Spatial Confounding in Generalized Estimating Equations

Francis K.C. Hui*

Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, Australia

and

Howard D. Bondell

School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia

Abstract

Spatial confounding, where the inclusion of a spatial random effect introduces multicollinearity with spatially structured covariates, is a contentious and active area of research in spatial statistics. However, the majority of research into this topic has focused on the case of spatial mixed models. In this article, we demonstrate that spatial confounding can also arise in the setting of generalized estimating equations (GEEs). The phenomenon occurs when a spatially structured working correlation matrix is used, as it effectively induces a spatial effect which may exhibit collinearity with the covariates in the marginal mean. As a result, the GEE ends up estimating a so-called unpartitioned effect of the covariates. To overcome spatial confounding, we propose a restricted spatial working correlation matrix that leads the GEE to instead estimate a partitioned covariate effect, which additionally captures the portion of spatial variability in the response spanned by the column space of the covariates. We also examine the construction of sandwich-based standard errors, showing that the issue of efficiency is tied to whether the working correlation matrix aligns with the target effect of interest. We conclude by highlighting the need for practitioners to make clear the assumptions and target of interest when applying GEEs in a spatial setting, and not simply rely on the robustness property of GEEs to misspecification of the working correlation matrix.

Keywords: marginal models; restricted spatial regression; sandwich covariance; spatial correlation; working correlation; unconditional effect

*Corresponding author: Francis K.C. Hui, Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, ACT, 2600, Australia. Email: francis.hui@anu.edu.au
This work was supported by the Australian Research Council, under Grant DE200100435. Thanks to Thomas Lumley and Tao Zou for useful discussions.

1 Introduction

Generalized estimating equations (GEEs, Liang and Zeger, 1986) are a well-established and studied approach for analyzing spatial data (Albert and McShane, 1995; Lin and Clayton, 2005). The method consists of a model relating the marginal mean to a set of observed covariates, an assumption on the marginal mean-variance relationship, and a working correlation matrix characterizing the marginal spatial correlation of the responses. A key feature of (spatial) GEEs is their robustness to misspecification of the working correlation matrix: the estimated coefficients converge to the same true parameter values, with the choice of working correlation only affecting the efficiency of the estimate.

In this article, we demonstrate that when GEEs are applied to spatially indexed data, spatial confounding (Hodges and Reich, 2010; Paciorek, 2010; Hanks et al., 2015) can arise, with the main consequence being that changing the working correlation can, in fact, change the target quantity that the GEE is estimating. At its core, spatial confounding in GEEs occurs because assuming a spatially structured working correlation effectively induces a spatial effect in the marginal mean, which may be collinear with other spatially indexed covariates. This results in the GEE estimating a so-called unpartitioned effect of the covariates. As an alternative and to alleviate for spatial confounding, we propose a restricted spatial working correlation matrix based on the idea of partitioning the induced spatial effect into a component that can be explained by the covariates along with a residual projection component, and then moving the former into the marginal mean. We show that the resulting, restricted spatial GEE estimates a so-called partitioned effect of the covariates, which contains the portion of spatial variability in the response lying in the direction of the covariates. In the case where a constant mean-variance relationship is assumed, restricted spatial GEEs simplify to independent estimating equations (IEEs, Liang and Zeger, 1986), with the implication that adjusting for spatial confounding in this setting produces the same estimates as ignoring the spatial correlation entirely (see also Khan and Calder, 2020; Zimmerman and Ver Hoef, 2021). We further demonstrate how spatial confounding has implications for inference in GEEs, specifically, statistical efficiency is tied to whether the choice of the working correlation matrix reflects the inferential target itself, and it is not simply a matter of how close the working correlation is to the true marginal correlation.

This paper makes an important contribution to the area of spatial confounding, as almost all research so far has been devoted to its occurrence in (Bayesian) spatial mixed models where the problem is relatively explicit i.e., both the fixed effects and spatial random effect are posited directly as part of the linear predictor. We refer the reader to Nobre et al. (2021) and Reich et al. (2021) for recent reviews on the topic, and Hodges and Reich (2010); Paciorek (2010); Hughes and Haran (2013); Hanks et al. (2015); Khan and Calder (2020); Dupont et al. (2021) among many others for examples of research into spatial confounding in the mixed models framework.

To our knowledge, spatial confounding has not been previously raised as an issue for GEEs; in fact, Paciorek (2010) conjectured that the estimating equation approach was not capable of reducing bias from unmeasured spatial confounding, while Hodges and Reich (2010) noted as a aside that GEEs adjusts standard errors for clustering but has little effect on point estimates unless the working correlations are very large. Our proposed restricted spatial GEE can be interpreted as an estimating equation version of restricted spatial regression (Hodges and Reich, 2010). That is, because spatial confounding occurs indirectly in a GEE, then to alleviate this we propose to adjust the working correlation rather than the marginal mean itself. Interestingly, the presence of a (typically) non-constant mean-variance relationship means that this adjustment is a function of both the observed covariates *and* the coefficients in the GEE. This contrasts to the mixed model setting where the adjustment usually depends solely on the former.

More generally, the concept of confounding in GEEs has been raised before in the setting of longitudinal data (see for instance, Gromping, 1996; Crouchley and Davies, 1999). Recently, Bible et al. (2019) went so far as to say that, in the context of GEEs for marginal transition models, practitioners have been arbitrarily choosing working correlation matrices and then mistakenly citing the works of Liang and Zeger (1986) among others for the robustness properties of GEEs. This article is the first to address similar issues arising when GEEs are applied to spatial data. At the same time, it is important to emphasize that we are by no means advocating the proposed restricted spatial GEEs as a necessarily superior method of inference in the estimating equation setting. Restricted spatial regression is not a universally accepted approach to alleviate for spatial confounding (Khan and Calder, 2020), and there exists active discussion in the spatial statistics literature regarding what exactly various approaches to alleviating spatial confounding are estimating and assuming (Hanks et al., 2015; Hefley et al., 2017; Khan and Calder, 2020; Papadogeorgou, 2021). It is not the aim of this article to settle this debate in the context of GEEs, and ultimately we do not believe there is a single best approach under all data settings. Rather, our main message is one of caution: spatial confounding *can* occur in GEEs, and while we have proposed one approach to alleviating this, this may not necessarily be what the practitioner wants. Rather, we must be more careful about the choice of the working correlation when applying GEEs to spatially indexed data, and understand whether it aligns with the target of interest and the assumptions regarding the true data generation mechanism.

The rest of this article is structured as follows. Section 2 establishes the concept of spatial confounding in GEEs and proposes the restricted spatial working covariance matrix. Section 2.1 provides some interpretation and insight behind the unpartitioned and partitioned effects, while Section 3 discusses the construction of standard errors. Sections 4 and 5 demonstrates the presence and impact of spatial confounding in GEEs through simulation and a real application to a dataset on pelagic fish species richness. Section 6 offers some concluding thoughts.

