Bayesian Wombling

Curvilinear Gradient Assessment Under Spatial Process Models

Sudipto Banerjee & Alan E Gelfand

Bayesian Wombling: Curvilinear Gradient Assessment Under Spatial Process Models

Sudipto BANERJEE and Alan E. GELFAND

Large-scale inference for random spatial surfaces over a region using spatial process models has been well studied. Under such models, local analysis of the surface (e.g., gradients at given points) has received recent attention. A more ambitious objective is to move from points to curves, to attempt to assign a meaningful gradient to a curve. For a point, if the gradient in a particular direction is large (positive or negative), then the surface is rapidly increasing or decreasing in that direction. For a curve, if the gradients in the direction orthogonal to the curve tend to be large, then the curve tracks a path through the region where the surface is rapidly changing. In the literature, learning about where the surface exhibits rapid change is called wombling, and a curve such as we have described is called a wombling boundary. Existing wombling methods have focused mostly on identifying points and then connecting these points using an ad hoc algorithm to create curvilinear wombling boundaries. Such methods are not easily incorporated into a statistical modeling setting. The contribution of this article is to formalize the notion of a curvilinear wombling boundary in a vector analytic framework using parametric curves and to develop a comprehensive statistical framework for curvilinear boundary analysis based on spatial process models for point-referenced data. For a given curve that may represent a natural feature (e.g., a mountain, a river, or a political boundary), we address the issue of testing or assessing whether it is a wombling boundary. Our approach is applicable to both spatial response surfaces and, often more appropriately, spatial residual surfaces. We illustrate our methodology with a simulation study, a weather dataset for the state of Colorado, and a species presence/absence dataset from Connecticut.

KEY WORDS: Arc-length measure; Bayesian modeling; Directional derivative; Flux; Gaussian process; Line integral; Parametric curve; Wombling.

1. INTRODUCTION

Spatially referenced datasets and their analysis arise in diverse areas of scientific interest, including geological and environmental sciences (Webster and Oliver 2001), ecological systems (Scheiner and Gurevich 2001), digital terrain cartography (Jones 1997), computer experiments (Santner, Williams, and Notz 2003), image analysis (Winkler 2003), and public health (Waller and Gotway 2004). Very often, such data will be referenced over a fixed set of locations in a region of study. These locations can be with regions or areas with well-defined neighbors (e.g., pixels in a lattice, counties in a map), in which case they are called *areally-referenced* or *lattice* data. Alternatively, they may be simply points with coordinates (e.g., latitude–longitude, easting–northing), in which case they are called *point-referenced* or *geostatistical* data. Statistical theory and methods for modeling and analyzing such data depend on these configurations. The last decade has seen enormous developments in such modeling (see, e.g., Cressie 1993; Chilès and Delfiner 1999; Møller 2003; Schabenberger and Gotway 2004; Banerjee, Carlin, and Gelfand 2004 for various methods and applications).

Customary inferential interest resides in estimation of model parameters, in spatial prediction, and, more generally, in assessing the estimated spatial surface over the spatial domain. Statistical estimation and interpolation of the spatial surface often help detect "features" such as cliffs, edges, and boundaries in the topography. In the fields of image analysis and pattern recognition, there has been much research on using statistical models for capturing "edge" and "line" effects. Such models are based on probability distributions (such as Gibbs distributions or Markov random fields) that model pixel intensities as conditional dependencies using the neighborhood structure (see, e.g., Chellappa and Jain 1993). Explicit modeling of the edges or pixel boundaries using probability distributions has been discussed by Geman and Geman (1984) and, more recently, by Dass and Nair (2003) with multivariate observations and missing data (see also Winkler 2003). Modeling objectives include identification of edges based on distinctly different image intensities in adjacent pixels. Another line of approach, often called image segmentation, uses a constructive approach to represent image "interfaces" that partition images into disjoint sets of pixels. This approach builds on work of Mumford and Shah (1989) using optimization of PDE-based objective functionals, unsupervised learning methods (e.g., Caillol, Wojciech, and Hillion 1997), or transforms such as the Hough transform (see, e.g., Walsh and Raftery 2002) to find interface curves. These algorithms help segment images by constructing contour lines and identifying image discontinuities using variational techniques. Morel and Solimini (1994) have provided a comprehensive review of these algorithms.

Much of the foregoing development uses the pixel-based spatial domain, where the edges are well-defined lines that separate adjacent pixels. In fact, the Markov random field models are constructed on these adjacency structures. However, in such disciplines as geology, environmental sciences, and ecology, data arise from irregular spatial locations that do not have natural neighborhoods but are referenced by their coordinates. Such data are widely modeled using spatial processes that model association as a function of intersite distance. Spatial process models assume, for a region of study D, a collection of random variables $Y(\mathbf{s})$, where $\mathbf{s}$ indexes the points in D. The set $\{Y(\mathbf{s}) : \mathbf{s} \in D\}$ can be viewed as a randomly realized surface over D, which in practice is observed only at a finite set of locations in $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\}$. Once such an interpolated surface has been obtained, investigation of rapid change on the surface may be of interest. For instance, environmental scientists

are interested in ascertaining whether natural boundaries (e.g., mountains, forest edges) represent a zone of rapid change in weather, ecologists are interested in determining curves that delineate differing zones of species abundance, and public health officials want to identify changes in health care delivery across municipal boundaries, counties, or states.

The foregoing objectives require the notion of gradients and in particular assigning gradients to curves (*curvilinear gradients*) to identify curves that track a path through the region where the surface is rapidly changing. Such boundaries are commonly referred to as difference boundaries or *wombling boundaries*, after Womble (1951), who discussed their importance in understanding scientific phenomena (also see Fagan, Fortin, and Soykan 2003). As a concept, wombling is useful because it attempts to quantify spatial information in such objects as curves and paths that are not easy to model as regressors. It is similar to image analysis in that it also seeks to capture lurking "spatial effects" on curves. However, unlike images, where edges and lines represent discontinuities or breaks, wombling boundaries capture rapid surface change; cutting across a wombling boundary should tend to reveal a drop in elevation or, equivalently, a sharp gradient. Evidently, gradients are central to wombling, and the concerned spatial surfaces must be sufficiently smooth. This precludes such methods as wavelets that have been used to detect image discontinuities, such as ridges and cliffs (e.g., Csillag and Kabos 2002) but do not admit gradients.

Visual assessment of the surface over *D* often proceeds from contour and image plots of the surface fitted from the data using surface interpolators. Surface representation and contouring methods range from tensor product interpolators for gridded data (e.g., Cohen, Riesenfeld, and Elber 2001) to more elaborate adaptive control lattice or tessellation-based interpolators for scattered data (e.g., Akima 1996; Lee, Wolberg, and Shin 1997). Mitas and Mitasova (1999) have provided a review of several such methods available in Global Informational Systems (GIS) software (e.g., GRASS, available at *http://grass.itc.it/*). These methods are often fast and simple to implement and produce contour maps that reveal topographic features; however, they do not account for association and uncertainty in the data. Far from being competitive with statistical methods, they play a complementary role, creating descriptive plots from the raw data in the premodeling stage and providing visual displays of estimated response or residual surfaces in the postmodeling stage. It is worth pointing out that although contours often provide an idea of the local topography, they are not the same as wombling boundaries. Contour lines connect points with the same spatial elevation and may or may not track large gradients, so they may or may not correspond to wombling boundaries.

Existing wombling methods for point-referenced data concentrate on finding points with large gradients and attempting to connect them in an algorithmic fashion, which then defines a boundary. Such algorithms have been widely used in computational ecology, anthropology, and geography. For example, Barbujani, Jacquez, and Ligi (1990) and Barbujani, Oden, and Sokal (1989) used wombling on red blood cell markers to identify genetic boundaries in Eurasian human populations by different processes restricting gene flow; Bocquet-Appel and

Bacro (1994) investigated genetic, morphometric, and physiologic boundaries; Fortin (1994, 1997) delineated boundaries related to specific vegetation zones; Fortin and Drapeau (1995) applied wombling on real environmental data; and Jacquez and Greiling (2003) estimated boundaries for breast, lung, and colorectal cancer rates in males and females in Nassau, Suffolk, and Queens Counties in New York. This last application is somewhat different from the others in that the data were areally referenced with counties. Unlike image pixels, these counties are not regularly spaced but still have a well-defined neighborhood structure (a topological graph), allowing direct application of the image analysis methods. The gradient is not explicitly modeled; boundary effects are considered edge effects and modeled using Markov random field specifications. A Bayesian framework for areal boundary analysis has been provided by Lu and Carlin (2006).

Our interest here lies in point-referenced curvilinear gradient or boundary analysis, a conceptually more difficult problem due to the lack of definitive candidate boundaries. Spatial process models help estimate not only response surfaces, but also residual surfaces after covariate and systematic trends have been accounted. Depending on the scientific application, boundary analysis may be desirable on either. Current methods treat statistical estimates of the surface as "data" and apply interpolation-based wombling to obtain boundaries. Although such methods produce useful descriptive surface plots, they preclude formal statistical inference. Indeed, boundary assessment using such reconstructed surfaces will suffer from inaccurate estimation of uncertainty. One might also contemplate predicting the surface on an overlaid pixel-based domain and applying edge detection methods for image analysis. Again, the uncertainty in spatial prediction will be difficult to account for, the optimal resolutions at which to predict become an issue, and implementation for spatial residual surfaces will be problematic.

Recent developments in GIS and related software have made it possible to extract points along natural boundaries and thus to achieve polygonal representations of curves. We then seek to test whether or not such a given curve is a wombling boundary. Therefore, from an implementation standpoint, we require a well-defined curve on which to construct our tests. Typically, these curves are supplied by the scientists; they may arise as features of interest, such as mountainous paths in weather analysis or forest edges in species abundance studies. In the absence of scientific candidates, some preliminary topographic analyses using simple surface interpolation may also suggest such curves. Once these curves are chosen, curvilinear gradient analysis can be performed on statistically estimated spatial surfaces to ascertain whether these features contain significant spatial information (in terms of local surface topography). Note that existing wombling algorithms construct curves algorithmically and do not accommodate evaluation of gradients along curves obtained externally from GIS or other software. Thus they do not use gradient information to address either question. We offer a statistical framework that incorporates inference on curvilinear boundaries within the diverse class of hierarchical spatial models for point-referenced data.

In the context of point-referenced data, we consider directional derivatives of spatial processes and subsequently use

formal vector analysis to develop a wombling measure that evaluates the average gradient along a curve. This is done in a continuous sense in the direction normal to the curve, and leads to a formal definition of whether or not the curve is a wombling boundary. Line integrals emerge, and, using Gaussian spatial processes, we provide distributional details for such integrals. We also show how to implement full inference in a computationally feasible manner. This is offered in a Bayesian framework (e.g., Gelman, Carlin, Stern, and Rubin 2004) for the following reasons. First, it allows exact inference, avoiding possibly inappropriate asymptotics associated with likelihood methods (Stein 1999). Second, it supplies an entire posterior distribution for arbitrary functions of these gradients, such as those introduced in Section 3. A particularly attractive feature of our approach is that all desired gradient analysis can be done after the model for the surface is fitted. Assuming that Markov chain Monte Carlo (MCMC) methods (Gelman et al. 2004) are used to fit the model, we need to retain only the posterior samples of the model parameters. This greatly enhances the applicability of our proposed methods.

The article is organized as follows. Section 2 provides a brief overview of directional derivative processes. Section 3 formalizes curvilinear gradients for spatial processes with a vector analytic treatment. Section 4 discusses the relevant distribution theory for implementing relevant statistical inference for curvilinear gradients. Section 5 provides a simulated illustration that helps validate our methods with topographic maps and provides a comparison with the purely algorithmic methods. Section 6 follows by illustrations in spatial regression contexts, first with a weather dataset supplied by the National Center for Atmospheric Research (NCAR) in Boulder, Colorado and then with a binary regression model for species presence/absence from Connecticut. Section 7 concludes the article with a summary and discussion.

2. SPATIAL GRADIENTS AND DIRECTIONAL DERIVATIVE PROCESSES

As mentioned earlier, the concept of rapid change in the spatial surface is central to wombling and is mathematically formalized using spatial gradients. Adler (1981) and Mardia, Kent, Goodall, and Little (1996) discussed derivatives (more generally, linear functionals) of Gaussian processes, and Banerjee, Gelfand, and Sirmans (2003) laid out an inferential framework for directional gradients on a spatial surface. The latter methodology can be used for statistical estimation (in the form of a posterior distribution) of nonlinear functionals, such as the value of the maximum slope at any point and the direction of this maximum slope. This can be done over the entire spatial domain, irrespective of the configuration of the sampling locations. We seek to extend framework this to gradient processes along a curve, but first we provide a brief review of the statistical theory and inferential framework for gradients and directional derivative processes that arise from spatial processes.

