



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Spatial+: a novel approach to spatial confounding

Citation for published version:

Dupont, E, Wood, SN & Augustin, N 2022, 'Spatial+: a novel approach to spatial confounding', *Biometrics*, vol. 78, no. 4, pp. 1279-1290. <https://doi.org/10.1111/biom.13656>

Digital Object Identifier (DOI):

[10.1111/biom.13656](https://doi.org/10.1111/biom.13656)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Biometrics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Spatial+: a novel approach to spatial confounding

Emiko Dupont^{1,*}, Simon N. Wood², and Nicole H. Augustin²

¹Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, U.K

²School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, U.K.

**email*: e.dupont@bath.ac.uk

SUMMARY: In spatial regression models, collinearity between covariates and spatial effects can lead to significant bias in effect estimates. This problem, known as spatial confounding, is encountered modeling forestry data to assess the effect of temperature on tree health. Reliable inference is difficult as results depend on whether or not spatial effects are included in the model. We propose a novel approach, spatial+, for dealing with spatial confounding when the covariate of interest is spatially dependent but not fully determined by spatial location. Using a thin plate spline model formulation we see that, in this case, the bias in covariate effect estimates is a direct result of spatial smoothing. Spatial+ reduces the sensitivity of the estimates to smoothing by replacing the covariates by their residuals after spatial dependence has been regressed away. Through asymptotic analysis we show that spatial+ avoids the bias problems of the spatial model. This is also demonstrated in a simulation study. Spatial+ is straightforward to implement using existing software and, as the response variable is the same as that of the spatial model, standard model selection criteria can be used for comparisons. A major advantage of the method is also that it extends to models with non-Gaussian response distributions. Finally, while our results are derived in a thin plate spline setting, the spatial+ methodology transfers easily to other spatial model formulations.

KEY WORDS: Bias reduction; Collinearity; Forestry; Partial thin plate spline regression; Spatial confounding.

1. Introduction

Regression models for spatially referenced data use spatial random effects to account for residual spatial correlation in the response data. As first noted by Clayton et al. (1993), these models can be problematic when estimation of individual covariate effects are of interest. So-called spatial confounding arises because spatial random effects may be correlated with spatially dependent covariates in the model and therefore interfere with their effect estimates. Reich et al. (2006) analyzed the issue using an example modeling the effect of socio-economic status on stomach cancer incidence in the municipalities of Slovenia. When spatial effects are added to the model, the covariate effect disappears, suggesting the spatial effects have taken over a disproportionate part of the explanatory power. While in this example, the spatial effects take the form of an Intrinsic Conditional Auto-Regressive (ICAR) random effect, spatial confounding is widely acknowledged as an issue that affects spatial models in general (see e.g. Hodges and Reich, 2010; Paciorek, 2010).

In this paper we model data from the Terrestrial Crown Condition Inventory (TCCI) forest health monitoring survey which has been carried out yearly since 1983 by the Forest Research Institute Baden-Württemberg. Crown defoliation (an indicator of poor tree health) has generally been worsening over time, and there is growing interest in understanding the effects of climate change in order to decide on forest management strategies for mitigation. Here, using a linear regression model, we consider the effect of temperature on crown defoliation. However, our results are highly dependent on whether or not we include spatial random effects in the model. As illustrated in Figure 1, in the null model (with no spatial effects), the estimated covariate effect is positive but not significant, whereas in the corresponding spatial model, the covariate effect is significant and the effect size more than triples. This behaviour suggests there is spatial confounding and makes reliable inference difficult.

[Figure 1 about here.]

A commonly used method for dealing with spatial confounding is restricted spatial regression (RSR), introduced by Reich et al. (2006) for the ICAR model, and further developed by Hanks et al. (2015) for continuous space models. In RSR the spatial random effects are restricted to the orthogonal complement of the covariates while keeping the overall column space of the model matrix in the regression unchanged. RSR directly eliminates collinearity and is designed to preserve the estimate of the null model while still accounting for residual spatial correlation. However, in the presence of unmeasured spatial confounders, the RSR estimate of the covariate effect may be significantly biased as it reflects not only the effect of the covariate but also that of the confounders (see e.g. Hanks et al., 2015; Khan and Calder, 2020). Here, we define a spatial confounder in the classical sense of an unmeasured spatial variable (causal or otherwise) that is associated with both the covariate and the response (Kirkwood and Sterne, 2010; McNamee, 2003).

Paciorek (2010) and Page et al. (2017) study the behaviour of the estimates in the spatial model when the covariate, like the response variable, has a spatial covariance structure. Intuitively, the model cannot distinguish the covariate from an unmeasured spatial effect, and the apportionment of effects between the covariate and spatial parts of the model may therefore be somewhat arbitrary. The analysis shows that the size of the resulting bias in the covariate effect estimate depends on the relative spatial scales of the covariate and spatial effects and, when the spatial scales agree, the bias is the same as that of RSR. Thus, while the estimate in the spatial model differs from RSR, it may be just as biased.

In many practical applications, however, the covariate of interest is spatially dependent but not fully determined by spatial location (*i.e. the covariate is not fully explained when a spatial model is fitted to it*). This form of the covariate is assumed in Thaden and Kneib (2018) who propose the geoaddivitive structural equations model (gSEM). Here, spatial dependence is regressed away from both the response and the covariates, and a regression involving

the residuals only is used to identify the original covariate effect. Simulations show that the bias in the covariate effect estimate of the spatial model is broadly removed using the gSEM. However, it is not immediately clear why the method works, and when the variables of interest are naturally spatially dependent, it seems undesirable to eliminate all spatial information from the modeling. The change in response variable also means that standard model selection criteria cannot be used for comparisons with the spatial and null models.

The structure of the covariate is usually not highlighted in the spatial confounding literature, but is important, as non-spatial information in a covariate can be used to distinguish it from the spatial effects without the need for considering differences in spatial scales. In this paper we show that, when a covariate is not fully spatial (*i.e. not fully explained when a spatial model is fitted to it*), unmeasured spatial confounders may still lead to significant bias in its effect estimate in the spatial model, however, the bias can be avoided in a relatively straightforward way. We propose a novel approach, spatial+, that is a simple modification of the spatial model in which the covariate is replaced by its residuals after spatial dependence has been regressed away. Similar to RSR, spatial+ retains the column space of the model matrix while reducing collinearity, but by adjusting the covariate rather than the spatial part of the model. Using asymptotic analysis as well as a simulation study we show that the estimates in spatial+ avoid the bias problems of the spatial model. We note that our asymptotic analysis applied to the gSEM estimates confirm the results of Thaden and Kneib (2018) and, for completeness, these derivations are included in Web Appendix D. An advantage of spatial+, however, is that all spatial information is retained in the model. Moreover, while the main properties of spatial+ are studied for models with a Gaussian response variable, we show that the method generalizes naturally to any response distribution from the exponential family of distributions.

Key to our analysis is that we formulate the spatial model as a partial thin plate spline

model. Here, spatial correlation is modeled by imposing a smoothing penalty on the spatial effects in the fitting process. We then see that the bias in the covariate effect estimate arises as a direct result of smoothing, and `spatial+` is a modification of the model matrix that makes the covariate part of the model less sensitive to this. Although our results are derived in the thin plate spline context, the methodology of `spatial+` can be directly applied to other commonly used spatial models including, for example, Gaussian Markov random field (GMRF) models and the (discrete space) ICAR model. In fact, it can be shown that modeling spatial random effects through the use of a smoothing penalty is equivalent to a Bayesian model formulation in which the spatial correlation structure is determined by a prior distribution. This equivalence is explained, for example, in Kimeldorf and Wahba (1970), Section 6.1 of Silverman (1985), pages 239-240 of Wood (2017) and Fahrmeir et al. (2004). Thus, while different spatial models correspond to different smoothing penalties, the underlying idea of reducing collinearity in this way to keep covariate effect estimates unaffected by spatial smoothing would apply in general.

