Published in final edited form as:
Biometrics. 2012 September ; 68(3): 849–858. doi:10.1111/j.1541-0420.2011.01721.x.

A Geostatistical Approach to Large-Scale Disease Mapping with Temporal Misalignment

Lauren Hund[1,*], **Jarvis T. Chen**[2], **Nancy Krieger**[2], and **Brent A. Coull**[1]

[1]Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.

[2]Department of Society, Human Development, and Health, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.

Summary

Temporal boundary misalignment occurs when area boundaries shift across time (e.g., census tract boundaries change at each census year), complicating the modeling of temporal trends across space. Large area-level datasets with temporal boundary misalignment are becoming increasingly common in practice. The few existing approaches for temporally misaligned data do not account for correlation in spatial random effects over time. To overcome issues associated with temporal misalignment, we construct a geostatistical model for aggregate count data by assuming that an underlying continuous risk surface induces spatial correlation between areas. We implement the model within the framework of a generalized linear mixed model using radial basis splines. Using this approach, boundary misalignment becomes a nonissue. Additionally, this disease-mapping framework facilitates fast, easy model fitting by using a penalized quasilikelihood approximation to maximum likelihood estimation. We anticipate that the method will also be useful for large disease-mapping datasets for which fully Bayesian approaches are infeasible. We apply our method to assess socioeconomic trends in breast cancer incidence in Los Angeles between the periods 1988–1992 and 1998–2002.

Keywords

Area-level data; Cancer incidence; Disease mapping; Health disparities; Spatial misalignment; Spatial regression

1. Introduction

Area-level aggregated count data arise frequently in the disease-mapping setting (Best, Richardson, and Thompson, 2005; Wakefield, 2007); for instance, in this article, we assess the impact of socioeconomic disparities on breast cancer incidence by linking census data

*lhund@hsph.harvard.edu.

Supplementary Materials

Web Tables and Figures referenced in Sections 5.1 and 5.2 are available under the Paper Information link at the *Biometrics* website http://www.biometrics.tibs.org.

from multiple time points to cancer registry data. Our dataset is large (~2000 areas at each time point) and contains temporally misaligned boundaries, because census tract (CT) boundaries change over time. These types of data are becoming increasingly common in practice, due to our ability to merge census data and data from other large databases, such as disease registries.

Area-level data are most frequently modeled using hierarchical Bayesian models, with spatial correlation between areas incorporated through area-specific random effects having conditional Markov random field (MRF) priors (Besag, York, and Mollie, 1991). In the spatiotemporal setting where boundaries change over time, the use of area-specific random effects is not applicable because the areas are not well defined over the course of the study (Chen et al., 2008).

Most statistical approaches for handling misalignment in spatial datasettings address misalignment over multiple scales of space/geography(Gotway and Young, 2002). To our knowledge, few approaches to spatial regression exist that allow for temporal boundary misalignment. Mugglin, Carlin, and Gelfand (2000) and Zhu et al. (2000) address the boundary misalignment issue using hierarchical Bayesian models, with conditional MRF priors on the area-specific random effects. These existing methods are computationally intensive, and typically bog down for large datasets. Zhu et al. (2000) suggest including time- and area-specific random effects in the linear predictor of the model, implicitly assuming that the area-level random effects are independent across time points. Because this independence assumption is typically violated in standard longitudinal settings, the resulting inferences on changes over time can be inefficient.

In this article, we propose a geostatistical disease-mapping model that allows for spatially misaligned boundaries over time. We model the underlying spatial continuous risk surface as a Gaussian random field (Muller, Stadtmuller, and Tabnak, 1997; Best, Ickstadt, and Wolpert, 2000; Kelsall and Wakefield, 2002).We reduce the computational burden of spatial smoothing by modeling spatial correlation using bivariate low-rank, penalized splines (Kammann and Wand, 2003; Ruppert, Wand, and Carroll, 2003). Area-level data are sometimes treated as point referenced based on the centroid of an area, and the penalized-spline/mixed model approach is then used to model spatial variability, e.g., Lee and Durban (2009). However, these models also do not directly incorporate information about the size and shape of each area and can perform poorly (Best et al., 2005). By modeling the underlying spatial risk surface and aggregating to the area level, we overcome these limitations.

We implement the model within the generalized linear mixed model (GLMM) framework, modeling the underlying spatial surface using radial basis splines (Kammann and Wand, 2003; Ruppert et al., 2003), facilitating fitting a reduced-rank, computationally fast version of the model. We estimate model parameters using a penalized quasilikelihood (PQL) approximation to maximum likelihood estimation (Breslow and Clayton, 1993). Similar to the Kelsall and Wakefield model, our approach has the desirable property that smaller areas have larger prior variances. Additionally, the actual shape of each area, as opposed to only the neighborhood structure, is incorporated into the covariance between areas, avoiding any

problems that could arise from oddly shaped areas. Our method is easy to program in standard statistical software packages and is not computationally intensive relative to MRF formulations.

Section 2 introduces the motivating study for our methodology. Section 3 describes the formulation of the geostatistical disease-mapping model, and Section 4 presents spatiotemporal extensions of the model. Sections 5 and 6 give the results of a simulation study and data analysis, respectively.

2. Motivating Study: Breast Cancer Incidence in Los Angeles

Breast cancer is presently the leading cause of cancer among U.S. women (excluding nonmelanoma skin cancers), accounting for 28% of the diagnosed cases (American Cancer Society, 2010). Breast cancer typically has been portrayed as a disease of affluence. As secular changes in the socioeconomic distribution of breast cancer risk factors occur, incidence rates in poorer countries and among poorer women in more affluent countries may be catching up over the long term (Krieger et al., 2006). Examining the changes in the socioeconomic distribution of breast cancer incidence is important for public perception and policies regarding the disease, as well as to gauge the population mortality burden of breast cancer (Krieger et al., 2006). We investigate the hypothesis that the socioeconomic gradient in breast cancer incidence is decreasing over time by examining data associations between socioeconomic measures and breast cancer incidence rates across two decades.

We apply our method to assess changes in the socioeconomic gradient of breast cancer in women over time in Los Angeles County, California, focusing on the time periods 1988–1992 and 1998–2002, which precedes the change in breast cancer incidence rate attributed to declining use of hormone therapy. Krieger et al. (2006) originally analyzed these data by calculating age-standardized breast cancer incidence rates, stratified by decade, race/ethnicity, and socioeconomic status, and ignoring spatiotemporal correlation between areas. Our analysis parallels this original report, but incorporates spatiotemporal information into a regression model, yielding a more efficient analysis.

We quantify the socioeconomic gradient by calculating the difference in the breast cancer log-incidence rate ratios corresponding to an area-based socioeconomic measure (ABSM) for the time periods 1988–1992 and 1998–2002. We obtain total population counts of women by age and race/ethnicity and poverty indicators (ABSMs) from U.S. census data at the CT level in Los Angeles County for 1990 and 2000. There are a total of 1,642 CTs in 1990 and 2,056 CTs in 2000, reflecting a large number of CT boundary changes between the two time periods. We obtained the breast cancer case data from the Los Angeles Cancer Surveillance Program cancer registry. We appended the CT geocode to each cancer registry record, based on the location and date of residence at diagnosis. We link incident cases between 1988 and 1992 to the 1990 census population data and cases between 1998 and 2002 to the 2000 census data.

U.S. CT boundaries are redefined over time as necessary to maintain an average population between 3000 and 4000 in each CT, with each tract relatively socioeconomically homogeneous (U.S. Bureau of the Census 1994). Figure 1 illustrates different types of

changes in CT boundaries. Because changes do not always take the form of simple splitting or merging, there is not a one-to-one correspondence between CTs over time.

