

Estimating the Effect of Spatial Dependency by modeling a Hierarchical Bayesian ICAR Prior on Archaeological Site Location Data

Matthew D. Harris

 md_harris

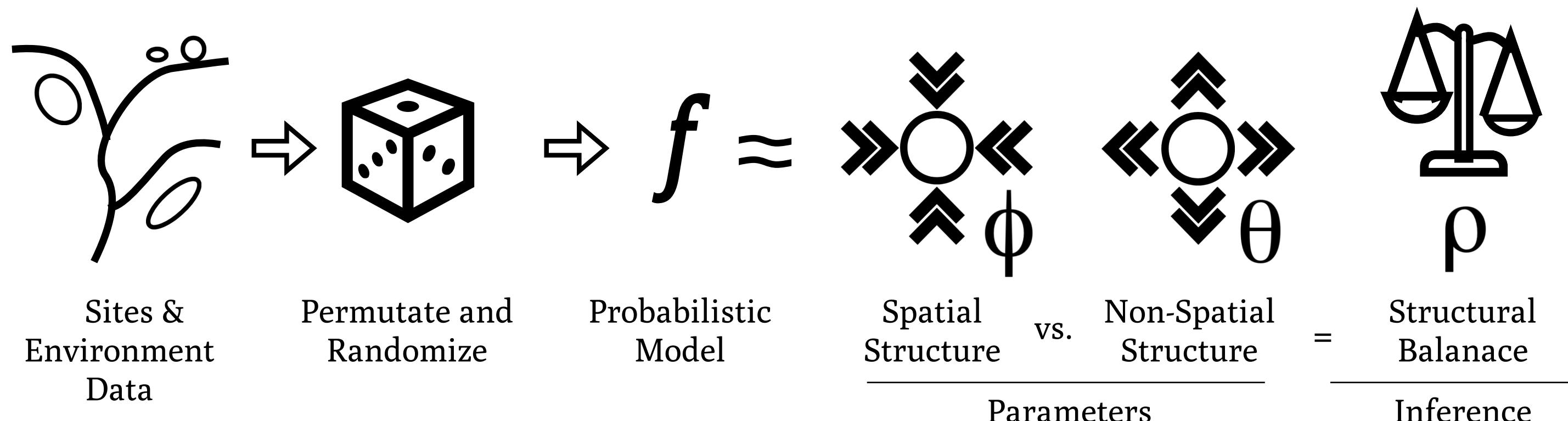
Mary E. Lennon

 MaryELennon

AECOM
aecom.com

Abstract

Q: Should spatial diffusion be used in landscape scale archaeology models?

