STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, September 2026, pp 2690–2702.
Published online in International Academic Press (www.IAPress.org)

SUR-SAR Modelling of the Human Development Index (HDI) of Regencies and Municipalities in West Java Province, Indonesia

Muhammad Aji Alrasyid [1], Suci Astutik [1,*], Ahmad Zia Ul-Saufie [2], Rahma Fitriani [1]

[1] *Department of Statistics, Brawijaya University, Indonesia*
[2] *Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, Malaysia*

Abstract This study examines regional disparities in the Human Development Index (HDI) across 27 regencies and municipalities in West Java Province from 2022 to 2024 using a Spatial Seemingly Unrelated Regression–Spatial Autoregressive (SUR–SAR) model. HDI in West Java exhibits both spatial clustering and temporal persistence, indicating that development outcomes are shaped by interregional interactions and structural continuity across years. To capture these dependencies, the analysis incorporates eight socio-economic variables such as poverty, availability of health facilities, population density, inequality, unemployment, consumption level, GRDP per capita, and regional minimum wage into a multi-equation spatial system estimated using Maximum Likelihood. Four model classes are evaluated sequentially that is OLS, SUR, SAR, and SUR–SAR. The results demonstrate that the SUR–SAR specification provides the best overall fit and effectively accounts for both spatial dependence and inter-year error correlations. Population density consistently emerges as the strongest positive and highly significant determinant of HDI, while poverty shows a stable negative effect across multiple models. Minimum wage becomes significant only within the SUR–SAR framework, indicating that the influence of labour income on human development is more apparent once spatial and temporal linkages are jointly modelled. The spatial lag parameter is positive and significant across all equations, confirming the presence of HDI spillover effects among neighboring regions.

Keywords Spatial Econometrics, Seemingly Unrelated Regression (SUR), Spatial Autoregressive Model (SAR), SUR–SAR Model

AMS 2010 subject classifications 62M30, 91B82

DOI: 10.19139/soic-2310-5070-3261

1. Introduction

Human development is a multidimensional policy objective that integrates improvements in health, education, and living standards into a single composite indicator, the Human Development Index (HDI). In West Java Province, one of Indonesia's most populous and economically dynamic regions, HDI values have shown steady improvement indicating gradual progress in overall human welfare. In 2024, the HDI of West Java Province reached 74.92, an increase of 0.53 points from the previous year, maintaining its status in the "high" development category. This improvement was driven by progress across all dimensions such as life expectancy rose to 74.92 years, expected years of schooling increased to 14.16 years, and mean years of schooling reached 9.37 years. In terms of living standards, the adjusted real per capita expenditure climbed to around IDR 13.23 million per year. However, disparities remain significant, with Depok City recording the highest HDI at approximately 83.69, while Cianjur Regency had the lowest at around 69.59, highlighting ongoing challenges in achieving equitable human development across West Java [7].

*Correspondence to: Suci Astutik (Email: suci_sp@ub.ac.id). Department of Statistics, Brawijaya University. Veteran Road, Malang, East Java, Indonesia (65145).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

However, the pace of this improvement remains uneven across regencies and municipalities, reflecting structural disparities in income opportunities, labor markets, educational attainment, and access to services. Such disparities are geographically patterned rather than random, as neighboring jurisdictions often share transport corridors, labor flows, supply chains, and social networks. Consequently, development outcomes in one area may be related to those in adjacent areas, suggesting that human development in West Java is spatially interconnected rather than independent. Despite continued progress, HDI inequality across regions persists, posing both empirical and methodological challenges. Empirically, differences in income distribution, education quality, and infrastructure investment produce enduring gaps between urbanized and rural areas. These patterns imply that HDI outcomes exhibit spatial dependence, whereby districts with similar socio-economic characteristics tend to cluster geographically [2, 27]. Moreover, HDI values across consecutive years are likely temporally correlated, given that the effects of provincial policies, economic growth, and demographic transitions evolve gradually. Classical linear regression models, which assume independently distributed errors, become inadequate under these conditions because they fail to account for cross-location and inter-year dependencies [17].

