Investigation of Bayesian spatial models

Suggested citation

Cramb SM, Duncan EW, Baade PD, Mengersen KL, 2017. *Investigation of Bayesian Spatial Models*. Brisbane: Cancer Council Queensland and Queensland University of Technology (QUT).

Author affiliations

Viertel Cancer Research Centre, Cancer Council Queensland: Susanna Cramb and Peter Baade.

ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology (QUT): Susanna Cramb, Earl Duncan and Kerrie Mengersen.

Acknowledgements

The authors wish to thank Upeksha Chandrasiri for conducting citation searches and assisting with data compilation. Methodological advice and manuscript revisions from Nicole White, James McGree, Pamela Burrage and Gentry White are also gratefully acknowledged. Additional suggestions from Craig Anderson and Louise Ryan also improved the final report.

Overview

The popularity of Bayesian disease mapping is increasing, as is the variety of available models. The most commonly used prior for enabling spatial correlation within a Bayesian model is the intrinsic conditional autoregressive (CAR) distribution. This approach allows for local smoothing of estimates over neighbouring areas, but it assumes a common variance for the smoothing term over the whole region. This is applicable if there is a smooth spatial trend over the region, which may not be valid for large, spatially heterogeneous areas. The aim of this report is to critically review alternative Bayesian models, especially those that enable local variation in the smoothing.

The motivation for this study is the development of a national cancer atlas for Australia. Our focus is therefore on the performance on small-area cancer incidence models that allow for substantive differences in area-level variables, in particular area-specific population, demographics, land area and number of cases. For example, remote areas of Australia typically have relatively low population counts, few or no cancer cases in certain years, and very large land areas, compared with urban areas. Moreover, the number of cancer cases are heavily dependent on the size and structure of the population, which can vary substantially within and between areas.

The study commenced with a literature search to identify Bayesian models that have been commonly used in disease mapping. Two broad types of models were identified, namely 'global' spatial smoothing models that have a common spatial smoothing term across the region, and 'local' spatial smoothing models that allow for differential smoothing depending on neighbourhood characteristics. The specific models that were identified are listed below.

Global spatial smoothing:

- Intrinsic CAR/BYM model (Besag *et al.*, 1991)
- Proper CAR model (Besag, 1974)
- Leroux model (Leroux *et al.*, 2000)
- Geostatistical model (Clements *et al.*, 2006)
- Global spline models (Lang & Brezger, 2004)

Local spatial smoothing

- CAR dissimilarity models (Lee & Mitchell, 2012)
- Localised autocorrelation (Lee & Mitchell, 2013)
- Locally adaptive model (Lee & Sarran, 2015)
- Hidden Potts model (Green & Richardson, 2002)
- Spatial partition model (Knorr-Held & Raßer, 2000, Denison & Holmes, 2001)
- Weighted sum of spatial priors (Lawson & Clark, 2002)
- Leroux scale mixture model (Congdon, 2017)
- Local spline models (Goicoa *et al.*, 2012, Perperoglou & Eilers, 2010)
- Skew-elliptical areal spatial models (Nathoo & Ghosh, 2013).

Selected models were then applied to simulated Australian incidence data for liver, lung and all invasive cancers. These cancers were chosen to reflect different patterns of spatial heterogeneity and overall levels of incidence.

Criteria used to differentiate between models were the plausibility of estimates (range of standardised incidence ratios, posterior probabilities and the width of the credible intervals); model goodness of fit (Watanabe-Akaike information criterion (WAIC), Moran's I on the residuals along with visually checking residual spatial patterns, and the Deviance information criterion (DIC)). Finally, the computational time and practical feasibility of implementing the models was assessed.

No one model performed brilliantly across every type of cancer, and it was surprisingly difficult to balance the need for smoothing against obtaining sufficient variation to capture genuine differences. The greatest differences in modelled estimates were found for the simulated liver cancer data, representing a rarer cancer. The three best performing models overall were identified as being the P-

spline (radial), Leroux, and localised autocorrelation (G=5). Recommendations are therefore to identify data characteristics (the numbers and likely disparities in neighbouring areas, here considered to be likely if there is evidence of a trend across socioeconomic quintiles) and choose the most appropriate model out of these top three models.

Contents

List of Abbreviations

ABS	Australian Bureau of Statistics
ACIM	Australian Cancer Incidence and Mortality
ASGS	Australian Statistical Geography Standard
BUGS	Bayesian inference Using Gibbs Sampling
BYM	Besag, York and Mollié
CAR	Conditional Autoregressive
CI	Credible interval
DIC	Deviance Information Criterion
DP	Dirichlet process
ICAR	Intrinsic Conditional Autoregressive
INLA	Integrated Nested Laplace Approximation
IRSD	Index of Relative Socioeconomic Disadvantage
JAGS	Just Another Gibbs Sampler
MCMC	Markov chain Monte Carlo
MRF	Markov Random Field
PP	Posterior probability
PRIDE	Penalised Random Individual Dispersion Effects
RA	Remoteness Areas
RMSPE	Root Mean Squared Prediction Error
SA2	Statistical Area 2
SEIFA	Socioeconomic Indexes for Areas
SIR	Standardised Incidence Ratio
WAIC	Watanabe-Akaike Information Criterion

Introduction

There are many advantages to using a Bayesian approach when mapping disease. These include enabling direct probabilistic statements to be made, such as the probability that an area has an increased disease risk (Kang *et al.*, 2016). The use of prior distributions enables estimates to be reliable and robust (i.e. well-defined and stable), even when there are few cases in an area (Kang et al., 2016).

Many Bayesian spatial models have been proposed, most of which vary with respect to the representation of the spatial prior. One of the outstanding difficulties for a user is choosing the type of prior that will appropriately characterise the spatial nature of the data of interest. Often choices are made on the ease of implementation, which is part of the appeal of using a conditional autoregressive (CAR) prior.

Spatial priors are designed to perform some smoothing over areas. Some smoothing is desirable as it reduces uncertainty of estimates in the model, and it provides insight into the underlying spatial trend which may otherwise be obscured by noise and the effects of other variables. Undersmoothing is adverse because it diminishes the above benefits, but oversmoothing is also undesirable because it can conceal genuine deviations from the underlying smooth spatial surface, which may signify areas that are of clinical interest and importance. However, achieving a satisfactory level of smoothing is a difficult task, and it forms part of the criteria for comparing models, in particular, the plausibility of estimates.

Here, we investigate the performance of the most popular Bayesian spatial models identified through searching the literature. We examine their commonalities and differences and contrast their performance on three simulated datasets.

Our aim was to identify appropriate models to apply to the Australian Cancer Atlas. This will examine the incidence of and survival from around 20 different cancer types, with greatly varying numbers (some extremely sparse) and patterns, across more than 2,100 small areas. These areas have large differences in population size, demographic structure, land area size and shape. This diversity increases the challenges of appropriately cancer estimates.

Methods

Search strategy

Our aim was to identify popular Bayesian models for disease mapping.

Databases including Web of Science, PubMed, ScienceDirect and Google Scholar were searched electronically in August 2017. The same search terms were used across all databases and were: Bayesian disease AND (map OR mapping) AND (autoregressive OR CAR OR smoothing). Details on the search strategy are included in Table 1. Due to the very large number of items returned through Google Scholar and Science Direct, items were sorted by relevance and the first 200 items scanned from those databases. Titles and abstracts were screened first and the resulting papers identified were evaluated through reading the full text. Of the 15 primary references relating to the specific models considered in this report, 11 were identified from the search strategy directly, and of the remaining four, three were found indirectly from this literature while the other one was a textbook.

Focused clustering models (Diggle, 1990) aim to determine the pattern of events near an exposure source (such as a nuclear power station) so are not applicable to the Australian Cancer Atlas and were excluded. Methods more suited to exploratory analyses, rather than a formal model (such as geographically weighted regression (Brunsdon *et al.*, 1996)) were also not considered (Wheeler, 2014).

Table 1: The detailed search strategy

Database	Keywords	Restrictions	Items returned
Google Scholar	Bayesian disease AND (map OR mapping) AND (autoregressive OR CAR OR smoothing)	Exclude patents and citations; sort by relevance	~38200 (Only first 200 scanned)
Web of Science	Bayesian disease AND (map OR mapping) AND (autoregressive OR CAR OR smoothing)	Searching in Topic (default) with default settings	206
Science Direct	Bayesian disease AND (map OR mapping) AND (autoregressive OR CAR OR smoothing)	Advanced search, searching in all fields; rest default settings (sorted by relevance)	3,550 (Only first 200 scanned)
Pubmed	Bayesian disease AND (map OR mapping) AND (autoregressive OR CAR OR smoothing)	Default settings	69

Data simulation

Popular models identified in the literature were applied to simulated Australian incidence data for ages 15+ years aggregated over 2005-2014 for 3 types of cancer: male liver (rare, strong socioeconomic gradient, so disparities expected between neighbouring regions), male lung (more common, but still a strong socioeconomic gradient) and female all invasive cancers (comparatively high numbers with a smaller socioeconomic influence due to opposing socioeconomic gradients between cancer types). The areas used were statistical areas 2 (SA2s) based on the 2011 Australian Statistical Geography Standard (ASGS) boundaries (Australian Bureau of Statistics, 2011). After excluding some areas with no/nominal resident populations, the number of areas was 2,153. The median population of the

included SA2s was 9,055 (range: 3 to 50,251). Land area size varied from 0.8 to 520,000km^2, with a median of 16km^2.

Details on calculating the simulated data are available in Appendix F. Briefly, these are based on the broad socioeconomic-remoteness patterns observed in cancer data in Queensland, by each type of cancer and sex, with additional random adjustment of the numbers. The median number of cases by SA2 was 2 liver cancer cases (range 0-19), 25 lung cancer cases (range: 0-163) and 210 all invasive cancer cases (range: 0-1012).

Model comparison

Models were compared based on the following criteria:

- Plausibility of estimates (Appendices A and G):
    - Standardised incidence ratios (SIRs)
    - Posterior probabilities (PPs) and
    - Credible intervals (CIs)
- Model goodness of fit (Appendices B and H):
    - Watanabe-Akaike information criterion (WAIC)
    - Moran's I on residuals
    - Model residuals (values and spatial patterns)
    - Deviance information criterion (DIC)
- Computational time and feasibility (Appendix B)
    - Software available for easy implementation.

In addition, convergence of the SIR estimates was examined and convergence results based on the Geweke convergence diagnostic (Geweke, 1992) are summarised in Appendix C. The differences in neighbourhood matrices between models are summarised in Appendix D. Code for implementing the models is available in Appendix E, and further methodological details are available in Appendix F.

The plausibility of estimates considered the CI width (unreasonably large CIs suggested the estimate was not well-defined; while very precise estimates suggested uncertainty was not appropriately included. It additionally considered the amount of smoothing of the median posterior SI in comparison to the raw SIRs.

To compare how models perform on the simulated data, four model goodness of fit measures are considered. DIC (Spiegelhalter *et al.*, 2002)and WAIC (Watanabe, 2010) are both useful for comparing the predictive accuracy between models. Although DIC is a commonly used measure to compare Bayesian models, WAIC has several advantages over DIC, including that it closely approximates Bayesian cross-validation, it uses the entire posterior distribution and it is invariant to parameterisation (Vehtari *et al.*, 2017). For both these measures, smaller values indicate a better fitting model.

Residuals were also calculated and Moran's I (Moran 1950) applied to these to determine if spatial autocorrelation was present after fitting the models. As values of Moran's I close to 0 indicate very low or no spatial autocorrelation, here we consider values above 0.2 to be suggestive of some positive spatial autocorrelation. The closer Moran's I is to zero, the better the model accounts for spatial autocorrelation (Anderson & Ryan, 2017).

Bayesian spatial models

In this section, we provide a summary of popular Bayesian spatial models that were identified as potentially applicable to map small-area cancer incidence. The following three-stage hierarchical model is "a natural model for disease mapping" and has been widely used (Best et al., 2005):

Stage 1:	$Y_i \sim \text{Poisson}(E_i e^{\mu_i})$ for $i = 1, \dots, N$ areas
Stage 2:	$\mu_i = \alpha + \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta} + R_i$
Stage 3:	$\alpha \sim p(\cdot \mid \boldsymbol{\theta}_\alpha)$
	$\boldsymbol{\beta} \sim p(\cdot \mid \boldsymbol{\theta}_\beta)$
	$R_i \sim p(\cdot \mid \boldsymbol{\theta}_R)$

The first stage is the likelihood model. The Poisson distribution is used because $\{Y_1, \dots, Y_N\}$ are count data for a comparatively uncommon disease. E_i represents the expected cancer counts and are commonly defined using internal standardisation of risk (see Appendix F for calculation details).

The second stage is an expression for the log-relative risk μ_i. This is often expressed as a regression equation and typically includes an overall fixed effect (intercept, denoted α), covariate effects ($\boldsymbol{\beta}$) where $\boldsymbol{x}_i$ denotes a vector of covariates relating to area i, and spatial random effect(s) (R_i). As shown below, the 'spatial' random effects can be formed from multiple components, some of which may allow for extra-Poisson variation (Besag et al., 1991).

The third stage consists of the prior distributions for each of the unknown parameters, which, in the absence of external information, are usually specified as weakly informative, such as Gaussian distributions with zero mean and some large variance. The random effects may be assumed to follow a CAR (or alternative) prior to account for spatial smoothing (Best et al., 2005, Besag et al., 1991). If the parameters $\boldsymbol{\theta}_\alpha$, $\boldsymbol{\theta}_\beta$ or $\boldsymbol{\theta}_R$ are unknown, then the hyperpriors represent a fourth stage of the hierarchy.

Global spatial smoothing

Global spatial smoothing means the same smoothing parameters are applied consistently across the entire region (Lee & Mitchell, 2012). Although the global CAR-based models are easy to implement in a range of software, disadvantages of global models include the potential for oversmoothing, as discontinuities between adjacent areas are smoothed over. Oversmoothing is defined as obscuring too much of the underlying geographic patterns, although what is 'too much' may depend on the context and the aim of the analysis.

Intrinsic CAR and BYM

The intrinsic CAR (ICAR) model specifies the following set of conditional distributions for the spatial random effect parameter:

$$R_i = S_i$$

$$S_i | \boldsymbol{s}_{\setminus i} \sim \mathcal{N}\left(\frac{1}{\sum_j w_{ij}} \sum_j w_{ij} s_j \,,\ \frac{\sigma_s^2}{\sum_j w_{ij}} \right)$$

or in matrix notation

$$S_i | \boldsymbol{s}_{\setminus i} \sim \mathcal{N}(\{\mathbf{D}^{-1}\mathbf{W}\boldsymbol{s}\}_i,\ \sigma_s^2 \{\mathbf{D}^{-1}\}_{ii})$$

where w_{ij} is the element of a spatial weights matrix $\mathbf{W}$ corresponding to row i and column j (Besag et al., 1991, Besag, 1974, Lee, 2011, Best et al., 2005), and $\mathbf{D}$ is a diagonal matrix with elements

$\text{diag}\{\sum_j w_{ij}\}$. $\mathbf{W}$ determines the spatial proximity between the random effects, and it is most commonly defined as a binary, first-order, adjacency matrix, whereby

$$w_{ij} = \begin{cases} 1 & \text{if areas } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

This model implies that the conditional expectation of S_i is equal to the mean of the random effects at neighbouring locations.

The S_i can be regarded as structured spatial random effects. If $R_i = S_i + U_i$, so that unstructured spatial random effects $U_i \sim N(0, \sigma_U^2)$ are also included, the resulting model is referred to as the convolution model, or the BYM model in honour of Besag et al. (1991). However, the two separate random effects components cannot be individually identified – only their sum is identifiable (Eberly & Carlin, 2000). Note that for all CAR-based models, the strength of the partial autocorrelation depends on the number of neighbouring areas rather than on any underlying relationship (Lee & Mitchell, 2013).

Proper CAR

The full conditionals for the ICAR prior are proper, but the joint distribution is improper since the precision matrix is singular. The impropriety of the ICAR prior can be overcome by redefining the precision matrix

$$\mathbf{T} = \frac{1}{\sigma_s^2}(\mathbf{D} - \mathbf{W})$$

to

$$\mathbf{T} = \frac{1}{\sigma_s^2}(\mathbf{D} - \rho\mathbf{W})$$

such that the full conditionals are:

$$S_i | \boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{\rho}{\sum_j w_{ij}} \sum_j w_{ij} s_j \,,\ \frac{\sigma_s^2}{\sum_j w_{ij}} \right)$$

with the constraint $|\rho| < 1$, and using the terminology in Banerjee *et al.* (2003), ρ represents the expected proportional 'reaction' of S_i to $\frac{\sum_j w_{ij} s_j}{\sum_j w_{ij}}$. This ensures the covariance matrix $\mathbf{B}^{-1}$ is positive definite and ensures $\boldsymbol{S}$ has a proper joint distribution (Kandhasamy & Ghosh, 2017). This is the proper CAR prior, but it may have certain disadvantages, including potentially limiting the breadth of the posterior spatial pattern (Banerjee et al., 2003). Also, ρ will likely need to be very close to 1 for there to be a reasonable amount of spatial association (Banerjee et al., 2003).

Leroux CAR model

Another variation of the BYM model was proposed by Leroux et al. (2000),

$$S_i | \boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{\rho \sum_{j=1}^{N} w_{ij} s_j + (1-\rho)\mu_0}{\rho \sum_j w_{ij} + 1 - \rho}, \frac{\sigma_s^2}{\rho \sum_j w_{ij} + 1 - \rho} \right)$$

which only requires a single set of random effects (Lee, 2011). This avoids the difficulties in identifiability, and also selection of hyperpriors (given that in the BYM model the S_i variance is conditional on neighbouring areas, while the U_i has a marginal variance term) (Riebler *et al.*, 2016).

The precision matrix can be expressed as

$$\mathbf{T} = \frac{1}{\sigma_s^2}[\rho(\mathbf{D} - \mathbf{W}) + (1 - \rho)\mathbf{I}].$$

This mixture representation consists of correlated smoothing of the neighbouring random effects (weighted by ρ) as well as uncorrelated smoothing to a global mean μ_0 (weighted by $(1 - \rho)$) (Lee & Mitchell, 2012). Thus S_i has a conditional expectation based on a weighted average of both the independent random effects and the spatially structured random effects. The ICAR prior is therefore a limiting case of both the proper CAR and Leroux CAR models when ρ is set to 1.

Geostatistical model

Here, the residual spatial structure is modelled as a Gaussian process using a geostatistical design (Clements et al., 2006). Because this model incorporates distance, counts are assumed to be in the centroid of an area.

$$R_i \sim \mathcal{N}(S_i, \sigma^2)$$
$$S_i = \exp\left(-(\lambda d_{ij})^k\right), \ \ \lambda > 0$$

where λ controls the rate of decay, k is the "degree of spatial smoothing", and d_{ij} is the distance between points (e.g. centroids of areas) i and j (Clements et al., 2006). This expression is the exponential decay function with the addition of the power k. Rather than fix decay parameter λ a priori, a hyperprior is specified as a fourth stage:

$$\lambda \sim \mathrm{Uniform}(0.1, 6)$$

The justification for the bounds 0.1 and 6 are thus:

"…upper and lower bounds set at 0.1 and 6.0, which gave possible values for spatial correlation of 0.99–0.55 with a separating distance of 0.1 decimal degrees (the minimum distance between observed data points) and possible values for spatial correlation of 0.00–0.55 with a separating distance of 6.0 decimal degrees (the maximum distance between observed data points), assuming $\mathrm{k} = 1.0$" (Clements et al., 2006)

Alternative functions are possible, including the disc model (Richardson, 1992) (a linear decrease with increasing distance, where two discs of common radius are centred on centroids, and the correlation is proportional to the disc intersection area), or combining two parametric functions to obtain different shapes of decrease, such as the Matern class (Best et al., 2005). Note often limited information is available to guide the choice of functional form, or correlation parameters, especially as complexity increases (Best et al., 2005). It is vital that the choice of correlation function (and hyperpriors) gives near zero correlation at distances within the study region, to avoid nonidentifiability of the mean and correlation parameters (Best et al., 2005). Finally, these models can be computationally expensive due to inversion of the covariance matrix at each iteration (Best et al., 2005).

For the Australian cancer data sets, two adjustments were made to this model to provide a better fit. First, the priors for λ and k were changed according to the possible values of spatial correlation observed given different combinations of λ, k, and distances d_{ij}. This exploratory analysis suggested using

$$\lambda \sim \mathrm{Uniform}(0.01,1)$$
$$k \sim \mathrm{Uniform}(0.1, 20).$$

To allow for further flexibility, λ and k were replaced by one of $\{\lambda_1, \dots, \lambda_5\}$ and $\{k_1, \dots, k_5\}$ respectively according to the remoteness of the area (major city, inner regional, outer regional, remote, and very remote) to allow the degree of smoothing to vary between the five levels of remoteness. Second, to

make this model computationally feasible, the distance matrix $\{d\}_{ij}$ was modified by imposing a remoteness-specific radius of influence $\{r_1, \dots, r_5\}$ on each area, such that areas beyond this threshold are not considered neighbours. These radii are $\{50, 100, 200, 400, 800\}$ kilometres respectively. This induces a Markov random field (MRF) structure which should have only a negligible effect on parameter estimation while greatly increasing computational efficiency. Some remote and very remote areas are relatively close to major city and inner regional areas, which can lead to some areas having more than 1000 neighbouring SA2s, thereby drastically reducing any computational gains. Therefore, the imposed MRF was further modified to exclude major city areas as neighbours of remote areas, and to exclude both major city and inner regional areas as neighbours of very remote areas. This is achieved by setting the distances to these excluded areas to infinity. The result of these adjustments leads to

$$\{\mathbf{S}\}_{ij} = \begin{cases} \exp\left(-(\lambda_{z_i} d_{ij})^{k_{z_i}}\right) & \text{if } d_{ij} \le r_{z_i} \\ 0 & d_{ij} > r_{z_i} \end{cases}$$

$$S_i = f(\mathbf{S}) = \frac{1}{N_i} \sum_{j=1}^{N_r} \mathbf{S}_{ij}$$

where N_i is the number of areas within a radius of r_{z_i} units from the centroid of area i (including area i), $N_r = \max_i\{N_i\}$, and z_i represents the degree of remoteness for area i, where $z_i = 1$ corresponds to an area in a major city.

Global spline models

The spline model also assumes that the incidence cases (counts) are all located at the centroid of each area (Goicoa et al., 2012).

There are two main methods: smoothing splines and P-splines (MacNab, 2007). Smoothing splines are penalised splines which have knots on all data points. P-splines allow for a smaller number of knots, and are commonly formulated as a penalised spline regression under a 'difference penalty' based on the coefficients of adjacent B-spline bases or other spline bases (MacNab, 2007).

The "interaction" (dependency or correlation) between areas i and j can be modelled by a two-dimensional smooth surface (Goicoa et al., 2012). First, define the longitude and latitude pairs representing the centroid of each area, denoted (c_{1i}, c_{2i}).

$$R_i = f(c_{1i}, c_{2i})$$

where the smooth function $f(\cdot)$ is expressed as

$$f(c_{1i}, c_{2i}) = \theta_1 B_1(c_{1i}, c_{2i}) + \cdots + \theta_k B_k(c_{1i}, c_{2i})$$

which is estimated using P-splines with B-spline bases $B_1, \dots, B_k$, $\theta_1, \dots, \theta_k$ are unknown coefficients which are penalised to control for "wiggliness" through a penalty matrix, and k depdence on the number of knots and the degree of the B-spline bases.

Define $\boldsymbol{c}_1 = (c_{11}, \dots, c_{1N})^{\mathrm{T}}$ and $\boldsymbol{c}_2 = (c_{21}, \dots, c_{2N})^{\mathrm{T}}$ and univariate B-spline bases$\mathbf{B}_1 = \{B_{11}(\boldsymbol{c}_1), \dots, B_{1k_1}(\boldsymbol{c}_1)\}$ and $\mathbf{B}_2 = \{B_{21}(\boldsymbol{c}_2), \dots, B_{2k_2}(\mathrm{c}_2)\}$. The bivariate B-spline basis is then constructed as the row-wise Kronecker product (denoted by $\boxtimes$) of the marginal B-spline bases:

$$\begin{aligned} \mathbf{B} &= \mathbf{B}_2 \boxtimes \mathbf{B}_1 \\ &= \left(\mathbf{B}_2 \otimes \mathbf{1}_{k_1}^{\mathrm{T}}\right) \odot \left(\mathbf{1}_{k_2}^{\mathrm{T}} \otimes \mathbf{B}_1\right). \end{aligned}$$

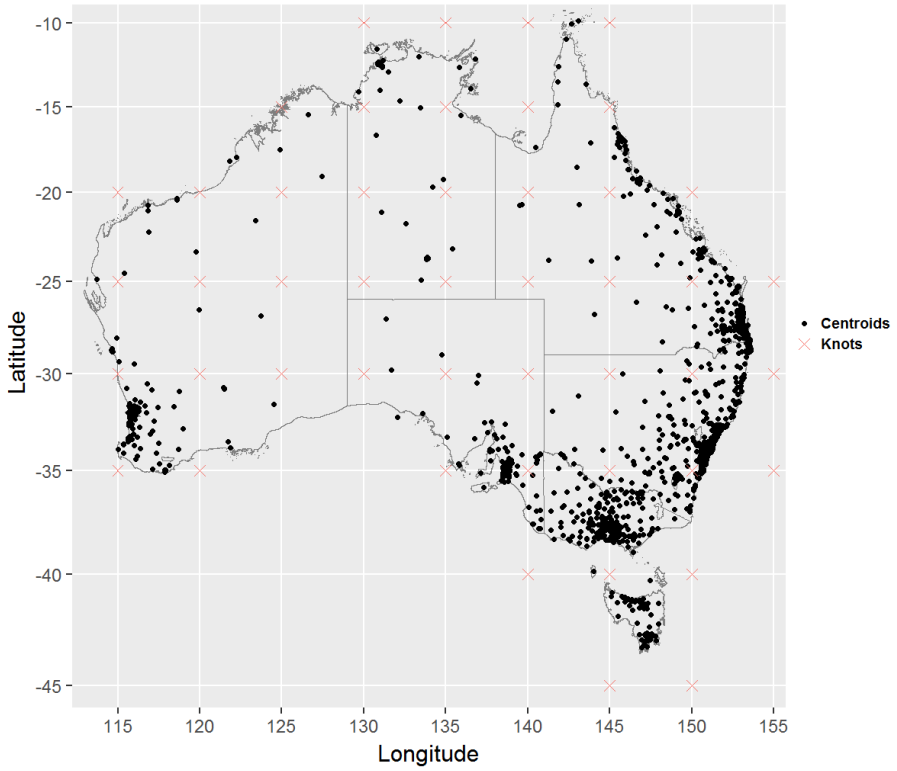
The basis $\mathbf{B}$ is of dimension $N \times k$ where $k = k_1 k_2$, the symbols $\otimes$ and $\odot$ represent the Kronecker product and "element-wise" matrix product respectively, and $\mathbf{1}_{k_1}$ and $\mathbf{1}_{k_2}$ are column vectors of ones of length k_1 and k_2" (Goicoa et al., 2012).

Overall this model provides a relatively smooth surface, as the covariance structure is impacted by long distance effects that influence the smoothing, which contrasts to the covariance structure of the CAR model where a region's mean depends on the mean of its neighbours (Goicoa et al., 2012).

The formulation of the P-spline model using the row-wise Kronecker product, or tensor product, is better suited to data which lie on a regular grid, or at least have similar distances between the centroids. An alternative formulation (Ruppert et al., 2003) is to define the B-spline bases in terms of the distances,

$$z_{ik} = \exp\left(-\frac{d_{ik}}{\Delta}\right)\left(1 + \frac{d_{ik}}{\Delta}\right)$$

where d_{ik} is the distance between the ith area and the kth knot, and Δ is a number used to normalise the distances so that the values of $\mathbf{B}$ are more evenly spread between the lower and upper limits. The knots were evenly spaced at intervals of 5 degrees of latitude and longitude, as shown in the figure below. Knots which were too distant from the centroids of SA2 areas were subsequently dropped. A total of 47 knots were retained for modelling. Based on these knots, Δ was set to 500.



This version of the P-spline uses a radial basis function which achieves rotational invariance (Ruppert *et al.*, 2003).

Local spatial smoothing

CAR dissimilarity model

Lee and Mitchell (2012) based this model on the Leroux CAR prior, with ρ set to be 0.99 to ensure strong global spatial smoothing which could then be altered locally through estimating $\{w_{ij}|i{\sim}j\}$. Here, the elements in **W** are modelled, so the partial autocorrelations can be reduced between certain adjacent random effects. This approach can have binary or non-binary elements in **W**.
The similarity between areas is determined by including non-negative dissimilarity metrics in the model, i.e. $\boldsymbol{z_{ij}} = (z_{ij1}, \ldots, z_{ijq})$ where $z_{ijk} = |z_{ik} - z_{jk}|/\sigma_k$ and σ_k is the standard deviation of $|z_{ik} - z_{jk}|$ over all pairs of contiguous areas.

The set of $\{w_{ij}\}$ are determined using regression parameters $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)$. These can be based on social or physical factors. Physical boundaries (e.g. river/railway line, or the distance between centroids) can be used if the aim is to explain the spatial pattern in the response and include covariates in the model. Alternatively, covariate information can be used to construct the dissimilarity metrics if the aim is to identify the locations of any boundaries (Lee, 2017).

$$R_i = S_i$$

$$S_i|\boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{0.99\sum_{j=1}^{N} w_{ij}(\boldsymbol{\alpha})s_j + 0.01\mu_0}{0.99\sum_j w_{ij}(\boldsymbol{\alpha}) + 0.01}, \frac{\sigma_s^2}{0.99\sum_j w_{ij}(\boldsymbol{\alpha}) + 0.01}\right)$$

The binary formulation:

$$w_{ij} = \begin{cases} 1 & \text{if} \exp\left(-\sum_{k=1}^{q} z_{ijk}\alpha_k\right) \geq 0.5 \text{ and } i \sim j \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_k \sim \text{Uniform}(0, M_k) \quad \text{for } k = 1, \dots, q$$

The non-binary formulation (which does not allow identification of hard boundaries, but does allow for localised smoothing):

$$w_{ij}(\boldsymbol{\alpha}) = \exp\left(-\sum_{k=1}^{q} z_{ijk}\alpha_k\right)$$

$$\alpha_k \sim \text{Uniform}(0,50) \quad \text{for } k = 1, \dots, q$$

Although this model can have covariates included, it appears as though there are advantages in excluding them so that the spatial structure is identical in both the risk surface and the random effects surface (Lee & Mitchell, 2012).

Localised autocorrelation

The spatially smooth random effects in this model are augmented with a piecewise constant intercept (cluster model). This allows for large jumps in the mean surface between adjacent areas if they are in different clusters. The approach by Lee and Sarran (2015) partitions the *I* areas into a maximum of G clusters, each with their own intercept term $(\lambda_1, \dots, \lambda_G)$ and is given by:

$$R_i = S_i + \lambda_{Z_i}$$

$$S_i|\boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{1}{\sum_j w_{ij}}\sum_j w_{ij}s_j, \frac{\sigma_s^2}{\sum_j w_{ij}}\right)$$

$$\lambda_g \sim \text{Uniform}(\lambda_{g-1}, \lambda_{g+1}) \quad \text{for } g = 1, \dots, G$$

$$f(Z_i) = \frac{\exp(-\delta(Z_i - G^*)^2)}{\sum_{r=1}^{G}\exp(-\delta(r - G^*)^2)}$$

$$\delta \sim \text{Uniform}(1, M)$$

where $f(Z_i)$ denotes a shrinkage prior on Z_i which penalises the values towards the middle intercept value so that the extremes of 1 or G may be empty. Label switching is prevented by ordering the cluster means $(\lambda_1, \dots, \lambda_G)$ so that $\lambda_1 < \lambda_2 <..< \lambda_G$. The penalty term $\delta(Z_i - G^*)^2$ where $G^* = (G+1)/2$ means that if G is odd then each data point will be shrunk towards a single intercept λ_{G*}, but if even, there may be two different intercept terms used even if there is a spatially smooth residual structure. Lee and Sarran (2015) thus recommend setting G to be a small odd number, such as 3 or 5. Area i is assigned to one of the G intercepts by $Z_i \in \{1, \dots, G\}$, and there is no spatial smoothing occurring on the indicator vector $\boldsymbol{Z}$.

The clustering is purely non-spatial, and it is the CAR prior on the S_i term that accounts for spatial autocorrelation (Lee & Sarran, 2015).

Locally adaptive model

A similar approach to the above dissimilarity model, except that here the boundaries are not identified by the use of additional information and the modelled w_{ij} are binary only. Lee and Mitchell (2013) again based this on the Leroux CAR model (with $\mu_0 = 0$):

$$S_i | \boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{\rho \sum_{j=1}^{N} w_{ij} s_j}{\rho \sum_{j=1}^{N} w_{ij} + 1 - \rho}, \frac{\sigma_s^2}{\rho \sum_{j=1}^{N} w_{ij} + 1 - \rho} \right)$$

Here ρ can be estimated in the model, or fixed at a specified value (Lee and Mitchell (2013) recommend 0.99).

The spatial weights matrix starts out as the binary, first-order, adjacency matrix where

$$w_{ij} = \begin{cases} 1 & \text{if areas } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases},$$

but this matrix is updated at each iteration which allows the weights corresponding to neighbours to be estimated as either 1 or 0 (but w_{ij} is fixed at zero for non-neighbouring areas). Lee and Mitchell (2013) recommend using INLA to make the computation more feasible. Because only weights corresponding to neighbouring areas are estimated, this approach should be more computationally feasible than areal wombling (Lu *et al.*, 2007) where all values in **W** are estimated. This approach "…allows the elements of **W** to be estimated, but without treating them as individual random quantities in an extra level of the Bayesian hierarchical model" (Lee & Mitchell, 2013). It is therefore not 'fully' Bayesian.

W is estimated as follows. For adjacent areas i and j: if the marginal 95% credible intervals of s_i and s_j overlap, then set $w_{ij} = 1$; else set $w_{ij} = 0$. For further details, refer to Lee and Mitchell (2013), who implemented this using INLA.

Hidden Potts model

This approach was proposed by Green and Richardson (2002) and is summarised by Best et al. (2005). The idea is to model the relative risk e^{μ_i}, or more generally, the spatial random effect on the log scale, as a K-component mixture model, where each component represents a different risk category, and the allocation of each area to a component follows a spatially correlated process. The number of components K is considered unknown and estimated by the model.

$$R_i = \log(S_{z_i})$$
$$S_k \sim \text{Gam}(a, b) \quad \text{for } k = 1, \dots, K$$
$$K \sim \text{Uniform}(1, K_{\max})$$

The Potts model is proposed as the allocation model.

$$p(\boldsymbol{z} \,|\psi, K) = \exp(\psi U(z) - \delta_k(\psi))$$

where $\psi > 0$ is the interaction parameter to be estimated and $U(z) = \sum_{i \sim j} \mathbb{I}(z_i = z_j)$ is the number of like labelled pairs of neighbouring areas.

Spatial partition model

Closely related to the above Hidden Potts model are the spatial partition models (Knorr-Held & Raßer, 2000, Denison & Holmes, 2001). These also have K non-overlapping clusters of areas, each with a constant relative risk, and K is unknown (Best et al., 2005).The key differences are in defining the clusters and the hyperprior specifications (Best et al., 2005). Both this model and the above hidden Potts model have been criticised for forcing discontinuities into a surface, and for assuming constant relative risk within a cluster (Lawson & Clark, 2002).

Weighted sum of spatial priors

The BYM model with its spatially structured component S_i and its unstructured spatial component U_i was extended to be able to detect discontinuities by Lawson and Clark (2002).

$$R_i = p_i S_i + (1 - p_i) Z_i + U_i$$

The $\mathbb{Z}$ component models abrupt discontinuities between areas. Although a range of options is possible, Lawson and Clark (2002) based this on the total absolute difference in risk between neighbouring areas, i.e.

$$\pi(Z_1, \dots, Z_I) \propto \frac{1}{\sqrt{\lambda}} \exp\left(-\frac{1}{\lambda} \sum_{i \sim j} |Z_i - Z_j| \right)$$

where λ acts as a constraining term. Note that if $p_i = 1$, then the model reverts to the BYM model. Conversely, if $p_i = 0$, then the model is entirely discontinuous.

Leroux scale mixture model

Congdon (2017) proposed using a scale mixture model within a Leroux prior, as follows:

$$S_i | \boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{\rho \sum_{j=1}^{N} \kappa_j w_{ij} s_j}{\rho \sum_{j=1}^{N} w_{ij} + 1 - \rho}, \frac{\sigma_s^2}{\kappa_i \left[\rho \sum_{j=1}^{N} w_{ij} + 1 - \rho \right]} \right)$$

If $\rho = 0$, this reduces to an unstructured iid scale mixture Student-t density. Small values of κ_j (<1) will indicate areas differ from their neighbours. The scale mixture is implemented by $\kappa_i \sim \mathrm{Gam}(0.5\nu, 0.5\nu)$, where ν is a hyperparameter.

The precision matrix has the following diagonal terms (Congdon, 2017):

$$\{\mathbf{T}\}_{ii} = \frac{1}{\sigma_s^2} \kappa_i \left[(1 - \rho) + \rho \sum_{j \neq i} w_{ij} \right]$$

and off-diagonal terms:

$$\{\mathbf{T}\}_{ij} = -\frac{1}{\sigma_s^2} \rho \kappa_i \kappa_j \mathrm{I}(i \sim j)$$

Local spline model

An extension to the global spline models described above that results in a less smooth surface was to incorporate unstructured random effects as in the penalised random individual dispersion effects (PRIDE) model, originally proposed by Perperoglou and Eilers (2010).

$$R_i = f(c_{1i}, c_{2i}) + \gamma_i$$

where γ_i is an area specific random effect, whose vector follows a multivariate normal distribution (Goicoa et al., 2012). This means that the covariance matrix captures the unstructured heterogeneity by containing an identity matrix multiplied by a variance component, in addition to the eigenvalues from the P-spline model component (Goicoa et al., 2012).

Skew-elliptical areal spatial model

Here

$$R_i = \eta_i^{-\frac{1}{2}}\left(\delta|Z_i| + S_i\right)$$

where $\delta|Z_i|$ is the skewing component where Z_i is a set of skewing variables each independently drawn from a standard normal distribution, η provides the scale mixing and S_i is from the CAR model, i.e.

$$S_i|\boldsymbol{s}_{\backslash i} \sim \mathcal{N}\left(\frac{\kappa}{\sum_j w_{ij}}\sum_j w_{ij}s_j\,,\ \frac{\sigma_s^2}{\sum_j w_{ij}}\right)$$

where κ is a spatial smoothing parameter (note that if it is set to 0 then the distribution corresponds to uncorrelated skew-t random effects) and other terms are defined as before.

Two versions were proposed by Nathoo and Ghosh (2013). The first aims to ensure each R_i has a skew-elliptical distribution, with the marginal distribution for each spatial effect belonging to the skew-t family of distributions.

The second is a semiparametric version that uses an approximation to a Dirichlet process to allow for data-driven departures from the parametric version. This accommodates uncertainty in the mixing structure, and gives greater flexibility in the tail behaviour of marginal distributions (Nathoo & Ghosh, 2013).

Summary of models

Of the models described above, we did not investigate further the localised form of splines, nor the proper CAR model. The disadvantages of the proper CAR formulation such as the potentially limited breadth of estimates have limited appeal for spatial modelling (Banerjee et al., 2003). The localised P-spline model was not investigated because the SIR estimates produced by the global smoothing P-spline models were already considerably less smooth than the BYM model (see Appendix A). Enabling additional variation therefore seemed unnecessary. Two formulations of the global P-spline model were implemented: the first uses a tensor product to define the basis, and the second uses a radial basis.

In addition, the CAR dissimilarity model can be applied in a variety of forms. The weighting matrix can be binary or non-binary, and the dissimilarity measure can be based on distance, geographical features (such as railways or mountains), or covariate information. In this report, we examine both binary and non-binary forms of this model based on Socioeconomic Indexes for Areas (SEIFA) dissimilarity, and also residual dissimilarity. While we are unaware of any published work using residuals for dissimilarity, we felt that this is an intuitive extension of the CAR dissimilarity model through identifying areas that differ from their neighbours under the specified model parameters.

We therefore attempted to apply each of the models in Table 2 to the simulated data. The BYM model was the most commonly cited model, but was also published much earlier than other models. A variety of software packages was used to run these models (Table 2).

Table 2: Citation counts and software used for attempted models

		Citations			
Models investigated	**Authors**	**Total**	**2017**	**2016**	**Software used**
Global spatial smoothing					
BYM	Besag et al., 1991	1486	113	137	R (CARBayes)
Leroux	Leroux et al., 2000	14	7	4	R (CARBayes)
Geostatistical model	Clements et al., 2006	115	6	12	JAGS (R2jags)
P-spline (tensor)	Lang & Brezger, 2004	302	34	22	JAGS (R2jags)
P-spline (radial)	Ruppert et al. 2003	58	9	7	
Local spatial smoothing					
CAR dissimilarity model	Lee & Mitchell, 2012	11	1	3	R (CARBayes)
Localised autocorrelation	Lee & Mitchell, 2013	16	3	4	R (CARBayes)
Locally adaptive model	Lee & Sarran, 2015	8	6	2	R (INLA)
Hidden Potts model	Green & Richardson, 2002	136	9	12	JAGS*
Spatial partition model	Knorr-Held & Raßer, 2000, Denison & Holmes, 2001	119 & 56	5 & 0	8 & 3	R*
Weighted sum of spatial priors	Lawson & Clark, 2002	50	3	7	WinBUGS
Leroux scale mixture model	Congdon, 2017	1	1	0	WinBUGS
Skew-elliptical areal spatial model	Nathoo & Ghosh, 2013	5	2	0	WinBUGS*

Notes: Citation counts as at 21st Dec 2017 using Web of Science (all databases).
*Model unable to run.

Results and Discussion

A summary of the models (and variants) run, along with key properties from each of the criteria (plausibility of estimates, model goodness of fit, and computational time) are in Tables 3 and 4. Further details on these components are available in Appendices A to D.

The small numbers for liver cancer means some smoothing is essential, yet it is expected that neighbouring areas will sometimes have genuine differences. Detecting these differences is problematic, and even many of the models designed to allow for local variation had results which were similar to that from the BYM and Leroux models, suggesting oversmoothing was occurring (Appendices A and G). Many of the models that obtained greater variation and less smoothing had excessive uncertainty around these estimates, such as the localised autocorrelation models (Appendix A). The appearance of maps differed quite substantially between models, whether examining the modelled SIRs or the posterior probabilities (PPs) and estimates for certain areas under different models could range from well below to well above the Australian average (Appendices A and G).

Although the amount of smoothing did vary between models for lung cancer, estimates were much more similar between different models overall than for liver cancer (Appendices A and G).

The amount of smoothing performed on all invasive cancers was surprising, given the high numbers. It is possible that the method of data simulation (see Appendix F) may have already somewhat smoothed the data, and resulted in less fluctuations in raw SIRs than would naturally be observed. For all invasive cancers, the raw SIRs ranged from 0.17 to 1.57, which would seem to not require much (if any) smoothing.

The models considered were all from the literature, albeit with certain modifications. For instance, we are not aware of others using residuals to determine which areas are dissimilar. However, these do show which areas are fitting poorly in comparison to their neighbours, and as such, may differ from them. The residuals used in the residual dissimilarity models were obtained from running a Leroux model with rho unfixed. This had high values for the male lung and liver cancers (rho of 0.87 and 0.93, respectively), and low values for female all invasive cancers (rho=0.04). Many alternatives are possible, including obtaining residuals from: a Leroux model where rho is fixed at 0.99, a BYM model, or a GLM model with no spatial structure incorporated. Determining the preferred form is currently being investigated.

Other model modifications tended to be more minor, and these are reflected in the supplied code provided in Appendix E. Models generated using MCMC were simplest to manipulate afterwards. Although INLA has the capacity to generate samples which can be used to approximate MCMC iterations, these are less straightforward, especially when wanting to combine multiple model parameters.

Table 3: Numeric summary of results across model criteria

	Mapped SIR		DIC	WAIC	Moran's I on residuals	Computation time (sec)
	Median	(Range)				
Liver cancer, males						
BYM	0.95	(0.38,1.34)	14,511.3	7,432.6	0.047	420.8
Leroux	0.95	(0.39,1.34)	14,523.9	7,429.1	0.047	309.5
Geostatistical	0.94	(0.51,1.84)	14,646.9	7,331.3	0.204	76,772.8
P-spline (tensor)	0.95	(0.13,1.87)	14,515.4	7,279.9	0.205	8,117.2
P-spline (radial)	0.93	(0.34,1.97)	14,760.3	7,402.6	0.241	1,960.1
SEIFA dissimilarity (binary)	0.94	(0.09,4.27)	14,360.3	7,354.9	0.042	8,462.3
SEIFA dissimilarity (non-binary)	0.94	(0.16,3.62)	14,315.4	7,361.2	0.031	5,316.0
Residual dissimilarity (binary)	0.86	(0.04,5.80)	13,162.8	6,779.9	0.091	8,650.5

	Mapped SIR		DIC	WAIC	Moran's I on residuals	Computation time (sec)
	Median	(Range)				
Residual dissimilarity (non-binary)	0.77	(0.13,5.41)	12,792.9	6,619.9	0.102	8,788.8
Localised autocorrelation (G=3)	0.96	(0.00,1.71)	13,571.9	8,107.0	0.054	1,035.6
Localised autocorrelation (G=5)	0.99	(0.00,1.61)	13,129.9	7,435.9	0.068	1,081.4
Locally adaptive (rho unfixed)	0.95	(0.38,1.34)			0.067	1,580.7
Locally adaptive (rho=0.99)	0.95	(0.38,1.34)			0.064	264.9
Weighted sum of spatial priors	0.98	(0.09,2.03)	14,421.8	7,414.2	0.094	6,815.7
Lung cancer, males						
BYM	1.00	(0.74,1.62)	23,680.8	12,148.9	0.131	419.4
Leroux	1.00	(0.74,1.61)	23,686.6	12,135.1	0.122	308.1
Geostatistical	0.98	(0.69,1.62)	22,213.4	11,218.7	0.220	98,412.3
P-spline (tensor)	0.97	(0.71,1.57)	22,173.0	11,110.1	0.228	7,640.8
P-spline (radial)	0.97	(0.71,2.38)	22,167.1	11,120.7	0.254	1,831.6
SEIFA dissimilarity (binary)	1.00	(0.74,1.61)	23,681.4	12,132.1	0.121	4,617.5
SEIFA dissimilarity (non-binary)	0.98	(0.65,2.53)	22,947.1	11,698.9	0.120	9,188.9
Residual dissimilarity (binary)	0.98	(0.69,2.03)	22,905.4	11,608.7	0.212	8,927.4
Residual dissimilarity (non-binary)	0.97	(0.75,1.72)	22,739.1	11,547.8	0.223	9,308.4
Localised autocorrelation (G=3)	1.00	(0.74,1.62)	23,690.2	12,131.6	0.124	898.4
Localised autocorrelation (G=5)	1.00	(0.74,1.61)	23,690.9	12,132.2	0.124	907.2
Locally adaptive (rho unfixed)	1.00	(0.74,1.61)			0.148	556.0
Locally adaptive (rho=0.99)	1.00	(0.74,1.61)			0.147	138.2
Weighted sum of spatial priors	1.00	(0.75,1.62)	23,761.2	12,218.1	0.182	6,553.6
Leroux scale mixture model	1.00	(0.78,1.52)	23,983.4	12,311.3	0.198	9,044.0
All invasive cancers, females						
BYM	1.00	(0.97,1.10)	31,410.9	15,911.1	0.055	412.6
Leroux	1.00	(0.98,1.08)	31,354.2	15,842.2	0.067	306.9
Geostatistical	1.00	(0.91,1.12)	31,504.9	15,868.3	0.171	91,365.4
P-spline (tensor)	1.00	(0.96,1.02)	31,409.0	15,825.6	0.108	7,634.6
P-spline (radial)	1.00	(0.66,1.30)	31,603.0	15,964.3	0.152	1,797.1
SEIFA dissimilarity (binary)	1.00	(0.98,1.08)	31,336.6	15,834.5	0.047	4,657.3
SEIFA dissimilarity (non-binary)	1.00	(0.98,1.08)	31,335.4	15,835.3	0.046	8,435.4
Residual dissimilarity (binary)	1.00	(0.90,1.43)	31,076.1	15,690.0	0.061	9,999.2
Residual dissimilarity (non-binary)	1.00	(0.91,1.38)	31,091.7	15,732.2	0.061	8,646.0
Localised autocorrelation (G=3)	1.00	(0.97,1.32)	31,245.1	15,938.0	0.042	768.0
Localised autocorrelation (G=5)	1.00	(0.97,1.36)	31,191.3	15,787.7	0.047	756.1
Locally adaptive (rho unfixed)	1.00	(0.99,1.04)			0.088	730.4
Locally adaptive (rho=0.99)	1.00	(0.99,1.04)			0.078	201.4
Weighted sum of spatial priors	1.00	(0.98,1.22)	31,305.4	15,890.2	0.100	6,291.5
Leroux scale mixture model	1.00	(0.98,1.32)	31,327.3	15,854.4	0.106	9,961.5

BYM= Besag, York and Mollié, DIC=Deviance Information Criterion, SEIFA= Socioeconomic Indexes for Areas, SIR=Standardised incidence ratio, WAIC= Watanabe-Akaike Information Criterion.

