

Conceptualizing Locational Uncertainty in Geographic Objects: Towards a More Balanced Approach

J. Fernando Rios¹

¹ Department of Geography, University at Buffalo, SUNY. jfernand@buffalo.edu

Abstract

A particular aspect relevant to GIScience is exploring the behavior of models subject to uncertainty associated with the spatial position of geographic objects. This work presents a way forward to address shortcomings in the way uncertainty in objects is conceptualized in spatial models by introducing a mixed probabilistic-fuzzy conceptual model of uncertainty that balances conceptual simplicity with flexibility. By taking a balanced approach, adoption of the new approach in practical applications will be encouraged.

Due to the different kinds of uncertainty that may be present in spatial models, it is envisioned that this conceptual model will serve as a foundation in analyzing mixed uncertainties in GIS applications. In order to forge a path towards this goal, challenges in systematic uncertainty and sensitivity analysis methodologies under mixed uncertainties are briefly presented.

Keywords: GIScience, positional uncertainty, fuzzy / probabilistic methods, GIS, error propagation, sensitivity analysis

1. Introduction

In order to encourage the wider adoption of uncertainty analysis of spatial models, the need for methodologies that are relatively easy to use and understand has been identified by the GIScience community (Aerts *et al.*, 2003; Karssenberg and De Jong, 2005; Zhang and Goodchild, 2002). This work presents a way forward to address that need by introducing a conceptualization of positional uncertainty in geographic objects that balances simplicity and flexibility.

2. Conceptualizing Locational Uncertainty in Objects

Due to its importance, the representation of uncertainty in objects has received attention over a long period of time. Epsilon bands (Perkal, 1956) which represent uncertainty as a simple buffer around a line eventually led to developments such as the probabilistic G-band model (Shi and Liu, 2000) in which uncertainty is calculated analytically, and approaches which combine simulation with analytical methods (Tong *et al.*, 2013). Fuzzy methods (Zadeh, 1965) have been used to model uncertainty in location, ranging from the “egg-yolk” representations (Cohn and Gotts, 1996) to fuzzy coordinate systems (Brimicombe, 1998) and many other theoretical and practical scenarios (Kronenfeld and Weeks, 2010; Usery, 1996; Zhan and Lin, 2003).

Each method has its strengths and drawbacks. For example, epsilon bands are conceptually simple however they contain comparatively less information about uncertainty than probabilistic or fuzzy approaches. Probabilistic methods are generally seen as more objective however they may have higher data demands compared to fuzzy methods. As a result, one can see that a framework which combines several approaches in a way that is easy to use is advantageous. General probabilistic-fuzzy methods for uncertainty analysis have been rigorously studied (Guyonnet *et al.*, 2003). Here, a flexible simulation-based approach for modeling uncertainty in geographic objects is presented with a view towards conducting probabilistic-fuzzy uncertainty and sensitivity analysis of spatial models.

3. An Improved Conceptual Model

To capture the advantages of probabilistic and fuzzy representations of uncertainty in a single conceptual model, a previously developed categorization of crisp objects which can easily represent varying degrees of probabilistic uncertainty (Heuvelink *et al.*, 2007) is extended to include non-crisp objects under non-probabilistic uncertainty. The first level of categorization assigns an object to be crisp or non-crisp. In this model, crisp objects are those whose location can be sharply and unambiguously defined at a particular scale of interest and whose dominant source of uncertainty is probabilistic. If those conditions are not met, then the dominant source of uncertainty is due to either a lack of knowledge about the object's uncertainty, or to indeterminacy in location. In such cases the object is categorized as non-crisp and modeled using fuzzy methods.

The second level of categorization combines the different sources of uncertainty with an object's primitive parts:

- Crisp points are single points which are well characterized by a single, but uncertain location in space which are represented by a joint probability density function (jpdf) of its X and Y coordinates (Figure 1a).
- Crisp rigid objects are objects whose boundary is defined by two or more crisp points whose relative locations are fixed when uncertainty is considered (Figure 1b). Uncertainty is characterized by a jpdf consisting of a single rotation angle and the XY coordinates of a single displacement applying to all constituent points.
- Crisp deformable objects consist of two or more crisp points whose relative locations are not fixed when uncertainty is considered (Figure 1c). Uncertainty is characterized by a jpdf of the XY coordinates of all comprising points.
- Crisp hybrid objects consist of two or more crisp points along with a rotation angle assigned to each point in which a single sequential subset of zero or more points are considered to be one of the following: (a) certain, (b) a rigid object, or (c) a deformable object. Uncertainty is characterized by a jpdf of the XY coordinates *and* rotation angle of each constituent point (Figure 1d). Rigid and deformable objects can be seen as special cases of hybrid objects.
- Non-crisp points are single points whose location, due to lack of knowledge or inherent indeterminacy of location, is not well represented by a crisp point (Figure 1e). Uncertainty is represented by fuzzy coordinates \tilde{X} , \tilde{Y} each with a membership function μ .
- Non-crisp objects are not well represented by a well-defined boundary. Uncertainty is represented differently depending on whether the boundary is considered as a collection of uncertain vertices or as a single entity (Brimicombe, 1998), resulting in a sub categorization into vertex-defined and edge-defined objects.