Let $Y(\mathbf{s})$ be a univariate stationary random field. (Stationarity is not required; we use it only to ensure smoothness of realizations and to simplify forms for the induced covariance function.) The process $\{Y(\mathbf{s}) : \mathbf{s} \in \Re^d\}$ is L_2 (or mean squared) continuous at $\mathbf{s}_0$ if $\lim_{\mathbf{s} \to \mathbf{s}_0} E(|Y(\mathbf{s}) - Y(\mathbf{s}_0)|)^2 = 0$. The notion

of a mean squared differentiable process can be formalized using the analogous definition of total differentiability of a function in $\Re^d$ in a nonstochastic setting (see, e.g., Banerjee et al. 2003). In particular, $Y(\mathbf{s})$ is mean-squared differentiable at $\mathbf{s}_0$ if it admits a first-order linear expansion for any scalar h and any unit vector $\mathbf{u} \in \Re^d$,

$$Y(\mathbf{s}_0 + h\mathbf{u}) = Y(\mathbf{s}_0) + h\langle \nabla Y(\mathbf{s}_0), \mathbf{u} \rangle + o(h), \qquad (1)$$

in the L_2 sense as $h \to 0$, where $\nabla Y(\mathbf{s}_0)$ is a $d \times 1$ vector called the gradient vector and $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product on $\Re^d$. In other words, for any unit vector $\mathbf{u}$, we require

$$\lim_{h \to 0} E\left(\frac{Y(\mathbf{s}_0 + h\mathbf{u}) - Y(\mathbf{s}_0)}{h} - \langle \nabla Y(\mathbf{s}_0), \mathbf{u} \rangle \right)^2 = 0. \qquad (1')$$

A directional gradient process on $\Re^d$ is defined as the following L_2 limit:

$$D_{\mathbf{u}} Y(\mathbf{s}) = \lim_{h \to 0} \frac{Y(\mathbf{s} + h\mathbf{u}) - Y(\mathbf{s})}{h}.$$

It immediately follows from $(1')$ that $D_{\mathbf{u}} Y(\mathbf{s}) = \langle \nabla Y(\mathbf{s}), \mathbf{u} \rangle$, where the equality is again in the L_2 sense. Note that there is no loss of generality in choosing $\mathbf{u}$ as a unit vector. Any vector $\mathbf{v}$ can be written as $\|\mathbf{v}\|\mathbf{u}$, where $\mathbf{u}$ is a unit vector collinear to $\mathbf{v}$, so the foregoing linearity implies that $D_{\mathbf{v}} Y(\mathbf{s}) = \|\mathbf{v}\| D_{\mathbf{u}} Y(\mathbf{s}) = \langle \nabla Y(\mathbf{s}), \mathbf{v} \rangle$. The gradient, $\nabla Y(\mathbf{s})$, is invariant over the choice of the direction vector, and we could consistently define

$$D_{\mathbf{v}} Y(\mathbf{s}) = \lim_{h \to 0} \frac{Y(\mathbf{s} + h\mathbf{v}) - Y(\mathbf{s})}{h}$$

for any vector $\mathbf{v}$. In fact, the physical interpretation of this directional gradient depends on the choice of $\mathbf{v}$. When $\mathbf{v}$ is a unit vector, the directional derivative evaluated at $\mathbf{s}_0$ is the slope at $\mathbf{s}_0$ of the curve traced out by slicing the surface Y in the direction $\mathbf{v}$. This geometric interpretation is desirable for spatial gradients; in contrast, derivatives arising in physical experiments may be better interpreted with nonunit vectors. For example, in heat dynamic experiments, if $Y(\mathbf{s})$ is taken as a spatial process that models temperature, then $\mathbf{v}$ is the velocity vector of the molecules in the domain at that point.

More generally, collecting a set of p directions in $\Re^d$ into the $d \times p$ matrix, $U = [\mathbf{u}_1, \ldots, \mathbf{u}_p]$, we can write $D_U Y(\mathbf{s})$ as the $p \times 1$ vector, $D_U Y(\mathbf{s}) = (D_{\mathbf{u}_1} Y(\mathbf{s}), \ldots, D_{\mathbf{u}_p} Y(\mathbf{s}))^T$, so that $D_U Y(\mathbf{s}) = U^T \nabla Y(\mathbf{s})$. In particular, setting $p = d$ and taking U as the $d \times d$ identity matrix (i.e., taking the canonical basis, $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$, as our directions), we have $D_I Y(\mathbf{s}) = \nabla Y(\mathbf{s})$, which gives a representation of $\nabla Y(\mathbf{s})$ in terms of the partial derivatives of the components of $Y(\mathbf{s})$. Explicitly,

$$\nabla Y(\mathbf{s}) = \left(\frac{\partial Y(\mathbf{s})}{\partial s_1}, \ldots, \frac{\partial Y(\mathbf{s})}{\partial s_d} \right)^T,$$

where $\mathbf{s} = \sum_{i=1}^{d} s_i \mathbf{e}_i$, so the s_i's are the coordinates of $\mathbf{s}$ with respect to the canonical basis and $D_U Y(\mathbf{s}) = U^T D_I Y(\mathbf{s})$. Thus the derivative process in a set of arbitrary directions is a linear transformation of the partial derivatives in the canonical directions, and all inference about the directional derivatives can be built from this relationship.

An inferential framework for directional gradients is developed by letting $Y(\mathbf{s})$ be a univariate stationary random field, $GP(\mu(\mathbf{s}, \boldsymbol{\beta}), K(\cdot; \phi))$, on $\Re^d$, where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi)$ is a set

of possibly unknown model parameters to be estimated from the data. We will often suppress the parameters for notational convenience, writing $\mu(\mathbf{s})$ and $K(\cdot)$ instead. Suppose that we have observed the random field at locations, say $\mathbf{Y} = (Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_n))^T$, and let $\mathbf{s}_0$ be a point at which we seek to investigate directional gradients. Based on our earlier discussion, it suffices to consider the predictive distribution $P(\nabla Y(\mathbf{s}_0)|Y)$, where $\nabla Y(\mathbf{s}_0)$ is the $d \times 1$ gradient vector. The joint process $(Y(\mathbf{s}_0), \nabla Y(\mathbf{s}_0))$ has a valid stationary cross-covariance matrix function (see, e.g., Banerjee et al. 2003),

$$\begin{pmatrix} \text{cov}(Y(\mathbf{s}_0), Y(\mathbf{s}_0 + \boldsymbol{\delta})) & \text{cov}(Y(\mathbf{s}_0), \nabla Y(\mathbf{s}_0 + \boldsymbol{\delta})) \\ \text{cov}(\nabla Y(\mathbf{s}_0), Y(\mathbf{s}_0 + \boldsymbol{\delta})) & \text{cov}(\nabla Y(\mathbf{s}_0), \nabla Y(\mathbf{s}_0 + \boldsymbol{\delta})) \end{pmatrix}$$
$$= \begin{pmatrix} K(\boldsymbol{\delta}) & -(\nabla K(\boldsymbol{\delta}))^T \\ \nabla K(\boldsymbol{\delta}) & -H_K(\boldsymbol{\delta}) \end{pmatrix},$$

where $H_K(\boldsymbol{\delta}) = ((\partial^2 K(\boldsymbol{\delta})/\partial \delta_i \, \partial \delta_j))$ is the $d \times d$ Hessian matrix of $K(\boldsymbol{\delta})$. The cross-covariance matrix induces a valid joint distribution $P(\mathbf{Y}, \nabla Y(\mathbf{s}_0)|\boldsymbol{\theta})$, allowing predictive inference not only for the gradient at arbitrary points, but also for functions thereof, including the direction of the maximal gradient $(\nabla Y(\mathbf{s}_0)/\|\nabla Y(\mathbf{s}_0)\|)$ and the size of the maximal gradient $(\|\nabla Y(\mathbf{s}_0)\|)$.

Simplifications arise when the mean surface, $\mu(\mathbf{s})$, admits a gradient $\nabla\mu(\mathbf{s})$. Let $\boldsymbol{\mu} = (\mu(\mathbf{s}_1), \ldots, \mu(\mathbf{s}_n))$, let Σ_Y denote the $n \times n$ dispersion matrix for the data $\mathbf{Y}$, and let $\boldsymbol{\gamma}^T = (\nabla K(\boldsymbol{\delta}_{01}), \ldots, \nabla K(\boldsymbol{\delta}_{0n}))$ be the $d \times n$ matrix with $\boldsymbol{\delta}_{0j} = \mathbf{s}_0 - \mathbf{s}_j$. Then $P(\mathbf{Y}, \nabla Y(\mathbf{s}_0)|\boldsymbol{\theta})$ is distributed as the $(d+n)$-dimensional normal distribution

$$\mathrm{N}_{d+n}\left(\begin{pmatrix} \boldsymbol{\mu} \\ \nabla\mu(\mathbf{s}_0) \end{pmatrix}, \begin{pmatrix} \Sigma_Y & \boldsymbol{\gamma} \\ \boldsymbol{\gamma}^T & -H_K(\mathbf{0}) \end{pmatrix} \right), \qquad (2)$$

and the conditional predictive distribution for the gradient, $P(\nabla Y(\mathbf{s}_0)|\mathbf{Y}, \boldsymbol{\theta})$, is the d-dimensional normal distribution

$$\mathrm{N}_d\big(\nabla\mu(\mathbf{s}_0) - \boldsymbol{\gamma}^T \Sigma_Y^{-1}(\mathbf{Y} - \boldsymbol{\mu}), -H_K(\mathbf{0}) - \boldsymbol{\gamma}^T \Sigma_Y^{-1} \boldsymbol{\gamma}\big). \quad (3)$$

The posterior predictive distribution $P(\nabla Y(\mathbf{s}_0)|\mathbf{Y})$ is obtained as $\int P(\nabla Y(\mathbf{s}_0)|\mathbf{Y}, \boldsymbol{\theta}) P(\boldsymbol{\theta}|\mathbf{Y}) \, d\boldsymbol{\theta}$. Simulation from this distribution is by composition; for each $\boldsymbol{\theta}_l$ obtained from $P(\boldsymbol{\theta}|\mathbf{Y})$, we draw $\nabla Y^{(l)}(\mathbf{s}_0)$ from (3). As long as $\nabla\mu(\mathbf{s}_0)$ is computable, obtaining samples from the foregoing distribution is routine. In practice, we could have $\mu(\mathbf{s}, \boldsymbol{\beta}) = \mu$, a constant, in which case $\nabla\mu(\mathbf{s}) = \mathbf{0}$. More generally, we would have $\mu(\mathbf{s}, \boldsymbol{\beta}) = \mathbf{f}(\mathbf{s})^T \boldsymbol{\beta}$, where $\mathbf{f}(\mathbf{s})$ may be a vector of spatial covariates. If $\mathbf{f}(\mathbf{s})^T \boldsymbol{\beta}$ describes a trend surface, then explicit calculation of $\nabla\mu(\mathbf{s})$ will be possible. For a continuous covariate such as elevation, we can interpolate a surface and approximate $\nabla\mu(\mathbf{s})$ at any location $\mathbf{s}$. Note that it might suffice to consider gradients of a residual (or intercept) spatial process, say $W(\mathbf{s})$, where $Y(\mathbf{s}) = \mu(\mathbf{s}; \boldsymbol{\beta}) + W(\mathbf{s}) + \epsilon(\mathbf{s})$, where $W(\mathbf{s}) \sim \mathrm{GP}(0, K(\cdot; \phi))$ is a mean-0 Gaussian process and $\epsilon(\mathbf{s}) \sim \mathrm{N}(0, \tau^2)$ is a white noise process, often called a nugget. Inference will then proceed from the posterior distribution of $P(\nabla W(\mathbf{s}_0)|\mathbf{Y})$.

Based on the foregoing distribution theory, formal statistical inference on gradients can be performed. For instance, given any direction $\mathbf{u}$ and any location $\mathbf{s}_0$, a statistically "significant" directional gradient would mean that a 95% posterior credible interval for $D_{\mathbf{u}}W(\mathbf{s}_0)$ would not include 0. Because $D_{-\mathbf{u}}W(\mathbf{s}_0) = -D_{\mathbf{u}}W(\mathbf{s}_0)$, inference for $\mathbf{u}$ is the same as that for $-\mathbf{u}$. In addition, assessing significance of the spatial residual

process $W(\mathbf{s})$ is more general than for the parent process $Y(\mathbf{s})$. Indeed, when $\nabla\mu(\mathbf{s}_0)$ exists and there is no nugget ($\tau^2 = 0$), then the former is equivalent to testing the significance of $D_{\mathbf{u}}Y(\mathbf{s}_0)$ as a departure from the trend surface gradient $D_{\mathbf{u}}\mu(\mathbf{s}_0)$ (the *null value*). But even when $\nabla\mu(\mathbf{s}_0)$ is inaccessible, or there is a nugget $\tau^2 > 0$, assessment of spatial gradients for $W(\mathbf{s}_0)$ is still legitimate.

3. WOMBLING MEASURE ASSOCIATED WITH A CURVE

We now extend the theory of the previous section to an inferential framework for curvilinear boundaries. The conceptual challenge in moving from points to curves is the formulation of a measure to associate with a curve to assess whether it can be declared a wombling boundary. In this regard, we can consider open or closed curves. In Section 3.1 we formally develop the notion of an average gradient to associate with an open curve. In Section 3.2 we turn to closed curves, the notion of flux, and special results for this case. In applications such curves might be proposed as, for instance, topographic or legislated boundaries or perhaps as level curves arising from a contouring routine. Finally, in Section 3.3 we offer a formal definition of a wombling boundary.

We use differential geometric notions for parametric boundaries as developed by, for example, Rudin (1976) or Frankel (2003). Because most spatial modeling is done on domains in $\Re^2$, we restrict our attention to that case, focusing on a real-valued process $Y(\mathbf{s})$ with the spatial domain as an open subset of $\Re^2$. Thus we offer an independent development of gradients along planar curves without resorting to geometry on manifolds. For hypercurves in general $\Re^d$, the theory is more complicated (especially if $d > 3$) and must involve the development of calculus on abstract manifolds.