Notes:

- The lowest DIC, WAIC, Moran's I and computation time by type of cancer are shaded purple. Any Moran's I above 0.2 are shaded orange, as this is considered suggestive of some spatial correlation existing in the residuals.
- DIC and WAIC are unavailable for the locally adaptive results (which used INLA), as it was not possible to obtain approximate posterior samples to use in our calculations.
- Graphical comparisons of these goodness of fit and computational time results are available in Appendix B. SIR maps and graphs are available in Appendices A and G.

Table 4: Summary of model comparison based on criteria

Model [1]	Model specs (variant)	Plausibility of estimates	Model Fit	Model Complexity	Estimation Difficulty	Computation Speed [2]
Single global smoothing						
Intrinsic CAR (BYM)		Plausible, although consistently over-smooths.	Generally poor relative to other models.	Very simple	Easy	Very fast
Leroux		Plausible, although consistently over-smooths.	Slight improvement on BYM for all invasive cancers, otherwise about the same.	Very simple	Easy	Very fast
Geostatistical model		Plausible but overly precise; tends to have more extreme SIR values (less smoothing) than BYM model, but more smoothing than P-spline model.	Quite unpredictable - can be significantly better than BYM according to DIC and WAIC, but may also perform worse.	Somewhat complex (if modified)	Easy	Very slow (requires simplified model to achieve this very slow speed)
P-spline model	Tensor	Tends to have more extreme SIR values (less smoothing) than BYM model.	As above.	Complex	Easy	Average
	Radial	Possible under-smoothing.	As above.	Somewhat complex	Easy	Fast
Locally adaptive smoothing						
CAR dissimilarity	SEIFA Z, binary	Plausible; possible under- and over-smoothing.	Slight improvement on BYM in some cases.	Simple	Easy	Average
	SEIFA Z, non-binary	Plausible.	About the same as the binary version of dissimilarity model, but improved model fit for lung cancer.	Simple	Easy	Average
	Residuals Z, binary	Plausible; possible under-smoothing.	Significantly better model fit than BYM in all cases.	Simple	Easy	Average
	Residuals Z, non-binary	Plausible.	Significantly better model fit than BYM in all cases, and generally better than the all other versions of the dissimilarity model.	Simple	Easy	Average

Table 4 continued: Summary of model comparison based on criteria

Model [1]	Model specs (variant)	Plausibility of estimates	Model Fit	Model Complexity	Estimation Difficulty	Computation Speed [2]
Locally adaptive smoothing						
Localised autocorrelation	G=3	Mostly plausible, but large uncertainty around estimates and potential oversmoothing.	Offers potential improvement on BYM according to DIC, but very poor according to WAIC.	Simple	Easy	Very fast
	G=5	As above.	Same or better than G=3 version, and offers potentially significant improvement on BYM.	Simple	Easy	Very fast
Locally adaptive	Rho determined in model	Plausible, although consistently over-smooths	Based on DIC (results not included in report), better than BYM, poorer than dissimilarity.	Complex	Easy, but difficult to obtain sample estimates	Very fast
	Rho fixed at 0.99	Plausible, although consistently over-smooths	As above.	Complex	As above.	Very fast
Weighted sum of spatial priors		Plausible, although consistently over-smooths	About the same as BYM, or potentially slightly better.	Very simple	Easy	Average
Leroux scale mixture model		Plausible, although consistently over-smooths	About the same as BYM.	Simple	Easy, but would not run for liver cancer	Average

BYM= Besag, York and Mollié, CAR= Conditional Autoregressive, DIC=Deviance Information Criterion, SEIFA= Socioeconomic Indexes for Areas, SIR=Standardised incidence ratio, WAIC= Watanabe-Akaike Information Criterion.

[1] The Proper CAR mode, Hidden Potts model, spatial partition model, and skew-elliptical model are omitted since they were not run.

[2] Computation speed is based on average computation time across all three cancer groups. <1000s = very fast, 1000s to 5000s = fast, 5000s to 10000s = average, 10000s to 15000s = slow, and > 15000s = very slow.

Plausibility of estimates

'Well-behaved', reliable estimates (in terms of reasonable CIs) for every SA2 tend to result from models with more smoothing occurring (Appendix A). In contrast, some of the local models could have enormous CIs for many SA2s, resulting in little useful information. This differs from the geostatistical model, which seemed to produce unreasonably precise estimates.

The binary dissimilarity model is unlikely to produce reliable estimates for each area, as it tended to remove too many neighbours. This was true when basing on SEIFA, true to an even greater extent when using residuals, and likely to hold for other formulations, such as distance-based.

Model goodness of fit

The DIC and WAIC (Table 3 and Appendix B) measures of goodness of model fit were generally in consensus for a given cancer type. Some models seemed to fit the data well for certain types of cancer, but not others. For example, the geostatistical and P-spline models fit the lung cancer and all invasive cancer data sets quite well, but result in poor to average model fit for liver cancer. In fact, no model appears to fit all three cancer types consistently well. Models which provide a noticeably better fit to the data across at least two of the cancer data sets include the geostatistical and P-spline models, and the two residual dissimilarity models.

Moran's I statistic (Table 3 and Appendix B) was also computed on the model residuals to determine the degree of spatial autocorrelation not accounted for by the models, and generally indicated that the residual spatial autocorrelation is quite small. This measure can be very sensitive to the spatial weights matrix used to define the spatial dependencies between areas, and other spatial weights matrices (inverse-distance, third-order neighbours etc) were considered. However, it is difficult to gauge an appropriate structure for the spatial dependencies for the purpose of calculating residual spatial autocorrelation, and therefore this measure is given little importance in the process of ranking models by their goodness of model fit.

Computational time and feasibility

Although computational time is an important consideration, given that most models have multiple software options, this was our least important item. We ran models in the simplest software option, and found that CARBayes was generally very fast, and even WinBUGS produced model estimates in a timely manner (Tables 3 to 4 and Appendix B). The slowest model was the geostatistical model.

However, the feasibility of increasing model complexity – whether through the introduction of additional covariates or the extension of spatial data to spatio-temporal data – varies considerably between models. Models which exhibit relatively long computational times and are necessarily complex even without these extensions, such as the P-spline model, indicate poor feasibility. The mathematical structure of the geostatistical model is relatively simple, yet the long computation times also suggests that this model is not ideal for consideration of such model extensions.

Summary and Recommendations

We have compared 15 model variants within the Australian context of small-area cancer incidence mapping. See Table 5 for our initial recommendations.

The BYM and Leroux models both tend to oversmooth, but provide a reasonable model fit, and are computationally efficient to implement. The Leroux model may be preferred over the BYM model to avoid the inability of the BYM model to identify both the structured and unstructured spatial random effects separately. However, some of the other models demonstrated that model fit can be improved.

The geostatistical model may be able to achieve a better model fit than either the BYM or Leroux models. However, this model is prohibitively slow for the type of data being analysed here, both in terms of the number of areas, and diversity of area sizes. The unpredictable model fit and computational inefficiency do not make this model an attractive option.

The P-spline model had the potential to provide a reasonable option, and a comparison by Adin *et al.* (2017) of these against moving average and CAR models found the P-spline performed well for

sparse disease mapping. Although a local variant is available, we found the global P-spline model had much more variation in modelled estimates than any of the local models considered. However, although it is capable of detecting areas genuinely at higher risk, Goicoa et al. (2012) found the global P-spline model was prone to also detecting more false high-risk areas than either the CAR or a local P-spline model. Many spatio-temporal spline models are available (Anderson & Ryan, 2017), but currently the capacity to increase the P-spline model complexity in the Australian context is unclear. In terms of implementation, this model is rather complex, requiring specification of a penalty matrix and the number of knots, which are subjective and can have a large impact on model fit. Although guidelines are available for choosing the number of knots (see Wand (2000), Section 3), it is difficult to anticipate how these choices will impact fit. The main concern with the P-spline model, however, was the specification of the basis matrix using the tensor product, which does not adequately address the fact that the SA2s are irregular in shape and the distances between their centroids can be vastly different. The radial basis version of the P-spline model was designed to address this, but aside from being computationally faster, it provided similar levels of smoothing and a worse model fit.

A non-binary dissimilarity model may also be an option, as this smooths more than a P-spline but less than BYM or Leroux. The non-binary dissimilarity formulation using the SEIFA covariate worked quite well, however the application would preclude any further adjustment of the spatial estimates by a similar measure as a covariate within the model. Since one of the objectives of the Australian Cancer Atlas is to provide spatial estimates adjusted for area-level socioeconomic and remoteness, the socioeconomic dissimilarity model will not be further considered. The residual non-binary dissimilarity model appears promising but does have convergence issues for some cancer types, and the most appropriate way to generate the initial residual matrix is unclear. Further investigation will be undertaken.

In conclusion, the results of these analyses do not enable us to make a definitive statement about the best model for the Australian Cancer Atlas overall. Instead, after removing those that would not be considered further due to theoretical/practical issues, the three better performing models were found to be: P-spline (radial), Leroux and localised autocorrelation (G=5) (Table 6).

Certain models were excluded from Table 6. The geostatistical model was excluded due to practical difficulties with determining appropriate threshold values, as well as prohibitively slow computational times. The P-spline (tensor) model was excluded due to the assumptions about the geographical dependencies inferred by the tensor basis matrix. The SEIFA dissimilarity models were excluded on the grounds that it is preferential that SEIFA be included in the model as a covariate, rather than a variable for defining the dissimilarity metric. The locally adaptive models were excluded since they are difficult to evaluate using criteria such as model fit, plausibility of estimates, and over- and under-smoothing given the posterior estimates are obtained using INLA, and the standard INLA commands to obtain samples are not available due to it being called from within the CARBayes package. Additionally, the Leroux scale mixture model was excluded for the liver cancer in males because it was unable to run.

The goals of the Australian Cancer Atlas require a delicate balance between conservative, appropriately smoothed estimates, but ones with sufficient spatial variation between those estimates to enable the impact of suitable ecological covariates to be assessed. This report has demonstrated the dramatic influence that different models can have on very rare diseases, such as those for liver cancer among males. Given the different performances of models on different cancer types, we have proposed a preferred model given specific data characteristics from our top models (Table 6). The data characteristics considered were counts (low/moderate/high) and the presence of a trend across SEIFA quintiles (weak/strong), likely to represent evidence that neighbouring areas of different socioeconomic estimates could have genuine disparities in cancer influences and estimates. The criteria used to rank models were based on a modified form of our original criteria, emphasising the importance of convergence, plausible estimates and model fit. The preferred model for those with low/moderate counts and a strong SEIFA trend was the P-spline radial model, while the preferred model for high counts and weak SEIFA trend was the localised autocorrelation (G=5) model, closely followed by Leroux. When conducting spatial modelling, it is vital to consider data and area characteristics to ensure an appropriate model is used.

Table 5: Model comparison and recommendations

Model [1]	Model specs (variant)	Strengths	Limitations	Verdict
Single global smoothing				
Intrinsic CAR (BYM)		Straightforward, easy to implement, fast.	Tendency to oversmooth estimates. Non-identifiability of both random effects.	Consider further application.
Leroux		Straightforward, easy to implement, fast.	Tendency to oversmooth estimates.	Consider further application.
Geostatistical model		Simple model (unless modified), easy to implement.	A mixture model may be necessary for some parameters to improve model fit if areas are vastly different sizes. Model fit can vary between datasets dramatically. Computationally very slow, even after major simplification of the model. Does not scale well with number of areas.	Not recommended.
P-spline model	Tensor basis	Can provide a reasonable model fit due to flexibility of splines.	Rather complex model which can be quite difficult to understand and implement. Specification of certain parameters such as number of knots is subjective. Computational speed not particularly fast, and extensions to spatio-temporal data could make this prohibitively slow.	Not recommended.
	Radial basis	Fast	Rather complex model which can be quite difficult to understand and implement. Specification of certain parameters such as location and number of knots is subjective. Extensions to spatio-temporal data could reduce the fast computation substantially.	Consider further application.
Locally adaptive smoothing				
CAR dissimilarity	SEIFA Z, binary	Simple model, easy to implement.	Requires appropriate covariate information. May isolate areas to the extent that resulting estimates are very uncertain, or unable to converge. Not viable to use the same covariates in the model as are used in the dissimilarity matrix.	Not recommended.
	SEIFA Z, non-binary	Simple model, easy to implement.	Requires appropriate covariate information.	Consider further application only if no interest in adjusting for SEIFA in the model.
	Residuals Z, binary	Simple model, easy to implement; good model fit.	Prone to isolating areas, may have convergence difficulties. This is an introduced (and untested) model variant.	Not recommended.
	Residuals Z, non-binary	Simple model, easy to implement; superior model fit.	Moran's I on residuals suggested some spatial autocorrelation for lung cancer. This is an introduced (and untested) model variant.	Consider further application.

Table 5 continued: Model comparison and recommendations

Model [1]	Model specs (variant)	Strengths	Limitations	Verdict
Locally adaptive smoothing				
Localised autocorrelation	G=3	Simple model, easy to implement.	Uncertainty of parameter estimates can be very large, which can adversely affect model fit.	Not recommended.
	G=5	Simple model, easy to implement.	Uncertainty of parameter estimates can be very large, which can adversely affect model fit.	Consider further application.
Locally adaptive	Rho determined in model	Fully automatic.	Lacks the flexibility of post-estimation analyses available with MCMC chains. Tendency to oversmooth estimates. Not fully Bayesian estimation of W elements.	Not recommended.
	Rho fixed at 0.99	Very fast.	Lacks the flexibility of post-estimation analyses available with MCMC chains. Tendency to oversmooth estimates. Not fully Bayesian estimation of W elements.	Not recommended.
Weighted sum of spatial priors		Simple model.	Tendency for model to oversmooth estimates.	Not recommended.
Leroux scale mixture model		Allows for non-normality.	Tendency for model to oversmooth estimates. Only recently introduced.	Not recommended.

BYM= Besag, York and Mollié, CAR=Conditional autoregressive, SEIFA= Socioeconomic Indexes for Areas.

[1] The Proper CAR model, Hidden Potts model, spatial partition model, local spline model and skew-elliptical model are omitted since they were not run.

Table 6: Ranking of selected models

Model [1]	WAIC [2]	Computation Time [3]	Under-Smoothing [4]	Over-Smoothing [5]	CI Plausibility [6]	Convergence [7]	Consensus [8]
Liver cancer, males. Data characteristics are low counts, strong SEIFA quintile gradient.							
Intrinsic CAR (BYM)	11	1	1	11	0	1	25
Leroux	11	1	1	11	0	1	25
P-spline (radial)	11	4	3	1	1	0	20
CAR dissimilarity (residuals Z, binary)	3	20	6	5	1	17	52
CAR dissimilarity (residuals Z, non-binary)	1	20	5	4	1	2	33
Localised autocorrelation (G=3)	20	2	1	8	2	17	50
Localised autocorrelation (G=5)	11	2	1	9	2	1	26
Weighted sum of spatial priors	11	16	2	10	1	1	41

Table 6 Continued: Ranking of selected models

Model [1]	WAIC [2]	Computation Time [3]	Under-Smoothing [4]	Over-Smoothing [5]	CI Plausibility [6]	Convergence [7]	Consensus [8]
Lung cancer, males. Data characteristics are moderate counts, strong SEIFA quintile gradient.							
Intrinsic CAR (BYM)	18	1	2	7	0	1	29
Leroux	18	1	2	10	0	1	32
P-spline (radial)	1	4	9	3	1	0	18
CAR dissimilarity (residuals Z, binary)	9	20	7	5	0	16	57
CAR dissimilarity (residuals Z, non-binary)	8	20	5	4	0	1	38
Localised autocorrelation (G=3)	18	2	2	9	0	1	32
Localised autocorrelation (G=5)	18	2	2	8	0	1	31
Weighted sum of spatial priors	19	14	3	12	0	1	49
Leroux scale mixture model	20	20	1	13	0	0	54
All invasive cancers, females. Data characteristics are high counts, weak SEIFA quintile gradient.							
Intrinsic CAR (BYM)	17	1	4	5	0	1	28
Leroux	12	1	2	6	0	1	22
P-spline (radial)	20	4	7	2	0	0	33
CAR dissimilarity (residuals Z, binary)	1	20	6	4	0	1	32
CAR dissimilarity (residuals Z, non-binary)	4	18	5	3	0	1	31
Localised autocorrelation (G=3)	19	2	4	6	0	3	34
Localised autocorrelation (G=5)	8	2	4	6	0	1	21
Weighted sum of spatial priors	15	13	4	7	0	1	40
Leroux scale mixture model	12	20	3	7	0	0	42

BYM= Besag, York and Mollié, CAR=Conditional autoregressive.

[1] The Proper CAR model, Hidden Potts model, spatial partition model, local spline model and skew-elliptical model are omitted since they were not run. The Geostatistical model, P-spline (tensor) model, both SEIFA dissimilarity models, both locally adaptive models, and in the case of liver cancer, the Leroux scale mixture model, were excluded on theoretical/practical grounds.
[2] Rank based on ventiles (20-quantiles) of WAIC.
[3] Rank based on ventiles (20-quantiles) of computation time in seconds, excluding the time for the Geostatistical model.
[4] Rank based on the number of posterior SIRs > 98th percentile of raw SIRs plus number of zero posterior SIRs (since 2nd percentile is zero for all cancers). Smallest rank corresponds to smallest number.

[5] Rank based on the number of posterior SIRs > 80th percentile of raw SIRs plus number of posterior SIRs < 20th percentile of raw SIRs. Smallest rank corresponds to largest number.
[6] Ranks: 0 = reasonable; 1 = ok, but some areas are too wide/precise; 2 = bad (either too precise/wide.
[7] Ranks: 0 = ok (<1% <0.01); 1 = poor (1-<10% <0.01); 2 - 20 = bad: 10%+ <0.01 (Geweke % divided by 5 and rounded)
[8] The three best models by type of cancer based on the consensus rank are shaded purple.

References

Adin A, Martinez-Beneito MA, Botella-Rocamora P, Goicoa T, Ugarte MD, 2017. Smoothing and high risk areas detection in space-time disease mapping: a comparison of P-splines, autoregressive, and moving average models. *Stochastic Environmental Research and Risk Assessment,* 31, 403-415.

Anderson C, Ryan L, 2017. A Comparison of Spatio-Temporal Disease Mapping Approaches Including an Application to Ischaemic Heart Disease in New South Wales, Australia. *International Journal of Environmental Research and Public Health,* 14, 146.

Australian Bureau of Statistics, 2011. Australian Statistical Geography Standard (ASGS): Volume 1 - Main structure and greater capital city statistical areas, July 2011. *ABS Cat. No. 1270.0.55.001.* ABS, Canberra.

Australian Bureau of Statistics, 2013a. Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure, July 2011 *ABS Cat. No. 1270.0.55.005.* ABS, Canberra.

Australian Bureau of Statistics, 2013b. Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011. *ABS Cat. No. 2033.0.55.001.* ABS, Canberra.

Banerjee S, Carlin BP, Gelfand AE, 2003. *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press, London, UK.

Besag J, 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological),* 36, 192-236.

Besag J, York J, Mollie A, 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics,* 43, 1-59.

Best N, Richardson S, Thomson A, 2005. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research,* 14, 35-59.

Brunsdon C, Fotheringham AS, Charlton ME, 1996. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis,* 28, 281-298.

Celeux G, Forbes F, Robert CP, Titterington DM, 2006. Deviance information criteria for missing data models. *Bayesian Anal.,* 1, 651-673.

Clark PJ, Stuart KA, Leggett BA, Crawford DH, Boyd P, Fawcett J, et al., 2015. Remoteness, race and social disadvantage: disparities in hepatocellular carcinoma incidence and survival in Queensland, Australia. *Liver Int,* 35, 2584-2594.

Clements ACA, Lwambo NJS, Blair L, Nyandindi U, Kaatano G, Kinung'hi S, et al., 2006. Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Trop. Med. Int. Health,* 11, 490-503.

Congdon P, 2017. Representing spatial dependence and spatial discontinuity in ecological epidemiology: a scale mixture approach. *Stochastic Environmental Research and Risk Assessment,* 31, 291-304.

Denison DGT, Holmes CC, 2001. Bayesian partitioning for estimating disease risk. *Biometrics,* 57, 143-149.

Diggle PJ, 1990. A Point Process Modelling Approach to Raised Incidence of a Rare Phenomenon in the Vicinity of a Prespecified Point. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 153, 349-362.

Eberly LE, Carlin BP, 2000. Identifiability and convergence issues for Markov chain Monte Carlo fitting of spatial models. *Stat Med,* 19, 2279-2294.

Gelman A, Hwang J, Vehtari A, 2014. Understanding predictive information criteria for Bayesian models. *Statistics and Computing,* 24, 997-1016.

Geweke J, 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith (eds.), *Bayesian Statistics 4.* Oxford University Press, Oxford.

Goicoa T, Ugarte MD, Etxeberria J, Militino AF, 2012. Comparing CAR and P-spline models in spatial disease mapping. *Environmental and Ecological Statistics,* 19, 573-599.

Goovaerts P, 2005. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *Int J Health Geogr,* 4, 31.

Green PJ, Richardson S, 2002. Hidden Markov models and disease mapping. *Journal of the American Statistical Association,* 97, 1055-1070.

Kandhasamy C, Ghosh K, 2017. Relative risk for HIV in India – An estimate using conditional auto-regressive models with Bayesian approach. *Spatial and Spatio-temporal Epidemiology,* 20, 27-34.

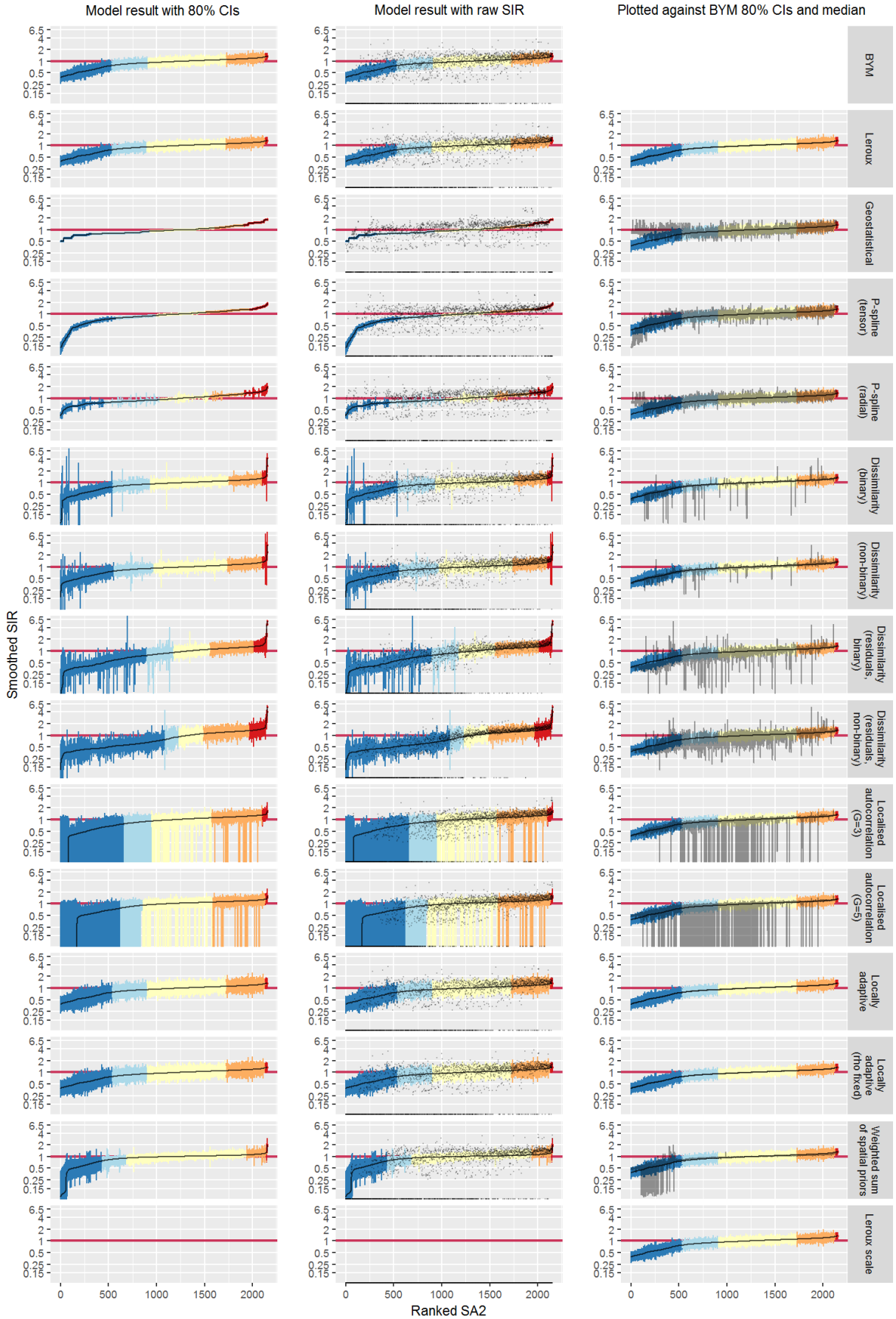
Kang SY, Cramb SM, White NM, Ball SJ, Mengersen KL, 2016. Making the most of spatial information in health: a tutorial in Bayesian disease mapping for areal data. *Geospatial Health,* 11, 428.

Knorr-Held L, Raßer G, 2000. Bayesian Detection of Clusters and Discontinuities in Disease Maps. *Biometrics,* 56, 13-21.
Lang S, Brezger A, 2004. Bayesian P-Splines. *Journal of Computational and Graphical Statistics,* 13, 183-212.
Lawson AB, Browne WJ, Vidal Rodeiro CL, 2004. *Disease Mapping with WinBUGS and MLwiN*. John Wiley & Sons, Ltd, Chichester.
Lawson AB, Clark A, 2002. Spatial mixture relative risk models applied to disease mapping. *Stat Med,* 21, 359-370.
Lee D, 2011. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial and Spatiotemporal Epidemiology,* 2, 79-89.
Lee D, 2017. CARBayes version 4.7: An R Package for Spatial Areal Unit Modelling with Conditional Autoregressive Priors. University of Glasgow, Glasgow.
Lee D, Mitchell R, 2012. Boundary detection in disease mapping studies. *Biostatistics,* 13, 415-426.
Lee D, Mitchell R, 2013. Locally adaptive spatial smoothing using conditional auto-regressive models. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 62, 593-608.
Lee D, Sarran C, 2015. Controlling for unmeasured confounding and spatial misalignment in long-term air pollution and health studies. *Environmetrics,* 26, 477-487.
Leroux BG, Lei X, Breslow N, 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: M. E. Halloran and D. Berry (eds.), *Statistical models in epidemiology, the environment and clinical trials.* Springer, New York.
Lu H, Reilly CS, Banerjee S, Carlin BP, 2007. Bayesian areal wombling via adjacency modeling. *Environmental and Ecological Statistics,* 14, 433-452.
MacNab YC, 2007. Spline smoothing in Bayesian disease mapping. *Environmetrics,* 18, 727-744.
Nathoo FS, Ghosh P, 2013. Skew-elliptical spatial random effect modeling for areal data with application to mapping health utilization rates. *Stat Med,* 32, 290-306.
Perperoglou A, Eilers PHC, 2010. Penalized regression with individual deviance effects. *Computational Statistics,* 25, 341-361.
Richardson S, 1992. Statistical methods for geographical correlation studies. In: P. Elliott, J. Cuzick, D. English and R. Stern (eds.), *Geographical and environmental epidemiology.* Oxford University Press, Oxford.
Riebler A, Sørbye SH, Simpson D, Rue H, 2016. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research,* 25, 1145-1165.
Ruppert D, Wand MP, Carroll RJ, 2003. *Semiparametric Regression*. Cambridge University Press, Cambridge.
Vehtari A, Gelman A, Gabry J, 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing,* 27, 1413-1432.
Wand MP, 2000. A Comparison of Regression Spline Smoothing Procedures. *Computational Statistics,* 15, 443-462.
Wheeler DC, 2014. Geographically Weighted Regression. In: M. M. Fischer and P. Nijkamp (eds.), *Handbook of Regional Science.* Springer Berlin Heidelberg, Berlin, Heidelberg.

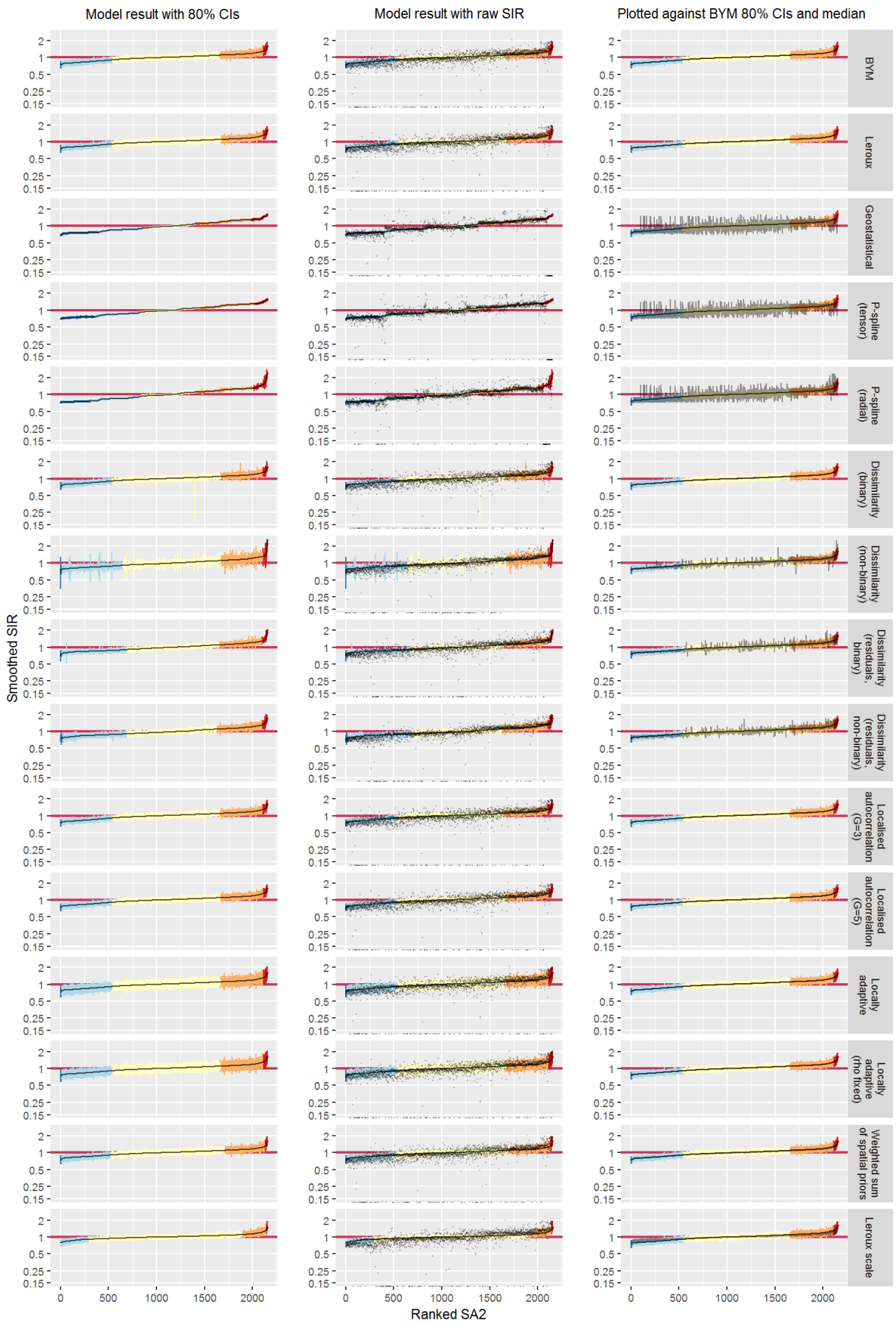
Appendix A: Graphs of model results

Note: Axes are consistent by cancer type.

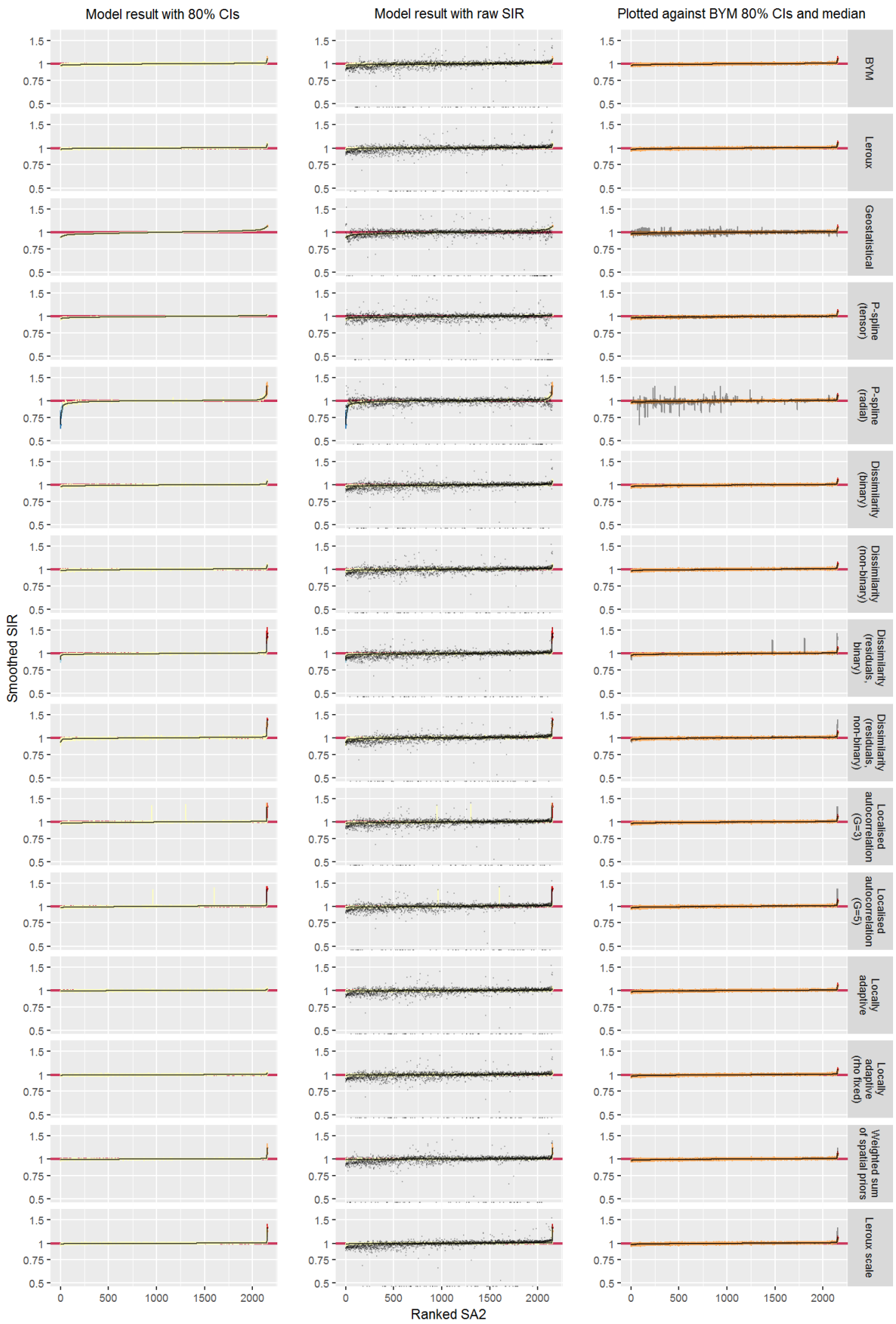
Liver cancer, males, modelled SIR



Lung cancer, males, modelled SIR



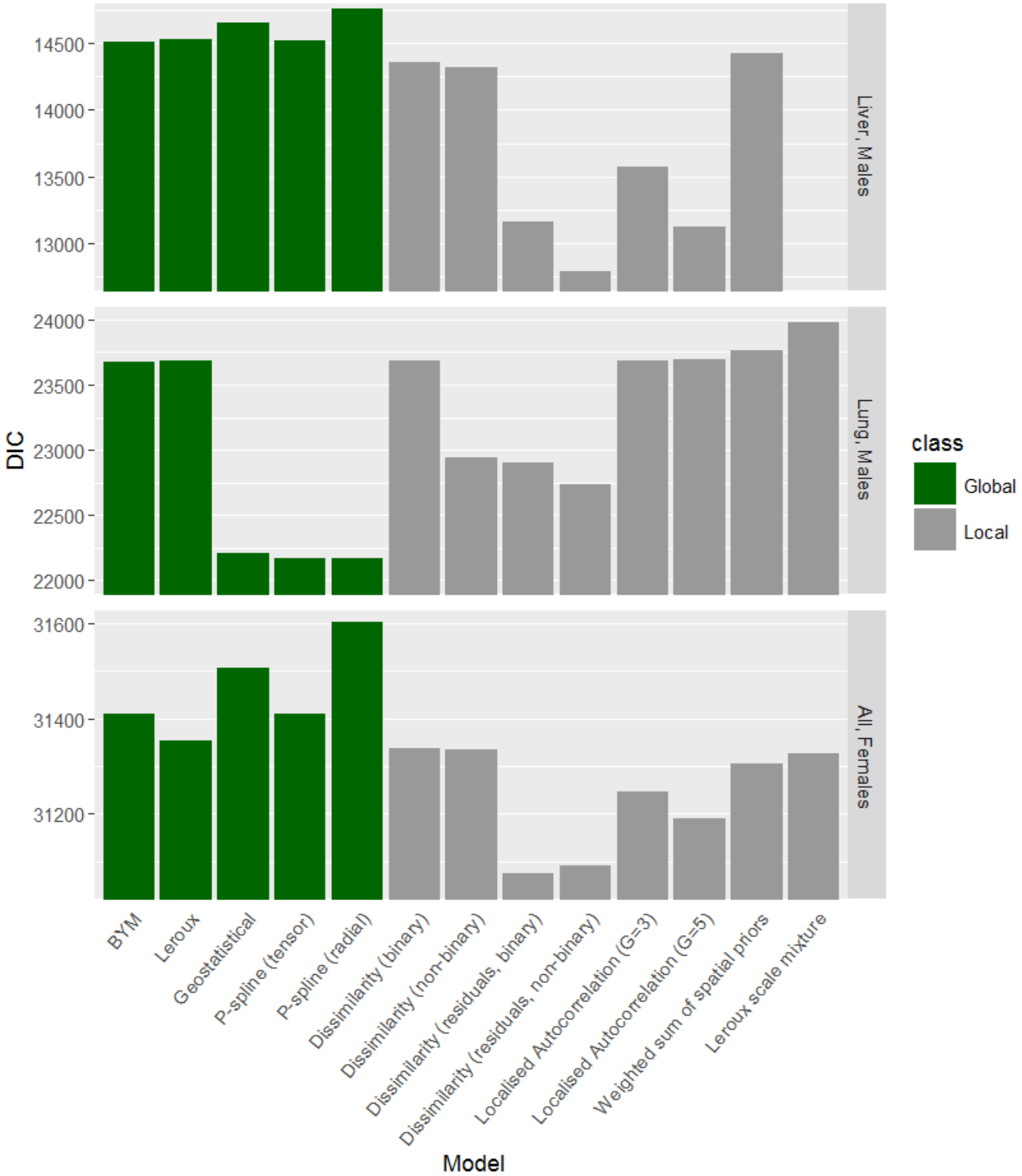
All invasive cancers, females, modelled SIR



Appendix B: Graphs of model fit and computation time

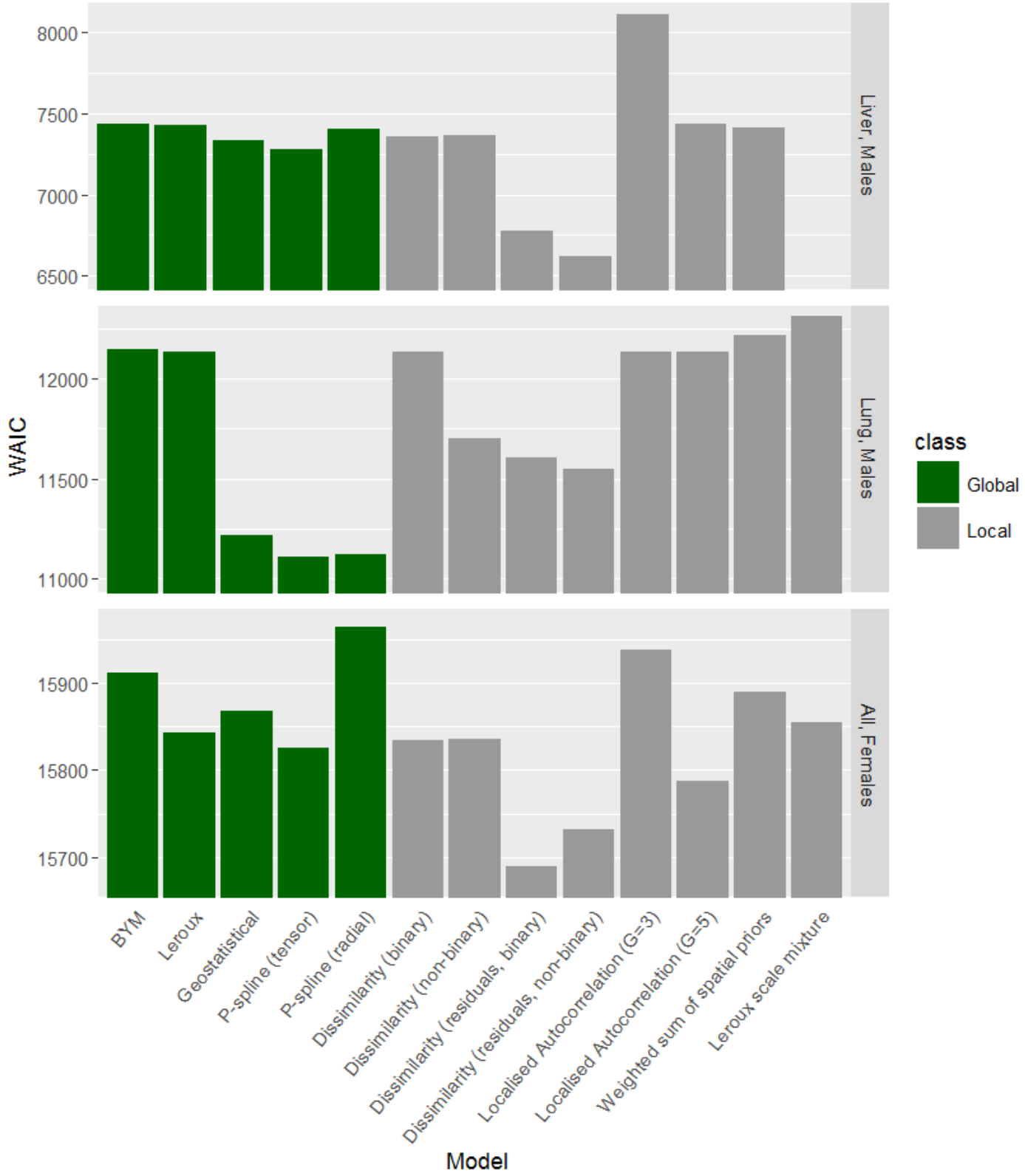
Deviance Information Criterion (DIC)

Note: smaller values of DIC indicate better model fit.



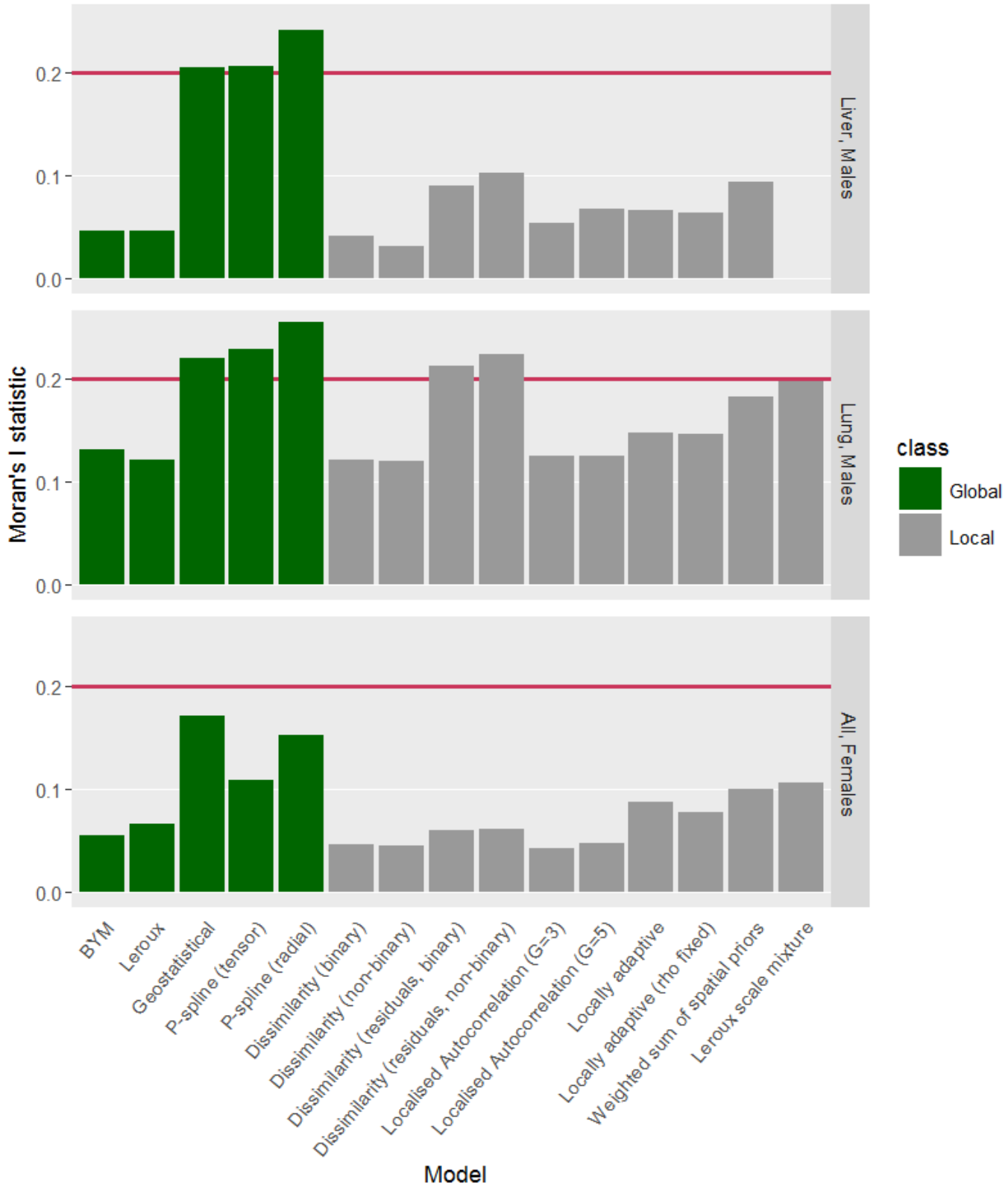
Watanabe-Akaike Information Criterion (WAIC)

Note: smaller values of WAIC indicate better model fit.



