

# Error Models in Geographic Information Systems Vector Data Using Bayesian Methods

KIMBERLY R. LOVE, KEYING YE, ERIC P. SMITH, and STEPHEN P. PRISLEY

*Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0439*

Geographic Information Systems, or GIS, has been an evolving science since its introduction. Recently, many users have become concerned with the incorporation of error analysis into GIS map products. In particular, there is concern over the error in the location of features in vector data, which relies heavily on geographic  $x$ -,  $y$ - coordinates. Current work in the field is based on bivariate normal distributions for these points, and their extension to line and polygon features. We propose here to incorporate Bayesian methodology into this existing model, which presents multiple advantages over existing methods. Bayesian methods allow for the incorporation of expert and historical knowledge and reduce the number of observations required to perform an accurate analysis. This is essential to the field of GIS where multiple observations are rare and outside knowledge is often very informative. Bayesian methods also provide results that are more easily understood by the average GIS user. We explore this addition and provide several examples based on our calculations. We conclude by discussing the advantages of Bayesian analysis for GIS vector data, and discuss our ongoing work, which is being conducted under a research grant from the National Geospatial Intelligence Agency.

KEY WORDS: Error, GIS, Vector Data, Bayesian Methods

DISCLAIMER: The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied, of the National Geospatial-Intelligence Agency or the U.S. Government.

## Introduction

Geographic Information Systems, or GIS, has become a popular means of producing, referencing, and manipulating map data. There are many current applications such as planning, engineering, land management, and environmental study ([4]). The widespread use of GIS has led to a recent concern with the error inherent in a geographic product. This is a special concern with computerized systems because users perceive the data as extremely accurate, when in fact there are often limitations on accuracy due to the processes involved in creating any map product. Among other disciplines, GIS users are turning to statistics and probability modeling to help describe and calculate risks based on this inevitable data error.

One particular type of GIS data is *vector data*, which uses a system of points based on geographic coordinates to describe point, line, and polygon features in the map setting. Shi et al. ([10], [13], [14], [12], [11]) have developed methods for analyzing the error in a vector GIS. They have based their model on a bivariate normal distribution for coordinate points. We propose here the introduction of Bayesian methodology to this model, which could improve accuracy of the error analysis as well as understanding of the results. This work is being conducted under a research grant from the National Geospatial-Intelligence Agency.

## Geographic Information Systems Background

The field of Geographic Information Systems, or GIS, is very popular in today's geographical and environmental disciplines. A GIS can be defined as a database system for storing, analyzing and manipulating spatially-referenced geographic data. Although some might expand the term to include paper maps and information, most would agree that is a computerized database structure. The majority of users encounter a GIS in the form of a computerized map and an attached database containing information about individual map features. GIS has found many applications, including research, government, military, and private uses. See D. J. Maguire([6]) or D. F. Marble ([7]) for additional general information on GIS.

There are several ways information can be recorded and presented in a GIS map. Here we will focus on one particular type of data storage known as *vector* data storage. In a vector GIS, geographic objects are represented by *points*, *lines*, and *polygons*. Each *point* that appears in a vector map represents a feature located at a specific coordinate pair. *Lines* connecting points across a map represent larger and more complicated linear features, such as rivers or roads. A *polygon* is defined as a closed set of lines, and is used to represent features with significant area. Polygons can represent features with sharp boundaries, like buildings and countries. Another common use for polygon features is to denote areas that are classified differently from other nearby areas, which in reality may have very rough boundaries. Examples of this include maps of soil types or land cover. Each polygon, line, and point feature is linked to a database through a unique identification number, which

enables the user to find information about that feature.

GIS users employ vector data for several reasons. It has the flexibility to model straight lines with a small number of vertices, or more vertices can be used to plot a more complicated line. It is a very good format to describe objects with well-defined boundaries. It also does not require the use of a lot of computing storage space, since a relatively small number of points need to be stored ([4]). Vector data has the potential to be infinitely precise, since coordinates can be stored with any desired number of decimal points.

Figure 1 is an example of a vector GIS map. This map depicts several campgrounds at a park. Polygons represent the various campgrounds, while individual cabins are represented by point features. A stream that flows through the park is represented by a line feature. In practice, information possibly including the names of the campgrounds, cabins and streams, the area of the campgrounds, and the length or average depth of each of the rivers would be available in a database. In most software applications, one can access this information by clicking on a feature.

### **Error in vector GIS data**

As the field of GIS continues to expand, users are paying more attention to error that may be present in a data set. As early as 1984, Chrisman ([3]) and others began to recognize that computers were capable of storing data much more precise than the source of the data could provide. It has perhaps taken us so long to acknowledge this fault because many users may

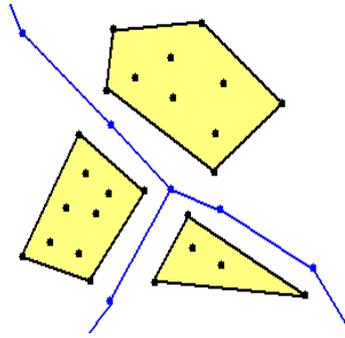


Figure 1: Example of a vector GIS map.

not wish to admit that there is some variability or uncertainty in their data, since it may affect how their clients or the public view their product ([8], [9]). Most people in the GIS community now realize, however, that there must be a system in place to describe and quantify error in GIS data, since it can be misleading and sometimes even dangerous to disregard geographic error (for example, in a military or natural disaster planning application).

There are many possible sources of error in a vector GIS map. For example, many current computerized GIS maps are paper maps that have been adapted through scanning and other digitization technology. Map providers must often manually denote particular features, and interpretation sometimes varies by operator. Another example of error in the map process occurs during field work. Instruments that measure GPS coordinates through use of satellites may not be accurate at all times of day or in all locations, and can impart error to the final map product. For a more complete description of error in a GIS map, see Bolstad and Smith ([2]).

## Current Models

When the academic community surrounding GIS initially became concerned with methods for handling error in a map product, Blakemore ([1]) created the epsilon band model for use with polygon features. While this was not a statistical model and did not allow for specific numerical calculations, it certainly laid the groundwork for more complicated models. Figure 2 demonstrates the epsilon band model for a polygon. The figure is broken into four regions; point (a) is considered to be *definitely in* the polygon, point (b) is considered *possibly in* the polygon, point (c) is considered *possibly out* of the polygon, and point (d) is classified as *definitely out*. Notice that the classification of these points is based entirely on distance from the polygon boundary as shown.

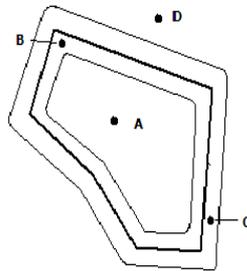


Figure 2: Blakemore's error model for a polygon.

This sort of pictorial representation is highly desirable, since maps contain primarily visual information and most users are able to interpret this very naturally. The current widely-recognized statistical model for vector data errors was initially proposed by Wenzhong

Shi in 1998 ([10]) and has a similar visual interpretation, while also relying on statistically sound principles that allow for error calculations.

Because the point is the basic unit of all features in a vector GIS, all error in any vector map feature can be traced back to points (either because the points are depicted at the wrong coordinates, or because enough points were not included to accurately describe the geographic feature). Shi et al. begin by describing a point as having a two-dimensional normal distribution as follows:

$$\mathbf{Q}_0 = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \sim N_2 \left[ \begin{bmatrix} \mu_{X_0} \\ \mu_{Y_0} \end{bmatrix}, \begin{bmatrix} \sigma_{X_0}^2 & \sigma_{X_0Y_0} \\ \sigma_{Y_0X_0} & \sigma_{Y_0}^2 \end{bmatrix} \right],$$

where  $\begin{bmatrix} \mu_{X_0} \\ \mu_{Y_0} \end{bmatrix}$  is the true value of the point coordinates,  $\begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$  is an observation taken of that point, and  $\begin{bmatrix} \sigma_{X_0}^2 & \sigma_{X_0Y_0} \\ \sigma_{Y_0X_0} & \sigma_{Y_0}^2 \end{bmatrix}$  is the variance-covariance matrix of the observations.

Note that this equation allows for variation in the X and Y directions as well as correlation between the two, which may be quite likely in some GIS applications. They went on to use this model as the basis for error analysis of lines and polygons in this and future papers ([10], [13], [14], [12], [11]). Shi and others often use the normal  $(1-\alpha)$  error ellipsoid, given

by  $(\mathbf{z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \leq \chi_{2,1-\alpha}^2$ , as a means of visually portraying the distribution of the point.

Their corresponding model for line segments is a direct extension of the point model as we have described it. Shi and Liu ([14]) begin by describing a line segment  $\mathbf{Z}_0\mathbf{Z}_1$  as a line connecting two endpoints  $\mathbf{Z}_0$  and  $\mathbf{Z}_1$ . We can geometrically represent a point on the line,  $\mathbf{Z}_t = (X_t, Y_t)$ , with the equations

$$\begin{cases} X(t) = (1-t)X_0 + tX_1 \\ Y(t) = (1-t)Y_0 + tY_1 \end{cases},$$

where  $0 \leq t \leq 1$ .

Suppose now that each endpoint has the bivariate normal point distribution, that is,

$$\mathbf{z}_i \sim N_2(\boldsymbol{\mu}_{z_i}, \boldsymbol{\Sigma}_{z_i z_i})$$

where  $i = 0,1$ . As they point out, we can further generalize this concept by allowing for the two endpoints of the line segment to be correlated. Shi and Liu characterize the joint distribution of the two endpoints as

$$\mathbf{z}_{01} \sim N_4(\boldsymbol{\mu}_{z_{01}}, \boldsymbol{\Sigma}_{z_{01}z_{01}})$$

where

$$\mathbf{z}_{01} = \begin{pmatrix} x_0 \\ y_0 \\ x_1 \\ y_1 \end{pmatrix}, \quad \boldsymbol{\mu}_{\mathbf{z}_{01}} = \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \\ \mu_{x_1} \\ \mu_{y_1} \end{pmatrix}, \quad \text{and } \boldsymbol{\Sigma}_{\mathbf{z}_{01}\mathbf{z}_{01}} = \begin{bmatrix} \sigma_{x_0}^2 & \sigma_{x_0y_0} & \sigma_{x_0x_1} & \sigma_{x_0y_1} \\ \sigma_{y_0x_0} & \sigma_{y_0}^2 & \sigma_{y_0x_1} & \sigma_{y_0y_1} \\ \sigma_{x_1x_0} & \sigma_{x_1y_0} & \sigma_{x_1}^2 & \sigma_{x_1y_1} \\ \sigma_{y_1x_0} & \sigma_{y_1y_0} & \sigma_{y_1x_1} & \sigma_{y_1}^2 \end{bmatrix}.$$

Using some basic results from linear models, Shi and Liu derive the distribution of a point on the line segment to be

$$\mathbf{Z}(t) = (X(t), Y(t))' \sim N_2(\boldsymbol{\mu}_z(t), \boldsymbol{\Sigma}_{zz}(t)),$$

where  $0 \leq t \leq 1$ ,

$$\boldsymbol{\mu}_z(t) = \begin{bmatrix} \mu_x(t) \\ \mu_y(t) \end{bmatrix} = \begin{bmatrix} (1-t)\mu_{x_0} + t\mu_{x_1} \\ (1-t)\mu_{y_0} + t\mu_{y_1} \end{bmatrix}$$

and

$$\boldsymbol{\Sigma}_{zz}(t) = \begin{bmatrix} \sigma_x^2(t) & \sigma_{xy}(t) \\ \sigma_{yx}(t) & \sigma_y^2(t) \end{bmatrix},$$

where

$$\sigma_x^2(t) = (1-t)^2\sigma_{x_0}^2 + 2t(1-t)\sigma_{x_0x_1} + t^2\sigma_{x_1}^2,$$

$$\sigma_{xy}(t) = (1 - t)^2 \sigma_{x_0 y_0}^2 + t(1 - t)(\sigma_{x_1 y_0} + \sigma_{x_0 y_1}) + t^2 \sigma_{x_1 y_1}^2,$$

$$\sigma_{yx}(t) = (1 - t)^2 \sigma_{y_0 x_0}^2 + t(1 - t)(\sigma_{y_1 x_0} + \sigma_{y_0 x_1}) + t^2 \sigma_{y_1 x_1}^2, \text{ and}$$

$$\sigma_y^2(t) = (1 - t)^2 \sigma_{y_0}^2 + 2t(1 - t)\sigma_{y_0 y_1} + t^2 \sigma_{y_1}^2.$$

They use this distribution to develop what is referred to as the *generic error band* or *G-band* model. They place a bivariate normal error ellipsoid at each point along the line segment, resulting in an infinite number of ellipsoids along the segment. Figure 3 demonstrates this concept; Shi et al. refer to the collective outer bound of the confidence ellipsoids as the *G-band* (sometimes called a *confidence region*).

