Bayesian Spatiotemporal Varying Coefficients Model to the Mortality Rate of Ischemic Stroke in Taiwan

Kuo-Jung Lee

Department of Statistics, NCKU

March 12, 2020

Motivation

- Stroke remains a major global health problem and its significance is likely to increase in the future due to ongoing demographic changes, including aging of the population and health transitions observed in Taiwan.
- A Bayesian spatiotemporal generalized linear regression with varying-coefficient model is proposed to study the association of the relevant risk factors:
  - comorbidities;
  - medication use; and
  - environmental and social factors.
- To provide more effective medical service behaviors and improve the distribution of medical resources, thus preventing the second stroke and reduce mortality after stroke.

Mortality

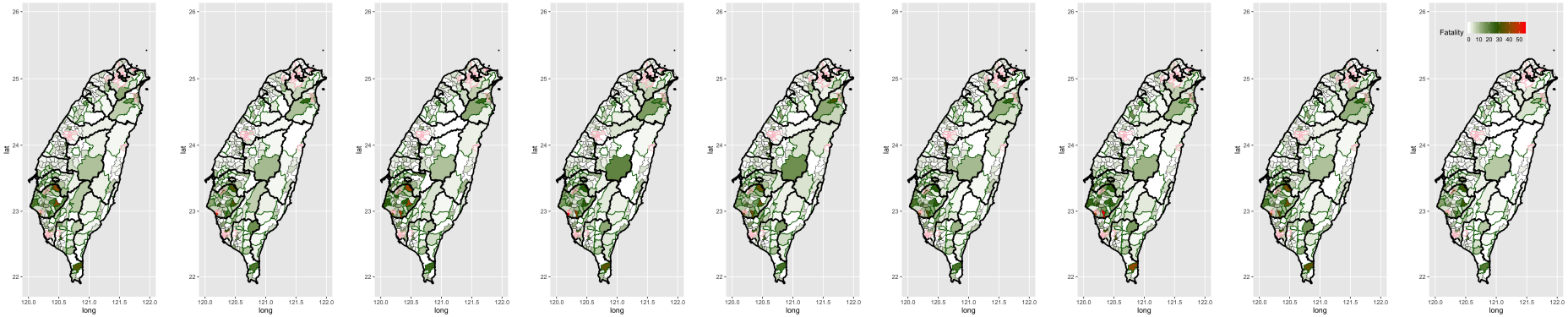


Figure 1: The distribution of fatality in ischemic stroke within 1 year for 349 townships in Taiwan from 2004 to 2012. The urbanized areas are surrounded in pink lines.

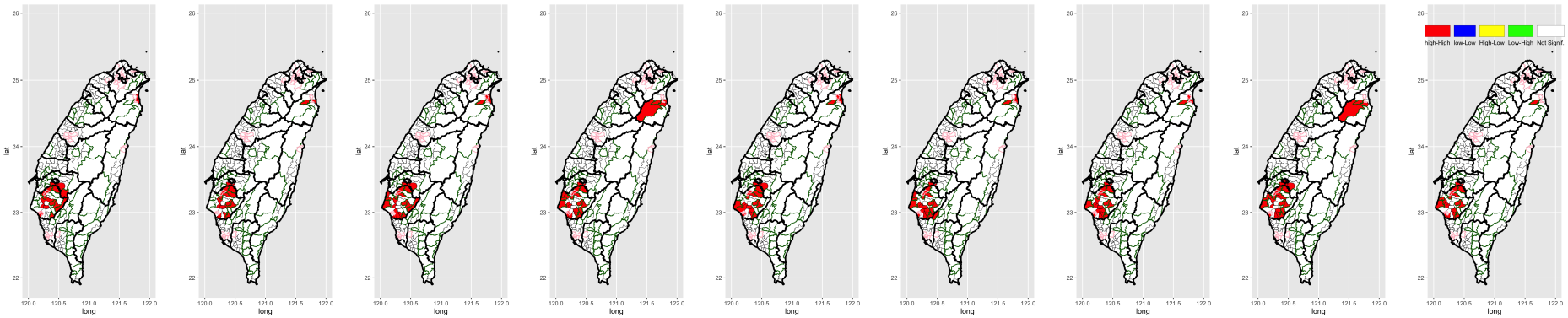


Figure 2: LISA cluster map of the fatality in ischemic stroke within 1 year for 349 townships in Taiwan from 2004 to 2012.

Mortality Rate

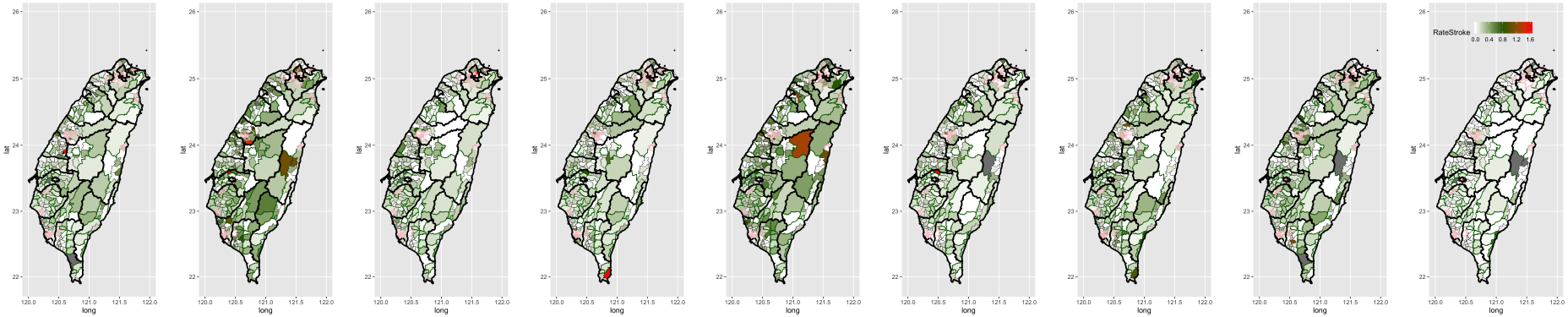


Figure 3: The distribution of fatality rate in ischemic stroke for 349 townships in Taiwan from 2004 to 2012.