3. Statistical Framework: Spatial Model

Best et al. (2005) and Wakefield (2007) review standard spatial disease-mapping models. We use the following notation for our disease-mapping model. Observed cases of a disease Y_i within an area A_i are modeled using a Poisson likelihood, $Y_i \sim \text{Poisson}(e^{S_i} E_i)$, for $i = 1, \ldots, M$. We model S_i, the log-relative risk of disease, as a function of covariates and spatial random effects. Assuming disease prevalence varies within certain strata j (such as age groups), we calculate the expected number of cases in a region E_i using the prevalence of disease and the population count in each strata (i.e., using internal or external standardization). Our model assumes that the disease is rare and that the risk associated with living in area i acts proportionally on the baseline risks for each stratum.

We now develop a geostatistical model for spatial correlation that is similar in nature to Kelsall and Wakefield (2002), which we refer to as KW throughout this article. Consider an area A that is partitioned into regions $\{A_i\}$. Specifically, let i index CTs in region A, $i = 1, \ldots, M$, and s_{ij} be a point location within A_i, $s_{ij} \in A_i$. Define $|A_i|$ as the area of A_i; Y_i as the number of events in A_i; $\lambda(s)$ as the intensity of the Poisson process at point s; and $f_i(s)$ as the population density in A_i at point s. If we assume the population density is uniform over A_i, then $f_i(s) = 1/|A_i|$. If more information is available about the population density within an area A_i, we can use a piecewise uniform surface to estimate $f_i(s)$.

The diseased cases $Y(s)$ follow a Poisson process with intensity $E_i f_i(s) R(s)$, where $R(s)$ is the relative risk of disease at location s. Aggregating to the area level, $Y_i \sim \text{Poisson}\{E_i \int_{A_i} f_i(s) R(s) ds\}$, and the average relative risk in area A_i is $R_i = \int_{A_i} f_i(s) R(s) ds$. Disregarding spatial and covariate effects, $Y_i \sim \text{Poisson}(E_i R_i)$.

We incorporate covariates and spatial random effects through modeling the log-relative risk as $S(s) = \log R(s) = S'(s) + \beta X(s)$, where $X(s)$ is the covariate surface and $S'(s)$ is a continuous surface inducing spatial correlation between areas. KW propose a multivariate normal model for the area-level log-relative risk, with the covariance between the two areas interpreted as the average covariance between two points chosen randomly from the two areas. Diverging from KW, we model $S'(s)$ using a penalized spline term.

3.1 Approximating the Log-Relative Risk

We construct our model as a GLMM using radial splines to model spatial correlation (Ruppert et al., 2003). The underlying model for the log-relative risk is $S(s) = \beta X(s) + S'(s) = \beta X(s) + \sum_l Z_l(s) u_l$, where we write the spatial terms of the model $S'(s)$ as a penalized spline term. The basis functions $\{Z_l(s)\}$ are known, derived from a set of knots on the area A and a standard spatial covariance function (which we discuss in Section 3.2), and the $\{u_l\}$ terms are basis coefficients assumed to be independent normal random effects estimated via model fitting. By using this penalized spline representation, we express the model for the underlying relative risk as a GLMM. We use a quadrature approximation to estimate the spatial random effects for each area:

$$S_i = \int_{A_i} f_i(s) \left\{ X(s)\beta + \sum_l Z_l(s)u_l \right\} ds \approx X_i\beta + \sum_l \sum_j w_{ij} Z_l(s_{ij})u_l. \quad (1)$$

where $\{s_{ij}\}_{j=1,\ldots,d_i}$ are the d_i design points selected area A_i, and $\{w_{ij}\}$ are the corresponding quadrature weights for each area ($\sum_j w_{ij} = 1$, where $j = 1, \ldots, d_i$). If $\{X_i\}$ is an aggregate-level covariate, then $X_i = \int_{A_i} X(s) f_i(s) ds$ whereas if $\{X_i\}$ is inheritable, $X_i = X(s)\ \forall s \in A_i$. Conclusions drawn based on inheritable covariates are subject to ecological bias if we extrapolate area-level results to individuals (Wakefield, 2007).

To fit our model, appropriate design points $\{s_{ij}\}$ and corresponding quadrature weights $\{w_{ij}\}$ must be selected. If the design points correspond to subareas with known population counts (such as the centroids of block groups within CTs), quadrature weights could be chosen to reflect the underlying population density, $w_{ij} = N_{ij}/N_i$, where N_{ij} is the total population size in subarea A_{ij} and $N_i = \sum_j N_{ij}$ is the total population size in area A_i. Alternatively, assuming the population density is constant within an area, the best choice of design points corresponds to a grid of equally spaced points within each area; the resulting quadrature weights are $w_{ij} = 1/d_i$. With only one design point (the centroid) per area, our approach reduces to the model in which areas are treated as point-referenced data based on the centroid of the area, and a standard covariance function is specified to model spatial correlation between centroids.

3.2 Defining the Spatial Correlation Structure

We write the *underlying* model for the log-relative risk in mixed model form as $\boldsymbol{S}^* = \boldsymbol{X}^* \boldsymbol{\beta} + \boldsymbol{Z}^* \boldsymbol{u}$, where $\boldsymbol{S}^* = \{S_{ij}\}_{j=1,\ldots,d_i;\ i=1,\ldots,M}$; X^* is a matrix of covariates; $\boldsymbol{Z}^*$ is a contrast matrix described below; and $\boldsymbol{u} \sim \mathrm{MVN}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$. S_{ij} is the log-relative risk at design point j in area i. We construct $\boldsymbol{Z}^*$ such that $\mathrm{Cov}(S_{ij}, S_{i'j'}) = C(|s_{ij} - s_{i'j'}|; \rho)$, where C is a standard spatial covariance function that depends on the distance between the design points and parameters ρ. Due to the aggregate nature of the data, choice of a spatial covariance function is less important in this setting, as the model is not as sensitive to misspecification of the correlation function (see Section 5.2). We recommend using the exponential covariance function, $\mathrm{Cov}(S_{ij}, S_{i'j'}) = \sigma^2 \exp(-\rho|s_{ij} - s_{i'j'}|)$ for its simplicity. We choose a value for the range parameter ρ by selecting a plausible value based on the fact that $3/\rho$ is the approximate distance at which the correlation between S_{ij} and $S_{i'j'}$ is less than 0.05 (Banerjee, Carlin, and Gelfand, 2003). Alternatively, we could select ρ by choosing a value that minimizes the model deviance.

We fit a reduced rank approximation of the model by choosing a set of knots $\{\kappa_g\}_{g=1,\ldots,G}$ and basing our spatial correlation structure on the distances between the design points s_{ij} and the G knots (Kammann and Wand, 2003). When computationally feasible, we define the knots as the centroids of the areas in the study (G equals the number of areas in the study). A more practical approach is to use a knot selection algorithm to choose G knots in the study region, e.g., Johnson, Moore, and Ylvisaker (1990), which performs well in practice (Wand, 2003).

