

Spatial Effects on the Human Development Index in South Sulawesi: A Residual Bootstrap Spatial Durbin Model Approach

Muh. Kemal Rydindo Irwan¹, Siswanto Siswanto², Andi M. Alfin Baso³

^{1,2,3}Statistics Study Program, Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Makassar, Indonesia

Article Info

Article history:

Received 06/05/2026

Revised 09/09/2026

Accepted 10/09/2026

Keywords:

Human Development Index

Residual Bootstrap

Spatial Durbin Model

Spatial Regression

ABSTRACT

This study examines spatial dependence in the Human Development Index (HDI) of South Sulawesi Province using a Spatial Durbin Model (SDM), which captures both the direct effects of regional predictors and the indirect (spillover) effects from neighboring regions. To improve the stability of parameter inference under a limited number of observations, a residual bootstrap procedure is integrated into the SDM estimation. Using 2023 cross-sectional data for 24 districts/cities and four predictors—percentage of poor population, per capita expenditure, life expectancy, and expected years of schooling—the residual bootstrap SDM yields a lower Akaike Information Criterion and a higher coefficient of determination ($AIC = 54.6039$, $R^2 = 0.9906$) than the conventional SDM ($AIC = 69.4245$, $R^2 = 0.9774$). Per capita expenditure, life expectancy, and expected years of schooling significantly affect HDI directly, while per capita expenditure and expected years of schooling also show significant spatial spillover effects on neighboring regions. These findings indicate that incorporating a residual bootstrap into the SDM framework improves estimation stability under limited sample conditions, offering more reliable evidence to support spatially targeted regional human development policies in South Sulawesi.



This work is licensed under a [CC-BY-NC](https://creativecommons.org/licenses/by-nc/4.0/)

Corresponding Author:

Siswanto Siswanto,

Department of Statistics,

Universitas Hasanuddin,

Jl. Perintis Kemerdekaan KM.10, Kec. Tamalanrea, Kota Makassar, Sulawesi Selatan.

Email: Siswanto@unhas.ac.id

1. INTRODUCTION

Spatial regression analysis is a statistical approach used to model data influenced by location or regional factors. In the context of spatial data, modeling does not only focus on the relationships among variables but must also account for spatial effects in the form of spatial heterogeneity and spatial dependence [1]. Spatial dependence indicates the existence of relationships between the value of a variable in one region and values in neighboring regions. Meanwhile, spatial heterogeneity reflects differences in characteristics

across regions that cannot be uniformly explained [2]. Therefore, location and geographical proximity play a crucial role in understanding patterns and relationships among social and economic phenomena [3]. Spatial modeling is essential to produce more accurate and contextually relevant analyses for geographically structured data [4].

The use of regression models that ignore spatial structure poses a risk of producing biased and inefficient parameter estimates. This occurs because the assumption of independence among observational units is violated in data characterized by geographical interdependence [5]. When variables within a system exhibit interrelationships based on geographical proximity, such cross-location relationships must be explicitly modeled through spatial regression approaches. One model capable of capturing these interactions is the Spatial Durbin Model (SDM). This model accommodates both the direct effects of variables within a region and the indirect effects of variables from neighboring regions [6].

The Spatial Durbin Model is a spatial regression approach that captures the relationships between dependent and independent variables through the use of a spatial weight matrix. This model can be applied to both cross-sectional and panel data [7]. SDM combines elements of the Spatial Lag Model and the Spatial Lag of X by incorporating spatial lags of both the dependent and independent variables. Consequently, SDM is able to explain the effects of variables within a region as well as the spillover effects from neighboring regions simultaneously [8]. This capability makes SDM more precise in representing complex and interconnected spatial dynamics across regions [9].

Despite its advantages in capturing spatial relationships, the application of spatial regression is not free from methodological challenges, particularly limitations related to sample size. Theoretically, the validity of regression inference relies heavily on the assumptions of the central limit theorem, which require an adequate sample size for the estimator distribution to approximate normality. When the sample size is relatively small, the parameter estimates may become unstable, and tests for spatial dependence may lose statistical power [10].

