A multi-scale positional uncertainty model that incorporates multi-scale modeling errors

Jianhong Cai

School of Surveying and Mapping Engineering Beijing University of Civil Engineering and Architecture,

Key Laboratory for Urban Geomatics of National Administration of Surveying, Mapping and Geoinformation,1 Exhibition Hall Road, Xicheng District, Beijing 100044, PR China

email: cjh@bucea.edu.cn

Peiliang Xu

Disaster Prevention Research Institute,

Kyoto University， Gokasho, Uji, Kyoto 611-0011， Japan

email: pxu@rcep.dpri.kyoto-u.ac.jp

Deren Li

State Key Laboratory of information Engineering in Surveying, Mapping and Remote Sensing,

Wuhan University, 129 Ruoyu Road, Wuhan 430079, PR China

email: drli@whu.edu.cn

Guang Zhu

School of Surveying and Mapping Engineering Beijing University of Civil Engineering and Architecture,

1 Exhibition Hall Road, Xicheng District, Beijing 100044, PR China

email: zhuguang@bucea.edu.cn

Abstract

Geographic information world may be interpreted as a digital multi-representation of the real world up to a finite resolving power of resolution. In the process of modeling the real world, multi-scale modeling errors become inevitable and accordingly change with scales. Although modeling errors should be an important part of uncertainty theory in multi-representation GIS, they have hardly been incorporated in any uncertainty models in the GIS literature. This paper develops an uncertainty model for geometric features in GIS at an arbitrary scale. It fully takes multi-scale modeling errors into account and properly reflects multi-scale complex characters through different mathematical models and rules in association with the reciprocal modeling errors at different scales. From this point of view, the new GIS uncertainty model should be more realistic to describe uncertainty of geographic information. Finally, we show that all the uncertainty models in the literature are special cases of our uncertainty model.

Keywords: uncertainty theory, modeling error, multi-scale, multi-representation GIS.

1. Introduction

Geometric uncertainty has been an important part in theoretical development and applications of geographical information systems (GIS). As a result, geometric uncertainty models for point, line and polygon have been developed.

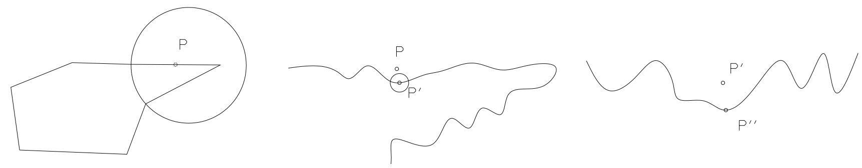
On the other hand, geographers have been aware of the fact that geographic phenomena are so complex that they depend on scale. During the process of map generalization, reducing redundant details, structures and the number of features at a specific scale can significantly improve readability and usability but at the expense of information loss in shape, size of length and area, which can be a big source of uncertainty. Coarser models always result in larger modeling errors; on the contrary, finer models have less modeling errors.

To visualize GIS uncertainty, a number of models such as epsilon-band model, e-band mode (Caspary and Scheuring, 1993) and g-band model (Shi, 1998; Shi and Liu, 2000) have been developed. Kilveri (1997) used a point uncertainty model to estimate and represent positional uncertainty of map boundaries (see also Seo and O'Hara, 2009). Lawford and Gordon (2010) estimated positional accuracy for linear features based on "time-series modelling and the other on moving block bootstrap analysis". Further development and usefulness of band models would have to consider serial correlation and modeling errors along (curved) lines. Multi-scaling adds more difficulty in building a more realistic GIS uncertainty model, which will be treated as randomized modeling errors in this paper.

The paper is organized as follows. In Section 2, we analyze multi-scaling and uncertainty in map features from the process of mathematical modeling (Cai, 2009), since digitizing phenomena can be reckoned as that of modeling relations of sample points. Section 3 will consider a statistical model for geographic features based on the assumption of random `distortions' and develop an uncertainty model for use in GIS, which fully accounts for random errors of data, modeling errors and scales. The model to be developed can also be applicable to vector data. Some simplified examples of the new uncertainty model will be discussed in Section 4, with all the uncertainty models published in the literature (see, e.g. Perkal, 1966; Blakemore, 1984, Chrisman, 1989; Dutton, 1992; Caspary and Scheuring, 1993; Shi and Tempfli, 1994; Shi, 1998; Shi and Liu, 2000) as its special cases. The simulation cases of points and lines are given in Sections 4 to demonstrate our uncertainty model. Finally, some conclusions will be summarized in Section 5.