3.1 Open Curves

Let C be an open curve in $\Re^2$, and suppose that we want to ascertain whether such a curve is a wombling boundary with regard to $Y(\mathbf{s})$. To do so, we seek to associate an average gradient with C. In particular, for each point $\mathbf{s}$ lying on C, we let $D_{n(\mathbf{s})}Y(\mathbf{s})$ be the directional derivative in the direction of the unit normal $n(\mathbf{s})$. (Again, the rationale for the choice of direction normal to the curve is that for a curve tracking rapid change in the spatial surface, lines orthogonal to the curve should reveal sharp gradients.) We can define the *wombling measure* of the curve either as the total gradient along C,

$$\int_C D_{n(\mathbf{s})}Y(\mathbf{s}) \, dv = \int_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle \, dv, \qquad (4)$$

or perhaps as the average gradient along C,

$$\frac{1}{v(C)} \int_C D_{n(\mathbf{s})}Y(\mathbf{s}) \, dv = \frac{1}{v(C)} \int_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle \, dv, \qquad (4')$$

where $v(\cdot)$ is an appropriate measure. For (4) and (4'), ambiguity arises with respect to the choice of measure. For example, $v(C) = 0$ if we take v as two-dimensional Lebesgue measure, and indeed this is true for any v that is mutually absolutely continuous with respect to Lebesgue measure. On reflection, an appropriate choice for v turns out to be arc-length. This can be made clear by a parametric treatment of the curve C.

In particular, a curve C in $\Re^2$ is a set parameterized by a single parameter $t \in \Re^1$, where $C = \{\mathbf{s}(t) : t \in \mathcal{T}\}$, with $\mathcal{T} \subset \Re^1$. We call $\mathbf{s}(t) = (s_{(1)}(t), s_{(2)}(t)) \in \Re^2$ the position vector of the curve; $\mathbf{s}(t)$ traces out C as t spans its domain. Then, assuming a differentiable curve with nonvanishing derivative $\mathbf{s}'(t) \neq 0$ (such a curve is often called *regular*), we obtain the (componentwise) derivative $\mathbf{s}'(t)$ as the "velocity" vector, with unit velocity (or tangent) vector $\mathbf{s}'(t)/\|\mathbf{s}'(t)\|$. Letting $n(\mathbf{s}(t))$ denote the parameterized unit normal vector to C, again if C is sufficiently smooth, then $\langle \mathbf{s}'(t), n(\mathbf{s}(t)) \rangle = 0$, a.e. $\mathcal{T}$. In $\Re^2$, we see that

$$n(\mathbf{s}(t)) = \frac{(s'_{(2)}(t), -s'_{(1)}(t))}{\|\mathbf{s}'(t)\|}. \qquad (5)$$

Under the foregoing parameterization (and the regularity assumption), the arc-length measure ν can be defined as

$$\nu(\mathcal{T}) = \int_{\mathcal{T}} \|\mathbf{s}'(t)\| \, dt. \qquad (6)$$

In fact, $\|\mathbf{s}'(t)\|$ is analogous to the "speed" (the norm of the velocity) at "time" t, so the foregoing integral is interpretable as the distance traversed or, equivalently, the arc-length $\nu(C)$ or $\nu(\mathcal{T})$. In particular, if $\mathcal{T}$ is an interval, say $[t_0, t_1]$, then we can write

$$\nu(\mathcal{T}) = \nu_{t_0}(t_1) = \int_{t_0}^{t_1} \|\mathbf{s}'(t)\| \, dt.$$

Thus we have $d\nu_{t_0}(t) = \|\mathbf{s}'(t)\| \, dt$, and, taking ν as the arc-length measure for C, we have the wombling measures in (4) (total gradient) and (4′) (average gradient) as

$$\Gamma_{Y(\mathbf{s})}(\mathcal{T}) = \int_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle \, d\nu$$

$$= \int_{\mathcal{T}} \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle \|\mathbf{s}'(t)\| \, dt \quad \text{and} \quad (7)$$

$$\bar{\Gamma}_{Y(\mathbf{s})}(\mathcal{T}) = \frac{1}{\nu(\mathcal{T})} \Gamma_{Y(\mathbf{s})}(\mathcal{T}).$$

This result is important because we want to take ν as the arc-length measure, but it will be easier to use the parametric representation and work in t space. Moreover, it is a consequence of the implicit mapping theorem in mathematical analysis (see, e.g., Rudin 1976) that any other parameterization $\mathbf{s}^*(t)$ of the curve C is related to $\mathbf{s}(t)$ through a differentiable mapping g such that $\mathbf{s}^*(t) = \mathbf{s}(g(t))$. This immediately implies [using (7)] that our proposed wombling measure is invariant to the parameterization of C and, as desired, is a feature of the curve itself.

For some simple curves, the wombling measure can be evaluated quite easily. For instance, when C is a segment of length 1 of the straight line through the point $\mathbf{s}_0$ in the direction $\mathbf{u} = (u_{(1)}, u_{(2)})$, we have $C = \{\mathbf{s}_0 + t\mathbf{u} : t \in [0, 1]\}$. Under this parameterization, $\mathbf{s}'(t)^T = (u_{(1)}, u_{(2)})$, $\|\mathbf{s}'(t)\| = 1$, and $\nu_{t_0}(t) = t$. Clearly, $n(\mathbf{s}(t)) = (u_{(2)}, -u_{(1)})$ (independent of t), which we write as $\mathbf{u}^\perp$, the normal direction to $\mathbf{u}$. Therefore, $\Gamma_{Y(\mathbf{s})}(\mathcal{T})$ in (7) becomes

$$\int_0^1 \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle \, dt = \int_0^1 D_{\mathbf{u}^\perp} Y(\mathbf{s}(t)) \, dt.$$

Another example is when C is the arc of a circle with radius r. For example, suppose that C is traced out by $\mathbf{s}(t) = $ ($r \cos t, r \sin t$) as $t \in [0, \pi/4]$. Then, because $\|\mathbf{s}'(t)\| = r$, the average gradient is more easily computed as

$$\frac{1}{\nu(C)} \int_0^{\pi/4} \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle r \, dt$$

$$= \frac{4}{r\pi} \int_0^{\pi/4} \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle r \, dt$$

$$= \frac{4}{\pi} \int_0^{\pi/4} \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle \, dt.$$

In either case, $n(\mathbf{s}(t))$ is given by (5).

Note that whereas the normal component, $D_{n(\mathbf{s})} Y(\mathbf{s})$, seems to be more appropriate for assessing whether a curve provides a wombling boundary, we may also consider the tangential direction, $\mathbf{u}(t) = \mathbf{s}'(t)/\|\mathbf{s}'(t)\|$, along a curve C. In this case the average gradient will be given by

$$\frac{1}{\nu(C)} \int_C \langle \nabla Y(\mathbf{s}(t)), \mathbf{u}(t) \rangle \|\mathbf{s}'(t)\| \, dt$$

$$= \frac{1}{\nu(C)} \int_C \langle \nabla Y(\mathbf{s}(t)), \mathbf{s}'(t) \rangle \, dt.$$

In fact, we have

$$\int_C \langle \nabla Y(\mathbf{s}(t)), \mathbf{s}'(t) \rangle \, dt = \int_{t_0}^{t_1} \langle \nabla Y(\mathbf{s}(t)), \mathbf{s}'(t) \rangle \, dt$$

$$= \int_{\mathbf{s}_0}^{\mathbf{s}_1} \langle \nabla Y(\mathbf{s}), d\mathbf{s} \rangle = Y(\mathbf{s}_1) - Y(\mathbf{s}_0),$$

where $\mathbf{s}_1 = \mathbf{s}(t_1)$ and $\mathbf{s}_0 = \mathbf{s}(t_0)$ are the endpoints of C. That is, the average directional gradient in the tangential direction is independent of the path C, depending only on the endpoints of the curve C.

3.2 Closed Curves

A related concept is that of the "flux" of a region bounded by C (e.g., C might be the boundary of a county, census unit, or school district). Then C is a closed curve, and the integral over a closed curve is denoted by $\oint_C$. The average gradient in the normal direction to the curve C is denoted by

$$\frac{1}{\nu(C)} \oint_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle d\mathbf{s}$$

$$= \frac{1}{\nu(C)} \oint_C \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle \|\mathbf{s}'(t)\| \, dt.$$

If we assume that the surface $Y(\mathbf{s})$ is twice mean-squared differentiable, then the closed-line integral can be written as a double integral over the domain of $\mathbf{s}$, and no explicit parameterization by t is required. In fact, let $F(\mathbf{s}) = \nabla Y(\mathbf{s}) = (F_1(\mathbf{s}), F_2(\mathbf{s}))^T$, so that $F_1(\mathbf{s}) = \partial Y/\partial s_1$ and $F_2(\mathbf{s}) = \partial Y/\partial s_2$. Then

$$\oint_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle \, d\mathbf{s}$$

$$= \oint_C \langle F(\mathbf{s}), n(\mathbf{s}) \rangle \, d\mathbf{s}$$

$$= \oint_C \langle F(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle \|\mathbf{s}'(t)\| \, dt$$

$$= \oint_C \left(F_1 \frac{s'_2(t)}{\|\mathbf{s}'(t)\|} - F_2 \frac{s'_1(t)}{\|\mathbf{s}'(t)\|} \right) \|\mathbf{s}'(t)\| \, dt$$

$$= \oint_C F_1 \, ds_2 - F_2 \, ds_1$$

$$= \int \int_{\mathcal{D}} \left(\frac{\partial F_1}{\partial s_1} + \frac{\partial F_2}{\partial s_2} \right) ds_1 \, ds_2,$$

where the last step follows from Green's theorem in the plane (see the App. for a simple proof). Noting that $F(\mathbf{s}) = \nabla Y(\mathbf{s})$, we rewrite the foregoing in terms of average directional gradients as

$$\frac{1}{\nu(C)} \oint_C \langle \nabla Y(\mathbf{s}), n(\mathbf{s}) \rangle \, d\mathbf{s} = \frac{1}{\nu(C)} \int \int_{\mathcal{D}} \langle \nabla^2 Y(\mathbf{s}), \mathbf{e}_1 + \mathbf{e}_2 \rangle \, d\mathbf{s},$$
(8)

where $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$. This result has computational advantages because the right-side integral can be computed by sampling within the region, which in general is simpler than sampling along a curve. However, if we work exclusively with line segments as described in Sections 4 and 5, then the computational advantage to (8) disappears. Finally, from path independence, it follows that the total gradient in the tangential direction along a closed boundary is 0 because $\mathbf{s}(t_0) = \mathbf{s}(t_1) = \mathbf{s}_0$, and so

$$\oint_C \langle \nabla Y(\mathbf{s}(t)), \mathbf{s}'(t) \rangle \, dt = \oint_{t_0}^{t_1} \langle \nabla Y(\mathbf{s}(t)), \mathbf{s}'(t) \rangle \, dt$$

$$= Y(\mathbf{s}_0) - Y(\mathbf{s}_0) = 0.$$

Thus, for a closed curve, with respect to total or average, gradient it does not make sense to consider the tangential direction.

3.3 Wombling Boundary

With the foregoing formulation in place, we now offer a formal definition of a curvilinear wombling boundary.

Definition. A *curvilinear wombling boundary* is a curve C that reveals a large wombling measure, $\Gamma_{Y(\mathbf{s})}(\mathcal{T})$ or $\bar{\Gamma}_{Y(\mathbf{s})}(\mathcal{T})$ [as given in (7)] in the direction normal to the curve.

If the surface were fixed, then we would have to set a threshold to determine what "large" (say, in absolute value) means. Because the surface is a random realization, $\Gamma_{Y(\mathbf{s})}(\mathcal{T})$ and $\bar{\Gamma}_{Y(\mathbf{s})}(\mathcal{T})$ are random. Hence we declare a curve to be a wombling boundary if, say, a 95% credible set for $\bar{\Gamma}_{Y(\mathbf{s})}(\mathcal{T})$ does not contain 0. It is worth pointing out that although one normal direction [as defined in (5)] is used in (7), $-n(\mathbf{s}(t))$ also would have been a valid choice. Because $D_{-n(\mathbf{s}(t))}Y(\mathbf{s}(t)) = -D_{n(\mathbf{s}(t))}Y(\mathbf{s}(t))$, we note that the wombling measure with respect to one is simply the negative of the other. Thus, in the foregoing definition, large positive as well as large negative values of the integral in (7) would signify a wombling boundary. Being a local concept, an uphill gradient is equivalent to a downhill gradient across a curve, as are the fluxes radiating outward or inward for a closed region.

We also point out that, being a continuous average (or sum) of the directional gradients, the wombling measure may "cancel" the overall gradient effect. For instance, imagine a curve C that exhibits a large positive gradient in the $n(\mathbf{s})$ direction for the first half of its length and a large negative gradient for the second half, thereby canceling the total or average gradient effect. A potential remedy is to redefine the wombling measure using absolute gradients, $|D_{n(\mathbf{s})}Y(\mathbf{s})|$, in (4) and (4'). The corresponding development does not entail any substantially new ideas,

but would destroy the attractive distribution theory in Section 4 and make the computation less tractable. In particular, it would make calibration of the resulting measure with regard to significance much more difficult. (For example, how do we select a threshold?) Moreover, in practice a descriptive contour representation is usually available in which sharp gradients will usually reflect themselves, and one could instead compute the wombling measure for appropriate subcurves of C. Although somewhat subjective, identifying such subcurves is usually unambiguous and leads to robust scientific inference. More fundamentally, in certain applications a signed measure may actually be desirable; one might want to classify a curve as a wombling boundary if it reflects either an overall "large positive" or a "large negative" gradient effect across it. For these reasons, we confine ourselves to working with $D_{n(\mathbf{s})}Y(\mathbf{s})$ and turn to the distribution theory for the wombling measure in the next section.

