Bayesian Spatial Risk Modelling of Suicide Mortality in South Korea Using the ICAR Model

Introduction

South Korea consistently reports one of the highest suicide rates among developed nations, with 24 to 28 deaths per 100,000 people in recent years—nearly twice the OECD average of 11 (Kim, 2024). Although rates declined modestly in the mid-2010s following national prevention efforts, they have since resurged, reaching 28.3 per 100,000 in 2024—the highest level in over a decade (YNA, 2025). This persistent trend positions South Korea as a global outlier and underscores the urgent need to investigate its underlying risk factors. A large body of research shows that suicide rates are shaped by a mix of socioeconomic, demographic, cultural, and psychological factors. Classic theories dating back to Durkheim (1951) emphasise the protective role of social integration, while economic hardship and mental illness remain key drivers. In South Korea, national crises such as the late 1990s Asian financial crash have been linked to suicide spikes, primarily through rising unemployment and social dislocation (Chang, 2009).

Among the many determinants examined, social isolation have been identified as a major risk factor. Individuals living alone or lacking family support are more vulnerable due to loneliness and reduced integration. In South Korea, areas with more single-person households or divorced residents tend to have higher suicide rates (Jang, 2022), reflecting the link between weaker social ties and suicide risk. Psychological stress is another significant factor. South Korea reports high levels of perceived stress, which is associated with depression and suicide ideation (Oh, 2020). The indicator "stress awareness"—the proportion of people experiencing significant stress—has been linked to higher suicide rates at the regional level (Jang, 2022). Unemployment is a well-documented driver. Job loss can trigger financial strain, identity loss, and psychological distress. Park and Lester (2006) identified unemployment as a key contributor to South Korea's suicide rate, particularly where social safety nets are weak. Lastly, unmet medical needs reflect gaps in healthcare access, including mental health services. Though 90% of people who die by suicide have a diagnosable condition, only about 15% ever receive treatment. Kim (2022) finds that regions with higher unmet medical needs tend to experience higher suicide mortality.

Given the spatially patterned nature of suicide, advanced statistical methods are needed to separate risk factors from geographic influences. We use a Bayesian Intrinsic Conditional Auto-Regressive (ICAR) model, which incorporates information from adjacent regions to produce more stable small-area risk estimates. Our analysis includes four key covariates discussed above, alongside spatial random effects to assess their contribution to geographic variation. This modelling framework accounts for unobserved spatial influences and estimates exceedance probabilites to identify municipalities where suicide risk surpasses a defined threshold. Our key outputs will provide localised insights into suicide risk and guide targeted public health interventions and efficient resource allocation across South Korea.

1. Data and Methods

1.1 Study Area and Spatial Units

The study area comprises all local government units in South Korea, including metropolitan cities, special administrative regions, and provinces, aggregated to the municipality level for analytical consistency. Administrative boundaries were derived from 2022 shapefiles and subsequently harmonised with statistical datasets from various sources.

1.2 Target and Predictor Variables

The primary outcome variable is the number of suicide casualties in 2022, computed using suicide rates (per 100,000 population) and registered population figures. Four predictor variables were selected based on literature and data availability:

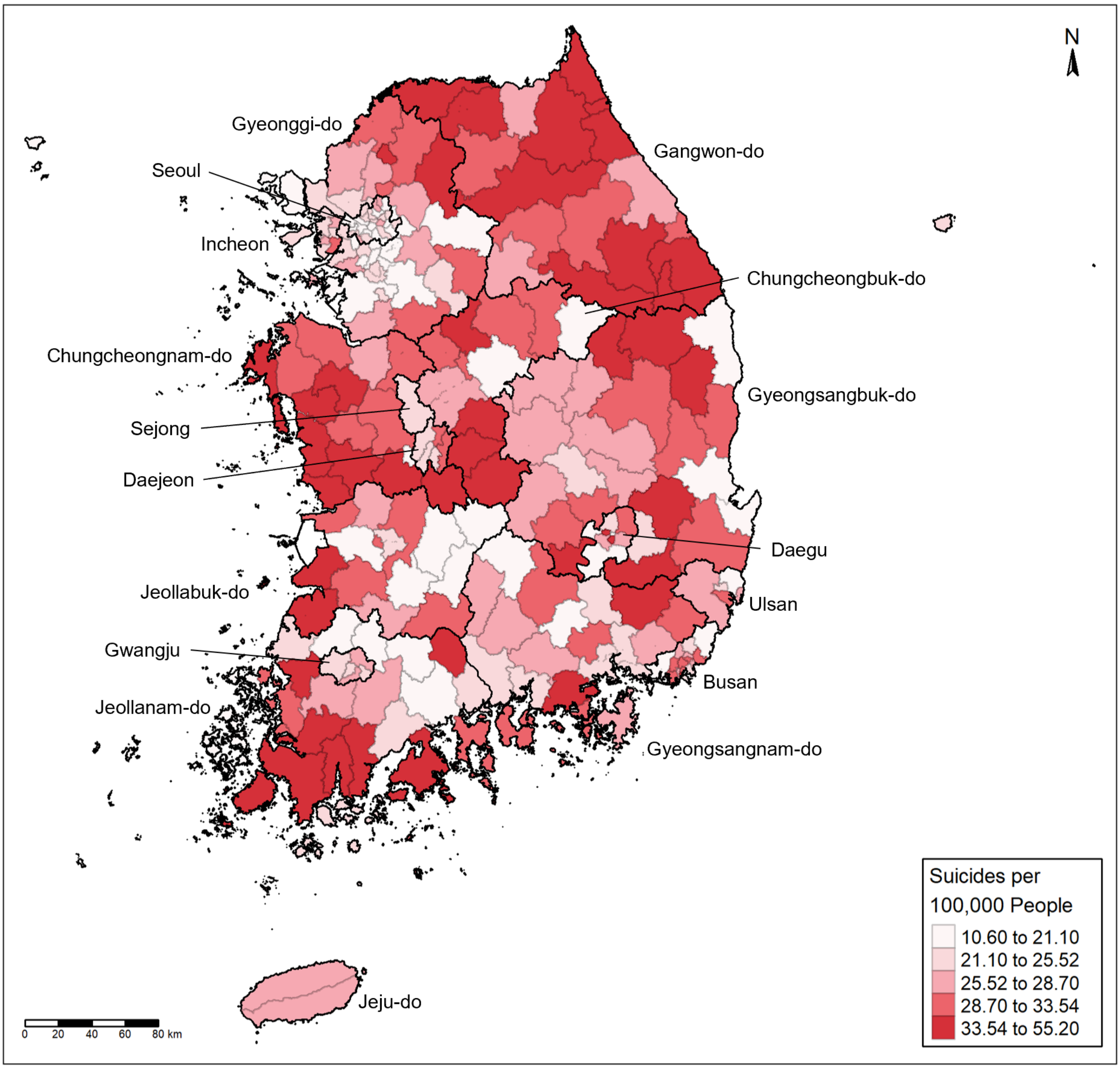
Define variables here.