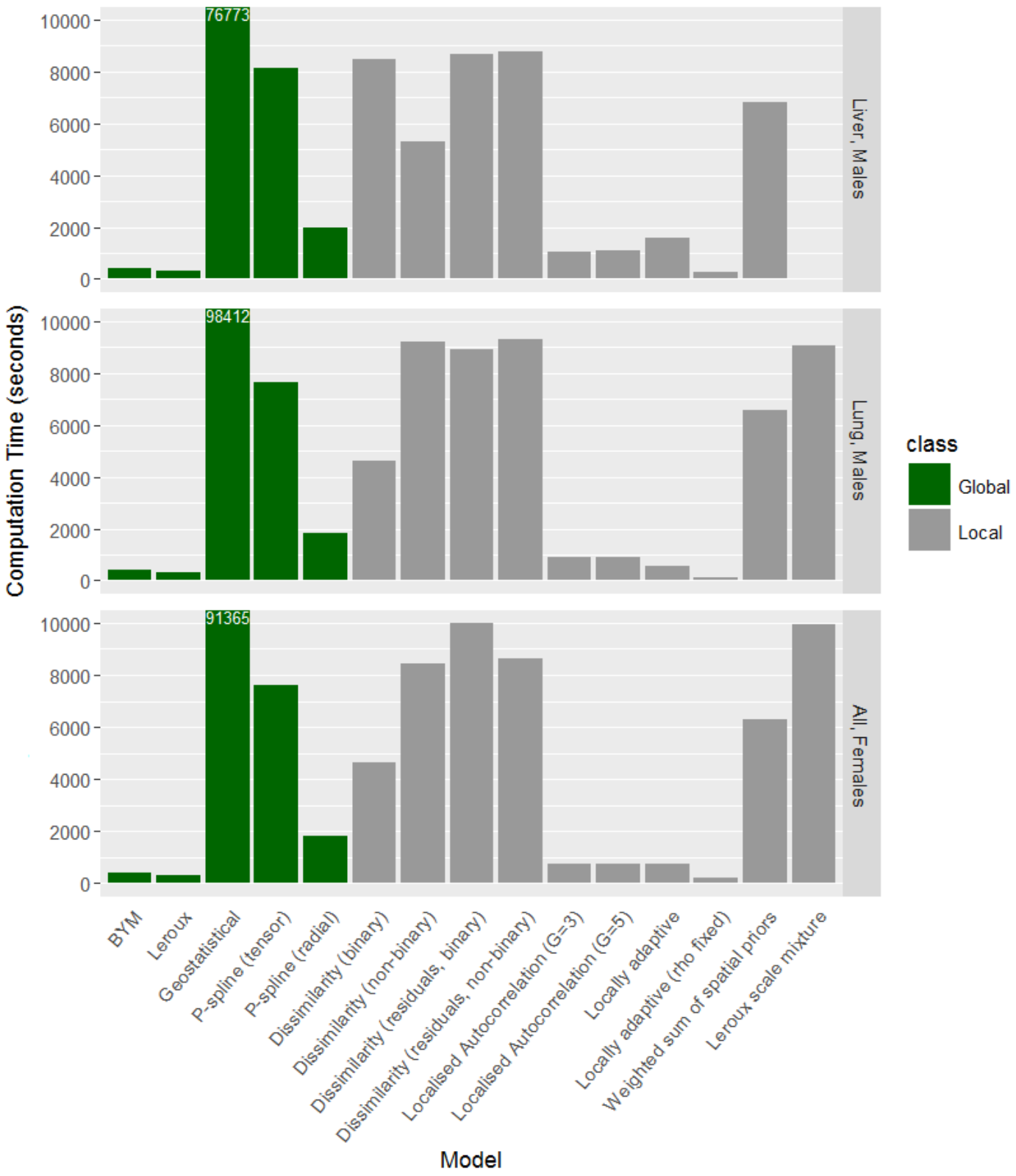
Moran's I statistic for the model residuals

Note: for the sake of comparability, the typical binary, first-order adjacency spatial weights matrix was used for all models. The closer Moran's I is to zero, the better the model fit. A horizontal line is shown at 0.2 as values above this may be indicative of spatial autocorrelation present in the residuals.



Computational time

Note: the computational time for the Geostatistical model greatly exceeds the time for any other model, even after modifying the model by imposing a MRF on the spatial structure. To facilitate comparisons between models with small computation time, the y-axis has been capped at 10000s while the computation time for the Geostatistical model, to the nearest second, is shown in the figure.



Appendix C: Convergence by model type

Geweke convergence diagnostic p-values on the SIR for each area out of a total of 2153 areas. Below 0.01 is unlikely to have converged, 0.01 to <0.05 may be worth examining plots. For further details on the Geweke convergence diagnostic, refer to Appendix F.

	<0.01		0.01-<0.05		0.05+	
Model	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**
Liver cancer, males						
BYM	93	4.3%	114	5.3%	1,946	90.4%
Leroux	100	4.6%	159	7.4%	1,894	88.0%
Geostatistical	3	0.1%	22	1.0%	2,128	98.8%
P-spline (tensor)	0	0.0%	78	3.6%	2,075	96.4%
P-spline (radial)	4	0.2%	32	1.5%	2,117	98.3%
SEIFA dissimilarity (binary weighting)	55	2.6%	187	8.7%	1,911	88.8%
SEIFA dissimilarity (non-binary weighting)	50	2.3%	123	5.7%	1,980	92.0%
Residual dissimilarity (binary weighting)	1,785	82.9%	96	4.5%	272	12.6%
Residual dissimilarity (non-binary weighting)	249	11.6%	179	8.3%	1,725	80.1%
Localised autocorrelation (G=3)	1,858	86.3%	72	3.3%	223	10.4%
Localised autocorrelation (G=5)	151	7.0%	168	7.8%	1,834	85.2%
Weighted sum of spatial priors	64	3.0%	141	6.5%	1,948	90.5%
Lung cancer, males						
BYM	25	1.2%	120	5.6%	2,008	93.3%
Leroux	33	1.5%	97	4.5%	2,023	94.0%
Geostatistical	14	0.7%	62	2.9%	2,077	96.5%
P-spline (tensor)	0	0.0%	116	5.4%	2,037	94.6%
P-spline (radial)	0	0.0%	9	0.4%	2,144	99.6%
SEIFA dissimilarity (binary weighting)	26	1.2%	100	4.6%	2,027	94.1%
SEIFA dissimilarity (non-binary weighting)	51	2.4%	106	4.9%	1,996	92.7%
Residual dissimilarity (binary weighting)	1,012	47.0%	210	9.8%	931	43.2%
Residual dissimilarity (non-binary weighting)	74	3.4%	135	6.3%	1,944	90.3%
Localised autocorrelation (G=3)	27	1.3%	101	4.7%	2,025	94.1%
Localised autocorrelation (G=5)	40	1.9%	104	4.8%	2,009	93.3%
Weighted sum of spatial priors	28	1.3%	104	4.8%	2,021	93.9%
Leroux scale mixture model	19	0.9%	88	4.1%	2,046	95.0%
All invasive cancers, females						
BYM	97	4.5%	159	7.4%	1,897	88.1%
Leroux	56	2.6%	132	6.1%	1,965	91.3%
Geostatistical	19	0.9%	149	6.9%	1,985	92.2%
P-spline (tensor)	40	1.9%	118	5.5%	1,995	92.7%
P-spline (radial)	0	0.0%	101	4.7%	2,052	95.3%
SEIFA dissimilarity (binary weighting)	97	4.5%	115	5.3%	1,941	90.2%

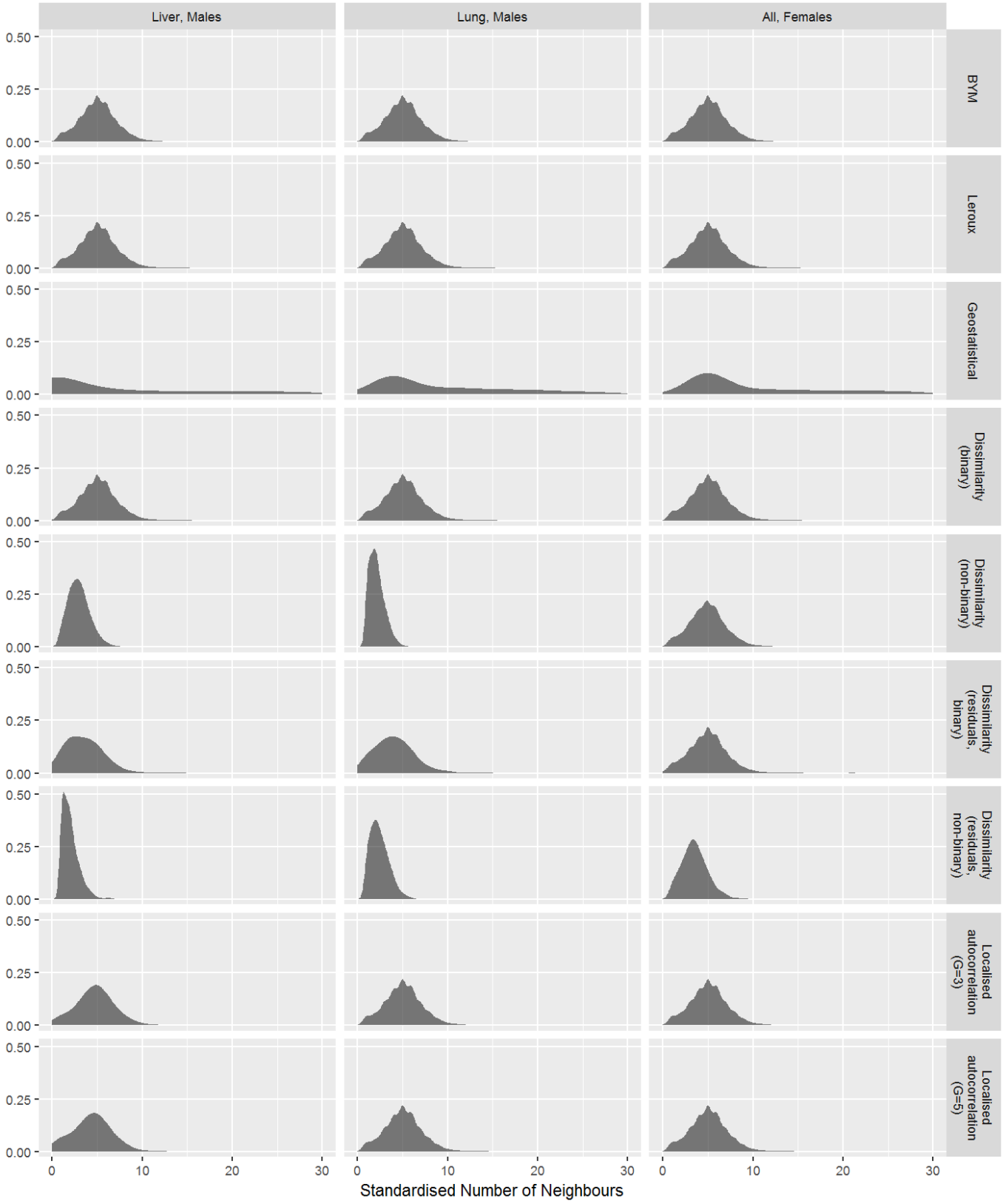
	<0.01		0.01-<0.05		0.05+	
Model	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**
SEIFA dissimilarity (non-binary weighting)	76	3.5%	127	5.9%	1,950	90.6%
Residual dissimilarity (binary weighting)	72	3.3%	191	8.9%	1,890	87.8%
Residual dissimilarity (non-binary weighting)	109	5.1%	162	7.5%	1,882	87.4%
Localised autocorrelation (G=3)	355	16.5%	218	10.1%	1,580	73.4%
Localised autocorrelation (G=5)	60	2.8%	136	6.3%	1,957	90.9%
Weighted sum of spatial priors	101	4.7%	190	8.8%	1,862	86.5%
Leroux scale mixture model	18	0.8%	95	4.4%	2,040	94.8%

Note: As the Poisson localised models were run in INLA, convergence is not applicable.

Appendix D: Changes to W for localised models

Standardised Number of Neighbours

Note: the values represent the sum of the weights in each row of the spatial weights matrix after standardising this matrix such that the maximum weight in each row is 1. This provides a more comparable and representative perspective of the neighbourhood structure used in each model.



Appendix E: Code for implementing selected models

The following code is supplied to:

1. Enable replication

2. Provide additional details on the models used, including the hyperprior specifications.

CARBayes R package, version 4.7

The default prior and hyperprior specifications were used for each of these models. This is inverse-gamma(1,0.01) for the variance of the spatial components.

BYM model

R code:

```
formula <-obs~offset(log(expect))

model.bym<-S.CARbym(formula=formula, data=file.input, family="poisson", W=W,
                    burnin=50000,
                    n.sample=150000, thin=10)
```

Leroux model

R code:

```
formula <-obs~offset(log(expect))

model.ler<-S.CARleroux(formula=formula, data=file.input, family="poisson", W=W,
                    burnin=50000,
                    n.sample=150000, thin=10)
```

SEIFA dissimilarity model (binary and non-binary)

The Z matrix was based on the socioeconomic index of relative disadvantage (IRSD) as a continuous score. See Appendix F for further details. For the non-binary version, W.binary is replaced with "FALSE" in the model call.

R code:

```
Z.irsd <- as.matrix(dist(cbind(irsd, irsd), method="maximum", diag=TRUE,
                    upper=TRUE))


formula <-obs~offset(log(expect))

model.diss<-S.CARdissimilarity(formula=formula, data=file.input, family="poisson", W=W,
                    Z=list(Z.irsd=Z.irsd), W.binary=TRUE, burnin=50000,
                    n.sample=150000, thin=10)
```

Residual dissimilarity model (binary and non-binary)

The Z matrix was based on the residuals output from the Leroux model as a continuous number. See Appendix F for further details.

R code:

```
Z.test <- as.matrix(dist(cbind(test, test), method="maximum", diag=TRUE,
                    upper=TRUE))

 formula <-obs~offset(log(expect))
 model.diss<-S.CARdissimilarity(formula=formula, data=file.input, family="poisson", W=W,
                                Z=list(Z.test=Z.test), W.binary=TRUE, burnin=50000,
                                n.sample=150000, thin=10)
```

Localised autocorrelation (G=3 and G=5)

For G=5, the G=3 is replaced with G=5 in the model call.

R code:

```
formula <-obs~offset(log(expect))
 model.local<-S.CARlocalised(formula=formula, data=file.input, family="poisson", G=3, W=W,
                             burnin=50000,
                             n.sample=150000, thin=10)
```

INLA

Locally adaptive

Note that the poisson.localisedINLA.R function is available from Duncan Lee on request. We made minor tweaks to this in the version run, including outputting the final W matrices.

R code:

```
# Rho is not fixed at a specific value
formula <-file.input$obs~offset(log(file.input$expect))

source("<filepath>/poisson.localisedINLA_.R")

model.ploc<-poisson.localisedINLA(formula=formula, W=W, fix.rho=FALSE)

#Rho is fixed at 0.99
model.ploc2<-poisson.localisedINLA(formula=formula, W=W, fix.rho=TRUE, rho=0.99)
```

WinBUGS version 1.4.3

Weighted sum of spatial priors

Based on code available on p. 134 (Lawson *et al.*, 2004).

WinBUGS code:

```
model {

for (i in 1:N) {
```

```
#Poisson likelihood for observed counts
O[i]~dpois(mu[i])

log(mu[i])<-log(E[i])+alpha+v[i]+p[i]*u[i]+(1-p[i])*fi[i]

#Prior distribution of the uncorrelated heterogeneity
v[i] ~ dnorm (0,tauv)

#Prior distribution of the p[i]
p[i]~dbeta(1,1) #Note that the original model had 0.5 as the parameters, but this was too
diffuse for WinBUGS, giving the error message "Cannot bracket slice for node p.."
}

# CAR prior distribution for spatial correlated heterogeneity
u[1 : N] ~ car.normal(adj[], weights[], num[], tauu)

# CAR prior distribution for spatial correlated heterogeneity
fi[1 : N] ~ car.l1(adj[], weights[], num[], taufi)

for(k in 1:sumNumNeigh) {
weights[k] <- 1
}

# Improper prior distribution for the mean RR in the study region
alpha ~ dflat()
mean<-exp(alpha)

#Hyperprior distributions on inverse variance parameter of random effects
tauu ~ dgamma(0.5, 0.0005)
tauv ~ dgamma(0.5, 0.0005)
taufi ~ dgamma(0.5, 0.0005)
}
```

Leroux scale mixture model

The following is a modified version of code obtained from Peter Congdon on request.

WinBUGS code:

```
model {

for (i in 1:N) {
O[i]~dpois(mu[i])

log(mu[i])<-log(E[i])+alpha+r[i]

r[i] ~dnorm(R[i],taur[i])
taur[i] <- tau * kap[i] * (1-lam+lam*num[i])
R[i] <- (lam/(1-lam+lam*num[i]))*sum(rneigh[cum[i]+1:cum[i+1]])
kap[i]~dgamma(nu2, nu2)
}

#  error vector over neighbours
for (k in 1:sumNumNeigh) {
rneigh[k] <- r[adj[k]]*kap[adj[k]]
}

#Other priors
nu~dexp(2.5) #Note the original dexp(0.1) caused the error message: cannot bracket slice for
node nu. And nu was set at 10 in a simulation study in Congdon's paper.
nu2<-nu/2
```

```
        tau ~dgamma(1,0.01)
        lam ~dunif(0,1)
        alpha ~ dnorm (0,1.0E-5)
}
```

Skew-elliptical areal spatial model

This code is based on that in the Appendix of Nathoo and Ghosh (2013). Note that this model would hang on the compile step in WinBUGS, so could not be run. It is possible a simplified version (non-DP) might work, but Farouk Nathoo was unable to supply the code for this version, and we were unable to modify the code within the necessary timeframes.

```
model {

for (i in 1:N) {
                O[i]~dpois(mu[i])
                log(mu[i])<-log(E[i])+alpha+b[i]-bmean
        }
                bmean<-mean(b[])

#Define skew-elliptical spatial random effects
for (i in 1:N) {
                b[i]<-(X[i]/sqrt(eta[i]))
                thetaX[i]<-deltaskew*abs(Z[i])
                m[i]<-1/num[i]
                #mixing variables based on semiparametric DP model
                Z[i]~dnorm(0,1)
                eta[i]<-wgeta[zg[i]]
                zg[i]~dcat(pi1[1:maxclus])

        for (j in 1:maxclus) {
                Memb1[i,j]<-equals(zg[i],j)
                }
        }

for(k in 1:sumNumNeigh){
        for(i in 1:N){
                pick[k,i]<-step(k-cum[i]-epsilon)*step(cum[i+1]-k)
                }
        C[k]<-1/inprod(num[], pick[k,]) #weight for each pair of neighbours
        }
epsilon<-0.0001

#no of nonempty clusters in DP
for (j in 1:maxclus) {
                TMemb1[j]<-sum(Memb1[,j])
                FMemb1[j]<-step(TMemb1[j]-1)
        }

                K1<-sum(FMemb1[])

#Base distribution for DP
for (j in 1:maxclus) {
                r1[j]~dbeta(1,alpha11)
                wgeta[j]~gen.gamma(df,df,1)
        }

#Stick-breaking for DP
pi1[1]<-r1[1]
for (j in 2:(maxclus-1)) {
        log(pi1[j])<-log(r1[j])+sum(R1[j,1:j-1])
        for (l in 1:j-1) {
                R1[j,l]<-log(1-r1[l])
                }
```

```
        }

#Ishwaran truncation approximating full DP
pi1[maxclus]<-1-sum(pi1[1:(maxclus-1)])

#Priors
#Hyperprarameter for DP
alpha11~dunif(0.5,4)
#Spatial smoothing
X[1:N]~car.proper(thetaX[], C[], adj[], num[], m[], taux, kappa)
#Spatial variability
taux~dgamma(0.5,0.0005)
#Skew parameter
deltaskew~dnorm(0,0.01)
#Spatial smoothing parameter in CAR model
kappa~dbeta(18,2)I(,0.99)
#Degrees of freedom in skew-t distribution
df<-nu/2
nu~dexp(lambdanu)I(2,)
lambdanu<-0.1
#Regression coefficient
alpha~dnorm(0,0.001)
}
```

JAGS using R2jags R package, version 0.5-7

Geostatistical Model

JAGS code:

```
model{
  for(i in 1:N){
    y[i] ~ dpois(E.RR[i])
    E.RR[i] = E[i] * exp(mu[i])
    mu[i] = alpha + R[i] + beta * x[i]
    R[i] ~ dnorm(S[i], tau)
    S[i] = sum(S.all[i,]) / N.i[i]          # Mean of non-zero values in each row
    for(j in 1:max(N.i)){
      S.all[i,j] = exp(-pow((lambda * d[i,j]), k[z[i]]))
    }
  }
  alpha ~ dnorm(0, 0.01)
  beta ~ dnorm(0, 0.01)
  lambda ~ dunif(0.001, 2)
  for(r in 1:R){
    k[r] ~ dunif(0.1, 20)
  }
}
```

P-spline Model (Global; tensor and radial)

JAGS code:

```
model{
  for (i in 1:N){
    y[i] ~ dpois(E.RR[i])
    E.RR[i] <- E[i] * exp(mu[i])
    mu[i] = alpha + S[i] + beta * x[i]
    S[i] = B[i, ] %*% theta[1:K]
  }
  alpha ~ dnorm(0, 0.01)
  beta ~ dnorm(0, 0.01)
```

```
  theta[1:K] ~ dmnorm(phi * ones, lambda * Q[1:K,1:K])

  phi ~ dnorm(0, 1e-6)
  lambda ~ dgamma(0.001, 0.001)
}
```

Appendix F: Model methods

Data simulation

Queensland cancer incidence data during 2005-2014 were summed together by cancer type (all, prostate, lung, kidney, cervical, liver), 5-year age group (to 85+ years), sex (males, females), the ABS's five remoteness areas (RAs, ranging from Major city to Very remote) (Australian Bureau of Statistics, 2013a) and (Australian) Socioeconomic Indexes for Areas (SEIFA) Index of Relative Socioeconomic Disadvantage (IRSD) quintiles (Australian Bureau of Statistics, 2013b). The corresponding population data were applied to obtain age-specific rates. These rates were then applied to each SA2 throughout Australia based on the SA2's remoteness and SEIFA classifications. When a SEIFA-RA combination existed in another state that was not in Queensland, the rate for the middle SEIFA quintile and the same RA was applied. Further, for all areas with a missing SEIFA value (even if the remoteness-SEIFA combination existed in Queensland), the rate for the middle SEIFA quintile and the corresponding RA was applied.

Rates were then multiplied by the population (by sex and age group) in each SA2 to generate counts. Within each state (including Queensland), counts were randomly adjusted by up to +/- 10%, and rounded to give integers.

Checks against ACIM suggested total Australian estimates (and allowing for different years) were reasonable, although cervical and especially liver cancers were rather low after rounding.

Expected counts

As internal standardisation was used, the Australian age-sex specific rate was multiplied by the SA2 age-sex specific population. These were then summed together to obtain the expected number of cases in each SA2 by sex.

Geographical areas

Lord Howe Island lies 600km off the coast of NSW, and was removed as considered too far away from the coastline to be influenced or exert influence. All SA2s with a zero or very low estimated resident population during 2005-2014 were also removed. These were often lakes, reservoirs, airports, military, industrial or national parks.

Dissimilarity model Z matrices

SEIFA IRSD

The SEIFA IRSD Score (Australian Bureau of Statistics, 2013b) for each SA2 was used. This is a continuous value that, across all of Australia, ranged from 440.67 to 1147.96.

Residuals

After calculating the Leroux CAR model, the residuals output from the CARBayes package (which are calculated as the observed values minus the fitted values) for each SA2 were used. For liver cancer these ranged from -3.23 to 6.71, lung had a range of -14.72 to 18.55 and all invasive of -54.38 to 115.36.

DIC, WAIC, Moran's I on residuals

These were generated using an author-written R-program. There are several variants of DIC, and the definition used was based on that in Celeux *et al.* (2006), specifically:

$$-4\mathbb{E}_\theta[\log f(y|\theta)|y] + 2\log f(y|\mathbb{E}_\theta[\theta|y])$$

where $f(y|\theta)$ is the likelihood, and the posterior mean is represented as: $\mathbb{E}_\theta[\theta|y]$.

WAIC was defined as follows (Gelman *et al.*, 2014):

$$-2 \times \left(\sum_{i=1}^{N} \log \int p(y_i|\theta)\, p_{\text{post}}(\theta)\, d\theta - \sum_{i=1}^{N} \text{var}_{\text{post}}(\log p(y_i|\theta)) \right)$$

where the predictive density is $p(y_i|\theta)$ and the posterior distribution is $p_{\text{post}}(\theta) = p(y_i|\theta)$ and var_{post} refers to the posterior variance.

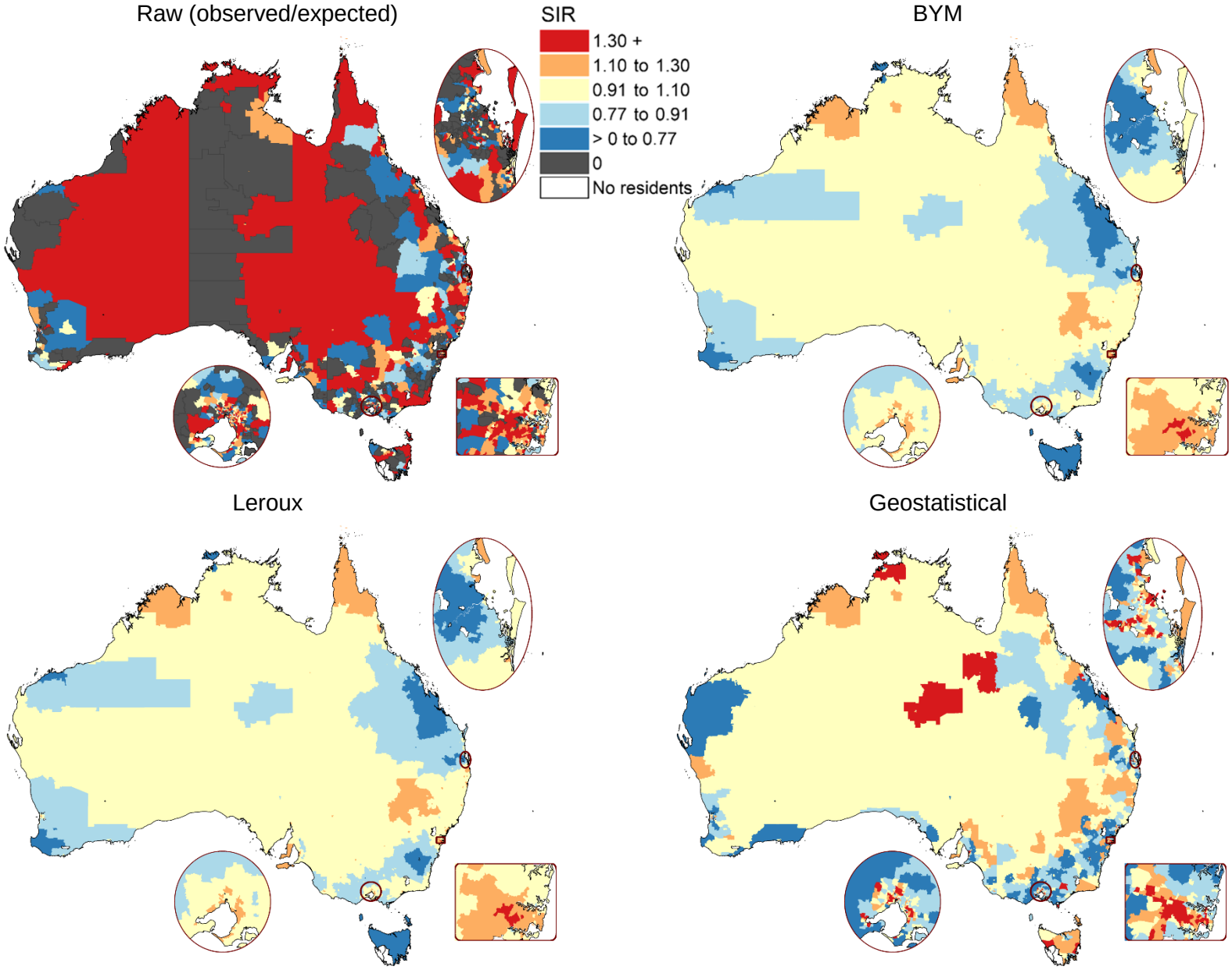
The DIC equation gave substantially higher estimates than obtained under default software options (CARBayes, JAGS or INLA). Nonetheless, as the exact formulation of this can differ, it was preferred to use these results to ensure consistency of calculation between models. In contrast, WAIC and Moran's I estimates were very similar to those obtained under default software options. For the Poisson localised models (run in INLA), the Moran's I reported is the default output from the software. All Moran's I estimates are based on the default first-order Queen neighbourhood adjacency matrix.

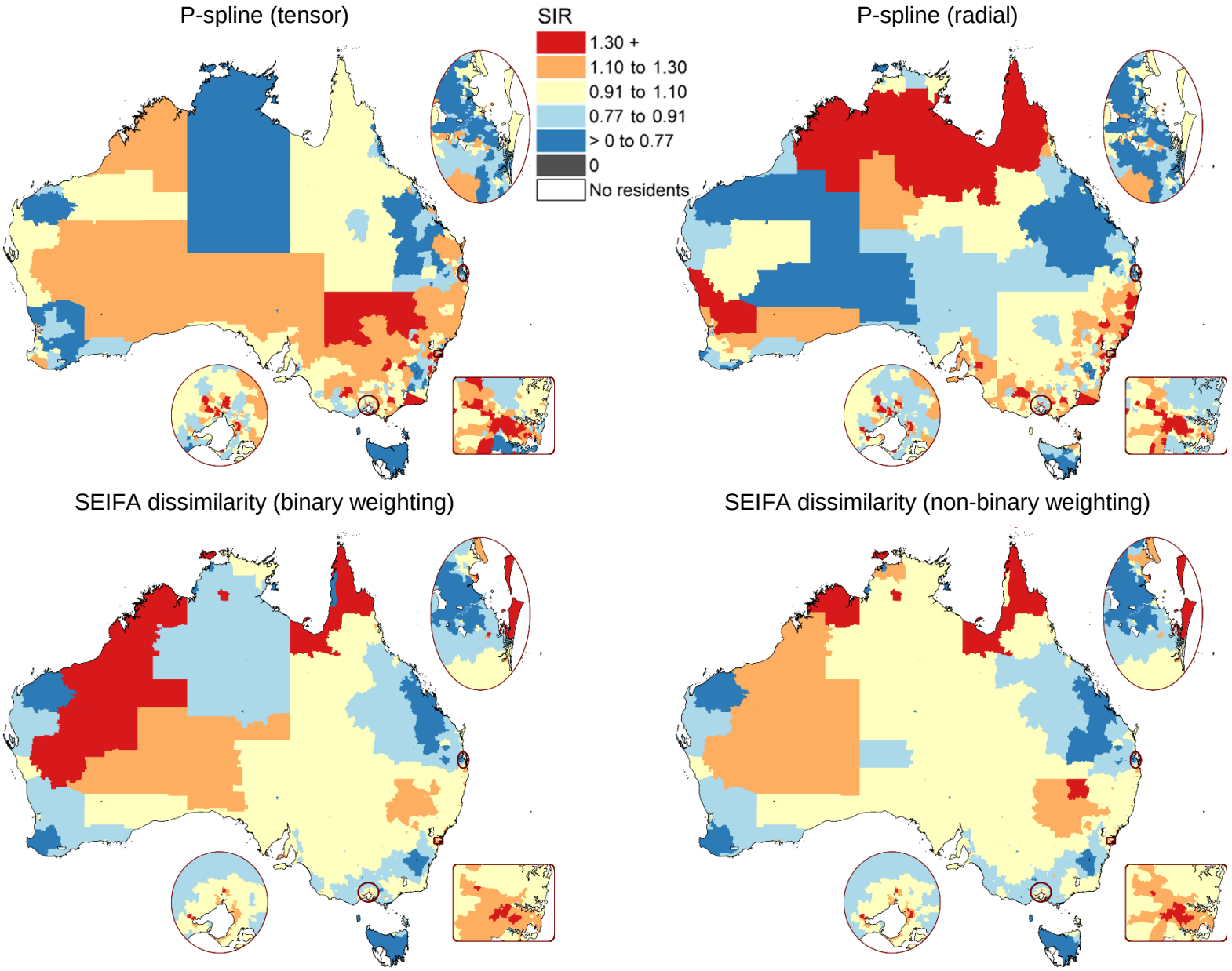
Geweke convergence diagnostic

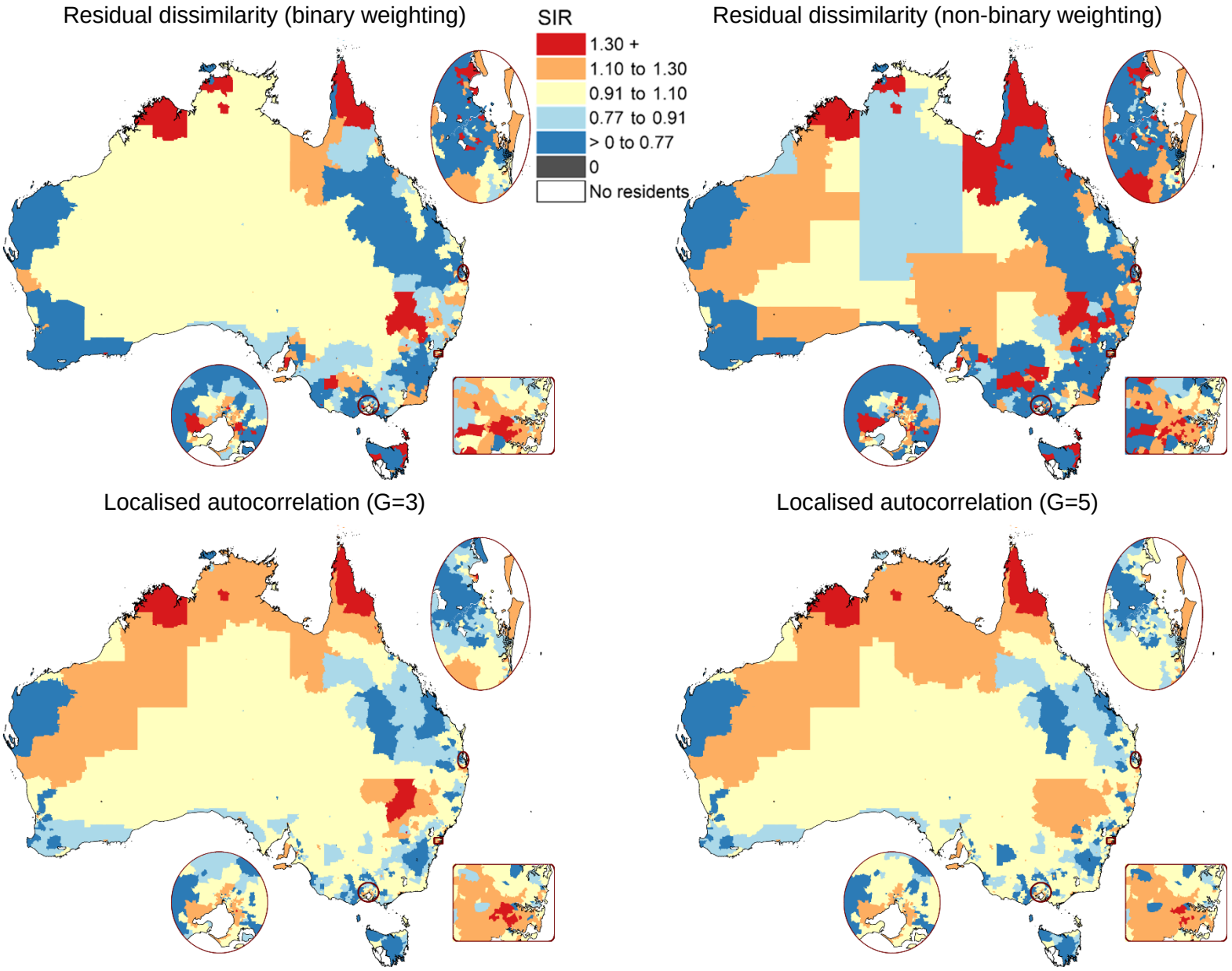
Some trace and autocorrelation plots were examined to check convergence, but given the high numbers of areas, and the use of single chains, the Geweke convergence diagnostic (Geweke, 1992) was calculated on the SIR for all included SA2s. This was calculated using the "wbgeweke" command in Stata v15.0, which compares the first 10% of iterations against the final 50%.

Appendix G: Maps of model results

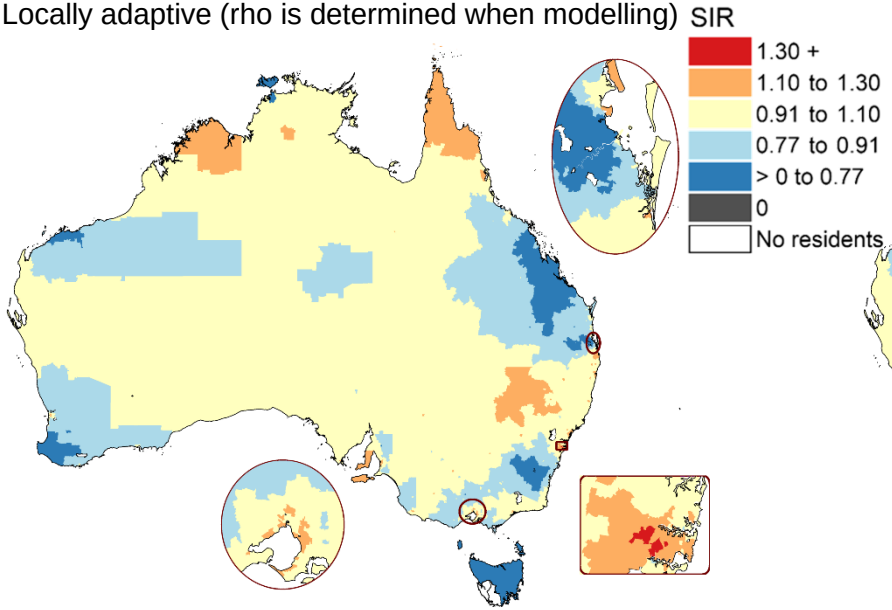
Liver cancer, males, modelled SIR



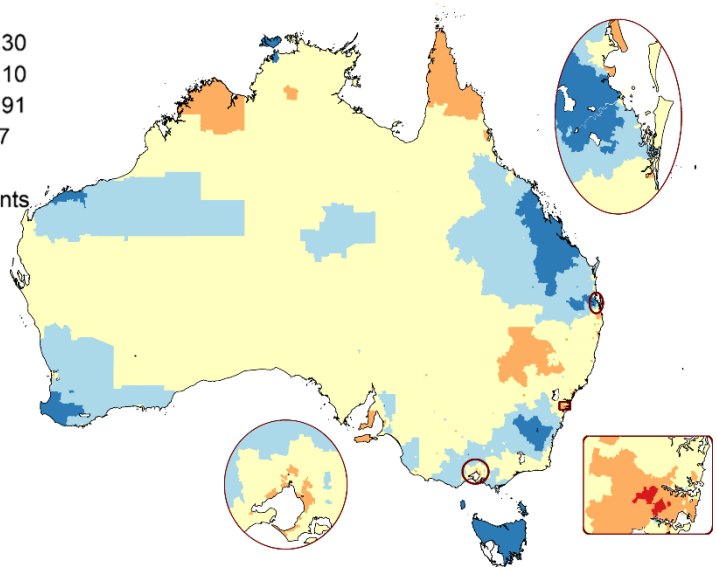

P-spline (tensor)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)


Residual dissimilarity (binary weighting)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
Residual dissimilarity (non-binary weighting)
Localised autocorrelation (G=3)
Localised autocorrelation (G=5)

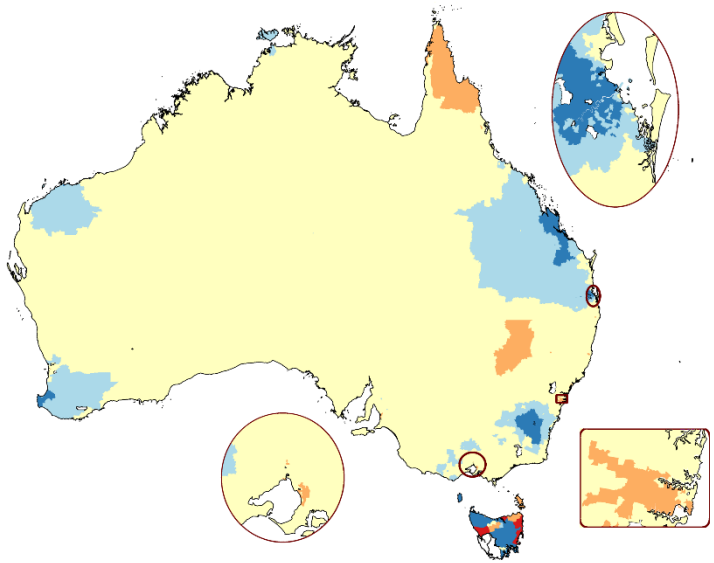
Locally adaptive (rho is determined when modelling)



Locally adaptive (rho is fixed at 0.99)



Weighted sum of spatial priors

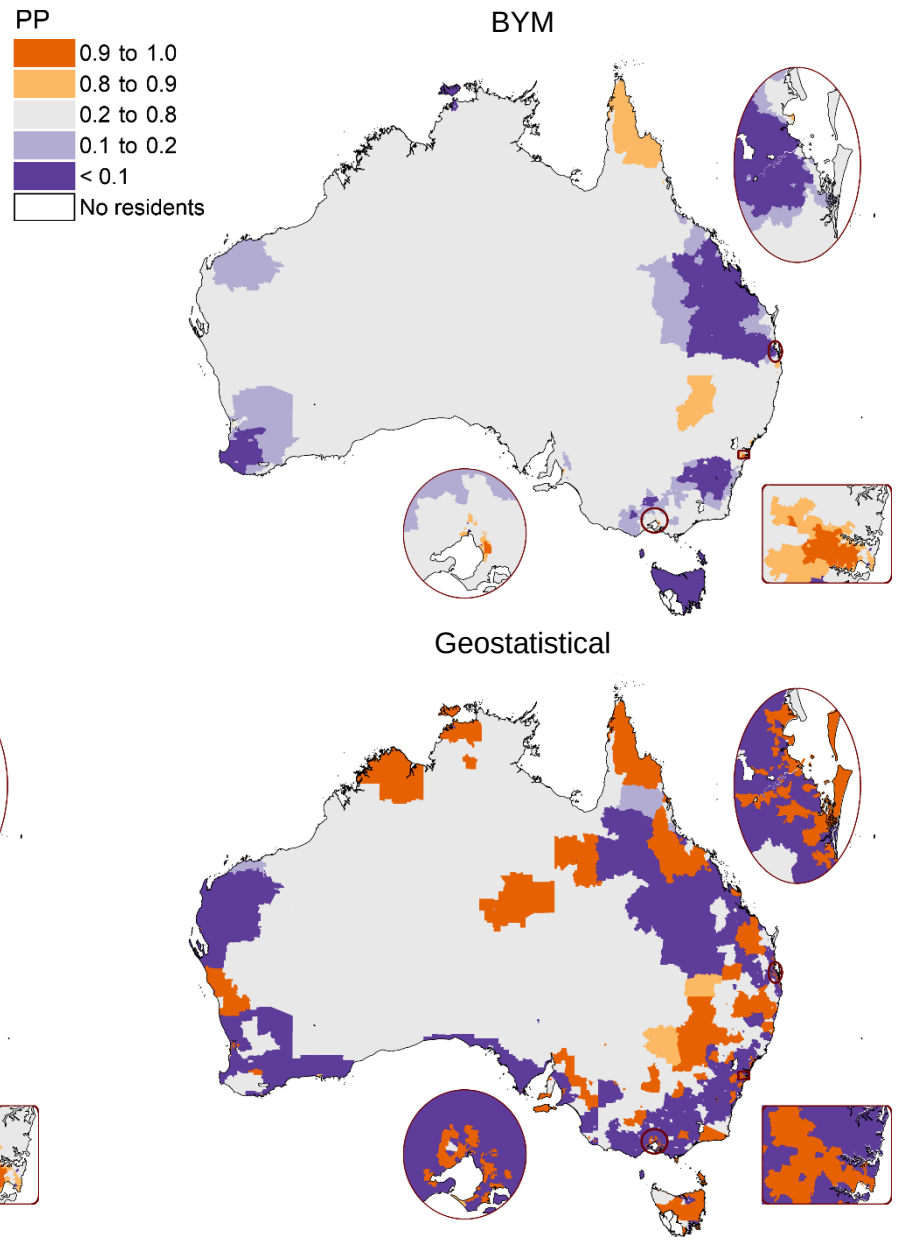


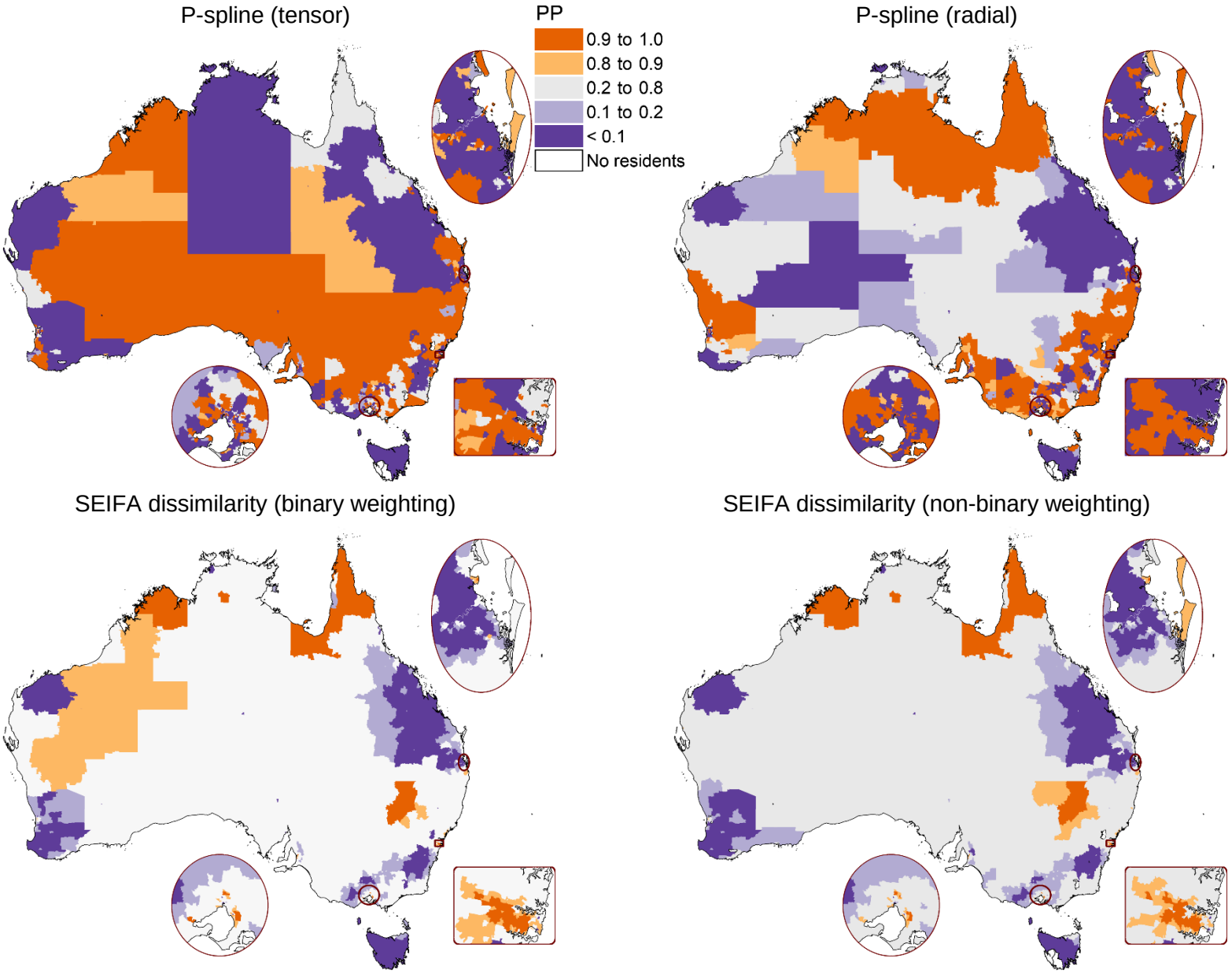
Leroux scale mixture model

Model would not run for liver cancer, males

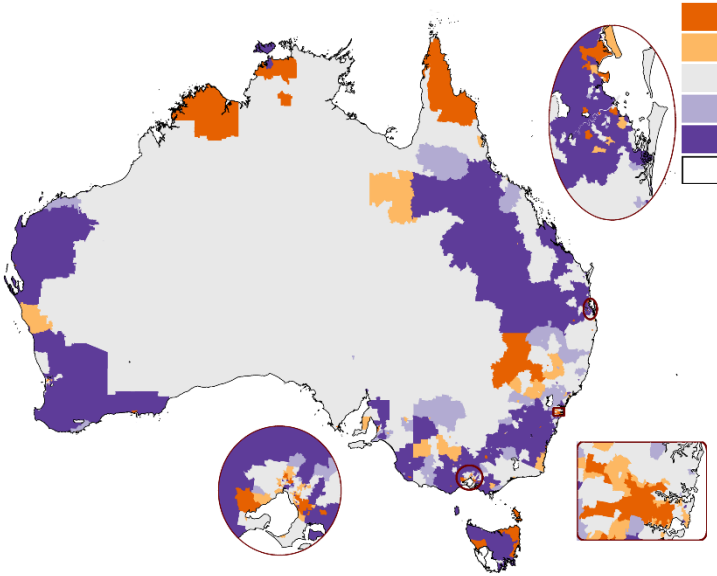
Liver cancer, males, posterior probability (PP) SIR >1

Maps are aligned to correspond with the layout of the modelled SIR maps.