To address these issues, this study employs a Spatial Seemingly Unrelated Regression – Spatial Autoregressive (SUR–SAR) framework that jointly captures spatial dependence among districts and cross-equation error correlation across multiple years. The theoretical foundations of this approach are drawn from two established econometric traditions. The Spatial Autoregressive (SAR) model explicitly incorporates spatial interactions in the dependent variable through a spatial weight matrix, recognizing that outcomes in one region may depend on those in neighboring regions [2, 23, 27]. The Seemingly Unrelated Regression (SUR) framework, introduced by Zellner [29] and further developed by Wickens [26], allows for contemporaneous correlations across related equations, enhancing efficiency and consistency in multi-equation estimation. Combining these two approaches yields the SUR–SAR model, estimated using Maximum Likelihood Estimation (MLE), which effectively captures both spatial and inter-year linkages within a unified inferential framework.

Accordingly, this study has three aims within a clearly delimited scope. First, it specifies and estimates a SUR–SAR model for HDI across the 27 regencies/municipalities of West Java for 2022, 2023, and 2024, using a policy-relevant set of socio-economic predictors drawn from established determinants of human development. Second, it adopts Maximum Likelihood Estimation (MLE) as the estimation method to remain consistent with foundational spatial econometric practice and to keep the inferential workflow transparent and replicable [2, 23, 26, 29]. Third, it positions the SUR–SAR model as an incremental and meaningful methodological refinement over single-equation SAR or non-spatial SUR approaches commonly used in regional HDI analysis in Indonesia. The novelty lies in deploying a SUR–SAR model for West Java's 2022–2024 HDI with MLE, an application that remains underrepresented in the current empirical literature for the province while preserving the integrity of classical inferential standards and the parsimony required for a small-scale study. This scope and design provide a clear, policy-relevant evidence base for provincial planning and a replicable template for future research that might extend the model to alternative weighting schemes, additional years, or complementary inferential frameworks.

2. Methodology

This section describes the data and modeling framework used to analyze the HDI across the 27 regencies and municipalities of West Java Province for the 2022–2024 period. The empirical approach integrates spatial and temporal dependencies using a SUR–SAR framework estimated by MLE. The analysis proceeds through four stages: (1) data acquisition and variable specification, (2) estimation of the SUR model, (3) estimation of the SAR model, and (4) integration of both approaches into the SUR–SAR model.

2.1. Data Acquisition and Variable Specification

This study employs secondary data obtained from the *Open Data Jabar* portal (`https://opendata.jabarprov.go.id/id`), managed by the West Java Provincial Government. The analysis covers 27 regencies and municipalities in West Java Province over the 2022–2024 period.The location of West Java in Indonesia, along with its 27 regencies and municipalities, can be seen in Figure 2.