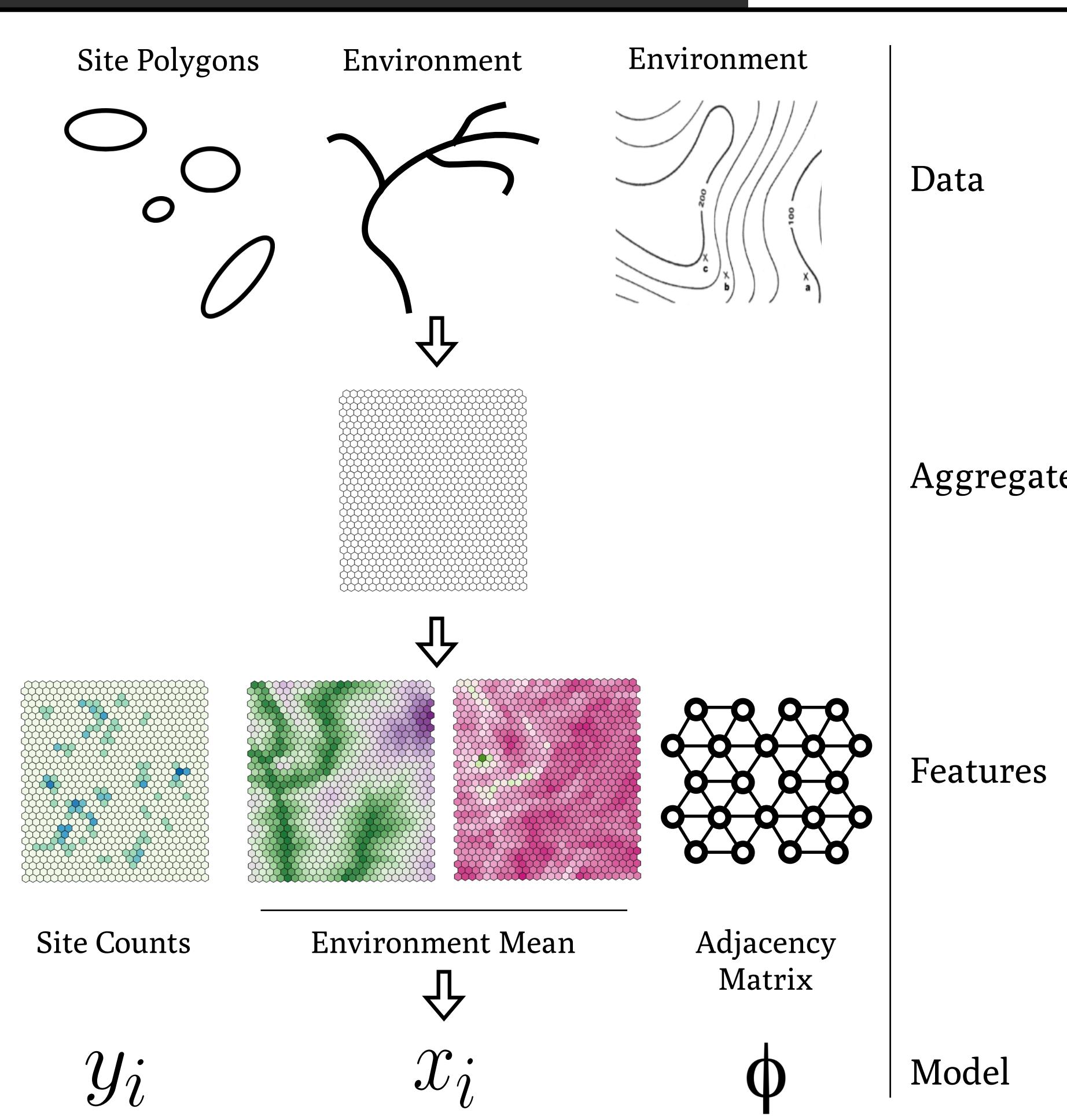


A: Yes! However, the effect is hard to discern from environmental autocorrelation. Models such as this help find a balance between the two.

This research seeks to find a method for determining an appropriate level of spatial diffusion to include in an archaeological predictive model (APM). APMs typically do not directly model spatial diffusion, because archaeologists do not want to assume that sites manifest in the vicinity of other sites without a cultural process or that a region has low archaeological sensitivity because none have yet been recorded in a near-by locale. However, substantial archaeological research relies on spatial diffusion to explain similarities in sites over space. Further, it is prudent to consider some level of spatial autocorrelation, the measure of similarity between observations, due to its crucial influence on cultural geography and Tobler's First law of Geography

The research presented here "learns" an appropriate level of spatial diffusion from known environmental and archaeological data and uses that to understand these related processes. This was accomplished through a Bayesian Hierarchical Poisson Model using an ICAR prior - generally referred to as a Besag, York and Mollié (BYM) model (Besag et al. 1991).

Data Creation



The data used in this study include archaeological site locations and four environmental predictor variables. The count of sites and the mean of all other environmental variables were aggregated to a 500-meter "fishnet" hexagonal grid. This analytical unit allows for: 1) more control over the analysis; 2) efficient for spatial coverage; and 3) its lattice characteristics translate to "better sampling efficiency" and reduced bias from edge effects (Lewin-Koh 2018). From these aggregations critical model parameters are created- the count of sites as the independent variable, predictor features, and adjacency matrix.