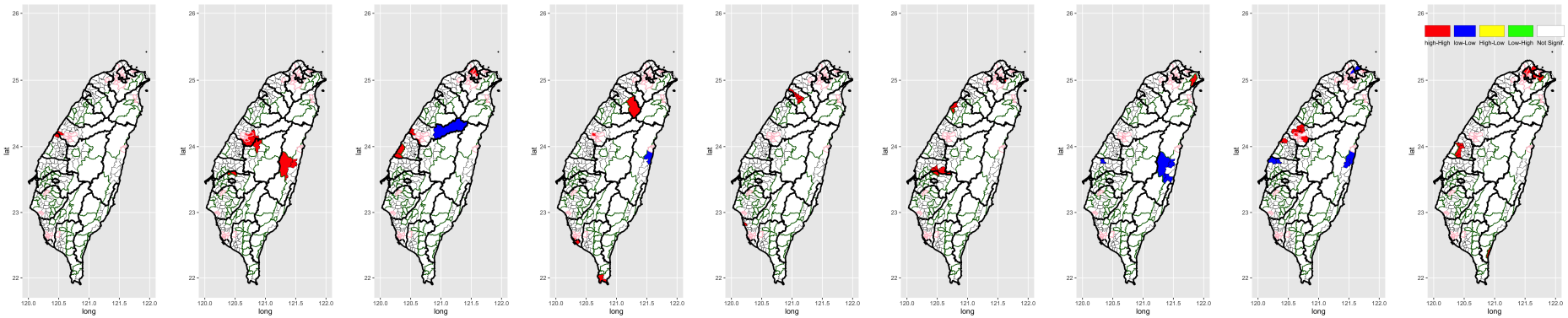


Figure 4: LISA cluster map of fatality rate in ischemic stroke for 349 townships in Taiwan from 2004 to 2012.

Statistical Modeling

Let

Y_{it}: observed number of deaths

E_{it}: expected number of deaths

caused by the ischemic stroke for town i at time point t within one year.

Assume Y_{it} follows a Poisson distribution

$$Y_{it} \sim \text{Poi}\left(E_{it} e^{\eta_{it}}\right), i = 1, \ldots, n, t = 1, \ldots, T.$$

We define the logarithm of the relative risk as

$$\eta_{it} = x_{it}' \beta + \phi_{it} + \psi_t,$$

where $x_{it} = (x_{it1}, \ldots, x_{itp})'$ and $\beta = (\beta_1, \ldots, \beta_p)'$ is the corresponding coefficient vector. We let $E_{it} = n_{it} \hat{p}_i$, where n_{it} is the number of observations at area i and time point t and $\hat{p}_i$ is the estimated incidence rate at area i.

變數說明

種類	項目	說明(括號內爲該項目之單位)
共病症		該城市首次中風病患送至醫院時合併有該共病症之病患比例
	AF	心房顫動(%)
	CHF	心衰竭(%)
	CKD	慢性腎臟病(%)
	DM	糖尿病(%)
	HTN	高血壓(%)
	IHD	缺血性心臟病(%)
	lipid	血脂異常(%)
病發後用藥		該城市首次中風病患病發後每人每日平均使用藥品定義日劑量
	beta	β — blocker, β 受體阻斷劑(DDD)
	ccb	Calcium channel blockers, 鈣離子通道阻斷劑(DDD)
	diu	Diuretics, 利尿劑(DDD)
	acei	ACEIs/ARBs, 降血壓藥物(DDD)
	coa	Anticoagulants, 抗凝血藥物(DDD)
	pla	Antiplatelet agents, 抗血小板藥物(DDD)
	statin	Statins, 降血脂劑(DDD)
	TPA	rt-PA, 合成組織胞漿素原活化劑(DDD)
環境社會因子		該城市內之各項環境與社會因子
	pm25	平均細懸浮微粒(PM 2.5) 濃度($\mu g/m^3$)
	copd	抽菸人口比例，以慢性肺阻塞病盛行率代表其抽菸比率(%)
	edu	15 歲以上受大專以上(含肄業) 教育人口比例(%)
	bed	總病床數占全國總病床數之比例(%)
	income	公勞保投保人之平均薪資(千元) (%)
	old	65 歲以上人口比例(%)
	ori	原住民人口比例(%)
	sex	男性人口比例(%)
	ssi	首次缺血性中風病患之平均中風嚴重度參數(SSI)

Temporal Dependency

Temporal Dependence: AR

Assume $\psi_t \sim N\left(\zeta\psi_{t-1}, \sigma_\psi^2\right)$ and $\zeta \sim U(-1,1)$ and $\sigma_\psi^2 \sim \mathcal{IG}\left(\frac{a_\psi}{2}, \frac{b_\psi}{2}\right)$.

Spatial Dependence: MCAR

Let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)'$ be the vector of spatial random effects.

$$\boldsymbol{\phi}|\lambda, \rho \sim \mathcal{N}_I \left(\mathbf{0}, \lambda \boldsymbol{Q}(\boldsymbol{W}, \rho)^{-1} \right), \tag{1}$$

where

$$\boldsymbol{Q}(\boldsymbol{W}, \rho) = \rho(\mathrm{diag}(\boldsymbol{W}\mathbf{1}) - \boldsymbol{W}) + (1 - \rho)\boldsymbol{I}.$$

$\boldsymbol{W}$ is the proximity matrix with the $(i,\ j)$th element $w_{ij} = 1$ if i and j are neighbor; otherwise $w_{ij} = 0$, and $\boldsymbol{I}$ is an $I \times I$ identity matrix. More precisely,

$$\phi_i|\boldsymbol{\phi}_{-i}, \boldsymbol{W}, \lambda, \rho \sim \mathcal{N}\left(\frac{\rho \sum_{i' \sim i} w_{i'i}\phi_{i'}}{\rho \sum_{i' \in \sim i} w_{i'i} + 1 - \rho}, \frac{\lambda}{(\rho \sum_{i' \in \sim i} w_{i'i} + 1 - \rho)} \right),$$

In addition,

$$\rho \sim U(0, 1) \quad \text{and} \quad \lambda \sim \mathcal{IG}\left(\frac{a_\lambda}{2}, \frac{b_\lambda}{2} \right).$$

