TMA4250 - Project 3

Jakob Heide and Bendik Waade

2024-03-11

Part 1

We consider two subdivisions of Nigeria, referred to as admin1 (37 areas) and admin2 (775 areas). Using each area as a node and connecting areas that share a border gives rise to a graph for each of the subdivisions, which we refer to as the admin1 graph and the admin2 graph.

1a)

The Besag model is an improper Gaussian Markov random field (GMRF) with respect to a connected graph, and it is defined through the probability density function

$$f(\boldsymbol{x};\tau) \propto \tau^{\frac{n-1}{2}} \exp(-\frac{\tau}{2}\sum_{i\sim j}(x_i - x_j)^2),\ \boldsymbol{x} \in \mathbb{R}^n, \tag{1}$$

where $\tau > 0$ is the precision parameter. To construct the improper precision matrix of the Besag model, we can use the graph matrix and the method of matching coefficients. We want an expression on the form

$$\exp(-\frac{1}{2}\boldsymbol{x}^T\mathbf{Q}\boldsymbol{x}) \Leftrightarrow \exp(-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i \mathrm{Q}_{ij} x_j) \tag{2}$$

To match the coefficients in Equation 1 and Equation 2, we note the different cases that arise:

- When there is no relation between i and j, there is no cross term between x_i and x_j in Equation 1, and so $\mathrm{Q}_{ij} = 0$.
- When there is a relation between i and j, we get two terms in Equation 2 of the form $x_i\mathrm{Q}_{ij}x_j$, and one term in Equation 1 of the form $-2\tau x_i x_j$, i.e. $Q_{ij} = -\tau$.
- When $i = j$, there is one term of the form $x_i^2\mathrm{Q}_{ii}$ in Equation 2. If node i has three neighbours in total, there will be three τx_i^2 terms in Equation 1. Thus, $Q_{ii} = \tau|ne(i)|$.

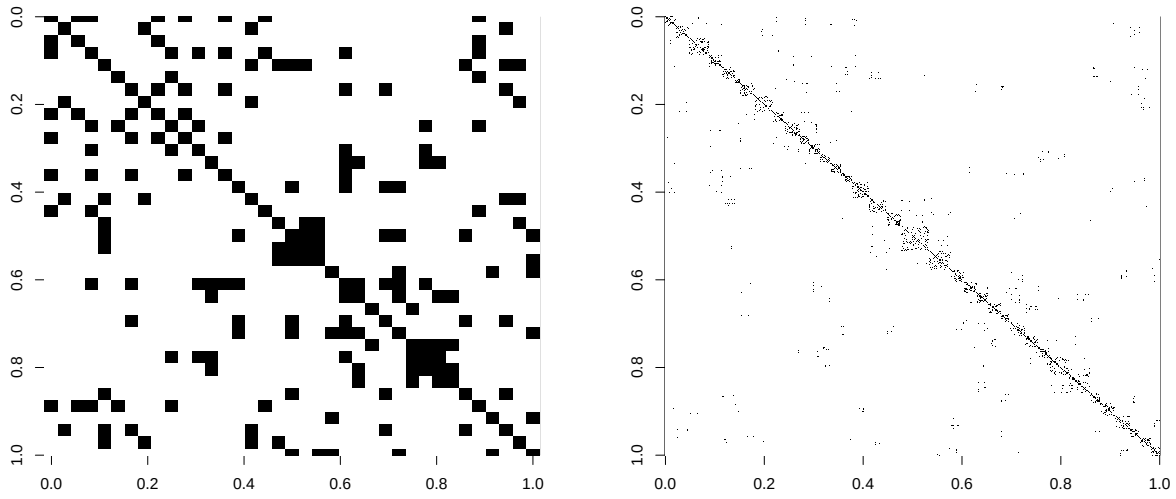
This means that we can construct the precision matrix by setting all the non-negative elements of the graph matrix equal to -1, setting the diagonal equal to $\{|ne(i)|\}_{i=1,2,\dots,n}$, and multiplying all entries by the precision parameter τ. We denote by $\mathbf{R}_1$ and $\mathbf{R}_2$ the structure matrices of the admin1 and admin2 areas, respectively, such that $\mathbf{Q}_1 = \tau_1\mathbf{R}_1$ and $\mathbf{Q}_2 = \tau_2\mathbf{R}_2$. The dimension of the precision matrices are equal to the number of areas in each subdivision, that is, $\mathbf{Q}_1 \in \mathbb{R}^{37\times37}$ and $\mathbf{Q}_2 \in \mathbb{R}^{775\times775}$. We note that if we add all the columns or rows in one of the precision matrices together, we get a vector of zeroes. This indicates that the precision matrices have rank $n - 1$.

We calculate the precision matrices for the admin1 graph and the admin2 graph. The percentage of non-zero elements in each precision matrix is shown below.

```
## The precision matrix of the admin1 graph has  15.26662 % nonzero entries.

## The precision matrix of the admin2 graph has  0.8792508 % nonzero entries.
```

We display the sparsity pattern for each the precision matrices.



(a) The admin1 graph

(b) The admin2 graph

Figure 1: Sparsity patterns for the precision matrices of the Besag model.

1b)

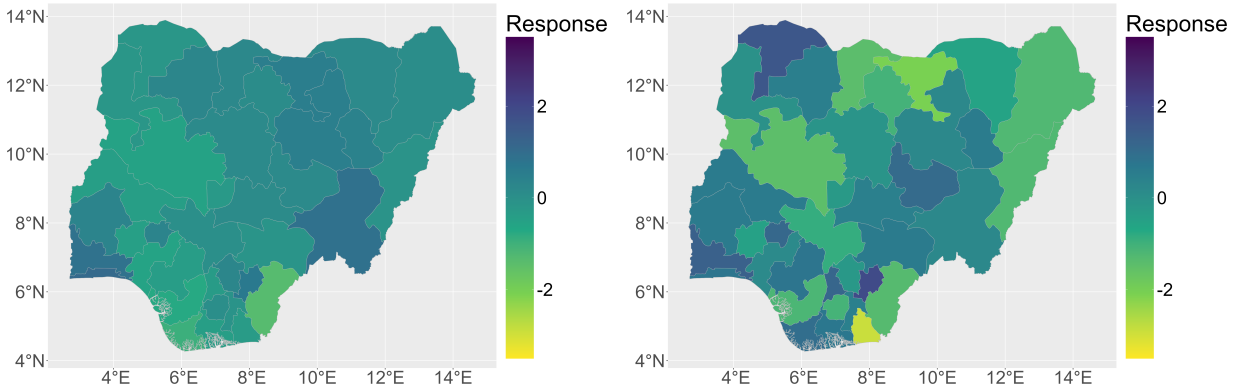
To simulate from the Besag model, it only makes sense to simulate from the proper part of the GMRF. The Besag model is an intrinsic GMRF of first order, which means that

$$\mathbf{Q}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} = \mathbf{0}$$