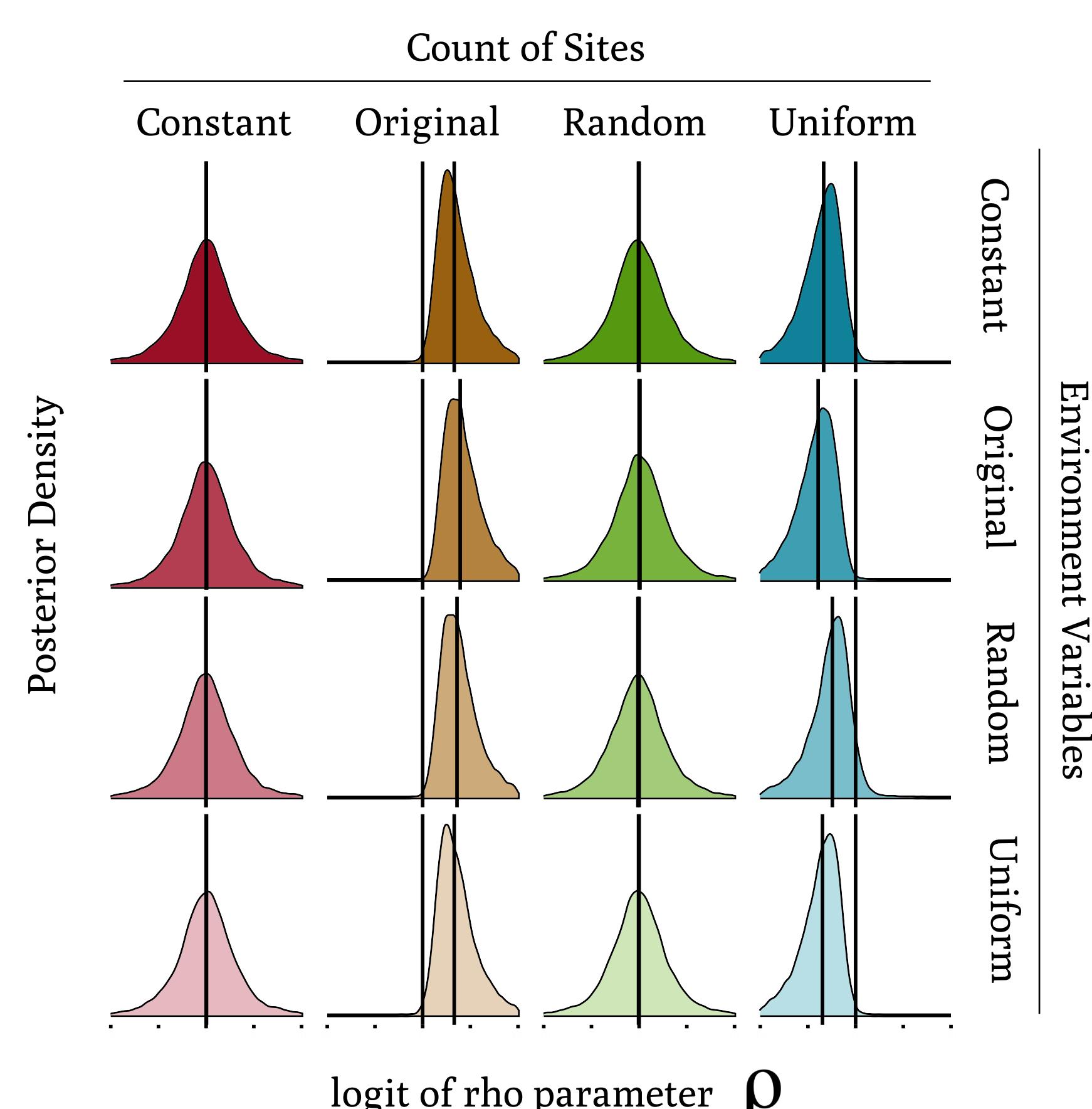
Resources:

- Besag Julian. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B (Methodological)*, 192-236.
- Besag J, York Y, Mollié A. (1991). Bayesian Image Restoration With Two Applications In Spatial Statistics. *Ann Inst Statist Math*. Vol. 43:1-5. 10.1007/BF00164666
- Lewin-Koh N. (2018). Hexagon Binning: An Overview. CRAN. https://cran.r-project.org/web/packages/hectbin/vignettes/hexagon-binning.pdf (accessed 26, March 2019).
- Morris, Mitzi. (2019). Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data. Website: https://mc-stan.org/users/documentation/case-studies/icar_stan.html (accessed 26, March 2019).
- Riebler, Andrea, Sigrunn H. Sørbye, Daniel Simpson, and Håvard Rue. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical methods in medical research* 25 no. 4: 1145-1169.
- Simpson DP, Rue H, Martínez TG, et al. (2015). Penalising model component complexity: a principled, practical approach to constructing priors. arXiv preprint arXiv:1403.4630v4. http://arxiv.org/pdf/1403.4630v4.pdf (accessed 26, March 2019).

Results

Permutations

To better understand the behavior and characteristics of the BYM model, in this setting, the site count and environmental features are permuted in three different ways, each indicating a different level of correlation; a constant value of one, random values between (0, 1), and uniformly random between their original min/max. The plot below depicts the posterior distribution of the parameter rho, the balance of both spatially structured and unstructured data, for each permutations.