2. Multi-scaling and uncertainty in map features

No matter in what way we obtain data --- either directly from the real world or indirectly from another map, the scale of a digital world differs from that of the real world. When scale changes from fine to coarse, details and characters are reduced, which may change geometric and semantic measurement in a digital world. On the contrary, when the real world is zoomed in immensely, various details and structures will be laid out up to the limit of information available, as shown in Fig. 1.



(a) scale A (b) scale B (c) scale C

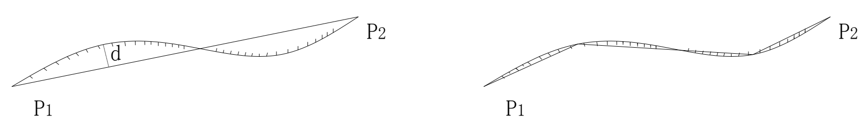
Fig.1: Illustration of modeling errors with the change of available information and scales.

Thus if a realistic uncertainty model is to be established for use in GIS, multi-scale modeling errors should be fully taken into account.

3. A multi-scale positional uncertainty model with multi-scale modeling errors

Uncertainty models for polylines and polygones such as the error band model (Dutton, 1992; Caspary and Scheuring, 1993; Shi, 2000) ignore modeling error or/and scale.

Since modeling errors always exist in any geographic modeling, multi-scale modeling errors should be considered fully in multi-representational GIS uncertainty model.



(a) original model (b) improved models with more information

Fig.2: Illustration of modeling improvement and reduced modeling errors with more information becoming available (after Li 2006)

3.1 A unified theory of positional uncertainty

An objective phenomenon W can be approximated as a model M_s plus a modeling error ξ_s, which can be written as follows:

$$W = M_s + \xi_s \tag{1}$$

where W is the real world phenomenon. Because the true model of the real world phenomena cannot be obtained, we can only use a mathematical model M_s to approximately describe W for a specific purpose at some scale and ξ_s is the corresponding modeling error. The magnitude of modeling error ξ_s may depend on the scale and model structures, however.

In this paper, we will treat ξ_s as a randomized phenomenon, and as a result, we will be able to develop a unified uncertainty theory for use in multi-representation GIS.

3.2 A point uncertainty model in GIS

Given a 2-D point *P* with the coordinate (x_s, y_s) and assuming a point incremental modeling function for a specific purpose at some scale to predict a point Q on the plane, then the predicted Q point can be written below:

$$\begin{cases} x_q = x_s + M^x_{pq} + \xi \\ y_q = y_s + M^y_{pq} + \xi \end{cases} \tag{2}$$

where (x_q, y_q) is the coordinates of the predicted point Q, (M^x_{pq}, M^y_{pq}) is the coordinate increments in the x- and y-axis, respectively, and are computable from the assumed modeling, (ξ_x, ξ_y) is the modeling errors of the x and y components due to the deviation of the assumed model from the true model. The modeling errors may change accordingly when the scale of the map changes.

We can obtain the variance-covariance matrix of the predicted point Q as follows:

$$\begin{bmatrix} \sigma^2_{qx} & \sigma_{qxy} \\ \sigma_{qyx} & \sigma^2_{qy} \end{bmatrix} = \begin{bmatrix} \sigma^2_{x} & \sigma_{xy} \\ \sigma_{yx} & \sigma^2_{y} \end{bmatrix} + \begin{bmatrix} \sigma^2_{\xi x} & \sigma_{\xi xy} \\ \sigma_{\xi yx} & \sigma^2_{\xi y} \end{bmatrix} \tag{3}$$

where $\begin{bmatrix} \sigma^2_{qx} & \sigma_{qxy} \\ \sigma_{qyx} & \sigma^2_{qy} \end{bmatrix}$ is the variance-covariance matrix of Q.

3.3 Geometric uncertainty of lines

Taking modeling errors into account, straight lines equation (4) can be more properly rewritten as:

$$\begin{cases} x_t = (1-t)x_1 + tx_2 + \xi_x \\ y_t = (1-t)y_1 + ty_2 + \xi_y \end{cases} \tag{4}$$

If the errors of the coordinates of the two end points are supposed to be independent, as is in Subsections 3.1 and 3.2, the variance-covariance matrix D_ε then becomes

$$D_\varepsilon = \begin{bmatrix} \sigma^2_{x_1} & \sigma_{x_1y_1} & \sigma_{x_1x_2} & \sigma_{x_1y_2} & 0 & 0 \\ \sigma_{y_1x_1} & \sigma^2_{y_1} & \sigma_{y_1x_2} & \sigma_{y_1y_2} & 0 & 0 \\ \sigma_{x_2x_1} & \sigma_{x_2y_1} & \sigma^2_{x_2} & \sigma_{x_2y_2} & 0 & 0 \\ \sigma_{y_2x_1} & \sigma_{y_2y_1} & \sigma_{y_2x_2} & \sigma^2_{y_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_{\xi_x} & \sigma^2_{\xi_x\xi_y} \\ 0 & 0 & 0 & 0 & \sigma^2_{\xi_y\xi_x} & \sigma^2_{\xi_y} \end{bmatrix}. \tag{5}$$

Some special cases of our unified uncertainty model will be given in the following.