This paper is structured as follows. In Section 2, we introduce the `spatial` and `spatial+` models that form the basis of our analysis. Section 3 summarizes our asymptotic analysis, details of which are in the supplementary web material. In Section 4, we illustrate our theoretical results in a simulation study which also compares `spatial+` with RSR and the `gSEM`. In Section 5, we demonstrate how `spatial+` can be implemented by applying it to our forestry example. Finally, in Section 6, we generalize the `spatial+` methodology to non-Gaussian response distributions and confirm that the method works in simulations for three different distributions.

2. Method

2.1 Spatial model

Our starting point is a spatial model formulated as a partial thin plate spline model (see e.g. Wahba, 1990) of the form

$$y_i = \beta x_i + f(\mathbf{t}_i) + \epsilon_i, \quad \epsilon_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2) \quad (1)$$

where $\mathbf{y} = (y_1, \dots, y_n)^T$ is the response, $\mathbf{x} = (x_1, \dots, x_n)^T$ an observed covariate with unknown effect β and f an unknown bounded function defined on an open bounded domain $\Omega \subset \mathbb{R}^d$ which includes the data locations $\mathbf{t}_1, \dots, \mathbf{t}_n$. The estimates $\hat{\beta}$ and \hat{f} in this model (known as the partial thin plate spline estimates of order $m > d/2$) are obtained as the minimizers of

$$\frac{1}{n} \sum_{i=1}^n (y_i - \beta x_i - f(\mathbf{t}_i))^2 + \lambda \sum_{i_1, \dots, i_m} \int_{\mathbb{R}^d} \left| \frac{\partial^m f(\mathbf{t})}{\partial t_{i_1} \dots \partial t_{i_m}} \right|^2 dt$$

where $\lambda > 0$ is an unknown smoothing parameter. Minimization here is over all $\beta \in \mathbb{R}$ and f is in a space of functions with square integrable partial derivatives of order m (for details see Duchon (1977)). The first term encourages fitted values that are close to the data while the second term induces smoothing by penalizing the wiggleness of the function f .

Duchon (1977) showed that the estimate of f can be obtained by estimating its coefficients in a basis known as the natural thin plate spline basis. This basis spans a finite-dimensional subspace in the space of functions defined on \mathbb{R}^d and has dimension $N = M + n$ where $M = \binom{m+d-1}{d}$. Using this basis, the partial thin plate spline estimates $\hat{\beta}$ and $\hat{\mathbf{f}} = (\hat{f}(\mathbf{t}_1), \dots, \hat{f}(\mathbf{t}_n))^T$ are the minimizers of

$$\|\mathbf{y} - \beta \mathbf{x} - \mathbf{f}\|^2 + n\lambda \mathbf{f}^T \mathbf{\Gamma} \mathbf{f} \quad (2)$$

with $\mathbf{\Gamma}$ an $n \times n$ diagonal penalty matrix. Solving the resulting normal equations, we see that

$$\hat{\beta} = (\mathbf{x}^T (\mathbf{I} - \mathbf{S}_\lambda) \mathbf{x})^{-1} \mathbf{x}^T (\mathbf{I} - \mathbf{S}_\lambda) \mathbf{y}, \quad \hat{\mathbf{f}} = \mathbf{S}_\lambda (\mathbf{y} - \hat{\beta} \mathbf{x}) \quad (3)$$

where $\mathbf{S}_\lambda = (\mathbf{I} + n\lambda\mathbf{\Gamma})^{-1}$ is known as the smoother matrix.

2.2 Spatial+ model

Starting with the model (1), in line with Rice (1986), we assume the covariate \mathbf{x} has the form

$$x_i = f^x(\mathbf{t}_i) + \epsilon_i^x, \quad \epsilon_i^x \stackrel{\text{iid}}{\sim} N(0, \sigma_x^2) \quad (4)$$

where f^x is a bounded smooth function defined on the spatial domain Ω . This means that \mathbf{x} is correlated with the smooth term f through the component f^x . Extending the two-stage smoothing spline model defined in Chen and Shiau (1991) to models of dimension $d \geq 1$, we define the spatial+ model as follows. Let $\hat{\mathbf{f}}^x = \mathbf{S}_{\lambda_x}\mathbf{x}$ and $\mathbf{r}^x = (\mathbf{I} - \mathbf{S}_{\lambda_x})\mathbf{x}$ be the fitted values and residuals in the thin plate spline regression (4) with smoothing parameter $\lambda_x > 0$. The spatial+ model is then the partial thin plate spline model

$$y_i = \beta r_i^x + f^+(\mathbf{t}_i) + \epsilon_i, \quad \epsilon_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2) \quad (5)$$

where β is the unknown effect of $\mathbf{r}^x = (r_1^x, \dots, r_n^x)^T$ and f^+ models the combined effect $f + \beta f^x$ in the original model (1). The spatial+ estimate $\hat{\beta}^+$ of β is its partial thin plate spline estimate in this model, i.e.

$$\begin{aligned} \hat{\beta}^+ &= (\mathbf{r}^{xT}(\mathbf{I} - \mathbf{S}_\lambda)\mathbf{r}^x)^{-1} \mathbf{r}^{xT}(\mathbf{I} - \mathbf{S}_\lambda)\mathbf{y} \\ &= (\mathbf{x}^T(\mathbf{I} - \mathbf{S}_{\lambda_x})(\mathbf{I} - \mathbf{S}_\lambda)(\mathbf{I} - \mathbf{S}_{\lambda_x})\mathbf{x})^{-1} \mathbf{x}^T(\mathbf{I} - \mathbf{S}_{\lambda_x})(\mathbf{I} - \mathbf{S}_\lambda)\mathbf{y}, \end{aligned}$$

and the spatial+ estimate $\hat{\mathbf{f}}^+$ of $\mathbf{f} = (f(\mathbf{t}_1), \dots, f(\mathbf{t}_n))^T$ is given by

$$\hat{\mathbf{f}}^+ = \mathbf{S}_\lambda(\mathbf{y} - \hat{\beta}^+\mathbf{x}) - \hat{\beta}^+\hat{\mathbf{f}}^x = \mathbf{S}_\lambda(\mathbf{y} - \hat{\beta}^+\mathbf{x}) - (\mathbf{I} - \mathbf{S}_\lambda)\mathbf{S}_{\lambda_x}\hat{\beta}^+\mathbf{x}.$$

2.3 Smoothness selection

Smoothing penalties introduce bias in estimates but reduce variance. The smoothing parameters λ and λ_x are usually estimated based on a separate smoothness selection criterion that balances this bias-variance trade-off.