Results

Type	Variable	Estimate	95%Credible Interval	
			2.5%	97.5%
Comorbidities	AF	0.0002	-0.0042	0.0047
	CHF	0.0036	-0.0002	0.0074
	CKD	0.0076	0.0024	0.0134
	DM	0.0013	-0.0014	0.0040
	HTN	0.0037	0.0020	0.0070
	IHD	0.0033	0.0005	0.0059
	lipid	-0.0056	-0.0083	-0.0030
Medication Use	**beta**	-0.80	-1.31	-0.31
	ccb	-0.03	-0.22	0.15
	diu	0.38	0.09	0.66
	acei	-0.16	-0.32	-0.01
	coa	-0.94	-2.15	0.28
	pla	-0.14	-0.30	0.04
	statin	0.12	-0.11	0.34
	TPA	-2.63	-3.93	-1.38
Environmental & Social Factor	pm25	0.0016	-0.0009	0.0041
	copd	0.0037	-0.0018	0.0092
	edu	-0.0013	-0.0062	0.0035
	bed	0.0584	0.0332	0.0834
	income	0.0049	-0.0007	0.0114
	old	-0.0008	-0.0020	0.0079
	ori	-0.3512	-0.5634	-0.1433
	sex	0.0153	-0.0161	0.0472
	ssi	0.1543	0.1321	0.1822

However

- In reality, the variable may have different effects on the mortality in different areas and time points.

- In practice, there may exist the similar spatial patterns for the time-varying coefficients not only in local but also in remote areas.

Statistical Modeling

Let

Y_{it}: observed number of deaths

E_{it}: expected number of deaths

caused by the ischemic stroke for town i at time point t.

Assume Y_{it} follows a Poisson distribution

$$Y_{it} \sim \text{Poi}\left(E_{it} e^{\eta_{it}}\right), i = 1, \ldots, n, t = 1, \ldots, T.$$

We define the logarithm of the relative risk as

$$\eta_{it} = x_{it}' \beta_{it} + \phi_i + \psi_t$$

where $x_{it} = (x_{it1}, \ldots, x_{itp})'$ and $\beta_{it} = (\beta_{it1}, \ldots, \beta_{itp})'$ is the corresponding time-varying coefficient vector.

Spatial and Temporal Dependencies

Spatial Dependence: MCAR

Let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)'$ be the vector of spatial random effects.

$$\boldsymbol{\phi}|\lambda, \rho \sim \mathcal{N}_I \left(\mathbf{0}, \lambda \boldsymbol{Q}(\boldsymbol{W}_S, \rho)^{-1} \right),$$

where

$$\boldsymbol{Q}(\boldsymbol{W}_S, \rho) = \rho(\mathrm{diag}(\boldsymbol{W}_S \mathbf{1}) - \boldsymbol{W}_S) + (1 - \rho)\boldsymbol{I}_S.$$

$\boldsymbol{W}_S$ is the proximity matrix with the $(i,\ j)$th element $w_{ij} = 1$ if i and j are neighbor; otherwise $w_{ij} = 0$, and $\boldsymbol{I}_S$ is an $I \times I$ identical matrix. In addition,

$$\rho \sim U(0, 1) \quad \text{and} \quad \lambda \sim \mathcal{IG}\left(\frac{a_\lambda}{2}, \frac{b_\lambda}{2} \right).$$

Temporal Dependence: AR

Assume $\psi_t \sim N\left(\zeta\psi_{t-1}, \sigma^2_\psi\right)$ and $\zeta \sim U(-1, 1)$ and $\sigma^2_\psi \sim \mathcal{IG}\left(\frac{a_\psi}{2}, \frac{b_\psi}{2}\right)$.

Variable Selection

- One important goal in this study is to detect which variables play an crucial role in affecting the mortality in stroke, that is which $\beta_{it} \neq 0$.
- A vector of categorical random variables, $\boldsymbol{\gamma}_t = (\gamma_{11}, \ldots, \gamma_{nt})^T$ to indicate whether β_{it} equals to 0 or not, that is,

$$\beta_{it} \begin{cases} = 0 & \gamma_{it} = 0; \\ \neq 0 & \gamma_{it} \neq 0. \end{cases}$$

- The time-varying regression coefficients β_{it} follow

$$\beta_{it}|\gamma_{it} \sim \prod_{g_t=1}^{G_t-1} \left[N\left(\alpha_{g_t}, \sigma_{g_t}^2\right)\right]^{\mathrm{I}\{\gamma_{it}=g_t\}} \delta_0^{\mathrm{I}\{\gamma_{it}=G_t\}},$$

where δ_0 is the point mass at zero and $P(\gamma_{it} = g_t) = \pi(\gamma_{it}|\theta_t) = \tau_{igt}$. We assume $\sigma_{g_t}^2 \sim \mathcal{IG}\left(\frac{a_\sigma}{2}, \frac{b_\sigma}{2}\right)$.

Remote & Local Patterns

Remote Pattern: Mixture Model

The mixture part $\sum_{g_t=1}^{G_t} \tau_{gt} N\left(\alpha_{g_t}, \sigma_{g_t}^2\right)$ is used to describe the remote areas that show similar patterns to a variable.

Local Pattern: Potts Model

The Potts model for γ_t to account for the neighboring areas that behave similarly:

$$\pi(\gamma_t|\theta_{jt}) \propto \exp\left\{\sum_{i=1}^{n}\sum_{g_t=1}^{G_t} \kappa_{g_t} \mathrm{I}\left\{\gamma_{it} = g_t\right\} + \theta_t \sum_{i\sim i'} u_{i,i'} \mathrm{I}(\gamma_{it} = \gamma_{i't})\right\},$$

where

- $u_{i,i'}$: the wights of the interaction between neighboring locations.
- θ_t: the strength of the interaction between any two areas at time t.
- $\sum_{i,g} \kappa_{g_t} \mathrm{I}\left\{\gamma_{it} = g_t\right\}$: the "external field" to incorporate the prior info.

Posterior & MCMC

- We do a component-wise MCMC approach for model estimation.
- But some of these requires Metropolis-Hastings-within-Gibbs.