4. DISTRIBUTION THEORY

Curvilinear wombling amounts to performing predictive inference for a line integral parameterized over $\mathcal{T}$. Let us suppose that $\mathcal{T}$ is an interval, $[0, T]$, which generates the curve $C = \{\mathbf{s}(t) : t \in [0, T]\}$. For any $t^* \in [0, T]$, let $\nu(t^*)$ denote the arc length of the associated curve C_{t^*}. The line integrals for total gradient and average gradient along C_{t^*} are given by $\Gamma_{Y(\mathbf{s})}(t^*)$ and $\bar{\Gamma}_{Y(\mathbf{s})}(t^*)$ as

$$\Gamma_{Y(\mathbf{s})}(t^*) = \int_0^{t^*} D_{n(\mathbf{s}(t))} Y(\mathbf{s}(t)) \|\mathbf{s}'(t)\| \, dt \qquad \text{and}$$
(9)

$$\bar{\Gamma}_{Y(\mathbf{s})}(t^*) = \frac{1}{\nu(t^*)} \Gamma_{Y(\mathbf{s})}(t^*).$$

We seek to infer about $\Gamma_{Y(\mathbf{s})}(t^*)$ based on data $Y = (Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_n))$. Because $D_{n(\mathbf{s}(t))}Y(\mathbf{s}(t)) = \langle \nabla Y(\mathbf{s}(t)), n(\mathbf{s}(t)) \rangle$ is a Gaussian process (from Sec. 2), $\Gamma_{Y(\mathbf{s})}(t^*)$ is a Gaussian process on $[0, T]$, equivalently on the curve C. Note that although $D_{n(\mathbf{s})}Y(\mathbf{s})$ is a process on $\Re^d$, our parameterization of the coordinates by $t \in \mathcal{T} \subseteq \Re^1$ induces a valid process on $\mathcal{T}$. In fact, $\Gamma_{Y(\mathbf{s})}(t^*)$ is $GP(\mu_{\Gamma_{Y(\mathbf{s})}}(t^*), K_{\Gamma_{Y(\mathbf{s})}}(\cdot, \cdot))$, where

$$\mu_{\Gamma_{Y(\mathbf{s})}}(t^*) = \int_0^{t^*} D_{n(\mathbf{s}(t))} \mu(\mathbf{s}(t)) \|\mathbf{s}'(t)\| \, t$$

and

$$K_{\Gamma_{Y(\mathbf{s})}}(t_1^*, t_2^*) = \int_0^{t_1^*} \int_0^{t_2^*} n^T(\mathbf{s}(t_1)) H_K(\Delta(t_1, t_2)) n(\mathbf{s}(t_2))$$

$$\times \|\mathbf{s}'(t_1)\| \|\mathbf{s}'(t_2)\| \, dt_1 \, dt_2,$$

with $\Delta(t_1, t_2) = \mathbf{s}(t_2) - \mathbf{s}(t_1)$. In particular,

$$\text{var}(\Gamma_{Y(\mathbf{s})}(t^*)) = K_{\Gamma_{Y(\mathbf{s})}}(t^*, t^*)$$

$$= \int_0^{t^*} \int_0^{t^*} n^T(\mathbf{s}(t_1)) H_K(\Delta(t_1, t_2)) n(\mathbf{s}(t_2))$$

$$\times \|\mathbf{s}'(t_1)\| \|\mathbf{s}'(t_2)\| \, dt_1 \, dt_2.$$

Evidently, $\Gamma_{Y(\mathbf{s})}(t^*)$ is mean squared continuous. But from the foregoing, note that even if $Y(\mathbf{s})$ is a stationary process,

$\Gamma_{Y(\mathbf{s})}(t^*)$ is not. For any $\mathbf{s}_j$ in the domain of Y,

$$\mathrm{cov}\big(\Gamma_{Y(\mathbf{s})}(t^*), Y(\mathbf{s}_j)\big)$$

$$= \int_0^{t^*} \mathrm{cov}\big(D_{n(\mathbf{s}(t))}Y(\mathbf{s}(t)), Y(\mathbf{s}_j)\big) \|\mathbf{s}'(t)\| \, dt$$

$$= \int_0^{t^*} D_{n(\mathbf{s}(t))}K(\Delta_j(t)) \|\mathbf{s}'(t)\| \, dt, \qquad (10)$$

where $\Delta_j(t) = \mathbf{s}(t) - \mathbf{s}_j$. Based on data $\mathbf{Y} = (Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_n))$, we seek the predictive distribution $P(\Gamma_{Y(\mathbf{s})}(t^*)|\mathbf{Y})$, but note that $Y(\mathbf{s})$ and $\Gamma_{Y(\mathbf{s})}(t^*)$ are processes on different domains; the former is over a connected region in $\Re^2$, while the latter is on a parameterized curve, $\mathbf{s}(t)$, indexed by $\mathcal{T}$. Nevertheless, $\Gamma_{Y(\mathbf{s})}(t^*)$ is derived from $Y(\mathbf{s})$, and we have a valid *joint distribution* $(\mathbf{Y}, \Gamma_{Y(\mathbf{s})}(t^*))$ for any $t^* \in \mathcal{T}$, given by

$$N_{n+1}\left(\begin{pmatrix} \boldsymbol{\mu} \\ \mu_{\Gamma_{Y(\mathbf{s})}}(t^*) \end{pmatrix}, \begin{pmatrix} \Sigma_Y & \boldsymbol{\gamma}_{\Gamma,\mathbf{Y}}(t^*) \\ \boldsymbol{\gamma}_{\Gamma,\mathbf{Y}}^T(t^*) & K_\Gamma(t^*, t^*) \end{pmatrix}\right). \qquad (11)$$

Here $\boldsymbol{\mu} = (\mu(\mathbf{s}_1), \ldots, \mu(\mathbf{s}_n))$ and

$$\gamma_{\Gamma,\mathbf{Y}}^T(t^*) = \big(\mathrm{cov}\big(Y(\mathbf{s}_1), \Gamma_{Y(\mathbf{s})}(t^*)\big), \ldots, \mathrm{cov}\big(Y(\mathbf{s}_n), \Gamma_{Y(\mathbf{s})}(t^*)\big)\big),$$

each component being evaluated from (10).

Suppose that we have regression parameters $\boldsymbol{\beta}$ in $\mu(\mathbf{s}; \boldsymbol{\beta})$ and variance–covariance parameters $\boldsymbol{\eta}$ in $K(\cdot; \boldsymbol{\eta})$. For now, we assume that $\mu(\mathbf{s}; \boldsymbol{\beta})$ is a smooth function in $\mathbf{s}$ [as would be needed to do prediction for $Y(\mathbf{s})$]. Using MCMC, these model parameters, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\eta})$, are available to us as samples, $\{\boldsymbol{\theta}_l\}$, from their posterior distribution $P(\boldsymbol{\theta}|\mathbf{Y})$. Therefore, $P(\Gamma_{Y(\mathbf{s})}(t^*)|\mathbf{Y}) = \int P(\Gamma_{Y(\mathbf{s})}(t^*)|\mathbf{Y}, \boldsymbol{\theta})P(\boldsymbol{\theta}|\mathbf{Y})\,d\boldsymbol{\theta}$ will be obtained by sampling, for each $\boldsymbol{\theta}_l$, $\Gamma_{Y(\mathbf{s})}^l(t^*)$ from $P(\Gamma_{Y(\mathbf{s})}(t^*)|\mathbf{Y}, \boldsymbol{\theta}_l)$, which, using (11), is normally distributed as

$$N\big(\mu_{\Gamma_{Y(\mathbf{s})}}(t^*; \boldsymbol{\beta}_l) - \gamma_{\Gamma,\mathbf{Y}}^T(t^*; \boldsymbol{\eta}_l)\Sigma_Y^{-1}(\boldsymbol{\eta}_l)(\mathbf{Y} - \boldsymbol{\mu}),$$

$$K_{\Gamma_{Y(\mathbf{s})}}(t^*, t^*; \boldsymbol{\eta}_l) - \gamma_{\Gamma,\mathbf{Y}}(t^*; \boldsymbol{\eta}_l)^T \Sigma_Y^{-1}(\boldsymbol{\eta}_l)\gamma_{\Gamma,\mathbf{Y}}(t^*; \boldsymbol{\eta}_l)\big). \quad (12)$$

In particular, for rectilinear wombling (i.e., where $C_{t^*} = \{\mathbf{s}_0 + t\mathbf{u} : t \in [0, t^*]\}$ is a line segment of length t^* joining $\mathbf{s}_0$ and $\mathbf{s}_1 = \mathbf{s}_0 + t^*\mathbf{u}$), we have seen after (7) that $\Gamma_{Y(\mathbf{s})}(t^*)$ equals $\int_0^{t^*} D_{\mathbf{u}^\perp}Y(\mathbf{s}(t))\,dt$. Thus, defining $\Delta_{0j} = \mathbf{s}_0 - \mathbf{s}_j$, we have

$$\mu_{\Gamma_{Y(\mathbf{s})}}(t^*; \boldsymbol{\beta}) = \int_0^{t^*} \langle \mathbf{u}^\perp, \nabla\mu(\mathbf{s}(t); \boldsymbol{\beta})\rangle \, dt;$$

$$(\gamma_{\Gamma,\mathbf{Y}}(t^*; \boldsymbol{\eta}))_j = \int_0^{t^*} D_{\mathbf{u}^\perp}K(\Delta_{0j} + t\mathbf{u}; \boldsymbol{\eta})\,dt, \qquad j = 1, \ldots, n;$$

and

$$K_{\Gamma_{Y(\mathbf{s})}}(t^*, t^*; \boldsymbol{\eta}) = \int_0^{t^*}\int_0^{t^*} -(\mathbf{u}^\perp)^T H_{K(\boldsymbol{\eta})}(\Delta(t_1, t_2))\mathbf{u}^\perp \, dt_1 \, dt_2.$$

These integrals need to be computed for each $\boldsymbol{\theta}_l = (\boldsymbol{\beta}_l, \boldsymbol{\eta}_l)$. Although they may not be analytically tractable [depending on our choice of $\mu(\cdot)$ and $K(\cdot)$], they are one- or two-dimensional integrals that can be efficiently computed using quadrature. Furthermore, because the $\boldsymbol{\theta}_l$'s will already be available, the quadrature calculations (for each $\boldsymbol{\theta}_l$) can be performed ahead of the predictive inference, perhaps using a separate quadrature program, and the output stored in a file for use in the predictive program. The only needed inputs are $\mathbf{s}_0$, $\mathbf{u}$, and the value of t^*. For a specified line segment, we will know these. In fact, for a general curve C, its representation on a given map is as a polygonal curve. As a result, the total or average gradient for C can be obtained through the $\Gamma_{Y(\mathbf{s})}(t^*)$ associated with the line segments that compose C.

Specifically, using GIS software we can easily extract (at high resolution) the coordinates along the boundary, thus approximating C by line segments connecting adjacent points. Thus $C = \bigcup_{k=1}^M C_k$, where the C_k's are virtually disjoint (only one common point at the "join") line segments and $\nu(C) = \sum_{k=1}^M \nu(C_k)$. If we parameterize each line segment as before and compute the line integral along each C_k by the foregoing steps, then the total gradient is the sum of the piecewise line integrals. To be precise, if $\Gamma_k(\mathbf{s}(t_k^*))$ is the line-integral process on the linear segment C_k, then we will obtain predictive samples, $\Gamma_k^{(l)}(\mathbf{s}(t_k^*))$, from each $P(\Gamma_k(\mathbf{s}(t_k^*))|\mathbf{Y})$, $k = 1, \ldots, M$. Inference on the average gradient along C will stem from posterior samples of

$$\frac{1}{\sum_{k=1}^M \nu(C_k)} \sum_{k=1}^M \Gamma_k^{(l)}(\mathbf{s}(t_k^*)).$$

Thus, with regard to boundary analysis, wombling measure reduces a curve to an average gradient and inference to examination of the posterior of the average gradient.

For the flux of a closed region bounded by the curve C, the foregoing framework will still be applicable. Here C_M will join C_1. But, if $\nabla^2 Y(\mathbf{s})$ exists, then (8) suggests a simpler alternative. Because the average gradient is a double integral on the region $\mathcal{D}$, we may easily approximate the flux as a Monte Carlo average. Whether or not the consequent computational ease warrants the smoothness assumption would depend on the application.

When $Y(\mathbf{s})$ is an isotropic Gaussian process with constant mean, $\mu(\mathbf{s}) = \mu$, and correlation function $K(\|\Delta\|; \sigma^2, \phi) = \sigma^2 \exp(-\phi\|\Delta\|^2)$, the calculations are simplified. We have $\nabla\mu(\mathbf{s}) = 0$, $\mu_\Gamma(t^*) = 0$, and $H_K(\Delta) = -2\sigma^2\phi\exp(-\phi\|\Delta\|^2) \times (I - 2\phi\Delta\Delta^T)$. Further calculations reveal that $(\gamma_{YZ}(t^*; \sigma^2, \phi))_j$ can be computed as

$$c(\sigma^2, \phi, \mathbf{u}^\perp, \Delta_{0j})$$

$$\times \big(\Phi\big(\sqrt{2\phi}(t^* + \langle\mathbf{u}, \Delta_{0j}\rangle)\big) - \Phi\big(\sqrt{2\phi}\langle\mathbf{u}, \Delta_{0j}\rangle\big)\big), \quad (13)$$

where $\Phi(\cdot)$ is the standard Gaussian cdf and

$$c(\sigma^2, \phi, \mathbf{u}^\perp, \Delta_{0j})$$

$$= -2\sigma^2\sqrt{\pi\phi}\langle\mathbf{u}^\perp, \Delta_{0j}\rangle \exp\big(-\phi|\langle\mathbf{u}^\perp, \Delta_{0j}\rangle|^2\big).$$

These computations can be performed using the Gaussian cdf function, and quadrature is needed only for $\mathrm{var}(\Gamma_{Y(\mathbf{s})}(t^*))$. We use this model to demonstrate a simulation of curvilinear wombling. In real data settings, such as our example in Section 6, the Matérn is favored, providing more flexible control of surface smoothness.