For the analysis summarized in Section 3, in line with Rice (1986); Chen and Shiau (1991), we choose the value of the smoothing parameter that minimizes the average mean squared error (AMSE) of the estimated spatial effect. The AMSE for an estimated effect $\hat{\mathbf{f}} = (\hat{f}_1, \dots, \hat{f}_n)^T$ of $\mathbf{f} = (f(\mathbf{t}_1), \dots, f(\mathbf{t}_n))^T$ is defined as

$$\text{AMSE}(\hat{\mathbf{f}}) = \frac{1}{n} \sum_{i=1}^n \text{E} \left[(\hat{f}_i - f(\mathbf{t}_i))^2 \right] = B^2(f, \lambda) + V(f, \lambda)$$

where $B^2(f, \lambda) = \frac{1}{n} \sum_{i=1}^n (\text{E}(\hat{f}_i) - f(\mathbf{t}_i))^2$ and $V(f, \lambda) = \frac{1}{n} \sum_{i=1}^n \text{Var}(\hat{f}_i)$ are the average squared bias and the average variance, respectively. (Note that $\hat{\mathbf{f}}$ is a linear transformation of \mathbf{y} and therefore has a multivariate normal distribution).

For the simulations in Section 4 and the example in Section 5 we use the generalized cross validation (GCV) criterion, which is the default option in the R-package `mgcv` used for implementation. Asymptotically (as the sample size $n \rightarrow \infty$), GCV selects the optimal smoothing parameter for minimizing prediction error. Thus, GCV is not dissimilar to the criterion used for the theoretical derivations. Indeed, Chen and Shiau (1994) show that their asymptotic results in Chen and Shiau (1991) for one-dimensional models also hold for GCV and Mallows' C_L . In practice, the restricted maximum likelihood (REML) criterion is often used instead of GCV as, for finite samples, GCV usually has more uncertain estimates than REML and tends to undersmooth (i.e. overfit) the data (Wood (2017) p. 266-267). Repeating the simulations and the data example using REML gave similar results to GCV.

3. Asymptotic results

In the supplementary web material we analyze the asymptotic behaviour of effect estimates in the models defined in Sections 2.1 and 2.2 as the sample size $n \rightarrow \infty$. Without the smoothing penalty, the spatial model (1) is an ordinary linear model in which the estimates are unbiased. Therefore, bias in the covariate effect estimate $\hat{\beta}$ arises as a direct result of smoothing. In fact, for partial spline models (i.e. models where the domain of the spline, here the spatial domain, is one-dimensional) Rice (1986) identified this smoothing-induced bias and showed that it

can become disproportionately large unless the data is undersmoothed. More specifically, Rice's results show that if the smoothing parameter λ converges at the optimal rate (for minimizing the AMSE of the estimated smooth effect), it cannot be ensured that the bias of $\widehat{\beta}$ converges faster than its standard deviation. Spatial+ is a higher-dimensional version of a model introduced by Chen and Shiau (1991) to overcome this type of bias in dimension one.

Rice, Chen and Shiau use the Demmler-Reinsch basis for natural splines to diagonalize the smoother matrix \mathbf{S}_λ , enabling them to explicitly study the asymptotic behaviour of the estimates in one-dimensional models. Due to results by Utreras (1988), we are able to extend these derivations to models of arbitrary spatial dimension. The main results of our analysis are provided in Web Appendix C. We confirm that, as is the case in dimension one, when smoothing is chosen at an optimal rate for minimizing the AMSE of the estimated spatial effect, the bias in the covariate effect estimate in the spatial model can become disproportionately large. In contrast, in the spatial+ model, the bias converges to 0 strictly faster than the standard deviation. Our results are based on a number of technical lemmas and assumptions, details of which are provided in Web Appendices A and B.

4. Simulation

Partial thin plate spline models can be implemented in the R-package `mgcv` using the computationally efficient reduced rank approximation known as thin plate regression splines. We use this implementation (with GCV as the smoothness selection criterion) to compare the results of models fitted to simulated data for which we know the true underlying covariate and spatial dependence.

4.1 Data

We generate 100 independent replicates of covariate data $\mathbf{x} = (x_1, \dots, x_n)^T$ and response data $\mathbf{y} = (y_1, \dots, y_n)^T$, observed at $n = 1000$ randomly selected locations in the spatial domain $[0, 10] \times [0, 10]$ in \mathbb{R}^2 (using a 50×50 grid), as follows. Let $\mathbf{z} = (z_1, \dots, z_n)^T$ and

$\mathbf{z}' = (z'_1, \dots, z'_n)^T$ denote observations at the selected locations of independently generated Gaussian spatial fields with an exponential and a spherical covariance structure, respectively. That is, each spatial field is sampled from a multivariate normal distribution centered at $\mathbf{0}$ with covariance structure defined by $C(h) = \exp(-(h/R)^p)$ with $R = 5$ and $p = 1$ for the exponential field and $C(h) = -1 - 1.5h/R + 0.5(h/R)^3$ for $h \leq R$, $C(h) = 0$ for $h > R$ with $R = 1$ for the spherical field (where h denotes Euclidean distance). To ensure that the fields lie in the span of the spatial basis vectors used for the models in Section 4.2, each is replaced by the fitted values of a spatial thin plate regression spline fitted to them. We then let

$$\begin{aligned}\mathbf{x} &= 0.5\mathbf{z} + \boldsymbol{\epsilon}^x \text{ where } \boldsymbol{\epsilon}^x \sim N(\mathbf{0}, \sigma_x^2\mathbf{I}), \\ \mathbf{y} &= \beta\mathbf{x} + \mathbf{f} + \boldsymbol{\epsilon}^y \text{ where } \boldsymbol{\epsilon}^y \sim N(\mathbf{0}, \sigma_y^2\mathbf{I}),\end{aligned}$$

with true covariate effect $\beta = 3$, true residual spatial effect $\mathbf{f} = -\mathbf{z} - \mathbf{z}'$ and $\sigma_y = 1$, $\sigma_x = 0.1$. Thus, \mathbf{f} is directly correlated with the spatial pattern $0.5\mathbf{z}$ of the covariate. This approach is similar to Thaden and Kneib (2018), except we have added the component $-\mathbf{z}'$ so that \mathbf{f} could represent, for example, the combined effect of an unobserved covariate (with a similar spatial pattern to that of \mathbf{x}) as well as an independent short-range spatial process. Also, rather than treating the spatial fields as fixed, we generate new fields for each replicate in the simulation. This slight change in approach was chosen to show that results do not rely on any particular replicates of the spatial fields. We note that similar results were obtained when we repeated the simulations for a number of fixed spatial fields. Finally, we have chosen σ_x relatively small (such that the model matrix for the spatial model has nearly collinear columns) and σ_y relatively large (to encourage smoothing). This is the situation in which we would expect spatial confounding issues to arise which is also confirmed by the simulations in Thaden and Kneib (2018).

4.2 Models

To each replicate of simulated response data \mathbf{y} and covariate data \mathbf{x} , we fit the following models (with basis size $k = 300$ for the thin plate regression splines). Models 2 - 5 are fitted twice: once with smoothing penalties applied (i.e. where λ and λ_x are estimated from the data) and once without (i.e. where $\lambda = \lambda_x = 0$). In `mgcv`, smoothing penalties are applied by default but can be removed using the option `fx=TRUE`. This option means that the smoothing parameter is fixed rather than estimated, defaulting to 0 (i.e. no smoothing) when no value is specified.

1. Null model: The model with no spatial effects given by

$$y_i = \beta x_i + \epsilon_i, \quad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (6)$$

where β and σ^2 are estimated parameters.

2. Spatial model: The model given by

$$y_i = \beta x_i + f(\mathbf{t}_i) + \epsilon_i, \quad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (7)$$

where β and σ^2 are estimated parameters and f a thin plate regression spline with $\mathbf{t}_1, \dots, \mathbf{t}_n$ the observed data locations.