To address this issue, the bootstrap approach has been widely recommended in the statistical literature. Bootstrap is a resampling technique that generates an empirical distribution of an estimator by repeatedly resampling from the available data [11]. This approach is considered effective in improving the stability of parameter estimation and enhancing statistical inference under conditions of limited observations [12]. One commonly used bootstrap method in regression analysis is the residual bootstrap, which resamples residuals from the model to preserve the regression error structure [13].

Numerous empirical studies have demonstrated the successful application of bootstrap approaches in spatial analysis. Residual bootstrap to a Spatial Autoregressive model for poverty analysis and obtained more stable estimation results [14]. Spatial regression approaches have also been developed in the form of robust SDM to map disease distribution and capture spatial interactions across regions more accurately [15]. In addition, the application of bootstrap methods to spatial models has been proven to enhance the reliability of statistical inference when dealing with limited observational data [16]. However, existing bootstrap applications in spatial econometrics have largely been developed for the Spatial Autoregressive model or for GTWR-type frameworks, and residual bootstrap has not yet been specifically implemented within the SDM for regional human development analysis, leaving a gap between the growing body of bootstrap-based spatial methods and their application to HDI modeling. This study addresses this gap, and its novelty lies in integrating a residual bootstrap procedure directly into SDM estimation to obtain more stable inference for HDI determinants under the limited number of observational units available at the district/city level.

The Human Development Index (HDI) is a strategic indicator used to measure the success of human development and to determine the level of development of a region. Based on data from Statistics Indonesia, the HDI of South Sulawesi Province in 2023 was recorded at 73.46, which remains below the national average. Geographical proximity among districts and cities potentially induces spatial autocorrelation in human development achievements, rendering conventional regression analysis less appropriate for HDI modeling. Therefore, a Spatial Durbin Model approach is required to capture both direct and indirect interregional effects. Given the limited number of observational units, this study integrates the SDM with a residual bootstrap approach to obtain more stable and robust parameter estimates in identifying factors influencing HDI in South Sulawesi Province. Specifically, this study aims to (i) estimate the SDM with residual bootstrap for HDI in South Sulawesi Province and (ii) identify the socioeconomic factors that significantly affect HDI both directly and through spatial spillover from neighboring districts/cities. The main contribution of this study is a residual bootstrap procedure embedded within the SDM framework that improves the stability of parameter inference under a small number of spatial units, offering more reliable evidence to support spatially-targeted human development policy in South Sulawesi Province.

2. METHOD

The data used in this study are secondary cross-sectional data obtained from the official publication of Statistics Indonesia (Badan Pusat Statistik) for South Sulawesi Province. The units of analysis comprise all

24 districts/cities in the province; thus, the study constitutes a complete enumeration of the administrative units rather than a drawn sample, and no inclusion or exclusion criteria beyond administrative boundaries were applied.

This study incorporates one dependent variable and four independent variables to examine the determinants of human development and their spatial interaction effects across different regions. The selection of variables was based on theoretical considerations related to socioeconomic development and the multidimensional nature of human well-being. The dependent variable in this study is the Human Development Index (HDI) (Y), which is used to measure the level of human development based on three fundamental dimensions: a long and healthy life, knowledge, and a decent standard of living.

The independent variables consist of four indicators. The first is the Percentage of Poor Population (PPP) (X_1), defined as the proportion of individuals whose average monthly per capita expenditure falls below the national poverty line. The second variable is Expenditure per Capita (EC) (X_2), measured as the average total monthly expenditure per individual within a household to meet living needs. The third variable is Life Expectancy (LE) (X_3), defined as the average number of years a person is expected to live from birth under prevailing mortality patterns. The fourth variable is Expected Years of Schooling (EYS) (X_4), defined as the average number of years of formal education that a seven-year-old child is expected to receive in the future.

The empirical analysis was conducted in several systematic stages to ensure the robustness and validity of the spatial econometric estimation. The procedure includes model specification, spatial diagnostic testing, parameter estimation, and model evaluation, as described in the following steps

1. Data exploration was conducted using descriptive and spatial analyses.
2. Multiple linear regression modeling was conducted to obtain an initial overview of the relationships among variables and to test the classical assumptions in the regression model. The multiple linear regression model is expressed in Equation (1).