P-spline (tensor)
PP
0.9 to 1.0
0.8 to 0.9
0.2 to 0.8
0.1 to 0.2
< 0.1
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)

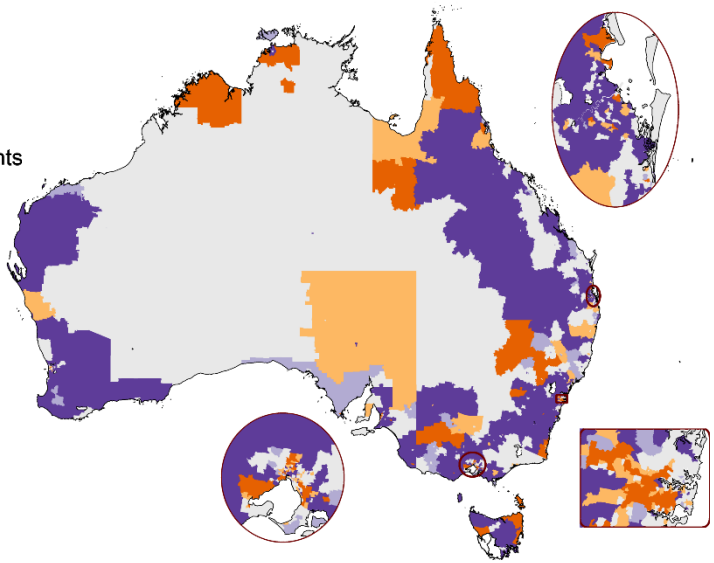
Residual dissimilarity (binary weighting)



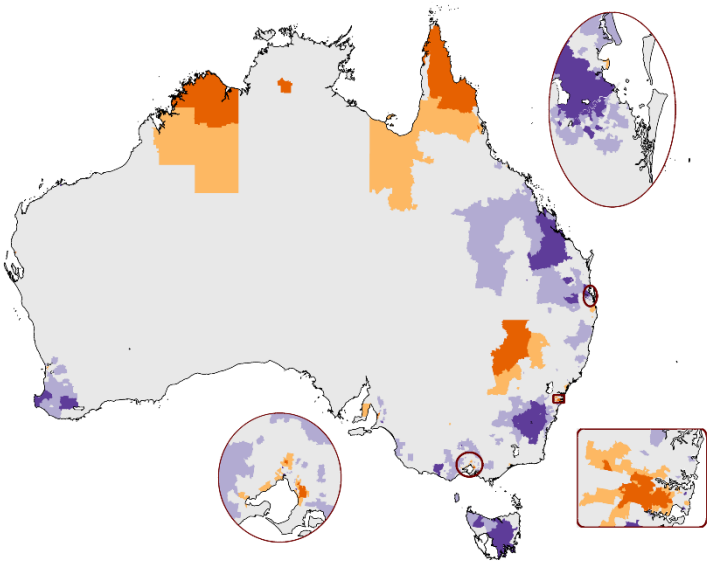
PP

- 0.9 to 1.0
- 0.8 to 0.9
- 0.2 to 0.8
- 0.1 to 0.2
- < 0.1
- No residents

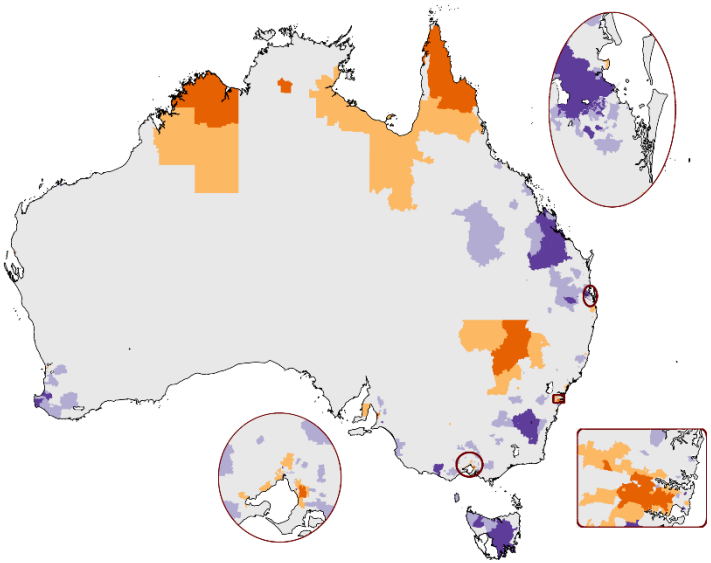
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)



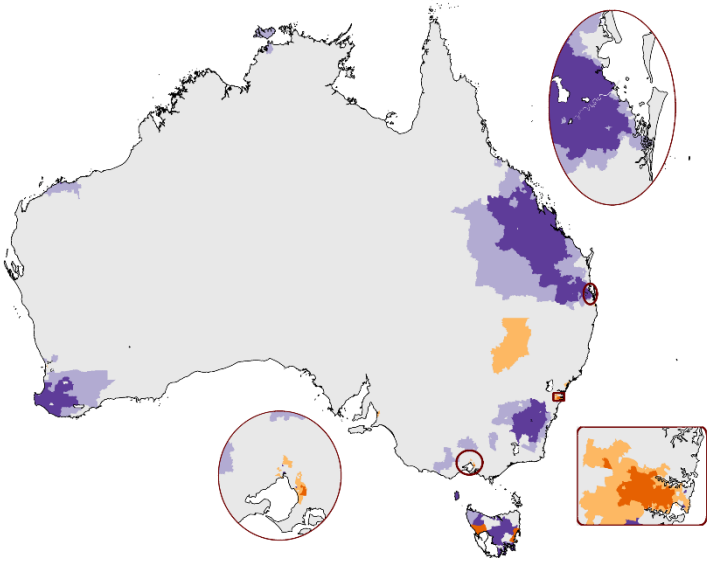
Locally adaptive (rho is determined when modelling)

[Unavailable]

Locally adaptive (rho is fixed at 0.99)

[Unavailable]

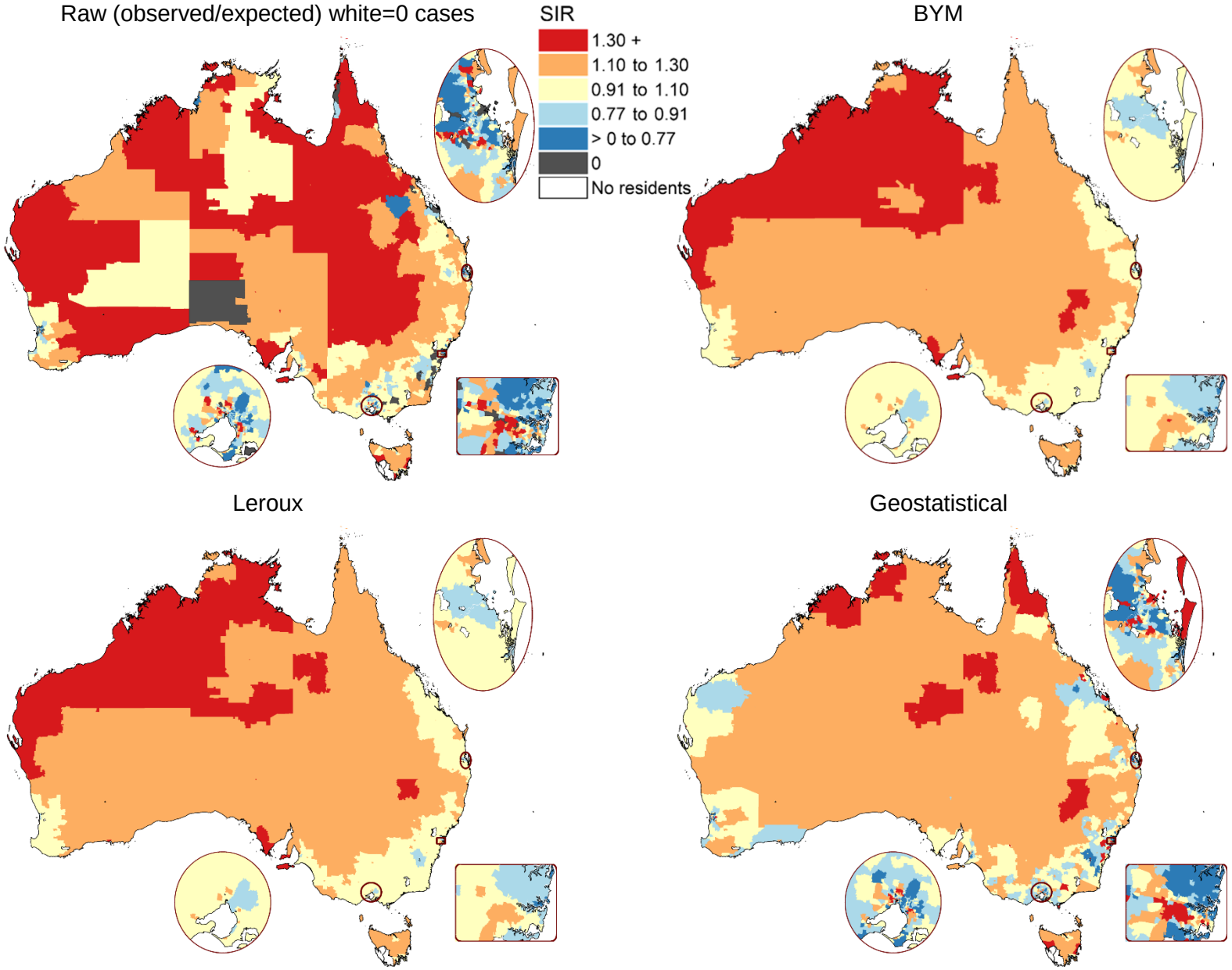
Weighted sum of spatial priors

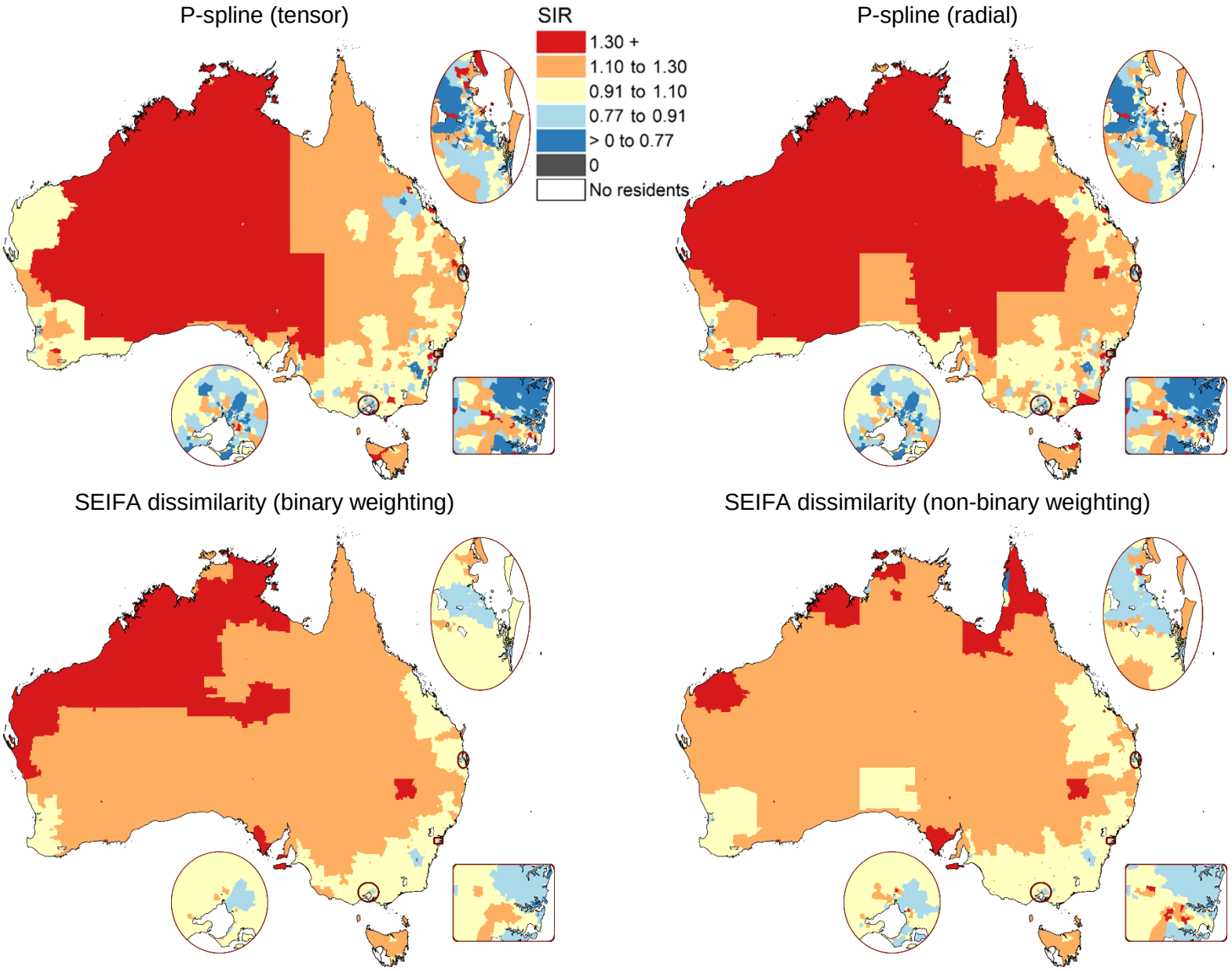


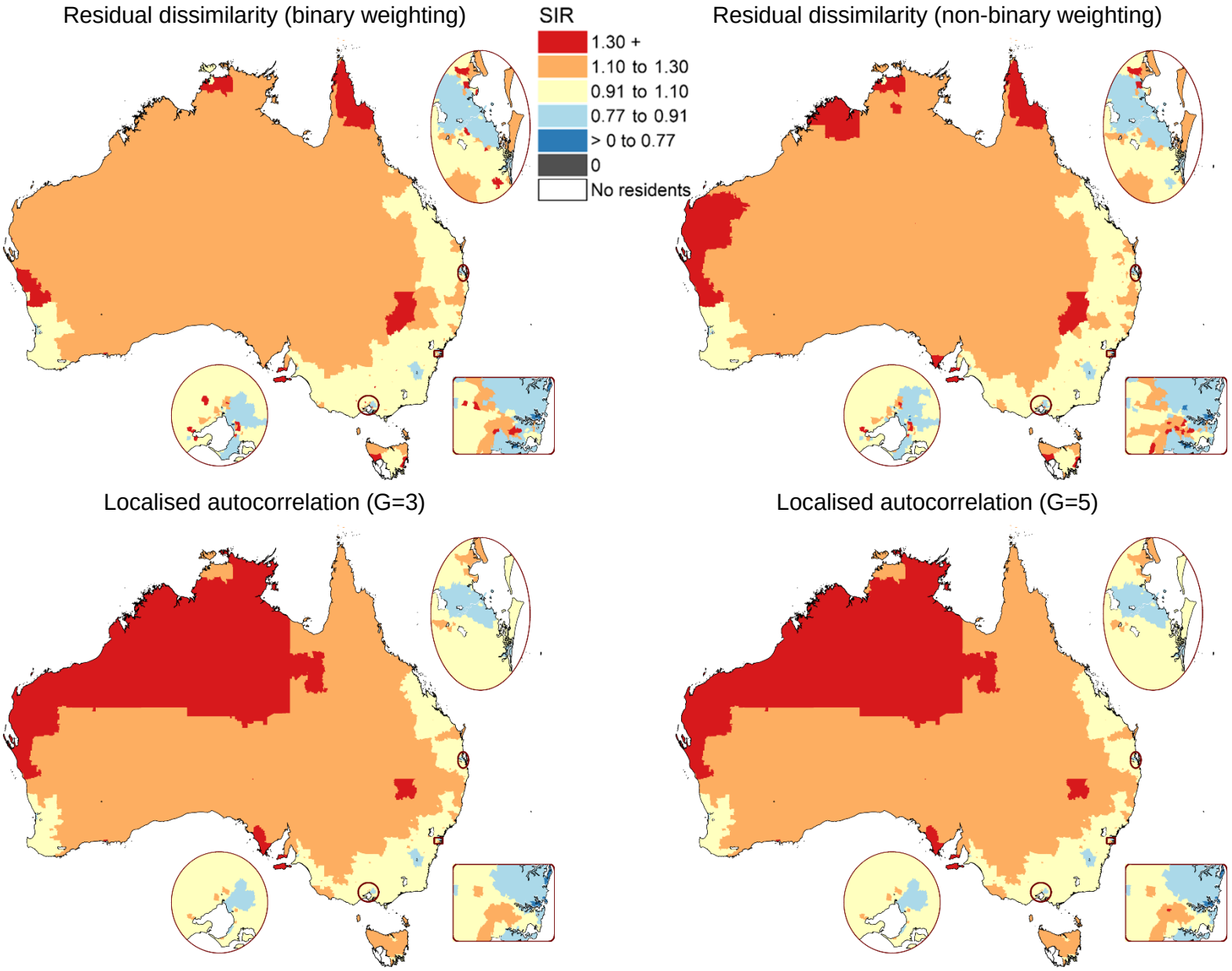
Leroux scale mixture model

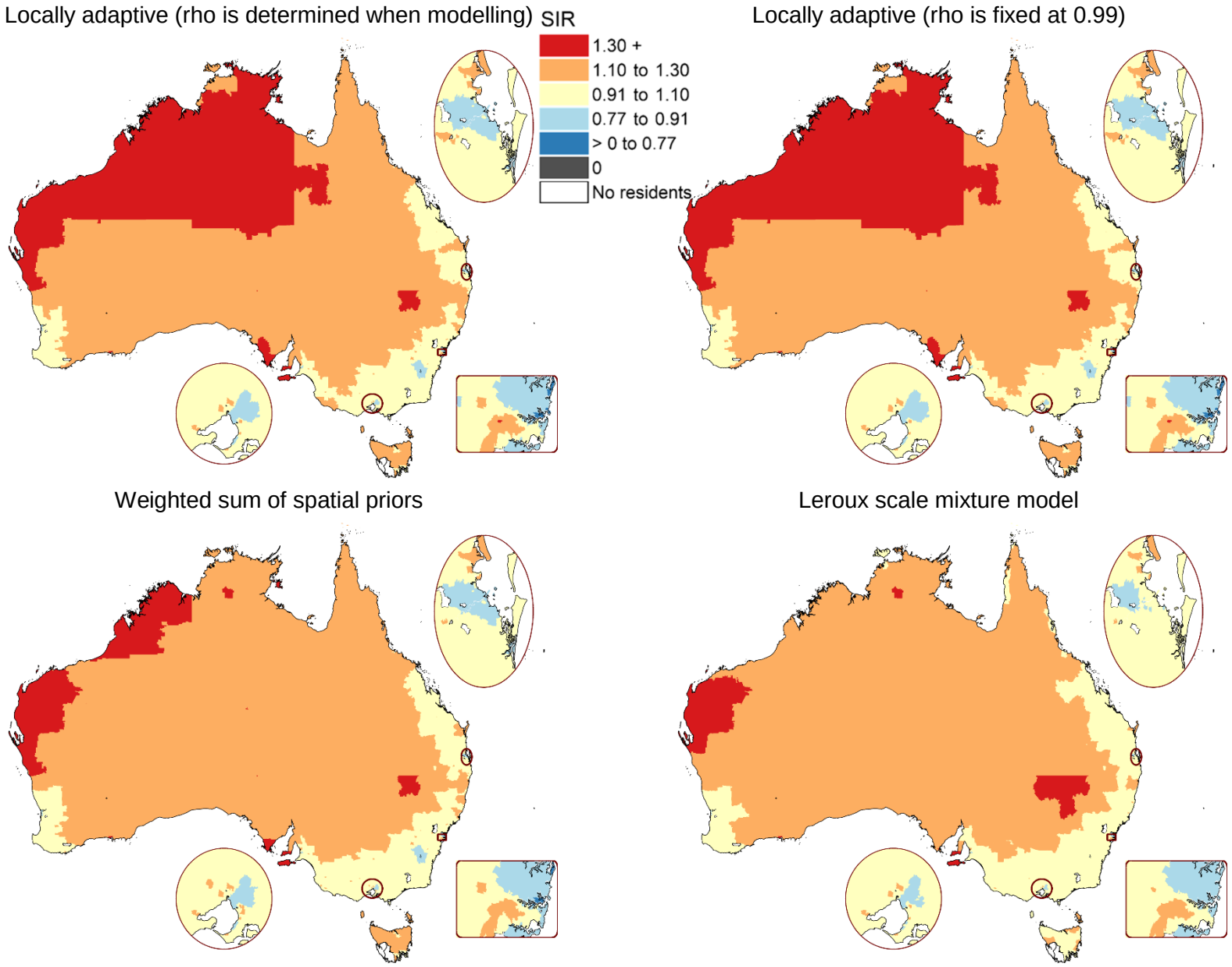
Model would not run for liver cancer, males

Lung cancer, males, modelled SIR



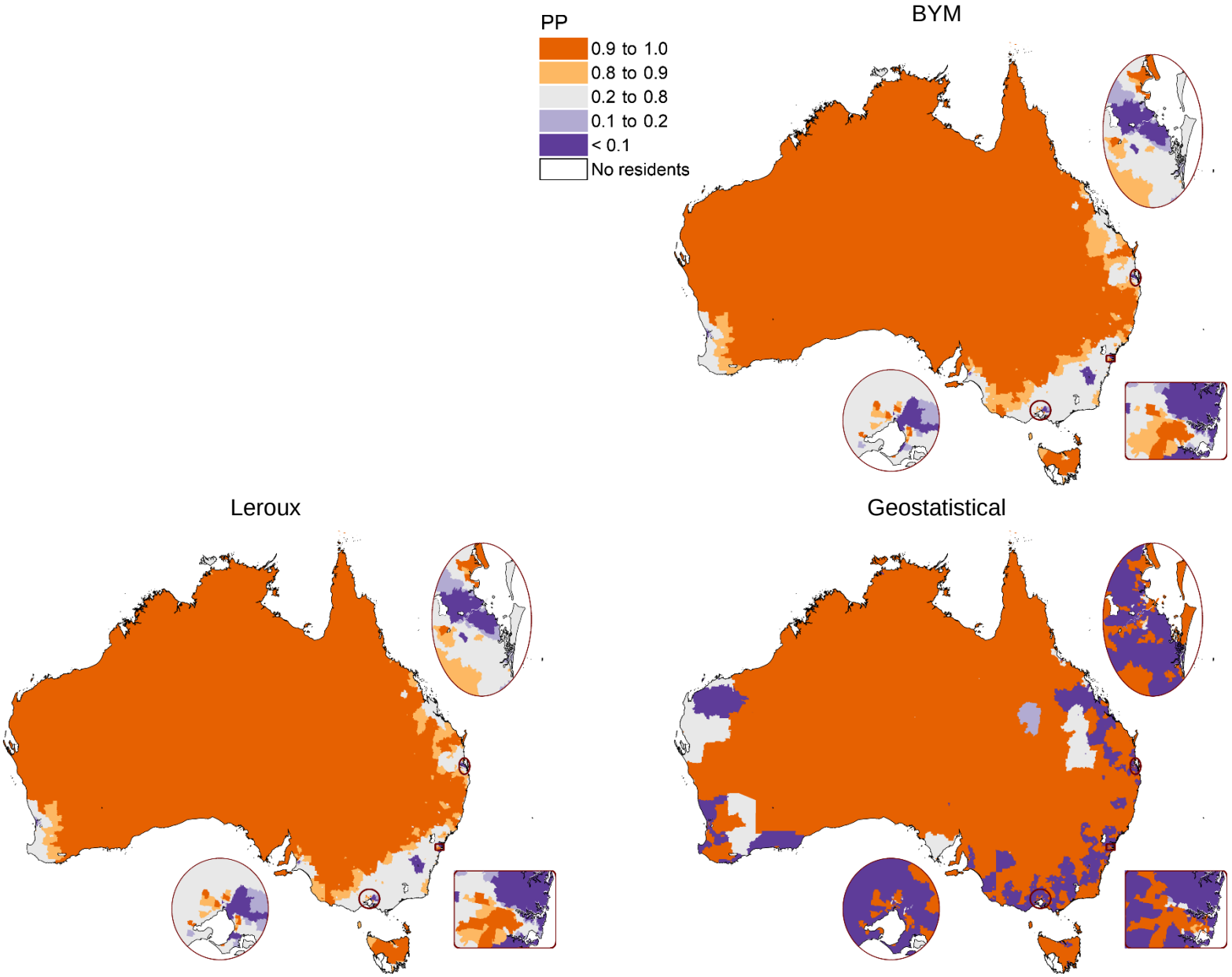

P-spline (tensor)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)

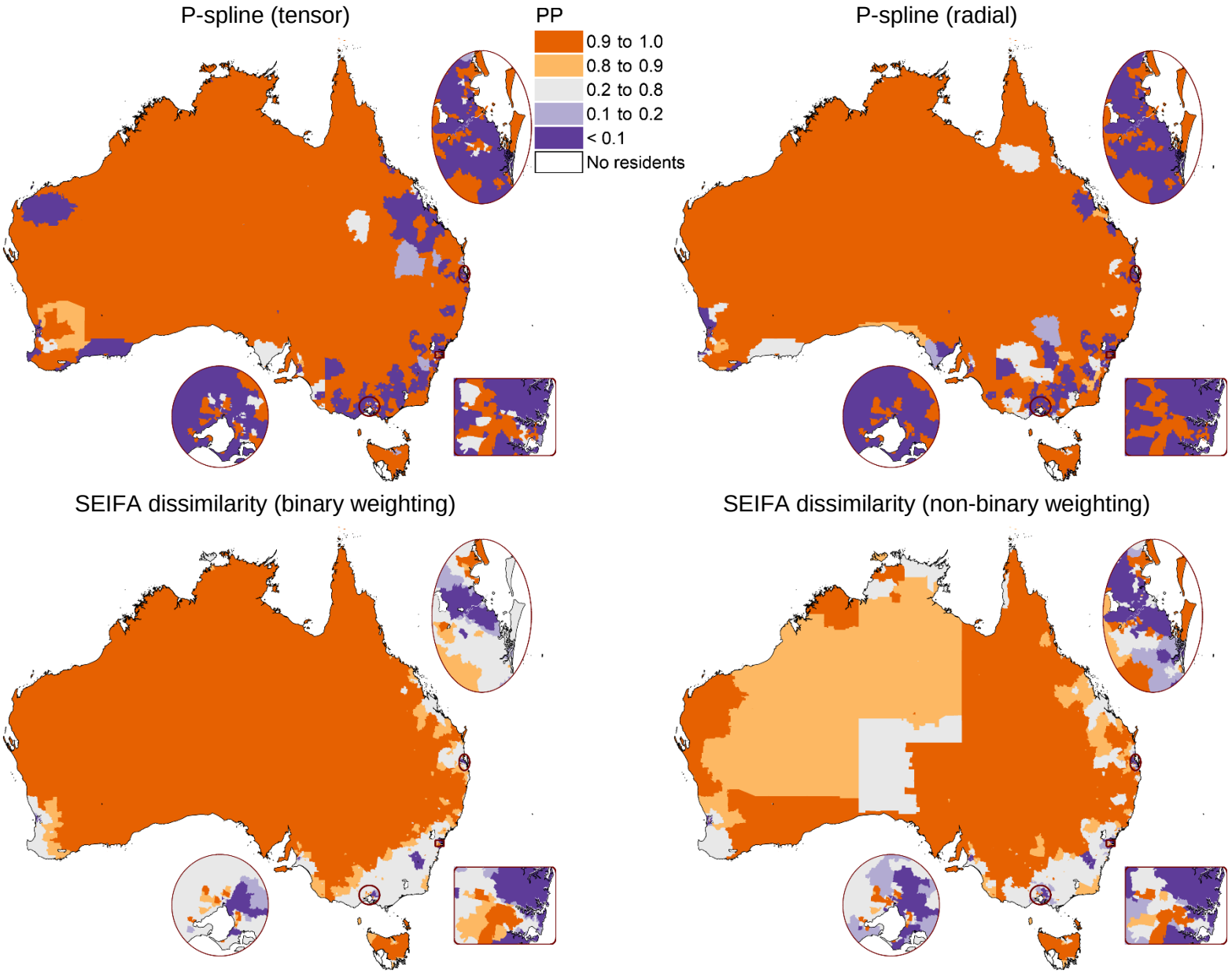

Residual dissimilarity (binary weighting)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
Residual dissimilarity (non-binary weighting)
Localised autocorrelation (G=3)
Localised autocorrelation (G=5)


Locally adaptive (rho is determined when modelling)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
Locally adaptive (rho is fixed at 0.99)
Weighted sum of spatial priors
Leroux scale mixture model

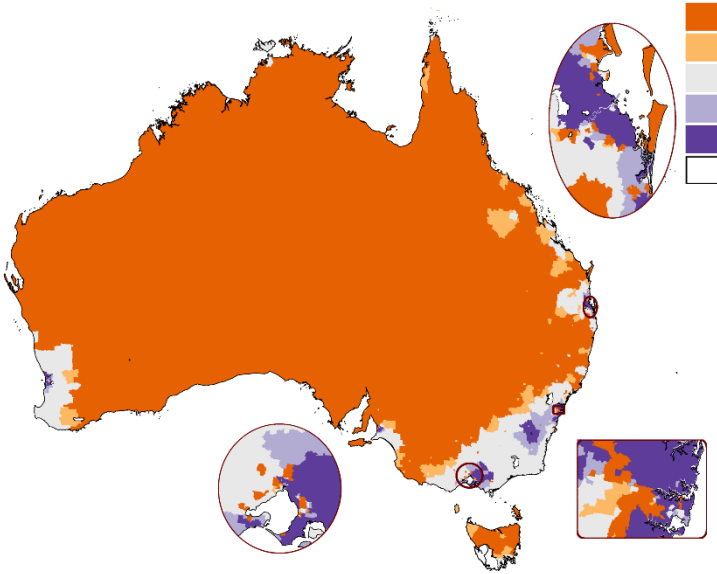
Lung cancer, males, PP SIR >1

Maps are aligned to correspond with the layout of the modelled SIR maps.




P-spline (tensor)
PP
0.9 to 1.0
0.8 to 0.9
0.2 to 0.8
0.1 to 0.2
< 0.1
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)

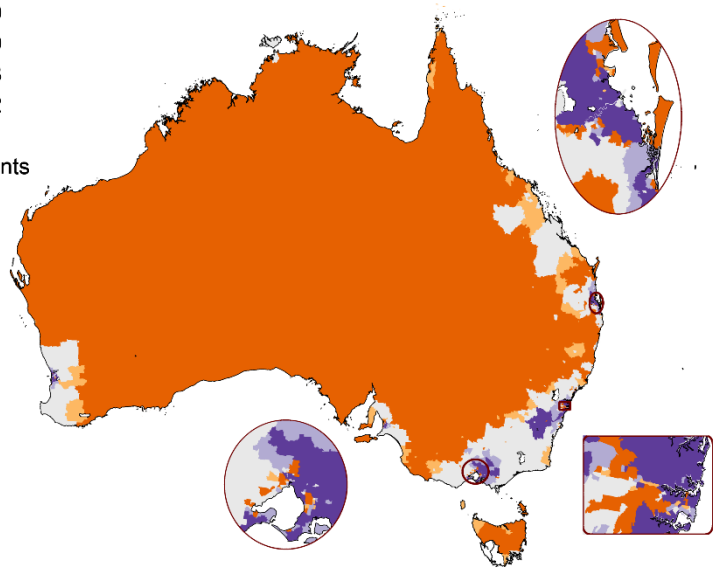
Residual dissimilarity (binary weighting)



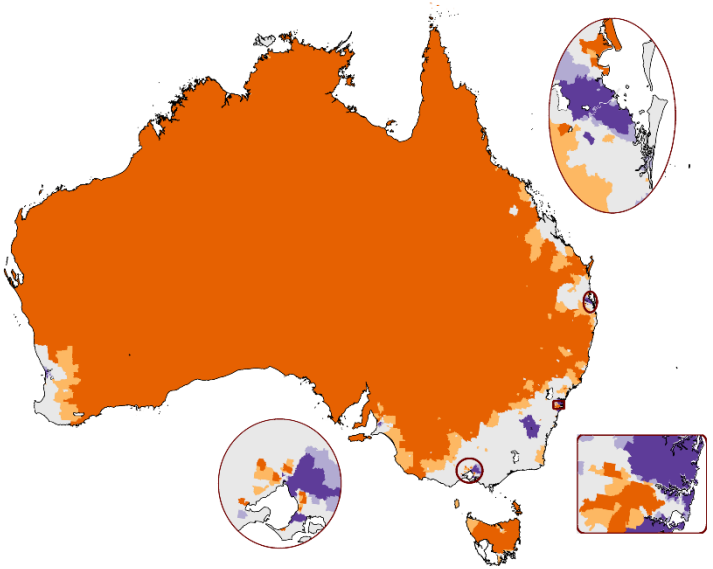
PP

- 0.9 to 1.0
- 0.8 to 0.9
- 0.2 to 0.8
- 0.1 to 0.2
- < 0.1
- No residents

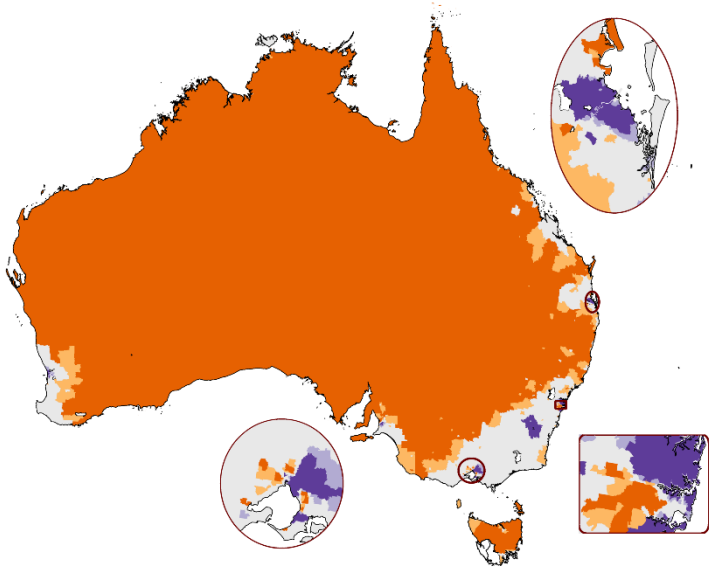
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)

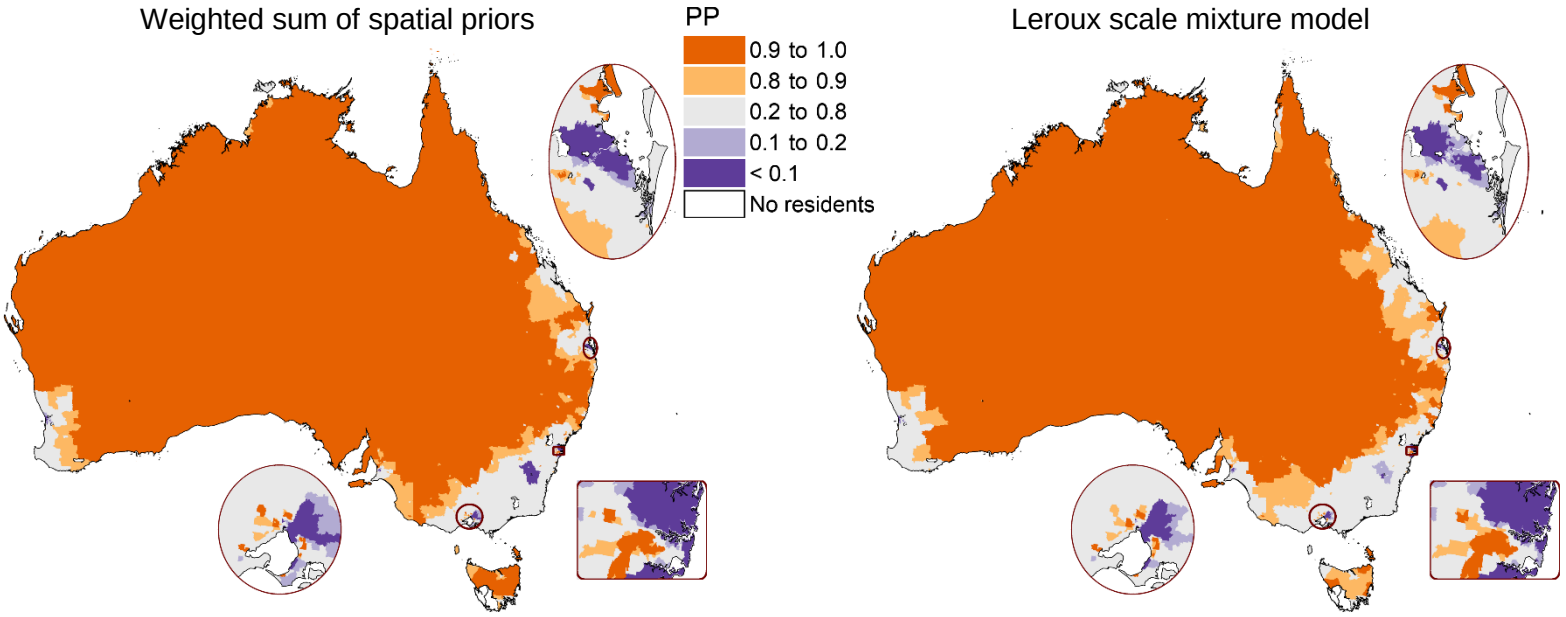


Locally adaptive (rho is determined when modelling)

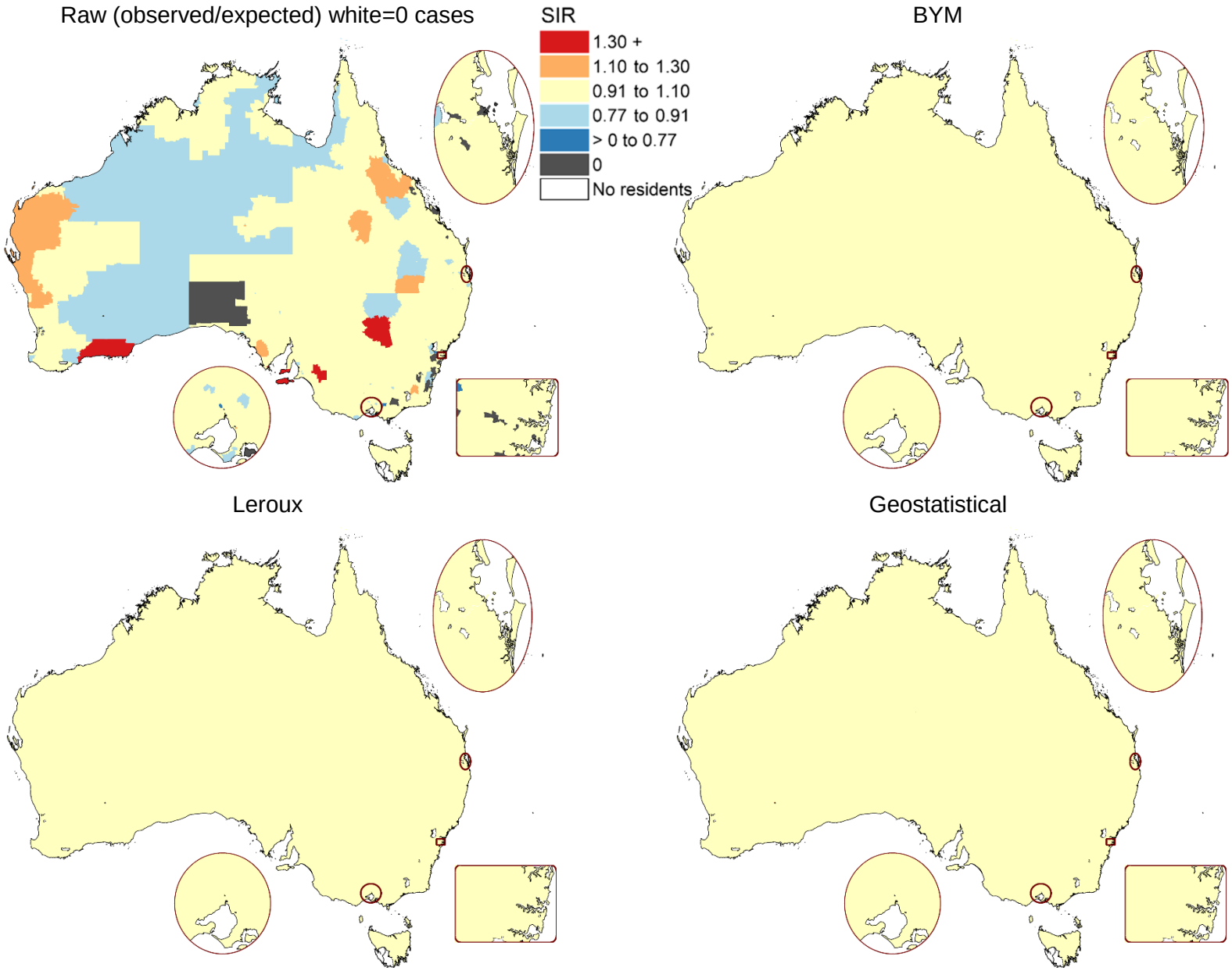
[Unavailable]

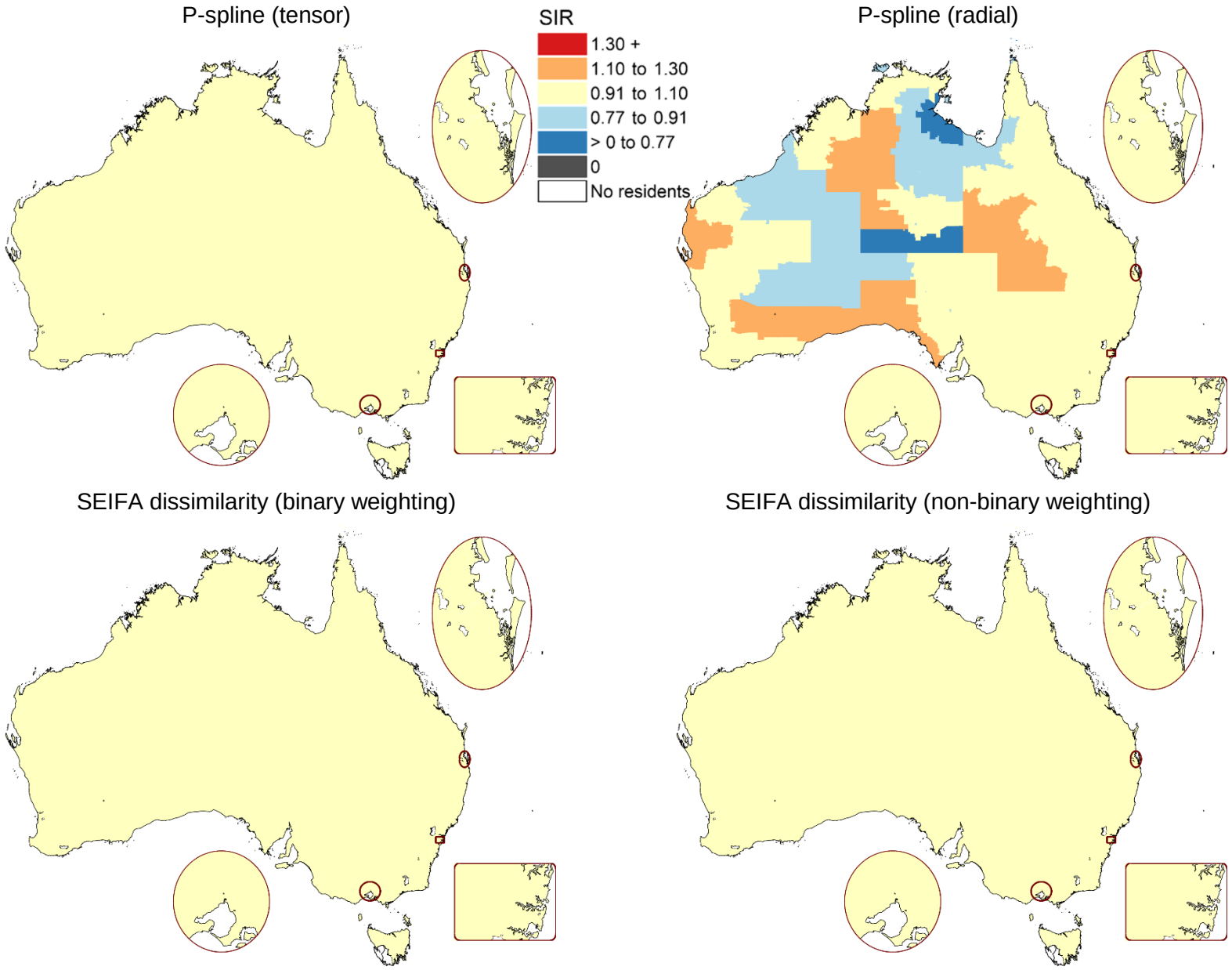
Locally adaptive (rho is fixed at 0.99)

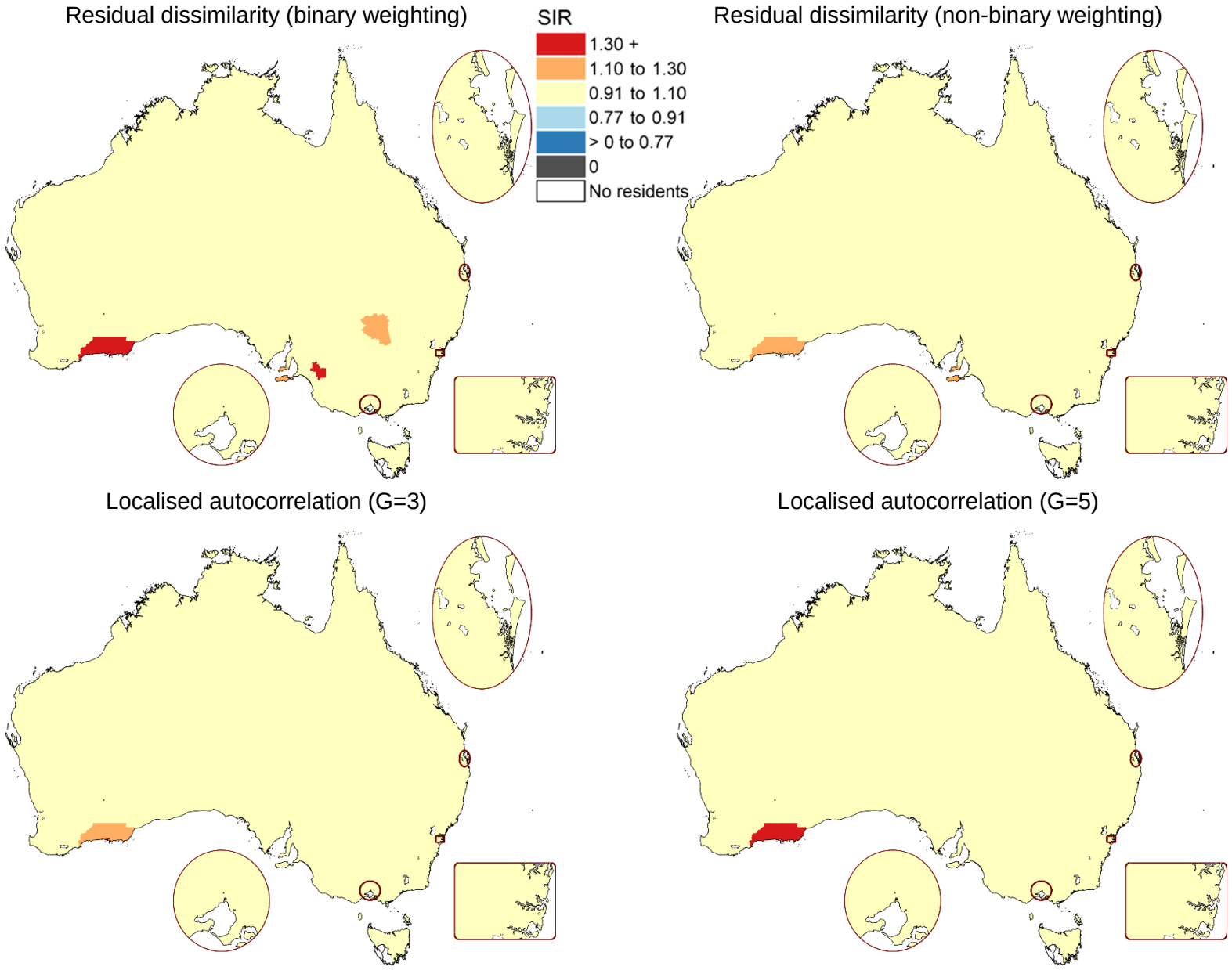
[Unavailable]