Update β and γ

- Update of β and γ simultaneously.
- The proposal distribution of β and γ is the joint prior distribution $q(\beta, \gamma|\theta, \alpha, \sigma)$.
- Assume β^* and γ^* are the candidate and and then accept the proposed value with probability $\min(1, r_{\beta,\gamma})$ and the Hastings ratio is

$$r_{\beta,\gamma} = \frac{p(\beta^*, \gamma^*|y)}{p(\beta, \gamma|y)} \frac{q(\beta, \gamma)}{q(\beta^*, \gamma^*)} = \frac{p(y|\beta^*, \gamma^*)q(\beta^*, \gamma^*)}{p(y|\beta, z)q(\beta, \gamma)} \frac{q(\beta, \gamma)}{q(\beta^*, \gamma^*)} = \frac{p(y|\beta^*, \gamma^*)}{p(y|\beta, \gamma)}.$$

Update θ: Spatial Dependence

$$p(\theta_t) = \frac{1}{Z(\theta_t, \kappa_{g_t})} \exp\left\{ \sum_{i=1}^{n} \sum_{g_t=1}^{G_t} \kappa_{g_t} \mathtt{I}\left\{\gamma_{it} = g_t\right\} + \theta_t \sum_{i \sim i'} u_{i,i'} I(\gamma_{it} = \gamma_{i't}) \right\}$$

where

$$Z(\theta_t, \kappa_{g_t}) = \sum_{\gamma} \exp\left\{ \sum_{i=1}^{n} \sum_{g_t=1}^{G_t} \kappa_{g_t} \mathtt{I}\left\{\gamma_{it} = g_t\right\} + \theta_t \sum_{i \sim i'} u_{i,i'} I(\gamma_{it} = \gamma_{i't}) \right\} .$$

Update θ (Cont'd)

Then generate a proposal θ_t^* from a proposal density $q(\theta)$.

Then set $\theta = \theta^*$ with probability the minimum of 1 and the Hastings ratio

$$\frac{Z(\theta_t, \kappa_{g_t})}{Z(\theta_t^*, \kappa_{g_t})} \frac{\exp\left\{\sum_{i=1}^{n} \sum_{g_t=1}^{G_t} \kappa_{g_t} \mathrm{I}\left\{\gamma_{it} = g_t\right\} + \theta_t^* \sum_{i \sim i'} u_{i,i'} I(\gamma_{it} = \gamma_{i't})\right\}}{\exp\left\{\sum_{i=1}^{n} \sum_{g_t=1}^{G_t} \kappa_{g_t} \mathrm{I}\left\{\gamma_{it} = g_t\right\} + \theta_t \sum_{i \sim i'} u_{i,i'} I(\gamma_{it} = \gamma_{i't})\right\}} \frac{q(\theta)}{q(\theta^*)}.$$

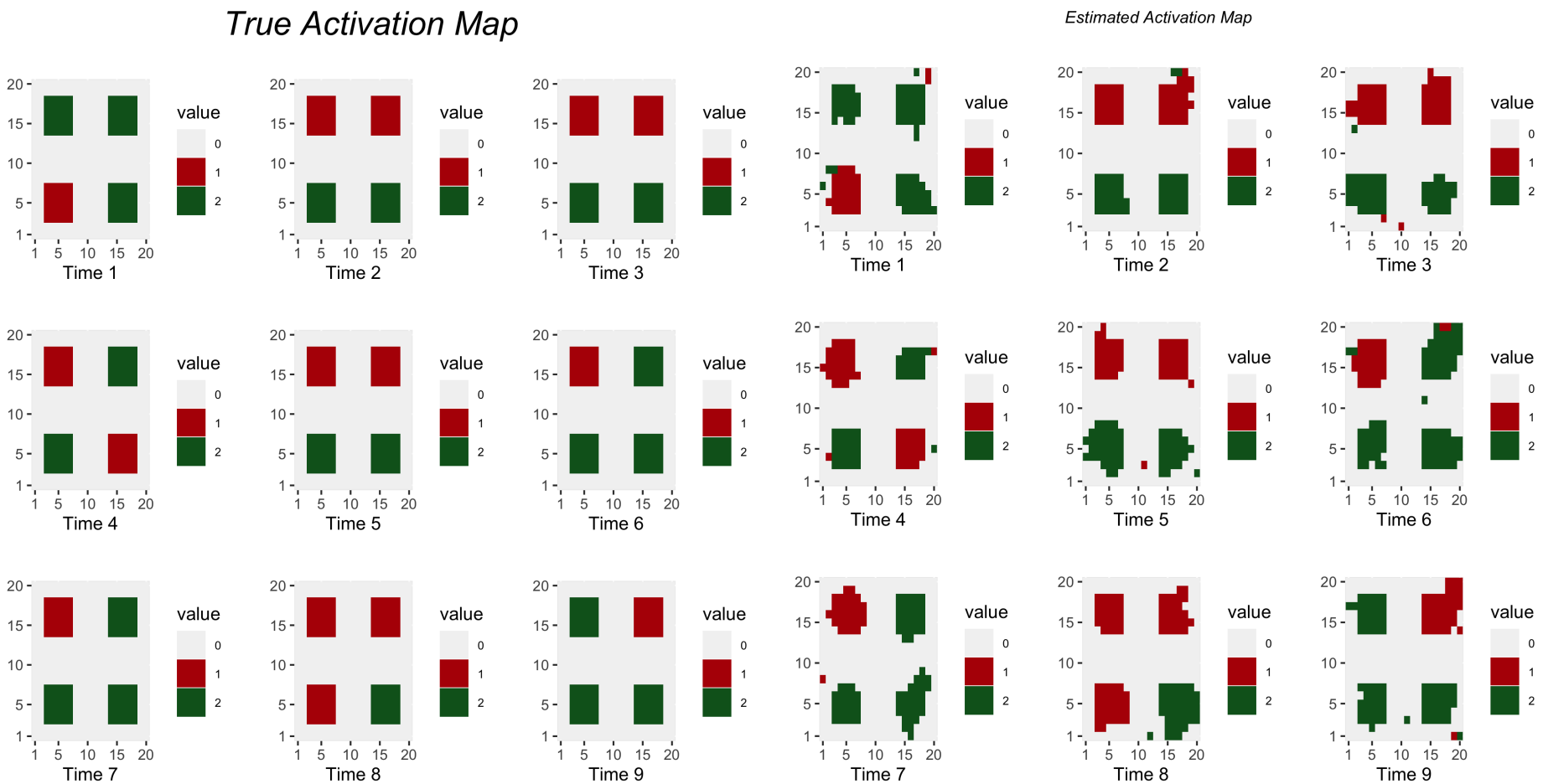
The ratio

$$\frac{Z(\theta, \kappa_{g_t})}{Z(\theta^*, \kappa_{g_t})}$$

is analytically intractable but can be estimated with path sampling.