2 Spatial GEEs

Consider a set of n spatially indexed observations $\{\boldsymbol{x}(\boldsymbol{s}_i), y(\boldsymbol{s}_i); i = 1, \ldots, n\}$, where $\boldsymbol{s}_i \in \mathcal{D}$ denotes the location of the i-th observation in some spatial domain $\mathcal{D}$, $y(\boldsymbol{s}_i)$ denotes a univariate response, and $\boldsymbol{x}(\boldsymbol{s}_i)$ denotes a p-vector of covariates. In this article we focus on the geostatistical setting where we have a continuous distance measure between spatial locations, although the developments below carry over to the case where we have areal data and the dependence is described through an associated adjacency matrix (say). Let $\boldsymbol{y} = \{y(\boldsymbol{s}_1), \ldots, y(\boldsymbol{s}_n)\}^\top$ denote the full n-vector of responses, and $\boldsymbol{X}$ denote the $n \times p$ model matrix formed from stacking the $\boldsymbol{x}(\boldsymbol{s}_i)$ as row vectors, and which is assumed to be of full column rank. We consider fitting spatial GEEs to such data, which involves the following three assumptions: 1) the marginal mean, $E(\boldsymbol{y}) = \boldsymbol{\mu}$, is modeled as $g(\boldsymbol{\mu}) = \boldsymbol{X}\boldsymbol{\tau}$, where $\boldsymbol{\tau}$ is a p-vector of regression coefficients and $g(\cdot)$ is a known link function applied element-wise to $\boldsymbol{\mu}$; 2) the marginal mean-variance relationship is given by $\text{Var}(\boldsymbol{y}) = \phi h(\boldsymbol{\mu})$ for some dispersion parameter $\phi > 0$ and variance function $h(\cdot)$ applied element-wise to $\boldsymbol{\mu}$; 3) a working correlation matrix $\boldsymbol{R}$ is used to describe the spatial correlation between observations. Based on these moment assumptions, a GEE then solves the system of equations

$$\frac{1}{\phi}\boldsymbol{D}^\top \boldsymbol{A}^{-1/2}\boldsymbol{R}^{-1}\boldsymbol{A}^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu}) = \boldsymbol{0},$$

where $\boldsymbol{D} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\tau}$ and $\boldsymbol{A} = \text{Diag}\{h(\boldsymbol{\mu})\}$. We alternate between solving the above to obtain updates of $\boldsymbol{\tau}$, and then separately updating for $\boldsymbol{R}$ and ϕ. For example, the latter can be updated via maximum pseudo-likelihood estimation, and we provide more details about this in the Supplementary Material.

For spatially indexed data, a typical form to assume for the working correlation is $\boldsymbol{R}_{\text{sp}} = \pi\boldsymbol{\Sigma}_{\text{sp}} + (1-\pi)\boldsymbol{I}$, that is, a weighted average of a spatial correlation matrix $\boldsymbol{\Sigma}_{\text{sp}}$ and a nugget effect as represented by the identity matrix $\boldsymbol{I}$, where $\pi \in (0,1)$ (see for example Albert and McShane, 1995; Lin and Clayton, 2005; Adegboye et al., 2018, noting the nugget effect is sometimes omitted). The precise form of the spatial correlation matrix, $\boldsymbol{\Sigma}_{\text{sp}}$, is not important here, and we only require it to be a positive definite matrix. In practice, a common choice is to parameterize $\boldsymbol{\Sigma}_{\text{sp}}$ via a Matérn correlation function with smoothness $\nu > 0$ and spatial scale parameter $s > 0$ (Lin and Clayton, 2005; Adegboye et al., 2018). Next, to ease discussion and make the notation more analogous to what is commonly seen in the spatial mixed model literature (e.g., Hanks et al., 2015), we adopt the alternate parameterization $\phi\pi = \sigma^2_{\text{sp}}$ and $\phi(1-\pi) = \sigma^2_e$, and subsequently define the *unrestricted spatial working covariance matrix* $\boldsymbol{V}_{\text{sp}} = \phi\boldsymbol{R}_{\text{sp}} = \sigma^2_{\text{sp}}\boldsymbol{\Sigma}_{\text{sp}} + \sigma^2_e\boldsymbol{I}$ along with the resulting unrestricted spatial GEE, $\boldsymbol{D}^\top\boldsymbol{A}^{-1/2}\boldsymbol{V}_{\text{sp}}^{-1}\boldsymbol{A}^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu}) = 0$. Estimation of $(\phi,\pi)^\top$ is then replaced with estimation of the variance parameters $(\sigma^2_{\text{sp}}, \sigma^2_e)^\top$.

Conditional on the unrestricted spatial covariance $\boldsymbol{V}_{\text{sp}}$, we can solve and interpret the resulting unrestricted spatial GEE as iteratively minimizing the quadratic loss function $2^{-1}\left\{(\boldsymbol{A}^{(0)})^{-1/2}\boldsymbol{z}^{(0)} - (\boldsymbol{A}^{(0)})^{-1/2}\boldsymbol{D}^{(0)}\boldsymbol{\tau}\right\}^\top \boldsymbol{V}_{\text{sp}}^{-1}\left\{(\boldsymbol{A}^{(0)})^{-1/2}\boldsymbol{z}^{(0)} - (\boldsymbol{A}^{(0)})^{-1/2}\boldsymbol{D}^{(0)}\boldsymbol{\tau}\right\}$, where $\boldsymbol{z}^{(0)} = \boldsymbol{D}^{(0)}\hat{\boldsymbol{\tau}}^{(0)} + (\boldsymbol{y} - \boldsymbol{\mu}^{(0)})$ is an n-vector of working responses based on the coefficient values at the current iteration, denoted as $\boldsymbol{\tau}^{(0)}$, and $\mu^{(0)} = g^{-1}(\boldsymbol{X}\hat{\boldsymbol{\tau}}^{(0)})$. It is straightforward to show that iteratively minimizing this loss function is equivalent to applying a Newton-Raphson method to the unrestricted spatial GEE. We can further interpret the quadratic loss function as iteratively solving the working linear model

$$\boldsymbol{A}^{-1/2}\boldsymbol{z} = \boldsymbol{A}^{-1/2}\boldsymbol{D}\boldsymbol{\tau} + \boldsymbol{\rho} + \boldsymbol{e}, \tag{1}$$

where the dependence on values at the current iteration is omitted for ease of presentation. The quantities $\boldsymbol{\rho}$ and $\boldsymbol{e}$ denote induced spatial and nugget effects respectively, which are independent of each other, and satisfy $\mathrm{E}(\boldsymbol{\rho}) = \mathrm{E}(\boldsymbol{e}) = \boldsymbol{0}$ and $\mathrm{Cov}(\boldsymbol{\rho}) = \sigma^2_{\mathrm{sp}}\boldsymbol{\Sigma}_{\mathrm{sp}}$, $\mathrm{Cov}(\boldsymbol{e}) = \sigma^2_e\boldsymbol{I}$. We emphasize that because we are working with GEEs, then neither the spatial or nugget effects are explicitly assumed as part of the model setup. Instead, the two effects are implied by the unrestricted spatial working covariance, $\boldsymbol{V}_{\mathrm{sp}}$. By iteratively solving the working linear model in (1), we obtain coefficient estimates from an unrestricted spatial GEE, which we denote here as $\hat{\boldsymbol{\beta}}$. Note we have deliberately chosen a different notation for the estimated coefficients to reflect a specific choice of the working covariance i.e., $\hat{\boldsymbol{\beta}}$ denote estimates based on the unrestricted spatial GEE using $\boldsymbol{V}_{\mathrm{sp}}$ as the form for the working covariance matrix.