Case (A): if we set $\sigma_{x_1x_2}, \sigma_{y_1y_2}, \sigma_{x_1x_2}, \sigma_{y_1y_2}, \sigma^2_{\xi_x}$ and $\sigma^2_{\xi_y}$ all to zero, and $\sigma^2_{x_1} = \sigma^2_{y_1} = \sigma^2_{x_2} = \sigma^2_{y_2} = \sigma^2$, then the positional uncertainty becomes

$$\begin{aligned}\sigma_{p_t}^2 &= 2\left[(1-t)^2\sigma^2 + t^2\sigma^2\right] \\ &= \left[4(t-\frac{1}{2})^2 + 1\right]\sigma^2\end{aligned} \tag{6}$$

Thus, we can gain an equation of the positional uncertainty of arbitrary point on a line, $\sigma^2 \le \sigma_{p_t}^2 \le 2\sigma^2$, which is actually the error band of a Line given by Dutton(1992), Caspary and Scheuring (1993), Shi and Tempfli (1994).

Case (B): if we set $\sigma_{\xi_x}^2$ and $\sigma_{\xi_y}^2$ all to zero, then the positional uncertainty becomes

$$\begin{aligned}\sigma_{p_t}^2 &= (1-t)^2\sigma_{x_1}^2 + (1-t)^2\sigma_{y_1}^2 + t^2\sigma_{x_2}^2 + t^2\sigma_{y_2}^2 \\ &+ 2(1-t)t\sigma_{x_1x_2} + 2(1-t)t\sigma_{y_1y_2}\end{aligned} \tag{7}$$

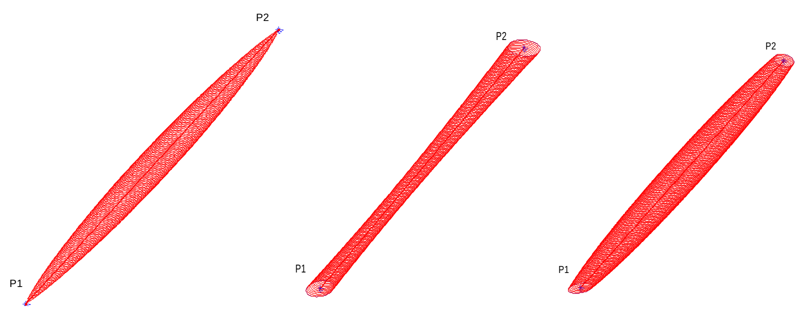
which is actually the general uncertainty model given by Shi (1998) and Shi and Liu (2000).

4. Uncertainty visualization

As the first example, let us assume that the coordinates (x_1, y_1) and (x_2, y_2) are error-free. In other words, we will only show the significance of uncertainty of the multi-scale modeling errors in our uncertainty model.

For the second example, let us assume that (1) the multi-scale modeling errors do not exist; and (2) there exists no correlation between (x_1, y_1) and (x_2, y_2). Actually, this is exactly what has been always discussed and seen in the literature of GIS uncertainty (see, e.g., Dutton 1992, Caspary and Scheuring 1993, Shi 1998).

The final example is to demonstrate the combined effect of the measurement errors of the coordinates (x_1, y_1) and (x_2, y_2) and the modeling errors (ξ_x, ξ_y) shown as Fig.3(c).



(a) no errors in the end points (b) no modeling errors (c) combined effect of both types of errors

Fig. 3 Illustration of error bands as derived from the new uncertainty model with modeling errors

5. Experiments and discussions

Here we conduct two separate experiments. They were designed to compare details of a polyline and an area between target features at the scale of 1:50000 and reference features at the scale of 1:5000 respectively, and the precision of the generation of abstract features can be both estimated.