Simulation Study: Setting 1

- We simulated data at all $n = 400$ areas on the 20×20 image for $t = 9$ time points with $p = 2$ covariates.
- For ease illustration, we consider $G_{jt} = 2$ and use 1 and 2 to denote the active group membership and 0 for inactive group.



Simulation Study: Setting 2

- We set $\alpha_{itg} = 1$ when $G_{tj} = 1$ and $\alpha_{itg} = -1$ when $G_{tj} = 2$.
- We simulate $\sigma^2_{itg} \sim \Gamma(1, 1)$ and $\beta_{itg} \sim N(\alpha_{itg}, \sigma^2_{itg})$.
- Given $\lambda \sim \Gamma(1, 1)$ and $\rho \sim U(0, 1)$, we can compuate $\boldsymbol{Q}(\boldsymbol{W}, \rho)$ and $\boldsymbol{\phi}|\lambda, \rho \sim \mathcal{N}_I\left(\mathbf{0}, \lambda \boldsymbol{Q}(\boldsymbol{W}, \rho)^{-1}\right)$.
- We simulate ψ from AR(1) with $\zeta = 0.5$.
- Then generate $Y_{it} \sim \text{Poi}\left(E_{it} e^{\eta_{it}}\right), i = 1, \ldots, n, t = 1, \ldots, T$, where $\eta_{it} = x'_{it}\beta_{it} + \phi_i + \psi_t$.

Simulation Study: Result

We calculate the accuracy of classifications of activations (ACA) for each time point by

$$\text{ACA}_t = \frac{1}{N}\sum_{i=1}^{N} \text{I}(\gamma_{it} = \hat{\gamma}_{it}).$$

where the assignment of an observation to a group is based on

$$\hat{\gamma}_{it} = \underset{g \in 1\ldots, G_t}{\text{argmax}} \left\{ p(\gamma_{it} = g | y) \right\}.$$

Table 1: The accuracy of classification of (in)active variables.

	Year								
ACA	2004	2005	2006	2007	2008	2009	2010	2011	2012
1st Q	0.86	0.85	0.84	0.82	0.85	0.83	0.83	0.84	0.85
median	0.92	0.90	0.92	0.90	0.91	0.88	0.88	0.89	0.90
3rd Q	0.94	0.93	0.94	0.93	0.93	0.92	0.92	0.91	0.94

Simulation Study: Model Comparison

In order to investigate the performance of the proposed model, we compared the proposed model with several competing models. The models under consideration are as follows:

1. Model 1: simple log-linear Poisson model with spatio-temporally constant coefficients ($\beta_{itp} = \beta_{ip}$) for **one**-component models.
2. Model 2: simple log-linear Poisson model with spatio-temporally constant coefficients ($\beta_{itp} = \beta_{ip}$) for **two**-component models.
3. Model 3: log-linear Poisson model with spatially **time varying** coefficients for **one**-component models.
4. Model 4: log-linear Poisson model with spatially **time varying** coefficients for **two**-component models.

Simulation Study: Model Comparison (Cont'd)

- The log marginal predictive likelihood (LMPL) is calculated through the log conditional predictive ordinate (CPO)

$$\text{LMPL} = \sum_i \sum_t \log \text{CPO}_{it} = \sum_i \sum_t \log f(y_{it}|y_{-it})$$

where CPO_{it} is the conditional predictive ordinate. A larger value of CPO indicates better prediction based on the model, and thus a model with a large value of LMPL implies better model of fit.

- Moreover, we calculate the mean square prediction error (MSPE) for the comparison of models in terms of prediction performance

$$\text{MSPE} = \frac{1}{nT} \sum_i \sum_t (Y_{it} - \hat{Y}_{it})^2,$$

where $\hat{Y}_{it}$ the predicted value for ith subject at time point t from the posterior predictive distribution.

Simulation Study: Model Comparison (Cont'd)

Table 2: Model comparison using AIC, BIC, DIC, MPL, and MSPE for the simulation study.

	Model			
Measurement	1	2	3	4
AIC	-2707	-5027	-3568	**-10219**
BIC	13293	13100	12433	**5782**
DIC	-28197	-30187	-29104	**-45781**
LMPL	720	541	744	**754**
MSPE	10.34	8.76	3.68	**2.91**

變數說明

種類	項目	說明(括號內爲該項目之單位)
共病症		該城市首次中風病患送至醫院時合併有該共病症之病患比例
	AF	心房顫動(%)
	CHF	心衰竭(%)
	CKD	慢性腎臟病(%)
	DM	糖尿病(%)
	HTN	高血壓(%)
	IHD	缺血性心臟病(%)
	lipid	血脂異常(%)
病發後用藥		該城市首次中風病患病發後每人每日平均使用藥品定義日劑量
	beta	**β − blocker, β 受體阻斷劑(DDD)**
	ccb	**Calcium channel blockers, 鈣離子通道阻斷劑(DDD)**
	diu	**Diuretics, 利尿劑(DDD)**
	acei	**ACEIs/ARBs, 降血壓藥物(DDD)**
	coa	**Anticoagulants, 抗凝血藥物(DDD)**
	pla	**Antiplatelet agents, 抗血小板藥物(DDD)**
	statin	**Statins, 降血脂劑(DDD)**
	TPA	**rt-PA, 合成組織胞漿素原活化劑(DDD)**
環境社會因子		該城市內之各項環境與社會因子
	pm25	平均細懸浮微粒(PM 2.5) 濃度($\mu g/m^3$)
	copd	抽菸人口比例，以慢性肺阻塞病盛行率代表其抽菸比率(%)
	edu	15 歲以上受大專以上(含肄業) 教育人口比例(%)
	bed	總病床數占全國總病床數之比例(%)
	income	公勞保投保人之平均薪資(千元) (%)
	old	65 歲以上人口比例(%)
	ori	**原住民人口比例(%)**
	sex	男性人口比例(%)
	ssi	首次缺血性中風病患之平均中風嚴重度參數(SSI)