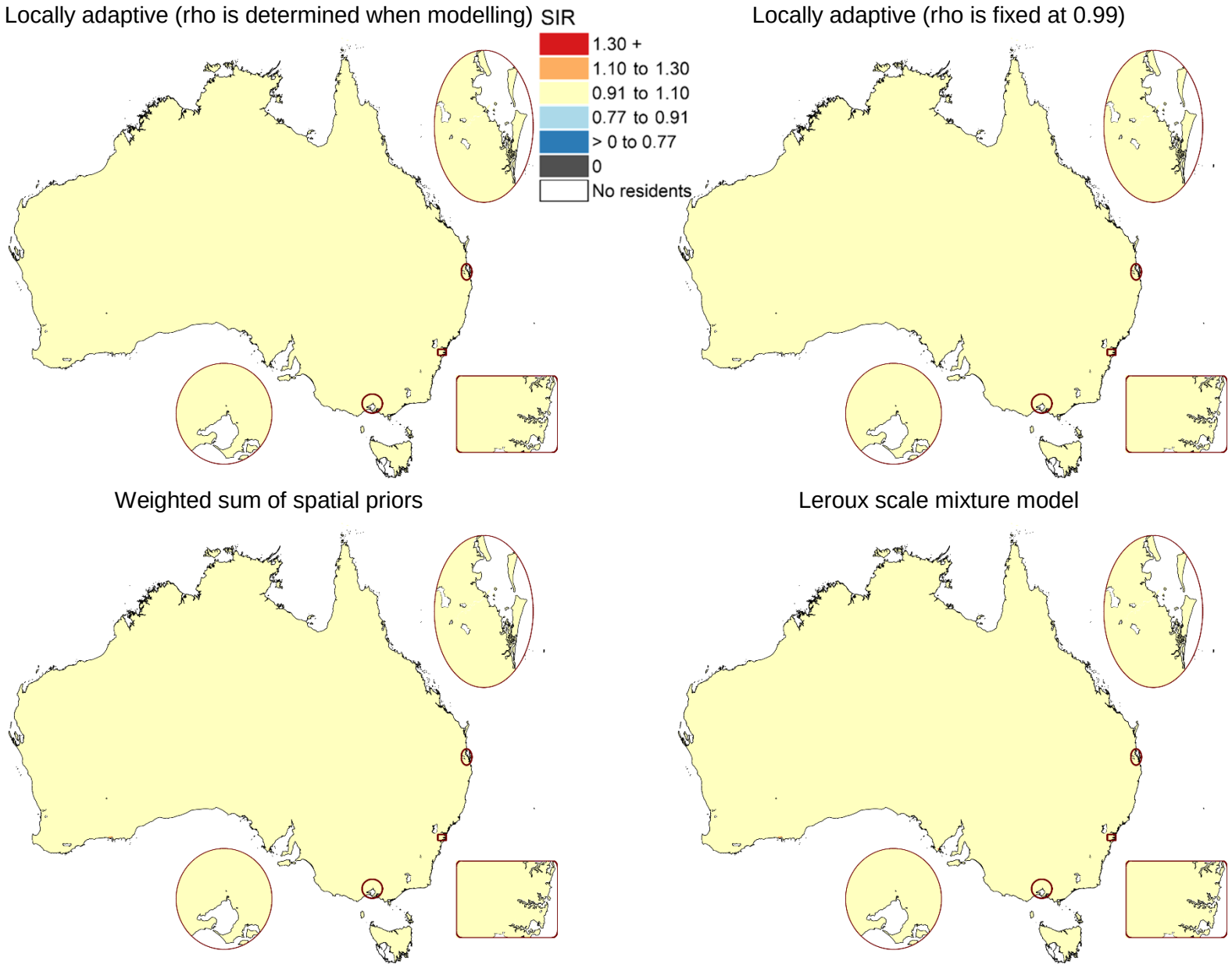

Weighted sum of spatial priors
PP
0.9 to 1.0
0.8 to 0.9
0.2 to 0.8
0.1 to 0.2
< 0.1
No residents
Leroux scale mixture model

All invasive cancers, females, modelled SIR



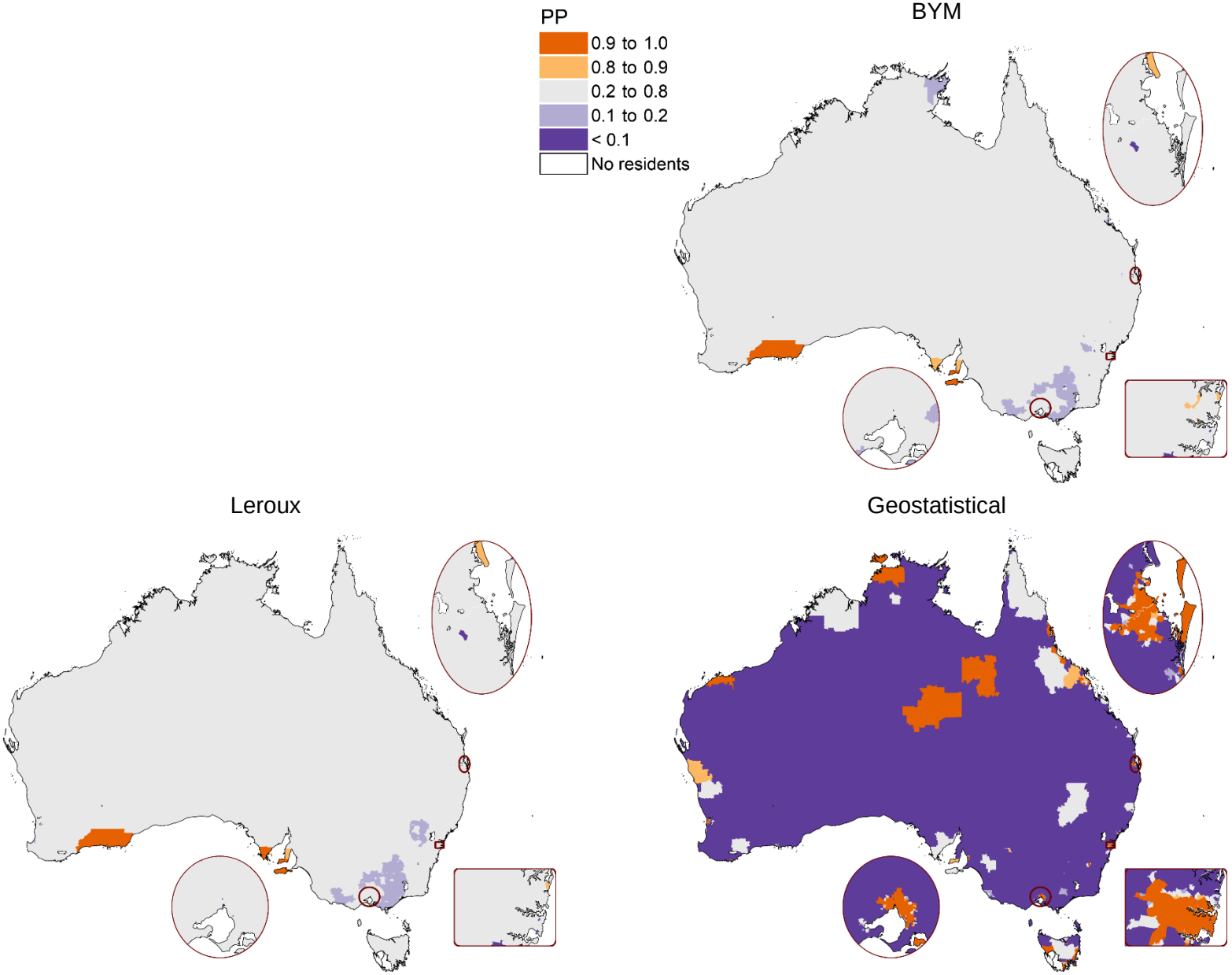

P-spline (tensor)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)

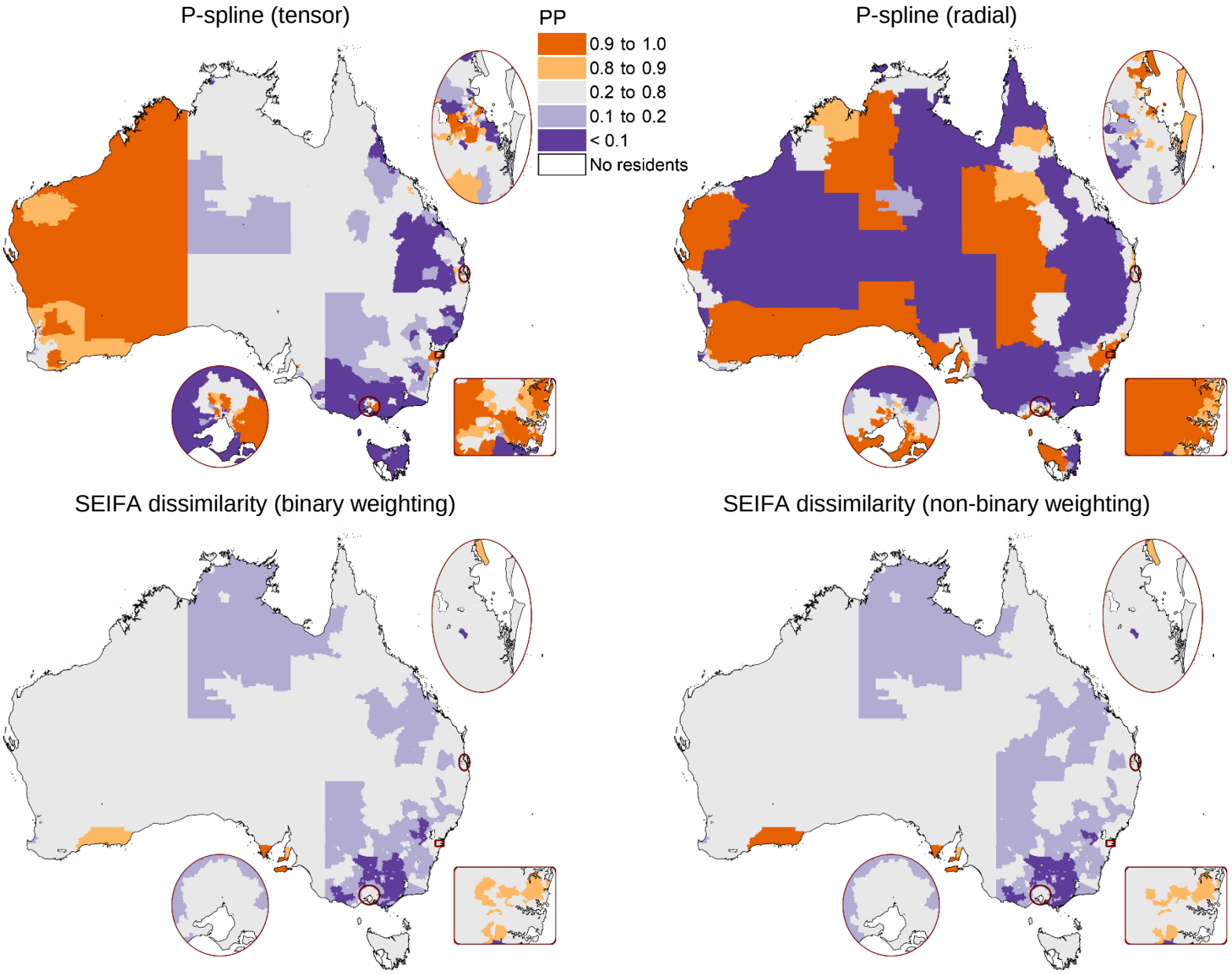

Residual dissimilarity (binary weighting)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
Residual dissimilarity (non-binary weighting)
Localised autocorrelation (G=3)
Localised autocorrelation (G=5)


Locally adaptive (rho is determined when modelling)
SIR
1.30 +
1.10 to 1.30
0.91 to 1.10
0.77 to 0.91
> 0 to 0.77
0
No residents
Locally adaptive (rho is fixed at 0.99)
Weighted sum of spatial priors
Leroux scale mixture model

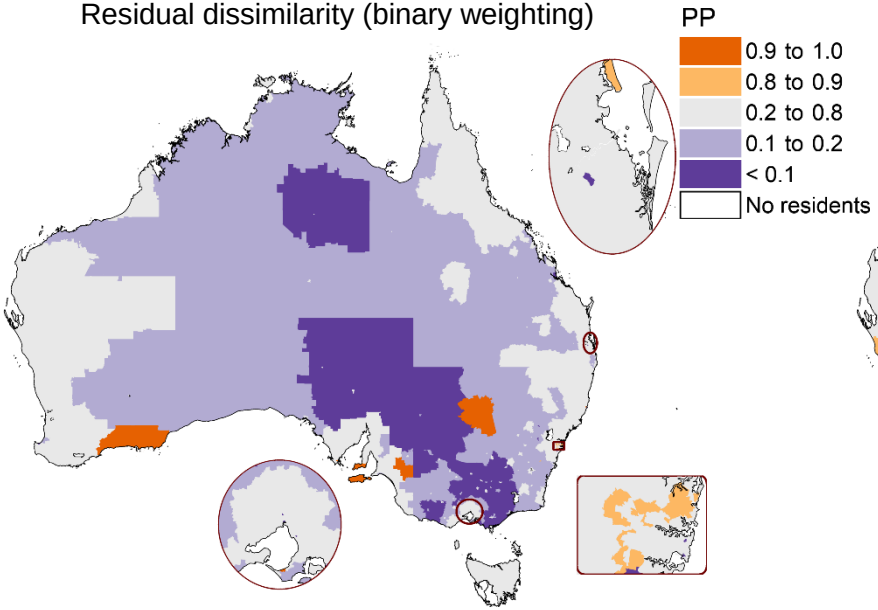
All invasive cancers, females, PP SIR >1

Maps are aligned to correspond with the layout of the modelled SIR maps.

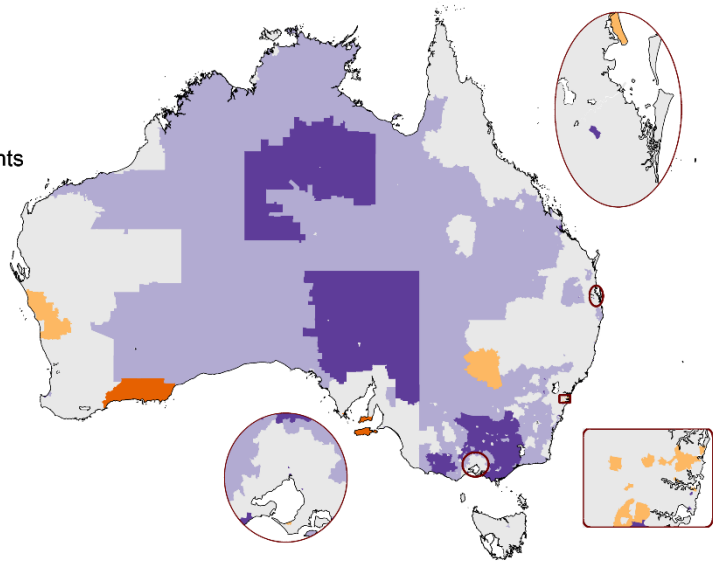



P-spline (tensor)
PP
0.9 to 1.0
0.8 to 0.9
0.2 to 0.8
0.1 to 0.2
< 0.1
No residents
P-spline (radial)
SEIFA dissimilarity (binary weighting)
SEIFA dissimilarity (non-binary weighting)

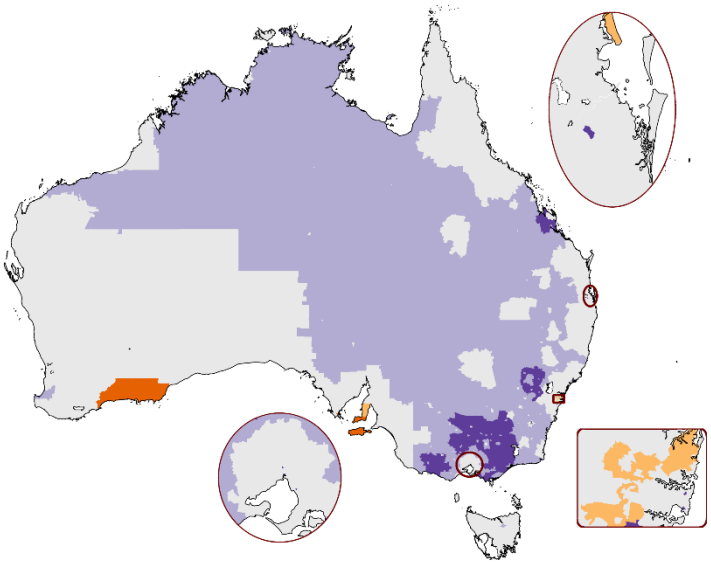
Residual dissimilarity (binary weighting)



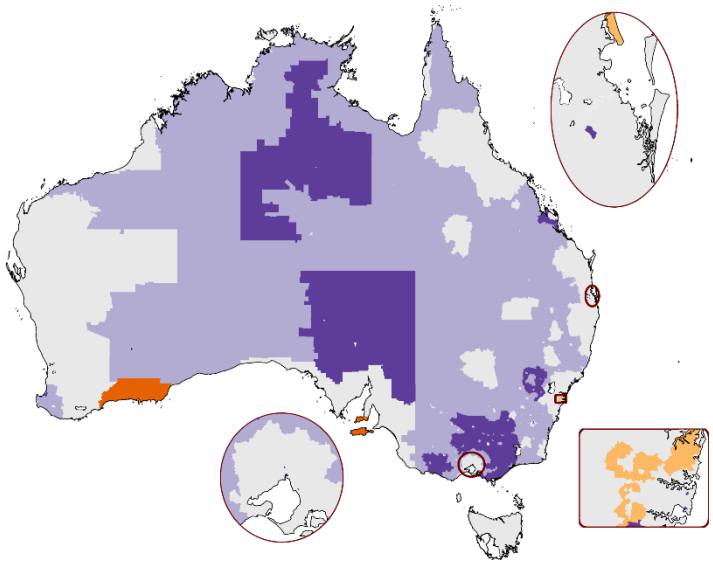
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)

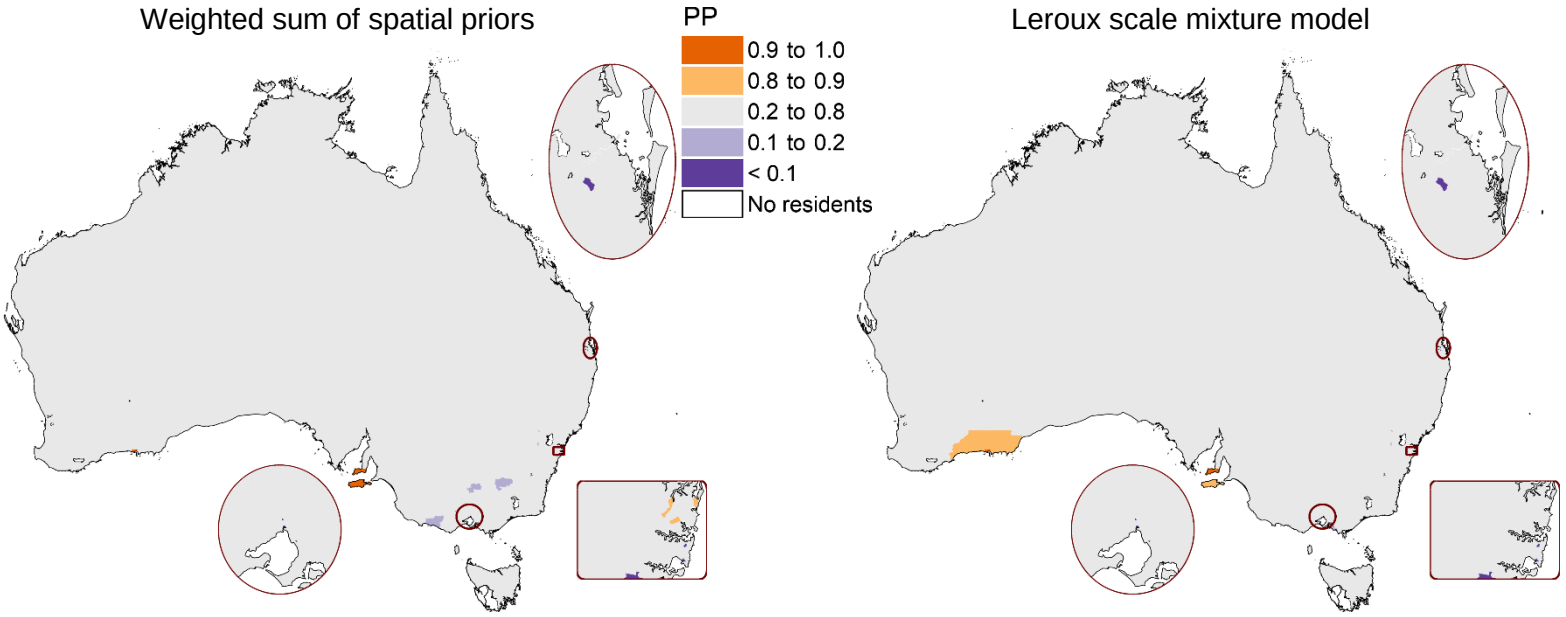


Locally adaptive (rho is determined when modelling)

[Unavailable]

Locally adaptive (rho is fixed at 0.99)

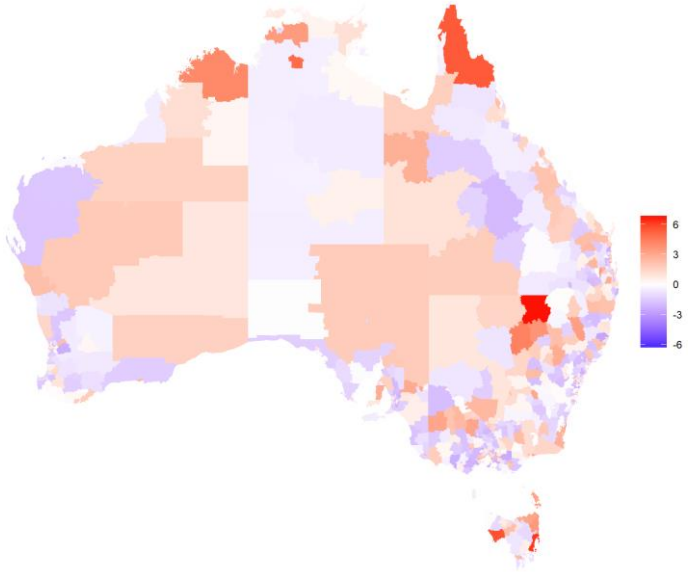
[Unavailable]

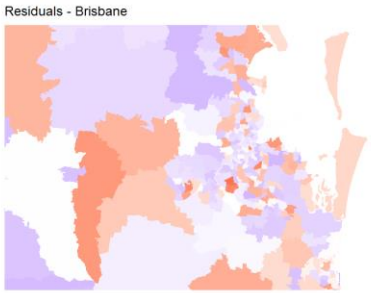

Weighted sum of spatial priors
PP
0.9 to 1.0
0.8 to 0.9
0.2 to 0.8
0.1 to 0.2
< 0.1
No residents
Leroux scale mixture model

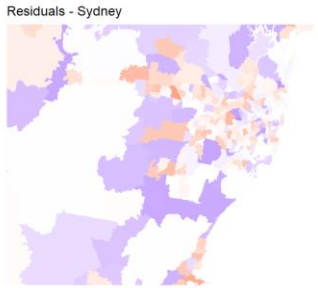
Appendix H: Maps of model fit

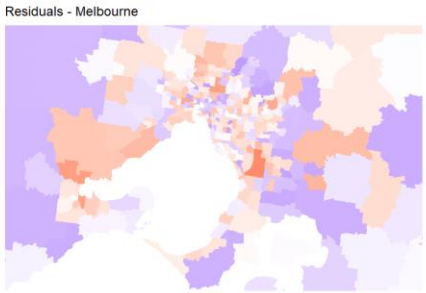
Liver cancer, males, residuals

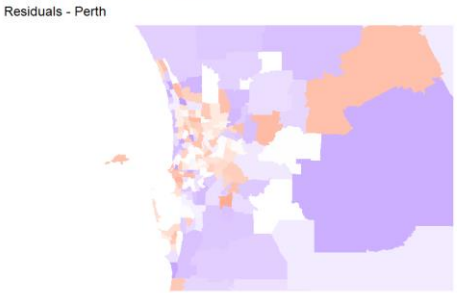
BYM



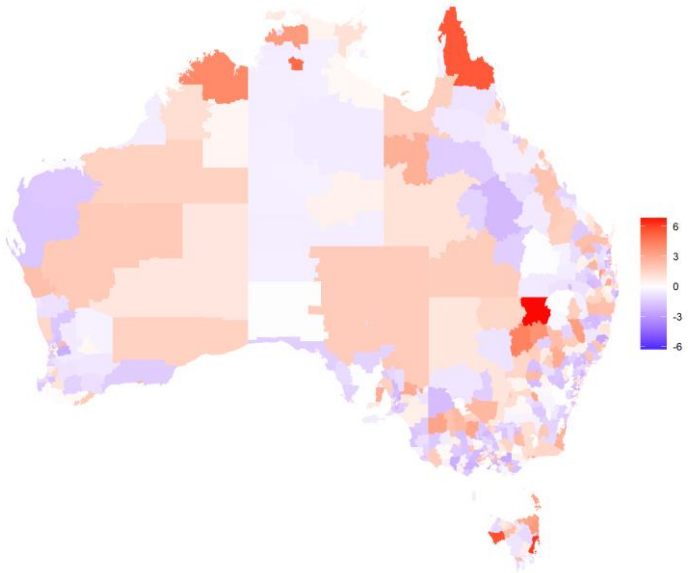


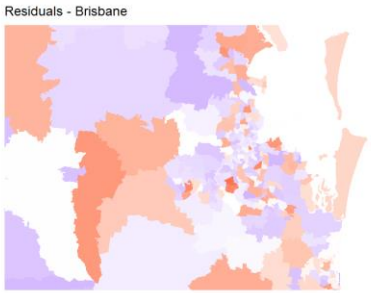


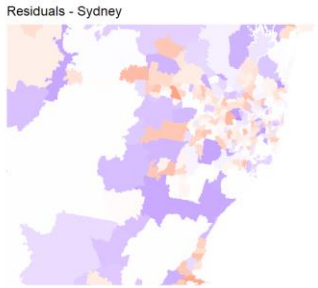


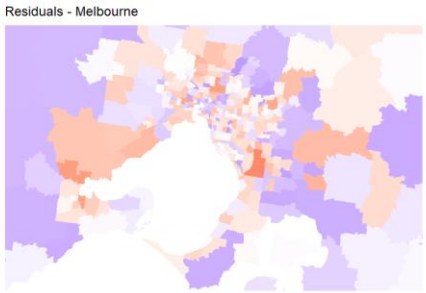


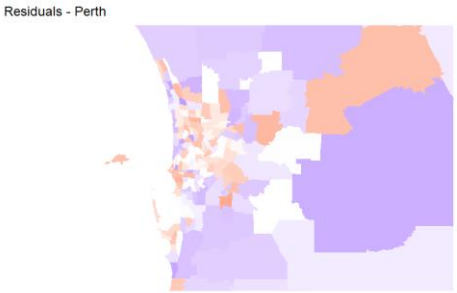
Leroux



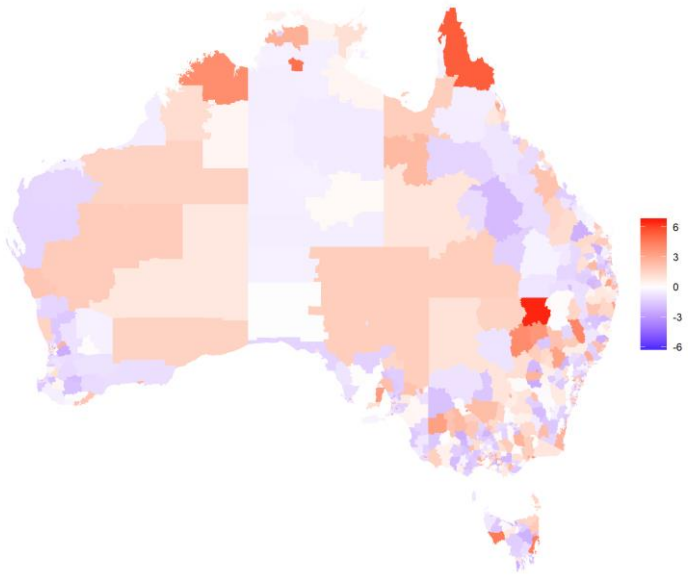


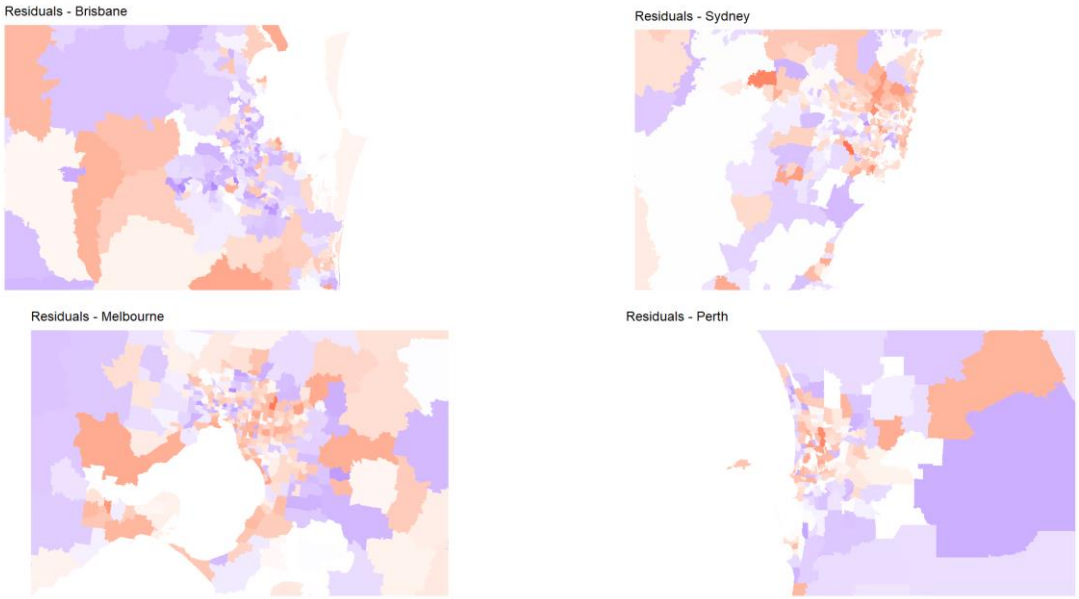




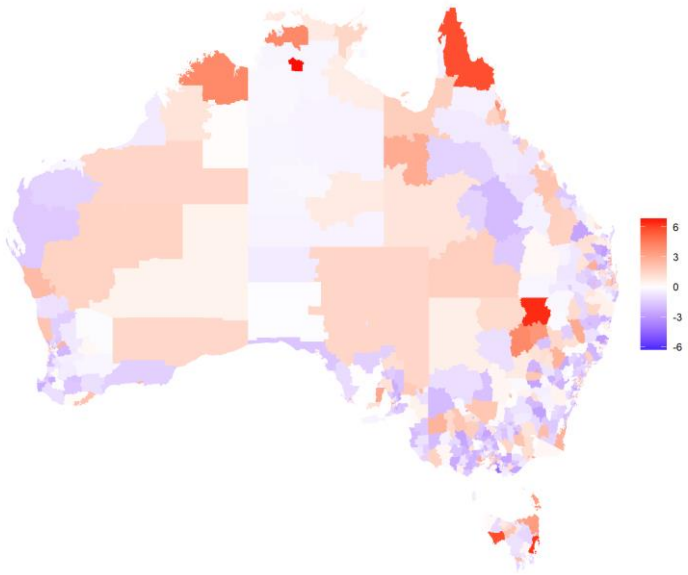


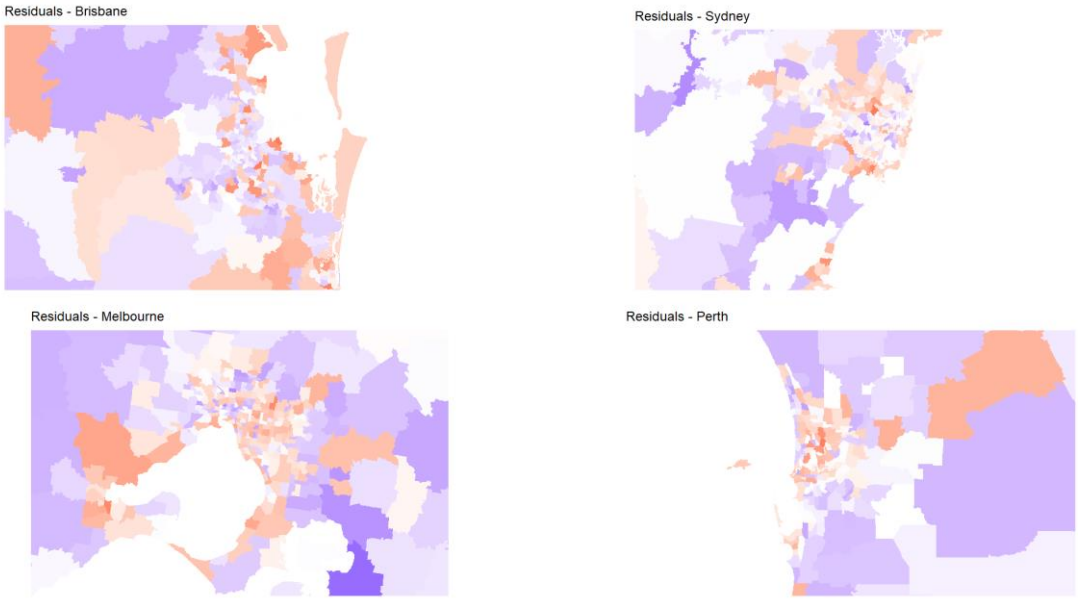
Geostatistical



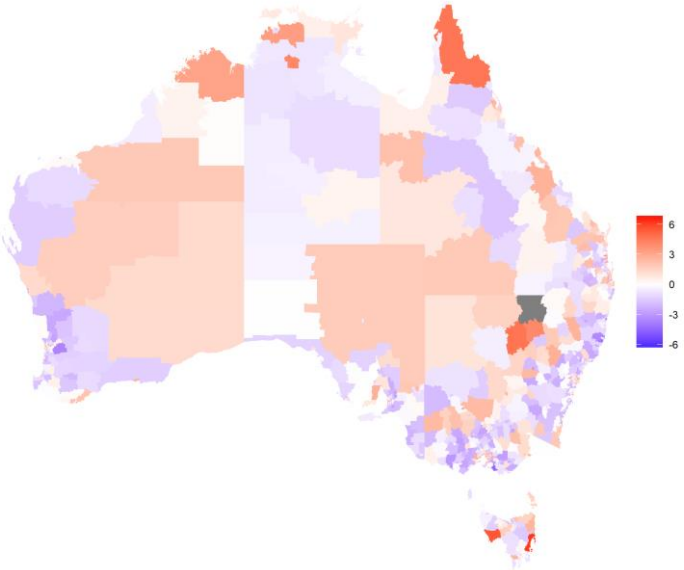


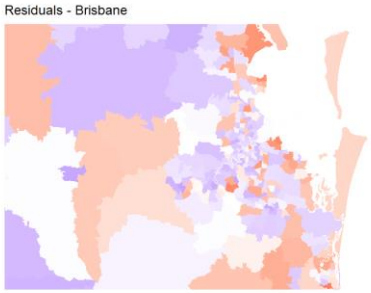
P-spline (tensor)

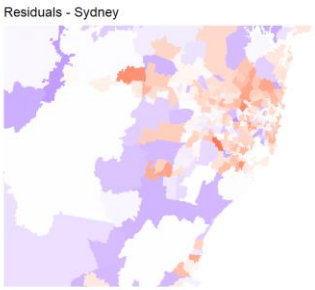


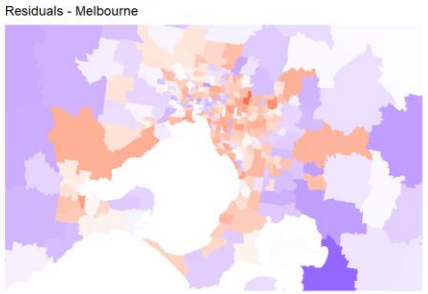


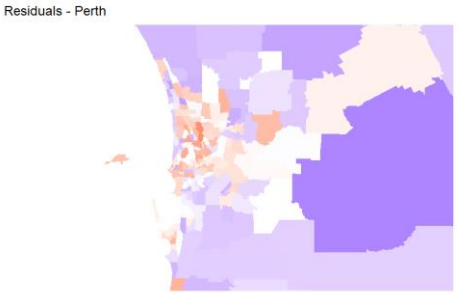
P-spline (radial)



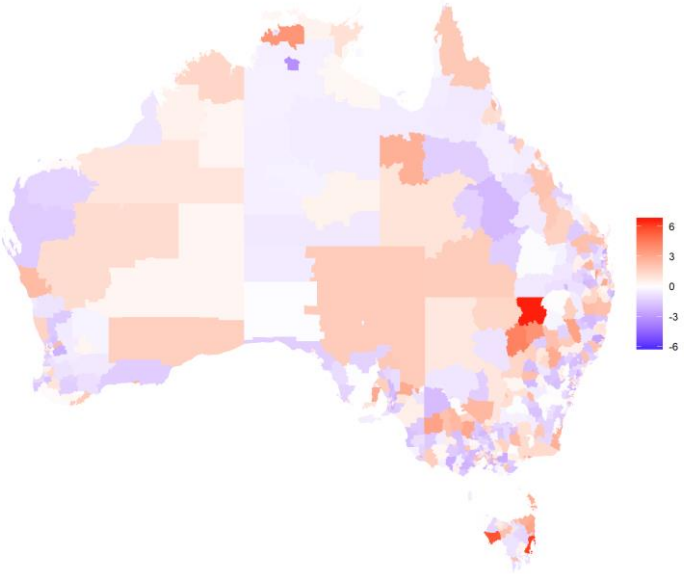


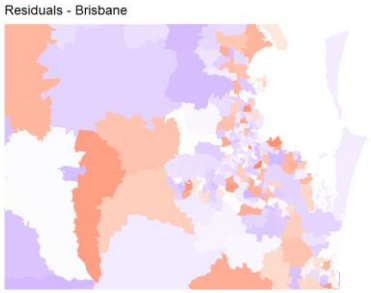


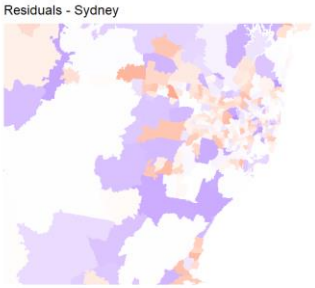


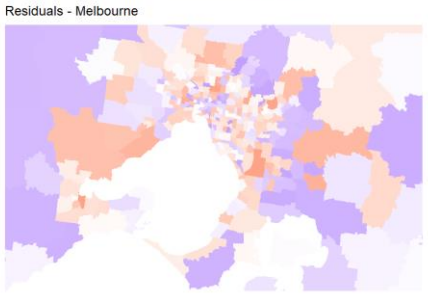


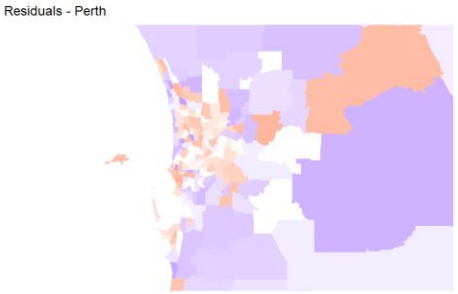
SEIFA dissimilarity (binary weighting)



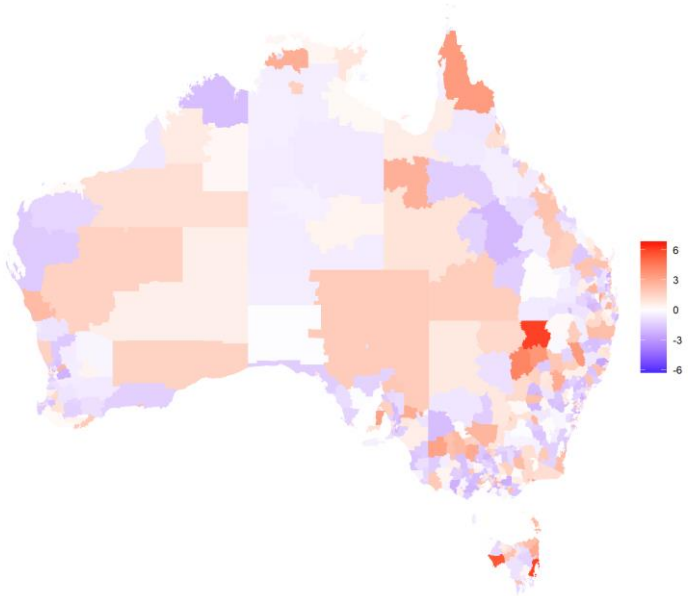


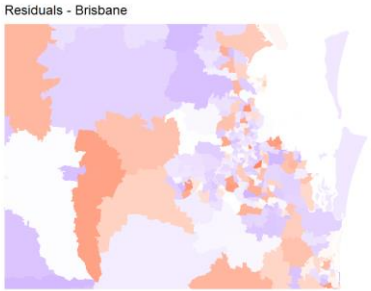


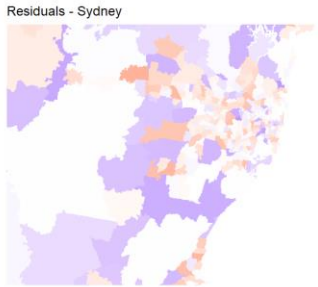


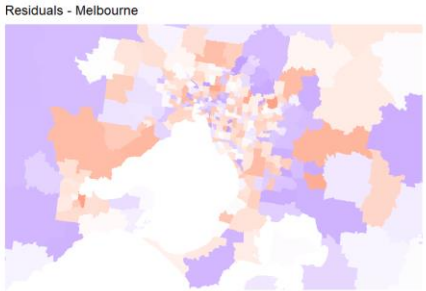


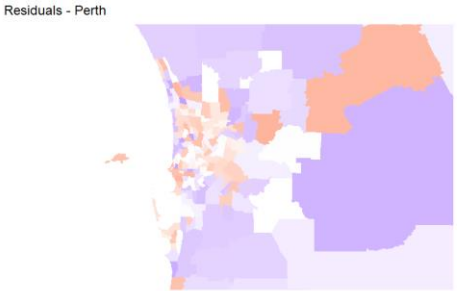
SEIFA dissimilarity (non-binary weighting)



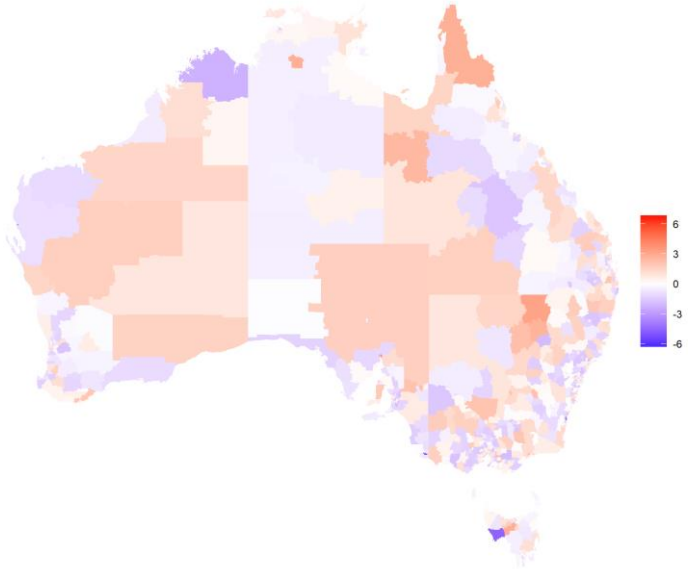


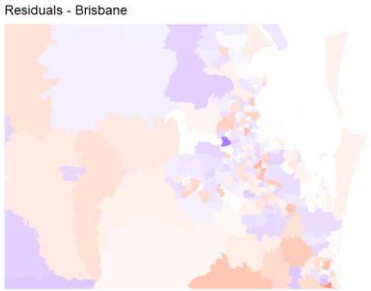


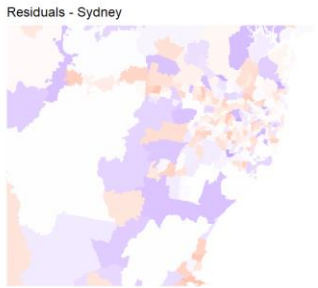


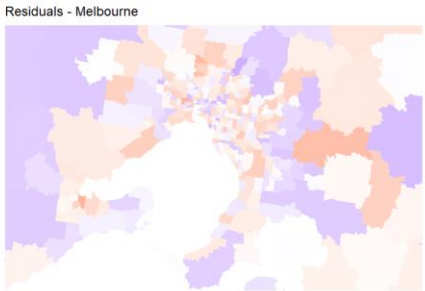


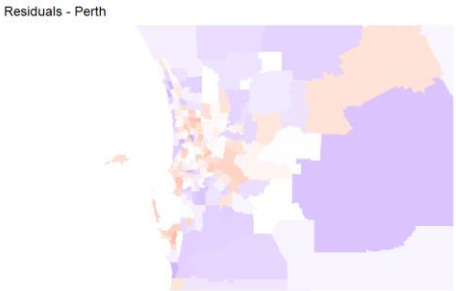
Residual dissimilarity (binary weighting)