3. RSR model: Let \mathbf{B}_{sp} be the matrix whose columns are the spatial basis vectors in the model matrix from (7) (i.e. the thin plate regression spline basis functions evaluated at the data locations) and let $\tilde{\mathbf{B}}_{\text{sp}} = (\mathbf{I} - \mathbf{x}(\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T) \mathbf{B}_{\text{sp}}$ be the projection of this onto the orthogonal complement of \mathbf{x} . The RSR model is given by

$$y_i = \beta x_i + \tilde{f}_i + \epsilon_i, \quad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (8)$$

where β and σ^2 are estimated parameters and $\tilde{\mathbf{f}} = (\tilde{f}_1, \dots, \tilde{f}_n)^T$ is modeled the same way as the spatial effect in (7) but with \mathbf{B}_{sp} replaced by $\tilde{\mathbf{B}}_{\text{sp}}$ in the model matrix.

4. gSEM: Let $\mathbf{r}^x = (r_1^x, \dots, r_n^x)^T$ and $\mathbf{r}^y = (r_1^y, \dots, r_n^y)^T$ denote the spatial residuals of \mathbf{x} and

\mathbf{y} , that is, $\mathbf{r}^x = \mathbf{x} - \widehat{\mathbf{f}}^x$ where $\widehat{\mathbf{f}}^x$ are the fitted values in the regression

$$x_i = f^x(\mathbf{t}_i) + \epsilon_i^x, \quad \epsilon^x \sim N(\mathbf{0}, \sigma_x^2 \mathbf{I}) \quad (9)$$

where σ_x^2 is estimated and f^x a thin plate regression spline, and \mathbf{r}^y is the same but replacing \mathbf{x} by \mathbf{y} . The gSEM model is then the linear model given by

$$r_i^y = \beta r_i^x + \epsilon_i, \quad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (10)$$

where β and σ^2 are estimated.

5. Spatial+: Let \mathbf{r}^x denote the spatial residuals of \mathbf{x} as above. The spatial+ model is then

$$y_i = \beta r_i^x + f^+(\mathbf{t}_i) + \epsilon_i, \quad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (11)$$

where β and σ^2 are estimated parameters and f^+ a thin plate regression spline.

4.3 Results

The results of the simulation are summarized in Figure 2. For each data replicate, the output is the estimated covariate effect and the mean squared error (MSE) of fitted values for each model fit. For ease of notation, in this section, we use $\widehat{\beta}$ to mean the estimated covariate effect in any of the fitted models (rather than the partial thin plate spline estimate alone). The MSE of fitted values is calculated as $\|\widehat{\mathbf{y}} - (\beta \mathbf{x} + \mathbf{f})\|^2$ where for models 1, 2, 3 and 5, $\widehat{\mathbf{y}}$ is the fitted values in the regressions (6), (7), (8) and (11), respectively, and for model 4, $\widehat{\mathbf{y}} = \widehat{\mathbf{f}}^y + \widehat{\mathbf{r}}^y$ where $\widehat{\mathbf{f}}^y$ and $\widehat{\mathbf{r}}^y$ are the fitted values in the regressions (9) and (10). Here $\beta = 3$ and $\mathbf{f} = -\mathbf{z} - \mathbf{z}'$ are the true values of the estimated effects with $\beta \mathbf{x} + \mathbf{f}$ the true mean of \mathbf{y} .

[Figure 2 about here.]

In the null model and the RSR model, the estimated covariate effect is the same (that is, for any given data replicate, the value of $\widehat{\beta}$ is identical) and has a noticeably larger bias than the estimates in the other models. This is expected as for these models, $\widehat{\beta}$ is the ordinary least squares estimate, which, in addition to the true effect β , includes a contribution from the part of \mathbf{f} that is correlated with \mathbf{x} . More specifically, the bias in $\widehat{\beta}$ is given by $E((\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{f})$

(see Web Appendix E). Note that in our simulations, since \mathbf{x} and \mathbf{f} are negatively correlated, the bias is negative, however, if the correlation had been positive, the bias would have been positive. The fitted values in RSR, however, differ from those of the null model as the larger model matrix explains a part of \mathbf{y} that is treated as random noise in the null model. In fact, the column space of the model matrix of the RSR model is the same as that of the spatial model, and it is therefore not surprising that, for any given data replicate, the fitted values in these two models are similar.

If no smoothing penalty is applied, models 2, 4 and 5 are essentially the same: for any given data replicate, they have the same fitted values and the same unbiased estimate for the covariate effect. This illustrates that spatial confounding bias is due to the combined effect of collinearity and smoothing, rather than collinearity alone. The spatial model is, in this case, an ordinary linear model where the columns in the model matrix are the covariate \mathbf{x} and the spatial basis vectors \mathbf{B}_{sp} . This is the model from which the data are generated and, therefore, it is not surprising that the spatial model is able to recapture the true effects. The spatial+ model is a reparametrization of the spatial model which preserves the overall column space, and simple linear model theory shows that the covariate effect estimate is preserved. Similarly, it is straightforward to show that the gSEM covariate effect estimate agrees with the spatial model in this case. (Derivations are included in Web Appendix E).

In the unsmoothed versions of models 2 - 5, for any given data replicate, the fitted values are the same across all models. When smoothing is applied, the MSE of fitted values reduces, indeed, this is the intended purpose of the smoothing penalty. Looking at the covariate effect estimate, in the RSR model, the (biased) ordinary least squares estimate is unaffected by smoothing. For the remaining three models, while the unsmoothed versions of the models give unbiased estimates of β , we see that smoothing introduces varying degrees of bias. In the spatial model, the bias is quite large illustrating our results from Section 3. In contrast,

while the covariate effect estimate is no longer the same in the gSEM and the spatial+ model, for both models, the bias is still negligible. This behaviour is therefore also consistent with what we would expect from our theoretical results.

4.4 Additional comments

Our analysis in Section 4.3 gives some intuition for why spatial+ works. If no smoothing penalty is applied we saw that for any given data replicate, spatial+ has the same unbiased estimate for the covariate effect as the spatial model. In fact, any decomposition $\mathbf{x} = \mathbf{v} + \mathbf{r}$ with \mathbf{v} in the column space of the spatial basis vectors \mathbf{B}_{sp} gives a reparametrization (replacing \mathbf{x} by \mathbf{r}) in which \mathbf{r} captures the original covariate effect (since $\beta\mathbf{x} = \beta\mathbf{v} + \beta\mathbf{r}$). However, by choosing \mathbf{r} to be broadly orthogonal to the column space of \mathbf{B}_{sp} (as it is in the spatial+ model), the estimates of the covariate and spatial effects are broadly decorrelated. Thus, the covariate effect estimate is largely unaffected when smoothing is applied to the spatial term and thereby remains broadly unbiased.

Although the asymptotic results in Section 3 are technically only expected to hold for large sample sizes n , the above intuition of spatial+ applies in general. In order to investigate the behaviour at moderate sample sizes, we repeated the simulations for $n = 300$, $n = 150$ and $n = 50$ (with spatial basis sizes $k_{\text{sp}} = 100$, $k_{\text{sp}} = 100$ and $k_{\text{sp}} = 30$, respectively). The results are included in Web Appendix F. We see that the smaller the sample size, the larger the variability of the estimate $\hat{\beta}$, particularly for the unsmoothed spatial model, gSEM and spatial+ model as well as the smoothed gSEM and spatial+ model. However, overall, the behaviour of the bias looks broadly similar to the results for the larger sample size and is in line with the asymptotic results. We note that, when the spatial basis size k_{sp} is large compared to n , the MSE of fitted values in the unsmoothed models becomes relatively high. The thin plate regression spline basis is generally ordered with lower frequency spatial patterns first so that adding more spatial basis functions increases the ability of the spatial

effect to model more complex spatial variation involving both lower and higher frequency spatial patterns. In practice, while a small basis size is preferable for reducing computation time for model fitting, a large basis may be necessary in order to capture the spatial variation in the data. However, when the flexibility of the spatial effect is increased, the model is also more likely to overfit the data. The purpose of smoothing is exactly to avoid this overfitting and reduce the effective degrees of freedom in the model.