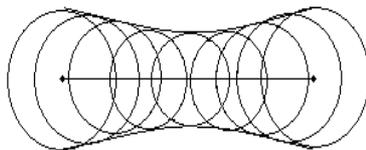


Figure 3: Shi et al.'s G-band concept.

The extension of the error model for lines to polygons is not a difficult transition. Because a polygon in GIS is, by definition, a closed set of lines, we can simply model the error in each line segment on the border of a polygon. Figure 4 shows this extension.

## Incorporating Bayesian Methods

We propose the incorporation of Bayesian methodology into the model for vector error

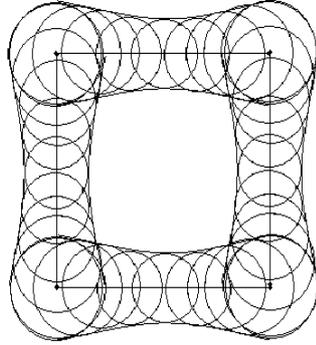


Figure 4: Extension of Shi et al.'s G-band model to a polygon.

proposed by Shi et al. The current frequentist model has many good features, including the ability to provide a visual interpretation of error, and to change the confidence level of the bivariate normal ellipsoids based on a map user's particular needs. There are several benefits, however, to the addition of Bayesian methodology, including less reliance on currently available data, the ability to incorporate expert and historical knowledge into coordinate estimates, and easier interpretation of confidence intervals.

In particular, it is very desirable to have less dependence on currently available data when analyzing GIS maps, specifically the *amount* of available data. In many cases, only one observation per point of interest is available; it is rare to have even two "identical" maps with slight variations in point coordinates at one's disposal. Map producers simply do not make many different versions of the same map. In fact, the only situations in which multiple observations are available are those involving training data (several map digitizers or field

surveyors in training are attempting to map the same area, for example) or in classroom exercises. Neither of these situations are likely to apply to most analysts' data.

Bayesian analysis is also a useful tool during the process of updating a map through the process of *ground truthing*. That is, we try to produce more accurate estimates of point coordinates by combining locations given on the map with new location data recorded on the ground. Here we consider the distribution of the coordinates on the map at hand to be the prior distribution for our points, and we update them with the distribution of the data from the field surveyors. For more discussion on these issues, see the Discussions and Conclusions section of this report.

Note that Bayes' rule will be central to our calculations. A common form of Bayes' rule is

$$p(\theta|y) \propto p(\theta)p(y|\theta),$$

which can be found in Gelman et al.'s introductory textbook on Bayesian data analysis ([5]). Incidentally, this textbook is a nice resource for learning more about general Bayes' methods.

We begin with a point  $\mathbf{P}_0$ . We assume the true value of the point's coordinates is  $\boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \end{pmatrix}$ , and that we have  $n$  random observations of this value,  $\mathbf{z}_{i0} = \begin{pmatrix} x_{i0} \\ y_{i0} \end{pmatrix}$ ,  $i = 1, \dots, n$ . We will assume that  $\boldsymbol{\mu}_0$  has a normal prior distribution with parameters  $(\boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0)$ , and that the data points  $\mathbf{z}_{i0}$  have a normal distribution with parameters  $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ .

To clarify the notation,

$$\boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \end{pmatrix}, \boldsymbol{\mu}_{0_0} = \begin{pmatrix} \mu_{x_{00}} \\ \mu_{y_{00}} \end{pmatrix}, \boldsymbol{\Lambda}_0 = \begin{pmatrix} \tau_{\mu_{x_0}}^2 & \tau_{\mu_{x_0}\mu_{y_0}} \\ \tau_{\mu_{y_0}\mu_{x_0}} & \tau_{\mu_{y_0}}^2 \end{pmatrix}, \boldsymbol{z}_{i0} = \begin{pmatrix} x_{i0} \\ y_{i0} \end{pmatrix},$$

$$\boldsymbol{Z}_0 = \begin{pmatrix} x_{10} & y_{10} \\ \vdots & \vdots \\ x_{n0} & y_{n0} \end{pmatrix}', \text{ and } \boldsymbol{\Sigma}_0 = \begin{pmatrix} \sigma_{x_0}^2 & \sigma_{x_0 y_0} \\ \sigma_{y_0 x_0} & \sigma_{y_0}^2 \end{pmatrix}.$$

**Theorem 1.** Under the above assumption, the posterior distribution of a point  $\boldsymbol{\mu}_0$  is

$$\boldsymbol{\mu}_0 | \boldsymbol{Z}_0, \boldsymbol{\Sigma}_0 \sim N(\boldsymbol{g}_0, \boldsymbol{H}_0),$$

where

$$\boldsymbol{g}_0 = (\boldsymbol{\Lambda}_0^{-1} + n\boldsymbol{\Sigma}_0^{-1})^{-1}(\boldsymbol{\Lambda}_0^{-1}\boldsymbol{\mu}_{0_0} + n\boldsymbol{\Sigma}_0^{-1}\bar{\boldsymbol{z}}_{i0}), \text{ and } \boldsymbol{H}_0 = (\boldsymbol{\Lambda}_0^{-1} + n\boldsymbol{\Sigma}_0^{-1})^{-1}.$$

*Proof.* This is an immediate result of Bayes' rule, and appears in Gelman's discussion of the multivariate normal model ([5]). □

**Corollary 1.** The equation for the  $100(1 - \alpha)\%$  confidence ellipsoid for a point  $\boldsymbol{\mu}_0$  is

$$(\boldsymbol{\mu}_0 - \mathbf{g}_0)' \mathbf{H}_0^{-1} (\boldsymbol{\mu}_0 - \mathbf{g}_0) \leq \chi_{2,1-\alpha}^2.$$

where  $\boldsymbol{\mu}_0$  is the set of points in the ellipsoid,  $\mathbf{g}_0$  is the Bayesian posterior mean for  $\boldsymbol{\mu}_0$ , and  $\mathbf{H}_0$  is the Bayesian posterior variance matrix.

*Proof.* Using the mean and variance of the posterior distribution for  $\boldsymbol{\mu}_0$ , and the fact that this distribution is bivariate normal, we calculate the bivariate normal confidence ellipse.

The corollary follows immediately. □

### Bayesian error model for line segments and polygons

The basic concept of the Bayesian error model for line segments is also very similar to the frequentist approach. Recall that any point on a line segment can be described as a function of its endpoints. Suppose then that the endpoints of a particular line segment are  $\mathbf{P}_0$  and  $\mathbf{P}_1$ , with coordinates  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ . We can then describe any point on the line with the equation

$$\boldsymbol{\mu}_\gamma = \gamma \boldsymbol{\mu}_1 + (1 - \gamma) \boldsymbol{\mu}_0,$$

where  $0 \leq \gamma \leq 1$ .

Suppose each of these coordinates has a normal prior distribution as above, with parameters  $(\boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0)$  and  $(\boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1)$  respectively. Suppose further that these two points may have

some correlation, so that the joint prior distribution of the endpoints is

$$\boldsymbol{\mu}_{01} \sim N(\boldsymbol{\mu}_{01_0}, \boldsymbol{\Lambda}_{01})$$

where

$$\boldsymbol{\mu}_{01} = \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \\ \mu_{x_1} \\ \mu_{y_1} \end{pmatrix}, \boldsymbol{\mu}_{01_0} = \begin{pmatrix} \mu_{x_{00}} \\ \mu_{y_{00}} \\ \mu_{x_{10}} \\ \mu_{y_{10}} \end{pmatrix},$$

and

$$\boldsymbol{\Lambda}_{01} = \begin{pmatrix} \tau_{\mu_{x_0}}^2 & \tau_{\mu_{x_0}\mu_{y_0}} & \tau_{\mu_{x_0}\mu_{x_1}} & \tau_{\mu_{x_0}\mu_{y_1}} \\ \tau_{\mu_{y_0}\mu_{x_0}} & \tau_{\mu_{y_0}}^2 & \tau_{\mu_{y_0}\mu_{x_1}} & \tau_{\mu_{y_0}\mu_{y_1}} \\ \tau_{\mu_{x_1}\mu_{x_0}} & \tau_{\mu_{x_1}\mu_{y_0}} & \tau_{\mu_{x_1}}^2 & \tau_{\mu_{x_1}\mu_{y_1}} \\ \tau_{\mu_{y_1}\mu_{x_0}} & \tau_{\mu_{y_1}\mu_{y_0}} & \tau_{\mu_{y_1}\mu_{x_1}} & \tau_{\mu_{y_1}}^2 \end{pmatrix}.$$

Assume we also have  $n$  independent observations on each of these endpoints,  $\mathbf{z}_{01_i} \sim$

$N(\boldsymbol{\mu}_{01}, \boldsymbol{\Sigma}_{01})$  where

$$\mathbf{z}_{01_i} = \begin{pmatrix} x_{0_i} \\ y_{0_i} \\ x_{1_i} \\ y_{1_i} \end{pmatrix}, \bar{\mathbf{z}}_{01} = \begin{pmatrix} \bar{x}_0 \\ \bar{y}_0 \\ \bar{x}_1 \\ \bar{y}_1 \end{pmatrix}, \boldsymbol{\mu}_{01} = \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \\ \mu_{x_1} \\ \mu_{y_1} \end{pmatrix},$$

and

$$\boldsymbol{\Sigma}_{01} = \begin{pmatrix} \sigma_{x_0}^2 & \sigma_{x_0 y_0} & \sigma_{x_0 x_1} & \sigma_{x_0 y_1} \\ \sigma_{y_0 x_0} & \sigma_{y_0}^2 & \sigma_{y_0 x_1} & \sigma_{y_0 y_1} \\ \sigma_{x_1 x_0} & \sigma_{x_1 y_0} & \sigma_{x_1}^2 & \sigma_{x_1 y_1} \\ \sigma_{y_1 x_0} & \sigma_{y_1 y_0} & \sigma_{y_1 x_1} & \sigma_{y_1}^2 \end{pmatrix}.$$

According to Bayes' rule, the joint conditional posterior distribution for the endpoints is

$N(\mathbf{g}_{01}, \mathbf{H}_{01})$ , where

$$\mathbf{g}_{01} = (\boldsymbol{\Lambda}_{01}^{-1} + n\boldsymbol{\Sigma}_{01}^{-1})^{-1}(\boldsymbol{\Lambda}_{01}^{-1}\boldsymbol{\mu}_{01_0} + n\boldsymbol{\Sigma}_{01}^{-1}\bar{\mathbf{z}}_{01}),$$

$$\mathbf{H}_{01} = (\boldsymbol{\Lambda}_{01}^{-1} + n\boldsymbol{\Sigma}_{01}^{-1})^{-1}.$$

For clarity of notation, we will indicate the individual elements of the posterior mean and

variance as follows:

$$\mathbf{g}_{01} = \begin{pmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \end{pmatrix} = \begin{pmatrix} g_{x_0} \\ g_{y_0} \\ g_{x_1} \\ g_{y_1} \end{pmatrix},$$

$$\mathbf{H}_{01} = \begin{pmatrix} h_{x_0}^2 & h_{x_0y_0} & h_{x_0x_1} & h_{x_0y_1} \\ h_{y_0x_0} & h_{y_0}^2 & h_{y_0x_1} & h_{y_0y_1} \\ h_{x_1x_0} & h_{x_1y_0} & h_{x_1}^2 & h_{x_1y_1} \\ h_{y_1x_0} & h_{y_1y_0} & h_{y_1x_1} & h_{y_1}^2 \end{pmatrix} = \begin{pmatrix} & & h_{x_0x_1} & h_{x_0y_1} \\ & \mathbf{H}_0 & & \\ & & h_{y_0x_1} & h_{y_0y_1} \\ h_{x_1x_0} & h_{x_1y_0} & & \\ & & & \mathbf{H}_1 \\ h_{y_1x_0} & h_{y_1y_0} & & \end{pmatrix}.$$

