Journal of Tropical Resources and Sustainable Science

journal homepage: jtrss.org

Analysis of spatial determinants of poverty in Kelantan

Siti Aisyah Nawawi*, Ibrahim Busu, Norashikin Fauzi and Mohamad Faiz Mohd Amin

Faculty of Earth Science, Universiti Malaysia Kelantan, 17600 Jeli, Kelantan, Malaysia

Received 12 July 2020
Accepted 13 July 2020
Online 28 December 2020

Keywords:
poverty, spatial autocorrelation, Conditional autoregressive models, neighbourhood matrices

⌧*Corresponding author:
Siti Aisyah Bt Nawawi
Faculty of Earth Science Faculty,
Universiti Malaysia Kelantan,
17600 Jeli, Kelantan, Malaysia
Email: aisyah.n@umk.edu.my

Abstract

This study examines socio-demographic effects on poverty and measures spatial patterns in poverty risk looking for high risk of areas. The poverty data were counts of the numbers of poverty cases occurring in each 66 districts of Kelantan. A Poisson Log Linear Leroux Conditional Autoregressive model with different neighbourhood matrices was fitted to the data. The results show that the contiguity neighbour was performed nearly similar to Delaunay triangulation neighbourhood matrix in estimate poverty risk. Apart from that, the variables average age, number of non-education of household head and number of female household head significantly associated with the number of poor households head. Kursial was found as the highest risk area of poverty among 66 districts in Kelantan.

1. INTRODUCTION

Many datasets are routinely available as summaries relating to non-overlapping areal units. Examples include disease counts, house prices, and household income. For example, in Scotland, there are Scottish statistics (http://statistics.gov.scot), which is an on-line government-run database of areal unit data on a wide variety of topics. This study used poverty data which the number of poor household counts for each areal unit.

The most common criterion in defining poverty is income insufficiency that is a household's income fails to meet a poverty line (Braithwaite & Mont, 2009; Brandolini et al., 2010). In Malaysia, Poverty Line Income (PLI) is used as a poverty threshold for classifying poor households. Even though the statistics have shown a steady decrease in poverty from one year to another, the inequality gap between the regions, states and rural-urban areas remain wide. For examples, in 2009, Sabah has the highest poverty rate with a value of 19.7%, while Melaka has the lowest poverty rate of 0.5% (Mat Zin, 2011). The difference is almost 19%. Similarly, Majid et al. (2016) found that areas with highest poverty concentrations were in northeast Kelantan and Hulu Terengganu while Klang valley had a low incidence of poverty. Siwar et al. (2016) in their study found that Kelantan proportion of the poor households in urban areas was 39.11% and the rural areas equal to 30.18%. The finding indicates that the incidence of poverty was higher in urban areas as compared to rural areas of the state. In 2012, the Department of Statistics

Malaysia revealed that Kelantan is the second-highest incidence of poverty after Sabah. Considering Kelantan is ruled by the opposition and not receiving oil royalties from the Federal Government compared with Sabah, Kelantan has potentially become the poorest state soon. Therefore, this study focuses on poverty in the state of Kelantan.

Recognising the incidence of poverty through standard statistical data tables alone is no longer adequate. The poverty data is in fact the number of poor household counts for each areal unit. This areal unit data typically exhibit spatial autocorrelation, with observations from areal units close together tending to have similar values. Ignoring this autocorrelation can result in biased parameter estimates and overly optimistic standard errors (Dormann, 2007). Most studies overcome this problem by adding a set of autocorrelated random effects on the linear predictor of the regression model. These random effects are most often modelled by a Conditional AutoRegressive (CAR) prior as part of a hierarchical Bayesian model. Examples of CAR models include the intrinsic, convolution and Leroux models. This study will use the commonly used CAR model, which is Leroux CAR with different neighbourhood matrices $\boldsymbol{W}$ to estimate poverty risk in Kelantan. We compare the different $\boldsymbol{W}$ specifications using poverty data and quantify their adequacy regarding covariate effects and fitted value estimation.