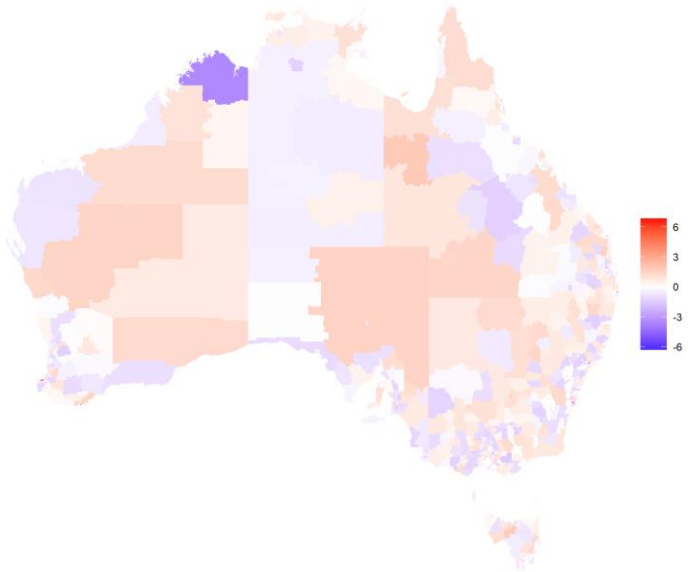


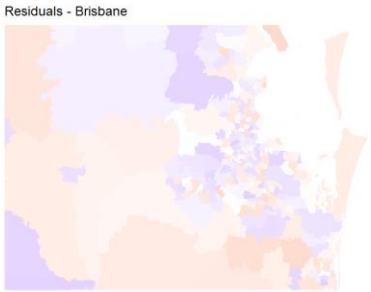


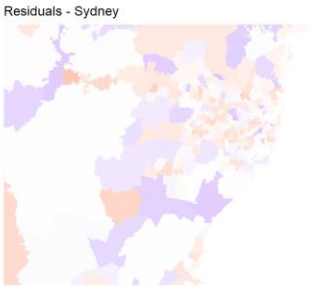


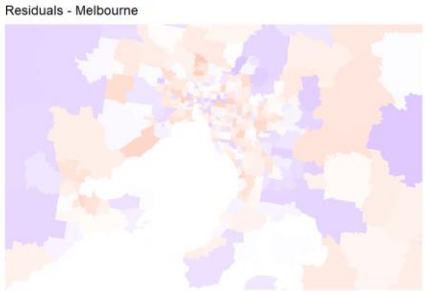


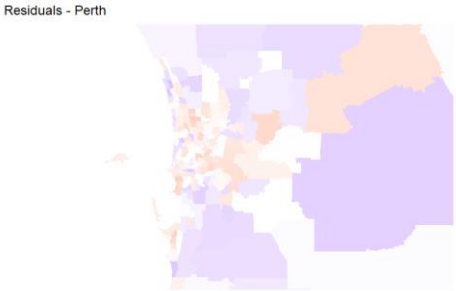
Residual dissimilarity (non-binary weighting)



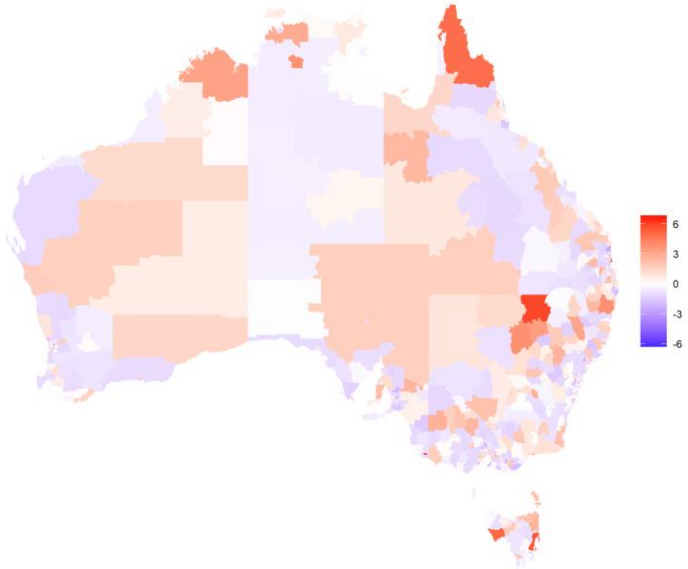


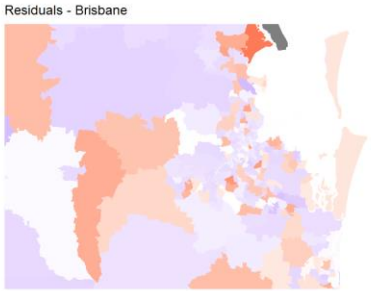


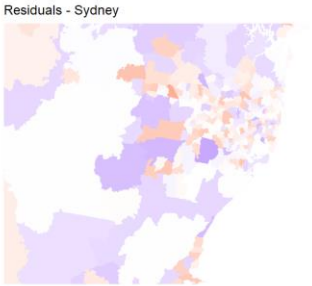


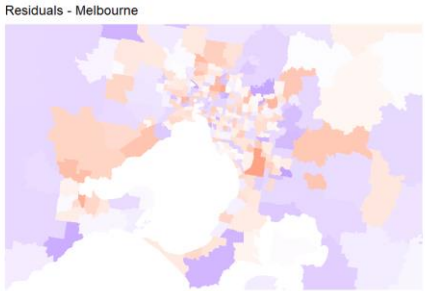


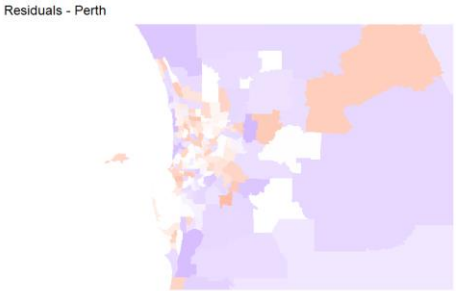
Localised autocorrelation (G=3)



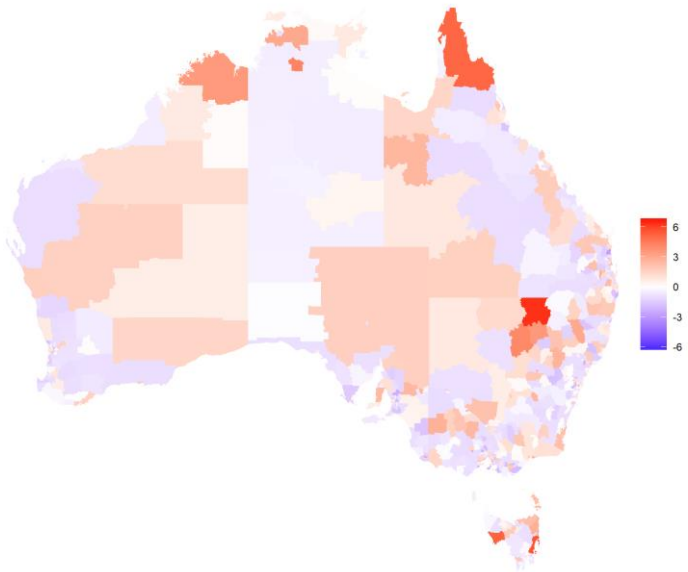


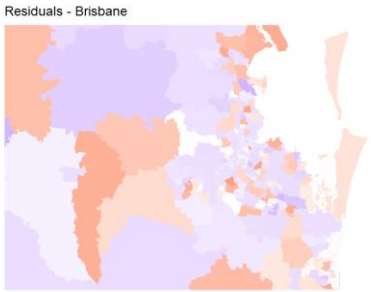


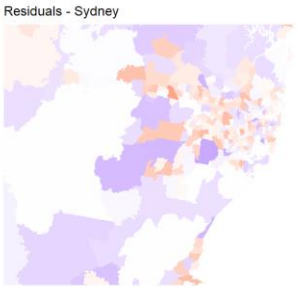


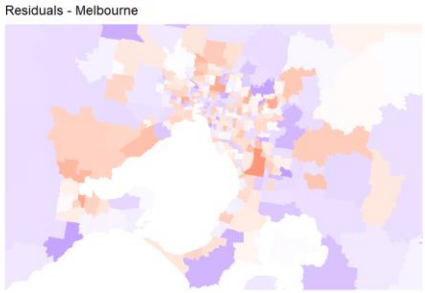


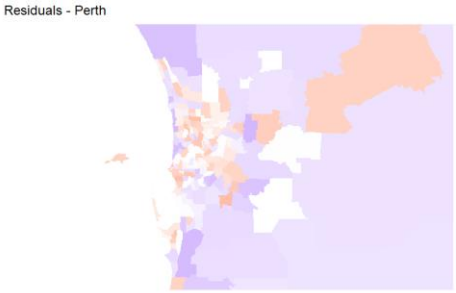
Localised autocorrelation (G=5)



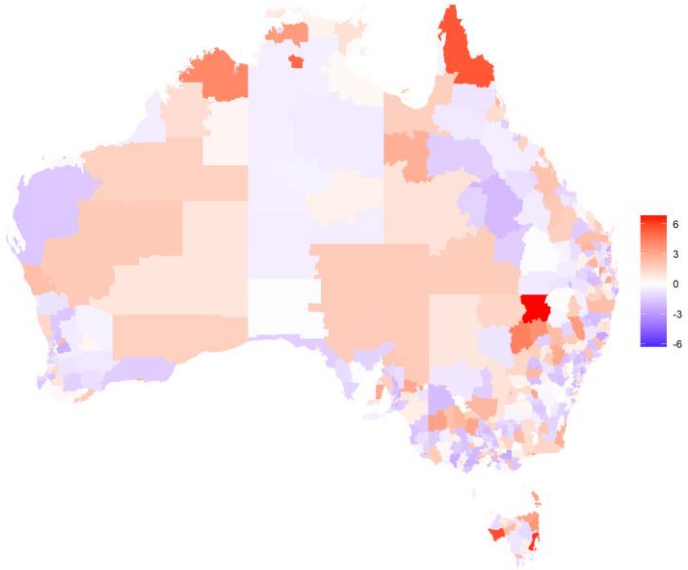


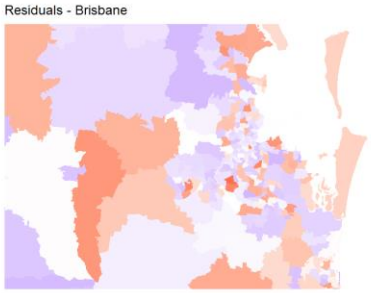


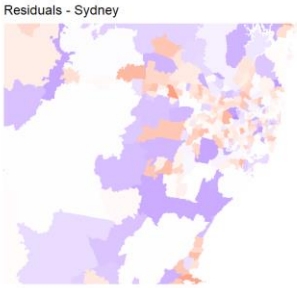


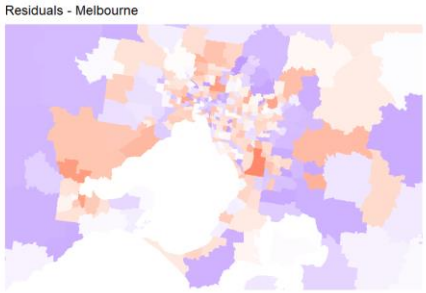


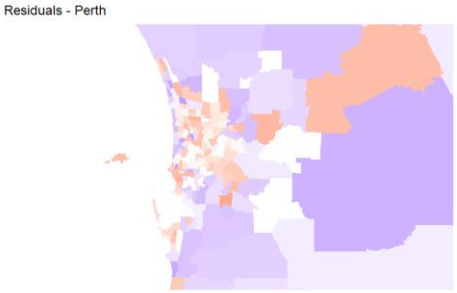
Locally adaptive (rho is determined when modelling)



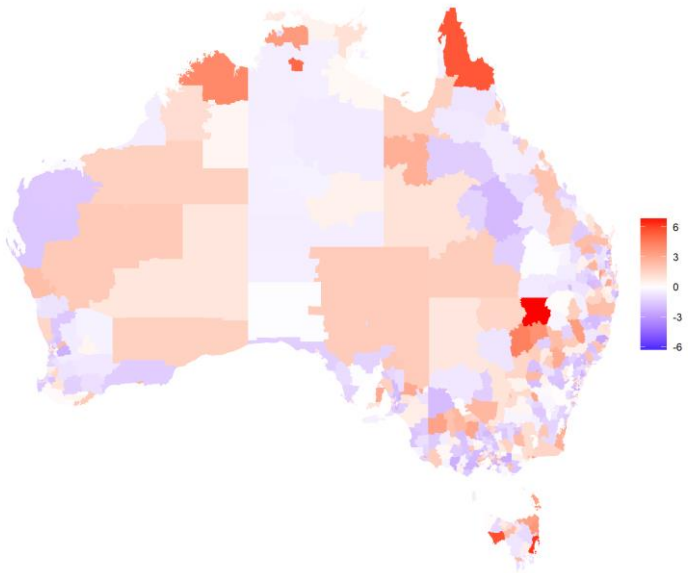


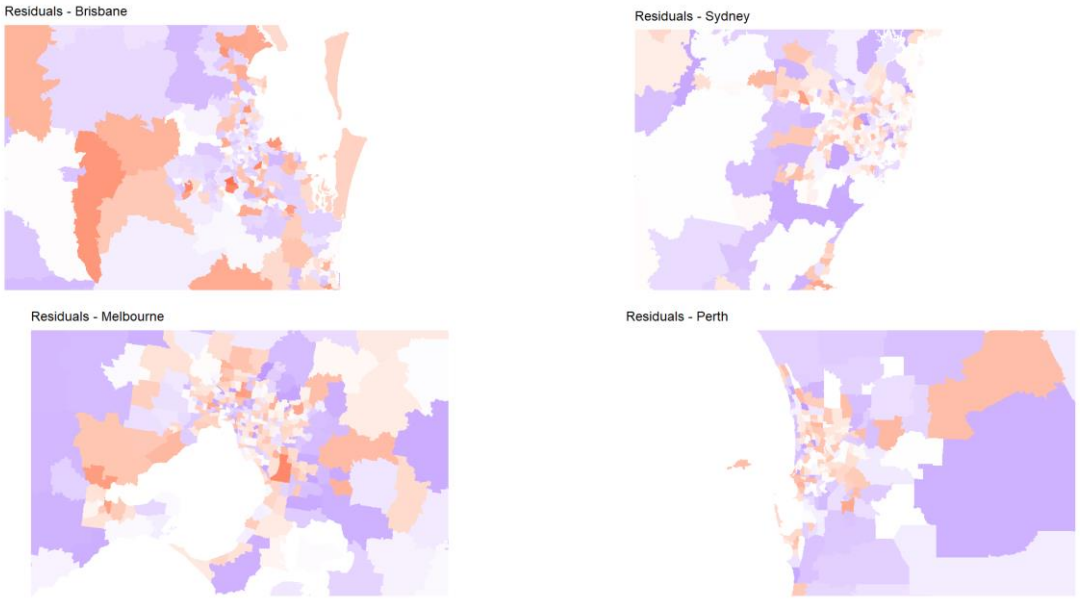




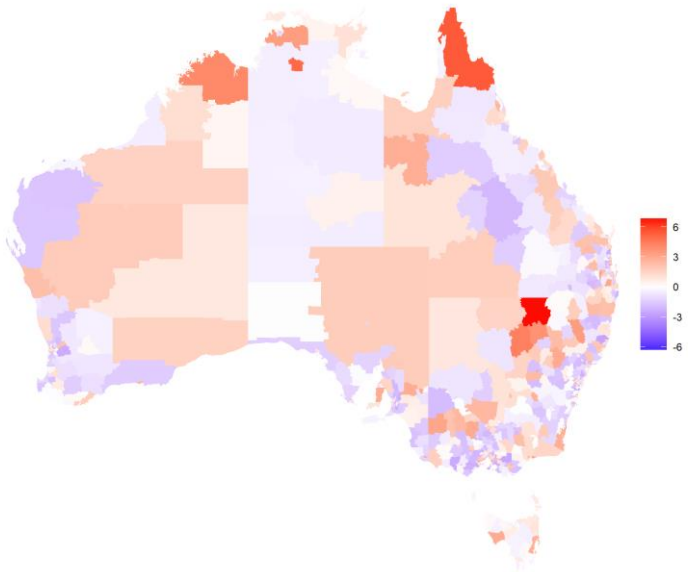


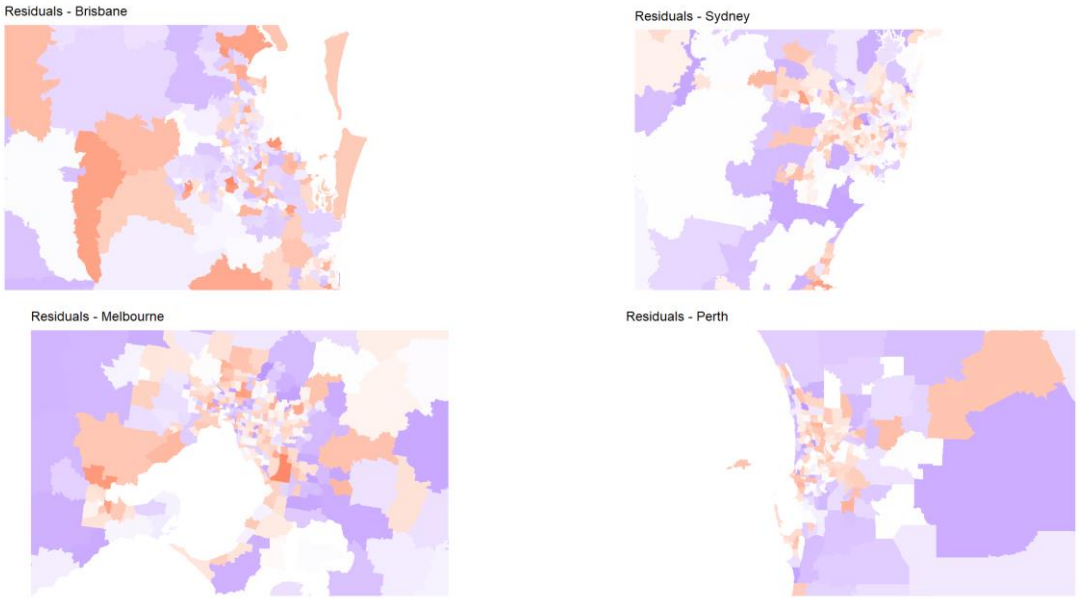
Locally adaptive (rho is fixed at 0.99)





Weighted sum of spatial priors



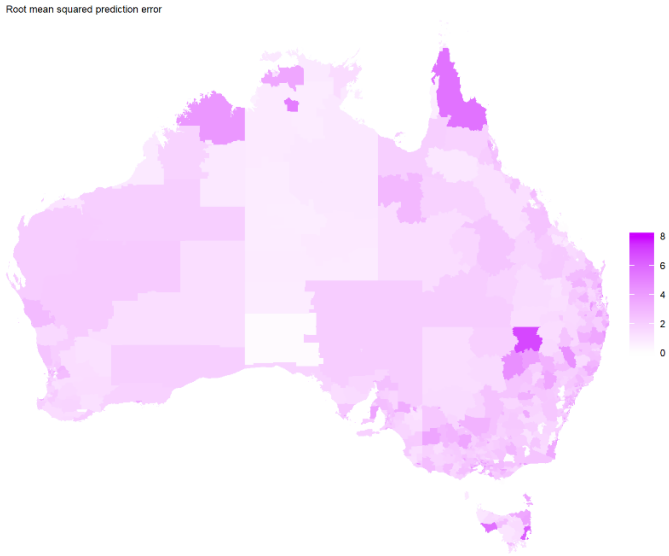


Leroux scale mixture model *Would not run for liver cancer, males*

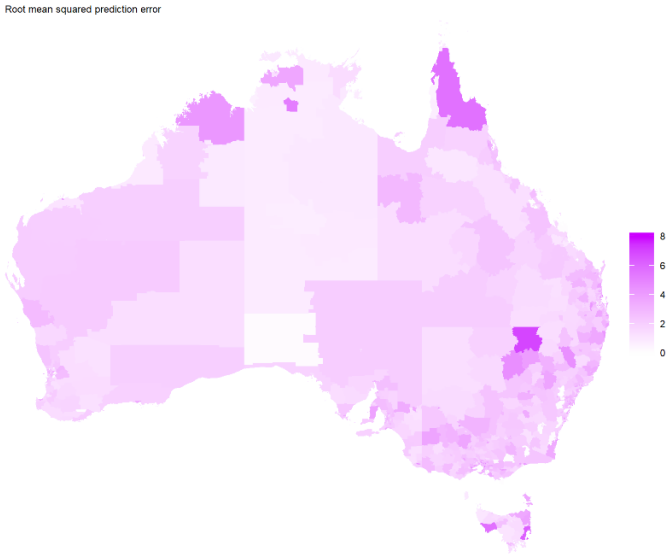
Note: The legend range was-6.0 to 6.7, and designed to enable differences between models to be visible. Areas shaded grey are outside this range.

Liver cancer, males, RMSPE

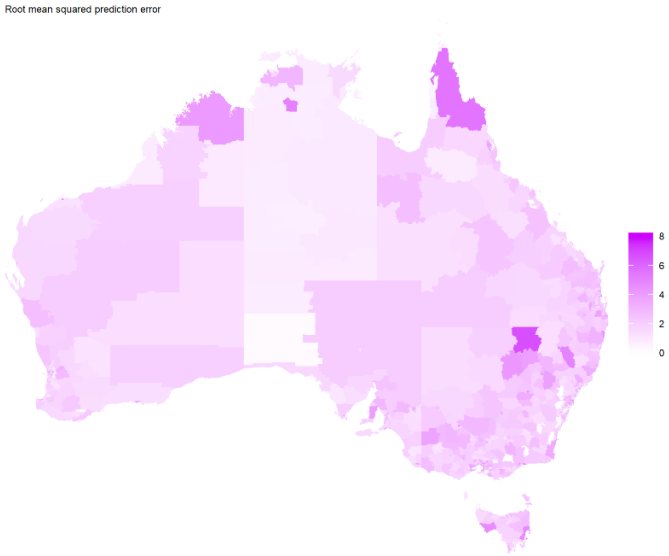
BYM



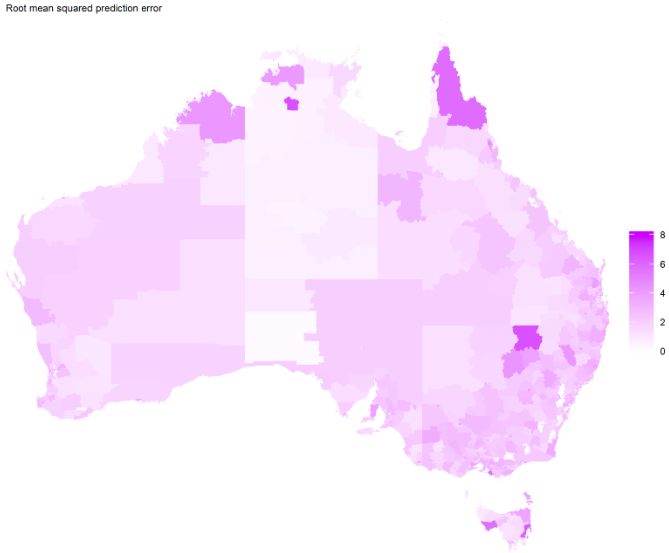
Leroux



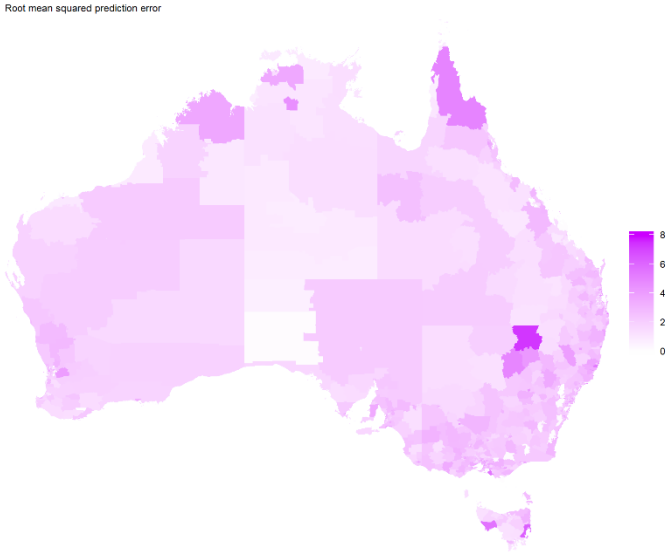
Geostatistical



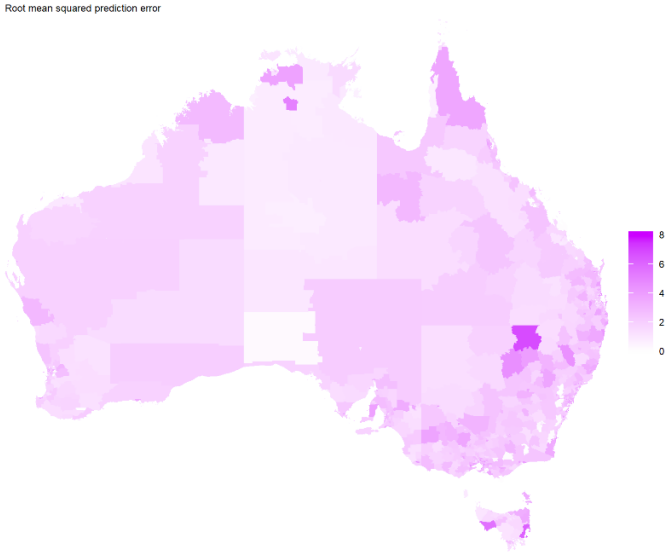
P-spline (tensor)



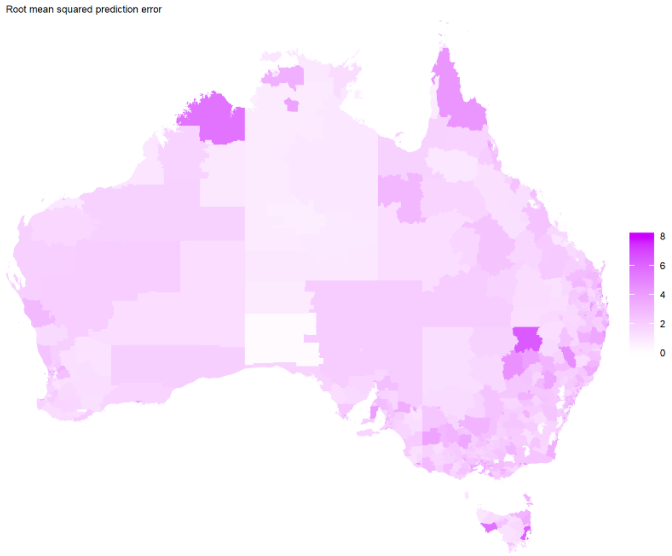
P-spline (radial)



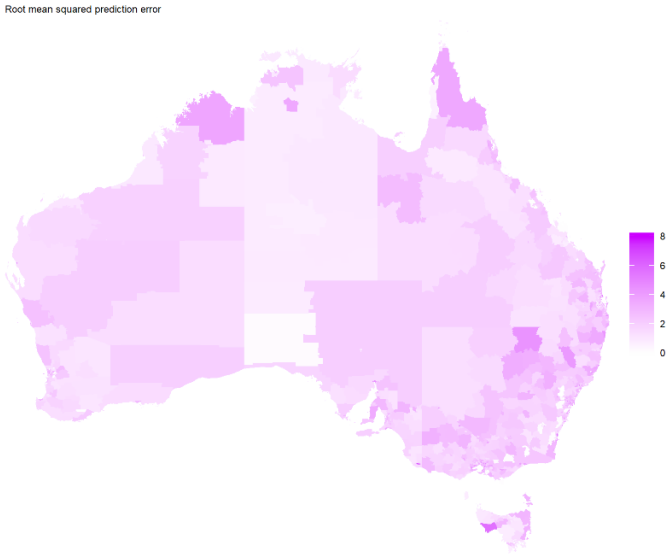
SEIFA dissimilarity (binary weighting)



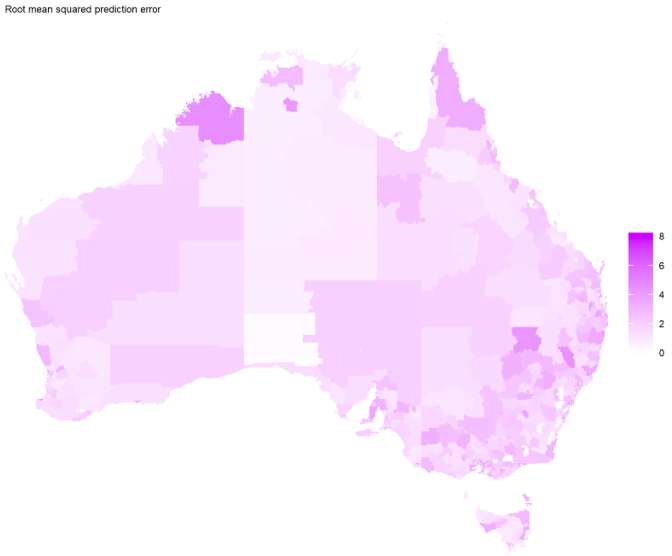
SEIFA dissimilarity (non-binary weighting)



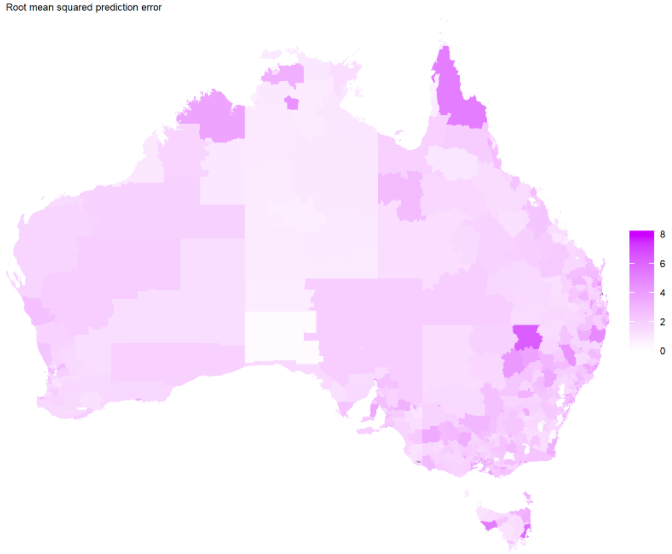
Residual dissimilarity (binary weighting)



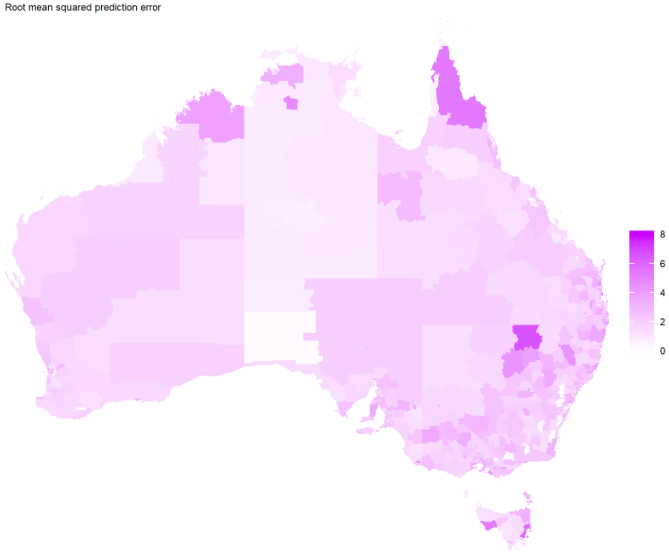
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)



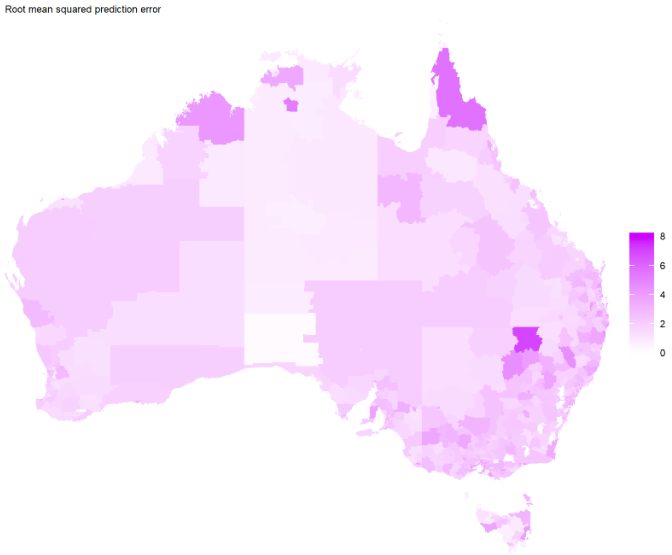
Locally adaptive (rho is determined when modelling)

[Unavailable from INLA output]

Locally adaptive (rho is fixed at 0.99)

[Unavailable from INLA output]

Weighted sum of spatial priors



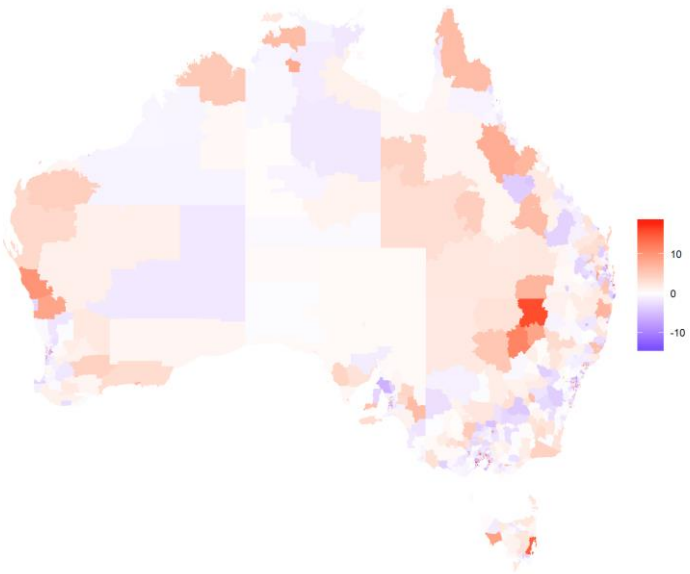
Leroux scale mixture model

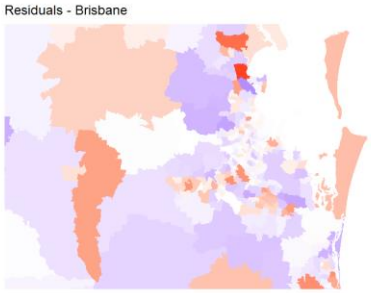
Model would not run for liver cancer, males

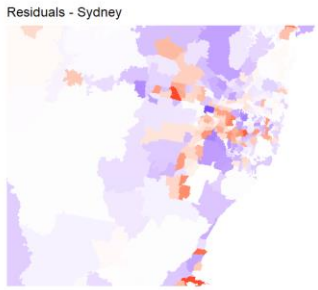
Note: The maximum legend value was set to 8. Areas shaded grey are higher than this value.

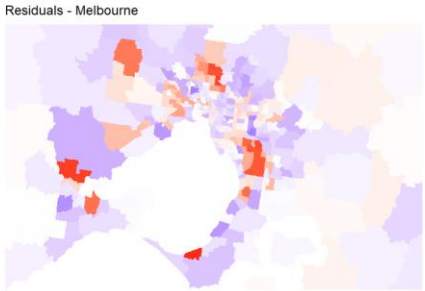
Lung cancer, males, residuals

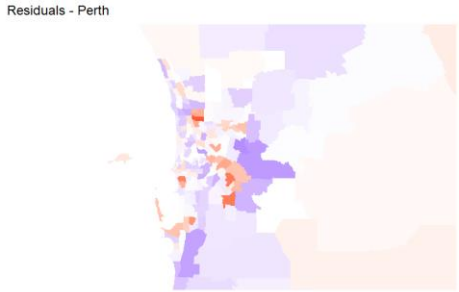
BYM



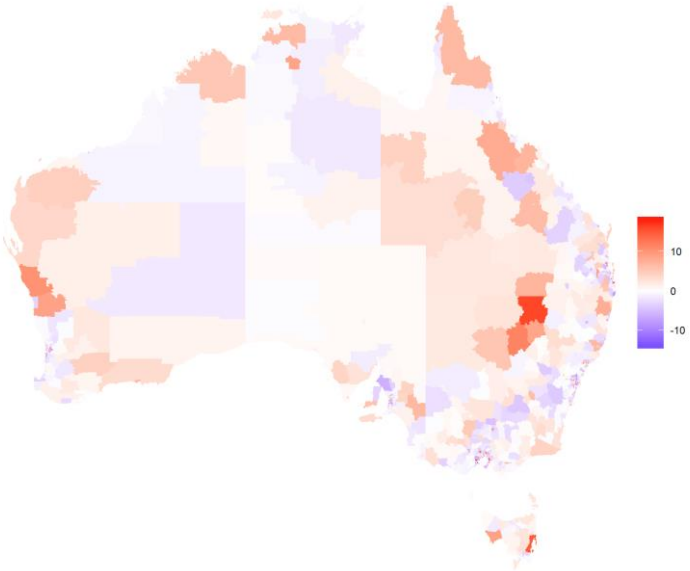


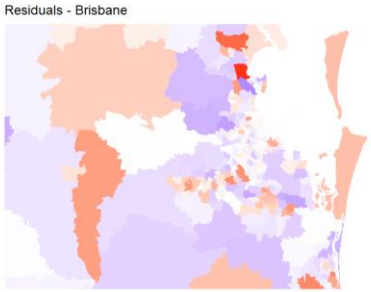


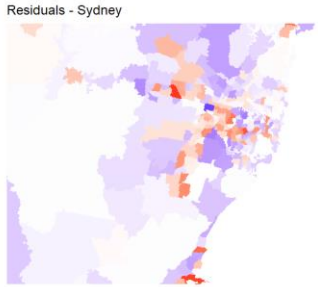


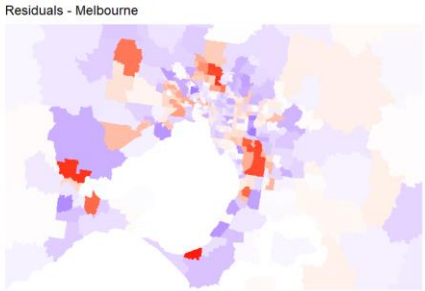


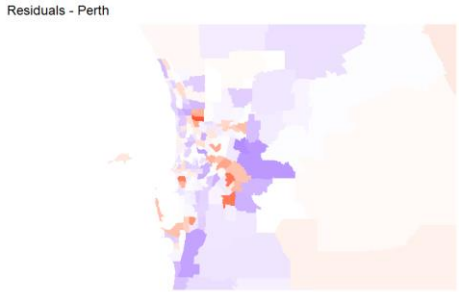
Leroux



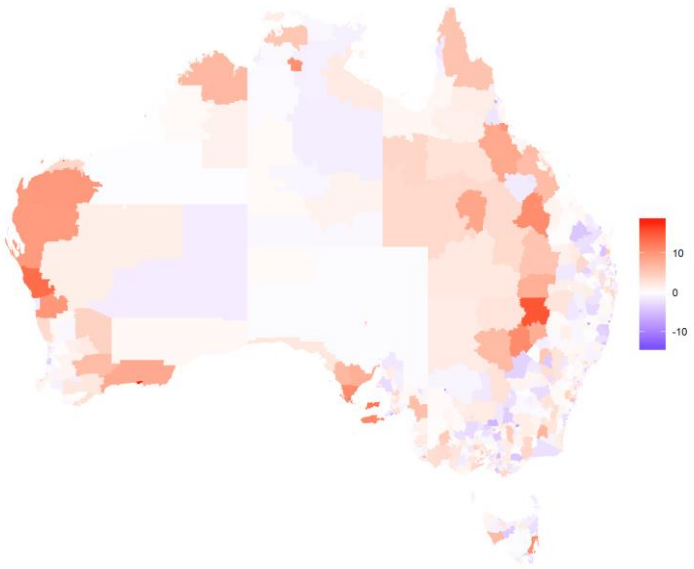


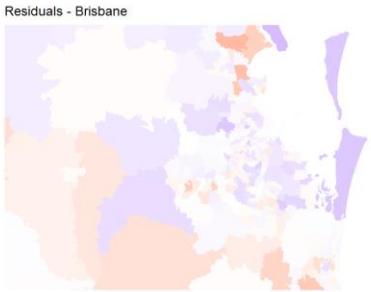


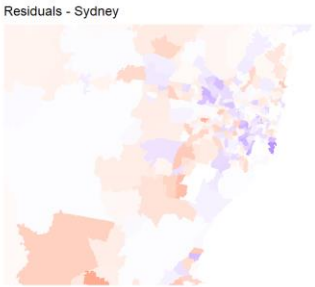


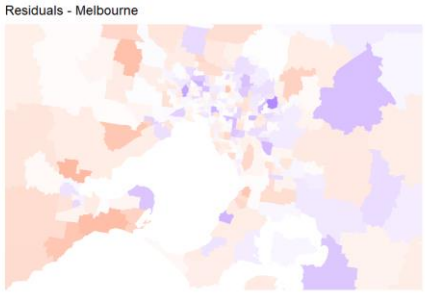


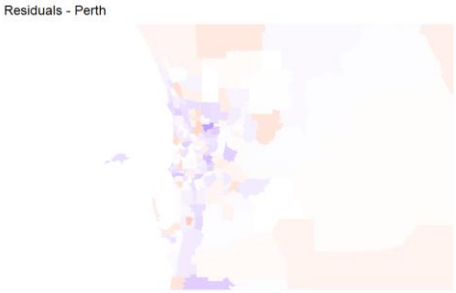
Geostatistical



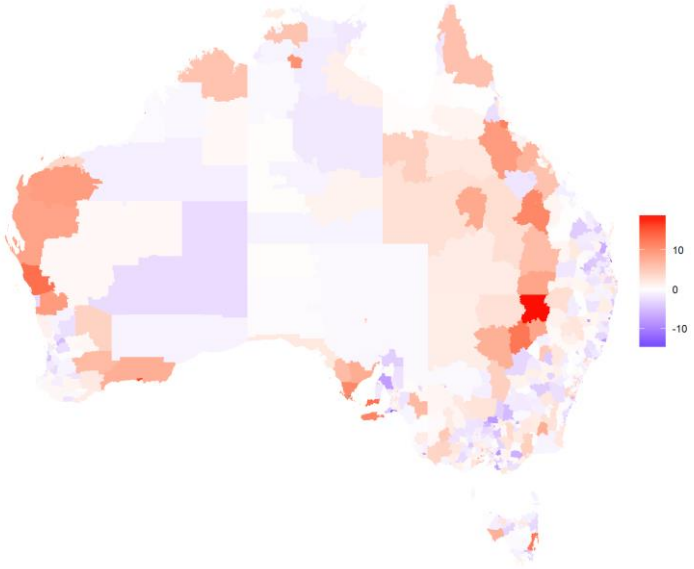


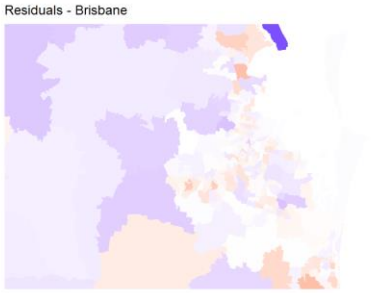


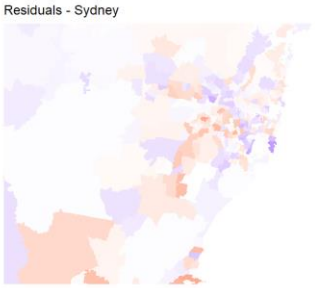


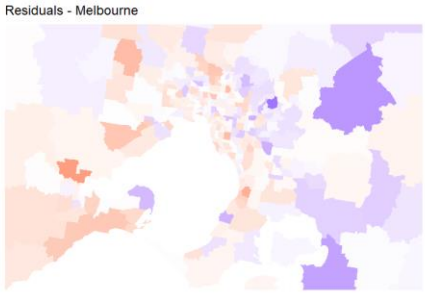


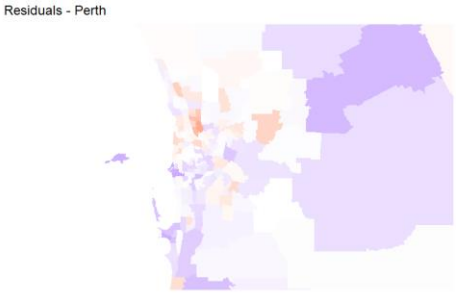
P-spline (tensor)



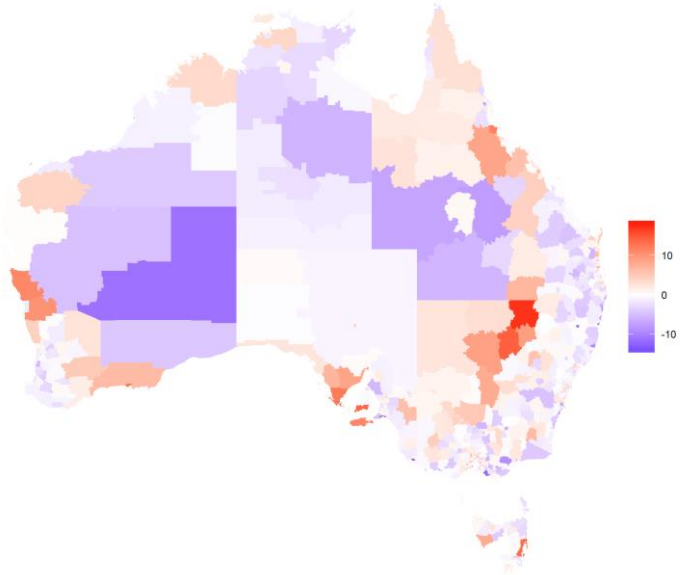


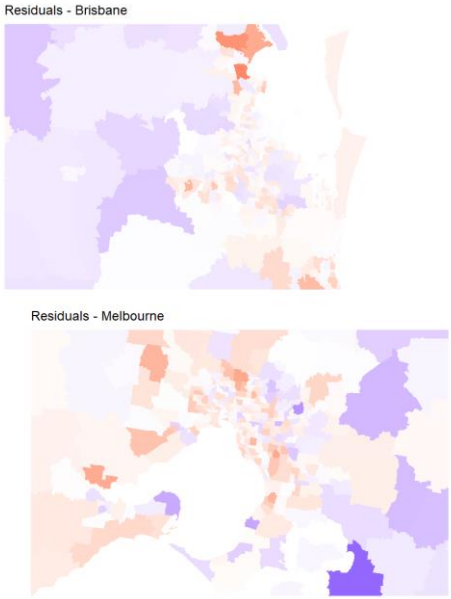


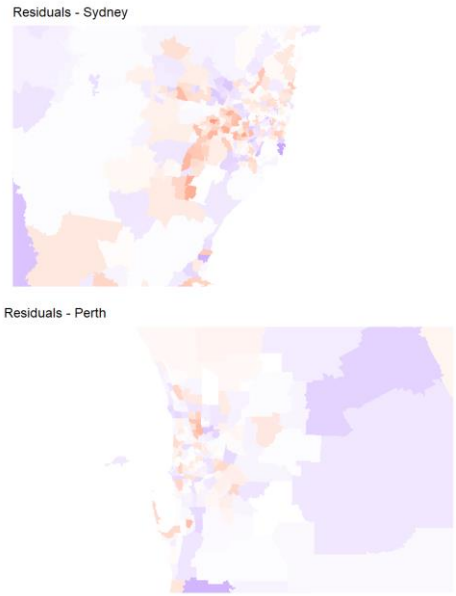




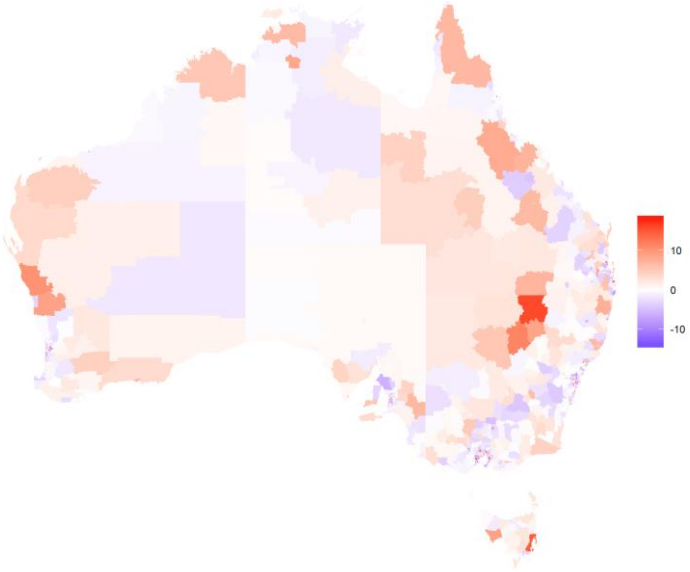
P-spline (radial)