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n; k = 1, 2, 3, \dots, p \quad (1)$$

The general equation of the linear regression model is expressed in matrix notation, as shown in Equation (2).

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

where $\mathbf{y}$ is a vector of dependent variables of size $n \times 1$, $\boldsymbol{\beta}$ is a vector of regression parameters of size $(p + 1) \times 1$, $\boldsymbol{\varepsilon}$ is a vector of errors of size $n \times 1$, and $\mathbf{X}$ is a matrix of independent variables of size $n \times (p + 1)$. The parameters of the multiple linear regression model were estimated based on the formula written in Equation (3).

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y}) \quad (3)$$

3. A spatial weight matrix was constructed using the inverse distance weighting (IDW) approach. The weights were calculated from the great-circle (Euclidean) distance between the geographic coordinates of the administrative center of each district/city, obtained from official BPS geospatial references, and the resulting matrix was row-standardized so that the weights for each district/city sum to one.
4. Spatial heterogeneity was tested using the Breusch-Pagan (BP) test. The BP test statistic was calculated using Equation (4). The BP test examines whether the error variance differs systematically across districts/cities (spatial heterogeneity), which is distinct from Moran's I, which instead examines whether values in one district/city are correlated with values in neighboring districts/cities (spatial dependence); both conditions are assessed because either can motivate the use of a spatial regression model.

$$BP = \frac{1}{2} \mathbf{f}'^T \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'^T \mathbf{f} \sim \chi^2 \quad (4)$$

5. Spatial dependence was examined using Moran's I statistic. The spatial dependence test used the formula in Equation (5).

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n \tilde{w}_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n \tilde{w}_{ij} (x_i - \bar{x})^2} \quad (5)$$

6. The parameters of the spatial Durbin model (SDM) were estimated using maximum likelihood estimation (MLE). The SDM specification includes a spatial lag of the dependent variable and spatial lags of all independent variables, as written in Equation (7), and MLE relies on the assumption that the error term is normally distributed with a mean of zero and constant variance, an assumption evaluated through the residual diagnostics reported in this study.
7. The residuals were calculated using Equation (6).

$$\boldsymbol{\varepsilon} = \mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{W}\mathbf{X}\boldsymbol{\theta} \quad (6)$$

8. The parameter significance was tested using the Wald test.

9. A residual bootstrap procedure with B replications was performed to obtain stable parameter estimates. The procedure resampled the SDM residuals with replacement, added the resampled residuals back onto the fitted values to generate a bootstrapped dependent variable, and re-estimated the SDM on this bootstrapped variable while holding the spatial weight matrix and the estimated spatial autoregressive coefficient fixed, so that the spatial dependence structure of the original model was preserved across replications. Replication sizes of 200, 500, 1,000, 2,000, 5,000, and 10,000 were compared, and stability was declared once the standard errors no longer changed appreciably between successive replication levels, which occurred at $B = 10,000$.
10. Model selection was based on the R^2 and Akaike Information Criterion (AIC). A higher R^2 indicates that a larger proportion of the variation in HDI is explained by the model, whereas a lower AIC indicates a better balance between goodness of fit and model complexity; the model with the higher R^2 and the lower AIC among those compared was selected as the best model.

3. RESULTS AND DISCUSSION

3.1 Parameter Estimation in the Spatial Durbin Model

The spatial Durbin model (SDM) is a form of spatial regression that resembles the spatial autoregressive (SAR) model, with the main characteristic being the presence of a spatial lag in the dependent variable (y). The SDM parameters were estimated using the maximum likelihood estimation (MLE) method. The likelihood function was constructed using the residual equation (ε), which followed a normal distribution, as presented in Equation (7).

$$\varepsilon = y - \rho W y - Z \delta \quad (7)$$

Based on the residual in Equation (7), the likelihood function was expressed as presented in Equation (8).

$$l(\rho, \delta, \sigma^2 | y) = -\frac{n}{2} (2\pi) - \frac{n}{2} \ln(\sigma^2) + \ln |I_n - \rho W| - \frac{1}{2\sigma^2} ((I_n - \rho W)y - Z\delta)' ((I_n - \rho W)y - Z\delta) \quad (8)$$

To estimate the parameter δ in the SDM, the log-likelihood function derived from Equation (8) was maximized. This procedure involved taking the derivative of the equation with respect to δ and equating it to zero.