Let $\boldsymbol{P}_D = \boldsymbol{D}(\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top$ denote the projection matrix onto the column space of $\boldsymbol{D}$. Then similar to Hanks et al. (2015), we can rewrite equation (1) as

$$\begin{aligned}\boldsymbol{A}^{-1/2}\boldsymbol{z} &= \boldsymbol{A}^{-1/2}\boldsymbol{D}\left(\boldsymbol{\tau} + (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{\rho}\right) + (\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\rho} + \boldsymbol{e} \\ &= \boldsymbol{A}^{-1/2}\boldsymbol{D}\left(\boldsymbol{\tau} + (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{\rho}\right) + \bar{\boldsymbol{\rho}} + \boldsymbol{e}, \end{aligned} \tag{2}$$

where $\bar{\boldsymbol{\rho}} = (\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\rho}$ is a so-called residual projected spatial effect with $\mathrm{E}(\bar{\boldsymbol{\rho}}) = \boldsymbol{0}$ and $\mathrm{Cov}(\bar{\boldsymbol{\rho}}) = \sigma^2_{\mathrm{sp}}(\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\Sigma}_{\mathrm{sp}}(\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)^\top$. Critically, equation (2) suggests that if we define a new *restricted spatial working covariance matrix*

$$\boldsymbol{V}_{\mathrm{rsp}} = \sigma^2_{\mathrm{sp}}(\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\Sigma}_{\mathrm{sp}}(\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)^\top + \sigma^2_e\boldsymbol{I}, \tag{3}$$

and subsequently solve the resulting restricted spatial GEE, $\boldsymbol{D}^\top\boldsymbol{A}^{-1/2}\boldsymbol{V}^{-1}_{\mathrm{rsp}}\boldsymbol{A}^{-1/2}(\boldsymbol{y} - \boldsymbol{\mu}) = \boldsymbol{0}$, then the target quantity being estimated is changed from $\boldsymbol{\tau}$ to $\boldsymbol{\tau} + (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{\rho}$.

More formally, we denote the estimated coefficients from a restricted spatial GEE as $\hat{\boldsymbol{\alpha}}$, as opposed to the estimates from an unrestricted spatial GEE, $\hat{\boldsymbol{\beta}}$. We will compare these two estimates in more detail later on. Stepping back however, the above developments demonstrate that spatial confounding can, in fact, occur in the setting of GEEs, with the major implication being that the choice of the working covariance matrix can have a profound impact on the target that the GEE is estimating.

One interesting feature of the restricted spatial working covariance matrix $\boldsymbol{V}_{\text{rsp}}$ is that it is a function of the regression coefficients. This contrasts to restricted spatial regression in spatial mixed effects models where, because the projection is done on the scale of the linear predictors, then the residual projection is (almost always) chosen to be $(\boldsymbol{I} - \boldsymbol{P}_X)$ and hence only depends on the measured covariates or some variation thereof (Hodges and Reich, 2010; Hughes and Haran, 2013). Also, note we can form the projection matrix $\boldsymbol{P}_D$ from only a subset of the covariates, and all the developments in this article can be adapted to such case. However for ease of presentation, we focus attention here on the projection formed from all the columns of $\boldsymbol{D}$.

We conclude this section by noting that in the special case of a constant mean-variance relationship i.e., $h(\boldsymbol{\mu}) = \mathbf{1}$, some simplifications arise in the case of restricted spatial GEEs. In the Supplementary Material we show that in this case the restricted spatial GEEs reduces to independent estimating equations (IEEs, Liang and Zeger, 1986), meaning both restricted spatial GEEs and IEEs produce the same estimates, and in fact the same inference if sandwich-based standard errors are used (see also Section 3 later on). The equivalence between the estimates produced from restricted spatial GEEs and IEEs in this special case, noting that the latter effectively amounts to a non-spatial GEE, is consistent with previous literature on restricted spatial regression in the spatial mixed model setting (e.g., Khan and Calder, 2020; Zimmerman and Ver Hoef, 2021). However, the fact that this equivalence holds provided a constant mean-variance relationship is assumed i.e., it does not depend on the choice of link function $g(\cdot)$, is a new finding and has some interesting implications. Given that, in practice, GEEs are primarily used for the situation with a non-constant mean-variance relationship, we defer the full details of these developments to the Supplementary Material.

2.1 To Restrict or Not to Restrict?

Consider the unrestricted spatial GEE, as encapsulated by the working linear model in (1). We interpret the coefficients $\boldsymbol{\beta}$ as an *unpartitioned* effect of the covariates, since this is based around *not* partitioning the induced spatial effect $\boldsymbol{\rho}$ and leaving it entirely in the working covariance $\boldsymbol{V}_{\text{sp}}$. In contrast, we interpret $\boldsymbol{\alpha}$ in the restricted spatial GEE as a *partitioned* effect of the covariates since, as seen from the working linear model in (2), it is motivated from the partitioning $\boldsymbol{\rho} = \boldsymbol{A}^{-1/2}\boldsymbol{P}_D\boldsymbol{\rho} + (\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\rho}$. That is, the induced spatial effect is decomposed into a component that can be explained by the covariates, and the residual partition lying in the orthogonal complement. In the restricted spatial GEE, the former is treated as fixed and pulled into the marginal mean, while the latter remains random and forms part of the restricted spatial working covariance matrix $\boldsymbol{V}_{\text{rsp}}$. The extent to which the unpartitioned and partitioned effects differ is determined by how collinear the induced spatial effect and the covariates are, as quantified by least squares type quantity $(\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{\rho}$, i.e., the regression of the induced spatial effect $\boldsymbol{\rho}$ on $\boldsymbol{D}$, noting that this quantity varies as a function of sample size n and covariates $\boldsymbol{X}$.

To summarize, spatial confounding in GEEs can be viewed as a form of multicollinearity: assuming a spatial working covariance matrix $\boldsymbol{V}_{\text{sp}}$ induces a spatial effect which may be collinear with the observed covariates, and the unpartitioned effect $\boldsymbol{\beta}$ arises as the consequence of this (see also Hanks et al., 2015; Hefley et al., 2017; Khan and Calder, 2020, for analogous explanations of spatial confounding in terms of multicollinearity for Bayesian spatial mixed models). The partitioned effect $\boldsymbol{\alpha}$ is an attempt to adjust for this collinearity, by moving the part of the spatial covariation in the response which can be explained by the covariates into the marginal mean. Put another way, in restricted spatial GEEs, all variation in $\boldsymbol{y}$ over which the covariate $\boldsymbol{X}$ and unrestricted spatial working covariance are competing over is attributed to the former.

In the context of spatial mixed models, Hanks et al. (2015) interpreted $\boldsymbol{\beta}$ as a conditional effect and $\boldsymbol{\alpha}$ as an unconditional effect, based on the idea that in the former one conditions on $\boldsymbol{\rho}$ while in the latter one does not. We choose not use this terminology for two reasons. First, the use of the term "conditional" is potentially confusing here because GEEs are usually thought of as estimating marginal or population-averaged effects, in contrast to the conditional effects derived from mixed models. Second, the interpretation of conditional versus unconditional effects brings about the connotation that the GEE either does or does not condition on the induced spatial effect $\boldsymbol{\rho}$. The above discussion however show that, in fact, the restricted spatial GEE *partly* conditions on $\boldsymbol{\rho}$, specifically, the part spanned by the column space of $\boldsymbol{D}$. The remaining residual projection, $\bar{\boldsymbol{\rho}} = (\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\rho}$, is still treated as random and forms part of the restricted spatial working covariance matrix. Moreover, this residual projection can still be spatially correlated. For example, consider a situation where we fail to include a spatially structured covariate that is informative for the response. Then the induced spatial effect $\boldsymbol{\rho}$ can be thought of as playing the role of this missing covariate (although there is controversy over this interpretation; see Hodges and Reich, 2010). If the missing covariate can not be entirely explained by the included covariates $\boldsymbol{X}$, then the residual projection $(\boldsymbol{I} - \boldsymbol{A}^{-1/2}\boldsymbol{P}_D)\boldsymbol{\rho}$ and thus the restricted spatial working covariance in (3) would still exhibit some sort of spatial structure.