Returning to the model $Y(\mathbf{s}) = \mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} + W(\mathbf{s}) + \epsilon(\mathbf{s})$ with $\mathbf{x}(\mathbf{s})$ a general covariate vector, $W(\mathbf{s}) \sim \mathrm{GP}(0, \sigma^2\rho(\cdot, \phi))$, and $\epsilon(\mathbf{s})$ a zero-centered white-noise process with variance τ^2, consider boundary analysis for the residual surface $W(\mathbf{s})$. In fact, boundary analysis on the spatial residual surface is feasible in generalized linear modeling contexts with exponential families, where $W(\mathbf{s})$ may be viewed as a nonparametric latent structure in the mean of the parent process; see Section 7.2.

Letting $\Gamma_{W(\mathbf{s})}(t)$ and $\bar{\Gamma}_{W(\mathbf{s})}(t)$ denote the total and average gradient processes [as defined in (7)] for $W(\mathbf{s})$, we seek the posterior distributions $P(\Gamma_{W(\mathbf{s})}(t^*)|\mathbf{Y})$ and $P(\bar{\Gamma}_{W(\mathbf{s})}(t^*)|\mathbf{Y})$. Note that

$$P\big(\Gamma_{W(\mathbf{s})}(t^*)|\mathbf{Y}\big)$$

$$= \int P\big(\Gamma_{W(\mathbf{s})}(t^*)|\mathbf{W}, \boldsymbol{\theta}\big) P(\mathbf{W}|\boldsymbol{\theta}, \mathbf{Y}) P(\boldsymbol{\theta}|\mathbf{Y}) \, d\boldsymbol{\theta} \, d\mathbf{W}, \quad (14)$$

where $\mathbf{W} = (W(\mathbf{s}_1), \ldots, W(\mathbf{s}_n))$ denotes a realization of the residual process and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \phi, \tau^2)$. Sampling of this distribution again proceeds in a posterior predictive fashion using posterior samples of $\boldsymbol{\theta}$ and is expedited in a Gaussian setting because $P(\mathbf{W}|\boldsymbol{\theta}, \mathbf{Y})$ and $P(\Gamma_{W(\mathbf{s})}(t^*)|\mathbf{W}, \boldsymbol{\theta})$ are both Gaussian distributions.

Formal inference for a wombling boundary is done more naturally on the residual surface $W(\mathbf{s})$ [i.e., for $\Gamma_{W(\mathbf{s})}(t^*)$ and $\bar{\Gamma}_{W(\mathbf{s})}(t^*)$] because $W(\mathbf{s})$ is the surface containing any nonsystematic spatial information on the parent process $Y(\mathbf{s})$. Because $W(\mathbf{s})$ is a mean-0 process, $\mu_{\Gamma_{W(\mathbf{s})}}(t^*; \boldsymbol{\beta}) = 0$, and thus we need to check for the inclusion of this null value in the resulting 95% credible intervals for $\Gamma_{W(\mathbf{s})}(t^*)$ or, equivalently, for $\bar{\Gamma}_{W(\mathbf{s})}(t^*)$. Again, this clarifies the issue of the normal direction mentioned in Section 3.3; significance using $n(\mathbf{s}(t))$ is equivalent to significance using $-n(\mathbf{s}(t))$. We need only select and maintain a particular orthogonal direction relative to the curve. In accordance with our remarks concerning absolute gradients in Section 3.3, we could compute (9) using $|D_{n(\mathbf{s}(t))}Y(\mathbf{s}(t))|$ using a Riemann sum, but this would be computationally expensive and would not offer a Gaussian calibration of significance.

5. A SIMULATED ILLUSTRATION OF ALGORITHMIC AND STATISTICAL WOMBLING

We illustrate curvilinear wombling with the following simulated data example. We first generated data from a stationary random field, $Y(\mathbf{s})$, following $GP(\mu, \sigma^2 \exp(-\phi\|\Delta\|^2))$. We use the Gaussian correlation function here simply for the convenience of closed-form (up to the Gaussian cdf) expressions for the posterior predictive distribution of $\Gamma(t^*)$ [see eq. (13)]. In particular, we generated 100 observations from a 10×10 square region with parameters $\mu = 5$, $\sigma^2 = 1$, and $\phi = .5$. The maximum observed intersite distance was 12.6 units.

We analyzed the data using a flat prior for μ, an inverted-gamma IG(2, .001) prior for σ^2 (so that $1/\sigma^2$ has mean 2,000 and variance 2.0E+6), and a G(.5, .01) prior for ϕ. An MCMC algorithm was designed with Gibbs updates for μ and σ^2, and a Metropolis update for ϕ was used. Burn-in was diagnosed at 5,000 iterations, and posterior samples of size 6,000 were retained. Table 1 shows that the posterior estimation was successful in capturing the true value for each of the parameters. Posterior predictive samples were used to predict the random field $Y(\mathbf{s})$ on a grid over the entire domain to produce the mean

Table 1. Parameter Estimates in the Simulated Example

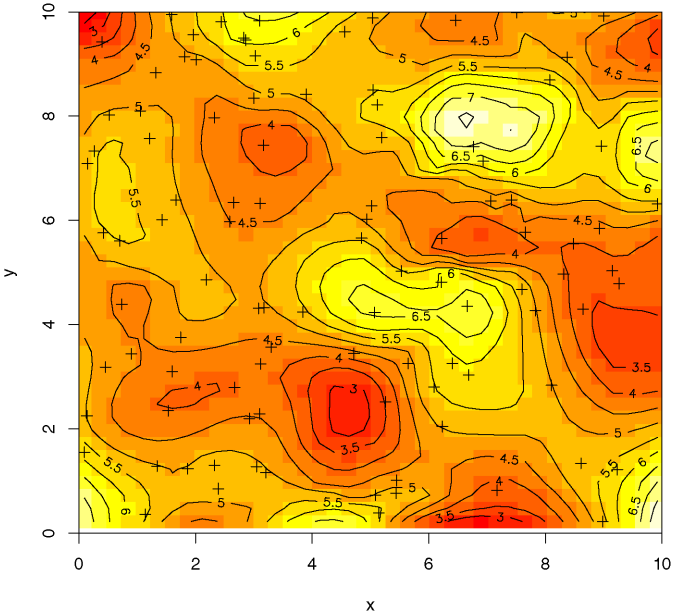| Parameters	50% (2.5%, 97.5%)	True value
μ	5.002 (4.661, 5.329)	5.0
σ^2	.887 (.662, 1.271)	1.0
ϕ	.526 (.431, .858)	.5

Figure 1. The Predicted Mean Response Surface Plot for the Simulated Random Field With Overlaid Contours and the 100 Observation Locations Indicated by "+".

response surface in Figure 1, where the "+" indicate our 100 observation locations. The darker shades represent lower values of $Y(\mathbf{s})$, and lighter shades represent higher values. Note that the overlaid contour lines provide additional details about the response surface topography and can help identify regions with high gradients. However, as mentioned in Section 1, contours are formed by connecting points with the same elevation and are not directly related to gradients.

Existing algorithmic wombling methods use deterministic interpolators. These interpolators are typically of the form $f(\mathbf{s}) = \sum_{i=1}^{n} w_i(\mathcal{S}; \mathbf{s}) f(\mathbf{s}_i)$, where $f(\mathbf{s}_i) = Y(\mathbf{s}_i)$'s are the observed values in $\mathcal{S}$ and $w_i(\mathcal{S}; \mathbf{s})$ are weight functions of $\mathbf{s}$ that also depend on the entire configuration of observed sites $\mathcal{S}$. The choice of the interpolator depends on the regularity in the configuration of the locations, a triangulation of the domain, the smoothness of the surface, the speed of implementation, and numerical error analysis (see Cohen et al. 2001; Phillips 2003 for further details regarding such procedures). Here we demonstrate a typical implementation, using the popular wombling software BoundarySEER (*www.terraseer.com*), where the domain is first triangulated using a Delaunay triangulation with the observed locations forming the vertices of the triangles [see Figs. 2(a) and 2(b)]. Next, the bivariate linear interpolator $f(s_1, s_2) = a + bs_1 + cs_2$ is fitted to each triangular plate, with the coefficients a, b, and c determined from the observed data at the three vertices. The gradient vector ∇f and the *maximal* gradient value over all directions, $\|\nabla f\| = [(\partial f/\partial s_1)^2 + (\partial f/\partial s_2)^2]^{1/2}$, are then computed at the centroids of the triangles, and those with the highest 5% or 10% of maximal gradients are identified as points that should form a part of the boundary. The corresponding triangles are then shaded as "zones" in which boundaries are likely to be found. Finally, these centroids are connected if they belong to adjacent zones (other angle-based connection rules also apply). Figures 2(a) and 2(b) show the output from such an implementation,
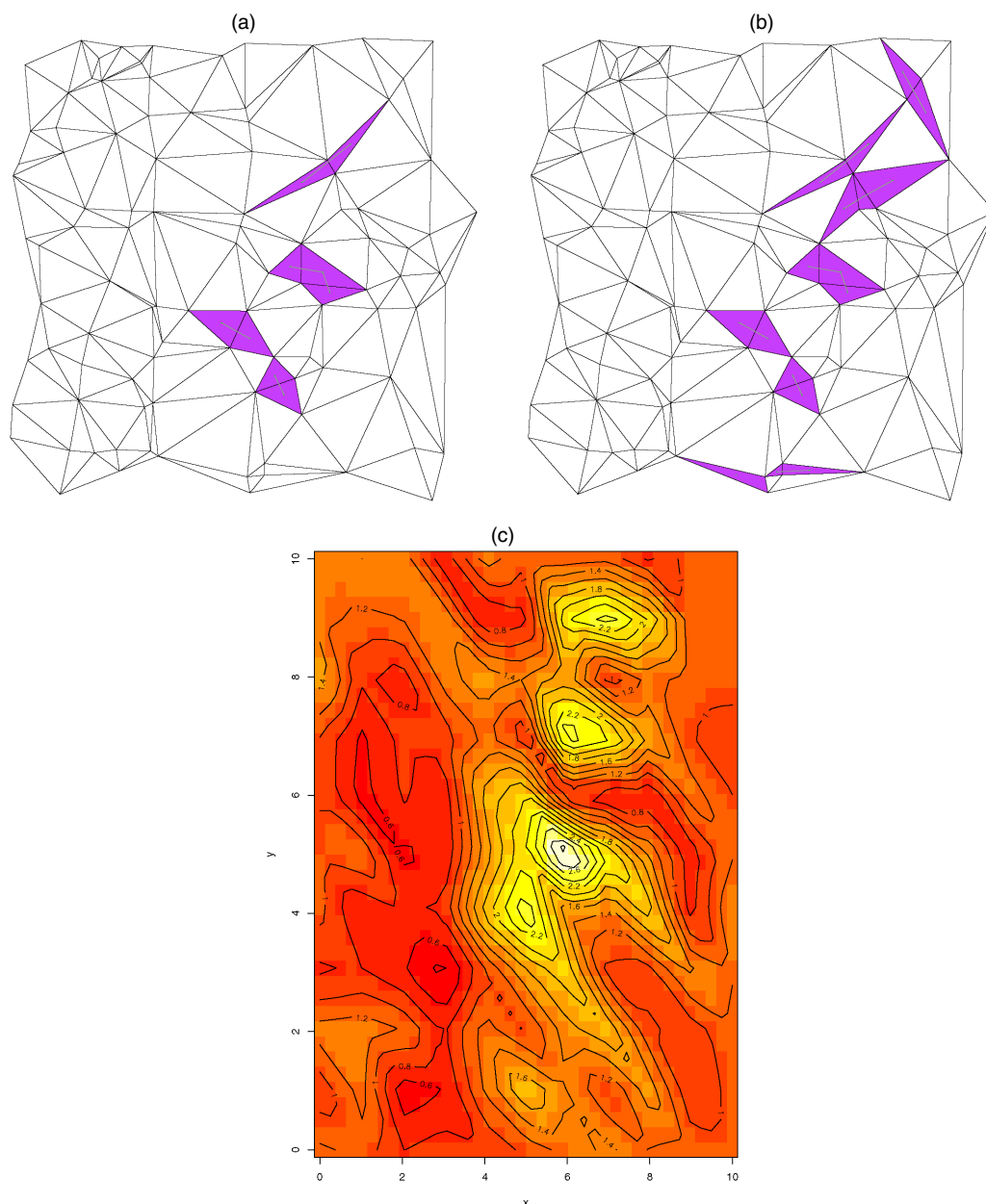
Figure 2. Algorithmic Point-Referenced Wombling Using BoundarySEER [(a) and (b)] and a Predictive Surface of the Maximal Gradient Process (c).

with (a) corresponding to the 5% threshold and (b) corresponding to the 10% threshold.

Although such algorithms are fast and simple to implement, they rely on the particular triangulation to obtain the zones of rapid change that determine the boundary representations. Thus boundaries are "internally constructed," and one cannot compute the average gradient along a curve obtained from GIS or external software. Furthermore, the construction of the boundaries is discrete and may not capture the local topography of the response surface. Uncertainty is not accounted for, and there is no way to test or validate these boundaries. Note that the directional derivative processes discussed in Section 2 can also be used to obtain zones of rapid change (also see Allard, Gabriel, and Bacro 2005). In fact, the image-contour plot in Figure 2(c) shows the mean predicted surface of $\|\nabla Y(\mathbf{s})\|$ over the spatial

domain. Unfortunately, it is not clear how to use these plots to ascertain curves that have large gradients.