$$\begin{aligned} \frac{\partial l}{\partial \delta} &= \frac{\partial \left(-\frac{1}{2\sigma^2} ((I_n - \rho W)y - Z\delta)' ((I_n - \rho W)y - Z\delta) \right)}{\partial \delta} = 0 \\ \frac{\partial \left(-\frac{1}{2\sigma^2} ((I_n - \rho W)y)' ((I_n - \rho W)y) - \delta' Z' ((I_n - \rho W)y) - ((I_n - \rho W)y)' Z \delta + \delta' Z' Z \delta \right)}{\partial \delta} &= 0 \\ \frac{\partial \left(-\frac{1}{2\sigma^2} (-2\delta' Z' (I_n - \rho W)y + \delta' Z' Z \delta) \right)}{\partial \delta} &= 0 \\ -\frac{1}{2\sigma^2} (-2Z' (I_n - \rho W)y + 2Z' Z \delta) &= 0 \\ \frac{1}{\sigma^2} Z' Z \delta &= \frac{1}{\sigma^2} Z' (I_n - \rho W)y \end{aligned} \quad (9)$$

Thus, the estimator for the parameter δ was obtained as shown in Equation (10).

$$\hat{\delta} = (Z' Z)^{-1} Z' y - \hat{\rho} (Z' Z)^{-1} Z' W y \quad (10)$$

3.2 Exploratory Data Analysis and Spatial Patterns of HDI

The Human Development Index (HDI) data of South Sulawesi Province in 2023 indicate considerable variation in human development achievements across districts/cities. The highest HDI value was recorded in Makassar City (83.52), while the lowest was observed in the Selayar Islands Regency (69.20). The average HDI of South Sulawesi Province in 2023 was 72.15.



Figure 1. Spatial Distribution Map of HDI in South Sulawesi 2023

The HDI distribution map reveals clear regional disparities in the level of human development. Makassar City stands out as the only region classified as having a very high HDI, distinguishing it markedly from other districts/cities. Most regions fall into the high HDI category, while several districts remain in the medium category. No region is classified as having a low HDI, indicating that overall human development in South Sulawesi is relatively good. Nevertheless, the observed interregional differences reflect development inequality that requires policy attention.

The spatial clustering pattern based on HDI levels shows that districts/cities with relatively similar HDI achievements tend to be geographically adjacent to each other. This finding suggests that human development is not independent across regions but is spatially interconnected. Regions with high HDI may exert positive spillover effects on neighboring areas and vice versa, highlighting the importance of a regional-based approach to human development analysis.

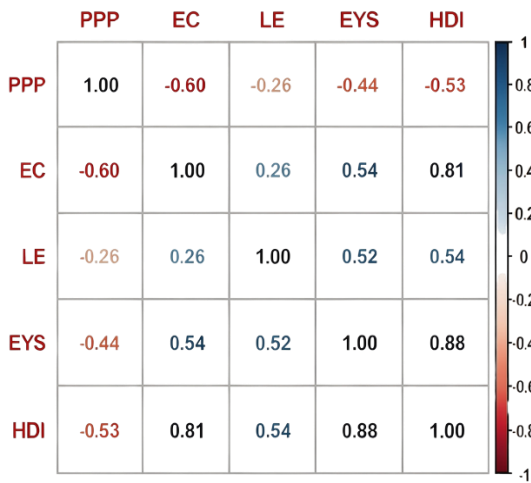


Figure 2. Correlation Plot among Research Variables

The correlation plot among variables shows that per capita expenditure and life expectancy are positively associated with HDI. This indicates that improvements in economic capacity and population health tend to be accompanied by higher human development achievements. Conversely, the percentage of the poor population exhibits a negative relationship with HDI, implying that higher poverty levels may reduce the quality of human development within a region. In addition, relatively high correlations are observed among several independent variables, reflecting the interrelated nature of socioeconomic indicators in shaping human development outcomes. These interconnections reinforce the argument that HDI analysis should be conducted comprehensively rather than partially. Overall, the exploratory data analysis provides an initial overview of HDI distribution patterns and the relationships among socioeconomic indicators in South Sulawesi Province. These findings form a critical basis for further analysis using a spatial regression approach, as they indicate that HDI variation across districts/cities is influenced not only by internal regional characteristics but also by spatial interactions with neighboring areas.