Prediction

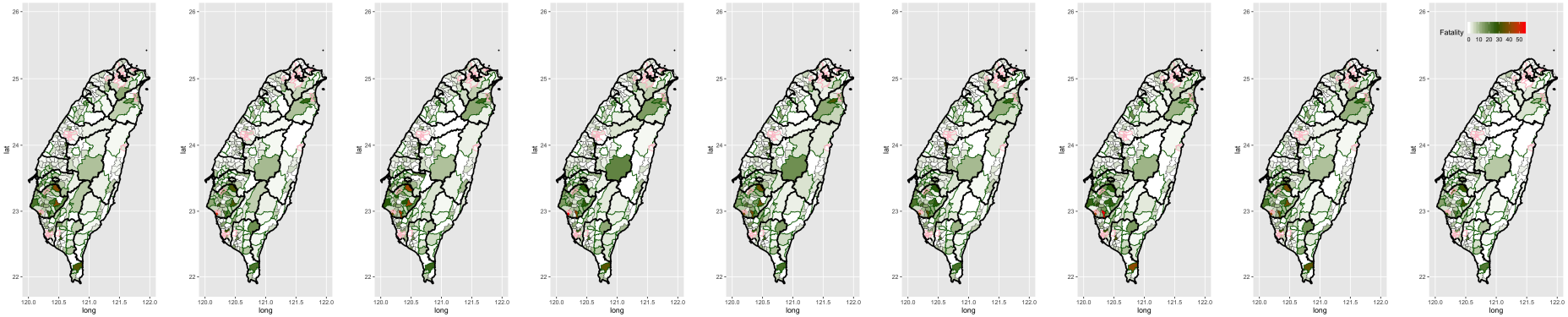


Figure 6: The distribution of fatality in ischemic stroke within 1 year for 349 townships in Taiwan from 2004 to 2012.

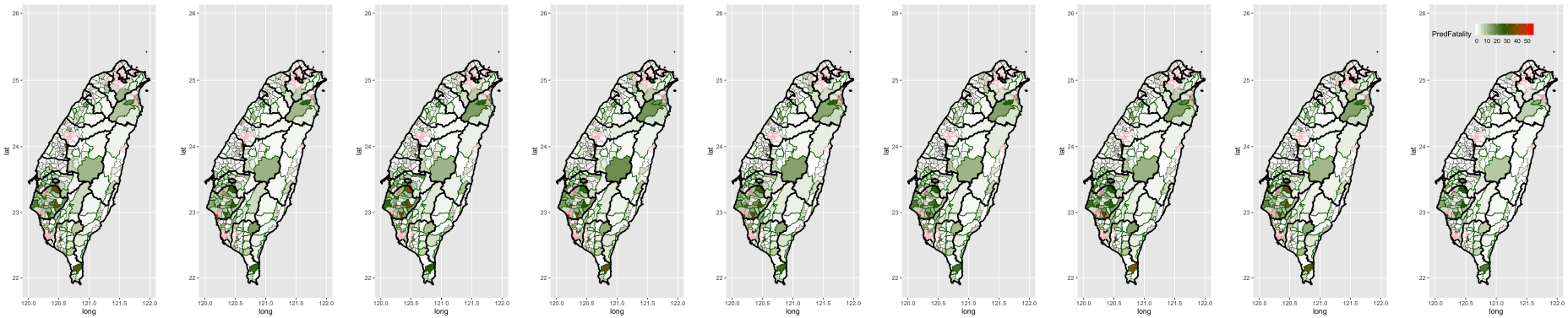
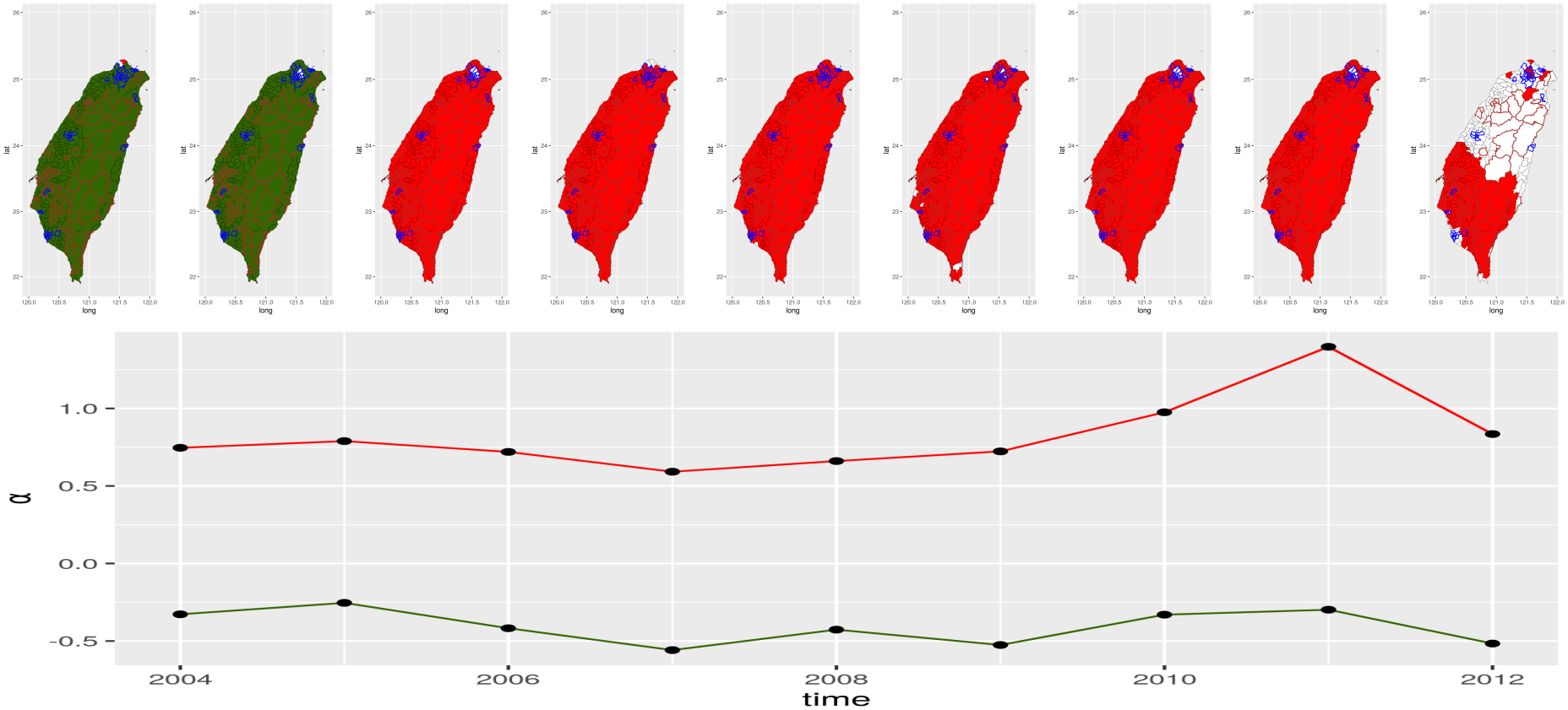