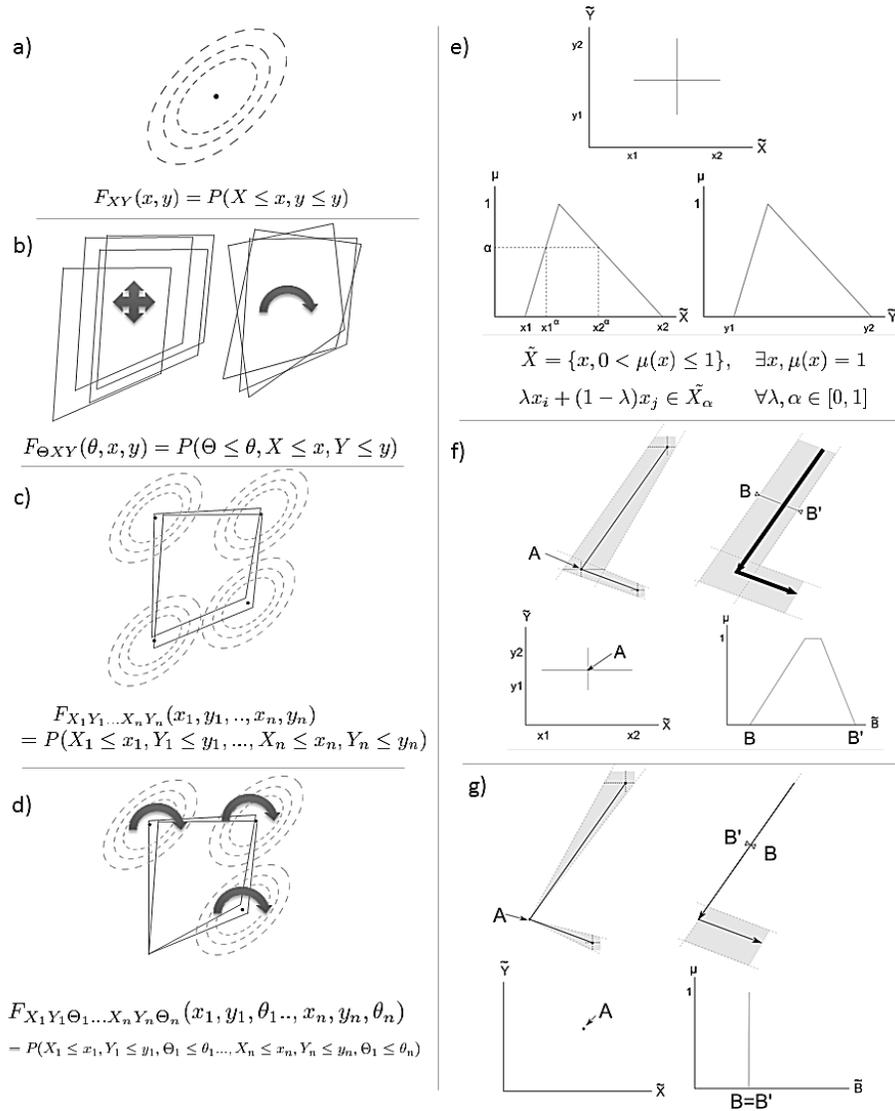


Figure 1: An improved, flexible conceptual model of positional uncertainty in objects. a) crisp point, b) crisp rigid object, c) crisp deformable object, d) crisp hybrid object (with the lower left vertex considered certain), e) non-crisp point with the corresponding membership functions of each coordinate below, f) vertex-defined non-crisp object (left), edge-defined non-crisp object (right), g) fuzzy hybrid vertex-defined and edge-defined objects.

- Vertex-defined non-crisp objects are represented by two or more non-crisp points (Figure 1f, left).
- Edge-defined non-crisp objects are defined per se as a collection of line segments. Uncertainty is represented as a fuzzy region around each segment, defined by a single fuzzy number or fuzzy interval perpendicular to the segment (Figure 1f, right).
- Non-crisp hybrid objects are vertex-defined or edge-defined objects in which a single sequential subset of vertices or edges is uncertain (Figure 1g).