3.3 Spatial Dependence of the Human Development Index

Spatial dependence analysis was conducted to identify the presence of spatial interrelationships in Human Development Index (HDI) achievements among districts/cities in South Sulawesi Province. This approach aims to determine whether HDI values in a given region are influenced by HDI values in neighboring regions. The results of the Moran's Index test on the response variable, independent variables, and regression residuals are presented in Table 1.

Table 1. Results of Moran's I Test

Variable	Moran's Index	<i>p</i> -value	Decision ($\alpha = 0.05$)
<i>Y</i>	0.0464	0.0108	Reject H_0
X_1	0.1015	0.0003	Reject H_0
X_2	0.0277	0.0363	Reject H_0
X_3	-0.0932	0.8815	Fail to Reject H_0
X_4	0.0390	0.0195	Reject H_0
Residual	-0.0951	0.8894	Fail to Reject H_0

The Moran's I test results indicate that the dependent variable (*Y*) has a *p*-value smaller than $\alpha = 0.05$ with a positive Moran's index, suggesting the presence of spatial dependence across locations for the dependent variable. Furthermore, several independent variables (X_1 , X_2 , and X_4) also exhibit *p*-values below $\alpha = 0.05$, indicating spatial dependence among these explanatory variables. The global spatial dependence test conducted on the residuals yields a *p*-value greater than $\alpha = 0.05$, implying no spatial dependence in the residuals. Since the dependent variable and at least one independent variable exhibit spatial dependence, the Spatial Durbin Model (SDM) is deemed appropriate for this study.

3.4 Estimation of the Spatial Durbin Model

The analysis proceeds with the estimation of the spatial Durbin model (SDM) to identify the factors influencing the human development index (HDI) in South Sulawesi Province while accounting for spatial interactions among regions. This model allows for the measurement of the effects of independent variables originating from both the region itself and neighboring regions. The estimated parameters of the SDM are presented in Table 2.

Table 2. Estimation Results of the Spatial Durbin Model for HDI

Parameter	Estimate	<i>z</i> -value	<i>p</i> -value
α	-341.9500	-1.1502	0.2501
β_1	0.0842	1.0887	0.2763
β_2	0.0009	10.3984	0.0000
β_3	1.0155	1.9730	0.0485
β_4	2.6330	8.4180	0.0000
θ_1	0.2051	0.2642	0.7916
θ_2	0.0034	2.9839	0.0028
θ_3	4.9935	1.2623	0.2069
θ_4	10.9080	3.3111	0.0009
ρ	-3.6497		0.0017

The SDM with a residual bootstrap approach can be expressed as follows:

$$\hat{y}_i = -3.6497 \sum_{j=1}^n w_{ij} y_j - 341.9500 + 0.0842 x_{i1} + 0.0009 x_{i2} + 1.0155 x_{i3} + 2.6330 x_{i4} \\ + 0.2051 \sum_{j=1}^n w_{ij} x_{1j} + 0.0034 \sum_{j=1}^n w_{ij} x_{2j} + 4.9935 \sum_{j=1}^n w_{ij} x_{3j} + 10.9080 \sum_{j=1}^n w_{ij} x_{4j}$$

Where $\sum_{j=1}^n w_{ij}$ represents the spatial interaction between the observed district/city i and its neighboring districts/cities j . The SDM parameter estimates indicate that independent variables associated with parameters β_3 ($p\text{-value} = 0.0485$) and β_4 ($p\text{-value} = 0.0000$) significantly influence the dependent variable at the 5% significance level. In addition, the spatially lagged variables corresponding to parameters θ_2 ($p\text{-value} = 0.0028$) and θ_4 ($p\text{-value} = 0.0009$) are also statistically significant at the 5% level. The positive coefficients (0.0034 and 10.9080) suggest that higher values in neighboring regions tend to increase the HDI in the focal region. The spatial autoregressive coefficient (ρ) is also statistically significant ($p\text{-value} = 0.0017$), indicating a strong spatial effect in the dependent variable itself. It should be noted that β and θ here represent, respectively, the direct effect of a predictor within its own district/city and the raw spillover association with neighboring districts/cities; because the model also includes the spatially lagged dependent variable (ρ), these raw coefficients do not equal the total effect, and a formal impact decomposition into direct, indirect, and total effects would be required to quantify the exact magnitude of spillover before drawing conclusions about the size of neighboring-region influence.