Turning to the illustration of statistical inference with the same data, Figure 3 displays the mean predicted response surface but with certain curves and segments marked on them. It is important to note that it does not make sense to talk about "true" wombling boundaries here, because we do not know what the "true" equation of the surface is. Hence validation of our proposed methods must use the topographic information from contour plots. For this, we consider five different curves or segments, which we test for average gradients. The segments AB, CD, and EF are all of unit length with differing local topographies. The segment AB lies in the space between two contour lines, CD lies on relatively flat land with not much contours, and EF cuts across a stream of contour lines. We
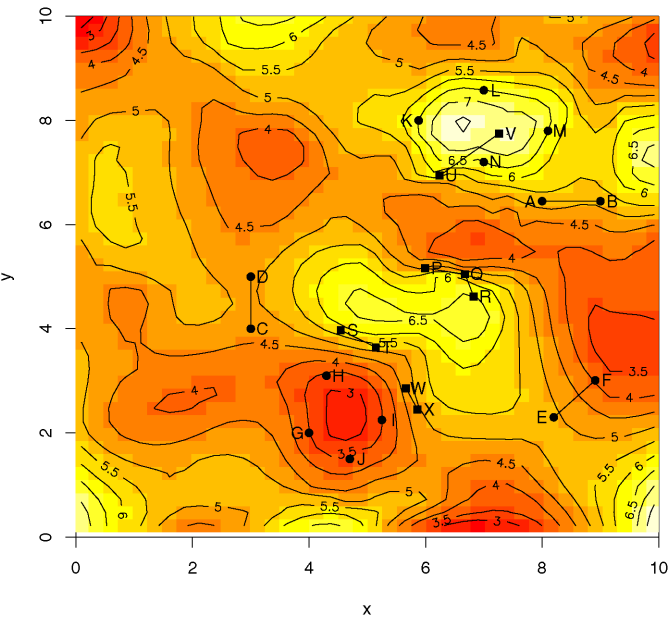
Figure 3. Curvilinear Boundary Analysis for Simulation Example.

also investigate two closed curves for gradients and flux that are actually shown as contours: GHIJG and KLMNK. Finally, we evaluate gradients along PQ, QR, ST, UV, and WX, which are the "boundaries" proposed by BoundarySEER.

Table 2 presents the results of our analysis. We present the average directional gradients along $n(\mathbf{s}(t))$. We take $n(\mathbf{s}(t))$ to be the uphill direction with positive median gradients. Again, we need not include the results for $-n(\mathbf{s}(t))$ because the latter is just a reflection of the former. Recall that a curve will form a wombling boundary if the 95% credible interval for its wombling measure does not include 0. The *true values* in Table 2 correspond to the conditional mean, $E[\Gamma(t)|\mathbf{Y}, \mu, \sigma^2, \phi]$, with the true parameter values plugged in. The results are quite consistent with the local topography of the curves. Consider the

Table 2. Curvilinear Gradient Assessment From Figure 3 for the Mean Response Surface in the Simulation Example

Curve	Average gradient [n($\mathbf{s}(t)$)]	True value [n($\mathbf{s}(t)$)]
AB	1.242 (.886, 1.574)	1.275
CD	.055 (−.648, .745)	.110
EF	.140 (−.780, 1.060)	.032
Curve GHIJG		
GH	.675 (.011, 1.312)	.735
HI	.899 (.545, 1.262)	.892
IJ	1.745 (1.313, 2.153)	1.799
JG	.935 (.590, 1.283)	.930
Flux	1.063 (.624, 1.493)	1.086
Curve KLMNK		
KL	1.015 (.424, 1.619)	.967
LM	1.646 (.473, 2.842)	1.777
MN	2.291 (1.721, 2.854)	2.310
NK	2.119 (1.411, 2.819)	2.178
Flux	1.794 (1.024, 2.569)	1.839
PQ	2.152 (1.776, 2.572)	2.124
QR	.322 (−.163, .827)	.289
ST	2.627 (2.117, 3.125)	2.634
UV	.315 (−.198, .773)	1.381
WX	1.843 (1.355, 2.360)	1.779

NOTE: For each segment or curve, the estimates in the direction corresponding to a positive gradient are presented.

rectilinear path of unit length, AB, going in the east–west direction. With $\mathbf{u} = (1, 0)$ and $\mathbf{u}^\perp = (0, -1)$, we see a significant average gradient cutting across AB; in other words, AB forms a wombling boundary. This is not surprising, because the contours reveal an uphill gradient as one moves northward, cutting across AB. In contrast, CD is a unit-length rectilinear segment on relatively flat land. There are no visible gradients, as evidenced by the absence of contour lines, and indeed from Table 2 we see an insignificant average gradient across CD. Generally, significant boundaries are found *parallel* to the contour lines, not cutting across contours. This is because significant directional gradients are found in directions cutting across contour lines, and we take the average directional gradients *perpendicular* to the curve. This is illustrated by the curve EF, which cuts across a flow of north–south contours. The gradient here is parallel to EF (i.e., along $\mathbf{u}$), and hence the directional gradients along $\mathbf{u}^\perp$ are insignificant, producing an insignificant average.

The next segment of Table 2 presents our results on curvilinear wombling for the contours GHIJG and KLMNK in Figure 3. The former lies in the slopes of a valley; the latter, on the slopes of a plateau. We first assess the significance of each curvilinear segment making up the contour, then evaluate the statistical flux of the enclosed region. Following Section 4, the statistical flux here is just a weighted average of the average gradients of each piece making up the contour. To be precise, suppose that C is the closed contour given by GHIJG. Once we have computed the curvilinear gradients for GH, HI, IJ, and JG, we can simply take the weighted average of the curvilinear average gradients of the four segments, with weights $\nu(C_k)/\nu(C)$ yielding the statistical flux reported in Table 2. Each of the segments as well as the flux are significant, again corroborated by the topography. Note that $n(\mathbf{s}(t))$ is always the uphill vector, which means that it is the "outward normal" for GHIJG and the "inward normal" for KLMNK. Evaluating the flux using Green's result, we obtained a median "outward" flux of 1.070 with a credible interval of (.633, 1.512) for GHIJG and corresponding values of 1.758 and (1.002, 2.571) for the "inward" flux of KLMNK. These values concur with those obtained as weighted averages in Table 2.

The last segment of Table 2 presents results on testing significance of the boundaries proposed by algorithmic wombling. Indeed P, Q, R, S, T, U, and V are the selected centroids forming the top 5% of maximal gradients, and the connection algorithm produces the five boundaries shown. The significance of PQ is quite expected, lying along a stream of contours. Cutting across PQ would indeed indicate a smooth jump in elevation. Similarly, ST and WX are also formally tested to be wombling boundaries, although WX is somewhat less pronounced. QR and UV, on the other hand, do not have significant gradients and fail the formal test of being a wombling boundary. Although both lie in zones of possibly rapid change, they tend to cut across contours (much like EF) rather than move along them. Even though the points Q, R, U, and V all have significant maximal gradients, the tessellation-based connection of these points is unable to capture the local topography and the continuous gradient along the curve.

6. SPATIAL REGRESSION EXAMPLES

We illustrate boundary analysis on the spatial residual surfaces from two spatial regression models. The first of

these models temperature against precipitation from a weather dataset obtained from the National Center for Atmospheric Research (NCAR) in Boulder, Colorado. The second uses a spatial logistic regression model for species presence/absence data from Connecticut. We note that Bayesian analysis of these models using posterior sampling is now available in user-friendly software such as WinBUGS (*www.mrc-bsu.cam.ac.uk/bugs*), and R packages such as geoR and geoRglm available from CRAN (comprehensive R archives network) (*http://cran.us.r-project.org*). Keep in mind, however, that the spatial processes modeled there usually assume a fixed Matérn smoothness parameter ν. Valid Bayesian boundary analysis can be performed by setting $\nu > 1$ (or $\nu > 2$ for evaluating the flux using Green's theorem). The powered exponential family is also usually available. However, we present our analysis keeping ν unknown, as in the previous section, implementing our models in C/C++ with posterior summarizations and graphics in R.

6.1 A Spatial Gaussian Regression Example

We have mean temperature measurements (in 10°C units) obtained at 50 sites in July 1997 as our dependent variable $Y(\mathbf{s})$. Also supplied are precipitation measurements (in 100-meter units) at each site, so that the covariate vector $\mathbf{x}(\mathbf{s})$ comprises an intercept and a precipitation. A univariate spatial model (see, e.g., Banerjee et al. 2004, chap. 5), $Y(\mathbf{s}) = \mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} + W(\mathbf{s}) + \epsilon(\mathbf{s})$, is used to explain temperature given precipitation, allowing for both pure error [$\epsilon(\mathbf{s}) \overset{iid}{\sim} N(0, \tau^2 I)$] and spatial correlation in the data [$W(\mathbf{s}) \sim GP(0, \sigma^2 \rho(\cdot; \phi, \nu))$].

We adopted a flat prior for $\boldsymbol{\beta}$ and relatively vague inverted-gamma, IG(2, .001), priors for σ^2 and τ^2. We used the Matérn correlation, $\rho(\phi, \nu; d)$, with a gamma prior for the correlation decay parameter, ϕ, specified so that the prior spatial range has a mean of about half of the maximum intersite distance in our data, and a U(1, 2) prior for the smoothness parameter ν. That is, $\nu > 1$ is needed for mean squared differentiability, whereas, following Stein (1999), the data will be unable to distinguish $\nu = 2$ from values of $\nu > 2$. Three parallel MCMC chains were run for 10,000 iterations. Convergence was diagnosed by monitoring mixing, Gelman–Rubin diagnostics, autocorrelations, and cross-correlations. In each case, 5,000 iterations were deemed enough to allow sufficient mixing of the chains, and the remaining 15,000 samples (5,000 × 3) were retained for posterior analysis.

Table 3 gives the parameter estimates for the spatial regression model. Precipitation seems to have a moderately positive impact on temperature. Note that the residual spatial story is quite strong here with $\sigma^2/(\sigma^2 + \tau^2) \approx .72$, implying that the spatial variance explains about 72% of the residual variation.
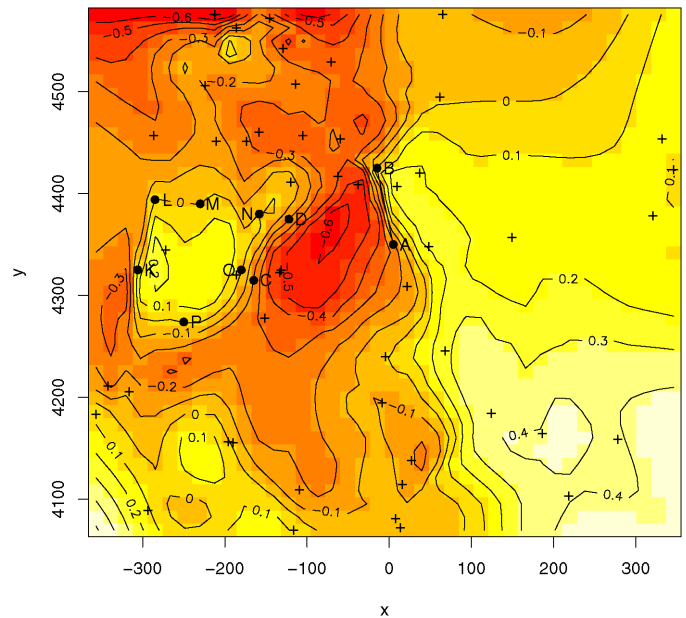

Figure 4. Curvilinear Boundary Analysis on the Spatial Residual Surface of Temperature Given Precipitation, Using GIS Extracted Boundary Segments for the Colorado Data.

This is corroborated by Figure 4, which shows an image plot with contours of the posterior mean residual surface $W(\mathbf{s})$ over the spatial domain, with the "+" indicating our 50 sampling locations. The coordinates are obtained from a sinusoidal planar projection of latitude and longitude and are in kilometers. The contours reveal an interesting topography with temperature residuals rich in spatial variation. Posterior samples of the spatial range are obtained by solving $\rho(\phi, \nu; d) = .05$ for each posterior draw of ϕ and ν. The estimated spatial range is about 262 km, which seems reasonable in our domain, with the maximum intersite distance of approximately 780 km.

It is expected that elevation can help explain the spatial temperature residuals. If elevation information is available over the entire domain, then one could, for example, model elevation as a regressor and study this impact. But when interest is focused on the impact of a curve or path of high elevation, estimating curvilinear gradients seems more appropriate. For choosing paths of high elevation to test for gradients, we used the open-source GIS software GRASS (*http://grass.itc.it/*) to identify topographic features. For example, in Figure 4 the linear segment AB and the curvilinear segment CD (along the illustrated contour) both represent mountain slopes, whereas the enclosed region bounded by KLMNOPK forms a valley.