Table 1: Summary of Datasets Used in the Analysis

Variable	Dataset	Usage	Source	Year
Suicide Rate	Cause of Death Statistics	Target variable	Statistics Korea (Population Trends Division)	2022
Single-Person Household Ratio	Population and Housing Census	Predictor variable	Statistics Korea (Population and Housing Census Division)	2022
Stress Awareness Rate	Community Health Survey	Predictor variable	Korea Disease Control and Prevention Agency (KDCA)	2022
Unemployment Rate	Local Employment Statistics	Predictor variable	Statistics Korea (Employment Statistics Division)	2022
Unmet Medical Needs Rate	Community Health Survey	Predictor variable	Korea Disease Control and Prevention Agency (KDCA)	2022
Estimated Population	Future Population Projections	Standardisation / Denominator	Statistics Korea (Regional Statistics Planning Division)	2022
Municipality Boundaries	Municipality Shapefiles	Spatial analysis	GIS Developer	2022
Province Boundaries	Province Shapefiles	Mapping / Labelling	GIS Developer	2022

All covariates reflect conditions in 2022. Municipality-level datasets were retrieved from national public databases, then cleaned and merged based on administrative codes and names.

Figure 1: Suicide Rates Across South Korean Municipalities



1.3 Model Specification

We applied a Bayesian Poisson model incorporating an ICAR prior to account for spatial autocorrelation among neighbouring municipalities. The model estimates relative risk (RR) of suicide for each area, using the expected number of cases (based on population) as an offset:

$$Y_i \sim \text{Poisson}(\lambda_i), \quad \log(\lambda_i) = \log(E_i) + \alpha + \mathbf{X}_i\boldsymbol{\beta} + \sigma \cdot \phi_i \tag{1}$$[1]

[1]In the full model, we include both structured and unstructured random effects: $\log(\lambda_i) = \log(E_i) + \alpha + \mathbf{X}_i\boldsymbol{\beta} + \sigma \cdot \left(\sqrt{1-\rho}\,\theta_i + \sqrt{\rho}\,\phi_i\right)$, where $\theta_i \sim \mathcal{N}(0,1)$ captures unstructured heterogeneity and ϕ_i follows an ICAR prior for spatial structure.

2. Results and Discussion

2.1 Global Parameter Estimates

The model converged successfully with all $\hat{R}$ values below 1.05 and high effective sample sizes. Table 1 summarises posterior estimates:

Table 2: Posterior Estimates of Global Parameters

Parameter	**Mean**	**2.5% CrI**	**97.5% CrI**	$\boldsymbol{n}_{\text{eff}}$	$\boldsymbol{\hat{R}}$
α	-0.1956	-0.4548	0.0685	25689.1	1.0005
β_1	0.0120	0.0067	0.0174	27152.2	1.0003
β_2	-0.0052	-0.0139	0.0035	26780.3	1.0001
β_3	-0.0385	-0.0706	-0.0065	9776.5	1.0007
β_4	0.0089	-0.0014	0.0190	28092.1	1.0001
σ	0.2118	0.1611	0.2685	7507.3	1.0006

Note. CrI = Credible Interval; n_{eff} indicates the effective sample size from the posterior. $\hat{R}$ denotes the Gelman-Rubin convergence diagnostic, with values close to 1.00 indicating good convergence.