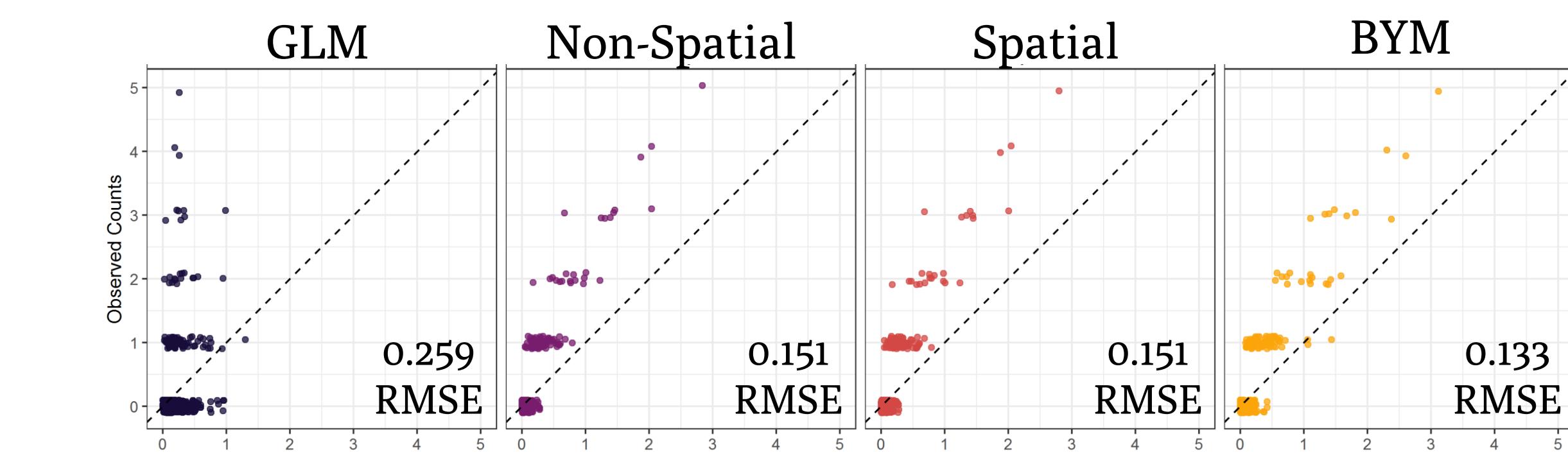
The results indicate that much of the variation in site counts can be explained by knowing which fishnet cells are connected via the adjacency matrix – highlighting the importance of spatial autocorrelation. Further, that more variance is explained by knowing both the spatial and non-spatial random effects rather than knowing either alone. The difference between the spatial and non-spatial random effects may indicate the degree of spatial interaction.

Model Comparison

When validating a model, it is important to compare its nested specifications. In this case these are a Poisson GLM with no random effects, a model with only non-spatial random effects, a model with only random spatial effects, and the BYM model with both random effects. The BYM model offers an improved explanation over all the other models.

	elpd_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
BYM	0.0	-292.9	22.1	78.8	7.6	585.8	44.3
Non-Spatial	-55.9	-348.8	27.4	116.3	12.6	697.6	54.9
Spatial	-56.1	-349.0	27.4	116.8	12.5	698.0	54.8
GLM	-79.7	-372.6	32.3	6.7	1.1	745.3	64.7

Comparing the Expected Log Predictive Density (ELPD), the BYM model is ~2 standard deviations better and has fewer parameters. The predictions have a lower RMSE (0.133) compared to the other models.



Model

Likelihood	$y_i \sim Poisson(\log(E)\Psi_i)$
Process Model	$\log(\Psi_i) = \beta_0 + \chi\beta_i + \theta + \phi$
Convolution	$\theta + \phi = \sigma(\sqrt{1 - \rho}\theta + \sqrt{\rho}\phi)$
iCAR Prior	$\phi = -\frac{1}{2}(\sum_{i \sim j}(\phi_i - \phi_j)^2)$
Priors	$\beta_0, \beta_i, \theta, \sigma \sim N(0, 1)$ $\rho \sim beta(0.5, 0.5)$