Define the $d_i \times G$ matrix $\boldsymbol{Z_i} = \{C(|s_{ij} - \kappa_1|), \ldots, C(|s_{ij} - \kappa_G|)\}_{j = 1, \ldots, d_i}$, which corresponds to the covariance between the design points in area i and the G knots. We stack the area-specific $\boldsymbol{Z_i}$ matrices to construct $\boldsymbol{Z} = (\boldsymbol{Z_i})_{i = 1, \ldots, M}$. Define the $G \times G$ matrix representing the covariance between the knots as $\boldsymbol{\Omega} = \{C(|\kappa_{g1} - \kappa_{g2}|)\}_{g1,g2 = 1, \ldots, G}$. Then, $\tilde{\boldsymbol{Z}^*} = \boldsymbol{Z\Omega^{-1/2}}$. From the definition of $\tilde{\boldsymbol{Z}^*}$, it follows that $\mathrm{Var}(\boldsymbol{S}^*) \approx \sigma_u^2 \tilde{\boldsymbol{Z}}^* \tilde{\boldsymbol{Z}}^{*\boldsymbol{T}} = \sigma_u^2 \boldsymbol{Z\Omega}^{-1} \boldsymbol{Z}^{\boldsymbol{T}}$.

Now, let S_i be the area-level log-relative risk for area A_i. For $\boldsymbol{S} = (S_1, S_2, \ldots, S_M)$, $\boldsymbol{S} = \boldsymbol{X\beta} + \boldsymbol{Z\tilde{u}}$, where $\boldsymbol{Z} = \boldsymbol{WZ}^*$, $\boldsymbol{X} = \boldsymbol{WX}^*$, and $\boldsymbol{W}$ is a $M \times \sum_{i=1}^{M} d_i$ block-diagonal matrix of the quadrature weights. Specifically, row i of $\boldsymbol{W}$ contains the d_i quadrature weights w_{ij} in the columns corresponding to area S_i and 0s everywhere else. Then, $\mathrm{Var}(\boldsymbol{S}) \approx \sigma_u^2 \tilde{\boldsymbol{Z}} \tilde{\boldsymbol{Z}}^{\boldsymbol{T}} = \sigma_u^2 \boldsymbol{W}\, \boldsymbol{Z\Omega}^{-1} \boldsymbol{Z}^{\boldsymbol{T}} \boldsymbol{W}^{\boldsymbol{T}}$. Because $\boldsymbol{S}^* \sim \mathrm{MVN}\{\boldsymbol{X}^*\boldsymbol{\beta}, \mathrm{Var}(\boldsymbol{S}^*)\}$ and $\boldsymbol{S}$ is a linear transformation of $\boldsymbol{S}^*$, it follows that $\boldsymbol{S} \sim \mathrm{MVN}\{\boldsymbol{X\beta}, \mathrm{Var}(\boldsymbol{S})\}$. Examining the covariance between individual areas clarifies that the covariance matrix in our model has the same interpretation as that in the KW model. The covariance between the log-relative risk for areas A_i and A_j is $\mathrm{Cov}(S_i, S_j) = \sum_{k=1}^{d_i} \sum_{l=1}^{d_j} w_{ik} w_{jl} \sigma_{kl} = \sum_{k=1}^{d_i} w_{ik} \sum_{l=1}^{d_j} w_{jl} \sigma_{kl}$, where σ_{kl} is the covariance between the log-relative risk at points s_{ik} and s_{jl}. That is, the covariance between two areas is a weighted average of the covariance between the design points in the area, and the variance of an area is the average covariance between the design points within an area.

3.3 Generalized Linear Mixed Model Construction

Using the above formulation of the log-relative risk within an area, we write the model as $Y_i \sim \mathrm{Poisson}(e^{S_i} E_i)$, where $\boldsymbol{S} = \boldsymbol{X\beta} + \boldsymbol{Z\tilde{u}}$. In this model, $\boldsymbol{\beta}$ are fixed effect parameters, $\boldsymbol{S} = (S_1, \ldots, S_M)^T$, $\boldsymbol{X} = (x_1, \ldots, x_M)^T$, $\mathbf{u} = (u_1, \ldots, u_G)^T \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ are independent random effects, and $\boldsymbol{Z}$ is the $M \times G$ matrix defined in Section 3.2.

We introduce an overdispersion parameter φ into the model to account for additional nonspatial variability in the data greater than that predicted by the Poisson distribution. We consider other methods for incorporating residual overdispersion and compare the performance of these various approaches in Web Appendix A. The likelihood for the model is:

$$L(\beta, \phi, \sigma^2; y_i) \propto (\sigma\sqrt{2\pi})^{-K} \int_{R^G} \times \exp\left\{ \sum_{i=1}^{M} \frac{1}{\phi}(-e^{\eta_i} + y_i \eta_i) + \sum_{i=1}^{G} - u_l^2 / 2\sigma^2 \right\} d\mathbf{u}.$$

The likelihood involves a G-dimensional integral, which is computationally expensive to evaluate. We approximate this integral using PQL (Breslow and Clayton, 1993). We have constructed an R package for fitting this model, available for download at http://www.hsph.harvard.edu/statinformatics/soft/areaglmm.html, and SAS code is available from the authors upon request.

In Section 5, we evaluate the performance of the PQL approximation for our model. Fitting our model using a Bayesian framework for the estimation of parameters is relatively

straightforward (Crainiceanu, Ruppert, and Wand, 2005), though much more computationally intensive.

3.4 Mapping the Relative Risk Surface

Constructing the smoothed predicted continuous relative risk surface or the smoothed predicted area-level relative risk surface is relatively straightforward. The pointwise relative risk estimates are $R(s) = \exp\{X(s)\hat{\beta} + \sum_l Z_l(s)\hat{u}_l\}$, and the area-level relative risk estimates are $R_i = \exp[\sum_j w_{ij}\{X(s_{ij})\hat{\beta} + \sum_l Z_l(s_{ij})\hat{u}_l\}]$, where the s_{ij} are design points in A_i with corresponding quadrature weights w_{ij}.

To obtain confidence bounds for the area-specific relative risk estimates, we estimate the standard errors for the fixed and random effects based on the PQL procedure (Ruppert et al., 2003). Specifically, $\mathrm{Cov}\{\begin{pmatrix}\hat{\beta}\\ \hat{\boldsymbol{u}}\end{pmatrix}|\boldsymbol{u}\} \simeq (\boldsymbol{C}^T\boldsymbol{V}\boldsymbol{C}+\sigma^2\boldsymbol{I})^{-1}\boldsymbol{C}^T\boldsymbol{V}\boldsymbol{C}(\boldsymbol{C}^T\boldsymbol{V}\boldsymbol{C}+\sigma^2\boldsymbol{I})^{-1}$, where $\boldsymbol{C} = (\boldsymbol{X}\boldsymbol{Z})$ and $\boldsymbol{V} = \mathrm{Var}(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{u}) = \boldsymbol{e}^{\boldsymbol{S}}\boldsymbol{E}$ and $\boldsymbol{E}$ is the expected count in each area. The standard error of the linear predictor $\boldsymbol{S}$ is $\sqrt{\boldsymbol{C}\mathrm{Cov}\{\begin{pmatrix}\hat{\beta}\\ \hat{\boldsymbol{u}}\end{pmatrix}|\boldsymbol{u}\}\boldsymbol{C}^T}$, which we can use to obtain pointwise or area-specific confidence intervals.

4. Spatiotemporal Extensions

Extending the model from the spatial to the spatiotemporal setting is straightforward using the spatial mapping methods of Wager, Coull, and Lange (2004). Assume we observe counts in M_t areas at time points $t = 1, \ldots, T$, where $Y_{it} \sim \mathrm{Poisson}(E_{it}R_{it})$ and $i = 1, \ldots, M_t$. For notational simplicity, we assume knot locations are the same across time points, though this assumption is not necessary.