3.5 Spatial Durbin Model with the Residual Bootstrap Approach

The Spatial Durbin Model (SDM) was estimated using a residual bootstrap approach. The initial stage of the bootstrap process was carried out by resampling the residuals obtained from the previously estimated SDM. Subsequently, the bootstrapped dependent variable was modeled with the independent variables using the SDM approach. The residual values used to generate the next replication were obtained through a resampling process of the residuals from the initial model. The procedure was repeated iteratively to obtain the smallest standard error by varying the bootstrap sample size through the parameter nb . This selection was performed by comparing the results from each iteration level, and the replication that produced the smallest standard error was used as the reference for presenting the results in Table 3.

Table 3. Comparison of Standard Residuals Values for Each Replication

Parameter	Without Replication	200	500	1000	2000	5000	10000
α	297.29	314.2946	301.15	298.9439	296.2107	294.4517	293.5987
β_1	0.0773	0.0699	0.0749	0.0759	0.0762	0.07	0.0754
β_2	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
β_3	0.5147	0.5175	0.5072	0.5043	0.5069	0.5037	0.5015
β_4	0.3128	0.2733	0.293	0.3033	0.3138	0.3102	0.3046
θ_1	0.7764	0.8464	0.6844	0.719	0.7255	0.7301	0.7204
θ_2	0.0011	0.001	0.001	0.001	0.001	0.001	0.001
θ_3	3.956	3.9619	3.8626	3.856	3.8272	3.8101	3.8081
θ_4	3.2943	3.1776	3.1848	3.2365	3.2662	3.2211	3.1575
ρ	0.73735	0.6995	0.691	0.688	0.6993	0.711	0.7063

Table 3 showed that when $B = 10,000$, the smallest standard error was obtained, and it was not substantially different from the standard error of the original model. Therefore, the best estimation result for the residual bootstrap resampling in the SDM modeling was obtained with $B = 10,000$ replications. The SDM modeling using the residual bootstrap approach on the analyzed data produced regression parameters as presented in Table 4.

Table 4. Estimation Results of the Spatial Durbin Model Using the Residual Bootstrap Approach

Parameter	Estimate	$z\text{-value}$	$p\text{-value}$
α	27.7866	0.0946	0.9246
β_1	0.0807	0.0946	0.2842
β_2	0.0009	1.0710	0.0000
β_3	1.0074	10.6203	0.0445
β_4	2.6436	2.0089	0.0000
θ_1	0.1980	8.6778	0.7835
θ_2	0.0038	0.2748	0.0003
θ_3	5.2212	3.5762	0.1704

Parameter	Estimate	z-value	p-value
θ_4	12.0741	1.3711	0.0001
ρ	-4.0902	3.8240	0.0000

The SDM with the residual bootstrap approach can be expressed as follows:

$$\hat{y}_i = -4.0902 \sum_{j=1}^n w_{ij} y_j + 27.7866 + 0.0807 x_{i1} + 0.0009 x_{i2} + 1.0074 x_{i3} + 2.6436 x_{i4} \\ + 0.1980 \sum_{j=1}^n w_{ij} x_{1j} + 0.0038 \sum_{j=1}^n w_{ij} x_{2j} + 5.2212 \sum_{j=1}^n w_{ij} x_{3j} + 12.0741 \sum_{j=1}^n w_{ij} x_{4j}$$