To test for significant gradients on $W(\mathbf{s})$ along the curves specified earlier, we compute the average gradients along AB, CD, and the curves forming the region KLMNOPK. These are computed using posterior predictive samples using (14). These gradients, along with the statistical flux of KLMNOPK, are presented in Table 4. As one might expect, we find a significant average gradient for both AB and CD, being in the middle of a dense flow of contour lines. As in the simulated example, difference boundaries on spatial residual surfaces are also likely to be seen parallel to dense flows of contours. The next part of Table 4 shows gradient evaluation of the curve KLMNOPK and the statistical flux of the enclosed region. The results seem

Table 3. Parameter Estimates for the Gaussian Spatial Regression Example

Parameters	50% (2.5%, 97.5%)
Intercept	2.827 (2.131, 3.866)
Precipitation	.037 (.002, .072)
σ^2	.134 (.051, 1.245)
ϕ	7.39E−3 (4.71E−3, 51.21E−3)
ν	1.577 (1.210, 1.931)
Range (in km)	261.6 (37.8, 418.3)
τ^2	.051 (.022, .092)

Table 4. Boundary Assessment on the
Spatial Residual Surface From Figure 4

Curve	Average gradient
AB	2.713 (2.328, 3.076)
CD	1.711 (1.280, 2.087)
KL	1.219 (.609, 1.830)
LM	.938 (−.003, 1.863)
MN	.842 (−.027, 1.750)
NO	1.668 (.727, 2.593)
OP	1.667 (.682, 2.542)
PK	1.129 (.587, 1.649)
Flux	1.257 (.448, 2.045)

quite consistent with the topography revealed in Figure 4. From the practitioner's viewpoint, statistically confirming boundaries with high gradients is a way of capturing spatial information that is difficult to incorporate in the trend surface. Such analysis assists scientists in formulating physical models incorporating such information.

6.2 A Spatial Logistic Regression Example

We consider data collected from 603 locations in Connecticut with presence/absence and abundance scores for some individual invasive plant species, plus environmental covariates. The covariates are available only at the sample locations, not on a grid. The response variable $Y(\mathbf{s})$ is a presence–absence binary indicator (0 for absence) for one species, *Celastrus orbiculatus*, at location $\mathbf{s}$. There are four categorical covariates: habitat class (representing the current state of the habitat) of four different types, land use and land cover (LULC) types (land use/cover history of the location; e.g., always forest, formerly pasture now forest) at five levels, and a 1970 category number (LULC at one point in the past, 1970; e.g., forest, pasture, residential) with six levels. In addition, we have an ordinal covariate, canopy closure percentage (percent of the sky blocked by "canopy" of leaves of trees; a location under mature forest would have close to 100% canopy closure, whereas a forest edge would have closer to 25%) with four levels in increasing order, a binary variable for heavily managed points (0 if "no"; "heavy management" implies active landscaping or lawn mowing), and a continuous variable measuring the distance from the forest edge in the logarithm scale. Figure 5 is a digital terrain image of the study domain, with the labeled curves indicating forest edges extracted using the GIS software ArcView (*http://www.esri.com/*). Ecologists are interested in evaluating spatial gradients along these 10 natural curves and identifying them as wombling boundaries.

We fit a logistic regression model with spatial random effects,

$$\log\left(\frac{P(Y(\mathbf{s}) = 1)}{P(Y(\mathbf{s}) = 0)}\right) = \mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} + W(\mathbf{s}),$$

where $\mathbf{x}(\mathbf{s})$ is the vector of covariates observed at location $\mathbf{s}$ and $W(\mathbf{s}) \sim \mathrm{GP}(0, \sigma^2 \rho(\cdot; \phi, \nu))$ is a Gaussian process with $\rho(\cdot; \phi, \nu)$ as a Matérn correlation function. Whereas $Y(\mathbf{s})$ is a binary surface that does not admit gradients, conducting boundary analysis on $W(\mathbf{s})$ is perfectly legitimate. The residual spatial surface reflects unmeasured or unobservable environmental features in the mean surface. Again, attaching curvilinear wombling boundaries to the residual surface tracks rapid changes in the departure from the mean surface.

We adopt priors similar to those in the preceding section: a completely noninformative flat prior for $\boldsymbol{\beta}$, an inverted-gamma IG(2, .001) prior for σ^2, and the Matérn correlation function with a gamma prior for the correlation decay parameter, ϕ, specified so that the prior spatial range has a mean of about half of the observed maximum intersite distance (the maximum distance is 11,887 m based on a UTM projection), and a U(1, 2) prior for the smoothness parameter ν. Again, we ran three parallel MCMC chains for 15,000 iterations each, 10,000 iterations revealed sufficient mixing of the chains, and the remaining 15,000 samples (5,000 × 3) were used for posterior analysis.

Table 5 presents the posterior estimates of the model parameters. We do not have a statistically significant intercept, but most of the categorical variables reveal significance. Types 2 and 4 for habitat class have significantly different effects than type 1, all four types of LULC show significant departure from the baseline type 1, and for the 1970 category number, category 2 shows a significant negative effect, whereas categories 4 and 6 show significant positive effects compared with category 1. Canopy closure is significantly positive, implying higher presence probabilities of *Celastrus orbiculatus* with higher canopy blockage, whereas points that are more heavily managed appear to have a significantly lower probability of species presence, as does the distance from the nearest forest edge. Posterior summaries of the spatial process parameters are also presented, and the effective spatial range is approximated to be around 1,109.3 meters. These produce the mean posterior surface of $W(\mathbf{s})$, shown in Figure 6 with the 20 endpoints of the forest edges from Figure 5 labeled to connect the figures.

Finally, Table 6 presents the formal curvilinear gradient analysis for the 10 forest edges in Figure 5. We see that 6 of the 10 edge curves (all except CD, EF, KL, and MN) are formally tested to be wombling boundaries. Our methodology proves useful here, because some of these edge curves meander along the terrain for very long distances. Indeed, although the residual surface in Figure 6 reveals a general pattern of spatial variation (higher residuals in the South), it is difficult to make visual assessments on the size (and significance) of average gradients for the longer curves. Furthermore, with nongridded data as here, the surface interpolators [in this case the Akima (1996) interpolator in R] often find it difficult to extrapolate beyond a convex hull of the site locations. Consequently, parts of the curve [e.g., endpoints C, G, and (almost) T] lie outside the fitted surface, making local visual assessment on them impossible.

Quickly and reliably identifying forest edges could be useful in determining boundaries between areas of substantial anthropogenic activity and minimally managed forest habitats. Such boundaries are important because locations where forest blocks have not been invaded by exotic plant species may be subject to significant seed rain from these species. These boundaries thus might form important "front lines" for efforts at monitoring or controlling invasive species.

7. SUMMARY AND DISCUSSION

Heretofore, a fully inferential approach to boundary analysis was not available in the literature. Although subjective assessments regarding wombling boundaries can be drawn from contour plots, formal inference (incorporating uncertainty) may
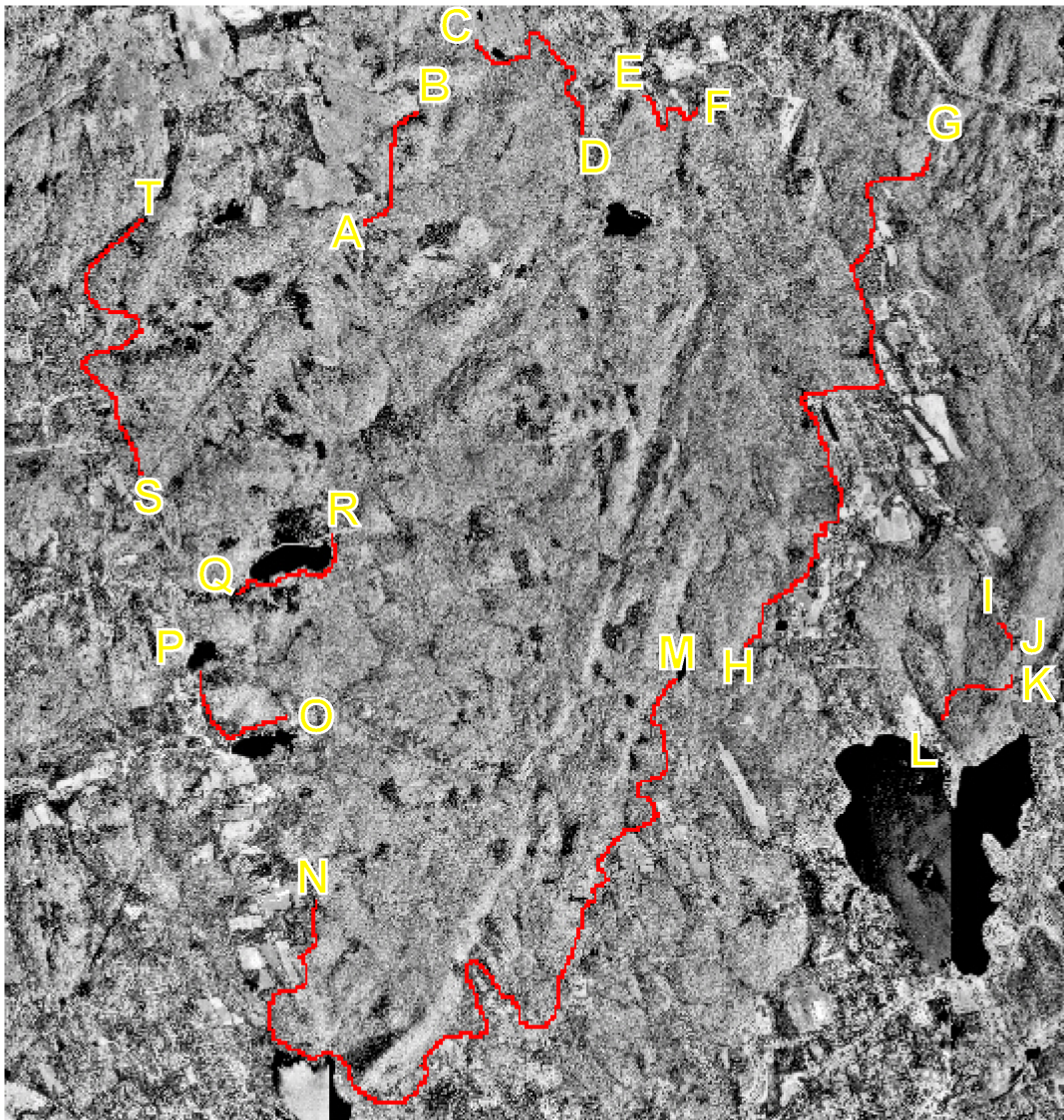
Figure 5. A Digital Image of the Study Domain in Connecticut Indicating the Forest Edges as Marked Curves. These are assessed for significant gradients. Eastings range from 699148 to 708961; northings range from 4604089 to 4615875 for the picture.

be desirable. In contrast to edge detection and contouring problems in image analysis and pattern recognition, our primary focus has been on the modeling of spatial gradients along curves from possibly irregularly spaced point-referenced data, without resorting to alignment algorithms for mapping such data into pixels. Starting with recently developed theory on directional derivatives at points on a random realization of a Gaussian process, we have moved from points to curves, using formal vector analytic methods to determine whether a proposed curve should be declared a curvilinear wombling boundary. Our methods are particularly geared toward statistical modelers, allowing boundary analysis on response and residual surfaces from spatial regression models. In fact, these can be easily extended to nonstationary models with weighted mixtures of Gaussian processes, and even to nonparametric spatial models using continuous mixtures of Gaussian processes.

The vector-analytic framework with parametric curves also opens up a number of new directions for further research. For instance in practice it would be of interest to develop a sto-

chastic or model-based method to construct wombling boundaries. As the earlier discussion suggests, it is both natural and easiest to construct boundaries using polygonal curves, and hence to consider only piecewise linear boundaries. Thus such a construction requires a starting point, say $\mathbf{s}_0^{(1)}$, and then travels an unknown distance $t_{(1)}^*$ in a direction $\mathbf{u}_{\text{opt}}^{(1)}$ to a new point $\mathbf{s}_0^{(2)}$, which becomes the new starting point. Here $\mathbf{u}_{\text{opt}}^{(1)} = \mathbf{u}_{\text{max}}^{\perp}$ would be, say, the perpendicular direction to $\mathbf{u}_{\text{max}} = \nabla Y(\mathbf{s}_0^{(1)})/\|\nabla Y(\mathbf{s}_0^{(1)})\|$, the direction of steepest spatial gradient at $\mathbf{s}_0^{(1)}$. However, the choice of starting point is not clear. It should be a point with a large maximum gradient, but this does not uniquely determine a point. In fact, if the maximum gradient at $\mathbf{s}_0$ is "significant," then our smoothness assumptions imply that this will be so for all points in a neighborhood of $\mathbf{s}_0$. Moreover, there will be many parts of the region where we may look for $\mathbf{s}_0$'s, many places to start. Next, even with $\mathbf{u}_{\text{opt}}^{(1)}$, according to sign, we have two choices for moving from $\mathbf{s}_0$. Furthermore, if the maximum gradient at $\mathbf{s}_0$ is the highest as we move

Table 5. Parameter Estimates for the Logistic Spatial
Regression Example

Parameters	50% (2.5%, 97.5%)
Intercept	.983 (−2.619, 4.482)
Habitat class (baseline: type 1)	
Type 2	−.660 (−1.044, −.409)
Type 3	−.553 (−1.254, .751)
Type 4	−.400 (−.804, −.145)
Land use land cover types (baseline: level 1)	
Type 2	.591 (.094, 1.305)
Type 3	1.434 (.946, 2.269)
Type 4	1.425 (.982, 1.974)
Type 5	1.692 (.934, 2.384)
1970 category types (baseline: category 1)	
Category 2	−4.394 (−6.169, −3.090)
Category 3	−.104 (−.504, .226)
Category 4	1.217 (.864, 1.588)
Category 5	−.039 (−.316, .154)
Category 6	.613 (.123, 1.006)
Canopy closure	.337 (.174, .459)
Heavily managed points (baseline: no)	
Yes	−1.545 (−2.027, −.975)
Log edge distance	−1.501 (−1.891, −1.194)
σ^2	8.629 (7.005, 18.401)
ϕ	1.75E−3 (1.14E−3, 3.03E−3)
ν	1.496 (1.102, 1.839)
Range (in m)	1109.3 (632.8, 1741.7)

in direction $\mathbf{u}_{\text{opt}}^{(1)}$, then, if we decided to select $t_{(1)}^*$ to maximize the average gradient along the line segment in this direction, we would not move; any departure from $\mathbf{s}_0$ would yield a smaller average. Finally, even if we did move and created a sequence of segments, if a closed boundary were appropriate, then such a construction would close with probability 0. To avoid endless circling, ad hoc intervention must be imposed. In summary, we have examined "algorithms" that resolve these various problems by reference to a contour plot and found that these perform well in comparison with currently available wombling boundary software. However, due to the ad hoc decisions that are required, we cannot offer a satisfactory version at present.