Finally, our results assume that the true spatial dependence of both the response variable and the covariate can be described by thin plate splines. (In our simulations this was ensured by fitting thin plate splines to the spatial fields \mathbf{z} and \mathbf{z}' used for the data generation.) If this is not the case, there may be additional bias caused by model mis-specification. In practical terms, the thin plate spline formulation assumes that spatial dependence is smooth and isotropic, but when these conditions hold it provides a fairly flexible way of modeling spatial effects. Thus, when we repeated our simulations using data based on spatial fields that were fitted with Gaussian process smooths rather than thin plate splines, the results were very similar. (See Web Appendix F).

5. Application

We illustrate how the spatial+ model can be used in practice by applying it to our forestry example. Details of the data can be found in Augustin et al. (2009); Eichhorn et al. (2017). We consider here the data for spruce for a single observation year, namely, 2013 which has measurements from $n = 186$ locations. We are interested in assessing the effect of the climate variable `tminmay` (minimum temperature in May) on the response variable `ratio` (crown defoliation expressed as a proportion). We expect a high minimum temperature in May to be indicative of a warmer and drier year in general which, in turn, is likely to lead to higher levels of tree defoliation (measured later in summer). We also expect older trees to have significantly more defoliation than younger trees and have therefore included the

variable `age` (age of trees) as an additional covariate in the models. Scatterplots of the data (not shown here) indicate the relationships between the covariates and the response variable are broadly as expected.

5.1 Models

A natural starting point is the null model

$$\mathbf{ratio}_i = \alpha + \beta_1 \mathbf{age}_i + \beta_2 \mathbf{tminmay}_i + \epsilon_i, \quad (12)$$

where $\epsilon_i \sim N(0, \sigma^2)$ is iid noise and α, β_1, β_2 and σ are estimated parameters. However, numerous spatially dependent predictors have not been included in the model, for example, soil characteristics such as soil depth and base saturation; other climatic variables such as those related to radiation and precipitation; water budget of the trees etc. Therefore, we would expect residual spatial correlation in the response variable, and a more appropriate model may therefore be a spatial model, which we define as

$$\mathbf{ratio}_i = \alpha + \beta_1 \mathbf{age}_i + \beta_2 \mathbf{tminmay}_i + f(\mathbf{t}_i) + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$ is iid noise, α, β_1, β_2 and σ are estimated parameters and f a thin plate regression spline (with basis size $k = 100$) with $\mathbf{t}_1, \dots, \mathbf{t}_n$ the observed data locations.

The covariate effects of interest are β_1 and β_2 but, as the results of Sections 3 and 4 show, the estimates of these effects may be highly biased in both the null model and the spatial model. This disproportionate bias is avoided in the spatial+ model. Let $\mathbf{r}^1 = (r_1^1, \dots, r_n^1)^T$ and $\mathbf{r}^2 = (r_1^2, \dots, r_n^2)^T$ be the residuals when a thin plate regression spline (with basis size $k = 100$) is fitted to `age` and `tminmay`, respectively. The spatial+ model is then

$$\mathbf{ratio}_i = \alpha + \beta_1 r_i^1 + \beta_2 r_i^2 + f^+(\mathbf{t}_i) + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$ is iid noise, α, β_1, β_2 and σ are estimated parameters and f^+ a thin plate regression spline (with basis size $k = 100$) with $\mathbf{t}_1, \dots, \mathbf{t}_n$ the observed data locations.

Finally, for comparison, we fit the gSEM as an alternative method for avoiding spatial

confounding bias. Let $\mathbf{r}^y = (r_1^y, \dots, r_n^y)^T$ be the residuals when a thin plate regression spline (with basis size $k = 100$) is fitted to the response variable `ratio`. The gSEM is then

$$r_i^y = \beta_1 r_i^1 + \beta_2 r_i^2 + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma^2)$ is iid noise and β_1, β_2 and σ are estimated parameters.

5.2 Results

The results of fitting the above four models to the data are summarized in Table 1.

[Table 1 about here.]

The spatial term in the spatial model is significant, which confirms there is residual spatial correlation in the data as expected. Furthermore, as the spatial term allows for more of the residual variation to be explained, the deviance explained is higher and the estimated standard deviation is lower than in the null model. As the AIC is also lower, we conclude that the spatial model is an overall better fitting model than the null model for this data. However, while the spatial model may be appropriate for overall predictions of the response variable, the estimate of any individual covariate effect may be biased. Using the spatial+ model, we expect to obtain similar fitted values as the spatial model but with covariate effect estimates that have only negligible bias. Indeed, in terms of overall fit, we see that the deviance explained, estimated standard deviation and AIC in the spatial+ model are similar to those of the spatial model. For completeness, we have also included the gSEM. Note, however, that in the gSEM, since the response variable in the regression differs from that of the other three models, the deviance explained, estimated standard deviation and AIC cannot be directly compared to the other models.

The covariate `age` is highly significant and has a positive effect as expected. This covariate does not appear to be affected by spatial confounding as the estimated effect and its p-value are largely robust to the choice of model. This happens, for example, if a covariate

is independent of the true underlying residual spatial effect. Also, in the case of `age`, not only is this a covariate that is not very well explained by spatial location (a spatial smooth fitted to this variable has deviance explained of only 13%), but its estimated spatial pattern looks dominated by linear spatial basis functions which are unpenalized in the spatial model. Therefore, penalization of the spatial term f in the spatial model is less likely to interfere with the covariate effect estimate (see Rice (1986) Proposition D).

In contrast, the estimated effect of the covariate `tminmay` is not significant in the null model but is significant in the spatial model and is even more significant in the spatial+ model. Furthermore, while in all models the effect estimate is positive as expected (i.e. higher temperature in May leads to more defoliation later in summer), the size of the estimate more than triples when a spatial effect is added to the null model and the estimate in the spatial+ model is more than double that in the spatial model. This shows that, if we were to use the spatial model for our inference, the effect of temperature on crown defoliation would likely be underestimated in both size and significance due to spatial confounding. Note that, as expected, the gSEM gives similar results to spatial+.

6. Non-Gaussian response data

In this section we generalize the models from Section 2 to response distributions from the exponential family, which includes, e.g. the Poisson, gamma and binomial distributions.

6.1 Spatial model

Suppose we have response data $\mathbf{y} = (y_1, \dots, y_n)^T$ where each y_i is assumed to be a random variable whose distribution is from the exponential family with $E(y_i) = \mu_i$, and suppose $\mathbf{x} = (x_1, \dots, x_n)^T$ and $\mathbf{t}_1, \dots, \mathbf{t}_n$ are covariate observations and spatial locations as before. A generalized version of (1) can then be formulated as

$$g(\mu_i) = \beta x_i + f(\mathbf{t}_i) \tag{13}$$

where β is an unknown parameter, f a thin plate spline and $g : \mathbb{R} \rightarrow \mathbb{R}$ a link function (i.e. a monotonic smooth function which ensures $g(\mu_i)$ is in the domain of the response variable). The partial thin plate spline estimates $\hat{\beta}$ and $\hat{\mathbf{f}} = (\hat{f}(\mathbf{t}_1), \dots, \hat{f}(\mathbf{t}_n))^T$ are found using the penalized iterative re-weighted least squares (PIRLS) algorithm (for details, see Web Appendix G). If no smoothing is applied, these estimates are the maximum likelihood estimates in a generalized linear model (GLM), which are asymptotically unbiased.