The bootstrap-based estimates indicate stable results, as reflected by low bias and relatively small standard errors across parameters. As shown in Table 3, the standard error of ρ changes by less than 0.02 between $B = 5,000$ and $B = 10,000$ (from 0.711 to 0.7063), and the standard errors of the remaining parameters likewise change only marginally at this range, which is the quantitative basis for regarding $B = 10,000$ as the point of stability rather than relying on the smallest standard error alone. The SDM with residual bootstrap, constructed using the percentage of poor population (X_1), per capita expenditure (X_2), life expectancy (X_3), and expected years of schooling (X_4), provides a comprehensive depiction of the socioeconomic factors influencing human development. Based on the estimation results, variables X_2 , X_3 , and X_4 have a statistically significant effect on improving HDI values. This is consistent with human capital and welfare theory, in which household consumption capacity, health, and education jointly determine long-run human development outcomes, and aligns with prior empirical findings that per capita expenditure, life expectancy, and years of schooling are among the strongest correlates of HDI at the district/city level [9], [16]. The spatial autoregressive coefficient (ρ) is also statistically significant, indicating the presence of spatial interdependence in HDI across regions. Moreover, indirect effects arising from neighboring regions through spatially lagged independent variables contribute significantly to the model.

3.6 Selection of The Best Models

Model selection was conducted to evaluate the appropriateness of the SDM with the bootstrap approach for the Human Development Index case in South Sulawesi Province. The evaluation was based on goodness-of-fit indicators, namely the coefficient of determination (R^2) and the Akaike Information Criterion (AIC). The results of the best model selection is presented in Table 5.

Table 5. Goodness-of-Fit Comparison

Model	R^2	AIC
SDM	0.9774	69.4245
SDM-Bootstrap	0.9906	54.6039

The comparison of model performance between the SDM and SDM-Bootstrap models based on R^2 and AIC values indicates that the SDM-Bootstrap model performs best. This model achieves an R^2 value of 0.9906, indicating that 99.06% of the variation in HDI can be explained by the explanatory variables, namely percentage of poor population (X_1), per capita expenditure (X_2), life expectancy (X_3), and expected years of schooling (X_4). In addition, the SDM-Bootstrap model records the lowest AIC value of 54.6039, reflecting the highest model efficiency and optimal goodness of fit. By comparison, using the same dataset and the same set of predictors, the conventional SDM model yields an R^2 of 0.9774 with an AIC of 69.4245 (Table 5). Therefore, the SDM-Bootstrap approach not only enhances the explanatory power of the model but also reduces statistical complexity, making it the most appropriate model for describing the spatial characteristics of HDI in South Sulawesi Province.

4. CONCLUSION

The results of the spatial regression modeling using the residual bootstrap approach indicate that the best model for analyzing the Human Development Index (HDI) in South Sulawesi Province in 2023 is the Spatial Durbin Model with residual bootstrap. This model is characterized by the smallest AIC value of 54.6039 and an R^2 of 0.9906. The estimated SDM-Bootstrap equation is expressed as follows:

$$\hat{y}_i = -4.0902 \sum_{j=1}^n w_{ij} y_j + 27.7866 + 0.0807 x_{i1} + 0.0009 x_{i2} + 1.0074 x_{i3} + 2.6436 x_{i4} \\ + 0.1980 \sum_{j=1}^n w_{ij} x_{1j} + 0.0038 \sum_{j=1}^n w_{ij} x_{2j} + 5.2212 \sum_{j=1}^n w_{ij} x_{3j} + 12.0741 \sum_{j=1}^n w_{ij} x_{4j}$$

The factors that significantly influence HDI include per capita expenditure, life expectancy, and expected years of schooling, all of which show a statistically significant direct effect within their own district/city. Among these, per capita expenditure and expected years of schooling additionally show a statistically significant spatial spillover association with neighboring regions, whereas life expectancy shows a significant direct effect only, with no significant spillover association. The residual bootstrap approach improves the stability of the SDM estimates, as reflected in smaller and more consistent standard errors and p-values for several key parameters across bootstrap replications (Table 3), suggesting more reliable inference under the study's limited number of spatial units. These findings suggest that regional human development policy in South Sulawesi Province may benefit from prioritizing per capita expenditure and years of schooling, given their combined direct and spillover influence on HDI. This study is limited by its reliance on 2023 cross-sectional data, so that the estimated relationships describe associations at a single point in time and do not capture how these effects may evolve over subsequent years.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to all parties who contributed to the completion of this research. Special thanks are extended to the academic supervisors for their guidance and valuable feedback throughout the research process, as well as to the relevant institutions for providing the necessary data and support. This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