eISSN Number: 2462-2389

2. MATERIALS AND METHODS

2.1. Data of study

Data for this study were obtained from the e-Kasih database from the Ministry of Women, Family and Community Development for the year 2010. Data in e-Kasih database was updated annually. The households that meet e-Kasih criteria such as income less than country's Poverty Line Income (PLI) or household's income less than Rm1000 per month at the rural area and RM1500 at the urban area were eligible to be inserted into e-Kasih. The state of Kelantan which comprises of 66 districts (see figure 1) was the study area. The poverty data will be the number of poor households, not including hardcore poor for each of the districts. The independent variables were the socio-demographic characteristics of the household head that comprises of the number of female household head, average age and number of non-education of the household head. These independent variables were chosen based on studies from Crimmins et al. (2009), Chapoto et al. (2011), Anyanwu (2014), Majid et al. (2016), Nawawi et al. (2019), Lastrapes & Rajaram (2016) and others.

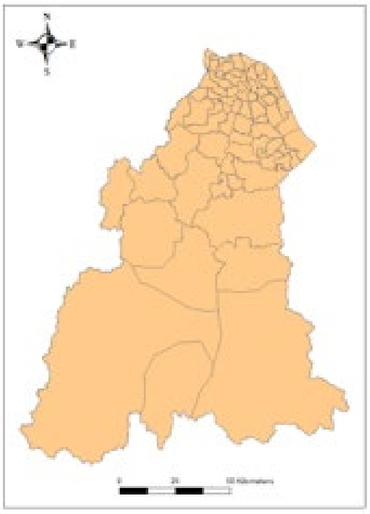

Figure 1: Map of 66 districts in Kelantan.

2.2. Standardised Poverty Rate (SPR)

The measure of poverty risk was the standardised poverty ratio (SPR), which is the ratio of the observed to the expected numbers of poverty cases (Majid et al., 2016). The formula is shown in equation 1 and 2.

$$\text{SPR} = \frac{y_k}{E_k}, \tag{1}$$

$$E_k = \frac{\sum y_k}{\sum p_k} \times p_k. \tag{2}$$

Here, y_k for $k = 1, \ldots, n$ is the number of poor households in district k. While P_k is the number of living households and E_k is expected poverty rate for each district k. For example, if SPR for an area is equal to 1.20, then this means there is a 20% increased poverty risk relative to the expected cases.

2.3. Neighbourhood matrix W

In the spatial analysis, when undertaking a test for autocorrelation and modelling data at an area level, it is essential to identify the neighbourhood structure of the data being analysed. The neighbourhood structure is defined by a neighbourhood matrix W. Given a study region with n distinct areas $\{A_1, \ldots, A_n\}$, the neighbourhood matrix W is a $n \times n$ matrix whose element w_{kj} represents a measure of closeness between area A_k and area A_j. According to O'Sullivan & Unwin (2010), the elements w_{kj} of the neighbourhood matrix, W can be binary or non-binary. In this study, contiguity neighbour and Delauney triangulation neighbourhood matrix are considered, which are the most common specification in the literature. The contiguity matrix is defined by whether two spatial units share a border or not.

$$w_{kj} \begin{cases} 1 & if\, A_k \text{ } shares\ a\ common\ border\ with\ A_j, \\ 0 & otherwise. \end{cases}$$

While the Delaunay triangulation is defined as follows:

$$w_{kj} \begin{cases} 1\ if\ centroid\ (c_k, c_j)\ are\ connected \\ \quad triangulation\ edge \\ 0\ otherwise \end{cases}$$

2.4. Moran's I

The most common statistic that is used to measure the strength of spatial autocorrelation among areal units is Moran's I. It can be defined as

$$I = \frac{n \sum_{k=1}^{n} \sum_{j=1}^{n} w_{kj}(y_k - \bar{y})(y_j - \bar{y})}{\sum_{k=1}^{n} \sum_{j=1}^{n} w_{kj} \sum_{k}^{n} (y_k - \bar{y})^2}, \tag{5}$$

where y_k is the observed value in area k, $\bar{y}$ is the overall mean and w_{kj} come from the **W** matrix. The value of I is in the interval $[-1,1]$ where a positive value indicates a positive autocorrelation and a negative value a negative or inverse correlation. The test statistic is Moran's I statistic. The p-value of Moran's I is computed using a Monte Carlo approach. Moran's I statistics are computed for K different random permutations of the data denoted $\{I_1, \ldots, I_K\}$. The observed value of Moran's I is then compared to the simulated sample distribution, $\{I_1, \ldots, I_K\}$. The null hypothesis is rejected if the probability of the observed value of Moran's I being bigger than the simulated sample distribution is less than 0.05. The hypotheses of this test are H_0 =no spatial autocorrelation and H_1 =some spatial autocorrelation