6.2 Spatial+ model

Starting with the model (13), let \mathbf{W} and \mathbf{z} denote the weights matrix and pseudodata at convergence of the PIRLS algorithm (described in Web Appendix G). We then define the corresponding spatial+ model as follows. Let $\hat{\mathbf{f}}^x$ and $\mathbf{r}^x = \mathbf{x} - \hat{\mathbf{f}}^x = (r_1^x, \dots, r_n^x)^T$ denote the fitted values and residuals in the weighted thin plate regression (4) with weights \mathbf{W} , i.e. $\hat{\mathbf{f}}^x$ is the minimizer of $\|\sqrt{\mathbf{W}}(\mathbf{x} - \mathbf{f}^x)\|^2 + n\lambda_x \mathbf{f}^{xT} \mathbf{\Gamma} \mathbf{f}^x$ with smoothing parameter $\lambda_x > 0$ and $\mathbf{\Gamma}$ defined as in (2). The spatial+ model is then the partial thin plate spline model defined by

$$g(\mu_i) = \beta r_i^x + f^+(\mathbf{t}_i) \quad (14)$$

where β and f^+ are estimated as in Section 6.1. For further details, see Web Appendix G.

6.3 Simulations

The models (13) and (14) can once again be implemented using thin plate regression splines in `mgcv`. To test the performance of the spatial+ model (14), we repeat the simulations from Section 4 for three different response distributions, namely, the Poisson distribution with canonical link function $g(\mu) = \log(\mu)$, the exponential distribution with (non-canonical) link function $g(\mu) = \log(\mu)$ and the binomial distribution with size parameter $n_{\text{bin}} = 10$ and canonical link function $g(\mu) = \log(\mu/(n_{\text{bin}} - \mu))$.

For each response distribution, we simulate 100 replicates of the response data $\mathbf{y} = (y_1, \dots, y_n)^T$ by independently sampling each y_i from the given distribution with mean

$\mu_i = g^{-1}(\eta_i)$ where $\eta_i = \beta x_i + f_i$ with $\mathbf{x} = (x_1, \dots, x_n)^T$ simulated as in Section 4.1, $\sigma_x = 0.1$, and true effects $\beta = 3$, $\mathbf{f} = (f_1, \dots, f_n)^T = -\mathbf{z} - \mathbf{z}'$ as before. The results of fitting the models (13) and (14) are summarized in Figure 3. For comparison we have also included the results of fitting the corresponding null model (i.e. the GLM defined by $g(\mu_i) = \beta x_i$) and the models (13) and (14) with no smoothing penalty applied. Finally, we have fitted a generalized version of the RSR model (for details see Web Appendix G). Note that we have not included the gSEM here as it is not immediately clear how to generalize this model to non-Gaussian response distributions.

We see that for all three response distributions, the overall behaviour of the models is similar to what we saw in the Gaussian case. As before, the null model and RSR model both have highly biased covariate effect estimates; however, note that unlike the Gaussian case, the estimate is not the same in the two models. This is because, while in both models the estimate is given by $\hat{\beta} = (\mathbf{x}^T \mathbf{W} \mathbf{x})^{-1} \mathbf{x}^T \mathbf{W} \mathbf{z}$, the fitted values, and hence the weights and pseudodata at convergence, differ. Without smoothing, as expected, the spatial and spatial+ models give the same results, however, while the covariate effect estimate looks unbiased for the Poisson and exponential response distributions, it looks slightly biased for the binomial distribution, though not materially. This is not surprising as GLMs are only asymptotically unbiased and may have some bias in practice, particularly, when the number of estimated parameters is relatively large as it is in this case (Cox and Snell, 1968). When smoothing is applied, MSE reduces as intended, but the covariate effect estimate in the spatial model becomes significantly biased while it remains broadly unbiased in the spatial+ model.

[Figure 3 about here.]

7. Discussion

We have shown that for covariates that are spatially dependent but not fully spatial (i.e. each covariate is not fully explained when a spatial model is fitted to it), the proposed spatial+ model can be used to avoid unreliable effect estimates in spatial regression with clear advantages over existing methods. Our analysis also gives a clearer understanding of spatial confounding in this context. Spatial models, whether formulated in terms of spatially induced prior distributions or smoothing penalties, usually apply some form of spatial smoothing to reflect spatial correlation in the data and avoid overfitting. However, from the model formulation (1), we see that it is exactly this smoothing that causes spatial confounding bias as, without smoothing, the spatial model has unbiased estimates. The non-spatial information in the covariate means that the model can distinguish it from an unmeasured spatial confounder. However, if the correlation between the covariate and the spatial confounder is high, the smoothing applied to the spatial term in the model can disproportionately affect the estimate of the covariate effect.

The excessive smoothing-induced bias is avoided in both spatial+ and the gSEM. If no smoothing penalty is applied, both models give the same unbiased covariate effect estimates as the unsmoothed spatial model. Spatial+ reparametrizes the spatial model so that spatial dependence is removed from the covariates and instead fully contained in the spatial term f^+ . This makes fixed effect estimates broadly independent of the spatial effects, in particular, they remain largely unbiased under spatial smoothing. The idea of decorrelating covariate and spatial terms is also used in RSR, however, restricting the spatial effects leads to bias by construction. In the gSEM, the elimination of all spatial information means that fixed effect estimates are once again decorrelated from the spatial effects and thereby protected from spatial smoothing. The resulting model of residuals only, however, seems less intuitive than spatial+ and, the change in response variable means that standard model selection criteria

cannot be used for comparisons with the other models. A major advantage of spatial+ is also that the method generalizes easily to models with non-Gaussian response distributions and our simulations illustrate that the method still works well here.

Our above discussion shows that the decorrelation of effect estimates is the underlying reason why the spatial+ approach works. As mentioned in Section 1, the modification of the model matrix that achieves this is easily transferable to other spatial model formulations, and we would therefore expect the method to work well in general. However, as our theoretical derivations are specific to thin plate spline estimates, similar derivations or simulations could be done to confirm our results in other settings. One limitation to the spatial+ approach is that the covariate effects in the model must be linear. This assumption is needed for the spatial residuals to capture the true covariate effects. The spatial model (1) is easily extended, using the generalized additive model (GAM) framework, to include non-linear covariate terms in the form of smooths (i.e. unknown functions of the covariates estimated from the data). It would be interesting to see if any of the ideas of spatial+, as well as our increased understanding of spatial confounding, can be used to develop methods for avoiding spatial confounding in this context.

Finally, applying spatial+ to the forestry example, we see that the effect of temperature on crown defoliation appears to be positive and significant as expected, and that this effect would likely be underestimated in both size and significance in the spatial model (and even more so in the null model). The other covariate, age of trees, in this example also illustrates that, if a covariate is not spatially confounded, this can be confirmed by showing that its effect estimate in the spatial and spatial+ models agree. It is possible that this idea could be used to develop a diagnostic or test that practitioners could use to identify spatial confounding in applications.

ACKNOWLEDGEMENTS

Emiko Dupont is supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/L015684/1. We thank The Forest Institute Baden-Württemberg (Germany) for making the 2013 Terrestrial Crown Condition Inventory (TCCI) forest health monitoring survey data available.