We propose three different spatiotemporal models for the underlying log-relative risk. First, if the underlying risk surface is the same shape at each time point and shifts by a constant over time, then we model the log-relative risk as (*model 1*):

$$S_{ijt} = X_{ijt}\beta + \sum_l Z_l(s_{ijt})u_l + \delta_t,$$

where $u_l \sim N(0, \sigma^2)$ and δ_t is an intercept for time t. We note that $\delta_t = \delta_t$, or some variation of this, might be more appropriate in some applications. Model 1 assumes perfect correlation between spatial random effects across time.

Another option for modeling spatial structure is to fit a model analogous to Zhu et al. (2000), where we do not allow for a common spatial surface across time points (*model 2*):

$$S_{ijt} = X_{ijt}\beta + \delta_t + \sum_l Z_{lt}(s_{ijt})u_{lt},$$

where $u_{lt} \sim N(0, \sigma_t^2)$. In model 2, spatial random effects are independent across time points, ignoring temporal correlation in the spatial random effects. $Z_{lt}(\cdot)$ are time-specific spline

terms defined in the same way as in Section 3.2, except that the range parameter (or correlation structure) does not necessarily have to be equal across time points.

By adding an additional time-specific spatial random effect, we can fit a more flexible spatiotemporal model (*model 3*):

$$S_{ijt}=X_{ijt}\beta+\delta_t+\sum_l\{Z_l(s_{ijt})u_l+Z_{lt}(s_{ijt})u_{lt}\},$$

where $u_l \sim N(0, \sigma^2)$ and $u_{lt}\sim N(0,\sigma_t^2)$ are independent random effects. In model 3, the spatial relative risk surface differs at each time point due to the inclusion of the $\{u_{lt}\}$ random effect terms. The shared spatial surface represented by the $\{u_l\}$s, which are constant across time, induce temporal correlation in the random effects. Unless data are extremely sparse, we recommend using model 3 in practice.

The area-level model for the log-relative risk at time *t* is $\boldsymbol{S_t}=\boldsymbol{W_t}\boldsymbol{S_t^*}$, where $\boldsymbol{S_t^*}=\{S_{ijt}\}_{i=1,\ldots,d_i;j=1,\ldots,M_t}$. $\boldsymbol{W_t}$ is a *time-specific* quadrature weight matrix, identical to $\boldsymbol{W}$, but specific to time *t*. For instance, the area-level log-relative risk for model 3 is $\boldsymbol{S_t} = \boldsymbol{X_t\beta_t} + \boldsymbol{W_t}\tilde{\boldsymbol{Z}^*}(\boldsymbol{u} + \boldsymbol{u_t})$, where $\boldsymbol{u_t} = (u_{1t}, \ldots, u_{Gt})^T$ and $\boldsymbol{u_t}\sim \mathrm{MVN}(\boldsymbol{0}, \sigma_{ut}^2\boldsymbol{I})$. Then, $\mathrm{Var}(\boldsymbol{S_t})=\boldsymbol{W_t}\tilde{\boldsymbol{Z}}^*(\boldsymbol{\Omega_0^{-1}}+\boldsymbol{\Omega_t^{-1}})\tilde{\boldsymbol{Z}}^{*\boldsymbol{T}}\boldsymbol{W_t^T}$, similar to the variance of the log-relative risk at a single time point.

Within this geostatistical framework, boundary misalignment between areas over time no longer requires complicated model fitting schemes. The quadrature weight matrix $\boldsymbol{W_t}$ is different between time points when boundaries are misaligned, because design points may lie in different areas across time as boundaries change. To understand how our method accounts for boundary misalignment, it is useful to think of the locations of the design points and knots that induce the underlying risk surface as being fixed across time (though this is not necessary in model fitting). Because we model the underlying risk surface through these reference design points and knots, changing boundaries are no longer problematic.

5. Simulation Study

We conduct a simulation study to assess performance of the model. Goals of the simulation study include: (1) quantifying gains in power for detecting changes in a covariate effect over time when we account for temporal correlation in spatial random effects; (2) confirming that bias is negligible in the fixed effects and variance components when we use the PQL approximation with sparse data; and (3) examining sensitivity of the model to choice of the range parameter ρ and to the number of knots used.

5.1 Design of Simulation Study

We briefly describe the design of our simulation study, but relegate the specific details to Web Appendix B in the web-based supplementary material.

To construct our datasets, we start with a (0, 1) × (0, 1) regular grid of equal size areas, but relax this assumption shortly. We divide the grid into 64, 256, or 1024 square blocks (areas). We fix the disease incidence p at 0.11 cases per 100 person years and the total population in the area at 9.5 million, similar to our data application. We define the expected number of cases in an area as the product of the disease incidence and the total population in the area.

We simulate a Poisson process with intensity λ_{ijt} at location j in area i at time t, where the log intensity is (analogous to model 3 in Section 4) $\log(\lambda_{ijt}) = \log(E_{ijt}) + \beta_x x_{it} + \beta_t t + \beta_{xt} x_{it} t + \xi(s_{ijt}) + \xi_t(s_{ijt})$. E_{ijt} is the expected number of cases at time t in area i at point j. $\xi(s_{ijt})$ and $\xi_t(s_{ijt})$ are shared and time-specific spatial log-relative risks, respectively, at location s_{ijt}, a point in area i at location j at time t, $t = \{0, 1\}$. We generate $\xi(\cdot)$ and $\xi_t(\cdot)$ as realizations from a smooth Gaussian process with a Matern($\nu = 0.3$, $\kappa = 2$) correlation structure, where ν is a range parameter and κ is a smoothness parameter (Web Figure 1). We generate the shared surface between time points, $\xi(\cdot)$, to induce spatiotemporal correlation in the data. We generate an area-level covariate x_{it} from a uniform distribution, and are interested in the parameter β_{xt}, which represents the change in the effect of the covariate across time. The true value for this covariate in our study is $\beta_{xt} = -0.5$.

At each time point, we generate a realization from the continuous Poisson process with rate λ_{ijt}, and aggregate the cases over each area to obtain area-level case counts. We run 2000 simulations for each scenario described. We model the area-level expected count μ_{it} using model 3:

$$\log(\mu_{it}) = \log(E_{it}) + \beta_x x_{it} + \beta_t t + \beta_{xt} x_{it} t + \sum_j \sum_l Z_l(s_{ij})(u_l + u_{lt}),$$

where $E_{it} = N_{it} p$; N_{it} is the population size in area i at time t; and the penalized spline terms are identical to those defined in Sections 3.2 and 4.We simulate our data assuming no residual overdispersion (unless otherwise stated). However, we fit the quasi-Poisson model in our simulations and estimate an overdispersion parameter φ, to investigate whether identifiability problems arise between spatial variance and overdispersion parameters.

5.2 Results of Simulation Study

First, we examine the power and type I error associated with the test $H_0 : \beta_{xt} = 0$ for models 1, 2, and 3 for two different settings. In setting 1, we simulate data from a model with $\sigma = 0.3$, $\sigma_1 = \sigma_2 = 0.2$ and compare the fits of models 1, 2, and 3; in setting 2, we eliminate time-specific heterogeneity by simulating from a model with $\sigma = 0.3$, $\sigma_1 = \sigma_2 = 0$ and compare the fits of models 1 and 2.

Figures 2 and 3 display power curves for settings 1 and 2, respectively, for testing $H_0 : \beta_{xt} = 0$, when the data contain 64 and 256 areas. These figures illustrate that incorporating temporal correlation in the spatial random effects increases the power to detect differences in a covariate effect across time. The amount of power gained increases as the amount of spatial heterogeneity or temporal correlation in the spatial random effects increases (results not shown) and as the number of areas at each time point decreases.