2.5. Poisson log-linear Leroux Conditional Autoregressive (CAR) model

eISSN Number: 2462-2389

A Bayesian Hierarchical model is adopted to model the data y_k using covariates information $x_k = (x_{1k}, \ldots, x_{pk})$ and a random effect ϕ_k, the latter representing the unmeasured spatial structure in the poverty cases. The random-effects $\phi = (\phi_1, \ldots, \phi_n)$ are included to model any spatial autocorrelation in the data, that persist after adjusting for the available covariate information. The random effects are modelled by a CAR prior distribution, which is a type of Gaussian Markov Random Field (GMRF) model. The model is determined by a set of n univariate full conditional distributions $f(\phi_k | \phi_{-k})$, where $\phi_{-k} = (\phi_1, \ldots, \phi_{k-1}, \phi_{k+1}, \ldots, \phi_n)$ for $k = 1, \ldots, n$. In this study, the random effects are given the Leroux CAR prior (Leroux et al. 2000). Therefore, the formulation of the Poisson log-linear Leroux CAR model used in this analysis is shown below:

$$Y_k \sim Poisson\,(E_k R_k) \quad \text{for } k = 1, \ldots, n,$$
$$\ln(R_k) = x_k^{\top} \beta + \phi_k,$$
$$\phi_k | \phi_{-k} \sim N\left[\frac{\rho \sum_{j=1}^{n} w_{kj}\phi_k}{\rho \sum_{j=1}^{n} w_{kj} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^{n} w_{kj} + 1 - \rho}\right],$$
$$\beta \sim N\,(\mu_\beta, V_\beta),$$
$$\rho \sim U\,(0,1),$$
$$\tau^2 \sim Inverse - gamma\,(0.001, 0.001).$$

In the above equation, R_k which is the poverty risk in area k will be estimated. If $R_k = 1$, then $\mathbb{E}(y_k) = E_k$ which is thus, the average risk. While if $R_k = 1.2$, then $\mathbb{E}(y_k) = 1.2E_k$ which means 20% more cases than expected. Besides, the value of the regression parameters β also will be estimated, which quantify the effects of the covariates on poverty risk. The covariates $x_k^{\top}$ are the number of female household head, average age and number of non-education of the household head. Here ρ is the level of spatial autocorrelation in the random effects, where $\rho = 1$ shows strong spatial autocorrelation between random effect and corresponds to the intrinsic model and $\rho = 0$ corresponds to independence ($\phi_k \sim N\,(0, \tau^2)$). Finally, τ^2 the conditional variance of $\phi_k | \phi_{-k}$. Inference for this type of model is typically based on Markov Chain Monte-Carlo (MCMC) simulation, using a combination of Gibbs sampling and Metropolis-Hasting steps. The software used for this study is *CARBayes* (Lee, 2013), which is an R package for Bayesian spatial modelling with conditional autoregressive priors.

3. RESULTS AND DISCUSSION

3.1 Exploratory analysis for poverty data

Herewith, the SPR values in Kelantan was found to be in the range between 0.27 and 6.40. Values above 1 represent areas with elevated levels of poverty risk, whilst values below 1 correspond to comparatively non-poverty areas. Figure 2 denotes the map of their spatial pattern. Kursial has the highest SPR (the darker area) with the SPR 6.40. Followed by Pengkalan Kubor and Perupok with the values of 3.82 and 3.57, respectively. Nevertheless, SPR was an unstable estimator of poor household risk especially when the expected counts E were small, which can occur when the population at risk was small or the cases in the location was rare. To substitute the instability of SPR, a Bayesian modelling approach was adopted to estimate poor household risk, using both covariate information and a set of random effects.