With two working covariance matrices producing two different covariate effects, a natural question to ask is which one should practitioners be (more) interested in (see Hanks et al., 2015; Hefley et al., 2017; Papadogeorgou, 2021, for similar discussions). In the context of GEEs, one could make the case for using a restricted spatial GEE and having more interest in the partitioned effect $\boldsymbol{\alpha}$, as it better aligns with what a marginal estimating equation approach to spatial analysis is designed to do, namely to explain the marginal mean using the observed covariates. Specifically, in fitting a GEE the aim is typically to have everything that the covariates can explain about the response to go into the marginal mean structure, and thus, if this is the goal, it could be argued that this should also include the portion of the spatial structure in the response that can be explained by the covariates. The role of the working correlation/covariance matrix should then be to explain any residual covariation between observations after accounting for this marginal mean. That is, it should be structured so as to not introduce any artificial multicollinearity with the covariates and take away part of their explanatory power in the marginal mean. This is precisely what the restricted spatial GEE sets out to achieve with the partitioned effect $\boldsymbol{\alpha}$.

At the same time, we emphasize that the above is by no means as a definitive argument for restricted spatial GEEs (noting that restricted spatial regression methods in general are by no means universally accepted, Khan and Calder, 2020), and much also depends on the quantity the practitioner is actually interested in estimating (Papadogeorgou, 2021). Rather, the main message of this article is really of caution: spatial confounding can arise in the GEE setting, and as a result we urge practitioners to think carefully about the choice of the working correlation when applying GEEs to spatially indexed data, and whether it aligns with their inferential quantity of interest and the assumptions they make relating to the true data generation mechanism.

3 Standard Errors

For both unrestricted and restricted spatial GEEs, sandwich covariance matrices can be constructed in a manner similar to standard applications of GEEs (Liang and Zeger, 1986). Let $\hat{\boldsymbol{B}}_{\text{sp}} = \hat{\boldsymbol{D}}^\top \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{sp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{D}}$ and $\hat{\boldsymbol{B}}_{\text{rsp}} = \hat{\boldsymbol{D}}^\top \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{rsp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{D}}$ denote the bread matrices based on the estimated unrestricted and restricted spatial GEE, respectively. Then the sandwich covariance matrices for $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ are respectively given by

$$\hat{\boldsymbol{G}}_{\text{sp}} = \hat{\boldsymbol{B}}_{\text{sp}}^{-1} \left(\hat{\boldsymbol{D}}^\top \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{sp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \tilde{\mathcal{V}}_0 \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{sp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{D}} \right) \hat{\boldsymbol{B}}_{\text{sp}}^{-1}, \tag{4a}$$

$$\hat{\boldsymbol{G}}_{\text{rsp}} = \hat{\boldsymbol{B}}_{\text{rsp}}^{-1} \left(\hat{\boldsymbol{D}}^\top \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{rsp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \tilde{\mathcal{V}}_0 \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{V}}_{\text{rsp}}^{-1} \hat{\boldsymbol{A}}^{-1/2} \hat{\boldsymbol{D}} \right) \hat{\boldsymbol{B}}_{\text{rsp}}^{-1}, \tag{4b}$$

where $\tilde{\mathcal{V}}_0 = \widehat{\text{Cov}}(\boldsymbol{y})$ generically denotes an estimate of the true marginal covariance, and quantities are calculated using the relevant parameter estimates. Based on the above, we can construct Wald confidence intervals and hypothesis tests for the estimates from unrestricted and restricted spatial GEEs e.g., for the latter a $(1-s)\times 100\%$ Wald confidence interval for the k-th coefficient is given as $(\hat{\boldsymbol{\alpha}}_k - q_{1-s/2}\hat{G}^{1/2}_{\text{rsp},kk}, \hat{\boldsymbol{\alpha}}_k + q_{1-s/2}\hat{G}^{1/2}_{\text{rsp},kk})$, where $\hat{G}_{\text{rsp},kk}$ denotes the k-th diagonal element of $\hat{\boldsymbol{G}}_{\text{rsp}}$ defined in (4b) and we set $q_{1-s/2}$ as the $(1-s/2)$-th quantile of the t-distribution with $(n-p)$ degrees of freedom. In the Supplementary Material, we provide further discussion of the special case of a constant mean-variance function, and how sandwich standard errors of IEE and restricted spatial GEE coincide in this setting.

Consider now the issue of statistical efficiency, as captured by the sandwich standard errors $\hat{G}^{1/2}_{\text{sp},kk}$ and $\hat{G}^{1/2}_{\text{rsp},kk}$. Commonly, discussions regarding the efficiency of GEEs come down to how close the form of the working covariance matrix matches that of the true marginal covariance $\mathcal{V}_0 = \text{Cov}(\boldsymbol{y})$. This is also the case here e.g., if $\text{Cov}(\boldsymbol{y})$ resembles that of an unrestricted spatial covariance matrix, then (4a) will produce smaller standard errors compared to using (4b), and vice versa. However, the more important but perhaps more subtle point here is that because the choice between the unrestricted and restricted working covariance matrix is tied to whether we are interested in estimating the unpartitioned or partitioned effect of covariates (see Section 2.1), then we see that efficiency is intimately connected to whether the working covariance is aligned with the inferential quantity of interest. To give an example of this, suppose we are interested in the partitioned effect of the covariates, $\boldsymbol{\alpha}$. If we fit the restricted spatial GEE and use the associated sandwich covariance matrix in (4b), then our standard errors will be comparatively small, because the restricted spatial working covariance is aligned with the target that we want to perform inference on. In fact, the estimated standard error here would (approximately) reduce to the naive model-based covariance estimator simply given by $\hat{\boldsymbol{B}}_{\text{rsp}}^{-1}$. On the other hand, if we are interested in $\boldsymbol{\alpha}$ but instead fit the unrestricted spatial GEE and use the associated sandwich covariance matrix as given by (4a), then our standard errors will be comparative larger because the unrestricted spatial working covariance is no longer aligned with the target of interest (since this type of GEE aims to estimate the unpartitioned effect instead). Put another way, even if the target quantity of interest and the working covariance matrix structure are not aligned, it is still possible to perform valid inference on the former e.g., confidence intervals with nominal coverage probability. But we pay the price of less efficiency. We confirm this result with our simulations in the next section. To summarize, because changing the working covariance matrix can affect the target that the GEE is estimating in the presence of spatial confounding, then statistical efficiency also becomes tied to how close the working covariance is aligned with the target quantity of interest.

4 Simulation Study

To empirically demonstrate the presence and implications of spatial confounding in GEEs, we simulated spatially indexed data from either an unrestricted or restricted marginal spatial model, and compared the estimation and inference performance of three types of

GEEs: 1) an IEE with working covariance matrix $\boldsymbol{V}_{\text{ind}} = \sigma_e^2 \boldsymbol{I}$ with σ_e^2 is estimated. We refer to this as GEE_{ind}; 2) An unrestricted spatial GEE characterized by (1), which we refer to as GEE_{sp}; 3) A restricted spatial GEE characterized by (2), which wee refer to as GEE_{rsp}. Maximum pseudo-likelihood estimation was used to estimate all parameters in each of the working covariance matrices; see the Supplementary Material for details. We considered three response types: continuous responses generated from a marginal Gaussian model, count responses from a marginal Poisson model, and responses from a marginal binomial model with trial size equal to five. For brevity, we only present results from the Poisson response case below; results for the Gaussian and binomial response case are provided in the Supplementary Material, and present broadly similar conclusions.