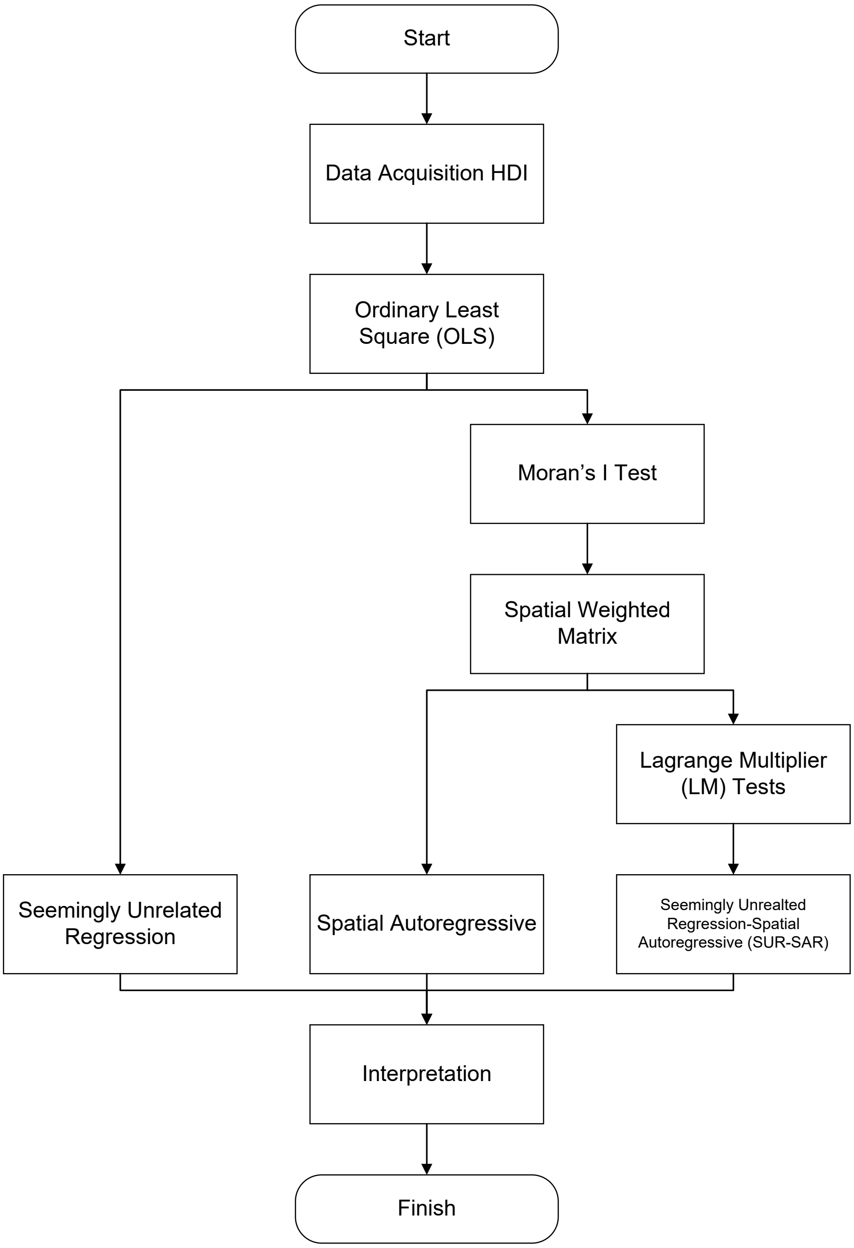

Figure 1. The Research Stages

The dependent variable is the Human Development Index (HDI) (Y) for each regency/municipality and year, denoted as HDI_t for $t = 2022,\ 2023,\ 2024$. The independent variables (X) are selected according to empirical evidence and theoretical relevance in explaining regional HDI in Indonesia. These variables are:

- POV_t : Poverty Rate (%) (X_{1t})
- NHC_t : Number of Health Centers (units) (X_{2t})
- PDS_t : Population Density (people/km²) (X_{3t})