Misspecifying the model by ignoring this temporal correlation can result incorrect inferences about the parameter β_{xt}, because we are examining the change in a covariate effect across time. In panel (a) in Figures 2 and 3, we see that the type I error deviates from 0.05 when the spatiotemporal correlation structure is misspecified. When temporal correlation is ignored (model 2), the type I error is less than 0.05, and the test is overly conservative. In setting 2, when time-specific heterogeneity is ignored, the type I error is greater than 0.05.

In our simulations, power gains and differences in type I error between the three models were negligible in the scenario with 1024 areas (results not shown). Existing models that handle temporal boundary misalignment will perform as well as our proposed model (in terms of the efficiency of β_{xt}). However, these existing approaches are fully Bayesian, and the computational efficiency of our frequentist parameter estimation framework is beneficial when the number of areas is large. Model fitting time can change from days (for alternative Bayesian models) to minutes (using the PQL approximation for parameter estimation).

Web Tables 2–4 in the web-based supplementary material show results of the simulation study assessing the sensitivity of the model to: (1) the sparseness of the data, (2) the choice of the range parameter ρ, and (3) the choice of the number of knots G.

In Web Table 2, we assess the performance of the PQL approximation when data are sparse, varying the expected area counts within an area. When testing $H_0 : \beta_{xt} = 0$, the type I error is near 0.05 and 95% Wald CI coverage is near 0.95 regardless of the expected area counts; variance components for the spatial random effects are also unbiased. We conclude that the PQL approximation performs well.

In Web Table 3, we examine sensitivity of the model to the choice of the range parameter ρ (assuming an exponential correlation structure and $\rho_t = \rho$). Zhang (2004) showed that ρ and σ^2 are not jointly identifiable in a spatial GLMM, and so misspecification of ρ leads to inconsistent estimates of σ, σ_1, and σ_2. Point estimates and standard errors of the fixed effects remain accurate, and type I error is near 0.05 when we misspecify the range parameter, insofar as the choice of the range parameter is reasonable (i.e., the average radius of an area $< 3/\rho <$ the maximum distance between areas).

Additionally, we note that the variance parameters σ^2, σ_1^2, and σ_2^2 are relatively unbiased when the range parameter is correctly specified (Web Table 3), suggesting that the model that allows for a common surface and time-specific spatial surfaces is identifiable with sufficient data. This result is consistent with Wager et al. (2004) and Coull, Ruppert, and Wand (2001), who fit similar models to model 3.

Further, we find that when data are generated from a smooth underlying surface, a model with as few as ~64 knots will perform as well as models with higher knot choices (Web Table 4). When the underlying spatial surface is less smooth, more knots are required to appropriately model the surface. For instance, if the range of the spatial correlation is less than the minimum distance between knots, oversmoothing of the spatial surface occurs. When we do not include enough knots in the model (for instance, the scenario with 64 areas

and $G < 64$), the data appear highly overdispersed ($\hat{\varphi} >> 1$) and the standard error of $\hat{\beta}_{xt}$ is underestimated.

For the scenarios with 64 and 256 areas, we repeat our simulations assessing sensitivity of the model to knot selection, choice of range parameter, and the number of design points using a nonregular, misaligned grid, shown in Web Figure 3. We exclude the 1024 case because we observed the greatest power differences and sensitivity to parameter choices for the 64 and 256 area scenarios; additionally, the setting with 1024 areas better approximates a regular grid. When estimating fixed effects and variance parameters, the model was not sensitive to the choice of the number of design points (results not shown).

In Web Table 5, we examine how well the model predicts the area-specific relative risks as a function of the number of knots included in the model, when the data are no longer on a regular grid. We present the average mean-squared error (defined in Web Appendix B), as well as estimates of β_{xt} and $se(\hat{\beta}_{xt})$. The results from the irregular grid are nearly identical to the regular grid, with a small inflation in the MSE when the data contain only 64 areas. The results from our simulations suggest that the model results and model validity do not change substantially, regardless of whether the data are misaligned over time.

Using data simulated on the irregular, misaligned grid, we evaluate how well our model performs when the range parameter changes across time, but we fit a model assuming that the range is constant over time (see Web Appendix C for detailed description of data generation). Under these conditions, we still obtain valid estimates of β_{xt} and $se(\hat{\beta}_{xt})$ (Web Table 6). Once again, correct specification of the range parameter is not important in obtaining valid model results, due to the lack of identifiability between the spatial variance parameters and the range parameter.

Lastly, the estimated overdispersion parameter $\hat{\varphi}$ in our simulations is often less than 1, suggesting that data are underdispersed. Specifically, data appear underdispersed when choice of G (number of knots) is high as well as when data are sparse (Web Tables 2 and 4). In this situation, unless there is a plausible reason for why the data are underdispersed, we recommend fixing $\varphi = 1$ or adjusting the number of knots such that $\hat{\varphi} \approx 1$, as suggested in Wager et al. (2004). Based on our simulations, when estimates of φ are less than 1, the model performance improves when we fix $\varphi = 1$ and do not estimate an overdispersion parameter (results not shown).

6. Analysis of the Los Angeles Breast Cancer Data

Using the spatiotemporal model described in Section 4, we reanalyze the Los Angeles cancer data presented in Krieger et al. (2006), restricting our attention to the time periods 1988–1992 and 1998–2002. Descriptive statistics for the Los Angeles cancer data are shown in Table 1. The number of CTs in Los Angeles county was 1,642 in 1990 and 2,056 in 2000; the total population count of women over 15 years old was 3,492,249 in 1990 and 3,625,360 in 2000. Following standard practice for cancer incidence rates centered around a census (Boyle and Parkin, 1991), we estimate person–time by assuming that the population counts are constant within each 5 year time period (1988–1992 and 1998–2002) and multiply the decennial population counts from the censuses by 5.

Age at diagnosis is categorized into 8 groups: 15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84, and 84+. Data are available for white non-Hispanic, Hispanic, black non-Hispanic, and Asian Pacific Islander populations. We use internal standardization to calculate the expected number of breast cancer cases by CT for each time period. We analyze the data combining all race/ethnicities (standardizing by age and race/ethnicity) and for each race/ethnicity group individually (standardizing only by age). Chen et al. (2008) emphasize that it may not be appropriate to assume a common spatial effect across racial/ethnic groups due to patterns of racial/ethnic segregation. We report results for all race/ethnicities combined and for the two largest subgroups, white non-Hispanics and Hispanics.

We select the percent of the population below the poverty level in a CT as our ABSM. Following Krieger et al. (2006), we model the relationship between the ABSM and the log-relative risks associated with predetermined epidemiologically meaningful poverty groups. Therefore, we model the percent of the population below poverty as a five-level categorical variable as follows: (i) among CTs with < 5% poverty, we distinguish between those with ≥ 10% high-income households (8.3% of CTs) and < 10% high-income households (9.2% of CTs); and (ii) among the remaining CTs, we distinguish between those with 5.0 − 9.9% (23.6% of CTs), 10.0 − 19.9% (26.6% of CTs), and ≥ 20% poverty (the federal definition of a poverty area and 32.4% of CTs). High-income households are defined as ≥ 4 times the U.S. median household income.

To model spatial variability, we define $\rho = 15/\Delta$ based on epidemiological plausibility, where Δ is the maximum distance between CTs in Los Angeles county. We use 30 design points per CT and select 100 knots throughout the study region using the space filling design described in Johnson et al. (1990) and implemented in the `R` package `FIELDS`.