3.1. Discussion

This categorization combines the strengths of probabilistic representations of uncertainty with the ability to consider a fuzzy approach if the type of uncertainty or

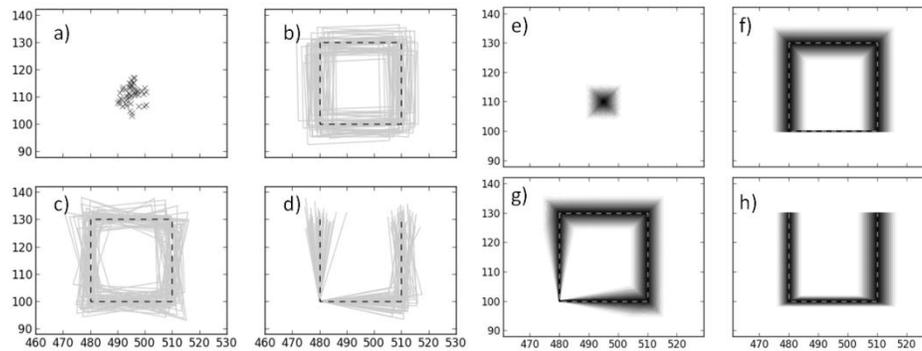


Figure 2: Crisp (a) point, (b) rigid object, (c) deformable object, (d) hybrid object with an open boundary and a certain vertex. Non-crisp (e) point, (f) edge-defined object with a certain edge, (g) vertex-defined object with a certain vertex (h) edge-defined object with open boundary and variable membership for each edge. Dashed line represents the mean value (a-d) or the maximum membership (e-h).

available data does not support a probabilistic view. Moreover, the categorization makes it easier to determine which kind of representation to use based on the type of uncertainty associated with an object. As a consequence reducing complexity, the categorization cannot consider nested uncertainties. For instance, it is not possible to account for fuzziness in an object's boundary while simultaneously considering a rotation/translation of the object (i.e., Figure 1b and f combined). Since these combinations are of limited utility, they do not greatly impact the flexibility of the conceptualization.

In order to illustrate how this conceptualization may be applied, a few simple examples are presented. A single but uncertain measurement of a point using a GPS device can readily be represented as a crisp point (Figure 2a) since error statistics can readily be obtained. The footprint of a building can be considered as a rigid object (Figure 2b) since it is fixed but there may be uncertainty in its absolute location. A deformable object (Figure 2c) can be used to model the seasonal extent of a wetland since it does not have a fixed form. If the aforementioned building footprint has uncertainty in its vertices but also an additional uncertainty which affects all vertices simultaneously, or if a vertex has been ground-truthed (Figure 2d) a hybrid object may be used. The location of an underground septic system may be modeled by a fuzzy point (Figure 2e) if the uncertainty in location cannot easily be characterized probabilistically. Lastly, edge-defined or vertex defined objects (Figure 2f,g,h) can both model the extent of a wetland when a probabilistic representation is not possible. The choice may be influenced by the specific representation of the object (e.g., a wetland polygon defined by vertices in a GIS) or the available information regarding uncertainty (e.g., it may be easier to define uncertainty with respect to edges as opposed to vertices).

5. Comparison to Other Conceptualizations

The conceptual model presented here extends a previous categorization of uncertain objects (Heuvelink *et al.*, 2007) by combining probabilistic and fuzzy representations of uncertainty. In addition, since crisp hybrid objects contain the properties of both rigid and deformable objects, by careful construction of the covariance matrix of a hybrid object, the rigid and deformable objects of Heuvelink *et al.* can be obtained.

Furthermore, by allowing for arbitrary vertices/edges (or sets of them) to be considered

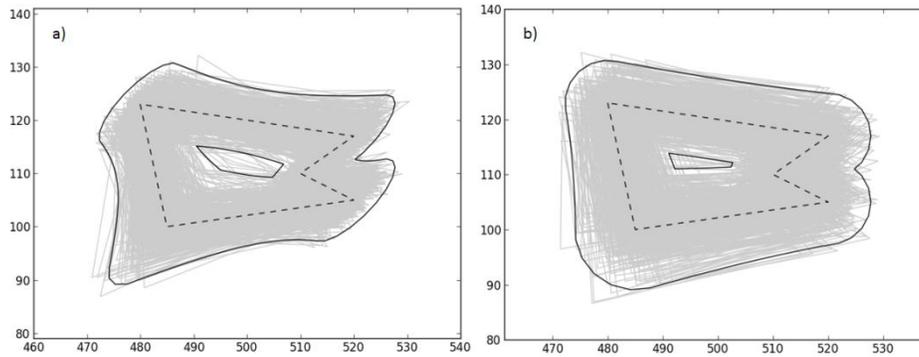


Figure 3: G-band at 95% confidence (solid line) compared with 500 realizations of a crisp deformable object (dashed line is the mean). (a) x, y coordinates of each vertex are normally distributed ($\sigma^2=2$ sq. units) and 0.8 correlated, no correlation between neighbors. (b) x, y are uncorrelated for each vertex but 0.8 correlated between immediate neighbors and 0.5 between second neighbors. The G-band only considers correlation between immediate neighbors.

as certain, greater flexibility is obtained compared to previous conceptual models (e.g., Brimicombe, 1998).