. This gives a sum-to-zero constraint on the samples. The algorithm used is presented in short here:

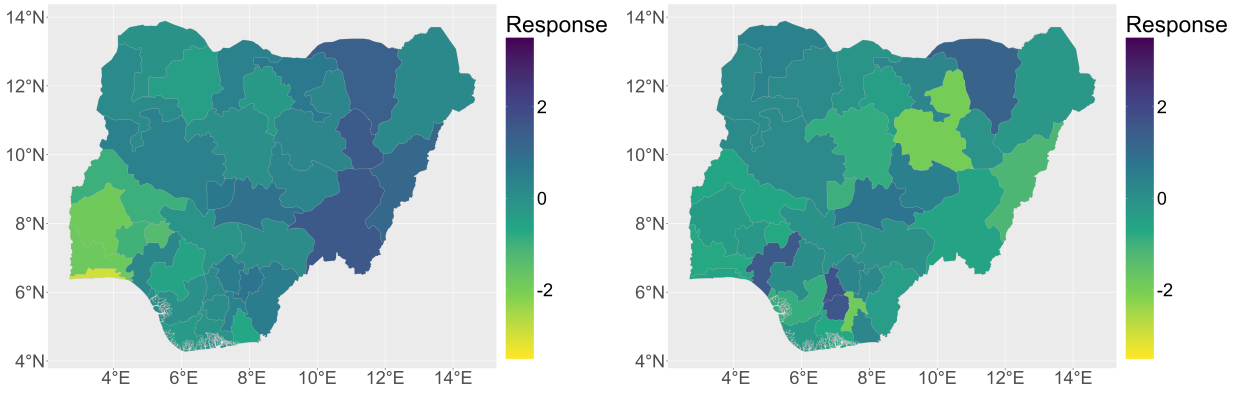
1. Set $\tilde{\mathbf{Q}} = \mathbf{Q} + \epsilon \mathbf{I}_n$ and compute $\tilde{\mathbf{L}}$ such that $\tilde{\mathbf{Q}} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T$
2. Sample $\boldsymbol{z} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{I}_n)$
3. Solve $\tilde{\mathbf{L}}\boldsymbol{v} = \boldsymbol{z}$
4. Compute $\boldsymbol{x} = \boldsymbol{v} - \text{mean}(\boldsymbol{v})\begin{pmatrix}1 & \cdots & 1\end{pmatrix}^T$

The sum-to-zero constraint becomes apparent in step 4 of the algorithm. The resulting sample is a sample from the proper part of the GMRF. We generate two realizations from the Besag model, and two realizations from the multivariate standard Gaussian distribution.



(a) Besag, realization 1

(b) Normal, realization 1



(c) Besag, realization 2

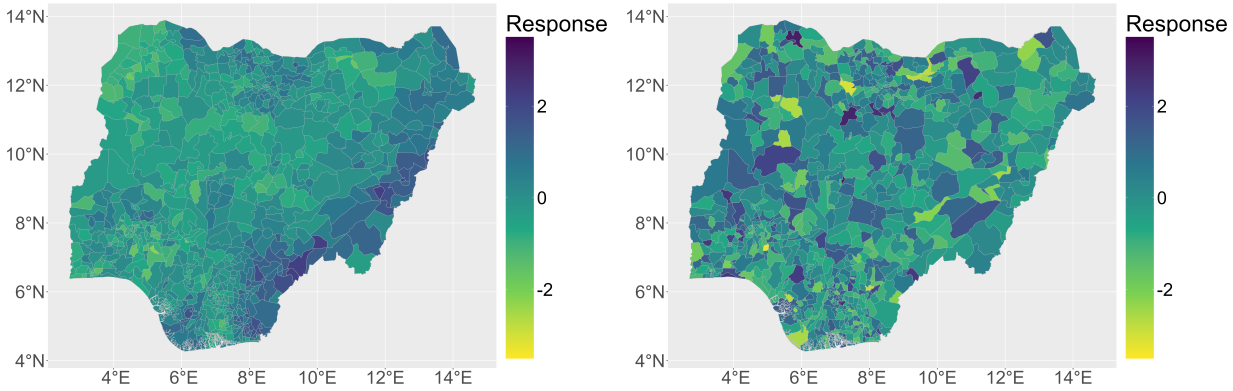
(d) Normal, realization 2

Figure 2: Two realizations of the Besag model (left) and of the multivariate normal (right) for the admin1 area.

In Figure 2, we see that the normal realizations indicate no correlation between the areas. The Besag realizations clearly indicate some spatial correlation, as the values in different areas are not so far apart, especially when the areas are close together. All of the realizations seem to have a mean value around zero.