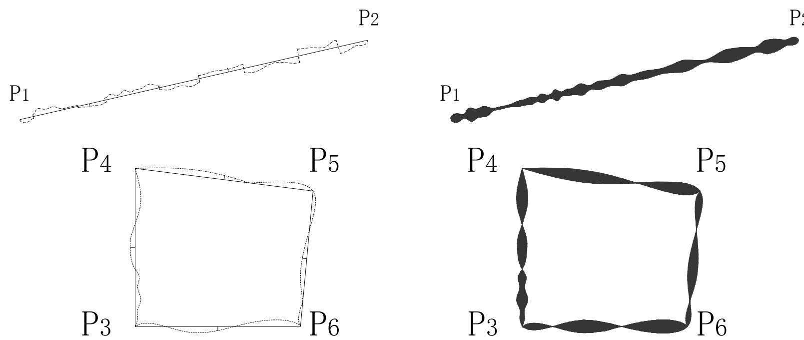


Fig. 6 Positional uncertainty estimation and visualization of a polyline and an area feature at the scale of 1:50 000

Aiming at a multi-purpose seamless master database, topographic objects at different scales, say at scale ranges from 1:5 000 to 1:500 000, the uncertainty of derivative map with 1:50 000 scale can be estimated compared to source map of 1:5 000 scale.

6. Conclusions

Although the importance of modeling errors in GIS uncertainty theory were illustrated by Alesheikh (1998) and Li (2006), no attempt has ever been made to develop a theory to automatically incorporate this type of errors. In this paper, we have treated multi-scale modeling errors as a kind of randomized phenomenon. As a result, we have successfully established a unified uncertainty model for use in multi-scale representation GIS. The new uncertainty model is most complete in the sense that it automatically takes both multi-scale modeling errors and measurement errors into account, not to mention that the correlation between the end points of a line can also be naturally implemented. By simplifying the unified geometric uncertainty model developed, we can readily obtain all the uncertainty models published in the literature as its special cases. Finally, we use visualization techniques to show the significance of multi-scale modeling errors.

References

Alesheikh A. A.(1998), *Modeling and Managing Uncertainty in Object-Based Geospatial Information System*. PhD thesis, University of Calgary, Canada.

Blakemore M.(1984), “Generalization and error in spatial data bases”.*Cartographic*, 21(2), 131-139.

Cai J.H. (2009), *Multi-scale Spatial Data Uncertainty Theory and Methods of GIS*, PhD thesis, Wuhan University, China.

Caspary W. and Scheuring R.(1993), "Positional accuracy in spatial databases", *Environment and Urban Systems*.17(2), 103-110.

Chrisman, N.R. and Yandell, B.S.(1988), "Effects of Point Error on Area Calculations: A Statistical Model", *Surveying and Mapping*, 48: 241-246.

Dunn, R., Harrison, A. R., and White, J. C.(1990), "Positional accuracy and measurement error in digital databases of land use: an empirical study". *International Journal of Geographical Information Systems*, 4, 385-98.

Dutton G.(1992),Handling Positional Uncertainty in Spatial Databases. In: *Proceedings of the 5th International Symposium on Spatial Data Handling*, vol. 2, 460–469.

Kyriakids P. C. and Goodchild M. F.(2006), "On the prediction error variance of three common spatial interpolation Schemes". *International Journal of Geographical Information Science*. 20(8), 823–855.

Lawford G. J., Gordon I.(2010), "The effect of offset correlation on positional accuracy estimation for linear features". *International Journal of Geographical Information Science*, Vol. 24, No. 1, 129–140

Li D. R.(2006), "Some Thoughts on Spatial Data Uncertainty in GIS". *Journal of Zhengzhou Institute of Surveying and Mapping*, 23(6), 391-295.

Li Z. L.(2009), *Algorithmic Foundation of Multi-Scale Spatial Representation*. CRC Press Taylor & Francis Group.

Perkal J. 1966a. On the length of empirical curves. Discussion Paper No. 10. Ann Arbor, MI: Michigan. Inter-University Community of Mathematical Geography.

Shi W. Z., and Tempfli K., 1994. Positional Uncertainty of Linear Features in GIS. In: *Proceedings of the ASPRS/ACSM Annual Meeting* (Bethesda: American Society of Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping), 696–705.

Shi W. Z., 1998. "A generic statistical approach for modelling error of geometric features in GIS". *International Journal of Geographical Information Science*, 12 (2), 131–143.

Shi W. Z.and Liu W. B., 2000. "A stochastic process-based model for the positional error of line segments in GIS". *International Journal of Geographical Information Science*, 14 (1), 51–66.

Seoa S., O'Hara C. G.. 2009, "Quality assessment of linear data". *International Journal of Geographical Information Science*, Vol. 23, No. 12, 1503–1525