Finally, a simulation-based comparison to the analytical G-band model (Shi and Liu, 2000) shows that this conceptualization agrees with analytical results. The G-band models uncertainty around a line by considering normally distributed vertices with correlation only between immediate neighbors. When this is the case, 95.1% of simulated lines fall within the 95% G-band (Figure 3a). When there is correlation between second neighbors (a case not supported by the G-band), 97.4% of lines are contained within, an overestimation by the G-band of the true uncertainty (Figure 3b).

6. Challenges in UA and SA and Future Work

A challenge in conducting systematic uncertainty and sensitivity analysis (UA and SA) is the need for the methodologies to be able to handle mixed uncertainties. Although there have been advances in this regard (Ferson and Tucker, 2006), their application in conjunction with uncertainty conceptualizations of geographic objects has yet to be explored. Another important issue that must be addressed is how can a systematic SA be performed so that it is capable of ranking inputs (locational and non-locational) with mixed uncertainties? This information would enlighten an analysis by answering questions such as: is locational uncertainty more important than non-locational uncertainty?

Finally, In order to remain in line with the goal of a balanced approach, it remains to investigate how the uncertainty conceptualization presented here can be integrated in UA and SA with mixed uncertainty.

Acknowledgement

The author acknowledges the support from the Natural Science and Engineering Research Council of Canada postgraduate scholarship (NSERC PGS).

References

- Aerts, J. C., Goodchild, M. F., and Heuvelink, G. (2003). "Accounting for spatial uncertainty in optimization with spatial decision support systems." *Transactions in GIS*, 7(2), 211-230.
- Brimicombe, A. (1998). "A fuzzy co-ordinate system for locational uncertainty in space and time." In S. Carver (Ed.), *Innovations in GIS*, London: Taylor & Francis 143-152.
- Cohn, A. G., and Gotts, N. M. (1996). "The 'Egg-Yolk' representation of regions with indeterminate boundaries." In P. Burrough, and A. M. Frank (Ed.), *Proceedings GISDATA Specialist Meeting on Geographical Objects with Undetermined Boundaries.*, Francis Taylor 2, 171-187.
- Ferson, S., and Tucker, W. T. (2006). "Sensitivity analysis using probability bounding." *Reliability Engineering System Safety*, 91(10), 1435-1442.
- Guyonnet, D., Bourguin, B., Dubois, D., Fargier, H., Come, B., and Chiles, J.-P. (2003). "Hybrid approach for addressing uncertainty in risk assessments." *Journal of Environmental Engineering*, 129(1), 68-78.
- Heuvelink, G. B., Brown, J. D., and van Loon, E. E. (2007). "A probabilistic framework for representing and simulating uncertain environmental variables." *International Journal of Geographical Information Science*, 21(5), 497-513.
- Karssenber, D., and De Jong, K. (2005). "Dynamic environmental modelling in GIS: 2. Modelling error propagation." *International Journal of Geographical Information Science*, 19(6), 623-637.
- Kronenfeld, B. J., and Weeks, A. (2010). "A sketch-based language for representing uncertainty in the locations of origin of herbarium specimens." In N. Tate, and P. Fisher (Ed.), *Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Leicester, UK: University of Leicester, 237-240.
- Perkal, J. (1956). "On epsilon length." *Bulletin de l'Academie Polonaise des Sciences*, 4, 399-103.
- Shi, W., and Liu, W. (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.
- Tong, X., Sun, T., Fan, J., Goodchild, M. F., and Shi, W. (2013). "A statistical simulation model for positional error of line features in Geographic Information Systems (GIS)." *International Journal of Applied Earth Observation and Geoinformation*, 21, 136-148.
- Usery, E. L. (1996). "A Conceptual Framework and Fuzzy Set Implementation for Geographic Features." In P. A. Burrough, and A. U. Frank (Eds.), *Geographic Objects with Indeterminate Boundaries* (71-85). Taylor and Francis, London.
- Zadeh, L. (1965). "Fuzzy sets." *Information and Control*, 8(3), 338-353.
- Zhan, F. B., and Lin, H. (2003). "Overlay of Two Simple Polygons with Indeterminate Boundaries." *Transactions in GIS*, 7(1), 67-81.
- Zhang, J., and Goodchild, M. F. (2002). *"Uncertainty in geographical information."* Taylor & Francis.