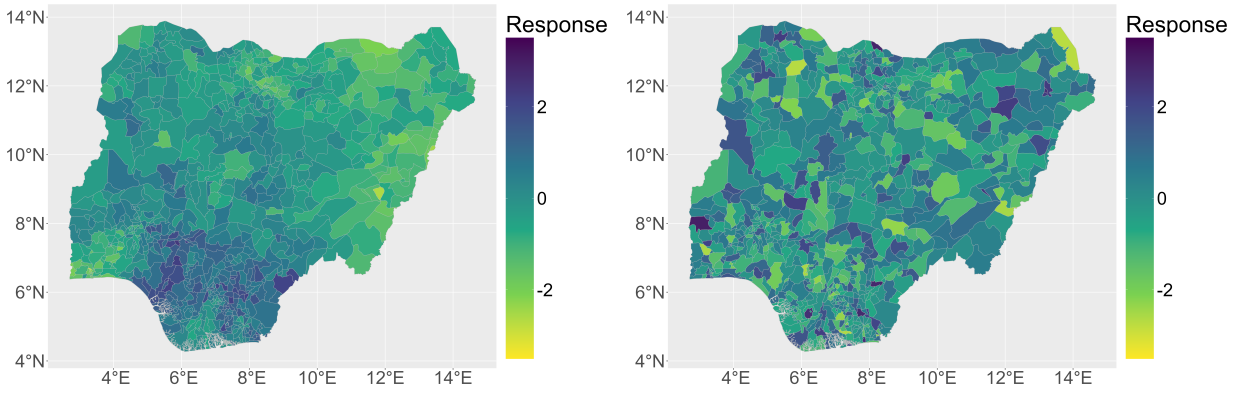
1c)

We repeat the previous exercise, but for the admin2 area.



(a) Besag, realization 1

(b) Normal, realization 1



(c) Besag, realization 2

(d) Normal, realization 2

Figure 3: Two realizations of the Besag model (left) and of the multivariate normal (right).

In Figure 3, we see the realizations of the Besag model and the multivariate Gaussian on the admin2 area. The spatial correlation of the Besag model becomes even more clear - when there are many small areas close together, the correlation creates a smoothing effect. In addition, the randomness of the Gaussian distribution is also perhaps more pronounced, as small areas can suddenly get spikes in values.

1d)

We simulate 100 realizations and compute the empirical variance in each of the areas in the admin2 subdivision.

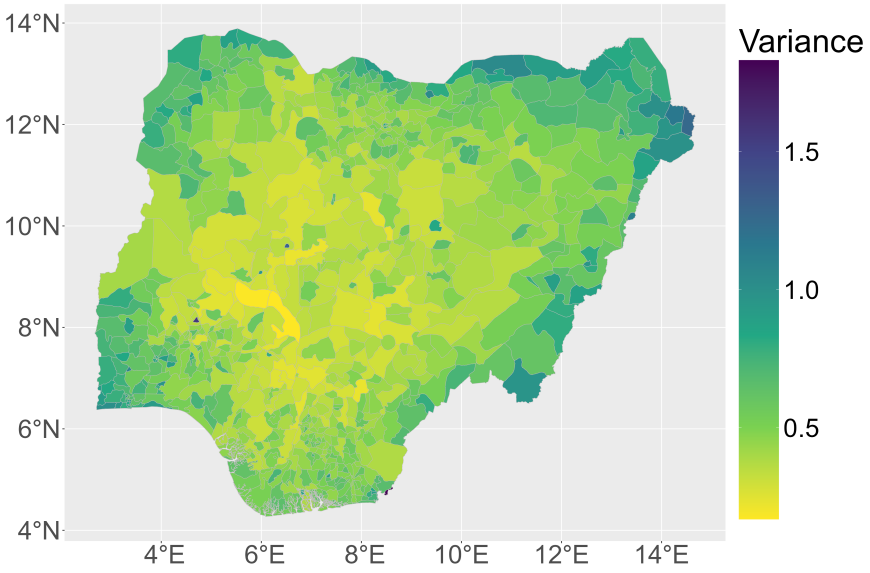


Figure 4: Empirical variance from 100 realizations of the Besag model.

Figure 4 shows clear signs of non-stationarity for the Besag model. Areas that are close to the boundary have fewer neighbours, and so they are allowed to fluctuate more in value, i.e. they have higher variance than areas towards the center.

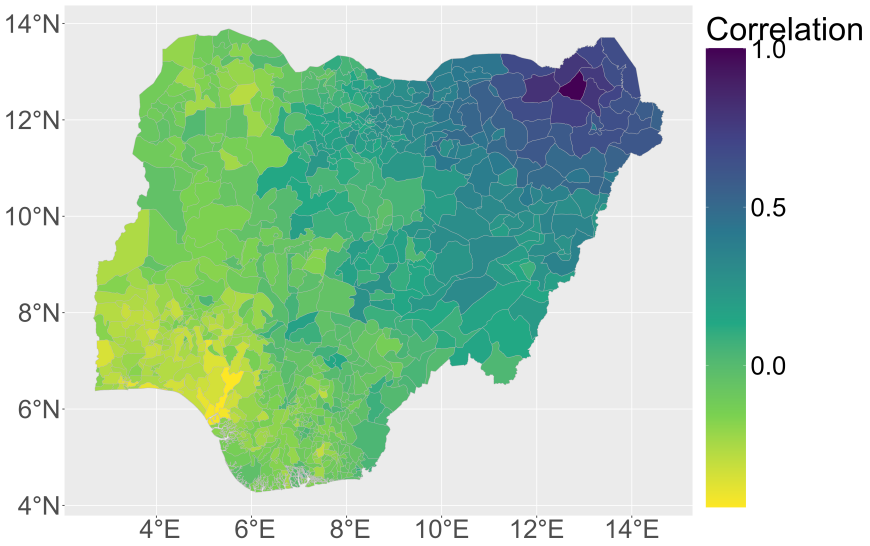


Figure 5: Empirical correlation between Gubio (area 150) and the rest.

Figure 5 displays the correlation between the Gubio area (to the top right) and the rest of the areas. The Besag model satisfies the pairwise Markov property, but since the GMRF is a positive distribution, the Besag model also satisfies the local and global Markov properties. This means that the value in Gubio is correlated with all the other areas. The correlation decreases the further away we are from Gubio. We have negative correlations in areas far from Gubio due to the sum-to-zero constraint.

Part 2

We consider the estimation of vaccine coverages for children in the 37 admin1 areas. Let p_a be the true number of vaccinated children, for $a = 1, \ldots, 37$, and $\hat{P}_a$ be the estimator for p_a. We assume that

$$\text{logit}(\hat{P}_a) \sim \mathcal{N}(\text{logit}(p_a), V_a), \;\; a = 1, \ldots, 37,$$

where $V_1, \ldots, V_{37}$ are known variances and $\hat{P}_1, \ldots, \hat{P}_{37}$ are independent. Let $\boldsymbol{X} = (\text{logit}(P_1), \ldots, \text{logit}(P_{37}))^T$ and $\boldsymbol{Y} = (\text{logit}(\hat{P}_1), \ldots, \text{logit}(\hat{P}_{37}))^T$.

