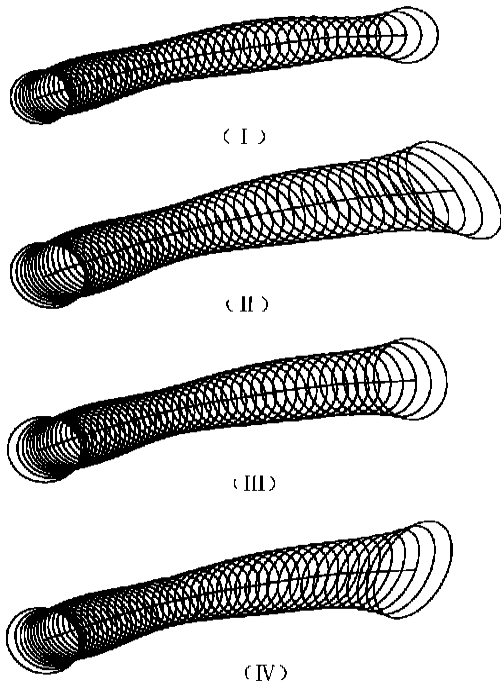


图 1 4 种误差状态下的误差曲线带

Fig. 1 Error bands of four types of geographical line

从曲线生成算法看,模型(1)在文献[12]中有系统讨论,其误差带是一个以均值曲线为中心,以某一置信值为宽度,向两侧分别扩展而形成的缓冲区域。模型(2)、(4)所对应的误差带是由地理曲线上所有点的误差椭圆相互叠加而形成的区域,能有效地顾及曲线的空间特征和空间相关性,兼有定量表达和图形显示的优点。由模型(1)、(2)生成的地理曲线不一定通过数字曲线的折点,因而对曲线形状和定位精度控制的自由度较小;而模型(4)不仅对地理曲线形状控制的自由度较大,而且使曲线形状与坐标系的选择无关。从统计学角度看,地理曲线 $\{z_t = (x_t, y_t)^T; t = t_1, t_2, \dots, t_n\}$ 的统计表现为一个二元随机过程<sup>[13]</sup>,因而模型(4)能有效地表达随机曲线的这种统计特性。

表 3 列出了利用(17)式计算的曲率  $\rho$  为 0, 0.20 和 0.36 的一条直线和两条曲线的线长误差。其中 I 为等精度独立情况, II 为等精度相关情况。分析数据可知,地理曲线长度的误差随着曲率的增大而

增大。在相同的曲率下,数字曲线折点坐标的相关性对曲线长度方差的影响较大。

表 3 曲线长度的方差

Table 3 Variance of line length/cm<sup>2</sup>

	$\rho$	I	II
		$\sigma_l^2$	$\sigma_l^2$
1	0	1.358	1.524
2	0.20	1.641	3.017
3	0.36	1.705	3.044