REFERENCES

- Augustin, N., Musio, M., von Wilpert, K., Kublin, E., Wood, S., and Schumacher, M. (2009). Modeling spatiotemporal forest health monitoring data. *Journal of the American Statistical Association* **104**, 899–911.
- Chen, H. and Shiau, J.-J. H. (1991). A two-stage spline smoothing method for partially linear models. *Journal of Statistical Planning and Inference* **27**, 187–201.
- Chen, H. and Shiau, J.-J. H. (1994). Data-driven efficient estimators for a partially linear model. *The Annals of Statistics* pages 211–237.
- Clayton, D. G., Bernardinelli, L., and Montomoli, C. (1993). Spatial correlation in ecological analysis. *International Journal of Epidemiology* **22**, 1193–1202.
- Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society: Series B (Methodological)* **30**, 248–265.
- Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in sobolev spaces. In *Constructive theory of functions of several variables*, pages 85–100. Springer.
- Eichhorn, J., Roskams, P., Potočič, N., Timmermann, V., Ferretti, Mues, V., Szepesi, A., Durrant, D., Seletković, I., H-W.Schröck, Nevalainen, S., Bussotti, F., Garcia, P., and Wulff, S., editors (2017). *ICP Forests manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests*. Thünen Institute of Forest Ecosystems, Eberswalde, Germany.

- Fahrmeir, L., Kneib, T., and Lang, S. (2004). Penalized structured additive regression for space-time data: a bayesian perspective. *Statistica Sinica* pages 731–761.
- Hanks, E. M., Schliep, E. M., Hooten, M. B., and Hoeting, J. A. (2015). Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics* **26**, 243–254.
- Hodges, J. S. and Reich, B. J. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician* **64**, 325–334.
- Khan, K. and Calder, C. A. (2020). Restricted spatial regression methods: Implications for inference. *Journal of the American Statistical Association* pages 1–13.
- Kimeldorf, G. S. and Wahba, G. (1970). Spline functions and stochastic processes. *Sankhyā: The Indian Journal of Statistics, Series A* pages 173–180.
- Kirkwood, B. R. and Sterne, J. AC. (2010). Essential medical statistics. *John Wiley & Sons*
- McNamee, R. (2003). Confounding and confounders. *Occupational and environmental medicine* **60** (3), 227–234.
- Paciorek, C. J. (2010). The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical science: a review journal of the Institute of Mathematical Statistics* **25**, 107.
- Page, G. L., Liu, Y., He, Z., and Sun, D. (2017). Estimation and prediction in the presence of spatial confounding for spatial linear models. *Scandinavian Journal of Statistics* .
- Reich, B. J., Hodges, J. S., and Zadnik, V. (2006). Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. *Biometrics* **62**, 1197–1206.
- Rice, J. (1986). Convergence rates for partially splined models. *Statistics & probability letters* **4**, 203–208.
- Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society: Series B (Methodologi-*

cal) **47**, 1–21.

Thaden, H. and Kneib, T. (2018). Structural equation models for dealing with spatial confounding. *The American Statistician* **72**, 239–252.

Utreras, F. I. (1988). Convergence rates for multivariate smoothing spline functions. *Journal of approximation theory* **52**, 1–27.

Wahba, G. (1990). *Spline models for observational data*, volume 59. Siam.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.

SUPPORTING INFORMATION

Web Appendices referenced in Sections 1, 3, 4 and 6, along with the R code for Sections 4, 5 and 6 are available with this paper at the Biometrics website on Wiley Online Library.

Received July 2020. Revised July 2020. Accepted July 2020.

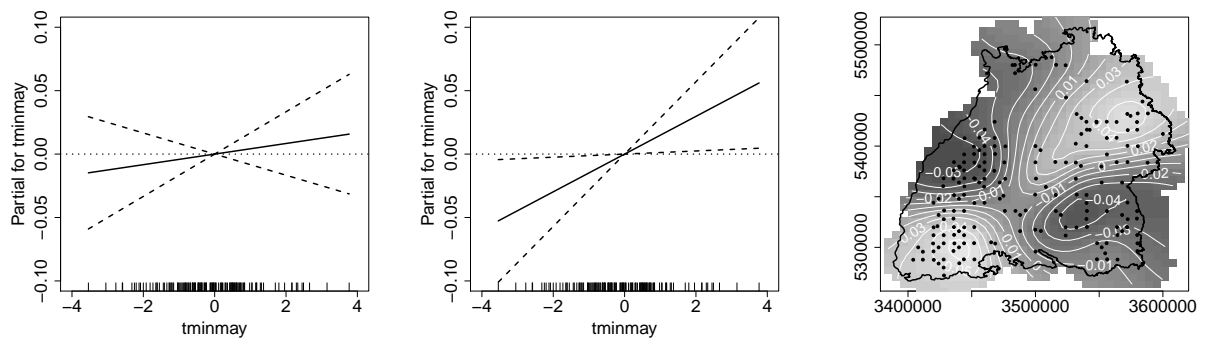


Figure 1. Forestry example. Estimated effect of minimum temperature in May on crown defoliation in the null model (left) and the spatial model (middle), where for each model the plot shows the contribution of the centered covariate to the predicted response (with two times standard error bands). Estimated spatial effect in the spatial model (right) with the border of Baden-Württemberg outlined and dots showing the data locations.