We display the observed proportions of vaccinations in Figure 6.

2a)

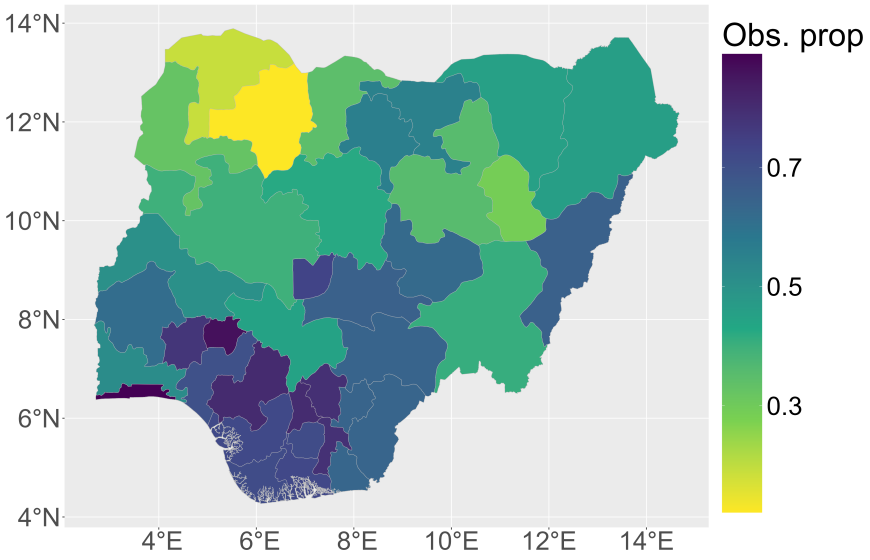


Figure 6: Observed proportion of children vaccinated in each area.

The goal of our model is to estimate the true proportions based on the observed proportions. In Figure 6, the observed proportions are higher towards the lower left areas, and smaller towards the upper left. This indicates some spatial correlation, so including some spatial correlation in our model to reduce uncertainty seems reasonable.

2b)

We have that $Y_a \sim N(\text{logit}(p_a), V_a),\ a = 1, \ldots, 37$, but we view $X_a = \text{logit}(p_a)$ as a stochastic variable. That is, the distribution of the conditional $\boldsymbol{Y}|\boldsymbol{X}$ is

$$\boldsymbol{Y}|\boldsymbol{X} = \boldsymbol{x} \sim \mathcal{N}_{37}(\boldsymbol{x}, \mathbf{V}),$$

where $\mathbf{V}$ is a diagonal matrix with elements $V_a,\ a = 1, \ldots, 37$.

We want to find the distribution of $\boldsymbol{X}|\boldsymbol{Y}$, when we assume a priori that $\boldsymbol{X} \sim \mathcal{N}_{37}(\mathbf{0}, \sigma^2\mathbf{I}_{37})$. We write $\mathbf{Q}_1^{-1} = \sigma^2\mathbf{I}_{37}$ and $\mathbf{Q}_2^{-1} = \mathbf{V}$ in the following.

The joint distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be written as

$$\begin{aligned} f(\boldsymbol{x},\boldsymbol{y}) &= f(\boldsymbol{x})f(\boldsymbol{y}|\boldsymbol{x}) \\ &\propto \exp(-\frac{1}{2}\boldsymbol{x}^T\mathbf{Q}_1\boldsymbol{x}) \cdot \exp(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\mathbf{Q}_2(\boldsymbol{y}-\boldsymbol{x})) \\ &= \exp(-\frac{1}{2}\boldsymbol{x}^T(\mathbf{Q}_1+\mathbf{Q}_2)\boldsymbol{x} - 2\boldsymbol{x}^T\mathbf{Q}_2\boldsymbol{y} + \boldsymbol{y}^T\mathbf{Q}_2\boldsymbol{y}) \end{aligned}$$

We know that the joint density should be proportional to an expression on the form

$$\begin{aligned} f(\boldsymbol{x},\boldsymbol{y}) &\propto \exp(-\frac{1}{2}\begin{pmatrix}\boldsymbol{x}^T & \boldsymbol{y}^T\end{pmatrix}\begin{pmatrix}\mathbf{Q}_{\boldsymbol{XX}} & \mathbf{Q}_{\boldsymbol{XY}} \\ \mathbf{Q}_{\boldsymbol{YX}} & \mathbf{Q}_{\boldsymbol{YY}}\end{pmatrix}\begin{pmatrix}\boldsymbol{x} \\ \boldsymbol{y}\end{pmatrix}) \\ &= \exp(-\frac{1}{2}(\boldsymbol{x}^T\mathbf{Q}_{\boldsymbol{XX}}\boldsymbol{x} + 2\boldsymbol{x}^T\mathbf{Q}_{\boldsymbol{XY}}\boldsymbol{y} + \boldsymbol{y}^T\mathbf{Q}_{\boldsymbol{YY}}\boldsymbol{y})) \end{aligned} \tag{3}$$

By matching the coefficients in the two terms, we find that

$$\begin{aligned} \mathbf{Q}_{\boldsymbol{XX}} &= \mathbf{Q}_1 + \mathbf{Q}_2 \\ \mathbf{Q}_{\boldsymbol{XY}} = \mathbf{Q}_{\boldsymbol{YX}} &= -\mathbf{Q}_2 \\ \mathbf{Q}_{\boldsymbol{YY}} &= \mathbf{Q}_2 \end{aligned}$$

Thus, $\boldsymbol{X}|\boldsymbol{Y}$ is a GMRF with expected value

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{X}|\boldsymbol{Y}} &= (\mathbf{Q}_1+\mathbf{Q}_2)^{-1}\mathbf{Q}_2\boldsymbol{y} \\ &= (\frac{1}{\sigma^2}\mathbf{I}_{37} + \mathbf{V}^{-1})^{-1}\mathbf{V}^{-1}\boldsymbol{y} \end{aligned}$$

and precision matrix

$$\begin{aligned} \mathbf{Q}_{\boldsymbol{X}|\boldsymbol{Y}} &= \mathbf{Q}_1 + \mathbf{Q}_2 \\ &= \frac{1}{\sigma^2}\mathbf{I}_{37} + \mathbf{V}^{-1} \end{aligned}$$