DAFTAR PUSTAKA

- [1] N. Paramita, M. Masjkur, and Indahwati, "Spatial Regression Model with Optimum Spatial Weighting Matrix on GRDP Data of Sulawesi Island Spatial Regression Model with Optimum Spatial Weighting Matrix on GRDP Data of Sulawesi Island," *J. Phys. Conf. Ser.*, 2021, doi: 10.1088/1742-6596/1863/1/012045.
- [2] D. I. Wulansari, M. F. Ghazali, and Armijon, "Perubahan Distribusi Spasial Permukiman Dan Relasinya Terhadap Pertumbuhan Penduduk Di Kota Bandar Lampung," *Geodika*, vol. 7, no. 1, pp. 11–22, 2023, doi: 10.29408/geodika.v7i1.6173.
- [3] O. Setiawan *et al.*, "The Egyptian Journal of Remote Sensing and Space Sciences Unraveling land use land cover change , their driving factors , and implication on carbon storage through an integrated modelling approach," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 27, no. 4, pp. 615–627, 2024, doi: 10.1016/j.ejrs.2024.08.002.
- [4] I. R. Ramadhani, R. Rahmawati, and A. Hoyyi, "Analisis Faktor-Faktor Yang Mempengaruhi Gizi Buruk Balita Di Jawa Tengah Dengan Metode Spatial Durbin Model," *Jural Gaussian*, vol. 2, no. 4, pp. 333–342, 2013.
- [5] L. Anselin, *Spatial econometrics: methods and models*. 1988.
- [6] J. LeSage and R. K. Pace, *Introduction to Spatial Econometrics*. CRC Press, 2009.
- [7] I. Ernawati, W. Somayasa, Arman, and Alfian, "Pemilihan Model Regresi Linier Berganda Dengan Kriteria AIC," *J. Mat. Komputasi dan Stat.*, vol. 3, no. 1, pp. 233–237, 2023.
- [8] C. E. Nababan and E. Simamora, "Bootstrap Estimation of Confidence Intervals of Multiple Regression Model Parameters in the Presence of Multicollinearity Using Principal Component Analysis," *Formosa J. Appl. Sci.*, vol. 2, no. 1, pp. 185–202, 2023.
- [9] J. E. Junior and J. Priyono, "Analisis Pengaruh Indeks Pembangunan Manusia terhadap Tingkat Kemiskinan di Kabupaten Lembata, Provinsi Nusa Tenggara Timur," *Ekon. J. Ilm. Manajemen, Ekon. Bisnis, Kewirausahaan*, vol. 12, no. 2, pp. 1231–1245, 2025.
- [10] N. Muhtasib, *Metode Fast Double Bootstrap Pada Regresi Spasial Data Panel Fast Double Bootstrap Method On Spatial Panel Data Regression*. 2015.
- [11] S. Gilda, B. Heidrich, and F. Kiraly, "Enhancing Time Series Analysis with Advanced Bootstrapping Techniques," *creativecommons*, vol. 1, no. 4, 2024.
- [12] L. M. Zeqo, "Some empirical results using block bootstrap in estimating the coefficients of a periodic autoregressive model," *Balk. J. Interdiscip. Res. Vol.*, vol. 9, no. 2, pp. 34–41, 2023.
- [13] F. Prihatmono, M. Y. Darsyah, and A. Karim, "Residual Bootstrap Resampling Method For Multiple,"

- J. Litbang Edusaintech*, vol. 1, no. 1, pp. 35–43, 2020.
- [14] P. W. Handayani *et al.*, *Konsep CB-SEM dan SEM-PLS disertai dengan contoh kasus*. 2019.
- [15] M. Wijana, U. Aryanti, S. Dewi, and N. I. Yusman, “Analisis Faktor-faktor yang Mempengaruhi Kesuksesan Sistem Elektronik PKH Menggunakan Model DeLone dan McLean,” *Internal*, vol. 6, no. 1, pp. 10–21, 2023.
- [16] M. H. Mukrom, H. Yasin, and A. R. Hakim, “Pemodelan angka harapan hidup provinsi jawa tengah menggunakan robust spatial durbin model,” *J. Gaussian*, vol. 10, no. 1, pp. 44–54, 2021.