The autocorrelations between the covariates were measured to assess collinearity. It was found that there was no clear relationship were observed between the covariates except for the number of non-education of household head and the number of the female household head. Note that the relationship between these two variables was highly autocorrelated, with an autocorrelation coefficient of (0.84). Hence, a natural log transformation was applied to the number of female household head to reduce the collinearity with the number of non-education of the household head (down to 0.64).

3.2 Residuals spatial autocorrelation

Therefore, to measure the existence of spatial autocorrelation, a Poisson log-linear model (Model 4) without any random effects or any spatial structure was fitted to the poverty data. Residuals from Model 4 were tested for the presence of spatial autocorrelation. Moran's I statistic with the commonly used neighbourhood matrix based on contiguity and neighbourhood matrices were used. The Moran's I p-value was found to be less than 0.05 denoting the residuals of the non-spatial model consisted of a strong spatial autocorrelation structure. Therefore, the null hypothesis was rejected. The assumption of independence in the model was not satisfied overall. As a consequence, Poisson log-linear Leroux CAR model with two different neighbourhood matrices was applied to the poverty data in order to find the best fitting model.

3.3 Poisson log-linear Leroux CAR model

The previous section has shown the existence of a spatial autocorrelation structure of the residuals for the Poisson log-linear model. Therefore, the data set was modelled using Model (5) with the contiguity and Delaunay triangulation matrix W. Inference for each model was based on 50,000 MCMC samples with a burn-in until convergence of the first 10,000 samples and the rest of the samples were thinned by 10, to reduce their autocorrelation resulting in 4,000 samples.

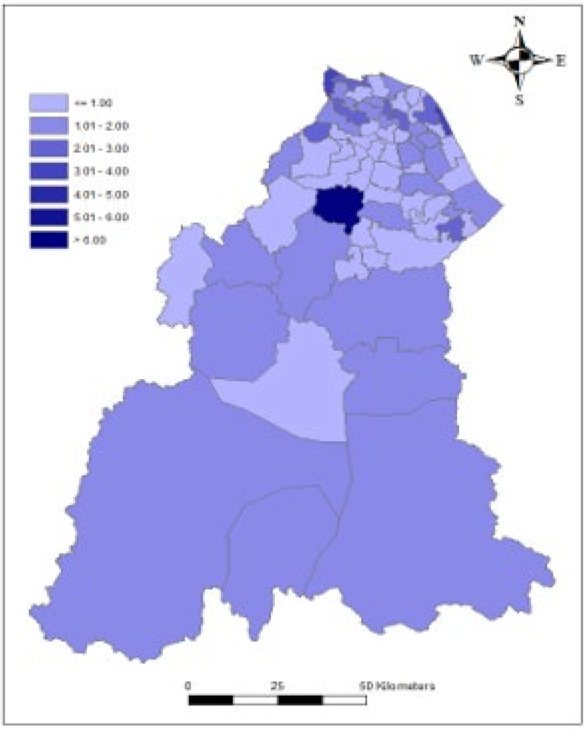

Figure 2: Spatial map of SPR for poor household head in Kelantan

Table 1 shows the results of DIC, pd and RMSE among the W matrices. The values for the DIC, pd and RMSE are also close to one another. Therefore both neighbourhood matrices perform the best in estimating poverty risk. The covariate impacts are illustrated in Table 2 and Table 3 for contiguity matrix and Delaunay triangulation, respectively. The tables show estimates (posterior median) and 95% credible intervals. The results are provided on the relative risk scale for an increase of one standard deviation in each covariate. The table shows that the relative risks and credible intervals are almost consistent among all the two W matrices. There is convincing evidence that an increase in the number of non- educated poor heads of household by 62, is associated with an increased poverty risk by 20%, indicating that no education is a very informative covariate. The poor female head of household (as measured by the log female household head) are also high at risk of poverty, with a relative risk of 1.12 for an increase of 0.16 in log female. The household head age appears to be associated with poverty reduction, with a relative risk of 0.93 for 2 years increase in the average age of the household head.
Figure 3 and 4 show the estimated poverty risks from the Poisson log-linear Leroux CAR model with contiguity and Delaunay triangulation, respectively. The estimated risk maps are smoother than the raw SPR values and are also less extreme. For example, the SPR for poverty ranges between 0.27 and 6.40, while the corresponding model estimates range between 0.28 and 6.32. However, the estimated risk surface exhibits a similar spatial pattern to the SPR map, with the highest risk is Kusial, followed by Pengkalan Kubor and Perupok. In order to determine the appropriateness of Poisson log-linear Leroux CAR model for the data, the residuals from the models were tested for the presence of spatial autocorrelation. The p-value of Moran's I statistic was greater than 0.05, indicates that there was no spatial autocorrelation as the model remove the spatial autocorrelation present in the data.