Let Y_{it} denote the observed number of incident breast cancer cases in CT i at time t, and assume $Y_{it} \sim \text{Poisson}(\mu_{it})$. We fit the model (analogous to model 3):

$$\log(\mu_{it}) = \log(E_{it}) + \beta_0 + \boldsymbol{\beta^p pov}_{it} + \beta^t I_{t=2000} + \boldsymbol{\beta^{pt} pov}_{it} I_{t=2000} + \sum_j \sum_l Z_l(s_{ij})(u_l + u_{lt}),$$

where $\boldsymbol{pov}_{it}$ is a 4 × 1 indicator variable for the poverty category of area i at time t; and $\sum_j \sum_l Z_l(s_{ij})(u_l + u_{lt})$ is the spatiotemporal spline term, defined in model 3 in Section 4. For model identifiability, we use the >20% poverty category as the reference category.

For each race/ethnicity, fitting model 3 takes approximately 15 minutes using the `glmmPQL` function in `R` and 3 minutes using `PROC GLIMMIX` in `SAS`. Based on the results from model 3 in Table 2, the socioeconomic gradient in breast cancer does not appear to be decreasing over the time period studied. Instead, consistent with the findings in Krieger et al. (2006), we observed that the incidence rate ratio (IRR) remained stable over time in the different racial/ethnic groups, and that the socioeconomic gradient was smaller among the white non-Hispanic women (among whom the catch up may have already occurred), and greater among Hispanic women, for whom cancer risk factors may still exhibit strong socioeconomic patterning.

In Table 3, we examine the estimated spatial variance parameters and compare the results from models 2 and 3. The standard errors of the estimated fixed effects $\hat{\boldsymbol{\beta}}^{pt}$ from model 3 are consistently smaller than those in model 2 in all analyses. In our analysis, incorporating correlation between the spatial random effects across time results in a substantial increase in power. Specifically, in the combined racial/ethnic group analysis, we have stronger evidence that there exists a change in the socioeconomic gradient of breast cancer over time using model 3 ($p = 0.03$) versus model 2 ($p = 0.15$). In Figure 4, we plot the common residual spatial surface across both time points for all races combined, as well as the residual spatial surface from 1990, estimated using model 3 (note that we do not detect any residual spatial variability for the 2000 time point, as $\hat{\sigma_2} \approx 0$). The similarity between the spatial surfaces across time drives the gain in efficiency obtained when we use model 3.

7. Discussion

Motivated by the temporal boundary misalignment issues in the Los Angeles breast cancer incidence study, we develop an area-level disease-mapping model that incorporates spatiotemporal correlation in the presence of temporal boundary misalignment. Anyone using U.S. census data from more than one decade inevitably encounters the same temporal boundary misalignment issues that we face. Previous solutions to this problem are computationally intensive and ignore temporal correlation in the spatial random effects, potentially leading to inefficient inferences.

The proposed model does require selecting a parametric form for the correlation structure for the underlying continuous relative risk surface. Our simulation study suggests that the exponential correlation structure performs well and that the choice of the range parameter ρ is not too important. While fixing the range parameter may seem arbitrary, Zhang (2004) prove that, in spatial GLMMs, it is impossible to consistently estimate ρ and the variance parameter σ^2, but that the ratio σ^2/ρ is both more stable and more important to interpolation than the individual parameters. Therefore, fixing one parameter (ρ) and estimating the other (σ^2) should provide a consistent estimate of the spatial random effects.

While we have emphasized the usefulness of our model in addressing the temporal boundary misalignment problem, it is important to note that this new method will be a very useful and computationally efficient alternative to the popular fully Bayesian disease-mapping models for data collected at a single time point. Most disease-mapping applications in the literature use study regions containing only a few hundred areas, and fitting fully Bayesian models is feasible in such cases. For larger datasets with thousands of areas, which are becoming more common in epidemiological applications, these Bayesian models are more difficult to implement. By using a PQL approximation to maximum likelihood inference, we reduce the computation time from hours to minutes for our dataset and avoid any issues associated with model convergence and prior selection. Our method is also easy to program in standard software (SAS and R), filling a gap in the available software for fitting GLMMs with area-level spatial correlation.

Furthermore, while constructed in a different manner, the area-level spatial prior in our model has the same interpretation as that proposed in Kelsall and Wakefield (2002). Their

model is often cited in disease-mapping reviews as a good option for modeling area-level spatial correlation, as it seems appropriate to model the area-level relative risk as arising from a continuous underlying surface. However, we could not find any articles that use this method in practice, presumably due to the challenges associated with model fitting. We hope that the simplicity of our model will facilitate its use in practice.

In the present model for spatiotemporal variability, following Wager et al. (2004), we assume that the correlation between the log-relative risk at a given location at different time points is the same. When boundaries are aligned across time, this corresponds to the assumption that $\text{Corr}\{\log(\mu_{it_j}, \mu_{it_k})\} = c$, where c is a constant, for all time points j, k. When data are available at only two time points, this model is appropriate. When data are available at more than two time points, one might develop more sophisticated longitudinal extensions of this model that induce more correlation between occasions closer together in time. One viable option is using a model-based approach and incorporating spatiotemporal correlation through placing relevant priors on the random effects $\{u_{lt}\}$, such as $u_{lt} \sim N\{0, \Sigma(\rho_t)\}$. For instance, we could specify an AR(1) prior on $\{u_{lt}\}$. Fitting a frequentist version of this model in standard software is also of interest.

Using the Los Angeles County breast cancer incidence data, we find no clear evidence supporting the hypothesis that the socioeconomic gradient in breast cancer incidence is decreasing over time, consistent with the findings in Krieger et al. (2006). Results were robust to the choice of model parameters, including as the range parameter, number of knots, or the ABSM included in the regression model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by NIH grants 5P01CA134294, ES012044, ES07142, and the Department of Defense through the National Defense Science & Engineering Graduate Fellowship Program.

References

American Cancer Society. Cancer Facts and Figures 2010. Atlanta, Georgia: America Cancer Society; 2010.

Banerjee, S.; Carlin, BP.; Gelfand, AE. Hierarchical Modeling and Analysis for Spatial Data. Boca Raton, Florida: Chapman & Hall; 2003.

Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics. 1991; 43:1–20.

Best NG, Ickstadt K, Wolpert RL. Spatial Poisson regression for health and exposure data measured at disparate resolutions. Journal of the American Statistical Association. 2000; 95:1076–1088.

Best N, Richardson S, Thompson A. A comparison of Bayesian spatial models for disease mapping. Statistical Methods in Medical Research. 2005; 14:35–59. [PubMed: 15690999]

Boyle, P.; Parkin, D. Cancer Registration Principles and Methods. Lyon, France: IARC; 1991. Statistical methods for registries; p. 126-158.

Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. Journal of the American Statistical Association. 1993; 88:9–25.

Chen JT, Coull BA, Waterman PD, Schwartz J, Krieger N. Methodologic implications of social inequalities for analyzing health disparities in large spatiotemporal data sets: An example using breast cancer incidence data (Northern and Southern California, 1988–2002). Statistics in Medicine. 2008; 27:3957–3983. [PubMed: 18551507]

Coull BA, Ruppert D, Wand MP. Simple incorporation of interactions into additive models. Biometrics. 2001; 57:539–545. [PubMed: 11414581]

Crainiceanu C, Ruppert D, Wand M. Bayesian analysis for penalized spline regression using WinBUGS. Journal of Statistical Software. 2005; 14:1–24.

Gotway C, Young L. Combining incompatible spatial data. Journal of the American Statistical Association. 2002; 97:632–648.