Figure 7: The distribution of predicted fatality in ischemic stroke within 1 year for 349 townships in Taiwan from 2004 to 2012.

降血脂劑(statin)



血栓溶解劑(rt-PA), 利尿劑(diu)

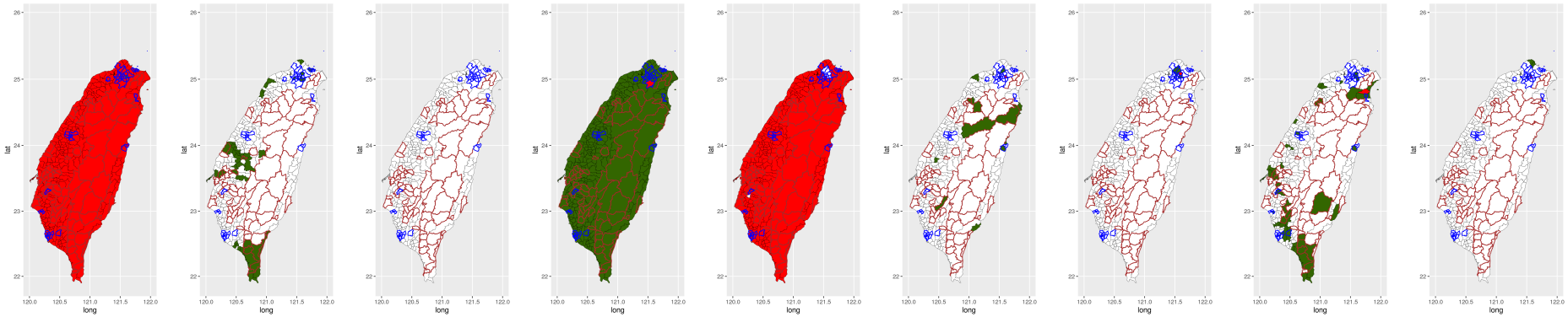


Figure 8: rt-PA. Red: Positive effect; Green: Negative Effect.

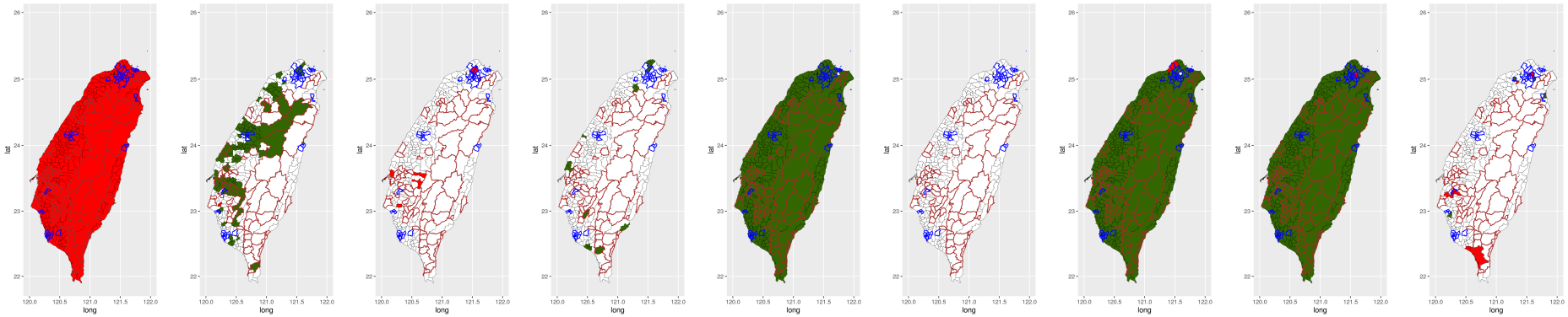


Figure 9: diu. Red: Positive effect; Green: Negative Effect.

抗血小板製劑(pla), 抗凝血劑(coa)

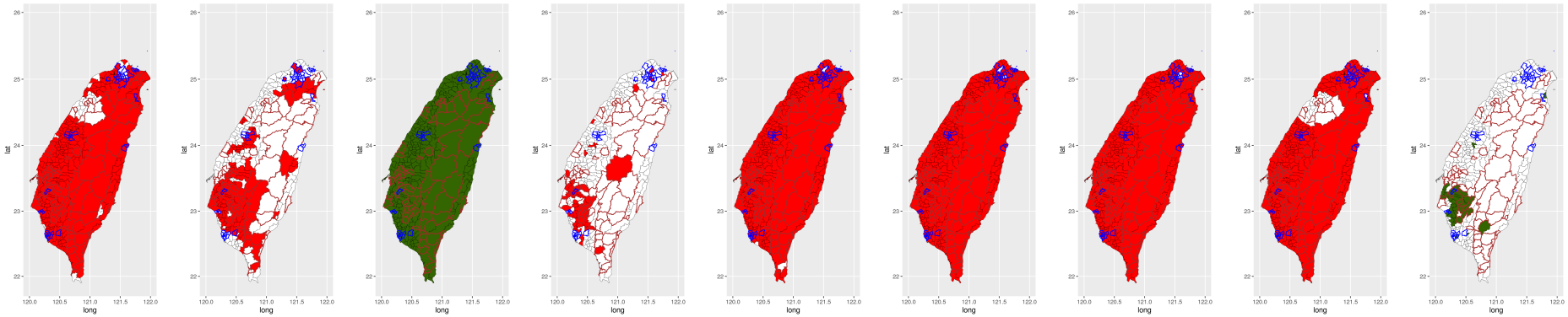


Figure 10: pla. Red: Positive effect; Green: Negative Effect.