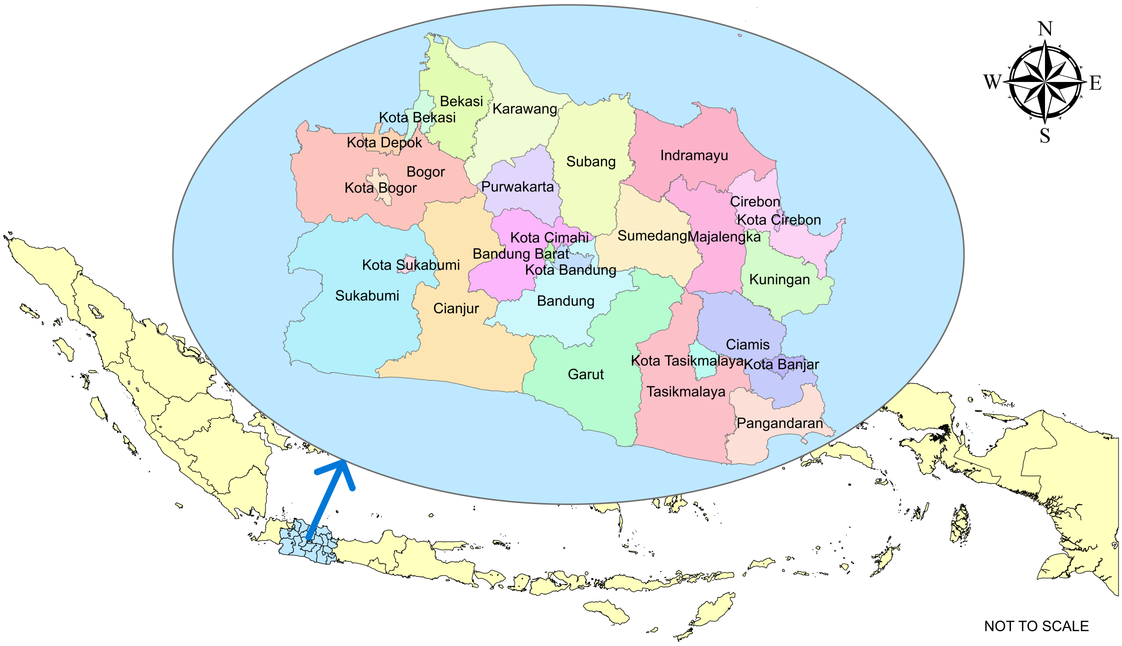

Figure 2. West Java Maps

- GNR_t : Gini Ratio (index) (X_{4t})
- UNE_t : Open Unemployment Rate (%) (X_{5t})
- CLI_t : Consumption Level Index (points) (X_{6t})
- GPC_t : Gross Regional Domestic Product (GRDP) per Capita (log) (Rupiah) (X_{7t})
- LMW_t : Minimum Wage (log) (log Rupiah) (X_{8t})

The data are measured consistently across the three years to ensure comparability. The spatial weight matrix ($\mathbf{W}$) is constructed based on economic distance, not merely geographical proximity. This approach considers the similarity of regional economic structures and GRDP levels, under the assumption that areas with closer economic conditions are more likely to influence one another. Each element w_{ij} of the matrix represents the inverse of the economic distance between regions i and j, and the matrix is row-standardized, such that the weights for each region sum to one.

The system of equations for the SUR–SAR model estimated by MLE is expressed as follows:

$$\begin{aligned} y_{22} = {} & \rho_1 \mathbf{W} y_{22} + \beta_{0,22} + \beta_{1,22} X_{1,22} + \beta_{2,22} X_{2,22} + \beta_{3,22} X_{3,22} + \beta_{4,22} X_{4,22} \\ & + \beta_{5,22} X_{5,22} + \beta_{6,22} X_{6,22} + \beta_{7,22} X_{7,22} + \beta_{8,22} X_{8,22} + \varepsilon_{22} \qquad (1) \\ y_{23} = {} & \rho_2 \mathbf{W} y_{23} + \beta_{0,23} + \beta_{1,23} X_{1,23} + \beta_{2,23} X_{2,23} + \beta_{3,23} X_{3,23} + \beta_{4,23} X_{4,23} \\ & + \beta_{5,23} X_{5,23} + \beta_{6,23} X_{6,23} + \beta_{7,23} X_{7,23} + \beta_{8,23} X_{8,23} + \varepsilon_{23} \qquad (2) \\ y_{24} = {} & \rho_3 \mathbf{W} y_{24} + \beta_{0,24} + \beta_{1,24} X_{1,24} + \beta_{2,24} X_{2,24} + \beta_{3,24} X_{3,24} + \beta_{4,24} X_{4,24} \\ & + \beta_{5,24} X_{5,24} + \beta_{6,24} X_{6,24} + \beta_{7,24} X_{7,24} + \beta_{8,24} X_{8,24} + \varepsilon_{24} \qquad (3) \end{aligned}$$

Each equation represents the spatial autoregressive HDI model for one year (2022, 2023, 2024). The term $\rho_t \mathbf{W} y_t$ captures spatial dependence through the economic distance matrix $\mathbf{W}$, while the error components ε_t are assumed to be correlated across equations, following:

$$Var(\varepsilon) = \mathbf{\Sigma} \otimes I_{27}, \tag{4}$$

where $\mathbf{\Sigma}$ is the cross-equation covariance matrix representing contemporaneous correlations across years, and I_{27} is the identity matrix for the 27 spatial units.