Johnson M, Moore L, Ylvisaker D. Minimax and maximin distance designs. Journal of Statistical Planning and Inference. 1990; 26:131–148.

Kammann E, Wand M. Geoadditive models. Applied Statistics. 2003; 52:1–18.

Kelsall J, Wakefield J. Modeling spatial variation in disease risk: A geostatistical approach. Journal of the American Statistical Association. 2002; 97:692–701.

Krieger N, Chen JT, Waterman PD, Rehkopf DH, Yin R, Coull BA. Race/ethnicity and changing U.S. socioeconomic gradients in breast cancer incidence: California and Massachusetts, 1978–2002. Cancer Causes and Control. 2006; 17:217–226. [PubMed: 16425100]

Lee D-J, Durban M. Smooth-car mixed models for spatial count data. Computational Statistics and Data Analysis. 2009; 53:2968–2979.

Mugglin AS, Carlin BP, Gelfand AE. Fully model-based approaches for spatially misaligned data. Journal of the American Statistical Association. 2000; 95:877–887.

Muller H-G, Stadtmuller U, Tabnak F. Spatial smoothing of geographically aggregated data, with application to construction of incidence maps. Journal of the American Statistical Association. 1997; 92:61–71.

Ruppert, D.; Wand, M.; Carroll, R. Semiparametric Regression. New York: Cambridge University Press; 2003.

U.S. Bureau of the Census. Geographical areas reference manual. Washington, DC: U.S. Department of Commerce; 1994. (http://www.census.gov/geo/www/garm.html) [accessed: December 8, 2011]

US Census Bureau. Geographical areas reference manual. Technical Report. 1994.

Wager C, Coull B, Lange N. Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging. Journal of the Royal Statistical Society, Series B. 2004; 66:429–446.

Wakefield J. Disease mapping and spatial regression with count data. Biostatistics. 2007; 8:158–183. [PubMed: 16809429]

Wand M. Smoothing and mixed models. Computational Statistics. 2003; 18:223–249.

Zhang H. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. Journal of the American Statistical Association. 2004; 99:250–261.

Zhu L, Carlin B, English P, Scalf R. Hierarchical modeling of spatio-temporally misaligned data: Relating traffic density to pediatric asthma hospitalizations. Environmetrics. 2000; 11:43–61.

Figure 1.
Examples of changes to CT boundaries from one census to the next. This figure appears in color in the electronic version of this article.

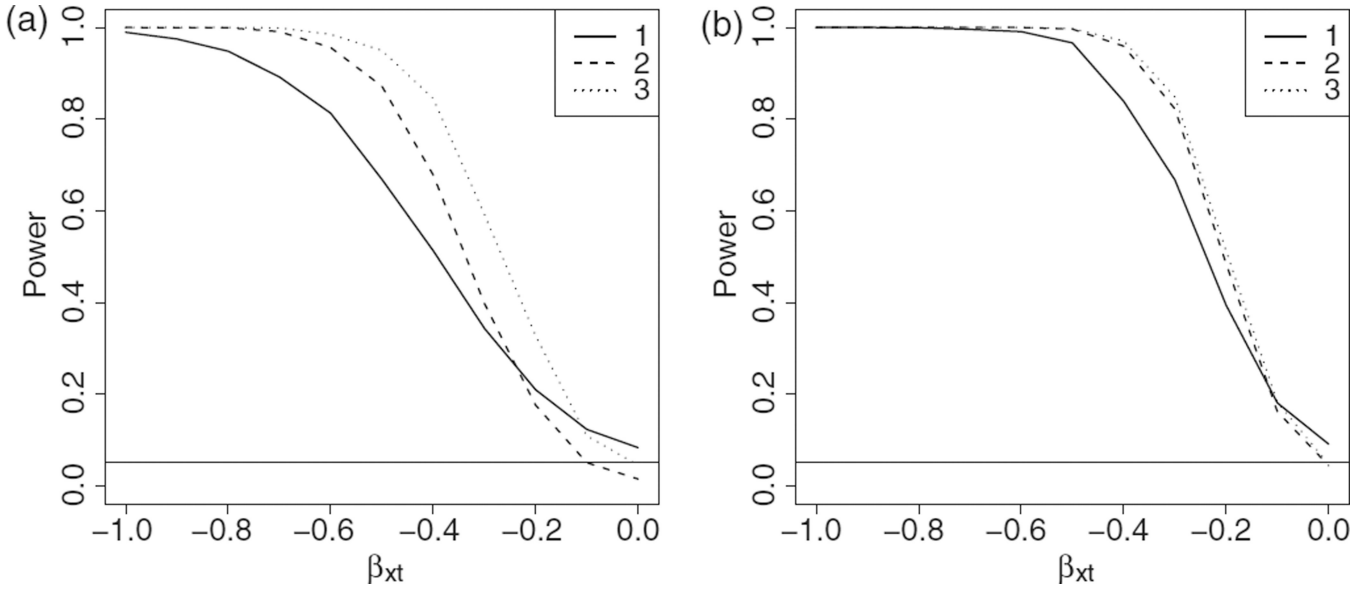


Figure 2.

Plot of the power for the test $H_0 : \beta_{xt} = 0$, $H_a : \beta_{xt} \neq 0$, at the $\alpha = 0.05$ level as a function of β_{xt}, when $\beta_x = 1$. (a) 64 and (b) 256 areas. $\sigma = 0.3$, $\sigma_1 = \sigma_2 = 0.2$. The lines are labeled according to the model that is being fit (e.g., 1 corresponds to model 1).