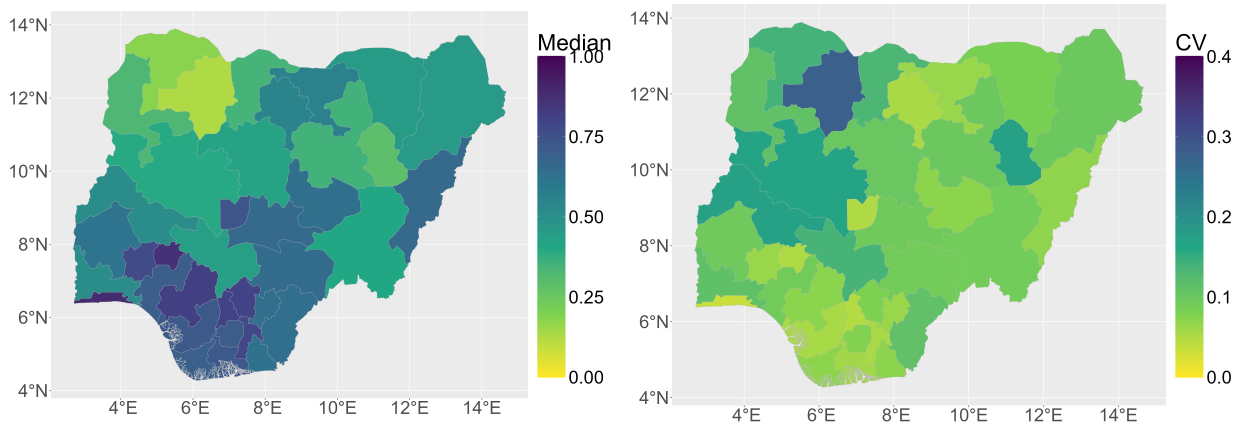
In the case that $\sigma^2 \to \infty$, we obtain

$$\begin{aligned} \lim_{\sigma^2\to\infty} \mu_{\boldsymbol{X}|\boldsymbol{Y}} &= \boldsymbol{y} \\ \lim_{\sigma^2\to\infty} \mathbf{Q}_{\boldsymbol{X}|\boldsymbol{Y}} &= \mathbf{V}^{-1} \end{aligned}$$

Since $\boldsymbol{X}|\boldsymbol{Y}$ follows a Gaussian distribution, we know that expit$(\boldsymbol{X}|\boldsymbol{Y})$ follows a logit-normal distribution. The inverse logit (expit) transform of $\boldsymbol{X}$ gives the vector $(P_1, P_2, \ldots, P_{37})$, so the marginal distributions are

$$P_a|\boldsymbol{Y} = \boldsymbol{y} \sim \text{Logitnormal}(y_a, V_a),\ a = 1, \ldots, 37.$$

We simulate 100 realizations with $\sigma^2 = 100^2$, and compute the median and coefficient of variation for each of the areas.



(a) (b)

Figure 7: Median values (left) and coefficient of variation (CV) (right).

In Figure 7a, we see that the estimated median proportions are higher towards the lower left and smaller towards the upper left, which is similar to that of the observed proportions. The area with the lowest median value also has the highest coefficient of variation, as seen in Figure 7b. Overall, the estimated median values seem to be very similar to that of the observed proportions in Figure 6. We assumed a prior with no spatial correlation and high variance, so it makes sense that the estimated proportions are similar to the observed proportions.

2c)

We place another prior on $\boldsymbol{X}$, namely the Besag model, which we defined in Equation 1. We repeat the procedure from above of matching coefficients.

The joint density takes the form

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{y}) &= f(\boldsymbol{x}) f(\boldsymbol{y}|\boldsymbol{x}) \\ &\propto \exp(-\frac{1}{2}(\boldsymbol{x}^T(\tau \mathbf{R}_1 + \mathbf{V})\boldsymbol{x} - 2\boldsymbol{y}^T\mathbf{V}^{-1}\boldsymbol{x} + \boldsymbol{y}^T\mathbf{V}^{-1}\boldsymbol{y})), \end{aligned}$$

which we match to the expression in Equation 3. We find

$$\begin{aligned} \mathbf{Q}_{\boldsymbol{XX}} &= \tau \mathbf{R}_1 + \mathbf{V}^{-1} \\ \mathbf{Q}_{\boldsymbol{XY}} = \mathbf{Q}_{\boldsymbol{YX}} &= -\mathbf{V}^{-1} \\ \mathbf{Q}_{\boldsymbol{YY}} &= \mathbf{V}^{-1} \end{aligned}$$

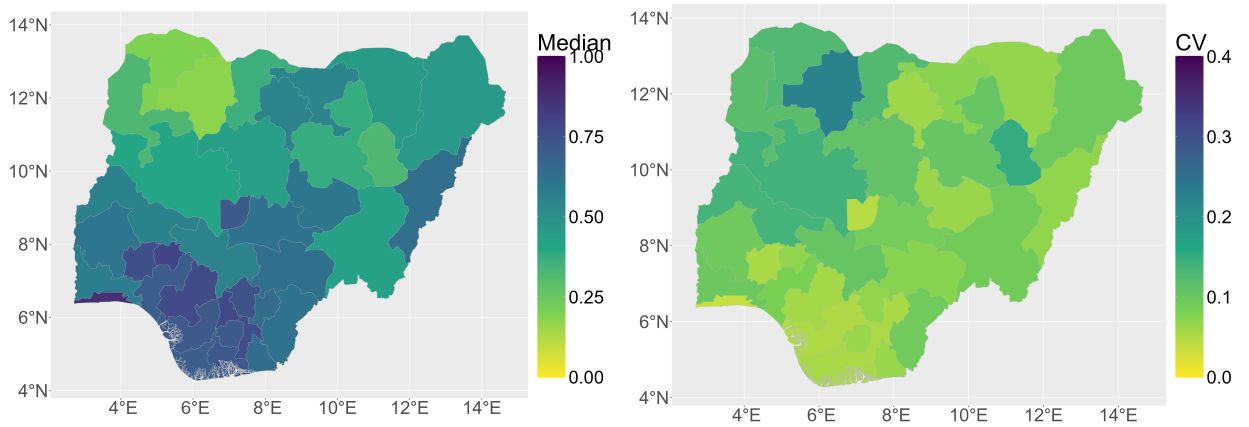
Thus, $\boldsymbol{X}|\boldsymbol{Y} = \boldsymbol{y}$ is normally distributed with expected value