2.2. Seemingly Unrelated Regression Model (SUR)

The Seemingly Unrelated Regression (SUR) model, first introduced by Zellner [29], extends the classical linear regression model to a system of multiple equations that may have different dependent and independent variables. Although each equation can appear unrelated, the error terms across equations are assumed to be correlated, allowing efficiency gains through joint estimation. According to Wickens [26], in general, the SUR model for M equations can be expressed as:

$$\begin{aligned} y_{1i} &= \beta_{11}X_{1i,1} + \beta_{12}X_{1i,2} + \cdots + \beta_{1K_1}X_{1i,K_1} + \varepsilon_{1i} \\ y_{2i} &= \beta_{21}X_{2i,1} + \beta_{22}X_{2i,2} + \cdots + \beta_{2K_2}X_{2i,K_2} + \varepsilon_{2i} \\ &\vdots \\ y_{Mi} &= \beta_{M1}X_{Mi,1} + \beta_{M2}X_{Mi,2} + \cdots + \beta_{MK_M}X_{Mi,K_M} + \varepsilon_{Mi} \\ i &= 1,\ 2,\ \ldots,\ N \end{aligned} \tag{5}$$

Using matrix notation, Equation (5) can be written as:

$$\mathbf{y}_m = \boldsymbol{\beta}_m \mathbf{X}_m + \varepsilon_m \tag{6}$$

where $\mathbf{y}_m$ is an $n_i \times 1$ vector of responses, $\mathbf{X}_m$ is an $n_i \times k_i$ matrix of predictors, $\boldsymbol{\beta}_i$ is a $k_i \times 1$ parameter vector, and ε_i is an $n_i \times 1$ error vector.

The error term follows:

$$E\left(\varepsilon_i\right) = 0, Var\left(\varepsilon_i\right) = \mathbf{\Sigma} \otimes \mathbf{I}_n, \tag{7}$$

where $\mathbf{\Sigma}$ is the cross-equation covariance matrix and I_n is the identity matrix. This indicates that while observations within an equation are independent, the residuals between equations may be correlated.

Parameter estimation in the SUR model is typically performed using MLE, which explicitly account for the covariance structure among equations to produce more efficient estimates [5, 20]. When the error terms across equations are uncorrelated (i.e., the covariance matrix $\mathbf{\Sigma}$ is diagonal), the SUR framework simplifies to a set of independent OLS regressions without efficiency loss [1]. However, if the error terms are contemporaneously correlated, SUR estimators such as MLE yield more efficient parameter estimates than OLS by leveraging cross-equation information [16, 22, 21]. Consequently, the SUR model enhances estimation accuracy in systems where different equations share unobserved shocks or contemporaneous disturbances, making it a valuable methodological tool in economics, regional development, and social science research [10, 13].

2.3. Spatial Regression Model (SAR)

The Spatial Regression Model extends classical regression by explicitly incorporating spatial dependence, the principle that observations located near one another tend to exhibit stronger relationships than those farther apart. Rooted in Tobler's First Law of Geography, which states that "everything is related to everything else, but near things are more related than distant things," spatial association must be recognized when modeling regional or geographically referenced data [24]. Consequently, spatial dependence represents a fundamental characteristic of regional phenomena and should not be ignored in empirical analysis.

Spatial dependence refers to the correlation among observations situated in proximate locations, where the value of a variable in one area is influenced by values in neighboring areas [2]. This relationship is commonly evaluated using Moran's I, a global measure of spatial autocorrelation. Positive values indicate clustering of similar attributes, whereas negative values reflect spatial dispersion. Statistical significance is typically assessed through the standardized Moran's I statistic, which approximates a normal distribution under the null hypothesis of no spatial dependence [27]. The magnitude of spatial dependence can thus be examined through the Moran's I index, as formulated in Equation (8).

$$\mathbf{I} = \frac{\widehat{\varepsilon}^{\mathbf{T}}\mathbf{W}\widehat{\varepsilon}}{\widehat{\varepsilon}^{\mathbf{T}}\widehat{\varepsilon}} \tag{8}$$

where $\widehat{\varepsilon}$ residual vector obtained by the OLS method and $\mathbf{W}$ is standardized spatial weighting matrix. To calculate the Moran's I index, a spatial weight matrix is required. Spatial weights matrices $(\mathbf{W})$ define the structure of these spatial relationships by specifying which regions are considered neighbors. Contiguity-based spatial relationships can be specified using rook, bishop, or queen criteria, in which neighboring units are identified by a shared boundary, a shared vertex, or a combination of both [2]. Alternatively, customized weights can incorporate additional dimensions like economic distance, transportation networks, or social factors [9]. With the presence of spatial dependence and the availability of a spatial weight matrix, spatial regression modeling can then be performed.