The details of the data generation process are as follows. For each simulated dataset, we first generated n random spatial locations uniformly from the unit square $[0, 1]^2$. We then constructed an $n \times 2$ model matrix $\boldsymbol{X}$ consisting of an intercept and one slope covariate $\boldsymbol{x} = (x_1, \ldots, x_n)^\top \sim N(0, \boldsymbol{\Sigma}_{x,0})$, where $\boldsymbol{\Sigma}_{x,0}$ was parameterized via an exponential correlation function with scale parameter $s_{x,0} = 0.8$ and using an Euclidean distance metric. This value of the spatial scale was chosen based on the formulas given in the simulation study of Hanks et al. (2015), and reflected a moderate spatial dependence. Next, we set up a spatial correlation matrix $\boldsymbol{\Sigma}_{\text{sp},0}$ that was also parameterized via an exponential correlation function with spatial scale $s_{\text{sp},0} = 0.8$. Based on the above quantities, we then simulated spatially structured from one of two potential models: i) an unrestricted marginal spatial model with true marginal mean vector given by $\boldsymbol{\mu}_0 = g^{-1}(\boldsymbol{X}\boldsymbol{\beta}_0)$ for a vector of true unpartitioned effects $\boldsymbol{\beta}_0$, and the true marginal spatial covariance matrix as $\boldsymbol{A}_0^{1/2}\boldsymbol{V}_{\text{sp},0}\boldsymbol{A}_0^{1/2} = \boldsymbol{A}_0^{1/2}(\boldsymbol{\Sigma}_{\text{sp},0} + 0.1\boldsymbol{I})\boldsymbol{A}_0^{1/2}$ where $\boldsymbol{A}_0 = \text{Diag}\{h(\boldsymbol{\mu}_0)\}$; ii) a restricted marginal spatial model with true marginal mean vector given by $\boldsymbol{\mu}_0 = g^{-1}(\boldsymbol{X}\boldsymbol{\alpha}_0)$ for a vector of true partitioned effects $\boldsymbol{\alpha}_0$, and the true marginal spatial covariance matrix as $\boldsymbol{A}_0^{1/2}\boldsymbol{V}_{\text{rsp},0}\boldsymbol{A}_0^{1/2} = \boldsymbol{A}_0^{1/2}\left\{(\boldsymbol{I} - \boldsymbol{A}_0^{-1/2}\boldsymbol{P}_{D_0})\boldsymbol{\Sigma}_{\text{sp},0}(\boldsymbol{I} - \boldsymbol{A}_0^{-1/2}\boldsymbol{P}_{D_0})^\top + 0.1\boldsymbol{I}\right\}\boldsymbol{A}_0^{1/2}$. Notice how that the forms of $\boldsymbol{V}_{\text{sp},0}$ and $\boldsymbol{V}_{\text{rsp},0}$ reflect the forms of the unrestricted and restricted spatial working covariance matrices defined in Section 2, respectively. The values of $\boldsymbol{\beta}_0$ and $\boldsymbol{\alpha}_0$ are discussed later on. We considered sample sizes $n = \{100, 225, 400, 625\}$, and for each n simulated 400 datasets. As discussed in Hanks et al. (2015), with both the slope covariate

$\boldsymbol{x}$ and marginal spatial covariance matrices exhibiting spatial structure, it means that in finite samples they can exhibit collinearity with each other (with the degree of collinearity varying across simulated datasets) and hence spatial confounding can arise under this data generation mechanism.

For the three GEEs fitted, we assessed performance as follows. When data were generated from the unrestricted marginal spatial model, we examined the bias and variability of the estimated slope coefficients relative to the true unpartitioned slope (i.e., second element in $\boldsymbol{\beta}_0$) as well as the true partitioned slope (we discuss the calculation of this shortly). We also calculated 95% Wald-type confidence intervals based on the sandwich covariance matrices in Section 3, and assessed inference on the true unpartitioned slope based on empirical coverage probability (averaged across the 400 simulated datasets) and interval widths. When data were generated from the restricted marginal spatial model, we examined the bias and variability of the estimated slope coefficients relative to the true partitioned slope (i.e., second element in $\boldsymbol{\alpha}_0$). Similar to the unrestricted case, we also calculated 95% Wald-type confidence intervals and assessed inference on the true partitioned slope. To construct the sandwich covariance matrices in all cases, we assumed the true marginal covariance was known e.g., for simulations based on the unrestricted marginal spatial model we set $\tilde{\mathcal{V}}_0 = \boldsymbol{A}_0^{1/2}\boldsymbol{V}_{\text{sp},0}\boldsymbol{A}_0^{1/2}$, and similarly $\tilde{\mathcal{V}}_0 = \boldsymbol{A}_0^{1/2}\boldsymbol{V}_{\text{rsp},0}\boldsymbol{A}_0^{1/2}$ for simulations using the restricted marginal spatial model.

Finally, note that in the case where data are generated from an unrestricted marginal spatial model, while the true unpartitioned effect $\boldsymbol{\beta}_0$ is known by design (the value we set it to is discussed later on), we do not know the true value of the partitioned effect $\boldsymbol{\alpha}_0$. We therefore propose to "estimate" the true $\boldsymbol{\alpha}_0$ as follows: consider the working linear model in (1), but evaluated at the true unpartitioned effect. Given $\text{E}(\boldsymbol{e}) = \boldsymbol{0}$ and $\text{Cov}(\boldsymbol{e}) = \sigma^2_{e,0}\boldsymbol{I}$, then we can rearrange the working linear model to produce a simple estimate of the induced spatial effect as $\hat{\boldsymbol{\rho}}_0 = \boldsymbol{A}_0^{-1/2}(\boldsymbol{y} - \boldsymbol{\mu}_0)$. An estimate of the true partitioned effect then follows as $\hat{\boldsymbol{\alpha}}_0 = \boldsymbol{\beta}_0 + (\boldsymbol{D}_0^\top\boldsymbol{D}_0)^{-1}\boldsymbol{D}_0^\top\hat{\boldsymbol{\rho}}$. This estimate obviously varies across simulated dataset, since the spatial locations and covariates change with each dataset. For the remainder of the simulation study, we treat $\hat{\boldsymbol{\alpha}}_0$ as being the actual true partitioned effect and denote it as $\boldsymbol{\alpha}_0$ for simplicity.