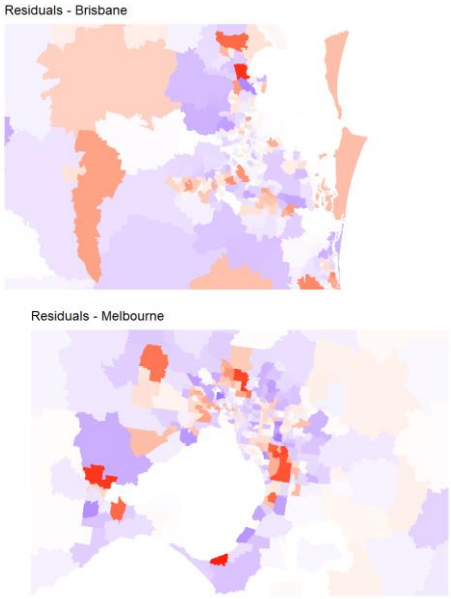


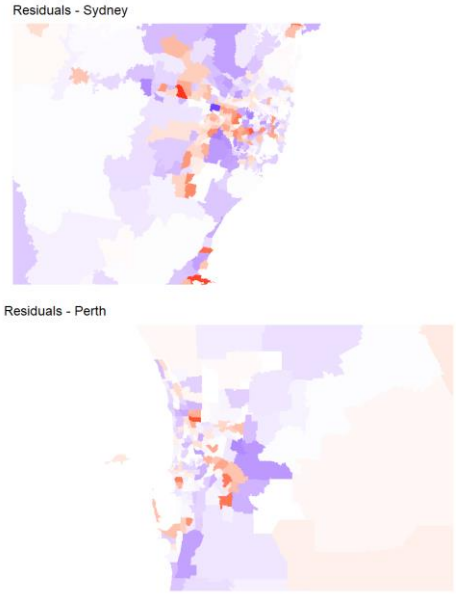




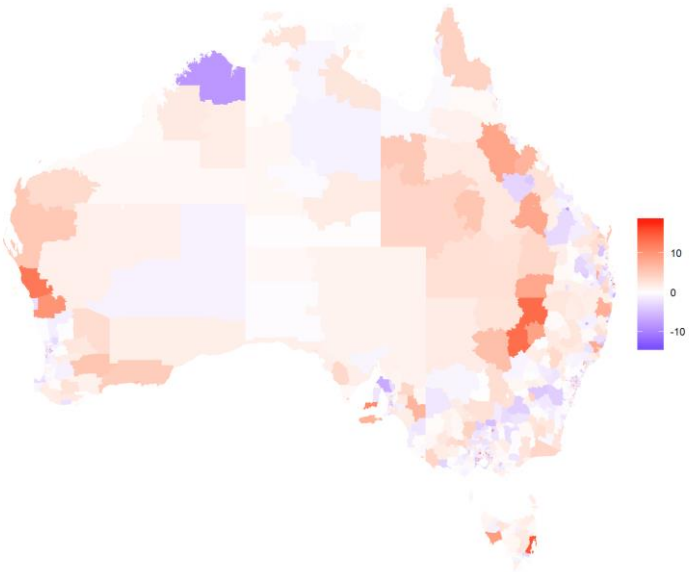
SEIFA dissimilarity (binary weighting)

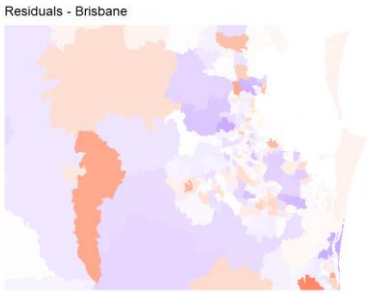


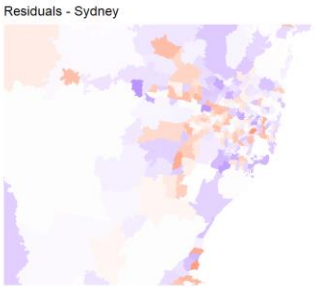


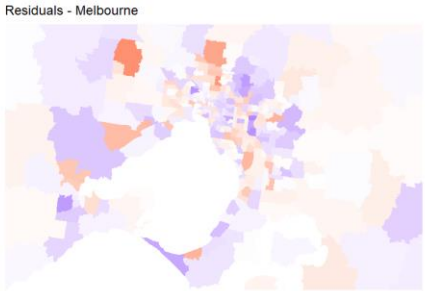


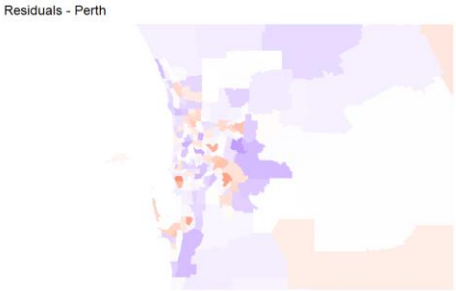
SEIFA dissimilarity (non-binary weighting)



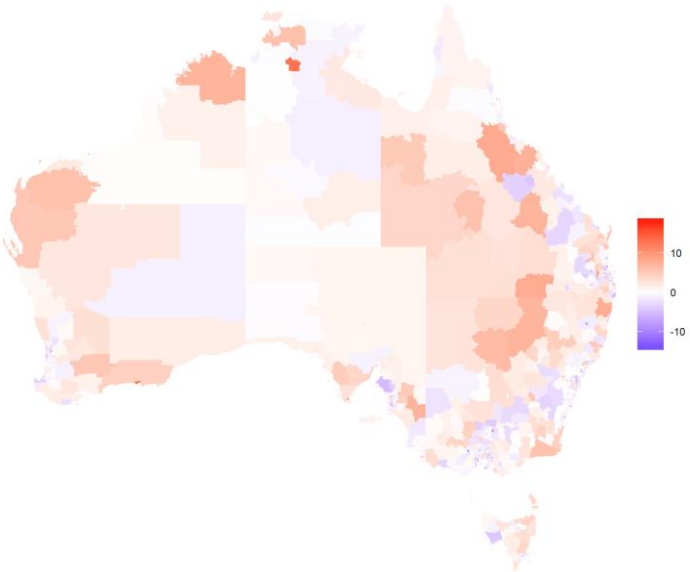


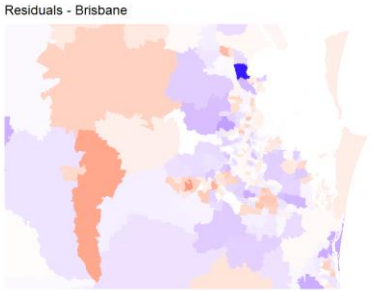


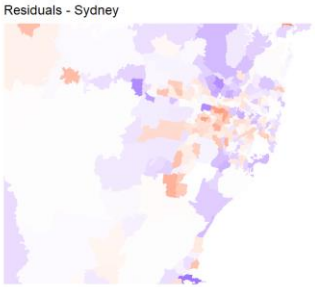


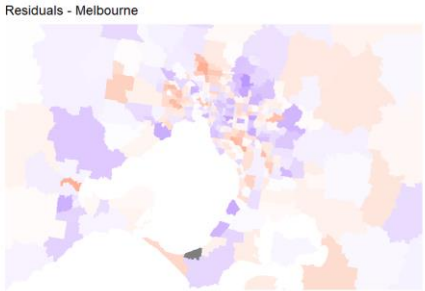


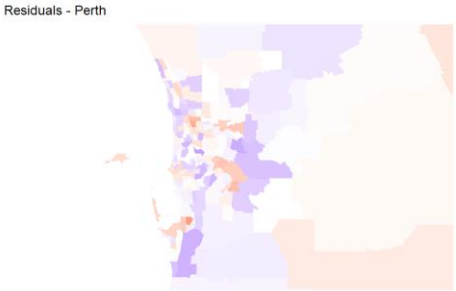
Residual dissimilarity (binary weighting)



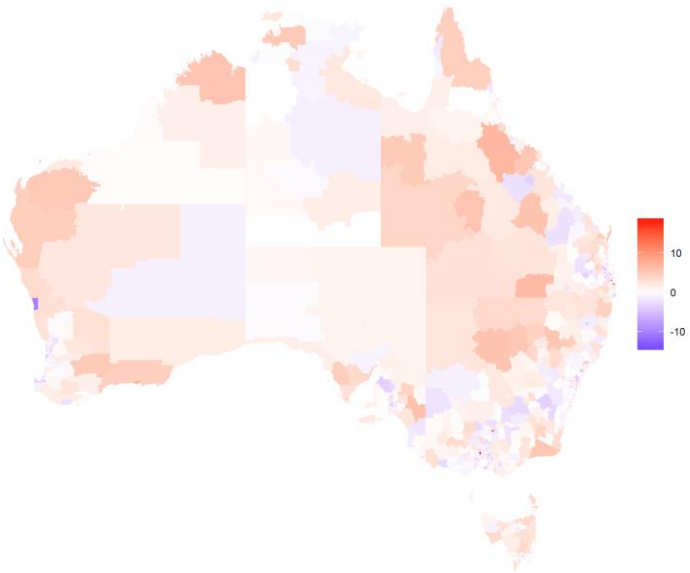


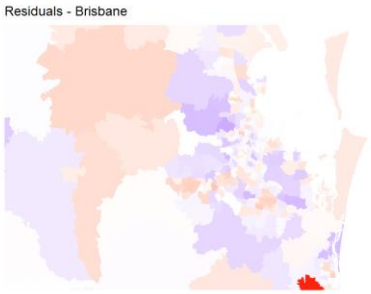


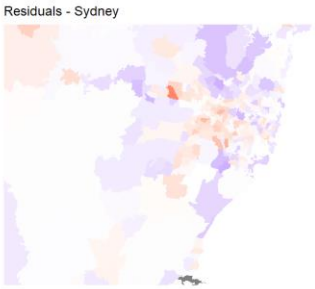


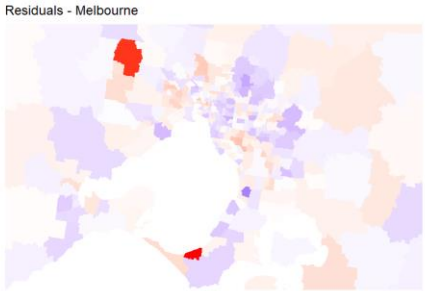


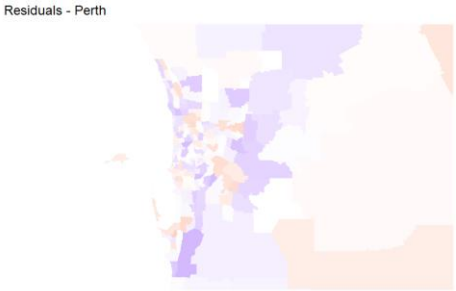
Residual dissimilarity (non-binary weighting)



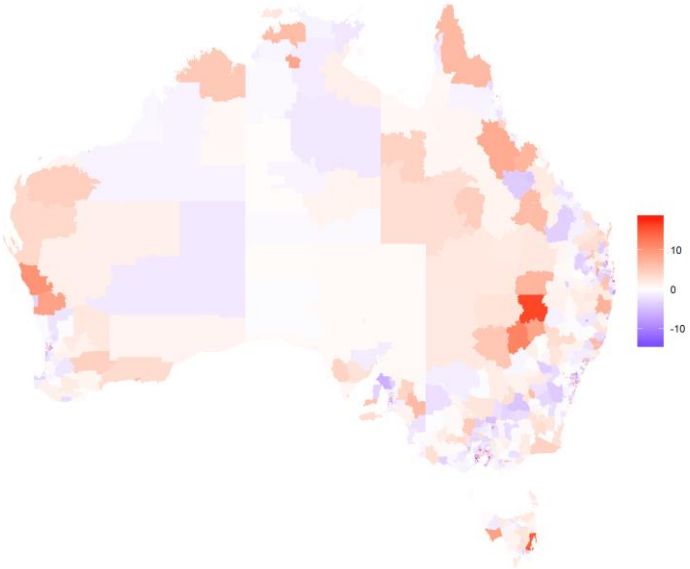


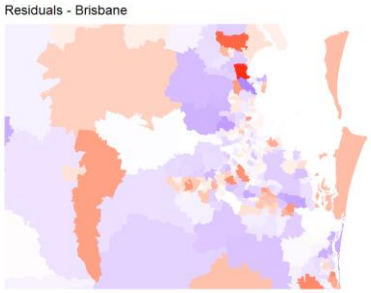


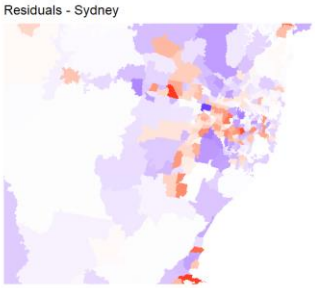


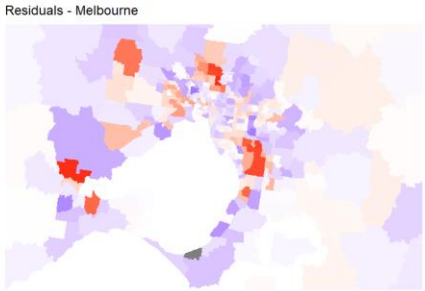


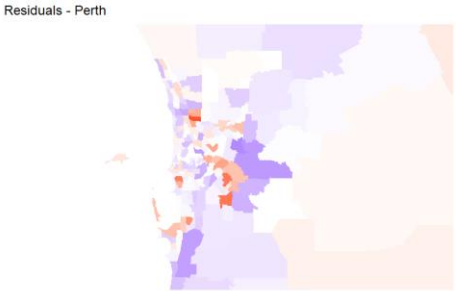
Localised autocorrelation (G=3)



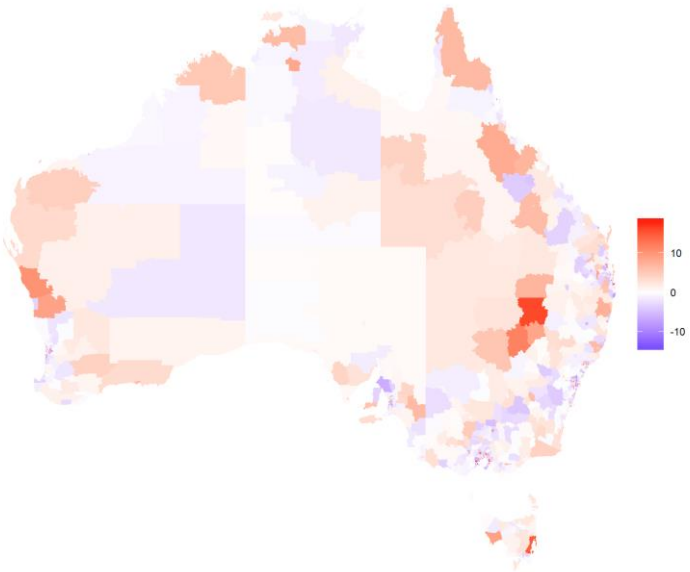


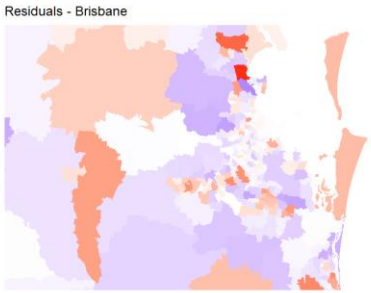


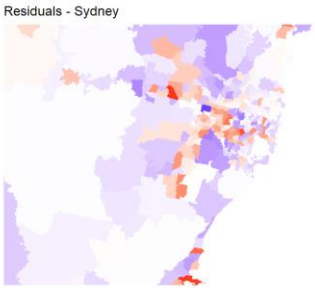


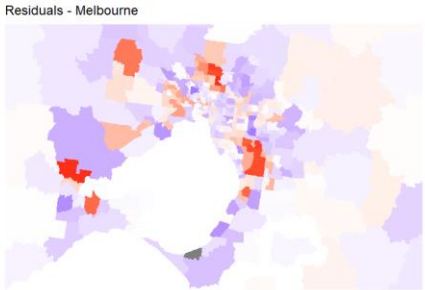


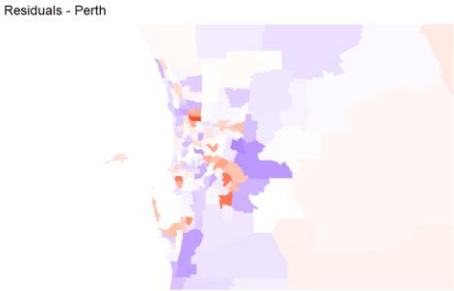
Localised autocorrelation (G=5)



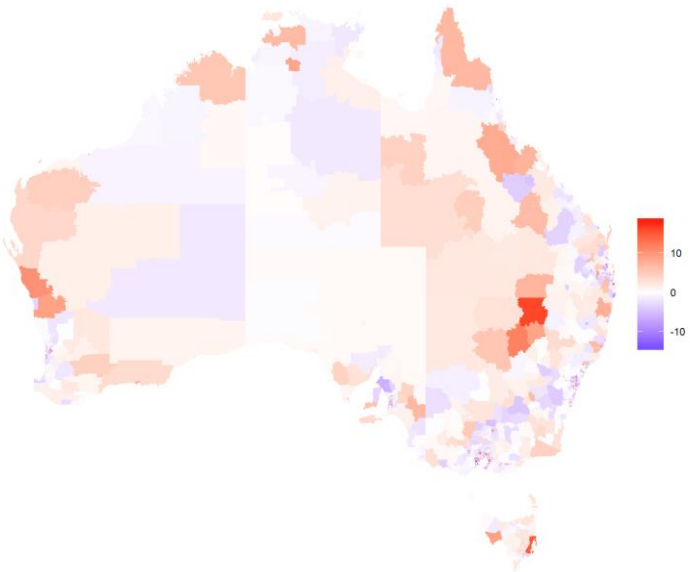


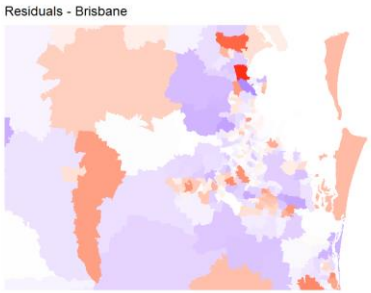


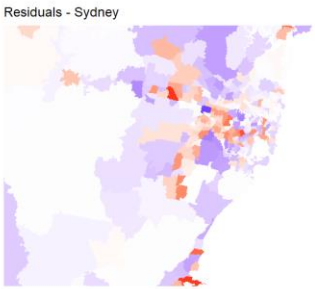


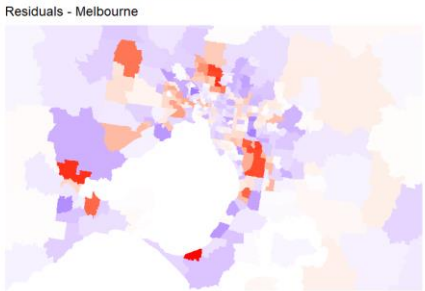


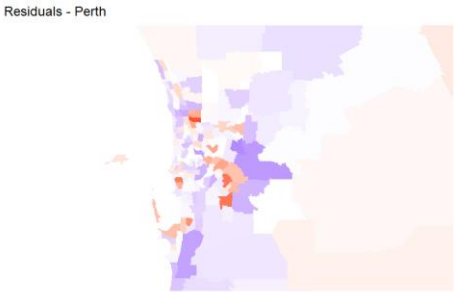
Locally adaptive (rho is determined when modelling)



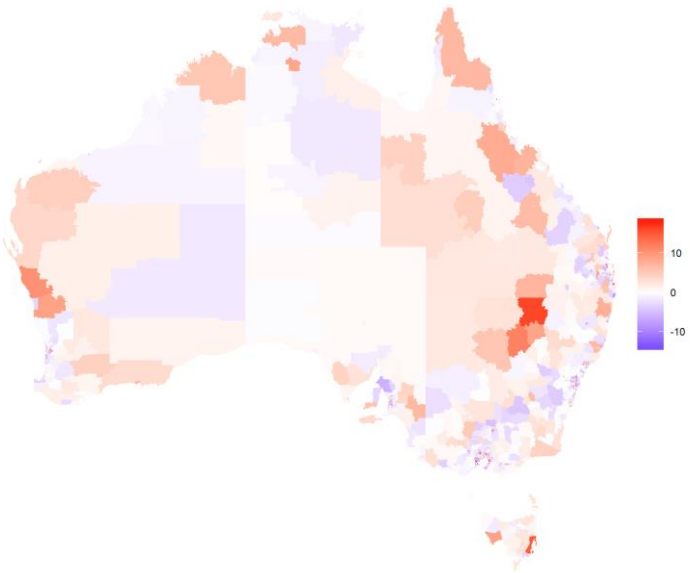


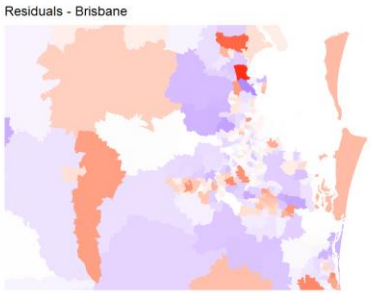


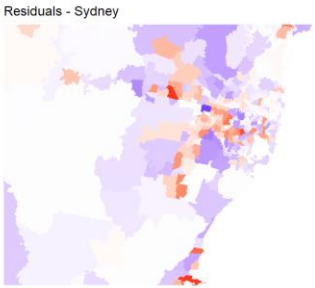


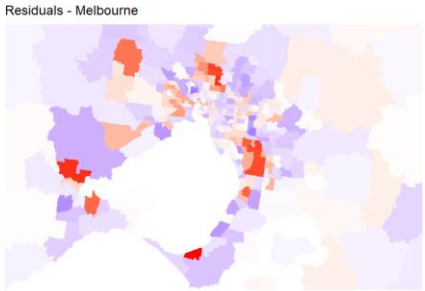


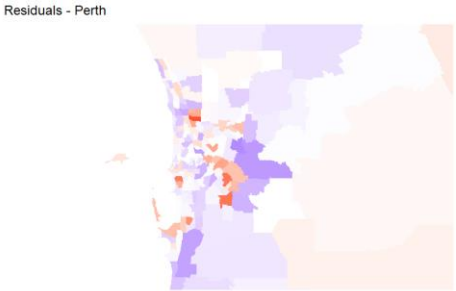
Locally adaptive (rho is fixed at 0.99)



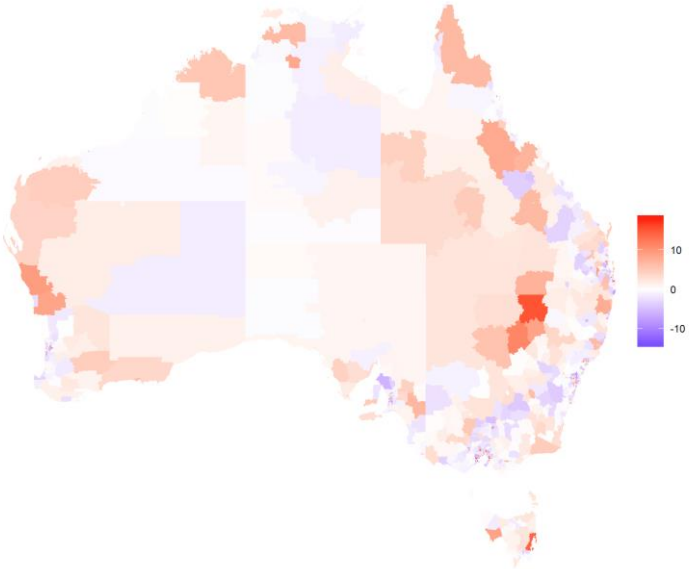


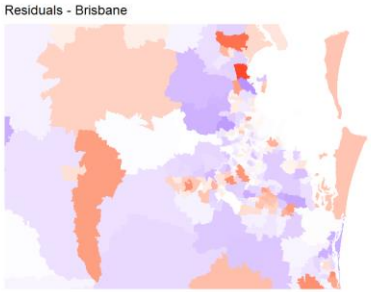


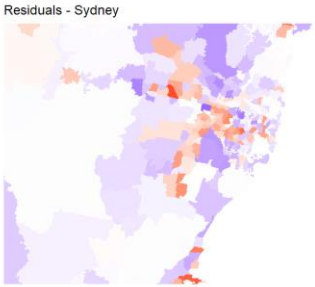


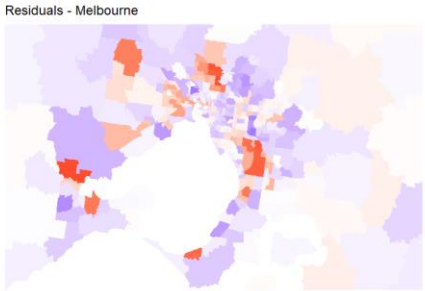


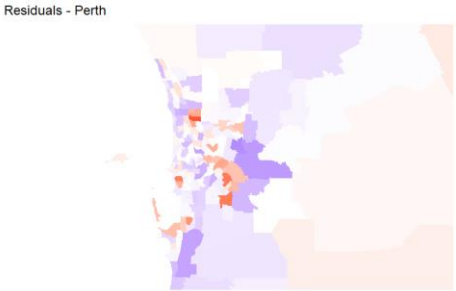
Weighted sum of spatial priors



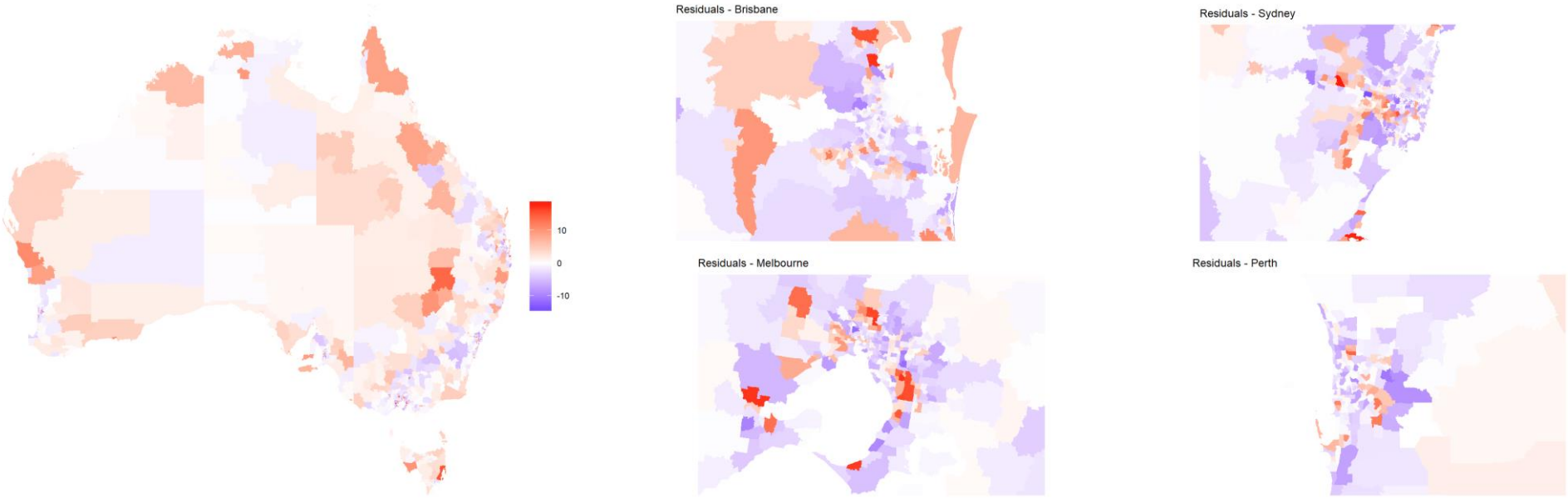








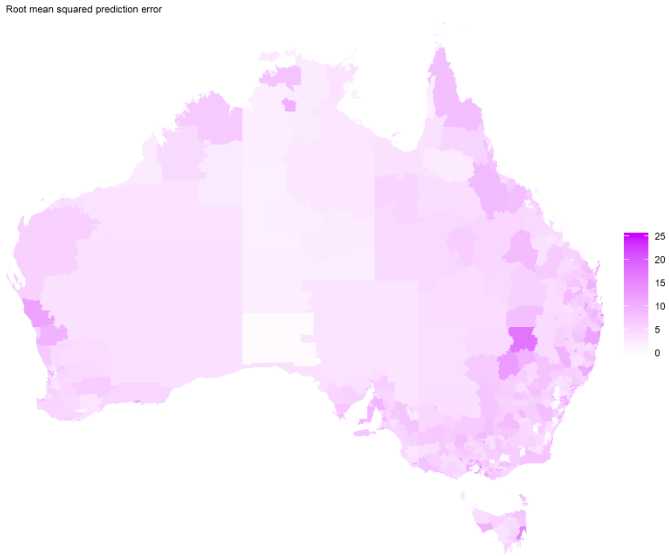
Leroux scale mixture model



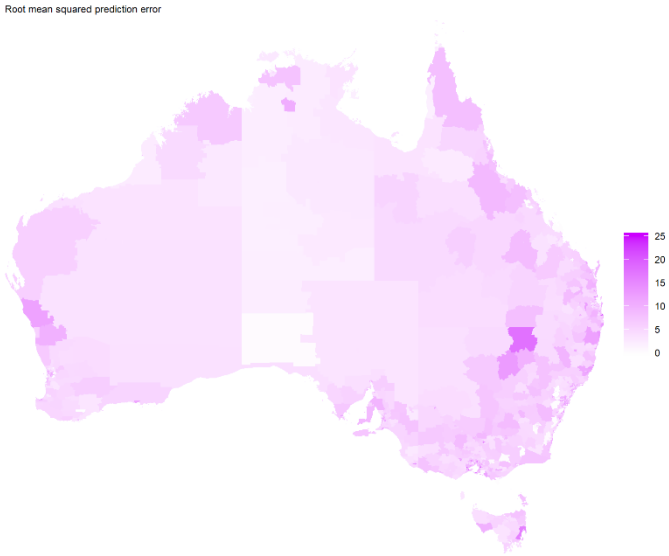
Note: The legend range was-15.1 to 18.7, and designed to enable differences between models to be visible. Areas shaded grey are outside this range.

Lung cancer, males, RMSPE

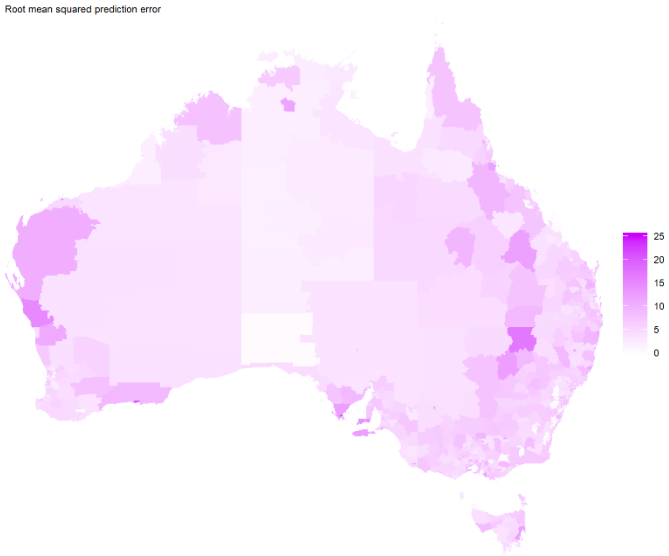
BYM



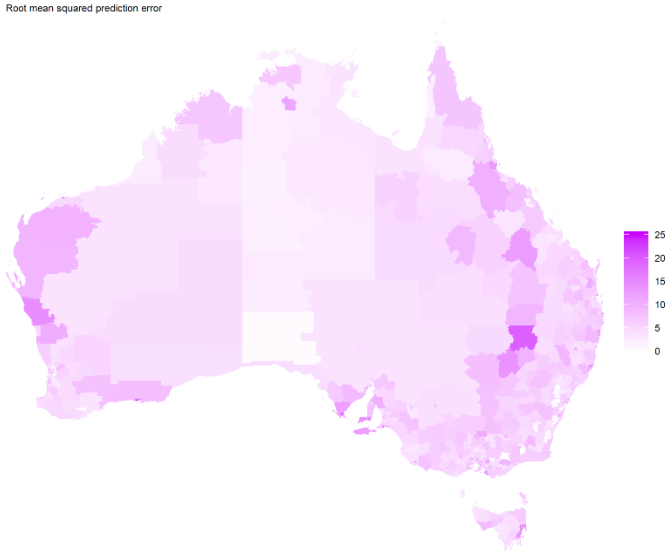
Leroux



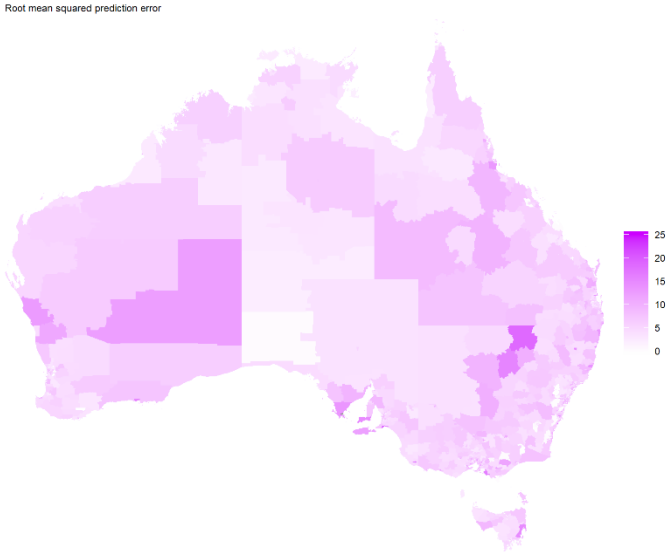
Geostatistical



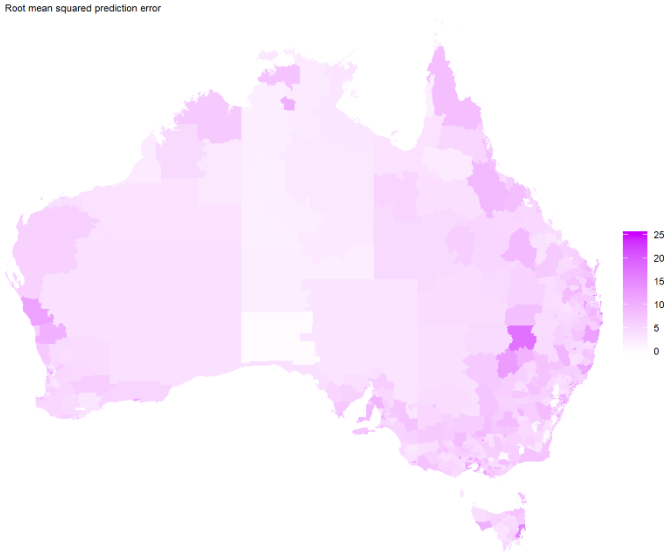
P-spline (tensor)



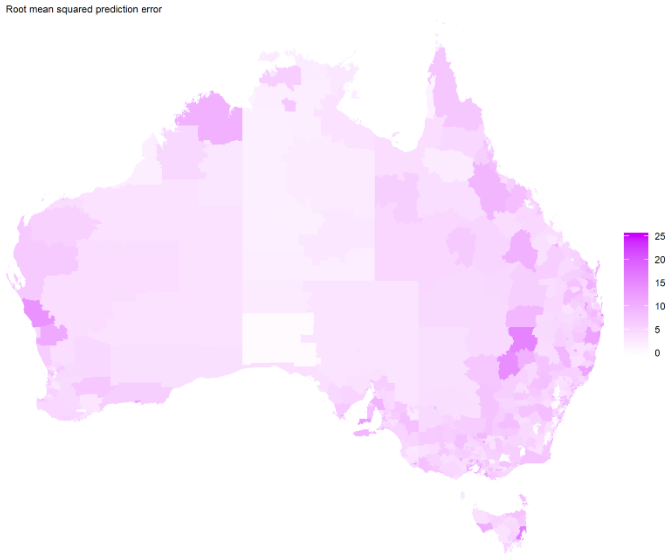
P-spline (radial)



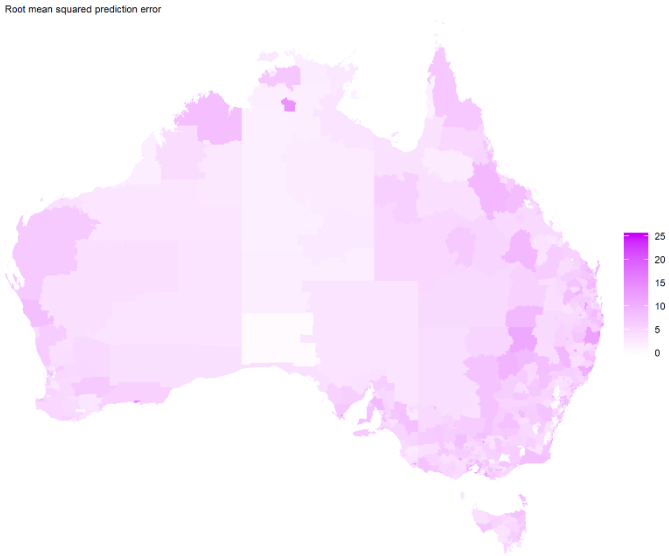
SEIFA dissimilarity (binary weighting)



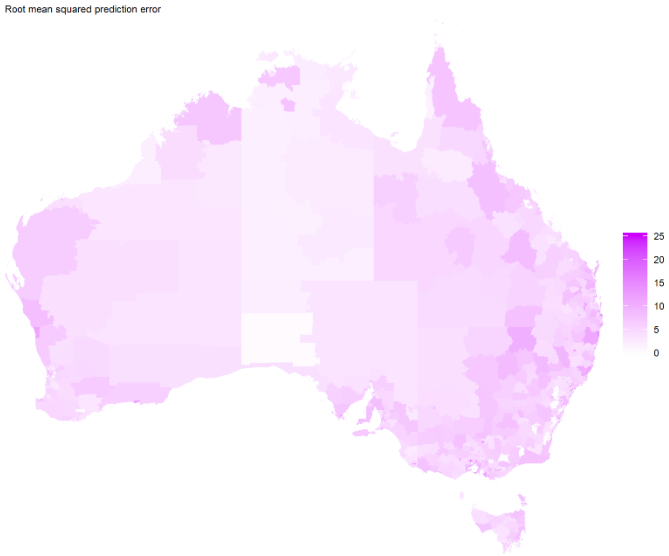
SEIFA dissimilarity (non-binary weighting)



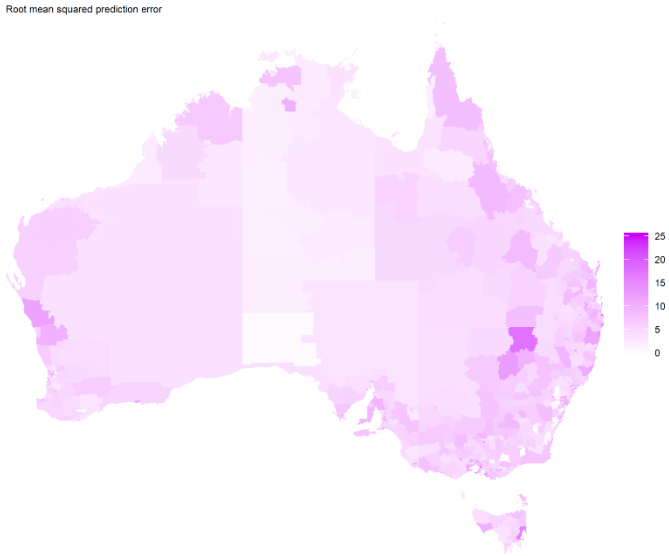
Residual dissimilarity (binary weighting)



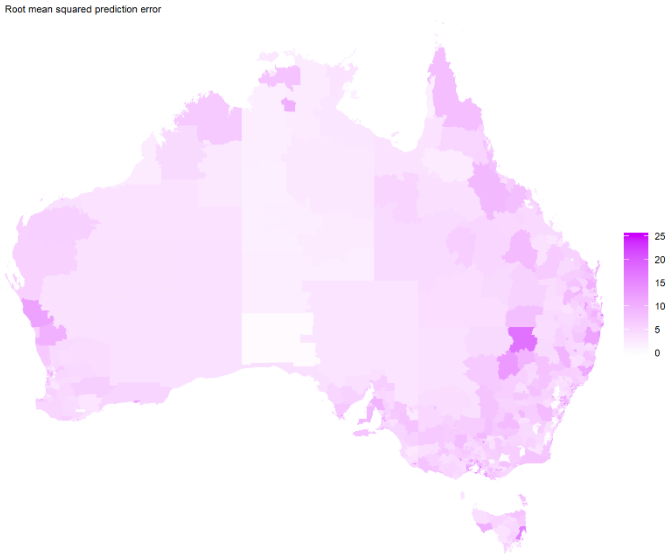
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)



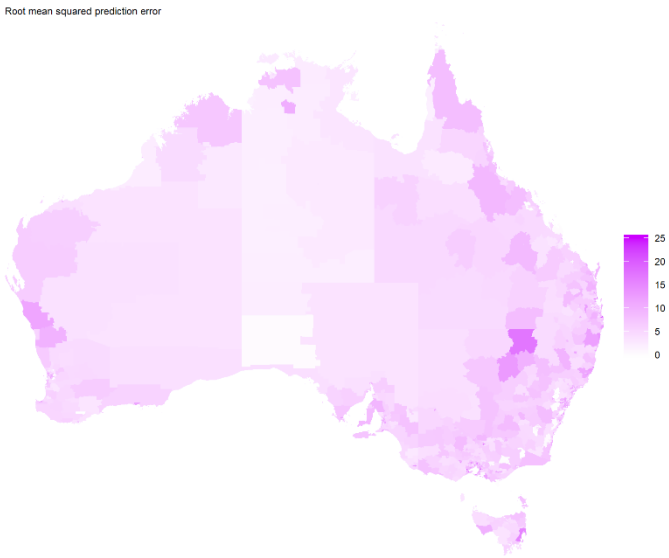
Locally adaptive (rho is determined when modelling)

[Unavailable from INLA output]

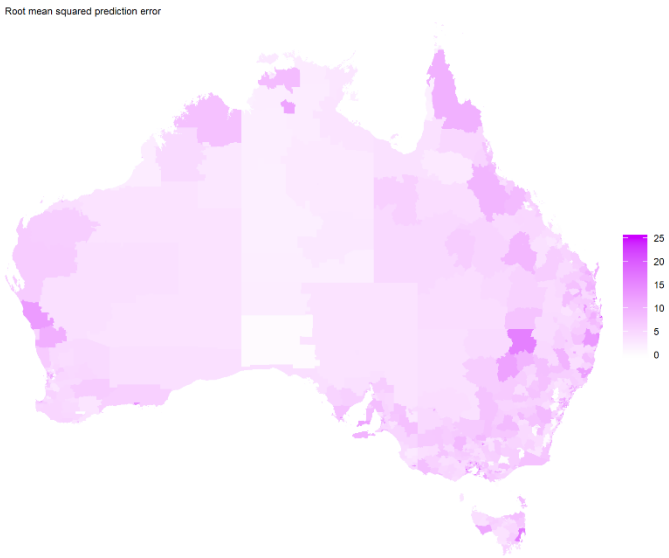
Locally adaptive (rho is fixed at 0.99)

[Unavailable from INLA output]

Weighted sum of spatial priors



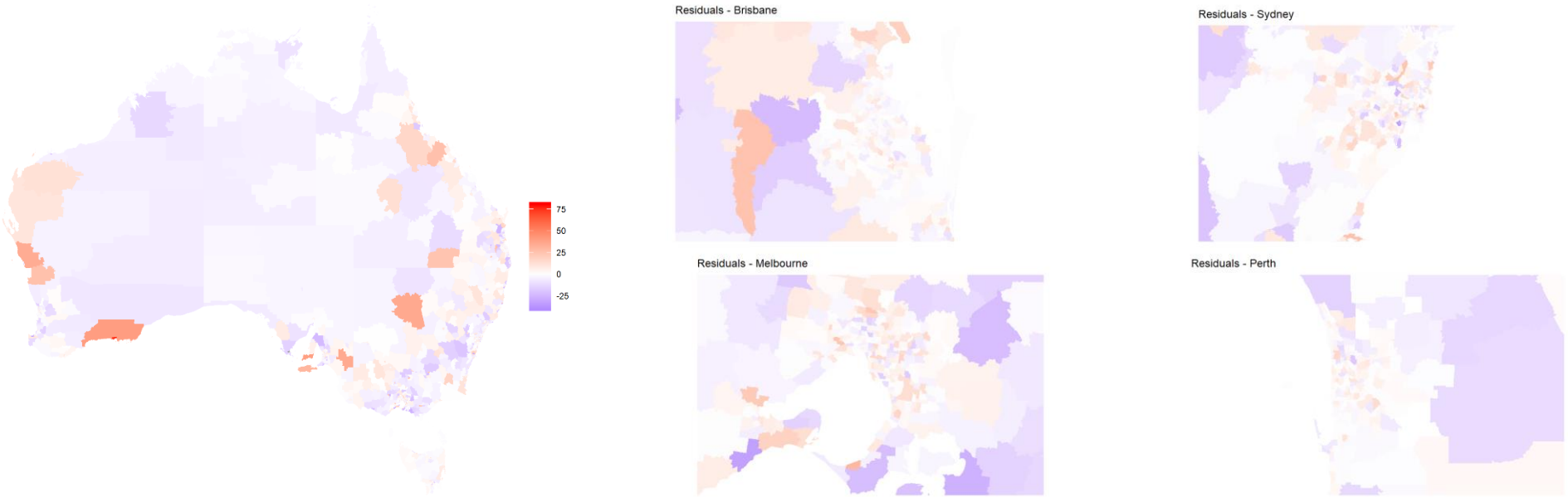
Leroux scale mixture model



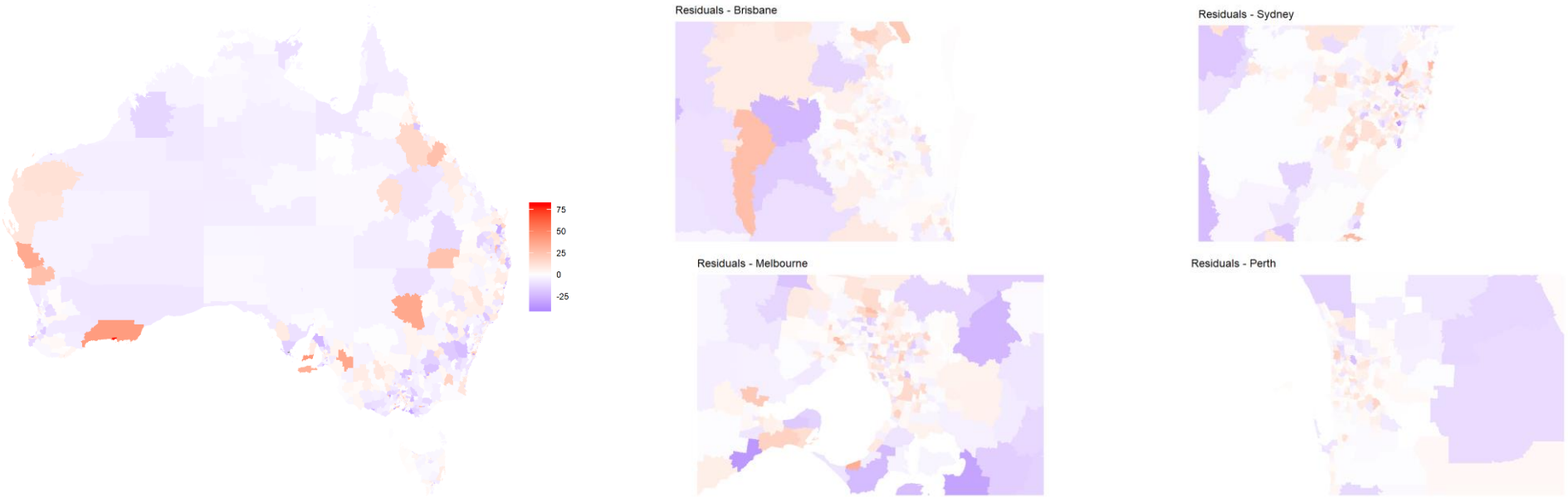
Note: The maximum legend value was set to 25. Areas shaded grey are higher than this value.