**Theorem 2.** Under the conditions above, the posterior conditional distribution of a point on a line segment,  $\boldsymbol{\mu}_\gamma = \gamma\boldsymbol{\mu}_1 + (1 - \gamma)\boldsymbol{\mu}_0$ , is

$$\boldsymbol{\mu}_\gamma | \boldsymbol{\Sigma}_{01}, \mathbf{Z}_{01} \sim N(\mathbf{g}_\gamma, \mathbf{H}_\gamma)$$

where

$$\begin{aligned}\mathbf{g}_\gamma &= (1 - \gamma)\mathbf{g}_0 + \gamma\mathbf{g}_1, \\ \mathbf{H}_\gamma &= \begin{pmatrix} h_{x_\gamma}^2 & h_{x_\gamma y_\gamma} \\ h_{y_\gamma x_\gamma} & h_{y_\gamma}^2 \end{pmatrix}, \\ h_{x_\gamma}^2 &= (1 - \gamma)^2 h_{x_0}^2 + 2\gamma(1 - \gamma)h_{x_0 x_1} + \gamma^2 h_{x_1}^2, \\ h_{x_\gamma y_\gamma} &= h_{y_\gamma x_\gamma} = (1 - \gamma)^2 h_{x_0 y_0} + \gamma(1 - \gamma)h_{x_0 y_1} + \gamma(1 - \gamma)h_{y_0 x_1} + \gamma^2 h_{x_1 y_1}, \\ h_{y_\gamma}^2 &= (1 - \gamma)^2 h_{y_0}^2 + 2\gamma(1 - \gamma)h_{y_0 y_1} + \gamma^2 h_{y_1}^2.\end{aligned}$$

*Proof.* Following some basic facts from linear models, the function  $\gamma\boldsymbol{\mu}_1 + (1 - \gamma)\boldsymbol{\mu}_0$  can be written as

$$\begin{pmatrix} (1 - \gamma) & 0 & \gamma & 0 \\ 0 & (1 - \gamma) & 0 & \gamma \end{pmatrix} \begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \\ \mu_{x_1} \\ \mu_{y_1} \end{pmatrix}.$$

Since  $\boldsymbol{\mu}_{01}$  is normally distributed, we know that this linear function of  $\boldsymbol{\mu}_{01}$  is normally

distributed. The mean of this distribution is

$$\begin{pmatrix} (1-\gamma) & 0 & \gamma & 0 \\ 0 & (1-\gamma) & 0 & \gamma \end{pmatrix} \begin{pmatrix} g_{x_0} \\ g_{y_0} \\ g_{x_1} \\ g_{y_1} \end{pmatrix} = (1-\gamma)\mathbf{g}_0 + \gamma\mathbf{g}_1.$$

The variance of this distribution is

$$\begin{pmatrix} (1-\gamma) & 0 & \gamma & 0 \\ 0 & (1-\gamma) & 0 & \gamma \end{pmatrix} \begin{pmatrix} h_{x_0}^2 & h_{x_0y_0} & h_{x_0x_1} & h_{x_0y_1} \\ h_{y_0x_0} & h_{y_0}^2 & h_{y_0x_1} & h_{y_0y_1} \\ h_{x_1x_0} & h_{x_1y_0} & h_{x_1}^2 & h_{x_1y_1} \\ h_{y_1x_0} & h_{y_1y_0} & h_{y_1x_1} & h_{y_1}^2 \end{pmatrix} \begin{pmatrix} (1-\gamma) & 0 \\ 0 & (1-\gamma) \\ \gamma & 0 \\ 0 & \gamma \end{pmatrix} \\ = \begin{pmatrix} h_{x_\gamma}^2 & h_{x_\gamma y_\gamma} \\ h_{y_\gamma x_\gamma} & h_{y_\gamma}^2 \end{pmatrix}$$

where the individual terms are as written in the theorem. □

Note that the result in Theorem 2 is very similar to the result in the frequentist case.

**Corollary 2.** The equation for the  $100(1 - \alpha)\%$  confidence ellipsoid for a point  $\boldsymbol{\mu}_\gamma$  is

$$(\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma)' \mathbf{H}_\gamma^{-1} (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma) \leq \chi_{2,1-\alpha}^2$$

where  $\boldsymbol{\mu}_\gamma$  is the set of points in the ellipsoid,  $\mathbf{g}_\gamma$  is the Bayesian Posterior mean for  $\boldsymbol{\mu}_\gamma$ , and  $\mathbf{H}_\gamma$  is the Bayesian posterior variance matrix.

*Proof.* Using the mean and variance of the posterior distribution for  $\boldsymbol{\mu}_\gamma$ , and the fact that this distribution is bivariate normal, we calculate a bivariate normal confidence ellipse. The corollary follows immediately.  $\square$

### Form of the posterior mean and variance for a point on a line segment

Suppose we want to find an explicit form for the posterior mean and variance of a point on a line,  $\boldsymbol{\mu}_\gamma$ . In the most general case, that is for correlated  $x$ - and  $y$ - data with

$$\boldsymbol{\Lambda}_{01} = \begin{pmatrix} \tau_{\mu_{x_0}}^2 & \tau_{\mu_{x_0}\mu_{y_0}} & \tau_{\mu_{x_0}\mu_{x_1}} & \tau_{\mu_{x_0}\mu_{y_1}} \\ \tau_{\mu_{y_0}\mu_{x_0}} & \tau_{\mu_{y_0}}^2 & \tau_{\mu_{y_0}\mu_{x_1}} & \tau_{\mu_{y_0}\mu_{y_1}} \\ \tau_{\mu_{x_1}\mu_{x_0}} & \tau_{\mu_{x_1}\mu_{y_0}} & \tau_{\mu_{x_1}}^2 & \tau_{\mu_{x_1}\mu_{y_1}} \\ \tau_{\mu_{y_1}\mu_{x_0}} & \tau_{\mu_{y_1}\mu_{y_0}} & \tau_{\mu_{y_1}\mu_{x_1}} & \tau_{\mu_{y_1}}^2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma}_{01} = \begin{pmatrix} \sigma_{x_0}^2 & \sigma_{x_0y_0} & \sigma_{x_0x_1} & \sigma_{x_0y_1} \\ \sigma_{y_0x_0} & \sigma_{y_0}^2 & \sigma_{y_0x_1} & \sigma_{y_0y_1} \\ \sigma_{x_1x_0} & \sigma_{x_1y_0} & \sigma_{x_1}^2 & \sigma_{x_1y_1} \\ \sigma_{y_1x_0} & \sigma_{y_1y_0} & \sigma_{y_1x_1} & \sigma_{y_1}^2 \end{pmatrix},$$

we can only substitute terms directly into the equations from Theorem 2 to calculate the posterior mean and variance. In many cases, however, it is possible to simplify our results,

and it is advantageous to do so. We present these cases as corollaries to our result in Theorem 2.

**Corollary 3.** Suppose the endpoints of a line segment,  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , are independent from one another, and the variance/covariance matrices of the prior and sampling distributions are

$$\boldsymbol{\Lambda}_{01} = \begin{pmatrix} \tau_{\mu_{x_0}}^2 & \tau_{\mu_{x_0}\mu_{y_0}} & 0 & 0 \\ \tau_{\mu_{y_0}\mu_{x_0}} & \tau_{\mu_{y_0}}^2 & 0 & 0 \\ 0 & 0 & \tau_{\mu_{x_1}}^2 & \tau_{\mu_{x_1}\mu_{y_1}} \\ 0 & 0 & \tau_{\mu_{y_1}\mu_{x_1}} & \tau_{\mu_{y_1}}^2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma}_{01} = \begin{pmatrix} \sigma_{x_0}^2 & \sigma_{x_0y_0} & 0 & 0 \\ \sigma_{y_0x_0} & \sigma_{y_0}^2 & 0 & 0 \\ 0 & 0 & \sigma_{x_1}^2 & \sigma_{x_1y_1} \\ 0 & 0 & \sigma_{y_1x_1} & \sigma_{y_1}^2 \end{pmatrix}.$$

The mean of the posterior distribution of  $\boldsymbol{\mu}_\gamma$  is then

$$\mathbf{g}_\gamma = \gamma(\boldsymbol{\Lambda}_1^{-1} + n\boldsymbol{\Sigma}_1^{-1})^{-1}(\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_{1_0} + n\boldsymbol{\Sigma}_1^{-1}\bar{\mathbf{z}}_1) + (1 - \gamma)(\boldsymbol{\Lambda}_0^{-1} + n\boldsymbol{\Sigma}_0^{-1})^{-1}(\boldsymbol{\Lambda}_0^{-1}\boldsymbol{\mu}_{0_0} + n\boldsymbol{\Sigma}_0^{-1}\bar{\mathbf{z}}_0),$$

and the posterior variance is  $\mathbf{H}_\gamma = \gamma^2\mathbf{H}_1 + (1 - \gamma)^2\mathbf{H}_0$ .

*Proof.* First, note that we can write  $\boldsymbol{\Lambda}_{01} = \begin{pmatrix} \boldsymbol{\Lambda}_0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_1 \end{pmatrix}$  and  $\boldsymbol{\Sigma}_{01} = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_1 \end{pmatrix}$ . Linear

models results tell us

$$\mathbf{\Lambda}_{01}^{-1} = \begin{pmatrix} \mathbf{\Lambda}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1^{-1} \end{pmatrix} \text{ and } \mathbf{\Sigma}_{01}^{-1} = \begin{pmatrix} \mathbf{\Sigma}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1^{-1} \end{pmatrix}.$$

Next, the Bayes' rule formula for the posterior mean tells us

$$\mathbf{g}_{01} = (\mathbf{\Lambda}_{01}^{-1} + n\mathbf{\Sigma}_{01}^{-1})^{-1}(\mathbf{\Lambda}_{01}^{-1}\boldsymbol{\mu}_{01_0} + n\mathbf{\Sigma}_{01}^{-1}\bar{\mathbf{z}}_{01}).$$

By writing the equation explicitly in terms of our assumed values for  $\mathbf{\Lambda}_{01}$  and  $\mathbf{\Sigma}_{01}$  and applying some linear algebra, we find

$$\begin{aligned} \mathbf{g}_{01} &= \left[ \begin{pmatrix} \mathbf{\Lambda}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1 \end{pmatrix}^{-1} + n \begin{pmatrix} \mathbf{\Sigma}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1 \end{pmatrix}^{-1} \right]^{-1} \left[ \begin{pmatrix} \mathbf{\Lambda}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1 \end{pmatrix}^{-1} \boldsymbol{\mu}_{01_0} + n \begin{pmatrix} \mathbf{\Sigma}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1 \end{pmatrix}^{-1} \bar{\mathbf{z}}_{01} \right] \\ &= \left[ \begin{pmatrix} \mathbf{\Lambda}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1^{-1} \end{pmatrix} + n \begin{pmatrix} \mathbf{\Sigma}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1^{-1} \end{pmatrix} \right]^{-1} \left[ \begin{pmatrix} \mathbf{\Lambda}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1^{-1} \end{pmatrix} \boldsymbol{\mu}_{01_0} + n \begin{pmatrix} \mathbf{\Sigma}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1^{-1} \end{pmatrix} \bar{\mathbf{z}}_{01} \right] \\ &= \begin{pmatrix} \mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1} \end{pmatrix}^{-1} \left[ \begin{pmatrix} \mathbf{\Lambda}_0^{-1}\boldsymbol{\mu}_{0_0} \\ \mathbf{\Lambda}_1^{-1}\boldsymbol{\mu}_{1_0} \end{pmatrix} + n \begin{pmatrix} \mathbf{\Sigma}_0^{-1}\bar{\mathbf{z}}_0 \\ \mathbf{\Sigma}_1^{-1}\bar{\mathbf{z}}_1 \end{pmatrix} \right] \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} (\Lambda_0^{-1} + n\Sigma_0^{-1})^{-1} & 0 \\ 0 & (\Lambda_1^{-1} + n\Sigma_1^{-1})^{-1} \end{pmatrix} \begin{pmatrix} \Lambda_0^{-1}\boldsymbol{\mu}_{00} + n\Sigma_0^{-1}\bar{\mathbf{z}}_0 \\ \Lambda_1^{-1}\boldsymbol{\mu}_{10} + n\Sigma_1^{-1}\bar{\mathbf{z}}_1 \end{pmatrix} \\
&= \begin{pmatrix} (\Lambda_0^{-1} + n\Sigma_0^{-1})^{-1}(\Lambda_0^{-1}\boldsymbol{\mu}_{00} + n\Sigma_0^{-1}\bar{\mathbf{z}}_0) \\ (\Lambda_1^{-1} + n\Sigma_1^{-1})^{-1}(\Lambda_1^{-1}\boldsymbol{\mu}_{10} + n\Sigma_1^{-1}\bar{\mathbf{z}}_1) \end{pmatrix}.
\end{aligned}$$

By applying the results of Theorem 2,  $\mathbf{g}_\gamma = (1 - \gamma)\mathbf{g}_0 + \gamma\mathbf{g}_1$ , we arrive at our stated conclusion.

