Bayesian Spatial Modeling of School Disparities in Brunei Darussalam

Alvin Bong Jia Lok

Visting Student (3rd Year Undergraduate), CEMSE Division, KAUST

bongbj@kaust.edu.sa

Introduction

Education is a key pillar of national development. Brunei's Vision 2035 (Wawasan Brunei 2035) and the UN Sustainable Development Goals both emphasize equitable access to quality education. However, the spatial distribution of schools remains understudied in Brunei, especially using quantitative methods. This study applies Bayesian spatial modeling to explore disparities in government school availability across Brunei's 39 mukims. Our goal is to **identify areas with potential under-provision of schools**, accounting for population, land area, and housing costs.

Data

- **Units of analysis:** Areal - 39 mukims (sub-districts)
- **Response variable:** Government school counts (primary & secondary, 2018)

Covariates	Units / Notes	Year
Population of youth	per 1000; aged 0-25	2021 census
Mukim area	10 km^2	-
Spatiotemporal trend in average house prices	socioeconomic proxy imputed via INLA-Gaussian model	2015-2023

Methods

Let Y_i and E_i denote the observed and expected counts of schools, respectively, in mukim $i \in \{1, \ldots, n\}$. Let θ_i represent the relative abundance of schools in mukim i. We model school counts across mukims using a **Bayesian spatial hierarchical Poisson model**, incorporating spatial dependence via the Besag-York-Mollie (BYM) model:

$$Y_i \mid \theta_i \sim \text{Poisson}(E_i \cdot \theta_i), \quad i = 1, \ldots, n$$

$$\log(\theta_i) = \beta_0 + \beta_1 \cdot \text{pop\_youth}_i + \beta_2 \cdot \text{area}_i + \beta_3 \cdot \text{hp}_i + u_i + v_i$$

Where:

- β_0: intercept,
- $\beta_1, \beta_2, \beta_3$: coefficients for standardized covariates (population, area size, house price)
- u_i is a structured spatial effect, modelled using an intrinsic conditional autoregressive (CAR) prior $u_i \mid u_{-i} \sim \mathcal{N}(\bar{u}_{\delta_i}, \frac{1}{\tau_u n_{\delta_i}})$
- v_i is an unstructured random effect, $v_i \sim \text{Normal}(0, \frac{1}{\tau_v})$

The expected counts are computed using indirect standardization:

$$E_i = \sum_{j=1}^{n} Y_j \cdot \frac{\text{pop}_i}{\sum_{j=1}^{n} \text{pop}_j}$$

- **Queen contiguity** is used to define spatial neighborhood structure.
- Model fitting is performed using **Integrated Nested Laplace Approximation (INLA)**.
- The **relative abundance** θ_i quantifies whether a mukim i has higher ($\theta_i > 1$) or lower ($\theta_i < 1$) school count than the national average.

Results

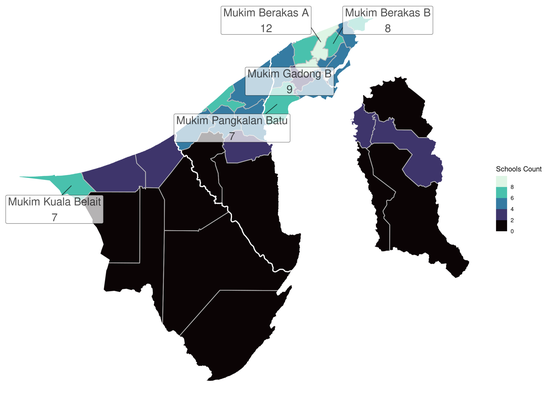
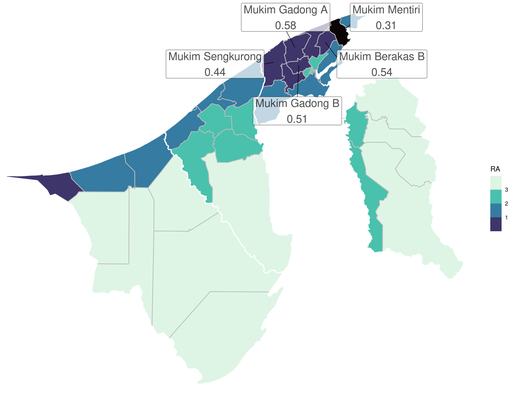
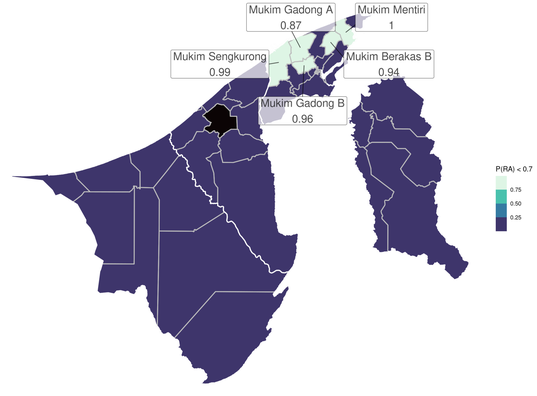


Figure 1: (a) observed school counts



(b) Estimates of Relative Abundance (RA)



(c) Non-exceedance Probability that RA is less than 0.7

Discussion

Term	Mean	2.5%	97.5%	Significant
(Intercept)	0.824	0.430	1.215	Yes
pop_youth	-0.123	-0.163	-0.083	Yes
area	0.018	0.001	0.036	Yes
hp	0.001	-0.002	0.003	No

- No evident economic disparity in school distribution.
- Larger mukims have more schools.
- Higher youth population is associated with fewer schools.
- Higher RA in rural mukims; lower RA in urban Brunei-Muara coastal region.
- **Underserved mukims** (RA < 0.7): Sengkurong, Gadong A & B, Berakas B, and Mentiri.
- Urban mukims exhibit signs of under-provision despite higher population densities.
- Underserved areas often coincide with new housing developments (e.g., Perpindahan Lugu).
- School infrastructure may lag behind residential expansion.

"التعليم في الصغر كالنقش على الحجر"

"Education at a young age is like engraving on stone."

Future Work

- Incorporate other spatial accessibility measures:
  i) distance to schools ii) school size / capacity
- Update with more recent data sources, with spatial-temporal models
- Comparison to private schools
- Explore alternative models (Negative Binomial, zero-inflated Poisson) to better handle overdispersion.

References

Alvin Bong (2025). *Bayesian Spatial Modeling of School Disparities in Brunei Darussalam.* https://github.com/alvinbjl/school-disparity-areal-model/