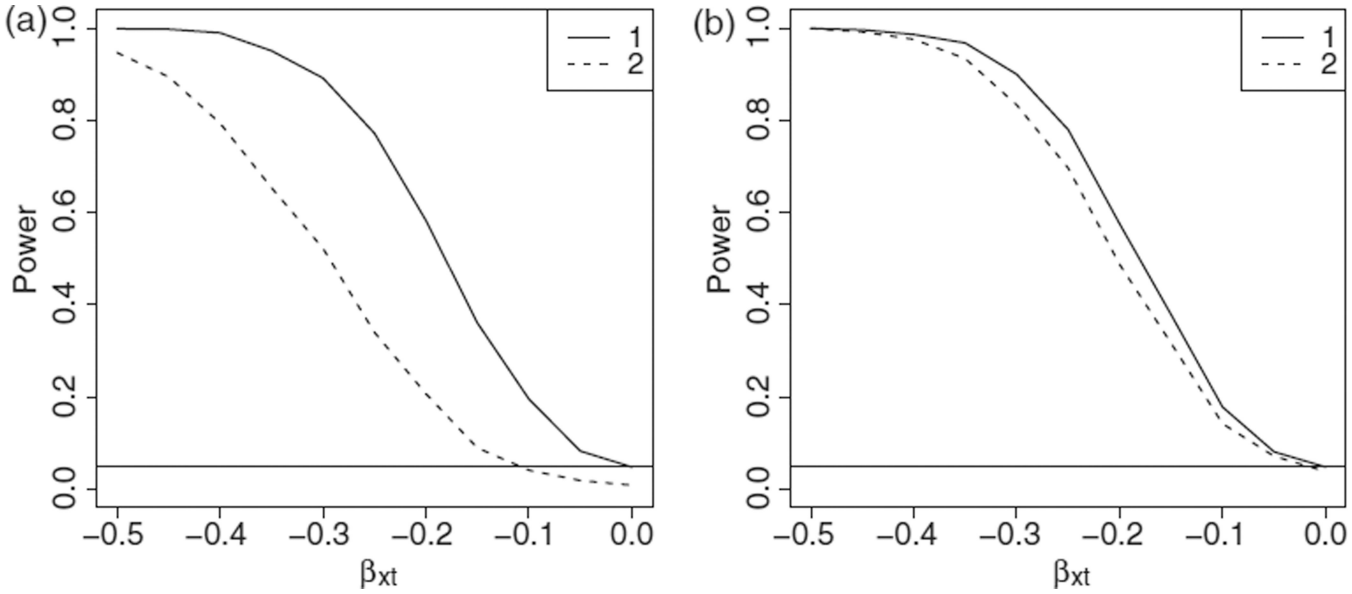
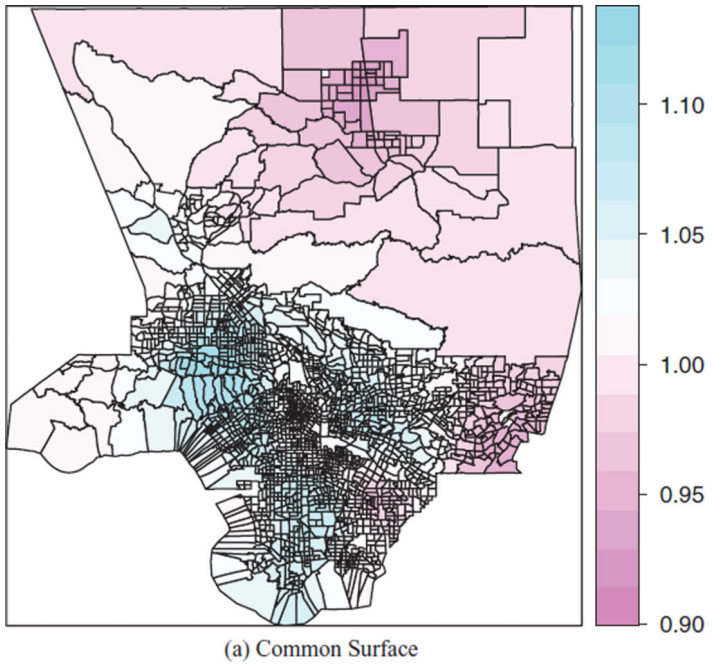
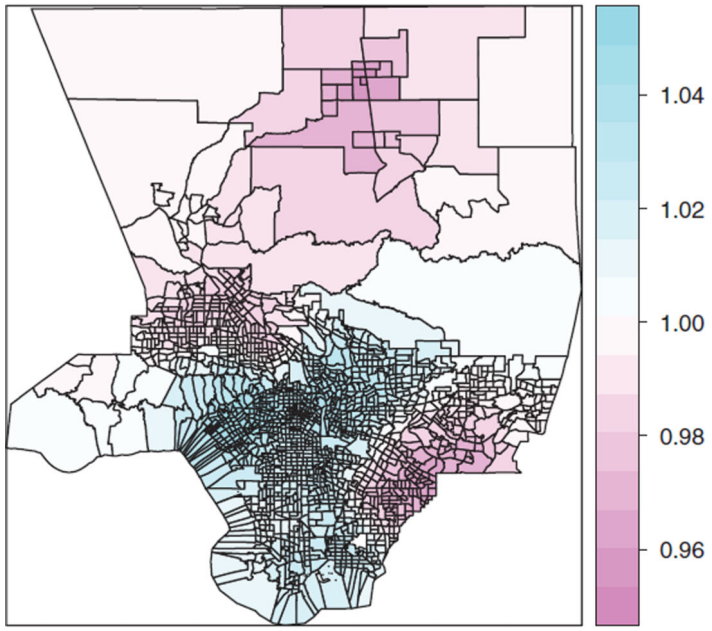


Figure 3.

Plot of the power for the test $H_0 : \beta_{xt} = 0$, $H_a : \beta_{xt} \neq 0$, at the $\alpha = 0.05$ level as a function of β_{xt}, when $\beta_x = 1$. (a) 64 and (b) 256 areas. $\sigma = 0.3$, $\sigma_1 = \sigma_2 = 0$. The lines are labeled according to the model that is being fit (e.g., 1 corresponds to model 1).



(a) Common Surface



(b) Time-specific surface for 1990

Figure 4.

Plot of the common spatial residual relative risk surface for the 1990 and 2000 time periods in Los Angeles county; and the additional time-specific spatial residual relative risk surface for the year 1990 in Los Angeles county. This figure appears in color in the electronic version of this article.

Table 1

Descriptive Statistics for Los Angeles Breast Cancer Data. Median (Interquartile Range) are presented

	Population size	Observed cases	Expected cases
All	9,150 (6,980, 12,005)	12 (7, 18)	12.2 (8.2, 17.4)
White non-Hispanic	3060 (710, 6061.3)	6 (1, 13)	7.0 (1.8, 13.6)
Hispanic	2742.5 (1120, 5180)	2 (1, 3)	1.7 (0.8, 3.1)

Table 2

Results from Los Angeles cancer data analysis. Incidence rate ratios relative to the >20% poverty category are shown for each time period and race/ethnicity group, with Wald p-values testing whether the log-IRR changes across time for each poverty category and for all categories combined

	Category	IRR 1988–1992	IRR 1998–2002	p-value
All	>20%	1	1	-
	10–20%	1.09 (1.05, 1.14)	1.09 (1.05, 1.13)	0.9829
	5–10%	1.13 (1.08, 1.18)	1.18 (1.14, 1.22)	0.1592
	<5% & <10% high inc.	1.15 (1.09, 1.20)	1.28 (1.22, 1.34)	0.0056
	<5% & ≥10% high inc.	1.25 (1.19, 1.30)	1.28 (1.23, 1.33)	0.5053
	Wald test, 4 df			0.0281
White	>20%	1	1	-
	10–20%	1.01 (0.93, 1.08)	0.992 (0.93, 1.06)	0.755
	5–10%	1.03 (0.96, 1.11)	1.073 (1.01, 1.14)	0.439
	<5% & <10% high inc.	1.07 (0.98, 1.15)	1.150 (1.07, 1.23)	0.178
	<5% & ≥10% high inc.	1.15 (1.07, 1.23)	1.190 (1.12, 1.26)	0.501
	Wald test, 4 df			0.2654
Hispanic	>20%	1	1	-
	10–20%	1.16 (1.07, 1.25)	1.23 (1.16, 1.31)	0.2752
	5–10%	1.33 (1.22, 1.44)	1.43 (1.34, 1.52)	0.2770
	<5% & <10% high inc.	1.38 (1.23, 1.53)	1.79 (1.64, 1.94)	0.0119
	<5% & ≥10% high inc.	1.63 (1.43, 1.83)	1.62 (1.43, 1.81)	0.9714
	Wald test, 4df			0.1419

Note: IRR = incidence rate ratio.

Table 3

Results from Los Angeles breast cancer data analysis, comparing spatial variance parameters and p-values testing $H_0 : \beta^{pt} = 0$ between model 2 (spatial random effects are independent across time) and model 3 (allows for temporal dependence in spatial random effects)

Race/ethnicity	Model	$\hat{\sigma}$	$\hat{\sigma_1}$	$\hat{\sigma_2}$	$\sqrt{\hat{\phi}}$	p-value
All	3	0.0870	0.0461	0.000	1.109	0.0281
	2	-	0.1008	0.0746	1.110	0.1497
White	3	0.1304	0.000	0.000	1.162	0.2654
	2	-	0.1572	0.0971	1.155	0.2904
Hispanic	3	0.1290	0.000	0.000	1.061	0.1419
	2	-	0.1411	0.0905	1.061	0.1906