4.1 Count Responses from an Unrestricted Spatial Model

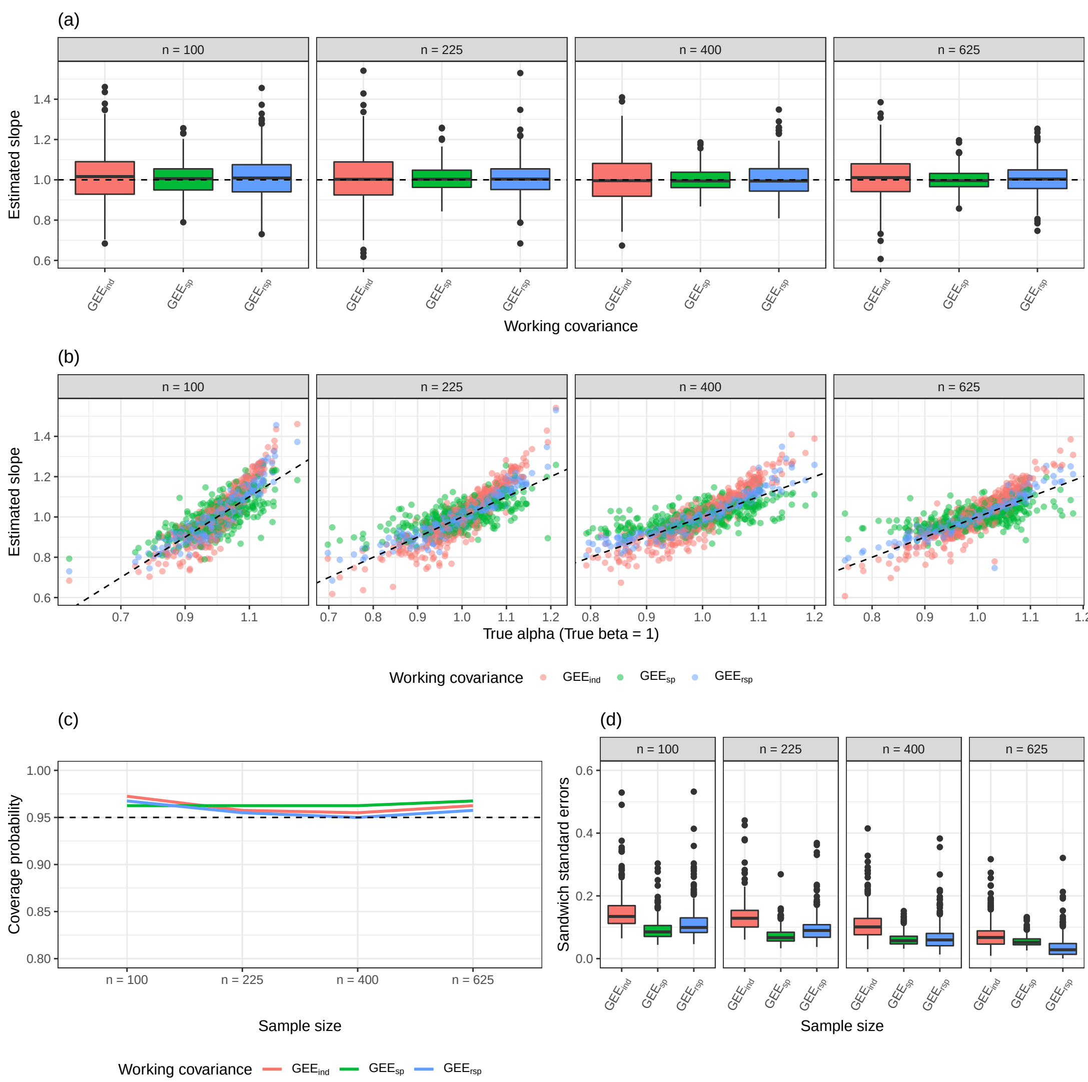
For generating count responses from an unrestricted marginal spatial model, we set the true vector of unpartitioned effects to $\boldsymbol{\beta}_0 = (-1, 1)^\top$ and $g(\cdot)$ to be the log link function. Then we used the algorithm implemented in the `R` package `PoisNor` (Amatya et al., 2019) to generate count responses from an unrestricted marginal spatial model.

From the comparative boxplots of the estimate slope, we see that across all four samples sizes tests, the three GEEs produced estimates centered around the true unpartitioned slope of one. However, GEE_{sp} exhibits much less variability compared to GEE_{ind} and GEE_{rsp} (Figure 1a), while its empirical variance also tended to zero the fastest. This is consistent with the idea that using a working covariance matrix which has a similar, or in this case the same, structure as the true marginal covariance leads to more efficient estimation. Overall then, one may (naively) conclude that changing the working covariance matrix only affects the precision of the estimates, not the quantity each GEE is estimating.

On the other hand, when we examine scatterplots of the estimated slopes versus the true partitioned slope, the evidence of spatial confounding start to become clearer (Figure 1b). In particular, we observe evidence that the GEE_{rsp} is in fact estimating the partitioned rather than the unpartitioned slope. This result empirically confirms one of the main consequences of spatial confounding in GEEs: for a given datset, changing from an unrestricted to a restricted spatial covariance matrix changes the target quantity being estimated in the GEE from an unpartitioned to a partitioned effect. As an aside, one could ask what quantity GEE_{ind} is estimating in this setting; we leave this as an avenue of future research (see Gromping, 1996, for related discussion).

Turning to inference, sandwich-based Wald intervals from all three types of GEEs achieved approximately nominal coverage probability for the true unpartitioned slope (Figure 1c). However, GEE_{ind} and GEE_{rsp} produce much wider confidence intervals compared to GEE_{sp} (Figure 1d). This is again in line with our discussion in Section 3. That is, if the working covariance matrix structure is not aligned with the target of inference, then to ensure valid inference we pay the price of lack of statistical efficiency and subsequently wider confidence intervals.

Figure 1: Results from fitting independent estimating equations (GEE_{ind}), unrestricted spatial GEEs (GEE_{sp}) and restricted spatial GEEs (GEE_{rsp}) to count responses simulated from an unrestricted marginal spatial model. Panel (a) shows boxplots of the estimated slopes, panel (b) shows scatterplots of the estimated slopes against the true unpartitioned effects, where the dashed line is the $y = x$ line, and panels (c) and (d) show the empirical coverage probability and boxplots of interval widths of 95% Wald confidence intervals, respectively, for the true unpartitioned slope.



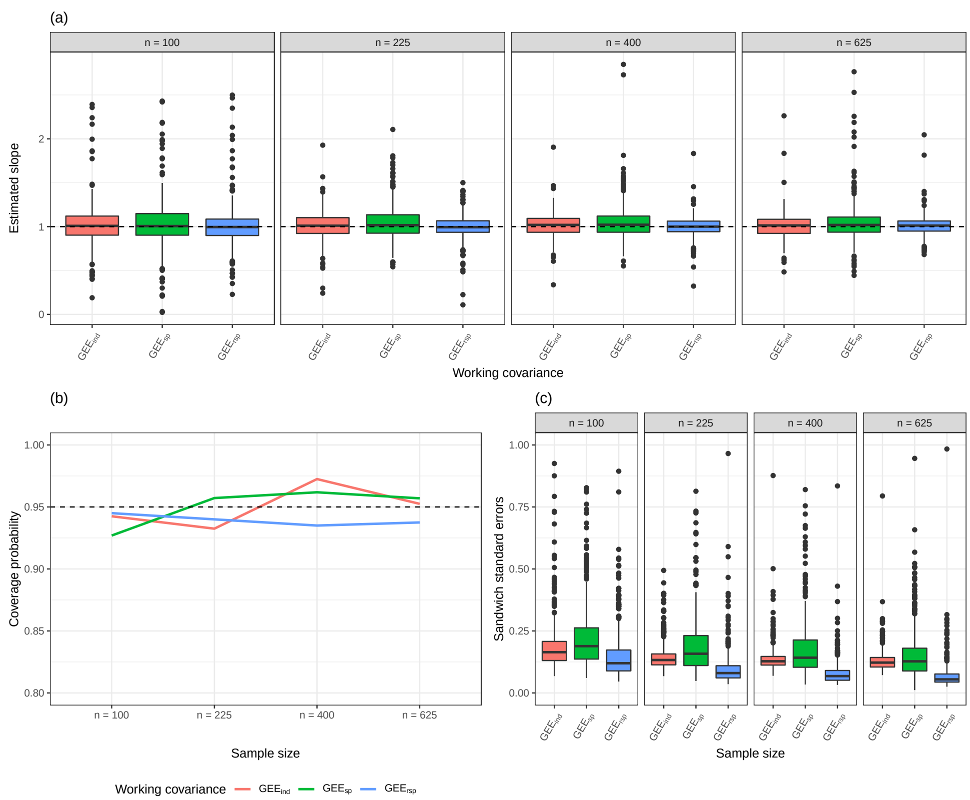
4.2 Count Responses from a Restricted Spatial Model