$$\boldsymbol{\mu}_{\boldsymbol{X}|\boldsymbol{Y}} = (\tau \mathbf{R}_1 + \mathbf{V}^{-1})^{-1}\mathbf{V}^{-1}\boldsymbol{y},$$

and precision matrix

$$\mathbf{Q}_{\boldsymbol{X}|\boldsymbol{Y}} = \tau \mathbf{R}_1 + \mathbf{V}^{-1}$$

This is a proper GMRF, since the matrix $\mathbf{V}^{-1}$ has full rank. We simulate 100 realizations with $\tau = 1$ and calculate the median and coefficient of variation in each of the areas.



(a) (b)

Figure 8: Median values (left) and coefficient of variation (CV) (right).

Figure 8 shows the effect of the Besag prior and the spatial correlation. If we compare Figure 7b to Figure 8b, we see that the variance is slightly lower across most of the areas, if not all. The median values are also smoother in Figure 8a compared to Figure 7a, which we attribute to the spatial correlation in the Besag prior.

2d)

We imagine that an independent survey gave rise to a much more precise estimate of the proportion in Kaduna. We assume that $Y_{38}|P_{Kaduna} \sim \mathcal{N}(\text{logit}(P_{Kaduna}), 0.1^2)$ and that $Y_{38}|\boldsymbol{P}$ is independent of $\boldsymbol{Y}|\boldsymbol{P}$. Let $\tilde{\boldsymbol{Y}} = (Y_1, \ldots, Y_{37}, Y_{38})^T$. Then,

$$\tilde{\boldsymbol{Y}}|\boldsymbol{X} = \boldsymbol{x} \sim \mathcal{N}_{38}(\mathbf{M}\boldsymbol{x}, \tilde{\mathbf{V}}),$$

where $\mathbf{M} \in \mathbb{R}^{38\times 37}$ is the identity matrix with an extra row, $(0, \ldots, 0, 1, 0, \ldots, 0)$, on the bottom. The only non-zero element 1 is placed on index 19, which corresponds to the Kaduna area. In addition, $\tilde{\mathbf{V}} \in \mathbb{R}^{38\times 38}$ is the diagonal matrix with elements $V_a,\ a = 1, \ldots, 37$ and $V_{38} = 0.1^2$.

We place the same Besag prior on $\boldsymbol{X}$ as before, and repeat the procedure of matching coefficients. The joint density takes the form

$$\begin{aligned} f(\boldsymbol{x}, \tilde{\boldsymbol{y}}) &= f(\boldsymbol{x}) f(\tilde{\boldsymbol{y}}|\boldsymbol{x}) \\ &\propto \exp(-\frac{1}{2}(\boldsymbol{x}^T \tau \mathbf{R}_1 \boldsymbol{x} + (\tilde{\boldsymbol{y}} - \mathbf{M}\boldsymbol{x})^T \tilde{\mathbf{V}}^{-1}(\tilde{\boldsymbol{y}} - \mathbf{M}\boldsymbol{x}))) \\ &= \exp(-\frac{1}{2}(\boldsymbol{x}^T(\tau \mathbf{R}_1 + \mathbf{M}^T\tilde{\mathbf{V}}^{-1}\mathbf{M})\boldsymbol{x} - 2\tilde{\boldsymbol{y}}^T\tilde{\mathbf{V}}^{-1}\mathbf{M}\boldsymbol{x} + \tilde{\boldsymbol{y}}^T\tilde{\mathbf{V}}^{-1}\tilde{\boldsymbol{y}})), \end{aligned}$$

which we match with Equation 3. We find

$$\begin{aligned}\mathbf{Q}_{\boldsymbol{XX}} &= \tau\mathbf{R}_1 + \mathbf{M}^T\tilde{\mathbf{V}}^{-1}\mathbf{M} \\ \mathbf{Q}_{\boldsymbol{YX}} &= \mathbf{Q}_{\boldsymbol{XY}}^T = -\tilde{\mathbf{V}}^{-1}\mathbf{M} \\ \mathbf{Q}_{\boldsymbol{YY}} &= \tilde{\mathbf{V}}^{-1}\end{aligned}$$

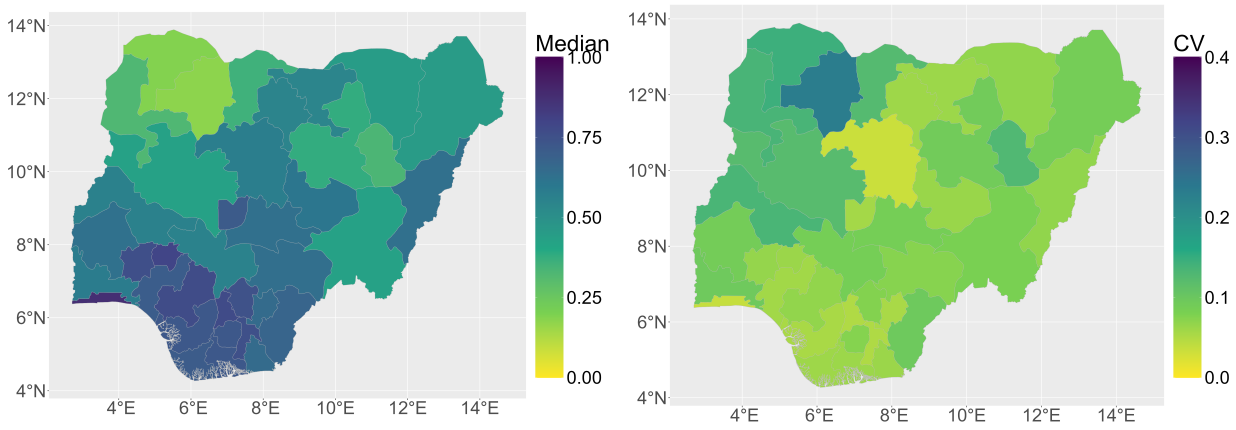
Thus, $\boldsymbol{X}|\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}}$ is a GMRF with expected value