Next, to get the result for the posterior variance, we again use Bayes' rule to find  $\mathbf{H}_{01} = (\Lambda_{01}^{-1} + n\Sigma_{01}^{-1})^{-1}$ . We can then write

$$\begin{aligned}
\mathbf{H}_{01} &= \begin{pmatrix} \Lambda_0^{-1} + n\Sigma_0^{-1} & \mathbf{0} \\ \mathbf{0} & \Lambda_1^{-1} + n\Sigma_1^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} (\Lambda_0^{-1} + n\Sigma_0^{-1})^{-1} & \mathbf{0} \\ \mathbf{0} & (\Lambda_1^{-1} + n\Sigma_1^{-1})^{-1} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \end{pmatrix}
\end{aligned}$$

where  $\mathbf{H}_0$  and  $\mathbf{H}_1$  are defined as in Theorem 1.

Our posterior variance is then

$$\begin{pmatrix} (1-\gamma) & 0 & \gamma & 0 \\ 0 & (1-\gamma) & 0 & \gamma \end{pmatrix} \begin{pmatrix} h_{x_0}^2 & h_{x_0y_0} & 0 & 0 \\ h_{y_0x_0} & h_{y_0}^2 & 0 & 0 \\ 0 & 0 & h_{x_1}^2 & h_{x_1y_1} \\ 0 & 0 & h_{y_1x_1} & h_{y_1}^2 \end{pmatrix} \begin{pmatrix} (1-\gamma) & 0 \\ 0 & (1-\gamma) \\ \gamma & 0 \\ 0 & \gamma \end{pmatrix},$$

from the proof of Theorem 2.

Multiplying through this equation gives us

$$\begin{pmatrix} (1-\gamma)^2 h_{x_0}^2 + \gamma^2 h_{x_1}^2 & (1-\gamma)^2 h_{x_0y_0} + \gamma^2 h_{x_1y_1} \\ (1-\gamma)^2 h_{y_0x_0} + \gamma^2 h_{y_1x_1} & (1-\gamma)^2 h_{y_0}^2 + \gamma^2 h_{y_1}^2 \end{pmatrix} = \gamma^2 \mathbf{H}_1 + (1-\gamma)^2 \mathbf{H}_0,$$

which is the result stated in the corollary. □

Corollary 3 provides the posterior mean and variance for a general situation that allows the  $x$ - and  $y$ - coordinates within each endpoint to be correlated, without correlation between endpoints. This may or may not be a valid assumption. In the case of GPS instrument error, for example, one instrument may be more likely to always read a high  $x$ - coordinate, or an instrument taking a different sample might always read a low one, meaning endpoints will have correlated error. If different instruments were used at each point, however, there may be no such correlation. Additionally, in the case of manual map digitization, a well-

trained technician may not demonstrate any trend in error between endpoints. The following corollaries provide details for the subset of situations in which there is no correlation between or within the endpoints.

**Corollary 4.** Suppose now that there is no correlation between the endpoints  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , and additionally, there is no correlation between the  $x$ - and  $y$ - coordinates at each endpoint. That is,

$$\boldsymbol{\Lambda}_{01} = \begin{pmatrix} \tau_{\mu_{x_0}}^2 & 0 & 0 & 0 \\ 0 & \tau_{\mu_{y_0}}^2 & 0 & 0 \\ 0 & 0 & \tau_{\mu_{x_1}}^2 & 0 \\ 0 & 0 & 0 & \tau_{\mu_{y_1}}^2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma}_{01} = \begin{pmatrix} \sigma_{x_0}^2 & 0 & 0 & 0 \\ 0 & \sigma_{y_0}^2 & 0 & 0 \\ 0 & 0 & \sigma_{x_1}^2 & 0 \\ 0 & 0 & 0 & \sigma_{y_1}^2 \end{pmatrix}.$$

The mean of the posterior distribution of  $\boldsymbol{\mu}_\gamma$  is

$$\begin{aligned} & \gamma \left[ \begin{pmatrix} \frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu_{x_1}}^2} & 0 \\ 0 & \frac{\sigma_{y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu_{y_1}}^2} \end{pmatrix} \boldsymbol{\mu}_{1_0} + n \begin{pmatrix} \frac{\tau_{\mu_{x_1}}^2}{\sigma_{x_1}^2 + n\tau_{\mu_{x_1}}^2} & 0 \\ 0 & \frac{\tau_{\mu_{y_1}}^2}{\sigma_{y_1}^2 + n\tau_{\mu_{y_1}}^2} \end{pmatrix} \bar{\boldsymbol{z}}_1 \right] \\ & + (1 - \gamma) \left[ \begin{pmatrix} \frac{\sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu_{x_0}}^2} & 0 \\ 0 & \frac{\sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu_{y_0}}^2} \end{pmatrix} \boldsymbol{\mu}_{0_0} + n \begin{pmatrix} \frac{\tau_{\mu_{x_0}}^2}{\sigma_{x_0}^2 + n\tau_{\mu_{x_0}}^2} & 0 \\ 0 & \frac{\tau_{\mu_{y_0}}^2}{\sigma_{y_0}^2 + n\tau_{\mu_{y_0}}^2} \end{pmatrix} \bar{\boldsymbol{z}}_0 \right] \end{aligned}$$

and the variance of the posterior distribution is

$$\gamma^2 \begin{pmatrix} \frac{\tau_{\mu x_1}^2 \sigma_{x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{\tau_{\mu y_1}^2 \sigma_{y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu y_1}^2} \end{pmatrix} + (1 - \gamma)^2 \begin{pmatrix} \frac{\tau_{\mu x_0}^2 \sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\tau_{\mu y_0}^2 \sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix}.$$

*Proof.* Because this is a special case of Corollary 3, we know that the mean of the posterior distribution is

$$\mathbf{g}_\gamma = \gamma(\mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1})^{-1}(\mathbf{\Lambda}_1^{-1}\boldsymbol{\mu}_{10} + n\mathbf{\Sigma}_1^{-1}\bar{\mathbf{z}}_1) + (1 - \gamma)(\mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1})^{-1}(\mathbf{\Lambda}_0^{-1}\boldsymbol{\mu}_{00} + n\mathbf{\Sigma}_0^{-1}\bar{\mathbf{z}}_0),$$

and the posterior variance is  $\mathbf{H}_\gamma = \gamma^2\mathbf{H}_1 + (1 - \gamma)^2\mathbf{H}_0$ . Imposing the additional condition that the  $x$ - and  $y$ - coordinates at each endpoint are uncorrelated, we know

$$\mathbf{\Lambda}_0^{-1} = \begin{pmatrix} \frac{1}{\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu y_0}^2} \end{pmatrix}, \quad \mathbf{\Lambda}_1^{-1} = \begin{pmatrix} \frac{1}{\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu y_1}^2} \end{pmatrix}, \quad \mathbf{\Sigma}_0^{-1} = \begin{pmatrix} \frac{1}{\sigma_{x_0}^2} & 0 \\ 0 & \frac{1}{\sigma_{y_0}^2} \end{pmatrix},$$

and

$$\mathbf{\Sigma}_1^{-1} = \begin{pmatrix} \frac{1}{\sigma_{x_1}^2} & 0 \\ 0 & \frac{1}{\sigma_{y_1}^2} \end{pmatrix}.$$

Inserting these results into our previous equations, we get

$$\begin{aligned} \mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1} &= \begin{pmatrix} \frac{1}{\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu y_0}^2} \end{pmatrix} + \begin{pmatrix} \frac{n}{\sigma_{x_0}^2} & 0 \\ 0 & \frac{n}{\sigma_{y_0}^2} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2}{\tau_{\mu x_0}^2 \sigma_{x_0}^2} & 0 \\ 0 & \frac{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2}{\tau_{\mu y_0}^2 \sigma_{y_0}^2} \end{pmatrix} \\ \Rightarrow (\mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1})^{-1} &= \begin{pmatrix} \frac{\tau_{\mu x_0}^2 \sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\tau_{\mu y_0}^2 \sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix}. \end{aligned}$$

Calculation of  $(\mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1})^{-1}$  is similar.

We can now calculate the posterior mean and variance for the general no-covariance case.

$$\begin{aligned}
\mathbf{g}_\gamma &= \gamma \begin{pmatrix} \frac{\tau_{\mu x_1}^2 \sigma_{x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{\tau_{\mu y_1}^2 \sigma_{y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu y_1}^2} \end{pmatrix} \left[ \begin{pmatrix} \frac{1}{\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu y_1}^2} \end{pmatrix} \boldsymbol{\mu}_{1_0} + n \begin{pmatrix} \frac{1}{\sigma_{x_1}^2} & 0 \\ 0 & \frac{1}{\sigma_{y_1}^2} \end{pmatrix} \bar{\mathbf{z}}_1 \right] \\
&+ (1 - \gamma) \begin{pmatrix} \frac{\tau_{\mu x_0}^2 \sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\tau_{\mu y_0}^2 \sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix} \left[ \begin{pmatrix} \frac{1}{\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu y_0}^2} \end{pmatrix} \boldsymbol{\mu}_{0_0} + n \begin{pmatrix} \frac{1}{\sigma_{x_0}^2} & 0 \\ 0 & \frac{1}{\sigma_{y_0}^2} \end{pmatrix} \bar{\mathbf{z}}_0 \right] \\
&= \gamma \left[ \begin{pmatrix} \frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{\sigma_{y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu y_1}^2} \end{pmatrix} \boldsymbol{\mu}_{1_0} + n \begin{pmatrix} \frac{\tau_{\mu x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{\tau_{\mu y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu y_1}^2} \end{pmatrix} \bar{\mathbf{z}}_1 \right] \\
&+ (1 - \gamma) \left[ \begin{pmatrix} \frac{\sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix} \boldsymbol{\mu}_{0_0} + n \begin{pmatrix} \frac{\tau_{\mu x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\tau_{\mu y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix} \bar{\mathbf{z}}_0 \right], \\
\mathbf{H}_\gamma &= \gamma^2 \begin{pmatrix} \frac{\tau_{\mu x_1}^2 \sigma_{x_1}^2}{\sigma_{x_1}^2 + n\tau_{\mu x_1}^2} & 0 \\ 0 & \frac{\tau_{\mu y_1}^2 \sigma_{y_1}^2}{\sigma_{y_1}^2 + n\tau_{\mu y_1}^2} \end{pmatrix} + (1 - \gamma)^2 \begin{pmatrix} \frac{\tau_{\mu x_0}^2 \sigma_{x_0}^2}{\sigma_{x_0}^2 + n\tau_{\mu x_0}^2} & 0 \\ 0 & \frac{\tau_{\mu y_0}^2 \sigma_{y_0}^2}{\sigma_{y_0}^2 + n\tau_{\mu y_0}^2} \end{pmatrix}.
\end{aligned}$$

□

Endpoints certainly may have correlation between their  $x$ - and  $y$ - coordinates. If, for example, similar instruments were used at the same endpoint to collect data at different times, this could cause some correlation between the errors in its coordinates. Depending on the time of day, for instance, GPS instruments relying on satellite information may have similar types of error in each coordinate, based on the changing positions of the satellites.