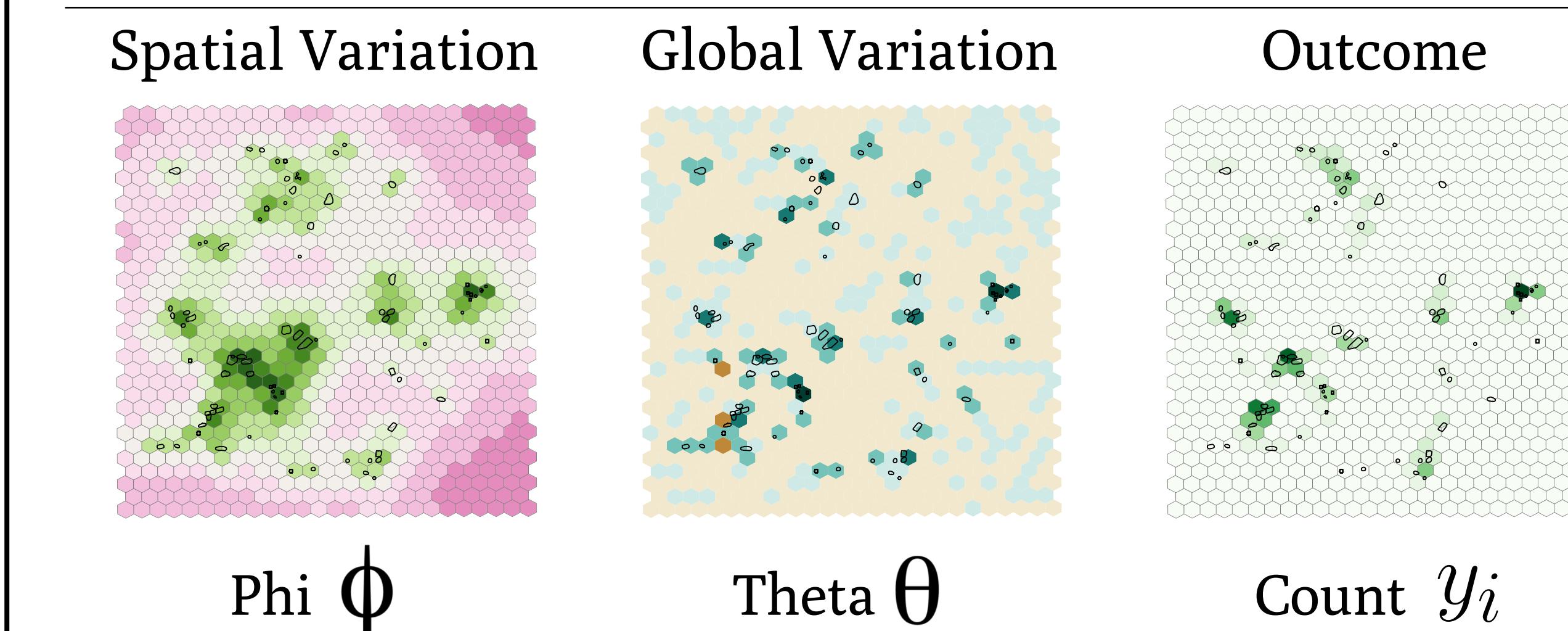
The model built for this study, a BYM with the addition of rho (Riebler et al 2016, Simpson 2014), was computed using Hamiltonian Monte Carlo with Stan in R and includes the following components:

- Poisson likelihood to model count data, such as the number of sites;
- A linear multivariate process model for linking Poisson site counts to predictor variables;
- A convolution model to balance spatial and non-spatial variance;
- An ICAR prior to model the spatial realtionship via the adjacency matrix;
- Wealy informative priors over model parameters and uniform prior on rho.

Findings

These results provisionally show that:

- Bayesian spatial modeling is applicable to archaeological data;
- The parameter rho indicates that these data exhibit spatial structure that contributes to understanding the count distribution;
- The BYM specification has promise to illuminate these structures;
- Model comparison indicates that random effects can be informative for archaeological data and hierarchical models;
- The spatial random effects learned for this model can be used to understand the potential for site sensitive areas and the justifiable amount of spatial smoothing given the data and other random effects.



- Future directions of this line of research:
- Apply to a variety of archaeological site datasets with known and unknown settlement dynamics;
 - Analyze the meaning of model priors in terms of archaeological site data;
 - Explore additional spatial and temporal random effects;
 - Test other adjacency structures;
 - Incorporate zero-inflation into the model specification.