Table 1: Comparison of model fit

Model	DIC	p.d	RMSE
Contiguity	549.03	28.4	.14
Delaunay triangulation	548.64	28.6	.72

Table 2: Estimates and 95% credible interval for the regression parameters using contiguity matrix. The results were presented on the relative risk scale for a standard deviation increase in each covariates value.

Variable	Median	2.5%	97.5%
Age	0.931	0.904	0.963
No Education	1.206	1.133	1.284
Log female head	1.117	1.100	1.132

Table 3: Estimates and 95% credible interval for the regression parameters using delauney triangulation matrix. The results were presented on the relative risk scale for a standard deviation increase in each covariates value.

Variable	Median	2.5%	97.5%
Age	0.931	0.904	0.963
No Education	1.207	1.133	1.284
Log female head	1.116	1.097	1.132

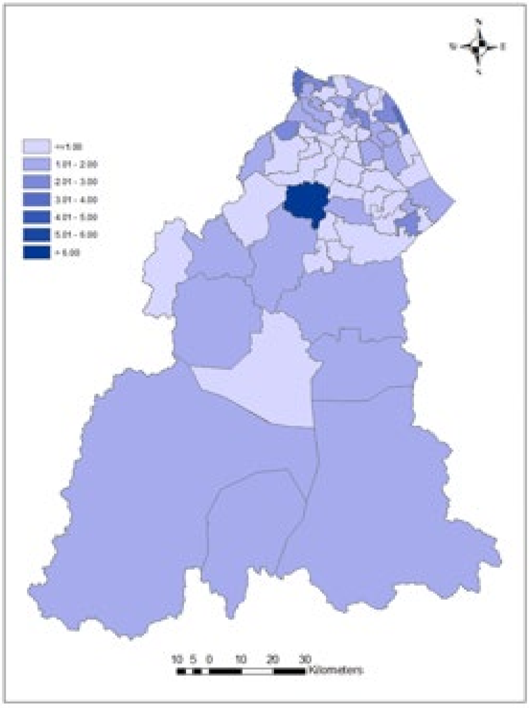

Figure 3: Poisson Log Linear Leroux CAR model with contiguity matrix indicated estimated poverty risks.

eISSN Number: 2462-2389

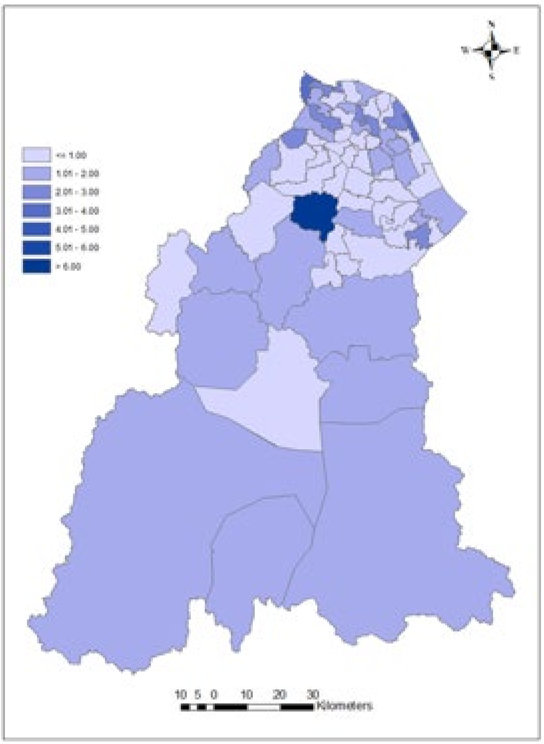

Figure 4: Poisson Log-Linear Leroux CAR model with Delaunay triangulation matrix indicated the estimated poverty risks.