$$\boldsymbol{\mu}_{\boldsymbol{X}|\tilde{\boldsymbol{Y}}} = (\tau\mathbf{R}_1 + \mathbf{M}^T\tilde{\mathbf{V}}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\tilde{\mathbf{V}}^{-1}\tilde{\boldsymbol{y}}$$

and precision matrix

$$\mathbf{Q}_{\boldsymbol{X}|\tilde{\boldsymbol{Y}}} = \tau\mathbf{R}_1 + \mathbf{M}^T\tilde{\mathbf{V}}^{-1}\mathbf{M}$$

The precision matrix has full rank, since $\tilde{\mathbf{V}}$ is positive definite, so $\boldsymbol{X}|\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}}$ is a proper GMRF.



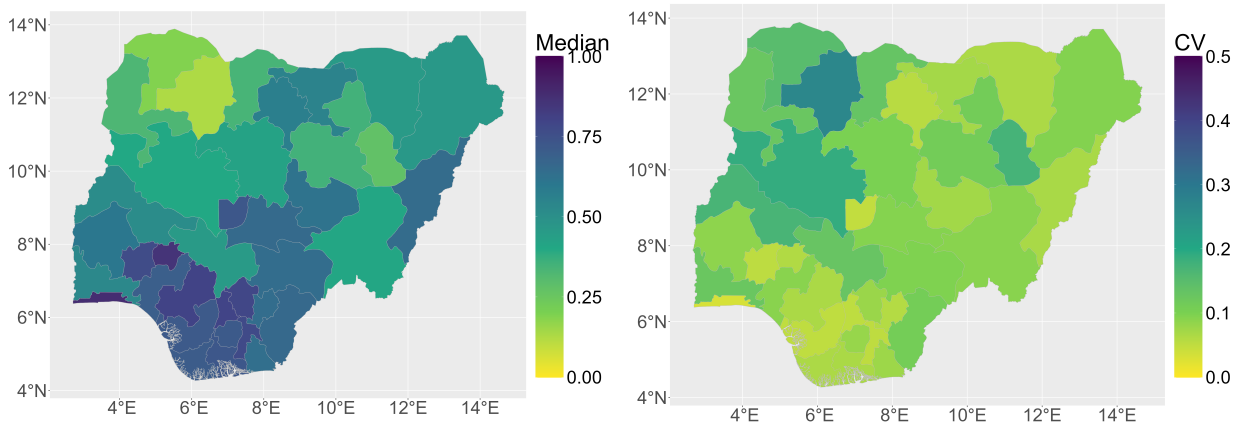
(a) (b)

Figure 9: Median values (left) and coefficient of variation (CV) (right).

In Figure 9, we see that the Kaduna region (large region in the center), has gained much lower variance with the inclusion of the extra survey. The surrounding regions also gain lower variance through the spatial correlation of the Besag prior. The median value of the Kaduna region is also lower than in the previous models. The other areas seem to have similar median values as before.

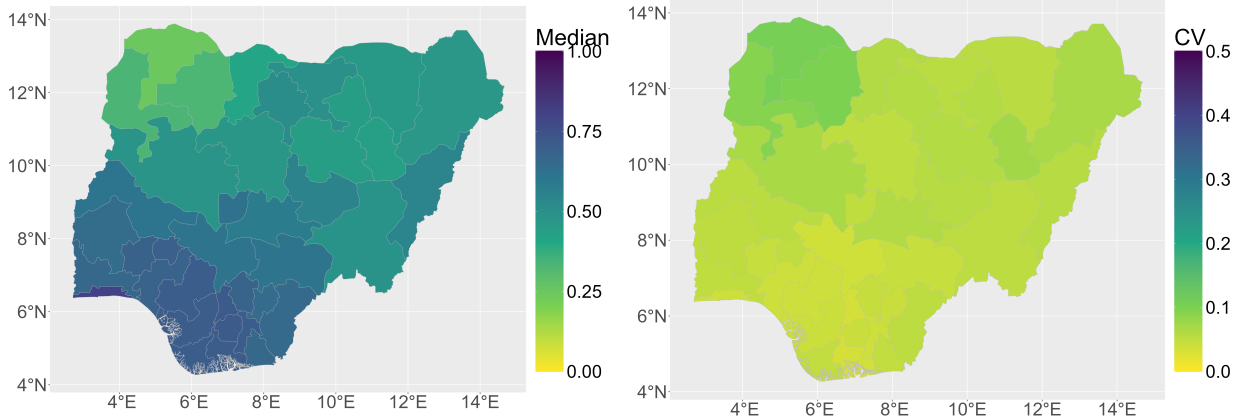
2e)

We investigate the effect of the precision parameter τ by repeating part 2c) for parameters $\tau = 0.1$ and $\tau = 10$.



(a) Median values for $\tau = 0.1$

(b) CV values for $\tau = 0.1$



(c) Median values for $\tau = 10$

(d) CV values for $\tau = 10$

Figure 10: Median values (left) and coefficient of variation (CV) (right).

Figure 10 shows the effect of the precision parameter τ. When we increase the parameter to $\tau = 10$, the coefficient of variation becomes small across all regions, and the range of the median values also becomes smaller. We can view the parameter as weighting the spatial correlation, where higher values of τ increases the spatial correlation and thus decreases the variance. We note that if we use $\tau = 0$, we get a model where the regions are independent, similar to the model part 2b. It is therefore important to estimate τ correctly, which we can do by method of maximum likelihood estimation.