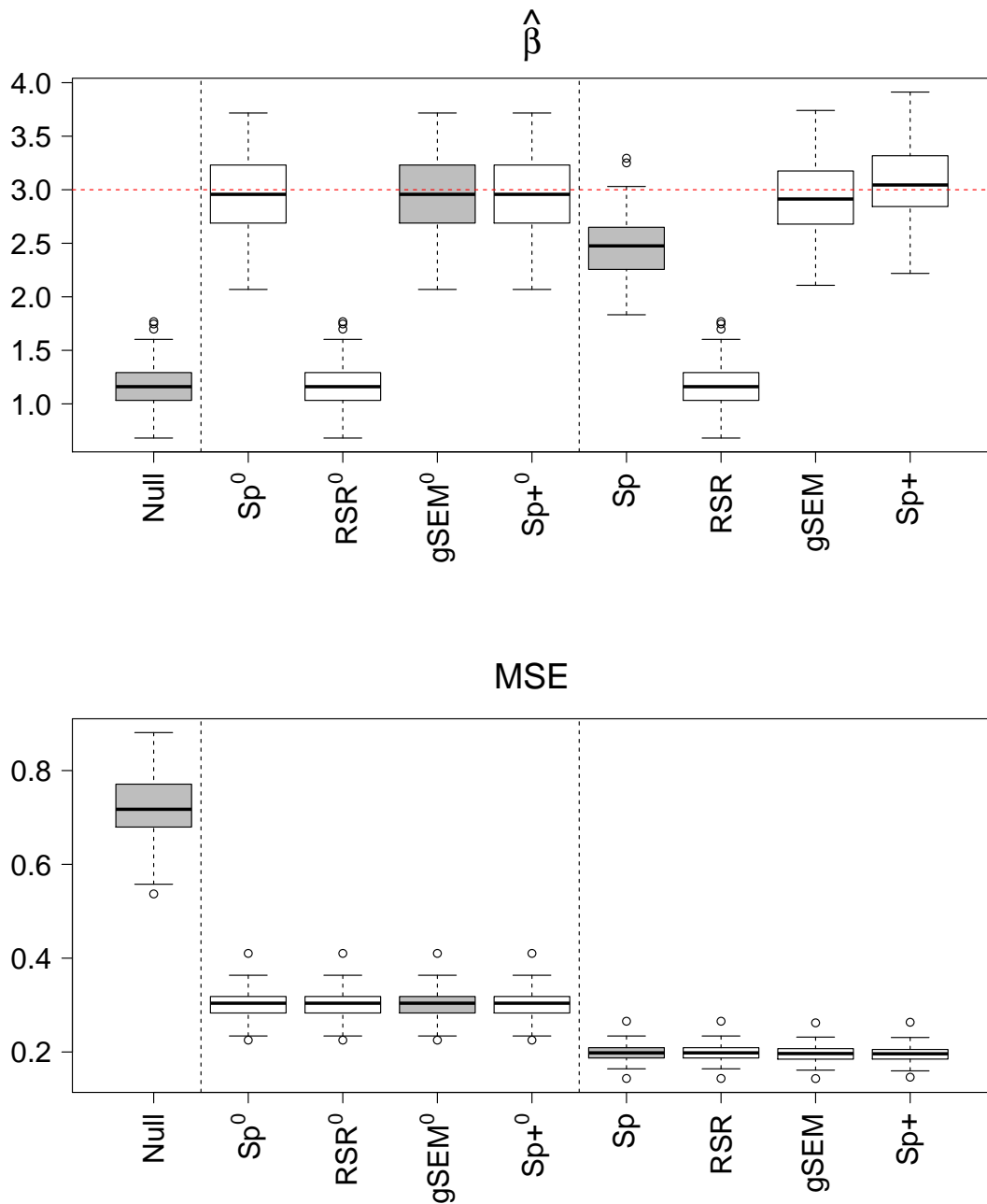


Figure 2. Estimated covariate effect $\hat{\beta}$ (top) and MSE of fitted values (bottom) for each model fitted to 100 data replicates, where the true covariate effect is $\beta = 3$. Sp and Sp+ denote the spatial and spatial+ models, respectively, and superscript 0 refers to an unsmoothed model (i.e. $\lambda = \lambda_x = 0$). Results in grey are the three models that correspond to those used in Thaden and Kneib's simulation study. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

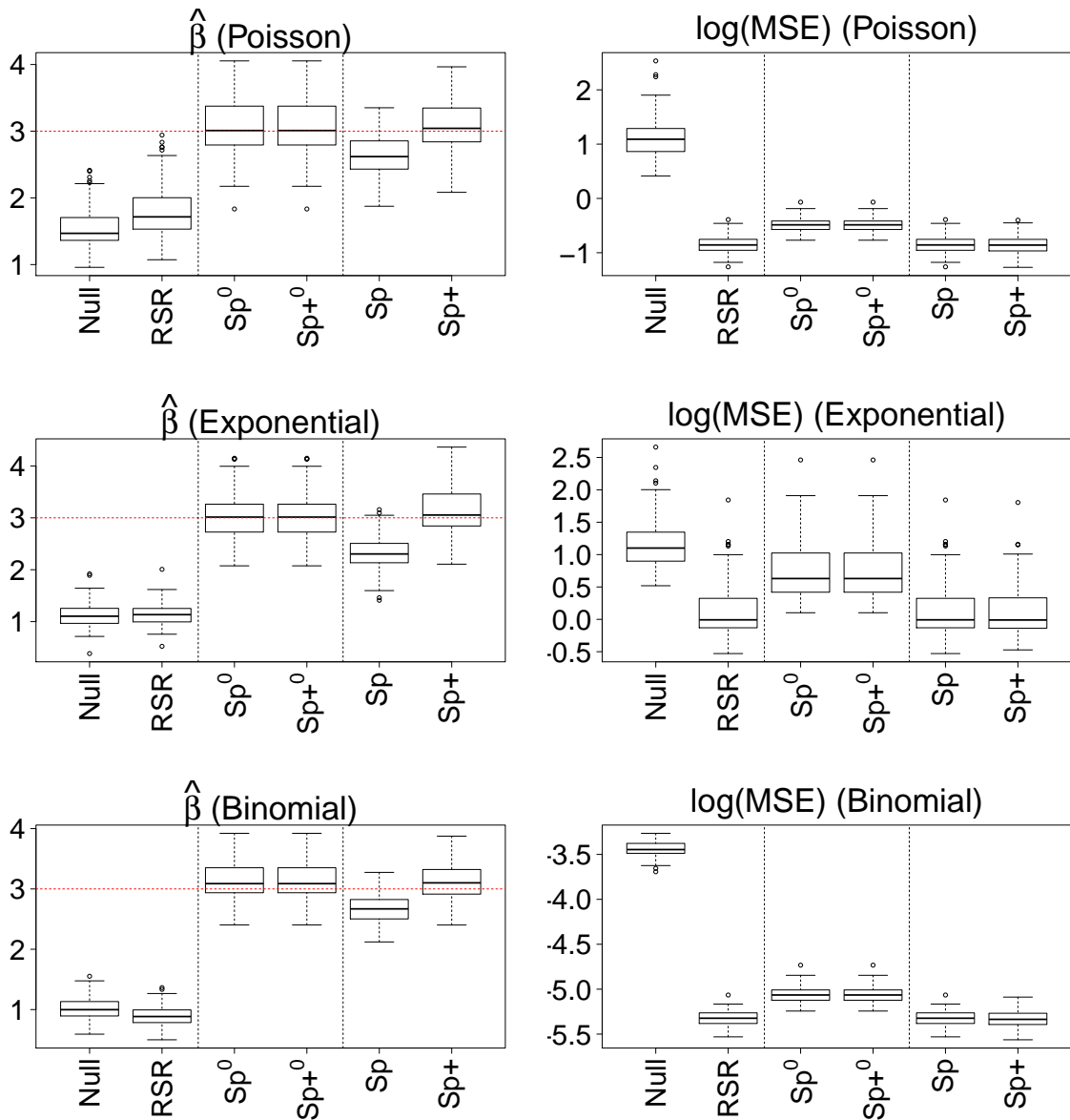


Figure 3. For each of the distributions Poisson (top), exponential (middle), binomial (bottom): the estimated covariate effect $\hat{\beta}$ (left) and $\log(\text{MSE})$ of fitted values (right) for each model fitted to 100 data replicates, where the true covariate effect is $\beta = 3$. Sp and Sp+ denote the spatial and spatial+ models, respectively, and superscript 0 refers to an unsmoothed model (i.e. $\lambda = \lambda_x = 0$). This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

Table 1

Forestry example: results of fitting models to the data. For each covariate: the estimate of the covariate effect β and its p -value. $s(\mathbf{x}, \mathbf{y})$ refers to the thin plate regression splines fitted to f in the spatial model and f^+ in the spatial+ model. For each of these: the effective degrees of freedom (edf) and the p -value. For each significant p -value we write *** if it is < 0.001 , ** if < 0.01 and * if < 0.05 . Note that in the gSEM, deviance explained, estimated standard deviation and AIC do not compare directly with the other models as the response variable is different.

	age			tminmay			s(x, y)			Dev expl	$\hat{\sigma}$	AIC
	$\hat{\beta}$	p-value		$\hat{\beta}$	p-value		edf	p-value				
Null	0.00247	$< 10^{-16}$	***	0.0042	0.5049					0.490	0.00940	-335
Spatial	0.00237	$< 10^{-16}$	***	0.0149	0.0307	*	14.2	0.0243	*	0.605	0.00789	-355
Spatial+	0.00237	$< 10^{-16}$	***	0.0316	0.0073	**	12.0	3.32e-05	***	0.598	0.00793	-356
gSEM	0.00232	$< 10^{-16}$	***	0.0317	0.0058	**						