In some cases, though, for instance manual digitization, it is probably a valid assumption that there is no correlation in error between coordinates.

**Corollary 5.** Suppose there is no correlation between the endpoints  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , and there is no correlation between the coordinates at each endpoint. Additionally, suppose each endpoint has equal variance at its  $x$ - and  $y$ - coordinates (although it may be different at each endpoint); that is, we know that  $\boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}$ ,  $\boldsymbol{\Sigma}_0 = \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_0^2 \end{pmatrix}$ ,  $\boldsymbol{\Lambda}_1 = \begin{pmatrix} \tau_1^2 & 0 \\ 0 & \tau_1^2 \end{pmatrix}$ ,

and  $\boldsymbol{\Lambda}_0 = \begin{pmatrix} \tau_0^2 & 0 \\ 0 & \tau_0^2 \end{pmatrix}$ .

The posterior mean is then

$$\mathbf{g}_\gamma = \gamma \left( \frac{\sigma_1^2}{\sigma_1 + n\tau_1^2} \boldsymbol{\mu}_{10} + \frac{n\tau_1^2}{\sigma_1 + n\tau_1^2} \bar{\mathbf{z}}_1 \right) + (1 - \gamma) \left( \frac{\sigma_0^2}{\sigma_0 + n\tau_0^2} \boldsymbol{\mu}_{00} + \frac{n\tau_0^2}{\sigma_0 + n\tau_0^2} \bar{\mathbf{z}}_0 \right),$$

and the posterior variance is

$$\mathbf{H}_\gamma = \left[ \gamma^2 \left( \frac{\tau_1^2 \sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) + (1 - \gamma)^2 \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \right] \mathbf{I}.$$

*Proof.* It is easy to first calculate  $(\mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1})^{-1}$  and  $(\mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1})^{-1}$ .

$$\mathbf{\Lambda}_1^{-1} = \begin{pmatrix} \frac{1}{\tau_1^2} & 0 \\ 0 & \frac{1}{\tau_1^2} \end{pmatrix}, \quad \mathbf{\Lambda}_0^{-1} = \begin{pmatrix} \frac{1}{\tau_0^2} & 0 \\ 0 & \frac{1}{\tau_0^2} \end{pmatrix},$$

$$\mathbf{\Sigma}_1^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_1^2} \end{pmatrix}, \quad \text{and} \quad \mathbf{\Sigma}_0^{-1} = \begin{pmatrix} \frac{1}{\sigma_0^2} & 0 \\ 0 & \frac{1}{\sigma_0^2} \end{pmatrix}.$$

We can then calculate

$$\mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1} = \begin{pmatrix} \frac{1}{\tau_1^2} + n\frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\tau_1^2} + n\frac{1}{\sigma_1^2} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_1^2 + n\tau_1^2}{\tau_1^2\sigma_1^2} & 0 \\ 0 & \frac{\sigma_1^2 + n\tau_1^2}{\tau_1^2\sigma_1^2} \end{pmatrix}$$

and

$$\mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1} = \begin{pmatrix} \frac{1}{\tau_0^2} + n\frac{1}{\sigma_0^2} & 0 \\ 0 & \frac{1}{\tau_0^2} + n\frac{1}{\sigma_0^2} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_0^2 + n\tau_0^2}{\tau_0^2\sigma_0^2} & 0 \\ 0 & \frac{\sigma_0^2 + n\tau_0^2}{\tau_0^2\sigma_0^2} \end{pmatrix}.$$

This gives us

$$(\mathbf{\Lambda}_1^{-1} + n\mathbf{\Sigma}_1^{-1})^{-1} = \begin{pmatrix} \frac{\tau_1^2\sigma_1^2}{\sigma_1^2 + n\tau_1^2} & 0 \\ 0 & \frac{\tau_1^2\sigma_1^2}{\sigma_1^2 + n\tau_1^2} \end{pmatrix} = \left( \frac{\tau_1^2\sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) \mathbf{I}$$

and

$$(\mathbf{\Lambda}_0^{-1} + n\mathbf{\Sigma}_0^{-1})^{-1} = \begin{pmatrix} \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} & 0 \\ 0 & \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \end{pmatrix} = \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \mathbf{I}.$$

We can now calculate the posterior mean and variance.

$$\begin{aligned} \mathbf{g}_\gamma &= \gamma \left[ \left( \frac{\tau_1^2 \sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) \mathbf{I} \right] \left[ \frac{1}{\tau_1^2} \boldsymbol{\mu}_{10} + \frac{n}{\sigma_1^2} \bar{\mathbf{z}}_1 \right] + (1 - \gamma) \left[ \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \mathbf{I} \right] \left[ \frac{1}{\tau_0^2} \boldsymbol{\mu}_{00} + \frac{n}{\sigma_0^2} \bar{\mathbf{z}}_0 \right] \\ &= \gamma \left( \frac{\sigma_1^2}{\sigma_1^2 + n\tau_1^2} \boldsymbol{\mu}_{10} + \frac{n\tau_1^2}{\sigma_1^2 + n\tau_1^2} \bar{\mathbf{z}}_1 \right) + (1 - \gamma) \left( \frac{\sigma_0^2}{\sigma_0^2 + n\tau_0^2} \boldsymbol{\mu}_{00} + \frac{n\tau_0^2}{\sigma_0^2 + n\tau_0^2} \bar{\mathbf{z}}_0 \right), \\ \mathbf{H}_\gamma &= \gamma^2 \left( \frac{\tau_1^2 \sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) \mathbf{I} + (1 - \gamma)^2 \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \mathbf{I} \\ &= \left[ \gamma^2 \left( \frac{\tau_1^2 \sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) + (1 - \gamma)^2 \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \right] \mathbf{I}. \end{aligned}$$

□

This may be a valid situation, especially in the case of human error. Suppose, for example, that two maps of adjoining areas are combined into a larger map. Each map may have been created by a different agency – possibly with different methods – and each agency may have emphasized a different standard of accuracy in terms of identifying exact point location. Therefore, some points on the large map may have small variance in error while others have larger error variance. It would, however, be likely within an agency that error in the  $x$ - and  $y$ - directions at each point would be quite similar.

**Corollary 6.** Suppose again that there is no correlation between the endpoints  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , and there is no correlation between the  $x$ - and  $y$ - coordinates at each endpoint. Suppose also that the variance in the  $x$  coordinates is similar between the endpoints, as is the variance in the  $y$  coordinates; that is,  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$ , and  $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_0 = \boldsymbol{\Lambda} = \begin{pmatrix} \tau_{\mu_x}^2 & 0 \\ 0 & \tau_{\mu_y}^2 \end{pmatrix}$ .

In this case, the posterior mean for a point  $\boldsymbol{\mu}_\gamma$  is

$$\mathbf{g}_\gamma = \begin{pmatrix} \frac{\sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} [\gamma \boldsymbol{\mu}_{1_0} + (1 - \gamma) \boldsymbol{\mu}_{0_0}] + \begin{pmatrix} \frac{\tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} [\gamma \bar{\mathbf{z}}_1 + (1 - \gamma) \bar{\mathbf{z}}_0],$$

and the posterior variance is

$$\mathbf{H}_\gamma = (2\gamma^2 - 2\gamma + 1) \begin{pmatrix} \frac{\sigma_x^2 \tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2 \tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix}.$$

*Proof.* Again calculating the individual terms involved in  $(\boldsymbol{\Lambda}^{-1} + n\boldsymbol{\Sigma}^{-1})^{-1}$ , we have

$$\boldsymbol{\Lambda}^{-1} = \begin{pmatrix} \frac{1}{\tau_{\mu_x}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu_y}^2} \end{pmatrix}, \text{ and } \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{pmatrix}.$$

We can then calculate

$$\mathbf{\Lambda}^{-1} + n\mathbf{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\tau_{\mu_x}^2} + n\frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\tau_{\mu_y}^2} + n\frac{1}{\sigma_y^2} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_x^2 + n\tau_{\mu_x}^2}{\tau_{\mu_x}^2 \sigma_x^2} & 0 \\ 0 & \frac{\sigma_y^2 + n\tau_{\mu_y}^2}{\tau_{\mu_y}^2 \sigma_y^2} \end{pmatrix}.$$

This gives us

$$(\mathbf{\Lambda}^{-1} + n\mathbf{\Sigma}^{-1})^{-1} = \begin{pmatrix} \frac{\tau_{\mu_x}^2 \sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2 \sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix}.$$

We can now calculate the posterior mean and variance.

$$\begin{aligned}
\mathbf{g}_\gamma &= \gamma \begin{pmatrix} \frac{\tau_{\mu_x}^2 \sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2 \sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} \left[ \begin{pmatrix} \frac{1}{\tau_{\mu_x}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu_y}^2} \end{pmatrix} \boldsymbol{\mu}_{1_0} + \begin{pmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{pmatrix} \bar{\mathbf{z}}_1 \right] \\
&+ (1 - \gamma) \begin{pmatrix} \frac{\tau_{\mu_x}^2 \sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2 \sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} \left[ \begin{pmatrix} \frac{1}{\tau_{\mu_x}^2} & 0 \\ 0 & \frac{1}{\tau_{\mu_y}^2} \end{pmatrix} \boldsymbol{\mu}_{0_0} + \begin{pmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{pmatrix} \bar{\mathbf{z}}_0 \right] \\
&= \begin{pmatrix} \frac{\sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} [\gamma \boldsymbol{\mu}_{1_0} + (1 - \gamma) \boldsymbol{\mu}_{0_0}] + \begin{pmatrix} \frac{\tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} [\gamma \bar{\mathbf{z}}_1 + (1 - \gamma) \bar{\mathbf{z}}_0], \\
&+ (1 - \gamma) \left[ \begin{pmatrix} \frac{\sigma_x^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} \boldsymbol{\mu}_{0_0} + \begin{pmatrix} \frac{\tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} \bar{\mathbf{z}}_0 \right] \\
\mathbf{H}_\gamma &= \gamma^2 \begin{pmatrix} \frac{\sigma_x^2 \tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2 \tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} + (1 - \gamma)^2 \begin{pmatrix} \frac{\sigma_x^2 \tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2 \tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix} \\
&= (2\gamma^2 - 2\gamma + 1) \begin{pmatrix} \frac{\sigma_x^2 \tau_{\mu_x}^2}{\sigma_x^2 + n\tau_{\mu_x}^2} & 0 \\ 0 & \frac{\sigma_y^2 \tau_{\mu_y}^2}{\sigma_y^2 + n\tau_{\mu_y}^2} \end{pmatrix}.
\end{aligned}$$

□

This is a possibility in a case where a map has been digitized in a situation where  $x$ - and  $y$ - distances are not displayed at the same scale. This can happen locally, for example, when using certain coordinate systems. A single technician digitizing a map will likely make

the same size error at all points based on the visual display available, in both the  $x$ - and  $y$ -directions. This means that the true size of error as measured on the ground will be different for  $x$ - and  $y$ - coordinates at a single point, but similar over the scope of the map.

**Corollary 7.** Suppose again that there is no correlation between the endpoints  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , and there is no correlation between the  $x$ - and  $y$ - coordinates at each endpoint. Suppose additionally that all coordinate variances are similar; that is,  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$ ,

and  $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_0 = \boldsymbol{\Lambda} = \begin{pmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{pmatrix}$ .