4. CONCLUSION

This study explored the most commonly used conditional autoregressive prior distributions, which was Poisson log-linear Leroux CAR model with different neighbourhood matrices. Previous studies commonly used contiguity matrix as neighbourhood matrix. This study denoted the performance of Poisson log-linear Leroux CAR with Delaunay triangulation was slightly similar to contiguity matrix. Therefore, both neighbourhood matrices fit the poverty data best. Overall, Kursial was recorded to have a higher risk of poverty than the other districts in Kelantan. These differences in risk appear to be partly by virtue of the covariates. Less numbers of female household head and non-education contribute to decreasing the poverty risk. In contrast, increase the age of poor household head contribute to reducing the poverty risk

ACKNOWLEDGEMENT

The authors would like to thank Universiti Malaysia Kelantan to give an approval to conduct this research.

REFERENCES

Anyanwu, J. C. (2014), 'Marital status, household size and poverty in Nigeria: evidence from the 2009/2010 survey data', *African Development Review 26*(1), 118–137.

Braithwaite, J. & Mont, D. (2009), 'Disability and poverty: a survey of world bank poverty assessments and implications', *ALTER-European Journal of Disability Research 3*(3), 219–232.

Brandolini, A., Magri, S. & Smeeding, T. M. (2010), 'Asset-based measurement of poverty', *Journal of Policy Analysis and Management 29*(2), 267–284.

Chapoto, A., Banda, D. J., Haggblade, S., & Hamukwala, P. (2011). *Factors Affecting Poverty Dynamics in Rural Zambia*, Retrieved from http://ageconsearch.umn.edu/record/109888/files /wp55.pdf

Crimmins, E. M., Kim, J. K. & Seeman, T. E. (2009), 'Poverty and biological risk: the earlier "aging" of the poor', Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences 64(2), 286–292.

Dormann, C. F. (2007), 'Effects of incorporating spatial autocorrelation into the analysis of species distribution data', *Global Ecology and Biogeography 16*(2), 129–138.

Lastrapes, W. D. & Rajaram, R. (2016), 'Gender, caste and poverty in india: evidence from the national family health survey', *Eurasian Economic Review 6*(2), 153–171.

Lee, D. (2013). CARBayes: an R package for Bayesian spatial modeling with conditional autoregressive priors. *Journal of Statistical Software*, *55*(13), 1-24.

Lee, D. & Mitchell, R. (2014), 'Controlling for localised spatio-temporal autocorrelation in long-term air pollution and health studies.', *Statistical Methods in Medical Research 0*(0), 1–19.

Lee, D. & Shaddick, G. (2010), 'Spatial modeling of air pollution in studies of its short-term health effects', *Biometrics 66*(4), 1238–1246.

Leroux, B. G., Lei, X., & Breslow, N. (2000). Estimation of disease rates in small areas: a new mixed model for spatial dependence. In *Statistical models in epidemiology, the environment, and clinical trials* (pp. 179-191). Springer, New York, NY.

Majid, M. R., Jaffar, A. R., Man, N. C., Vaziri, M., & Sulemana, M. (2016). Mapping poverty hot spots in peninsular Malaysia using spatial autocorrelation analysis. *Planning Malaysia*, *14*(4), 1-16.

Mat Zin, R. (2011). Measuring and monitoring poverty and inequality: The Malaysian experience. *Singapore: Institute of Southeast Asian Studies*.

Nawawi, S. A., Busu, I., Fauzi, N., & Amin, M. F. M. (2019). Determinants of Poverty: A Spatial Analysis. *J. Trop. Resour. Sustain. Sci*, *7*, 83-87.

O'Sullivan, D. & Unwin, D. J. (2010), Geographic information analysis. Hoboken, NewJersey: John Wiley & Sons.

Siwar, C., Ahmed, F., Zahari, S. Z., Idris, N. D. M., Mia, M. S. & Bashawir, A. (2016), 'Poverty mapping and assessing the relationship between poverty and socio demographic characteristics of households: A study in kelantan, malaysia', *Journal of Environmental Science and Technology 9*(6), 407–416.

eISSN Number: 2462-2389