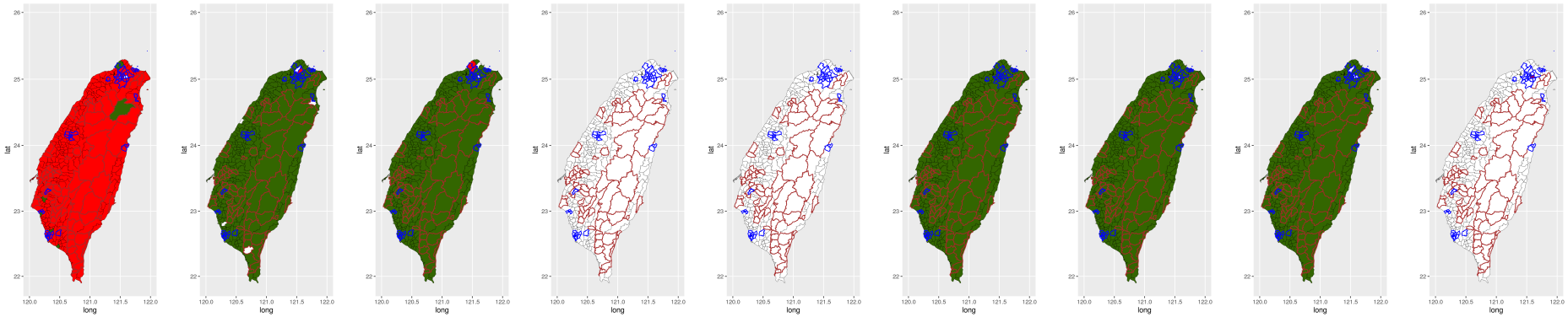


Figure 11: coa. Red: Positive effect; Green: Negative Effect.

原住民人口比例(ori)

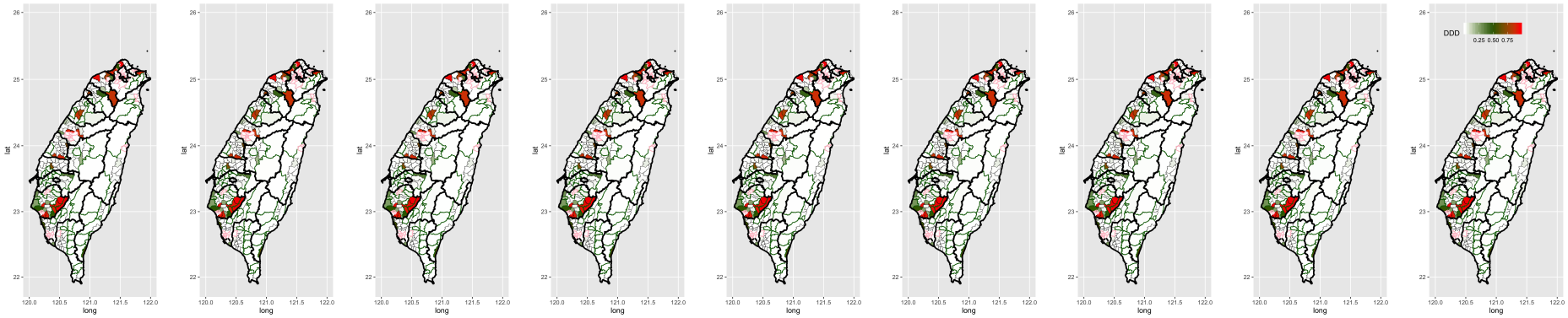


Figure 12: The distribution of originals.

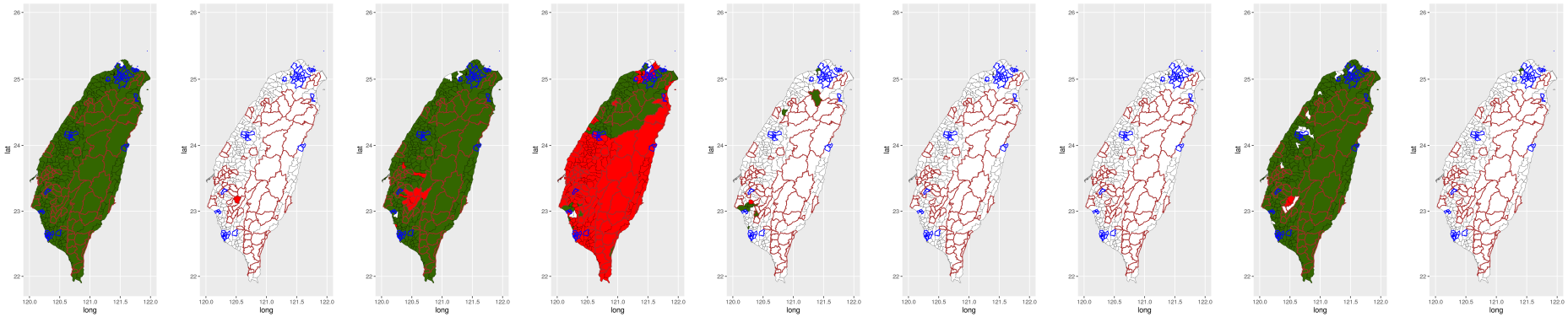


Figure 13: ori. Red: Positive effect; Green: Negative Effect.

Summary

- In varying coefficient models,
  - require computational effort;
  - allow to select important factors; and
  - model the local and remote spatial patterns.
- We conclude that
  - risks are spatially clustered, mostly in suburban areas; and
  - medical resources in remote areas should be strengthened to reduce the mortality.