In this case, the posterior mean for a point  $\boldsymbol{\mu}_\gamma$  is

$$\mathbf{g}_\gamma = \left( \frac{\tau^2 \sigma^2}{\tau^2 + n\sigma^2} \right) \left[ (1 - \gamma) \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{0_0} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_0 \right) + \gamma \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{1_0} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_1 \right) \right],$$

and the posterior variance is

$$\mathbf{H}_\gamma = (2\gamma^2 - 2\gamma + 1) \frac{\tau^2 \sigma^2}{\sigma^2 + n\tau^2} \mathbf{I}.$$

*Proof.* Again calculating the individual terms involved in the computations, we have

$$\mathbf{\Lambda}^{-1} = \begin{pmatrix} \frac{1}{\tau^2} & 0 \\ 0 & \frac{1}{\tau^2} \end{pmatrix}, \text{ and } \mathbf{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix}.$$

We can then calculate

$$\mathbf{\Lambda}^{-1} + n\mathbf{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\tau^2} + n\frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\tau^2} + n\frac{1}{\sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{\sigma^2 + n\tau^2}{\tau^2\sigma^2} & 0 \\ 0 & \frac{\sigma^2 + n\tau^2}{\tau^2\sigma^2} \end{pmatrix}.$$

This gives us

$$(\mathbf{\Lambda}^{-1} + n\mathbf{\Sigma}^{-1})^{-1} = \begin{pmatrix} \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} & 0 \\ 0 & \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \end{pmatrix} = \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I}. \quad (1)$$

We can now easily calculate the posterior mean and variance.

$$\begin{aligned} \mathbf{g}_\gamma &= \gamma \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I} \left[ \begin{pmatrix} \frac{1}{\tau^2} & 0 \\ 0 & \frac{1}{\tau^2} \end{pmatrix} \boldsymbol{\mu}_{10} + \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix} \bar{\mathbf{z}}_1 \right] \\ &+ (1 - \gamma) \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I} \left[ \begin{pmatrix} \frac{1}{\tau^2} & 0 \\ 0 & \frac{1}{\tau^2} \end{pmatrix} \boldsymbol{\mu}_{00} + \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix} \bar{\mathbf{z}}_0 \right] \\ &= \left( \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \right) \left[ \gamma \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{10} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_1 \right) + (1 - \gamma) \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{00} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_0 \right) \right], \\ \mathbf{H}_\gamma &= \gamma^2 \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I} + (1 - \gamma^2) \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I} \\ &= (2\gamma^2 - 2\gamma + 1) \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2} \mathbf{I}. \end{aligned}$$

□

This is something of a best-case scenario, and would certainly simplify subsequent calculations using error variance. It may even be a realistic situation, for example, when all maps involved have been digitized by a single technician. It seems reasonable that human error would tend to be randomly and evenly distributed at all points. There are many cases in which this assumption is not valid, however, which we have already mentioned. One example is multiple technicians digitizing separate parts of a map. Error in instrument readings is another example, since those types of idiosyncracies in an instrument are likely to be correlated across a map. Therefore, anyone hoping to accurately discuss positional error in a GIS product should seriously consider the types of correlation and variance that may occur, as it may change the model considerably.

## **Confidence Region Boundary Calculation**

Although pictures of the boundary region on a line segment are frequent in the available literature on error in vector GIS, to our knowledge an explicit formula for the confidence bound has never been calculated. (The bound is usually formed by placing confidence ellipses at a large number of points along the line segment.) We believe that finding the formula for the confidence bound will greatly aid us in solving current probability problems in vector data, for example, the point in polygon problem we have discussed here.

In order to compute the boundary, we must first assume that each of the infinite points on the line has a confidence ellipsoid around it. This means we assume that the joint prior distribution of the endpoints of the line segment is  $\boldsymbol{\mu}_{01} \sim N(\boldsymbol{\mu}_{01_0}, \boldsymbol{\Lambda}_{01})$ , and that the data distribution is  $\mathbf{z}_{01} \sim N(\boldsymbol{\mu}_{01}, \boldsymbol{\Sigma}_{01})$ .

Next, it is clear that the points on each error ellipsoid that are at the furthest perpendicular distance from the line segment will be on the boundary of its confidence region. For a demonstration of this fact, see Figure 5. Figure a) shows this for circular point error ellipsoids, and figure b) shows this in a more general case. A proof will follow in an upcoming paper.

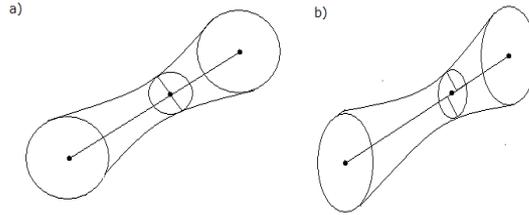


Figure 5: Points on the boundary of the confidence region.

While there are several ways to calculate this boundary point in the general case, they are complicated calculations we will include in a an upcoming paper. If, however, the error ellipsoid around  $\boldsymbol{\mu}_\gamma$  is circular, the boundary can be calculated fairly easily as follows.

**Lemma 1.** In the case covered by corollary 5, in which each endpoint of a line segment is normally distributed and has equal prior and sampling variance for both of its coordinates,

the error ellipsoid at  $\boldsymbol{\mu}_\gamma$  is circular.

*Proof.* We know from corollary 5 that  $\mathbf{H}_\gamma = \left[ \gamma^2 \left( \frac{\tau_1^2 \sigma_1^2}{\sigma_1^2 + n\tau_1^2} \right) + (1 - \gamma)^2 \left( \frac{\tau_0^2 \sigma_0^2}{\sigma_0^2 + n\tau_0^2} \right) \right] \mathbf{I} = h_\gamma \mathbf{I}$  (for simplified notation). We can then use corollary 2 which tells us that the boundary of the confidence ellipsoid at point  $\boldsymbol{\mu}_\gamma$  is  $(\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma)' \mathbf{H}_\gamma^{-1} (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma) = (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma)' h_\gamma^{-1} \mathbf{I} (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma) = \chi_{2,1-\alpha}^2$ .

This implies

$$\begin{aligned} h_\gamma^{-1} (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma)' (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma) &= \chi_{2,1-\alpha}^2 \\ \Rightarrow (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma)' (\boldsymbol{\mu}_\gamma - \mathbf{g}_\gamma) &= h_\gamma \chi_{2,1-\alpha}^2 \\ \Rightarrow ((\mu_{x_\gamma} - g_{x_\gamma})^2 + (\mu_{y_\gamma} - g_{y_\gamma})^2) &= h_\gamma \chi_{2,1-\alpha}^2 \\ \Rightarrow \sqrt{(\mu_{x_\gamma} - g_{x_\gamma})^2 + (\mu_{y_\gamma} - g_{y_\gamma})^2} &= \sqrt{h_\gamma \chi_{2,1-\alpha}^2}. \end{aligned}$$

The left side of this equation is the well-known geometric formula for the distance between  $\boldsymbol{\mu}_\gamma$  and  $\mathbf{g}_\gamma$ . The fact that this distance is equal to a constant,  $\sqrt{h_\gamma \chi_{2,1-\alpha}^2}$ , means that the distance to the boundary from  $\mathbf{g}_\gamma$  is constant, implying that the boundary is circular.  $\square$

Note that this lemma applies in particular to the subset of cases discussed in corollary 7, in which all coordinates at both endpoints have the same prior and sampling variance.

In the following theorem, we give the boundary for the line segment's confidence region in the event of a circular confidence ellipsoid at  $\boldsymbol{\mu}_\gamma$ .

**Theorem 3.** In the case covered by corollary 5, in which each endpoint of a line segment is

normally distributed and has equal prior and sampling variance for both of its coordinates, the point on the boundary of the line segment associated with point  $\boldsymbol{\mu}_\gamma$  is

$$\begin{pmatrix} b_{x_\gamma} \\ b_{y_\gamma} \end{pmatrix} = \begin{pmatrix} g_{x_\gamma} \pm \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \\ g_{y_\gamma} \mp \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right) \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \end{pmatrix}.$$

*Proof.* The points on the circular error ellipse at  $\boldsymbol{\mu}_\gamma$  that are farthest from the line segment (thus on the boundary of its confidence region) are the points on a line perpendicular to the line segment at  $\boldsymbol{\mu}_\gamma$ . This is not true in the general case if the error ellipse is not circular. For a heuristic demonstration, review Figure 5.

Given this fact, we next find the formula for a line perpendicular to the line segment going through  $\boldsymbol{\mu}_\gamma$ . First, we know from geometry that the line connecting  $g_1$  and  $g_0$  has slope  $\left(\frac{g_{y_0} - g_{y_1}}{g_{x_0} - g_{x_1}}\right)$ . We then know that the slope of the perpendicular line is the negative inverse of this slope, or  $-\left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)$ . The slope of the new line can also be written as  $\left(\frac{b_{y_\gamma} - g_{y_\gamma}}{b_{x_\gamma} - g_{x_\gamma}}\right)$  where  $\begin{pmatrix} b_{x_\gamma} \\ b_{y_\gamma} \end{pmatrix} = \mathbf{b}_\gamma$  is the boundary point we are searching for. This means we can write  $\left(\frac{b_{y_\gamma} - g_{y_\gamma}}{b_{x_\gamma} - g_{x_\gamma}}\right) = -\left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)$ , or  $b_{y_\gamma} = g_{y_\gamma} + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)(g_{x_\gamma} - b_{x_\gamma})$ .

Next, we know already that the formula for the circular error boundary at  $\boldsymbol{\mu}_\gamma$  is

$$\begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ b_{y_\gamma} - g_{y_\gamma} \end{pmatrix}' h_\gamma^{-1} \mathbf{I} \begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ b_{y_\gamma} - g_{y_\gamma} \end{pmatrix} = h_\gamma^{-1} \begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ b_{y_\gamma} - g_{y_\gamma} \end{pmatrix}' \begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ b_{y_\gamma} - g_{y_\gamma} \end{pmatrix}.$$

Because the boundary point is on the perpendicular line, we can substitute our information from above about  $b_{y_\gamma}$  and write

$$\begin{aligned} & h_\gamma^{-1} \begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} (g_{x_\gamma} - b_{x_\gamma}) \end{pmatrix}' \begin{pmatrix} b_{x_\gamma} - g_{x_\gamma} \\ \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} (g_{x_\gamma} - b_{x_\gamma}) \end{pmatrix} \\ &= \frac{(b_{x_\gamma} - g_{x_\gamma})^2}{h_\gamma} \left( 1 + \left( \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} \right)^2 \right) = \chi_{2,1-\alpha}^2. \end{aligned}$$

This means

$$\begin{aligned} (b_{x_\gamma} - g_{x_\gamma})^2 &= \frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left( 1 + \left( \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} \right)^2 \right)} \\ \Rightarrow b_{x_\gamma}^2 - 2b_{x_\gamma} g_{x_\gamma} + g_{x_\gamma}^2 &= \frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left( 1 + \left( \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} \right)^2 \right)} \\ \Rightarrow b_{x_\gamma}^2 - 2b_{x_\gamma} g_{x_\gamma} + \left( g_{x_\gamma}^2 - \frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left( 1 + \left( \frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}} \right)^2 \right)} \right) &= 0. \end{aligned}$$

This is a quadratic equation of the form  $ax^2 + bx + c$ . We can now solve for  $b_{x_\gamma}$  by finding the roots of the equation:

$$b_{x_\gamma} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{2g_{x_\gamma} \pm \sqrt{(-2g_{x_\gamma})^2 - 4 \left( g_{x_\gamma}^2 - \frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)} \right)}}{2}$$

$$= \frac{2g_{x_\gamma} \pm \sqrt{4 \left( g_{x_\gamma}^2 - g_{x_\gamma}^2 + \frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)} \right)}}{2}$$

gives us

$$b_{x_\gamma} = g_{x_\gamma} \pm \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}}.$$

Substituting this result for  $b_{x_\gamma}$  in the parallel line equation to find  $b_{y_\gamma}$ , a quick calculation tells us that the points of interception (and the points on the boundary of the confidence region of the line segment) are

$$\begin{pmatrix} b_{x_\gamma} \\ b_{y_\gamma} \end{pmatrix} = \begin{pmatrix} g_{x_\gamma} \pm \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \\ g_{y_\gamma} \mp \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right) \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \end{pmatrix}.$$

□

## Examples

We now provide some theoretical examples of the incorporation of Bayesian methodology into GIS vector data error analysis. For more examples using real data, look for an upcoming paper.