For generating count responses from a restricted marginal spatial model, we set the true vector of partitioned effects as $\boldsymbol{\alpha}_0 = (-1, 1)^\top$ and $g(\cdot)$ to be the log link function. From the comparative boxplots of the estimate slope, we see that across all four samples sizes, the three GEEs produced estimates again centered around the true partitioned slope of one (Figure 2a). However, this time it was GEE_{rsp} which exhibited the least variability, followed by compared to GEE_{ind} and GEE_{sp}. Regarding inference for the true partitioned slope, while sandwich-based Wald intervals from all three types of GEEs had approximately nominal coverage probability (Figure 2b), it was GEE_{rsp} whivh had consistently the smallest interval widths, followed by GEE_{ind} and GEE_{sp} (Figure 2c). The intervals from the restricted spatial GEE also tended to zero the fastest with increasing sample size. These results are again consistent with the notion that if the working covariance matrix structure is not aligned with the target of inference, then a trade off in statistical efficiency is made in order to ensure valid inference.

In the Supplementary Material, we present numerical results for the cases of the continuous responses and binomial responses, while also presenting further simulations where either the unpartitioned slope (in the case of the unrestricted models) or the partitioned slope (in the case of the restricted models) was set equal to zero. Results from these were very similar to those present above for the case of non-zero effects. We also performed further simulation studies (not presented) to examine scenarios where the covariate and/or the marginal covariance matrix exhibited little spatial structure. Not surprisingly, in such cases the degree of spatial confounding was reduced and so the differences in results between all three types of GEEs fitted was less pronounced.

To summarize, the results from this simulation study demonstrate how spatial confounding can arise in GEEs, and its consequences on estimation and statistical efficiency. Naively, one could examine Figures 1 and 2 and conclude that the results are entirely as expected: the closer the working covariance matrix structure to the true marginal covariance, the more efficient the inference from a GEE is. While this conclusion is correct, it belies how spatial confounding is driving these results. That is, the choice of the spatial working covariance matrix has an effect on both the target quantity that the GEE is estimating and

Figure 2: Results from fitting independent estimating equations (GEE_{ind}), unrestricted spatial GEEs (GEE_{sp}) and restricted spatial GEEs (GEE_{rsp}) to count responses simulated from a restricted marginal spatial model. Panel (a) shows boxplots of the estimated slopes, panel (b) shows scatterplots of the estimated slopes against the true unpartitioned effects, where the dashed line is the $y = x$ line, and panels (c) and (d) show the empirical coverage probability and boxplots of interval widths of 95% Wald confidence intervals, respectively, for the true unpartitioned slope.



the efficiency of the inference, thus reflecting the degree of alignment between this target and the true effect in the underlying data generation mechanism.

5 Application to Pelagic Fish Species Richness Data

As an example of the effects of spatial confounding in a real application of GEEs, we consider data collected as part of the 2016 fall bottom trawl survey by the US Northeast Fisheries Science Centre (Northeast Fisheries Science Center, 2021). Data from the survey are publicly available, and can be accessed along with more details about the survey design at `https://www.fisheries.noaa.gov/inport/item/22560`. As the response, we considered the recorded species richness of 20 pelagic fish species recorded at $n = 605$ spatial locations in the US Northeast Shelf marine ecosystem. That is, each element in the response vector $\boldsymbol{y}$ is a non-negative integer representing the number of different pelagic fish species recorded at that spatial location. Analysis was done by treating the response as an integer-valued count.

We modeled the distribution of species richness as a function of two covariates known to be key environmental drivers of the ecosystem, namely bathymetry (or depth) and sea surface temperature. Furthermore, to account for potential non-linearity in the relationship between species richness and these two covariates, we included both covariates as linear and quadratic terms along with their (linear) interaction. Collectively, all terms were via the `poly` function in `R`. Along with an intercept, this lead to a model matrix $\boldsymbol{X}$ of dimension 605×6. Next, from spatial plots of the species richness along with the two covariates (see Supplementary Material), all exhibited noticeable spatial patterns. Also, a histogram of the species richness suggested no strong evidence of overdispersion, and so in the GEEs below we used a log link function to relate mean species richness to the two covariates, and set $h(\boldsymbol{\mu}) = \boldsymbol{\mu}$ as the marginal mean-variance function.

We began by fitting an IEE to the data, and applying Moran's I test (Moran, 1950) to the corresponding Pearson residuals. The resulting test showed clear statistical evidence of residual spatial correlation in the data (p-value < 0.001). Next, we proceeded to fit both the unrestricted spatial GEE and restricted spatial GEE, and constructed 95% sandwich-based Wald confidence intervals for all three GEEs using the approach discussed in Section 3; see the Supplementary Material for details on estimation of the marginal covariance matrix $\tilde{\boldsymbol{V}}_0$. The three GEEs produced varying conclusions in terms of which coefficients were statistically different from zero (Table 1). For example, all three presented clear evidence of a strong negative effect for the linear effect of sea surface temperature (with similar magnitude of the estimated coefficients), as well as no evidence of a linear effect for depth (although the magnitudes and signs of the estimated coefficients differed substantially between the three GEEs). On the other hand, only the unrestricted spatial GEE exhibited evidence of the quadratic effect for depth being statistically different from zero, while the unrestricted and restricted spatial GEEs but not the IEE showed evidence

Table 1: Estimated regression coefficients and 95% sandwich-based Wald confidence intervals (in parentheses) for the IEE (GEE_{ind}), unrestricted spatial GEE (GEE_{sp}), and restricted spatial GEE (GEE_{rsp}), fitted to pelagic fish species richness data. The two covariates included in all models were sea surface temperature (Temp) and bathymetry (Depth). Confidence intervals that do not contain zero are bolded.

Covariate	GEE_{ind}	GEE_{sp}	GEE_{rsp}
Intercept	**1.428 (1.274, 1.581)**	**1.451 (1.307, 1.594)**	**1.438 (1.349, 1.528)**
Depth	-1.773 (-3.907, 0.360)	0.722 (-0.811, 2.254)	-0.415 (-1.861, 1.030)
$Depth^2$	-0.091 (-1.657, 1.475)	**-1.327 (-2.483, -0.171)**	-0.764 (-1.723, 0.196)
Temp	**-2.062 (-3.824, -0.300)**	**-1.849 (-2.809, -0.888)**	**-1.980 (-3.093, -0.868)**
$Temp^2$	1.032 (-0.453, 2.517)	**1.159 (0.263, 2.056)**	**1.059 (0.159, 1.959)**
Depth:Temp	**-2.155 (-4.067, -0.244)**	-0.729 (-1.690, 0.232)	**-1.449 (-2.427, -0.472)**

of a strong positive effect of the quadratic effect of temperature. Regarding the interaction term between depth and temperature, only the IEE and restricted spatial GEE found clear evidence of an effect. Overall, the differing conclusions suggests possible evidence of spatial confounding in this data. Indeed, across all the six terms included in the mean model, it was interesting to observe that the magnitude of the estimated coefficients from the restricted spatial GEE was always between that of the IEE and unrestricted spatial GEE. This was generally consistent with our simulation results for Poisson responses when the data were generated from an unrestricted spatial model (see Section 4 above), and with the impact of spatial confounding on GEEs.