All invasive cancers, females, residuals

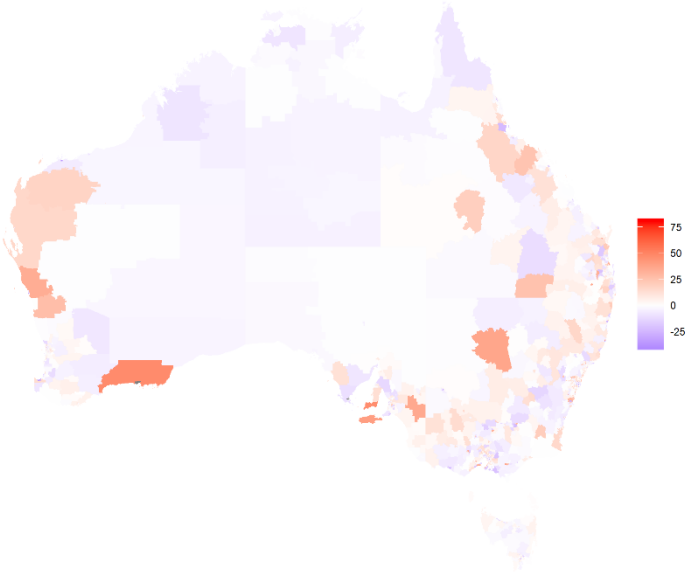
BYM

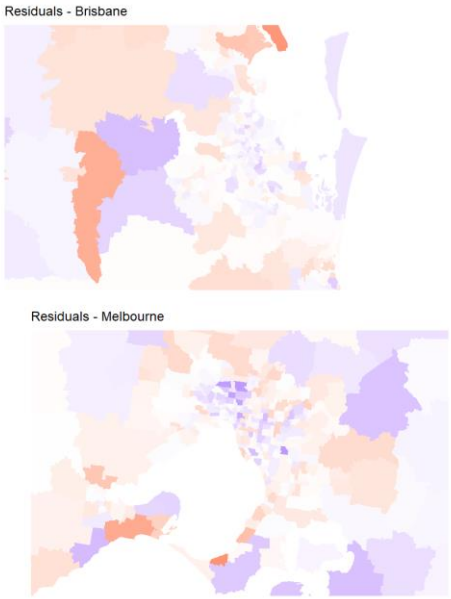


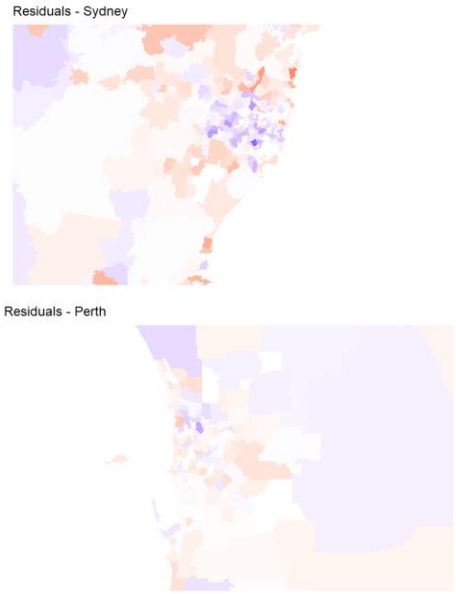
Leroux



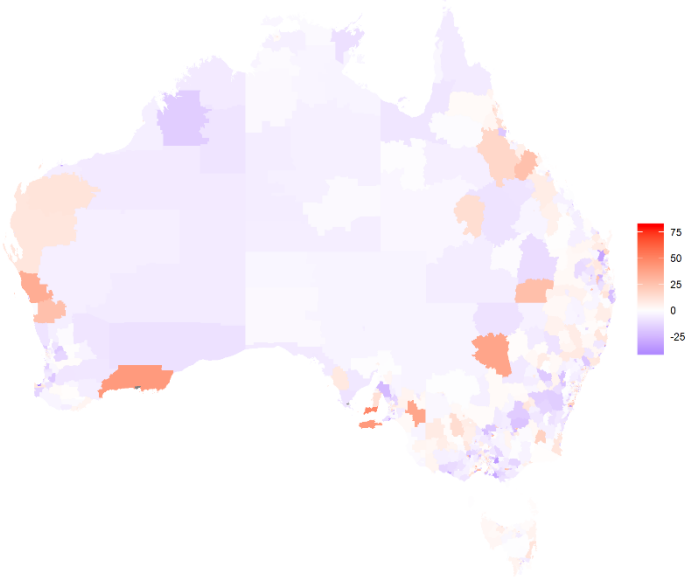
Geostatistical

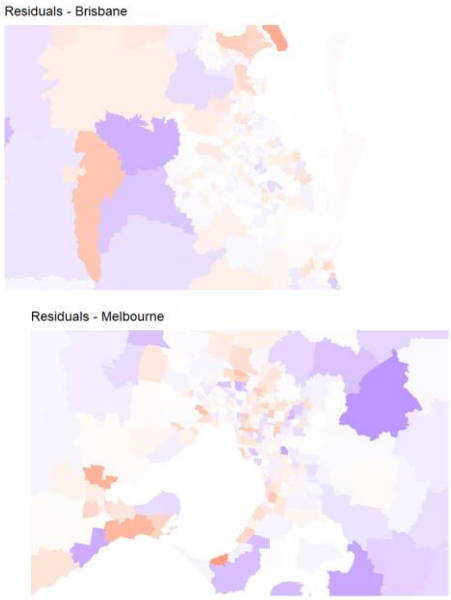


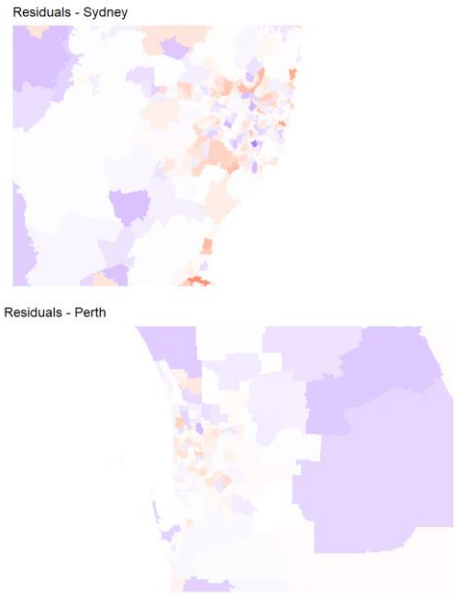




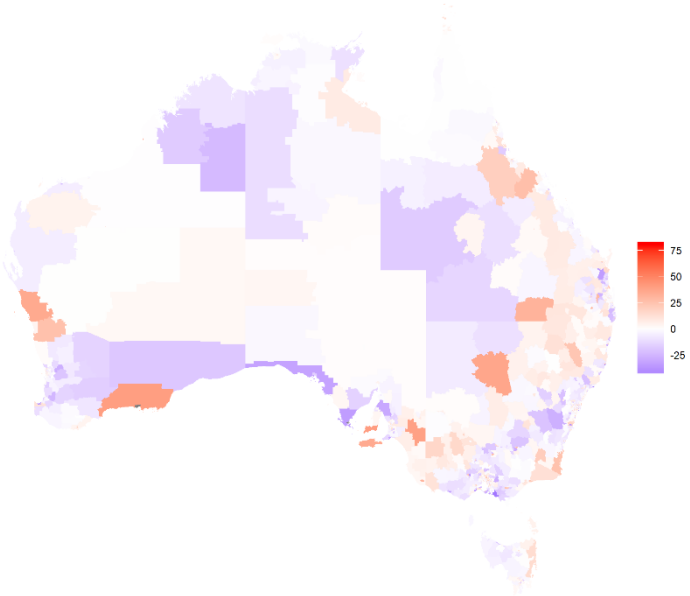
P-spline (tensor)



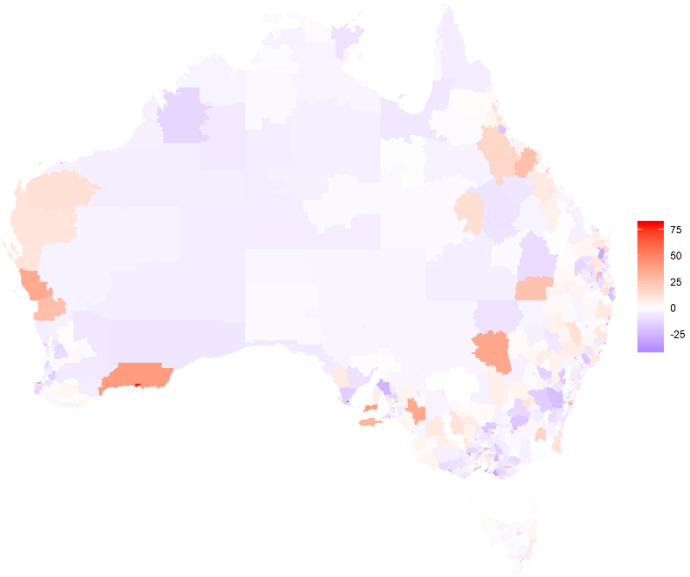


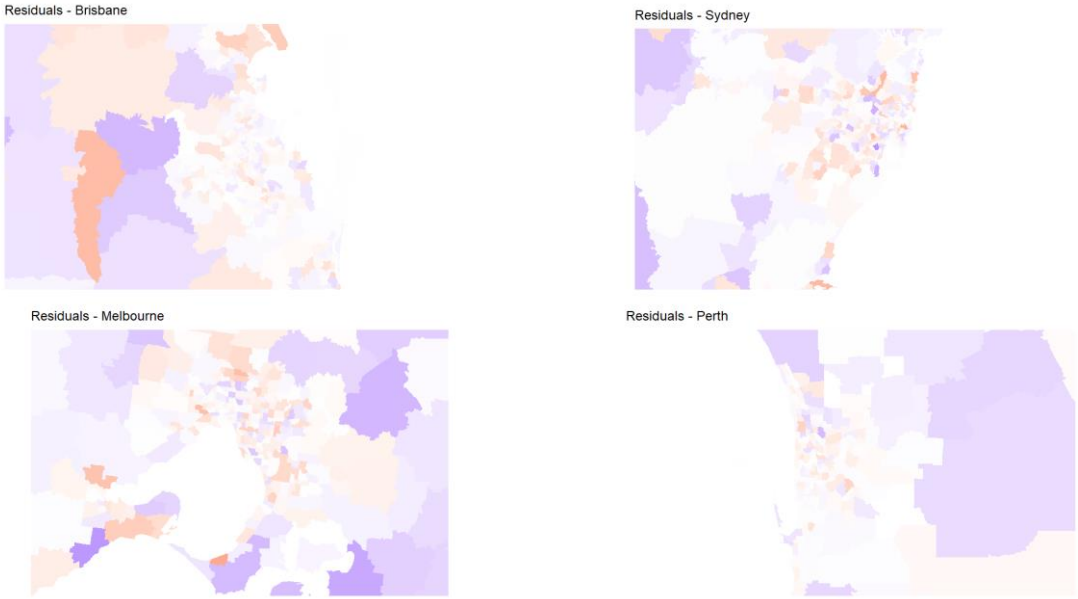


P-spline (radial)

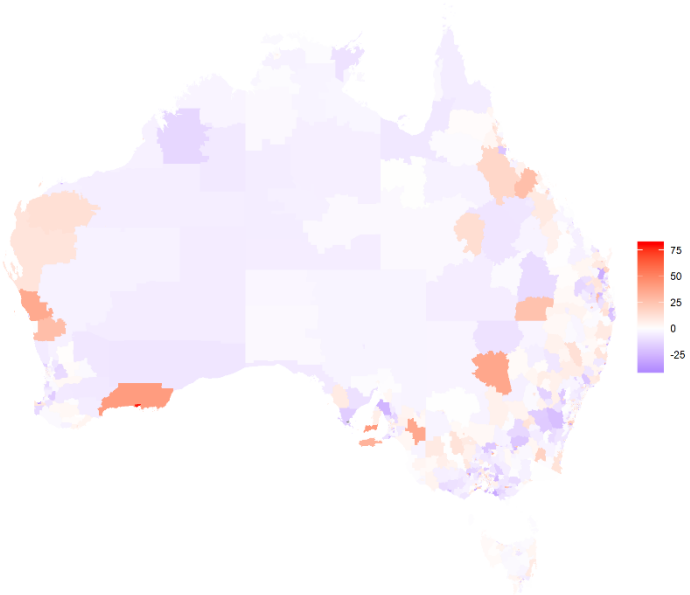


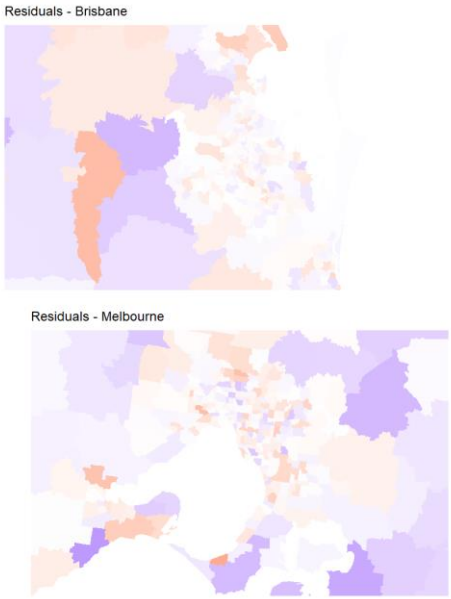
SEIFA dissimilarity (binary weighting)

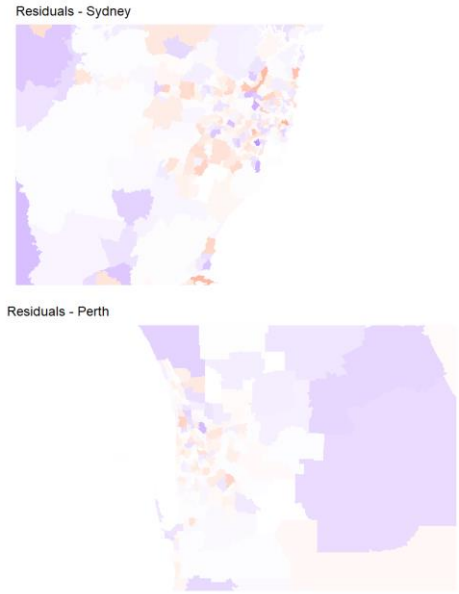




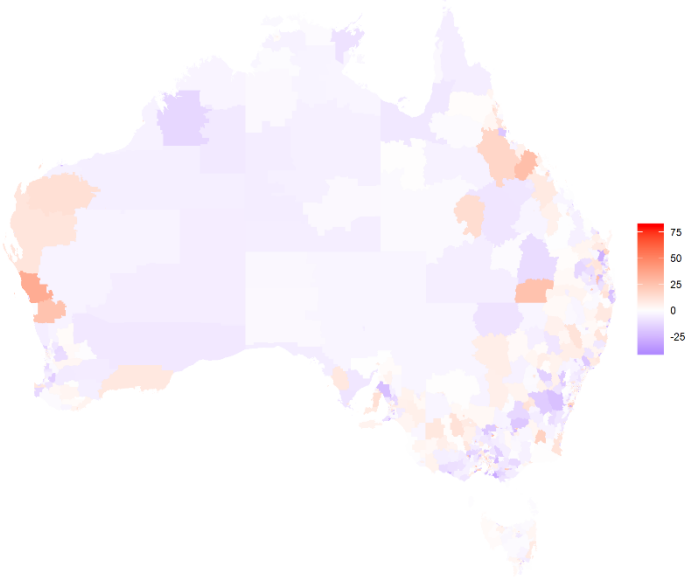
SEIFA dissimilarity (non-binary weighting)

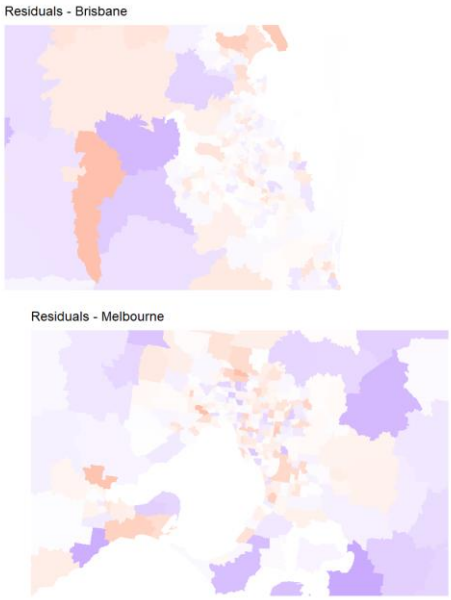


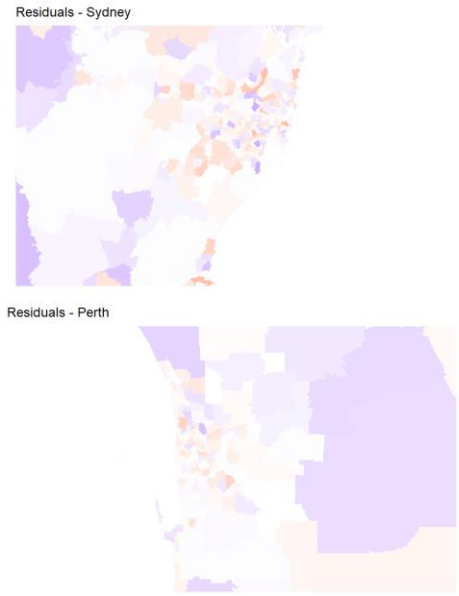




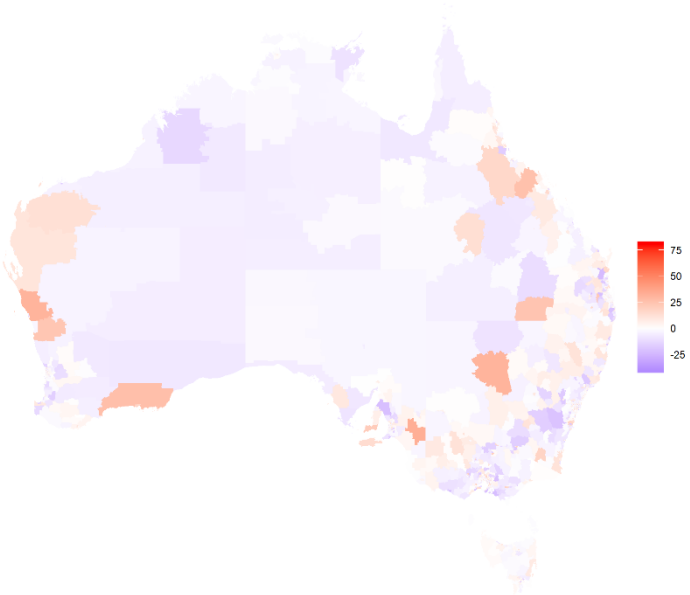
Residual dissimilarity (binary weighting)

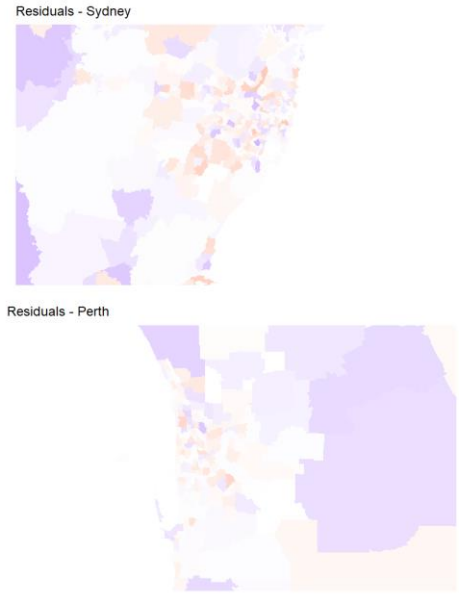




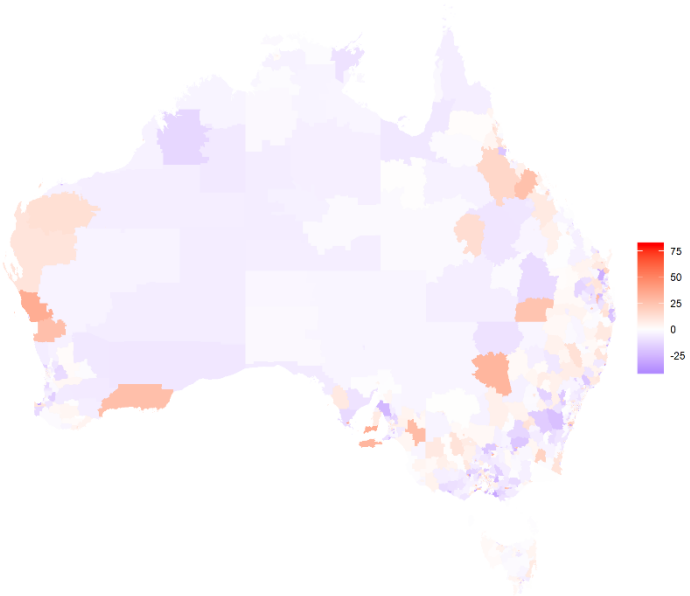


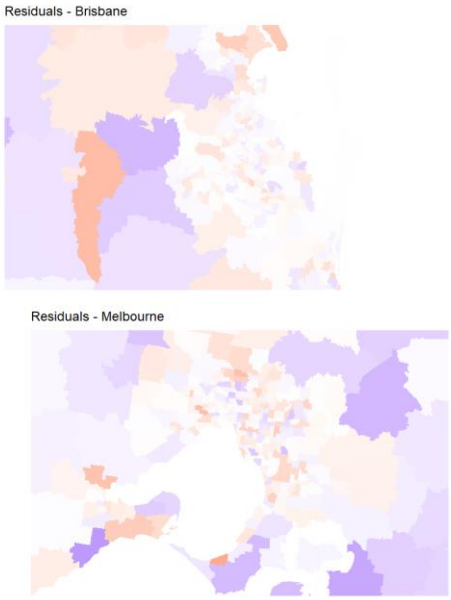
Residual dissimilarity (non-binary weighting)

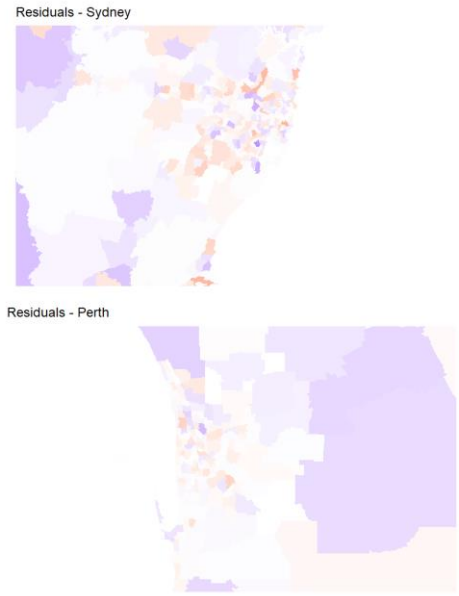




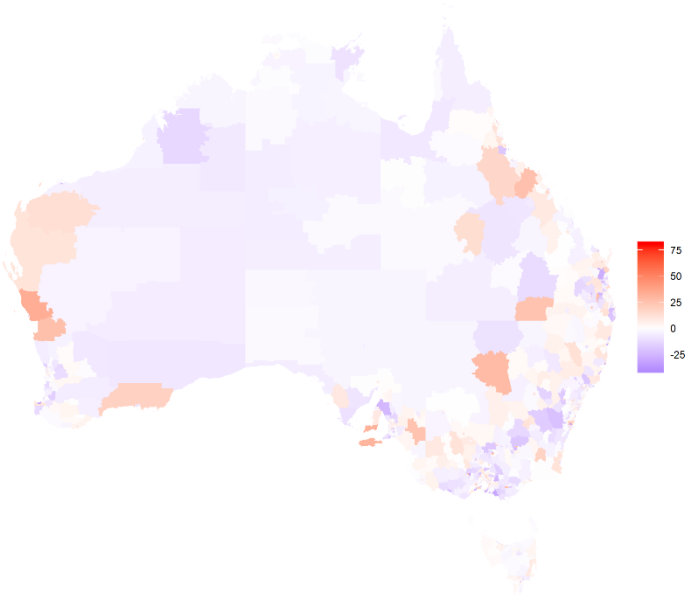
Localised autocorrelation (G=3)



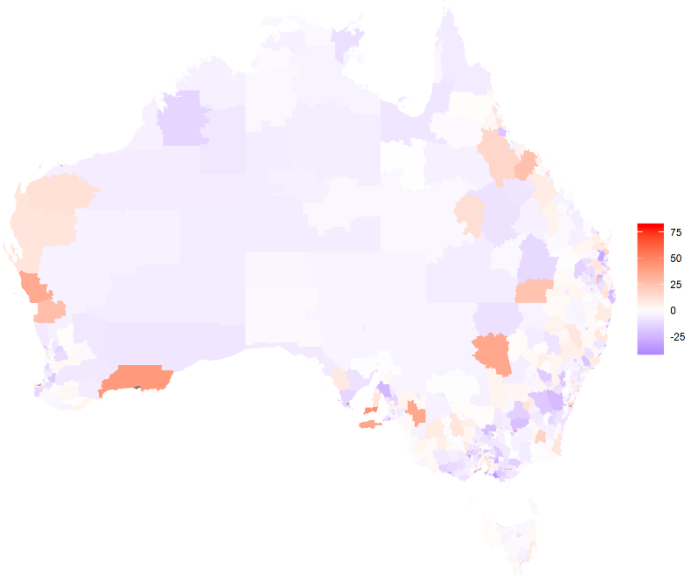




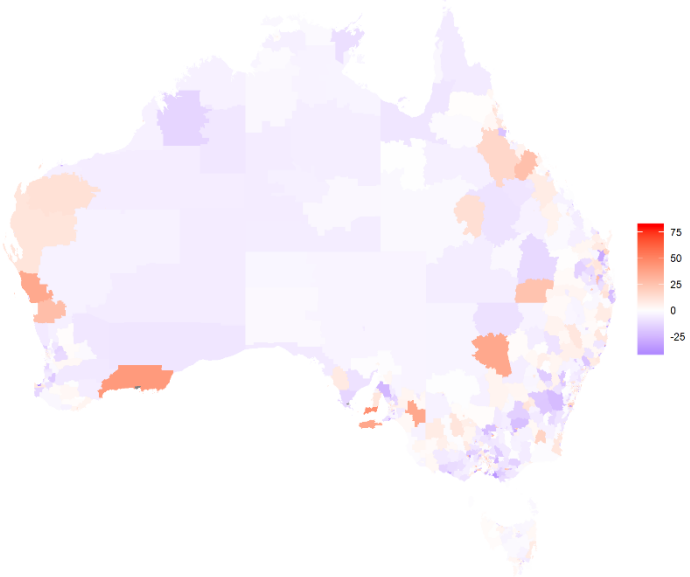
Localised autocorrelation (G=5)

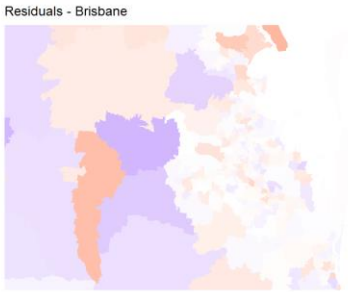


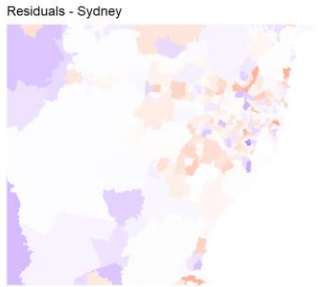
Locally adaptive (rho is determined when modelling)

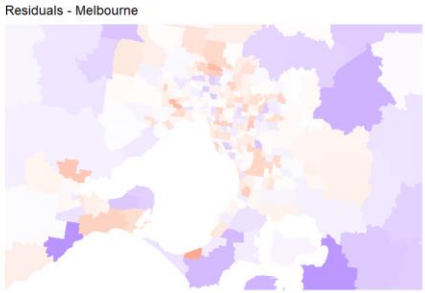


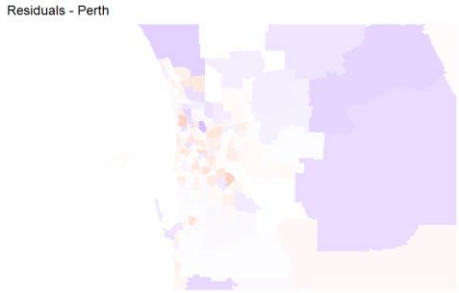
Locally adaptive (rho is fixed at 0.99)



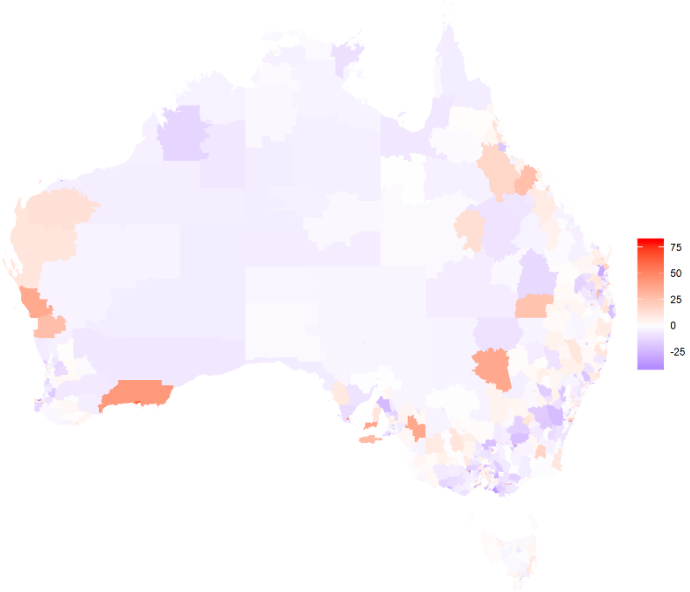


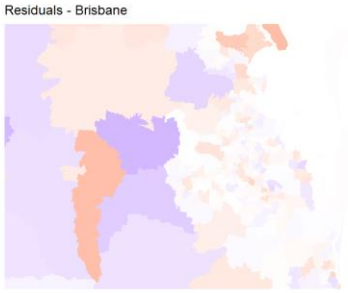


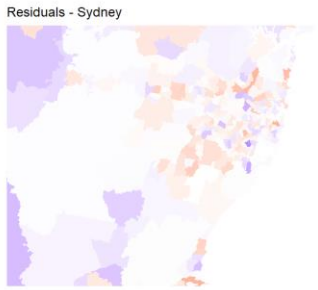


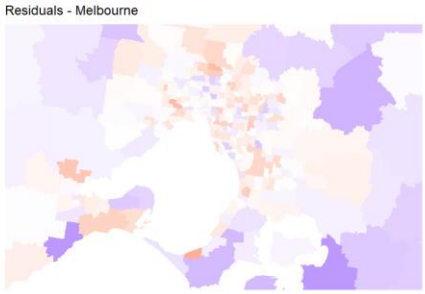


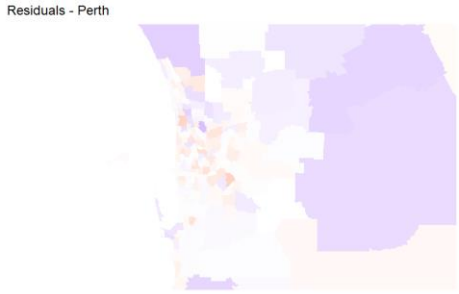
Weighted sum of spatial priors



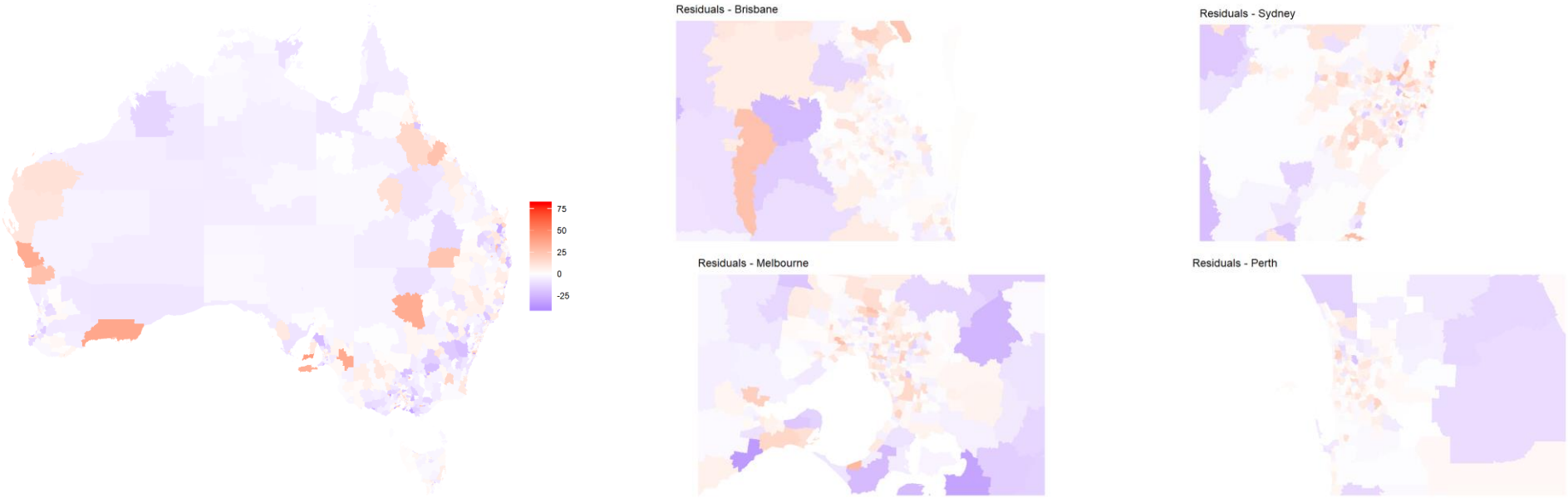








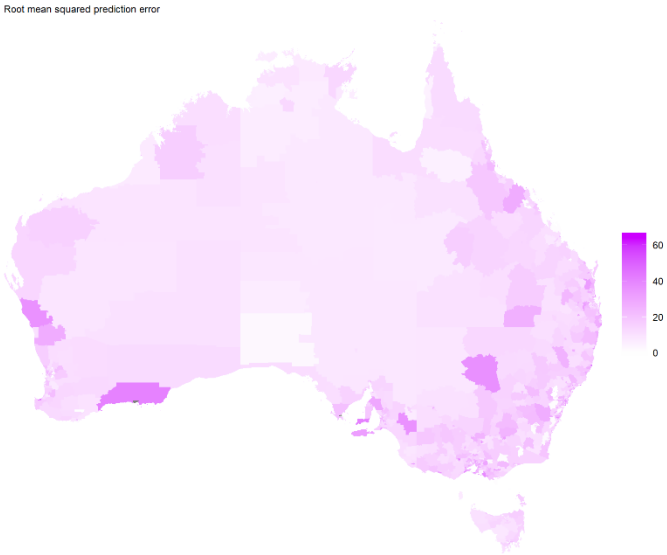
Leroux scale mixture model



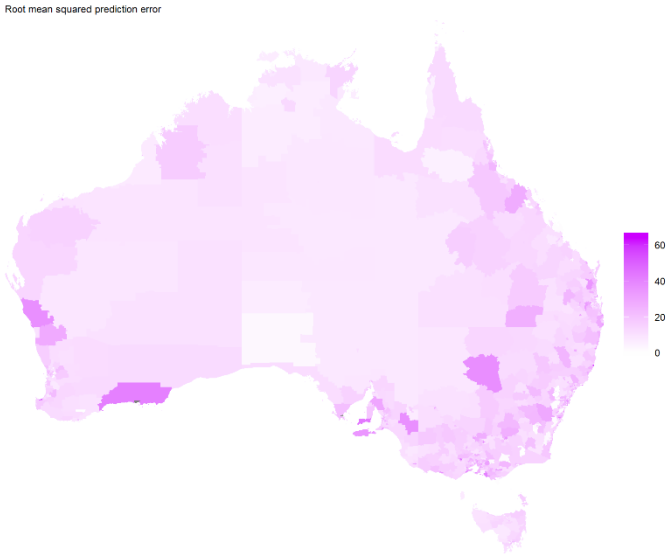
Note: The legend range was-44.8 to 80, and designed to enable differences between models to be visible. Areas shaded grey are outside this range.

All invasive cancers, females, RMSPE

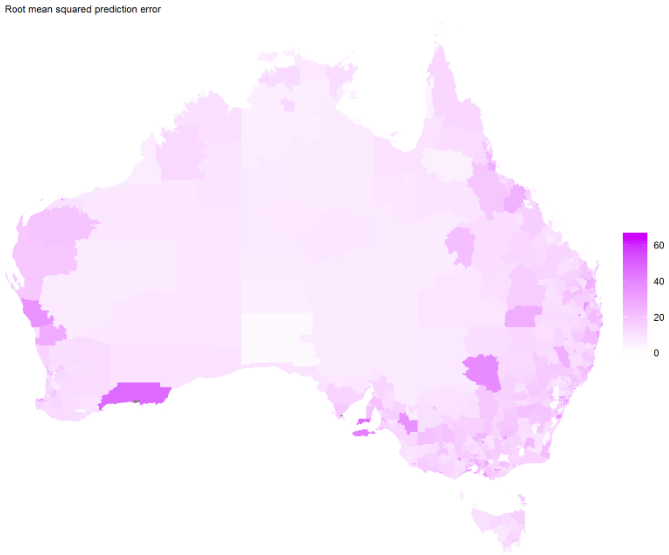
BYM



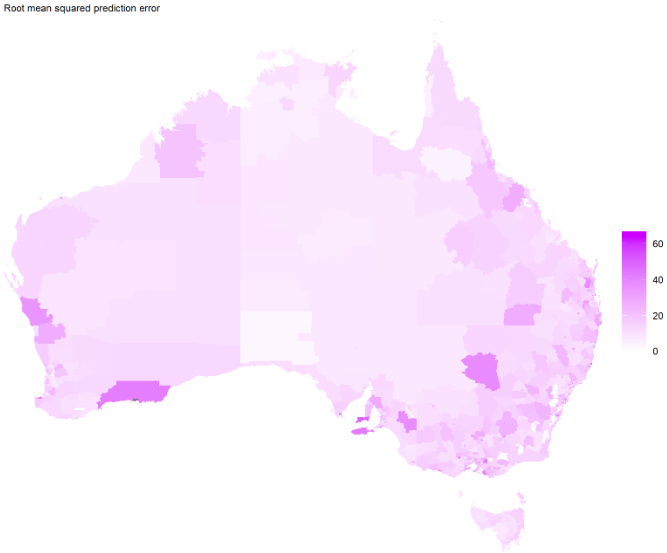
Leroux



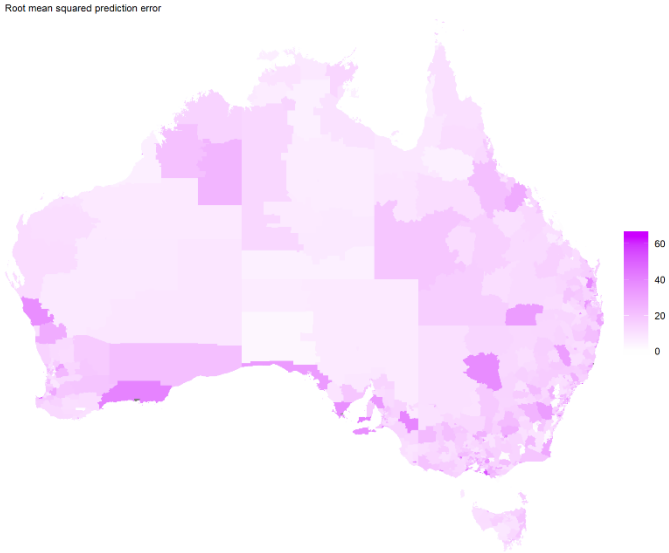
Geostatistical



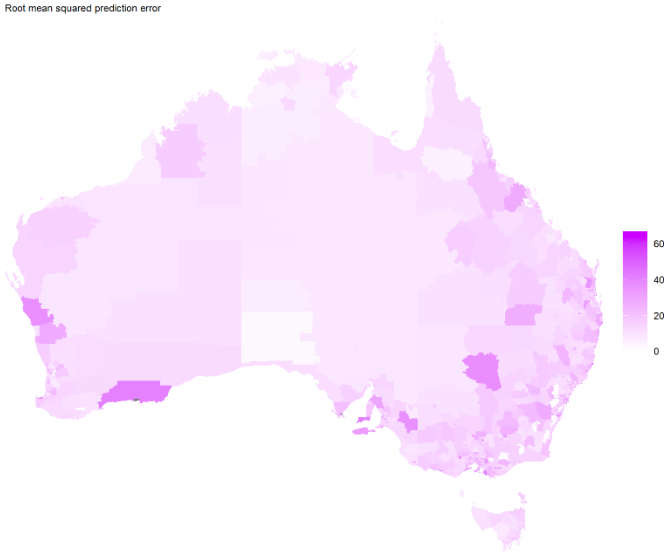
P-spline (tensor)



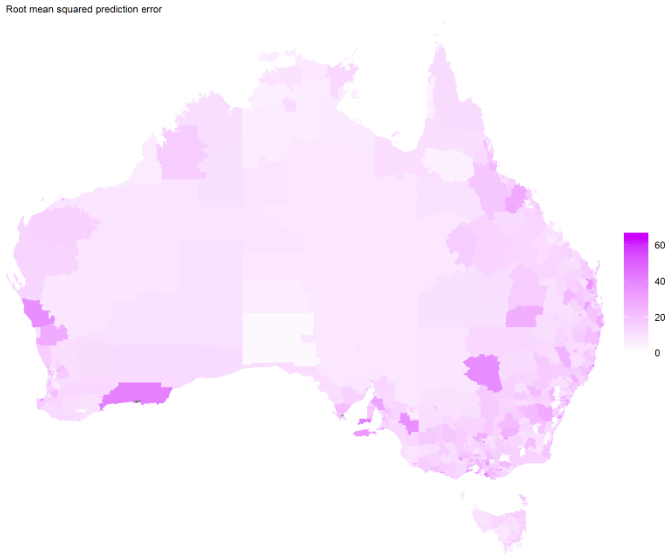
P-spline (radial)



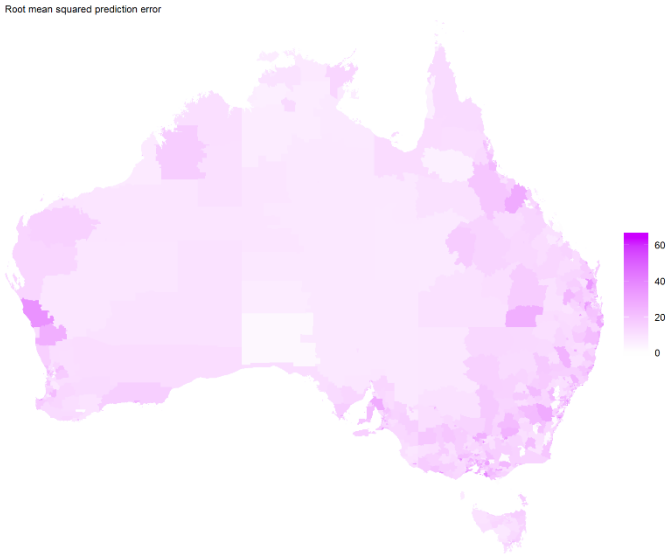
SEIFA dissimilarity (binary weighting)



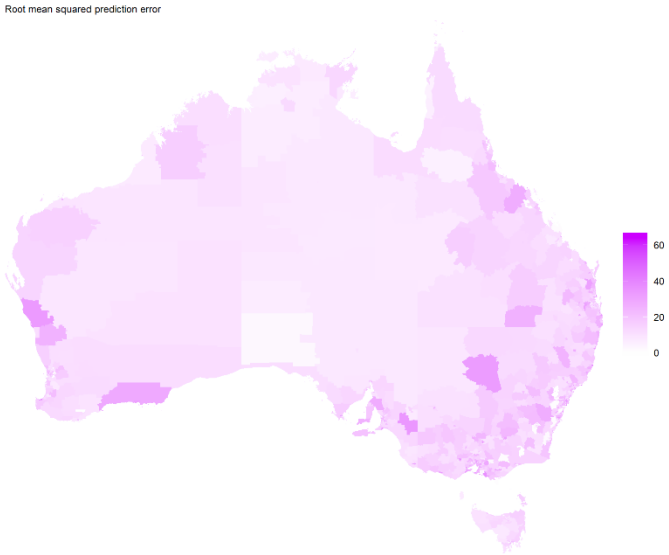
SEIFA dissimilarity (non-binary weighting)



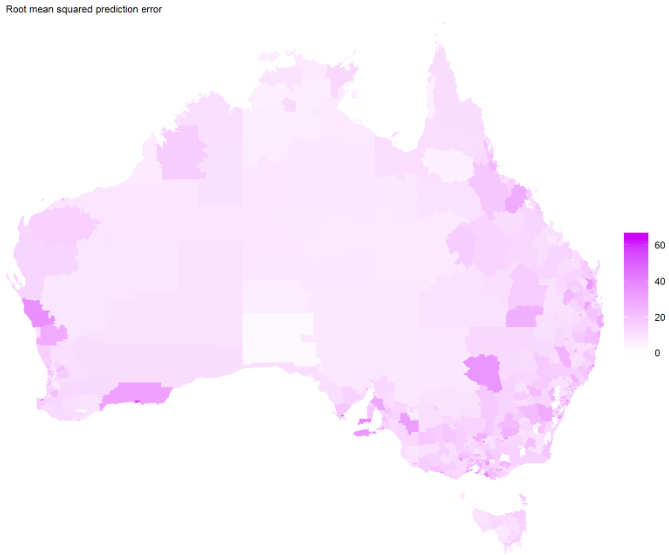
Residual dissimilarity (binary weighting)



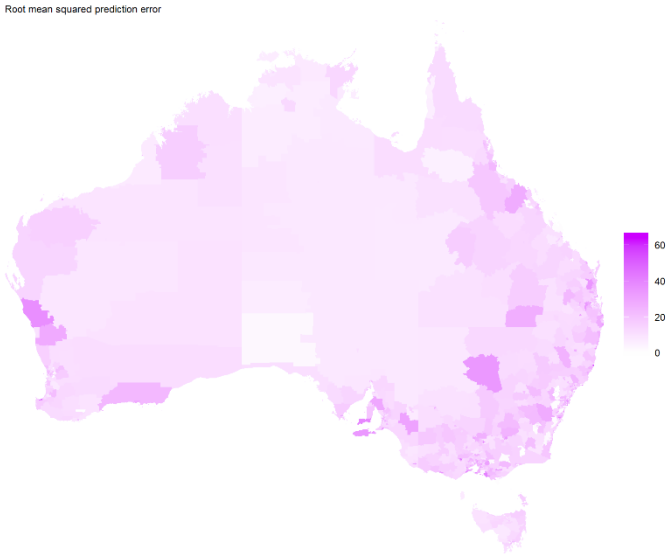
Residual dissimilarity (non-binary weighting)



Localised autocorrelation (G=3)



Localised autocorrelation (G=5)



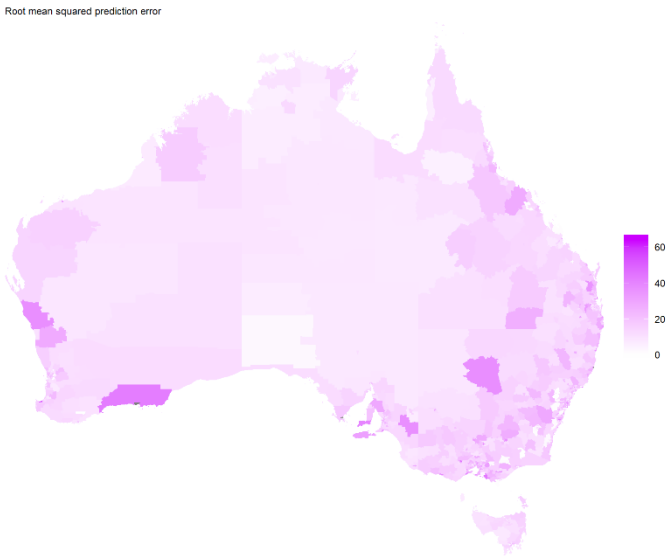
Locally adaptive (rho is determined when modelling)

[Unavailable from INLA output]

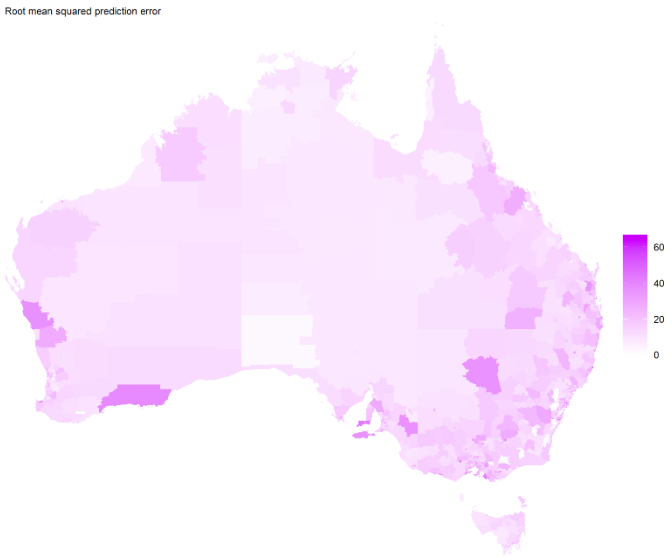
Locally adaptive (rho is fixed at 0.99)

[Unavailable from INLA output]

Weighted sum of spatial priors



Leroux scale mixture model



Note: The maximum legend value was set to 64. Areas shaded grey are higher than this value.