### *Example 1.*

The professor of a geography class assigns his five students to take GPS instruments and record the location of a “benchmark” at the university. (A *benchmark* is a location with coordinates that have been measured very precisely.) The students take the following coordinate measurements (in meters):

$i$	$x$	$y$
1	551470.3	4119766.4
2	551464.6	4119770.2
3	551472.8	4119763.1
4	551475.3	4119767.6
5	551468.7	4119769.9

The standard deviation of the error of the instruments used by the students is known to be 3 m. in the  $x$  direction, and 2 m. in the  $y$  direction, and the errors are uncorrelated. The

professor also happens to know that the benchmark's coordinates have been very carefully measured at (551469.1, 4119766.1) with a standard deviation of .1 m in either direction, and the errors are uncorrelated.

### *Frequentist Method*

Traditionally, only the current information on the coordinates would be used to estimate the true location of the benchmark. The estimate of the benchmark would be the average of the  $x$  and  $y$  observations,  $(\bar{x}, \bar{y}) = (551470.34, 4119767.44)$ . The variance of these estimates is  $\Sigma = \frac{1}{5} \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1.8 & 0 \\ 0 & 0.8 \end{bmatrix}$ . Finally, we can represent the error pictorially by drawing the error ellipsoid at the 95% confidence level,

$$(\bar{\mathbf{z}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{z}} - \boldsymbol{\mu}) = \begin{pmatrix} 551470.34 - \mu_{x_0} \\ 4119767.44 - \mu_{y_0} \end{pmatrix}' \begin{bmatrix} 1.8 & 0 \\ 0 & 0.8 \end{bmatrix}^{-1} \begin{pmatrix} 551470.34 - \mu_{x_0} \\ 4119767.44 - \mu_{y_0} \end{pmatrix} \leq \chi_{2,.95}^2,$$

shown in black in figure 6.

### *Bayesian Method*

The Bayesian methodology we have proposed here allows us to use the additional information we have about the benchmark - that it has been measured much more accurately

with a standard deviation of .1 m. We will use this as our prior distribution on  $\boldsymbol{\mu}_0$ :

$$\begin{pmatrix} \mu_{x_0} \\ \mu_{y_0} \end{pmatrix} \sim \left( \begin{pmatrix} 551469.1 \\ 4119766.1 \end{pmatrix}, \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix} \right).$$

Combining this prior distribution with the information from the data, we can give the posterior distribution of the point as  $\boldsymbol{\mu}_0 | \mathbf{Z}_0, \boldsymbol{\Sigma}_0 \sim N(\mathbf{g}_0, \mathbf{H}_0)$ , where

$$\begin{aligned} \mathbf{g}_0 &= \left( \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}^{-1} + 5 \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}^{-1} \right)^{-1} \left( \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}^{-1} \begin{pmatrix} 551469.1 \\ 4119766.1 \end{pmatrix} + 5 \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}^{-1} \begin{pmatrix} 551470.34 \\ 4119767.44 \end{pmatrix} \right) \\ &= \begin{pmatrix} 551469.1069 \\ 4119766.117 \end{pmatrix}, \end{aligned}$$

and

$$\mathbf{H}_0 = \begin{bmatrix} 0.00994 & 0 \\ 0 & 0.00988 \end{bmatrix}.$$

We can represent the 95% probability error ellipsoid in this situation with the equation

$$(\boldsymbol{\mu}_0 - \mathbf{g}_0)' \mathbf{H}_0^{-1} (\boldsymbol{\mu}_0 - \mathbf{g}_0) = \begin{pmatrix} \mu_{x_0} - 551469.1069 \\ \mu_{y_0} - 4119766.117 \end{pmatrix}' \begin{bmatrix} 0.00994 & 0 \\ 0 & 0.00988 \end{bmatrix}^{-1} \begin{pmatrix} \mu_{x_0} - 551469.1069 \\ \mu_{y_0} - 4119766.117 \end{pmatrix} \leq \chi_{2,.95}^2 = 5.99,$$

shown in red in figure 6.

### *Method Comparison*

When we compare the results from our two methods, we can see that the Bayesian result has markedly smaller variance, due to the inclusion of the much more accurate Bayesian prior. Evaluating the posterior estimate of the mean, it is obvious that this estimate more closely resembles what has already been carefully measured (the prior value on the benchmark) than the traditional estimate does. In situations where this type of information is available, it is clear how the Bayesian prior can improve the inference process. Figure 6 compares the error ellipsoids around the benchmark for the frequentist (black) and Bayesian (red) methods. Not only is the Bayesian ellipsoid much smaller, but it has an easier interpretation - rather than being a *confidence* region it is a *probability* region. Instead of being 95% “confident” that the coordinates are in the ellipsoid, we can say there is a 95% “probability” that the coordinates of the benchmark are in the Bayesian ellipsoid. It is also important to note that the Bayesian result is compatible and in agreement with the traditional result, since the probability ellipsoid here falls well within the boundary of the traditional confidence region.

*Example 2.*

Suppose you are interested in learning the coordinates of two adjacent corners of a particular building from a computerized GIS map. You know the map was digitized from an aerial photo; furthermore, based on the map scale, producer, and the program used to digitize the photo, you know the standard deviation of the coordinates is 10 meters in either direction. On the map, the coordinates have been placed at (376111, 2798632) and (376217, 2798650).

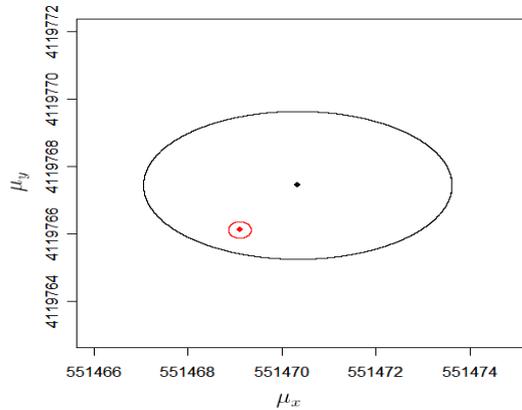


Figure 6: Comparison of traditional error ellipse (black) and Bayesian error ellipse (red).

Having been digitized by hand, there is no correlation between any coordinates.

You also have access to the original plans for the building, which instructed that the two corners of the building in question should be placed at  $(376100, 2798635)$  and  $(376220, 2798635)$ . The company has a recorded overall standard deviation in their accuracy of 5 meters. Assume the errors are uncorrelated.<sup>1</sup>

### *Frequentist Method*

From the frequentist standpoint, the estimates of the vertices of the side of this building are simply the parameters on the digitized map itself. This is true because we only have one observation. The standard deviation of the estimates is therefore the standard deviation of the digitizer's accuracy. We have

---

<sup>1</sup>This is probably not a realistic assumption since the construction company likely has limited materials and could not afford to accidentally extend the side of a building by say, 10 meters, but we will use this assumption here to simplify the calculations.

$$\mathbf{z}_{01} = \begin{pmatrix} 376111 \\ 2798632 \\ 376217 \\ 2798650 \end{pmatrix} \text{ and } \Sigma_{01} = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix}.$$

We can follow this up with a 95% confidence ellipse around each endpoint. The ellipse around the first endpoint is

$$(\boldsymbol{\mu}_0 - \mathbf{z}_0)' \Sigma_0^{-1} (\boldsymbol{\mu}_0 - \mathbf{z}_0) = \begin{pmatrix} \mu_{x_0} - 376111 \\ \mu_{y_0} - 2798632 \end{pmatrix}' \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \begin{pmatrix} \mu_{x_0} - 376111 \\ \mu_{y_0} - 2798632 \end{pmatrix} \leq \chi_{2,.95}^2 = 5.99,$$

and the ellipse around the second endpoint is

$$(\boldsymbol{\mu}_1 - \mathbf{z}_1)' \Sigma_1^{-1} (\boldsymbol{\mu}_1 - \mathbf{z}_1) = \begin{pmatrix} \mu_{x_1} - 376217 \\ \mu_{y_1} - 2798650 \end{pmatrix}' \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \begin{pmatrix} \mu_{x_1} - 376217 \\ \mu_{y_1} - 2798650 \end{pmatrix} \leq \chi_{2,.95}^2 = 5.99.$$

To create a confidence region around the entire line segment, we note that we can create a confidence region around any point on the line segment with mean estimate

$$\begin{pmatrix} (1-t)x_0 + (t)x_1 \\ (1-t)y_0 + (t)y_1 \end{pmatrix} = \begin{pmatrix} (1-t)376111 + (t)376217 \\ (1-t)2798632 + (t)2798650 \end{pmatrix}$$

and variance

$$\Sigma_{zz}(t) = \begin{bmatrix} \sigma_x^2(t) & \sigma_{xy}(t) \\ \sigma_{yx}(t) & \sigma_y^2(t) \end{bmatrix} = \begin{bmatrix} 100((1-t)^2 + (t)^2) & 0 \\ 0 & 100((1-t)^2 + (t)^2) \end{bmatrix}.$$

The ellipsoid-based picture of the confidence region around the line segment is shown in black in figure 7.

### *Bayesian Method*

The Bayesian method allows us to include information from both the original design plans and the current digitized map. According to our formulas, our Bayesian posterior estimate of the value of the coordinates for the first point is

$$\begin{aligned} \mathbf{g}_0 &= \left( \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}^{-1} + \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \right)^{-1} \left( \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}^{-1} \begin{pmatrix} 376100 \\ 2798635 \end{pmatrix} + \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \begin{pmatrix} 376111 \\ 2798632 \end{pmatrix} \right) \\ &= \begin{pmatrix} 376102.2 \\ 2798634.4 \end{pmatrix}, \text{ and the posterior variance of this estimate is } \mathbf{H}_0 = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}. \end{aligned}$$

Our Bayesian posterior estimate of the value of the coordinates for the second point is

$$\begin{aligned} \mathbf{g}_1 &= \left( \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}^{-1} + \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \right)^{-1} \left( \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}^{-1} \begin{pmatrix} 376220 \\ 2798635 \end{pmatrix} + \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}^{-1} \begin{pmatrix} 376217 \\ 2798650 \end{pmatrix} \right) \\ &= \begin{pmatrix} 376219.4 \\ 2798638.0 \end{pmatrix}, \text{ and the posterior variance of this estimate is } \mathbf{H}_1 = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}. \end{aligned}$$

We can now describe a 95% probability ellipse around each endpoint. The ellipse around the first endpoint is

$$(\boldsymbol{\mu}_0 - \mathbf{g}_0)' \mathbf{H}_0^{-1} (\boldsymbol{\mu}_0 - \mathbf{g}_0) = \begin{pmatrix} \mu_{x_0} - 376102.2 \\ \mu_{y_0} - 2798634.4 \end{pmatrix}' \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}^{-1} \begin{pmatrix} \mu_{x_0} - 376102.2 \\ \mu_{y_0} - 2798634.4 \end{pmatrix} \leq \chi_{2,.95}^2 = 5.99,$$

and the ellipse around the second endpoint is

$$(\boldsymbol{\mu}_1 - \mathbf{g}_1)' \mathbf{H}_1^{-1} (\boldsymbol{\mu}_1 - \mathbf{g}_1) = \begin{pmatrix} \mu_{x_1} - 376219.4 \\ \mu_{y_1} - 2798638.0 \end{pmatrix}' \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}^{-1} \begin{pmatrix} \mu_{x_1} - 376219.4 \\ \mu_{y_1} - 2798638.0 \end{pmatrix} \leq \chi_{2,.95}^2 = 5.99.$$

To create a probability region around the entire line segment, we note that we can create a probability region around any point on the line segment with mean estimate

$$\begin{pmatrix} (1 - \gamma)g_{x_0} + (\gamma)g_{x_1} \\ (1 - \gamma)g_{y_0} + (\gamma)g_{y_1} \end{pmatrix} = \begin{pmatrix} (1 - \gamma)376102.2 + (\gamma)376219.4 \\ (1 - \gamma)2798634.4 + (\gamma)2798638.0 \end{pmatrix}$$

and variance

$$\mathbf{H}_\gamma = \begin{bmatrix} h_{x_\gamma}^2 & h_{x_\gamma y_\gamma}(t) \\ h_{y_\gamma x_\gamma} & h_{y_\gamma}^2 \end{bmatrix} = \begin{bmatrix} 20((1 - \gamma)^2 + (\gamma)^2) & 0 \\ 0 & 20((1 - \gamma)^2 + (\gamma)^2) \end{bmatrix} = 20(2\gamma^2 - 2\gamma + 1)\mathbf{I}.$$