6 Discussion

The findings of this article have important implications for the use of GEEs in spatial analysis. To quote the recent work of Bible et al. (2019) who examined confounding for GEEs in a different context: "In practice, analysts rarely check for the misspecification of the working correlation but directly apply GEEs ..., falsely hoping that the sandwich variance estimator corrects for the correlation." In the presence of spatial confounding, we have demonstrated that GEEs can estimate different quantities depending on the choice of the working correlation matrix, and how spatial confounding affects efficiency of inference in a GEE based on the extent to which the choice of the working correlation is aligned with the inferential target. While the proposed restricted spatial GEE is an attempt to alleviate spatial confounding in this context, this does not necessarily translate to better performance all the time. Rather, we hope that this article will bring about a more cautious approach in the way GEEs are applied to spatial data: instead of “falling back” on its robustness to misspecification, in the presence of spatial confounding practitioners need to be more circumspect, using *a-priori* background knowledge along with careful consideration of the research questions and data generation process to determine the target quantity they are interested in, and from this decide on the form of the working covariance matrix to use. Moreover, we concur with Khan and Calder (2020) among others that more theoretical and empirical research needs to be done to better understand when it is appropriate to use methods that adjust versus do not adjust for spatial confounding.

It is interesting to compare our work with that of Hanks et al. (2015) and Khan and Calder (2020), who showed in the context of Bayesian spatial mixed models that one is almost always better off fitting a non-spatial model rather than a restricted spatial model, because the latter tends to suffer from severe undercoverage and inflated Type-S errors (the Bayesian equivalent of Type I errors). Khan and Calder (2020) showed this occurs because restricted spatial regression effectively amounts to using an overfitted fixed effects model, which reduces the posterior variance inappropriately such that covariates are deemed to be statistically significant even if they are truly unimportant. By contrast, we did not observe evidence in our simulations of such undercoverage for restricted spatial GEEs e.g., our confidence intervals were relatively well-calibrated irrespective of the working covariance matrix used. While a direct comparison between spatial mixed models and GEEs is not straightforward, we believe a large part of why such undercoverage did not occur is due to the use of the sandwich-based standard errors. That is, because GEEs are built on the idea that the working correlation may not be equal to the true marginal spatial correlation, then a necessary correction of the standard errors is made to ensure undercoverage will not occur, at least asymptotically. As explained in Section 3, the sandwich standard error adjusts for the misalignment between the type of GEE being used and the target quantity of inference. Such an adjustment does not occur in the Bayesian spatial mixed models explored by Khan and Calder (2020), although interestingly Hanks et al. (2015) had in fact earlier proposed sandwich-based standard errors for such mixed models, and empirically showed that it resolves the problem of undercoverage and Type-S error inflation, albeit it may end up being too conservative. A further avenue of research is to compare our developments with those for autologistc models for spatial data, where marginal intepretations of covariate effects are also commonly of interest and for which some research on spatial confounding has been done (e.g., Caragea and Kaiser, 2009; Hughes, 2014).

As a concluding point, in the Supplementary Material we provide an extensive discussion on the large sample properties of unrestricted and restricted spatial GEEs in the presence of spatial confounding (see also Zimmerman and Ver Hoef, 2021, for related research). In brief, the form of spatial confounding we have studied in this article arises due to a finite sample correlation between spatially structured covariates $\boldsymbol{X}$ and the induced spatial effect $\boldsymbol{\rho}$ (similar to that of Hanks et al., 2015). It is also possible for spatial confounding in GEEs to occur in a way that persists with increasing sample size, and we leave investigation of this as an avenue of future research (see also Paciorek, 2010; Dupont et al., 2021, in the context of spatial mixed models).

SUPPLEMENTARY MATERIAL

Additional discussion: Appendix A presents details for estimating the unrestricted and restricted spatial GEEs, Appendix B discusses the constant mean-variance function case in more detail, and Appendix F discusses large sample behaviour.

Additional Simulation and Application Results: Appendices C and D presents further numerical results for the simulation study in Section 4, and Appendix E presents further exploratory plots for the application to the pelagics species richness data.

References

Adegboye, O. A., D. H. Leung, and Y.-G. Wang (2018). Analysis of spatial data with a nested correlation structure. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 67*, 329–354.

Albert, P. S. and L. M. McShane (1995). A generalized estimating equations approach for spatially correlated binary data: applications to the analysis of neuroimaging data. *Biometrics 51*, 627–638.

Amatya, A., H. Demirtas, and R. Gao (2019). *PoisNor: Simultaneous generation of multivariate data with Poisson and normal marginals*. R package version 1.3.1.

Bible, J., P. S. Albert, B. G. Simons-Morton, and D. Liu (2019). Practical issues in using generalized estimating equations for inference on transitions in longitudinal data: What is being estimated? *Statistics in medicine 38*, 903–916.

Caragea, P. C. and M. S. Kaiser (2009). Autologistic models with interpretable parameters. *Journal of Agricultural, Biological, and Environmental Statistics 14*, 281–300.

Crouchley, R. and R. Davies (1999). A comparison of population average and random-effect models for the analysis of longitudinal count data with base-line information. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 162*, 331–347.

Dupont, E., S. N. Wood, and N. Augustin (2021). Spatial+: a novel approach to spatial confounding. *Biometrics In Press*.

Gromping, U. (1996). A note on fitting a marginal model to mixed effects log-linear regression data via GEE. *Biometrics 52*, 280–285.

Hanks, E. M., E. M. Schliep, M. B. Hooten, and J. A. Hoeting (2015). Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics 26*, 243–254.

Hefley, T. J., M. B. Hooten, E. M. Hanks, R. E. Russell, and D. P. Walsh (2017). The Bayesian group lasso for confounded spatial data. *Journal of Agricultural, Biological and Environmental Statistics 22*, 42–59.

Hodges, J. S. and B. J. Reich (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician 64*, 325–334.

Hughes, J. (2014). ngspatial: A Package for Fitting the Centered Autologistic and Sparse Spatial Generalized Linear Mixed Models for Areal Data. *R Journal 6*, 81–95.

Hughes, J. and M. Haran (2013). Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 75*, 139–159.

Khan, K. and C. A. Calder (2020). Restricted Spatial Regression Methods: Implications for Inference. *Journal of the American Statistical Association In press*.

Liang, K.-Y. and S. L. Zeger (1986). Longitudinal data analysis using generalized linear models. *Biometrika 73*, 13–22.

Lin, P.-S. and M. K. Clayton (2005). Analysis of binary spatial data by quasi-likelihood estimating equations. *The Annals of Statistics 33*, 542–555.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika 37*, 17–23.

Nobre, W. S., A. M. Schmidt, and J. B. Pereira (2021). On the effects of spatial confounding in hierarchical models. *International Statistical Review 89*, 302–322.

Northeast Fisheries Science Center (2021). Fall bottom trawl survey. https://www.fisheries.noaa.gov/inport/item/22560.

Paciorek, C. J. (2010). The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical Science 25*, 107–125.

Papadogeorgou, G. (2021). Discussion of the manuscript: Spatial+ a novel approach to spatial confounding.

Reich, B. J., S. Yang, Y. Guan, A. B. Giffin, M. J. Miller, and A. Rappold (2021). A review of spatial causal inference methods for environmental and epidemiological applications. *International Statistical Review In press*.

Zimmerman, D. L. and J. M. Ver Hoef (2021). On deconfounding spatial confounding in linear models. *The American Statistician In press*.