## 4 结 论

(1)GIS 图上地理曲线的统计特性取决于 GIS 空间数据库中数字曲线折点的统计特性。

(2)地理曲线误差带的形状和大小与 GIS 空间数据库中数字曲线折点的精度、数目、相关性和曲线的形状以及 GIS 算法有关。

(3)利用误差椭圆的叠加图形成的误差带显示地理曲线的定位误差是一种近似方法,但具有易可视化优点。

(4)地理曲线的长度误差与曲线的曲率和数字曲线折点的相关性有关。

## 参 考 文 献 (References)

- [1] Veregin H. Data quality parameters [A]. Longley P. A. *et al.* Geographical Information Systems, 2nd Eds. Vol. 1 [C]. New York: John Wiley & Sons, 1999, 177—189.
- [2] Gupta S C., Morrison J. L. Elements of Spatial Data Quality [M]. New York: Elsevier Science, 1995, 1—12.
- [3] Goodchild M. F. Modeling error in objects and fields [A]. Goodchild M. F., Gopal, S. Accuracy of Spatial Databases [C]. London: Taylor & Francis, 1989, 107—113.
- [4] Burrough P. A., McDonnell R. A. Principles of Geographical Information Systems [M]. Oxford: Oxford University Press, 1998, 220—240.
- [5] Goodchild M. F. Issues of quality and uncertainty [A]. Muller G. C. Advances in Cartography [C]. Oxford: Elsevier Science, 1991, 111—139.
- [6] Liu Wenbao. A Theory of Uncertainty in Spatial Data within GIS. Ph. D. Dissertation [D]. Wuhan Technical University of Surveying and Mapping, 1995, 1—93. [刘文宝. GIS 空间数据的不确定性理论 [D]. 博士学位论文. 武汉测绘科技大学, 1995, 1—93.]
- [7] Chrisman N. R. A theory of cartographic error and its measurement in digital databases [C]. Proceedings of AutoCarto 5, 1982, 159—168.
- [8] Perkal J. On epsilon length [J]. Bulletin de l'Academie Polonaise

- des Sciences*, 1956, **4**: 399—403.
- [9] Liu Wenbao *et al.* Analytical equations of positional error band for line features in GIS [J]. *ACTA GEODAETICA et CARTOGRAPHICA SINICA*, 1998, **27**(3): 231—237. [刘文宝, 戴洪磊等. 平面线位误差带几何形状的解析表达 [J]. 测绘学报, 1998, **27**(3): 231—237.]
- [10] Liu Wenbao *et al.* Models of positional error donut for planar polygon in GIS [J]. *Acta Geodaetica et Cartographica Sinica*, 1998, **27**(4): 338—344. [刘文宝, 戴洪磊等. GIS 中平面面位误差环的解析模型 [J]. 测绘学报, 1998, **27**(4): 338—344.]
- [11] Dai Honglei *et al.* Visualization models of positional uncertainties for polyline in vector GIS [J]. *Acta Geodaetica et Cartographica Sinica*, 1999, **28**(3): 239—243. [戴洪磊, 刘文宝等. 矢量 GIS 中随机折线定位不确定性的可视化模型 [J]. 测绘学报, 1999, **28**(3): 239—243.]
- [12] Gong Peng, Chen Jun. Boundary uncertainties in digitized maps I: some possible determination methods [A]. Proceedings of GIS/LIS'92 Annual Conference and Exposition [C]. California, 1992. 274—281.
- [13] Liu Wenbao. Models of Data Quality and Dynamic Spatial relations in GIS-T [D]. Postdoctoral Technical Report. Nanjing: Southeast University, 1998. 39—44. [刘文宝. GIS-T 中的数据质量与动态空间关系模型 [D]. 博士后出站报告. 南京: 东南大学, 1998, 39—44.]
- [14] Huang Youcai, Liu Wenbao. Building the estimation model of digitizing error [J]. *Photogrammetric Engineering & Remote Sensing*, 1997, **63**(10): 1203—1209.

## Modeling Positional Errors of Geographical Lines in Vector GIS

LIU Wen-bao<sup>1,2</sup>, DENG Min<sup>1</sup>, XIA Zong-guo<sup>2,3</sup>

(1. Department of Geosciences, Shandong University of Science and Technology, Taian 271019, China; 2. Department of Earth & Geographic Sciences, University of Massachusetts at Boston, Boston, MA 02125, USA; 3. Wuhan Technical University of Surveying and Mapping, Wuhan 430070, China)

**Abstract:** To analyze positional accuracy of geographical lines in GIS, this paper begins with a discussion on the representation and positional errors of linear features. We make a clear distinction between digital lines in spatial databases and continuous lines in the natural world. Using algorithms for generating continuous curves from digital lines in GIS, we develop several error band models for continuous curves and derive numerical relationships for estimating errors of line length. Examples are provided to show the visualization of error bands of geographical curves and the use of error equations of line length.

**Key words:** geographical line; positional error; visualization