Alternatively, we can find the formula for the posterior mean and variance of a point on the line segment by consulting corollary 7, and the results agree.

$$\begin{aligned}
\mathbf{g}_\gamma &= \left( \frac{\tau^2 \sigma^2}{\tau^2 + n\sigma^2} \right) \left[ (1 - \gamma) \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{0_0} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_0 \right) + \gamma \left( \frac{1}{\tau^2} \boldsymbol{\mu}_{1_0} + \frac{n}{\sigma^2} \bar{\mathbf{z}}_1 \right) \right] \\
&= \left( \frac{25 \cdot 100}{25 + 100} \right) \left[ (1 - \gamma) \left( \frac{1}{25} \begin{pmatrix} 376100 \\ 2798635 \end{pmatrix} + \frac{1}{100} \begin{pmatrix} 376111 \\ 2798632 \end{pmatrix} \right) + \gamma \left( \frac{1}{25} \begin{pmatrix} 376220 \\ 2798635 \end{pmatrix} + \frac{1}{100} \begin{pmatrix} 376217 \\ 2798650 \end{pmatrix} \right) \right] \\
&= 20 \left[ (1 - \gamma) \begin{pmatrix} 18805.11 \\ 139931.72 \end{pmatrix} + \gamma \begin{pmatrix} 18810.97 \\ 139931.9 \end{pmatrix} \right] = \begin{pmatrix} (1 - \gamma)376102.2 + (\gamma)376219.4 \\ (1 - \gamma)2798634.4 + (\gamma)2798638.0 \end{pmatrix}, \\
\mathbf{H}_\gamma &= (2\gamma^2 - 2\gamma + 1) \frac{\tau^2 \sigma^2}{\sigma^2 + n\tau^2} \mathbf{I} = (2\gamma^2 - 2\gamma + 1) \frac{25 \cdot 100}{100 + 25} \mathbf{I} = 20(2\gamma^2 - 2\gamma + 1) \mathbf{I}.
\end{aligned}$$

The ellipsoid-based picture of the credible region around the line segment is shown in red in figure 7. For a demonstration of our explicit confidence region formula, please see the section on *Confidence Region Calculation*.

#### *Method Comparison*

As in our first example, we see here that the variance of our estimate is considerably smaller when using the prior information about the building in combination with the data currently available on the map. We can also see that our final posterior estimate of the mean is much closer to that specified by the original business plans (thanks to the inclusion

of this information), and therefore probably more accurate. We can also examine figure 7, shown here, which demonstrates both the difference in placement of the line segment (where the two endpoint vertices appear) and the variance of our estimate (the region surrounding the posterior distribution's line segment is smaller than the region surrounding the data distribution's line segment). Notice that like with the ellipses in example 1, the Bayesian region is completely compatible with the traditional region since it falls entirely within the traditional boundary. Furthermore, we again consider the difference in the meaning of the two regions - the traditional region is a confidence region, and we are 95% “confident” that the line segment is contained in that region, and the Bayesian region is a “credible” region, and there is a 95% probability that the true line segment is contained in that region. Once more, the benefits of Bayesian methods are clear.

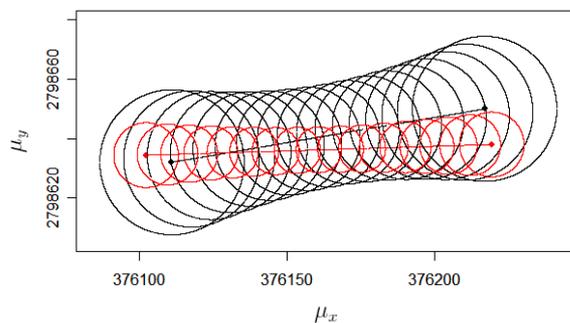


Figure 7: Comparison of traditional confidence region (black) and Bayesian credible region (red).

### *Confidence Region Calculation*

We have shown the confidence and probability regions around this line segment in terms of the collection of error ellipses along its points. We will here use the formula developed earlier in this paper to determine an explicit calculation of the probability region boundary in the Bayesian version of the above calculations.

Note that this case is indeed covered by theorem 3, since the variances involved are such that there is no correlation between the endpoints  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_1$ , and there is no correlation between the  $x$ - and  $y$ - coordinates at each endpoint; additionally, all coordinate variances are similar. That is,  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$ , and  $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_0 = \boldsymbol{\Lambda} = \begin{pmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{pmatrix}$ .

The formula for the boundary points at a point  $(\mu_{x_\gamma}, \mu_{y_\gamma})$  is

$$\begin{pmatrix} b_{x_\gamma} \\ b_{y_\gamma} \end{pmatrix} = \begin{pmatrix} g_{x_\gamma} \pm \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \\ g_{y_\gamma} \mp \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right) \sqrt{\frac{\chi_{2,1-\alpha}^2 h_\gamma}{\left(1 + \left(\frac{g_{x_0} - g_{x_1}}{g_{y_0} - g_{y_1}}\right)^2\right)}} \end{pmatrix}.$$

Here, from our previous results, we have

$$\begin{pmatrix} g_{x_0} \\ g_{y_0} \\ g_{x_1} \\ g_{y_1} \end{pmatrix} = \begin{pmatrix} 376102.2 \\ 2798634.4 \\ 376219.4 \\ 2798638.0 \end{pmatrix}, h_\gamma = 20(2\gamma^2 - 2\gamma + 1).$$

Assuming we want to create a 95% probability region around our line segment, we also know  $\chi_{2,.95}^2 = 5.99$  from a chi-square distribution table.

$$\begin{aligned} \text{First, } \mathbf{g}_\gamma &= (1 - \gamma)\mathbf{g}_0 + (\gamma)\mathbf{g}_1 = (1 - \gamma) \begin{pmatrix} 376102.2 \\ 2798634.4 \end{pmatrix} + (\gamma) \begin{pmatrix} 376219.4 \\ 2798638.0 \end{pmatrix} \\ &= \begin{pmatrix} (1 - \gamma)376102.2 + (\gamma)376219.4 \\ (1 - \gamma)2798634.4 + (\gamma)2798638.0 \end{pmatrix} \end{aligned}$$

We can now write the formula for the boundary points as

$$\begin{pmatrix} (1 - \gamma)376102.2 + (\gamma)376219.4 \pm \sqrt{\frac{5.99 \cdot 20(2\gamma^2 - 2\gamma + 1)}{\left(1 + \left(\frac{376102.2 - 376219.4}{2798634.4 - 2798638.0}\right)^2\right)}} \\ (1 - \gamma)2798634.4 + (\gamma)2798638.0 \mp \left(\frac{376102.2 - 376219.4}{2798634.4 - 2798638.0}\right) \sqrt{\frac{5.99 \cdot 20(2\gamma^2 - 2\gamma + 1)}{\left(1 + \left(\frac{376102.2 - 376219.4}{2798634.4 - 2798638.0}\right)^2\right)}} \end{pmatrix}.$$

Note that beyond the endpoints, the remainder of the boundary is calculated from the error ellipses around the endpoints. Figure 8 offers a graphic depiction of this boundary. The ellipse-based confidence region is included as well (in red) for comparison.

## Discussions and Conclusions

In this paper, we have explored the addition of Bayesian methodology to current methods for analyzing error in vector GIS data. There are multiple advantages to this addition. For

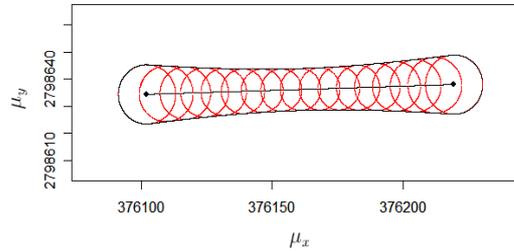


Figure 8: Boundary region of a line segment based on explicit calculation.

one thing, Bayesian analysis does not rely strongly on the present data to develop an error distribution or to estimate the location of a point. This is because the distribution of a point can be made to rely strongly on a prior distribution that can be based on expert or historical knowledge. This is not unreasonable in the geographic disciplines, where a lot of knowledge may already exist to indicate coordinate locations (for example, the possible path of a stream or the locations of certain well-studied landmarks). This is a big advantage because often only one sample observation is available for each point on a map. Additionally, Bayesian methods can increase the accuracy of an analysis when prior information is more reliable than the data distribution, which could often be the case in GIS applications. Bayesian methodology is also completely compatible with traditional methods, as seen in our examples.

Bayesian analysis is also good from the standpoint of understandability. The average non-statistician often finds the concept of a frequentist confidence region difficult to grasp, since it relies on the concept of confidence rather than probability; in fact, many non-statisticians

usually interpret a confidence interval as a probability interval. The interpretation of a Bayesian credible region, on the other hand, is direct and accessible to many users.

We are presently continuing to develop methods of Bayesian error analysis. Our current pursuits include finding an explicit formula for the general credible region surrounding a line segment, and using this information to do complicated probability calculations. This includes the probability that a particular point falls within the boundary of a particular polygon, and the probability that small extraneous “sliver” polygons created when two polygon maps are overlaid exist in reality.

We are also developing alternative methods of defining the confidence region around a line segment. The current model is certainly acceptable in situations where the line feature is straight between two endpoints, often correct when dealing with man-made objects, but unacceptable in situations where the line between two points is definitely not straight. This is much more likely in natural situations, such as a river or the boundary of a forest.

## References

- [1] BLAKEMORE, M. Generalisation and error in spatial data bases. *Cartographica* 21 (1984), 131–139.
- [2] BOLSTAD, P. V., AND SMITH, J. L. Errors in gis: Assessing spatial data accuracy. *Journal of Forestry* 90, 11 (1992), 21–29.
- [3] CHRISMAN, N. R. On storage of coordinates in geographic systems. *Geo-Processing* 2 (1984), 259–170.
- [4] CLARKE, K. C. *Getting Started with Geographic Information Systems*, fourth ed. Prentice Hall Pearson Education, Inc., 2003.
- [5] GELMAN, A., CARLIN, J. B., STERN, H. S., AND RUBIN, D. B. *Bayesian Data Analysis*, second ed. Chapman & Hall/CRC, 2004.
- [6] MAGUIRE, D. J. An overview and definition of gis. In *Geographical Information Systems*, D. J. Maguire, M. F. Goodchild, and D. W. Rhind, Eds. Longman Scientific and Technical, 1991, ch. 1, pp. 9–20.
- [7] MARBLE, D. F. Geographic information systems: an overview. In *Introductory Readings in Geographic Information Systems*, D. J. Peuquet and D. F. Marble, Eds. Taylor & Francis, 1990, ch. 1, pp. 8–17.

- [8] MOWRER, H. T. Accuracy (re)assurance: Selling uncertainty assessment to the uncertain. In *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*, K. Lowell and A. Jaton, Eds. Ann Arbor Press, 1999, ch. 1, pp. 3 – 10.
- [9] OPENSHAW, S. Learning to live with errors in spatial databases. In *Accuracy of Spatial Databases*, M. Goodchild and S. Gopal, Eds. Taylor and Francis, 1989, ch. 23, pp. 263–276.
- [10] SHI, W. A generic statistical approach for modelling error of geometric features in gis. *International Journal of GIS* 12, 2 (1998), 131–134.
- [11] SHI, W., CHEUNG, C.-K., AND TONG, X. Modelling error propagation in vector-based overlay analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* 59, 1-2 (2004), 47–59.
- [12] SHI, W., CHEUNG, C. K., AND ZHU, C. Modelling error propagation in vector-based buffer analysis. *International Journal of GIS* 17, 3 (2003), 251–271.
- [13] SHI, W., EHLERS, M., AND TEMPFLI, K. Analytical modelling of positional and thematic uncertainties in the integration of remote sensing and geographical information systems. *Transactions in GIS* 3, 2 (1999), 119–136.

- [14] SHI, W., AND LIU, W. A stochastic process-based model for the positional error of line segments in gis. *International Journal of GIS* 14, 1 (2000), 51–66.