2f)

To obtain the log-likelihood $l(\tau; \boldsymbol{y})$, we use Bayes' rule, which states that

$$f(\boldsymbol{x}|\boldsymbol{y};\tau) = \frac{f(\boldsymbol{y}|\boldsymbol{x};\tau)f(\boldsymbol{x}|\tau)}{f(\boldsymbol{y};\tau)}$$

This gives the log-likelihood

$$\begin{aligned} l(\tau;\boldsymbol{y}) &= \log(f(\boldsymbol{y};\tau) \\ &= \log\left(\frac{f(\boldsymbol{y}|\boldsymbol{x})f(\boldsymbol{x};\tau)}{f(\boldsymbol{x}|\boldsymbol{y};\tau)}\right) \\ = \log(f(\boldsymbol{y}|\boldsymbol{x})) + \log(f(\boldsymbol{x};\tau)) &- \log(f(\boldsymbol{x}|\boldsymbol{y};\tau)) \end{aligned}$$

We know that $\boldsymbol{Y}|\boldsymbol{X} = \boldsymbol{x} \sim \mathcal{N}_{37}(\boldsymbol{x}, \mathbf{V})$, so the first term becomes

$$\begin{aligned} \log(f(\boldsymbol{y}|\boldsymbol{x};\tau)) &\propto \log(\exp(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\mathbf{V}^{-1}(\boldsymbol{y}-\boldsymbol{x}))) \\ &= -\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\mathbf{V}^{-1}(\boldsymbol{y}-\boldsymbol{x}) \end{aligned}$$

The second term is the prior on $\boldsymbol{X}$, which in this case is the Besag model:

$$\begin{aligned} \log(f(\boldsymbol{x};\tau)) &\propto \log(\tau^{\frac{37-1}{2}}\exp(-\frac{1}{2}(\boldsymbol{x}^T(\tau\mathbf{R})\boldsymbol{x}))) \\ &= \frac{37-1}{2}\log(\tau) - \frac{1}{2}\boldsymbol{x}^T\tau\mathbf{R}\boldsymbol{x} \end{aligned}$$

We know that $\boldsymbol{X}|\boldsymbol{Y} = \boldsymbol{y} \sim \mathcal{N}_{37}(\boldsymbol{\mu}_C := (\tau\mathbf{R}+\mathbf{V}^{-1})^{-1}\mathbf{V}^{-1}\boldsymbol{y}, \mathbf{Q}_C := \tau\mathbf{R}+\mathbf{V}^{-1})$. The third term becomes

$$\begin{aligned} \log(f(\boldsymbol{x}|\boldsymbol{y};\tau)) &\propto \log(|\mathbf{Q}_C^{-1}|^{-1/2}\exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_C)^T\mathbf{Q}_C(\mathbf{x}-\boldsymbol{\mu}_C))) \\ &= -\frac{1}{2}\log(|\mathbf{Q}_C^{-1}|) - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_C)^T\mathbf{Q}_C(\mathbf{x}-\boldsymbol{\mu}_C) \end{aligned}$$

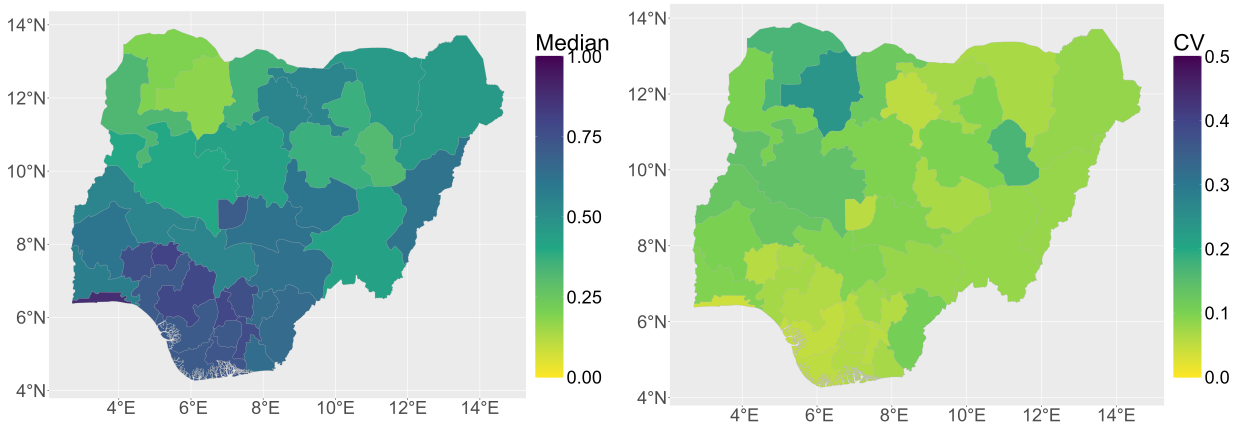
Putting all the terms together, we find the log-likelihood

$$\begin{aligned} l(\tau;\boldsymbol{y}) &= \log(f(\boldsymbol{y}|\boldsymbol{x})) + \log(f(\boldsymbol{x};\tau)) - \log(f(\boldsymbol{x}|\boldsymbol{y};\tau)) \\ = -\frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^T\mathbf{V}^{-1}(\boldsymbol{y}-\boldsymbol{x}) &+ \frac{37-1}{2}\log(\tau) - \frac{1}{2}\boldsymbol{x}^T\tau\mathbf{R}\boldsymbol{x} \\ -\frac{1}{2}\log(|\mathbf{Q}_C|) &+ \frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_C)^T\mathbf{Q}_C(\mathbf{x}-\boldsymbol{\mu}_C) + \text{Const} \end{aligned}$$

We optimize this function to find the maximum likelihood estimate of τ.

```
## [1] 0.8062634
```

The maximum likelihood estimate of τ is found to be $\hat{\tau} = 0.8063$. We use this value to again compute the median and coefficient of variations of 100 samples from $P_a|\boldsymbol{Y} = \boldsymbol{y}$.



(a) Median values for $\tau = \hat{\tau}$

(b) CV values for $\tau = \hat{\tau}$

Figure 11: Median values (left) and coefficient of variation (CV) (right).

Figure 11 shows the median values and coefficient of variation for 100 realizations, using the maximum likelihood estimate of τ. If we compare to Figure 8, where we used $\tau = 1$, the results are very similar, but the coefficient of variation is smaller in Figure 11.