More fundamentally, here we have confined ourselves to curves that track zones of rapid change. However, as we alluded to earlier, zones of rapid change are areal notions; description by a curve may be an unsatisfying simplification. Describing zones as areal quantities (i.e., as sets of nonzero Lebesgue measure in $\Re^2$) is an alternative. When proceeding, the crucial issue is to formalize shape-oriented definitions of a wombling boundary. Although much work has been done on statistical shape analysis, its use in the point-referenced spatial data context that we have set out herein is unexplored. There are many possibilities using formal differential geometry and calculus of variations, providing directions for future research. Finally, we also note that our approach is built entirely on the specification of a point-referenced spatial process model. One might examine the boundary analysis problem from an alternative modeling perspective, in which curves or zones arise as random processes. Possibilities include line processes, random tessellations, and random areal units.

APPENDIX: GREEN'S THEOREM

Green's theorem on the plane is a special case of Stokes's theorem on general manifolds (see, e.g., Rudin 1976). We provide a direct derivation for simple closed regions. Consider the function $F(\mathbf{s}) = (F_1(\mathbf{s}), F_2(\mathbf{s}))$, where each $F_i(\mathbf{s})$ is a real-valued function and $\mathbf{s} = (s_1, s_2)$ references the coordinates in $\Re^2$. Let C be a curve traced out by
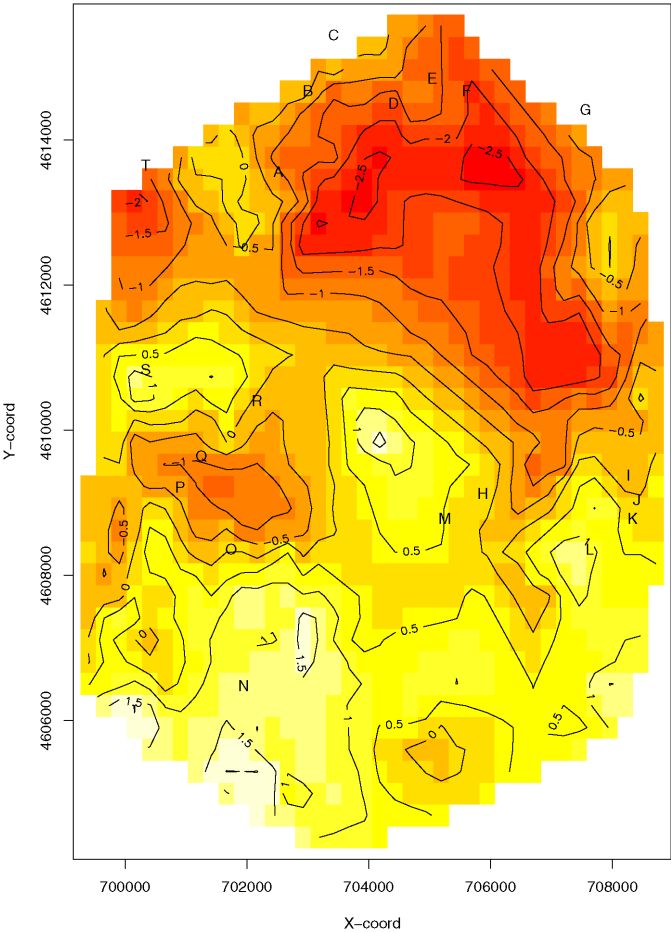

Figure 6. The Spatial Residual Surface From the Presence–Absence Application for the 603 Observed Locations (not shown) in Connecticut. Also shown are the 20 endpoints for the 10 curves in Figure 5 to connect the figures.

the parameter $t \in \Re^1$, so that $C = \{\mathbf{s}(t) : t \in \Re^1\}$. For the present context, assume that $C = C_1 \cup C_2$ is composed of two curves, C_1 and C_2, such that $C_1 = \{(t, g_1(t)) : a \le t \le b\}$ and $C_2 = \{(t, g_2(t)) : a \le t \le b\}$. Then the region enclosed by C is denoted by $\mathcal{D}$ and can be expressed as $\mathcal{D} = \{(s_1, s_2) : a \le s_1 \le b, g_1(s_1) \le s_2 \le g_2(s_1)\}$. Now consider the double integral over $\mathcal{D}$,

$$\int \int_{\mathcal{D}} \left(\frac{\partial F_1}{\partial s_1} + \frac{\partial F_2}{\partial s_2} \right) ds_1 \, ds_2$$

$$= \int \int_{\mathcal{D}} \frac{\partial F_1}{\partial s_1} \, ds_1 \, ds_2 + \int \int_{\mathcal{D}} \frac{\partial F_2}{\partial s_2} \, ds_1 \, ds_2.$$

Table 6. Curvilinear Gradient Assessment for the
10 Forest Edges Labeled in Figure 5 for the
Logistic Regression Example

Curve	Average gradient
AB	1.021 (.912, 1.116)
CD	.131 (−.031, .273)
EF	.037 (−.157, .207)
GH	1.538 (1.343, 1.707)
IJ	.586 (.136, .978)
KL	.036 (−.154, .202)
MN	.005 (−.021, .028)
OP	.227 (.087, .349)
QR	.282 (.118, .424)
ST	.070 (.017, .117)

Consider the second integral on the right side,

$$\int\int_{\mathcal{D}} \frac{\partial F_2}{\partial s_2} \, ds_1 \, ds_2$$

$$= \int_a^b \int_{g_1(s_1)}^{g_2(s_1)} \left(\frac{\partial F_2}{\partial s_2} \, ds_2 \right) ds_1$$

$$= \int_a^b \left(F_2(s_1, g_2(s_1)) - F_2(s_1, g_1(s_1)) \right) ds_1$$

$$= \int_a^b F_2(s_1, g_2(s_1)) \, ds_1 - \int_a^b F_2(s_1, g_1(s_1)) \, ds_1$$

$$= - \int_{C_2} F_2(s_1, g_2(s_1)) \, ds_1 - \int_{C_1} F_2(s_1, g_1(s_1)) \, ds_1$$

$$= - \oint_C F_2 \, ds_1.$$

Similarly,

$$\int\int_{\mathcal{D}} \frac{\partial F_1}{\partial s_1} \, ds_1 \, ds_2 = \oint_C F_1 \, ds_2.$$

Combining these, we obtain

$$\int\int_{\mathcal{D}} \left(\frac{\partial F_1}{\partial s_1} + \frac{\partial F_2}{\partial s_2} \right) ds_1 \, ds_2 = \oint_C F_1 \, ds_2 - F_2 \, ds_1. \qquad (A.1)$$

REFERENCES

Adler, R. J. (1981), *The Geometry of Random Fields*, Chichester, U.K.: Wiley.

Akima, H. (1996), "Algorithm 761: Scattered-Data Surface Fitting That Has the Accuracy of a Cubic Polynomial," *ACM Transactions on Mathematical Software*, 22, 362–371.

Allard, D., Gabriel, E., and Bacro, J. N. (2005), "Estimating and Testing Zones of Abrupt Changes for Spatial Data," research report, Unité de Biométrie, INRA.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004), *Hierarchical Modeling and Analysis for Spatial Data*, Boca Raton, FL: Chapman & Hall/CRC.

Banerjee, S., Gelfand, A. E., and Sirmans, C. F. (2003), "Directional Rates of Change Under Spatial Process Models," *Journal of the American Statistical Association*, 98, 946–954.

Barbujani, G., Jacquez, G. M., and Ligi, L. (1990), "Diversity of Some Gene Frequencies in European and Asian Populations, V: Steep Multilocus Clines," *American Journal of Human Genetics*, 47, 867–875.

Barbujani, G., Oden, N. L., and Sokal, R. R. (1989), "Detecting Areas of Abrupt Change in Maps of Biological Variables," *Systematic Zoology*, 38, 376–389.

Bocquet-Appel, J. P., and Bacro, J. N. (1994), "Generalized Wombling," *Systematic Biology*, 43, 442–448.

Caillol, H., Wojciech, P., and Hillion, A. (1997), "Estimation of Fuzzy Gaussian Mixture and Unsupervised Statistical Image Segmentation," *IEEE Transactions on Image Processing*, 6, 425–439.

Chellappa, R., and Jain, A. K. (1993), *Markov Random Fields*, New York: Academic Press.

Chilés, J. P., and Delfiner, P. (1999), *Geostatistics: Modelling Spatial Uncertainty*, New York: Wiley.

Cohen, E., Riesenfeld, R. F., and Elber, G. (2001), *Geometric Modeling With Splines: An Introduction*, Natick, MA: A.K. Peters.

Cressie, N. A. C. (1993), *Statistics for Spatial Data* (2nd ed.), New York: Wiley.

Csillag, F., and Kabos, S. (2002), "Wavelets, Boundaries and the Analysis of Landscape Pattern," *Ecoscience*, 9, 177–190.

Dass, S. C., and Nair, V. N. (2003), "Edge Detection, Spatial Smoothing, and Image Reconstruction With Partially Observed Multivariate Data," *Journal of the American Statistical Association*, 98, 77–89.

Fagan, W. F., Fortin, M. J., and Soykan, C. (2003), "Integrating Edge Detection and Dynamic Modeling in Quantitative Analyses of Ecological Boundaries," *BioScience*, 53, 730–738.

Fortin, M. J. (1994), "Edge Detection Algorithms for Two-Dimensional Ecological Data," *Ecology*, 75, 956–965.

——— (1997), "Effects of Data Types on Vegetation Boundary Delineation," *Canadian Journal of Forest Research*, 27, 1851–1858.

Fortin, M. J., and Drapeau, P. (1995), "Delineation of Ecological Boundaries: Comparisons of Approaches and Significance Tests," *Oikos*, 72, 323–332.

Frankel, T. (2003), *The Geometry of Physics: An Introduction*, Cambridge, U.K.: Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), Boca Raton, FL: Chapman & Hall/CRC.

Geman, S., and Geman, D. (1984), "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Jacquez, G. M., and Greiling, D. A. (2003), "Geographic Boundaries in Breast, Lung and Colorectal Cancers in Relation to Exposure to Air Toxins in Long Island, New York," *International Journal of Health Geographics*, 2, 4.

Jones, C. B. (1997), *Geographical Information Systems and Computer Cartography*, Harlow, Essex, U.K.: Addison-Wesley, Longman.

Lee, S., Wolberg, G., and Shin, S. Y. (1997), "Scattered Data Interpolation With Multilevel B-Splines," *IEEE Transactions on Visualization and Computer Graphics*, 3, 228–244.

Lu, H., and Carlin, B. P. (2006), "Bayesian Areal Wombling for Geographical Boundary Analysis," *Geographical Analysis*, to appear.

Mardia, K. V., Kent, J. T., Goodall, C. R., and Little, J. A. (1996), "Kriging and Splines With Derivative Information," *Biometrika*, 83, 207–221.

Mitas, L., and Mitasova, H. (1999), "Spatial Interpolation," in *Geographical Information Systems: Principles, Techniques, Management and Applications, GeoInformation International*, eds. P. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, New York: Wiley, pp. 481–492.

Møller, J. (Ed.) (2003), *Spatial Statistics and Computational Methods*, New York: Springer-Verlag.

Morel, J. M., and Solimini, S. (1994), *Variational Methods in Image Segmentation, With Seven Image Processing Experiments*, Boston: Birkhäuser.

Mumford, D., and Shah, J. (1989), "Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems," *Communications in Pure and Applied Mathematics*, 42, 577–685.

Phillips, G. M. (2003), *Interpolation and Approximation by Polynomials*, New York: Springer-Verlag.

Rudin, W. (1976), *Principles of Mathematical Analysis*, New York: McGraw-Hill.

Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, New York: Springer-Verlag.

Schabenberger, O., and Gotway, C. A. (2004), *Statistical Methods for Spatial Data Analysis*, Boca Raton, FL: Chapman & Hall/CRC.

Scheiner, S. M., and Gurevitch, J. (2001), *Design and Analysis of Ecological Experiments* (2nd ed.), Oxford, U.K.: Oxford University Press.

Stein, M. L. (1999), *Interpolation of Spatial Data: Some Theory of Kriging*, New York: Springer-Verlag.

Waller, L. A., and Gotway, C. A. (2004), *Applied Spatial Statistics for Public Health Data*, New York: Wiley-Interscience.

Walsh, D. C. I., and Raftery, A. E. (2002), "Accurate and Efficient Curve Detection in Images: The Importance Sampling Hough Transform," *Pattern Recognition*, 35, 1421–1431.

Webster, R., and Oliver, M. A. (2001), *Geostatistics for Environmental Scientists*, New York: Wiley.

Winkler, G. (2003), *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction*, New York: Springer-Verlag.

Womble, W. (1951), "Differential Systematics," *Science*, 114, 315–322.