Anselin [2] formulated the general spatial regression model as:

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{u}, \text{ where} \\ \mathbf{u} &= \lambda\mathbf{W}\mathbf{u} + \varepsilon \end{aligned} \tag{9}$$

where $\mathbf{y}$ is the dependent variable, $\mathbf{X}$ the matrix of independent variables, $\mathbf{W}$ the spatial weights matrix, ρ the spatial autoregressive coefficient, λ the spatial error coefficient, and ε the error term.

Depending on the parameters, this general form yields several models:

- If $\rho = 0$ and $\lambda = 0$: Classical Linear Regression.
- If $\lambda = 0$: Spatial Autoregressive Model (SAR), captures dependence in the dependent variable.
- If $\rho = 0$: Spatial Error Model (SEM), accounts for correlated errors.
- If both ρ and λ are nonzero: Spatial Autocorrelation Model (SAC).

Spatial regression provides a powerful and robust framework for modeling interaction effects across geographic locations, allowing researchers to identify spatial clustering, diffusion processes, and spillover patterns in economic, social, and environmental data [12]. Unlike standard regression models that assume independent observations, spatial regression explicitly accounts for spatial dependence and autocorrelation among neighboring units, leading to unbiased and more efficient parameter estimates [8, 14]. By incorporating spatial lags or spatially correlated error structures, these models capture interdependencies that reflect real-world processes such as regional policy spillovers, economic diffusion, or environmental externalities, thereby enhancing empirical validity, reliability, and explanatory power in spatial data analysis [28, 30].

2.4. SUR–SAR Model

The first step prior to estimating the SUR–SAR model is to to assess both the presence and the nature of spatial dependence through Lagrange Multiplier (LM) diagnostics, which help identify whether spatial effects arise in the dependent variable, the error term, or both. Following Anselin [2], the LM-Lag test evaluates spatial autocorrelation in the dependent variable, whereas the LM-Error test detects correlation in the disturbance term, their robust counterparts are designed to distinguish the dominant spatial process when multiple sources of dependence are present [3]. In multi-equation systems, these diagnostics are extended as LM-SUR tests to assess spatial lag (LM_{SAR}^{SUR}) or spatial error (LM_{SEM}^{SUR}) dependence across correlated equations [11, 19]. Although conventional LM tests may be oversized, the robust forms proposed Anselin [3] are widely recommended to select the appropriate model specification whether SUR-SAR, SUR-SEM, or SUR-SARAR [6, 15, 18]. Once the relevant spatial process is confirmed, the SUR–SAR model can then be estimated.

The SUR–SAR model combines two important econometric frameworks, the SUR model, which accounts for correlated errors across multiple equations, and the SAR model, which incorporates spatial dependence among regions [2, 29]. This integration allows the simultaneous estimation of several spatial regression equations that share cross-equation correlations and spatial interactions.

For M equations, the SUR–SAR system can be written as:

$$\begin{aligned} \mathbf{y}_1 &= \rho_1 \mathbf{W}\mathbf{y}_1 + \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1 \\ \mathbf{y}_2 &= \rho_2 \mathbf{W}\mathbf{y}_2 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2 \\ &\vdots \\ \mathbf{y}_M &= \rho_M \mathbf{W}\mathbf{y}_M + \mathbf{X}_M\boldsymbol{\beta}_M + \boldsymbol{\varepsilon}_M \end{aligned} \tag{10}$$

Using matrix notation, Equation (10) can be written as:

$$\mathbf{y}_m = \rho_m \mathbf{W}\mathbf{y}_m + \mathbf{X}_m\boldsymbol{\beta}_m + \varepsilon_m \tag{11}$$

where $\mathbf{y}_m$ is an $n_i \times 1$ vector of responses, $\boldsymbol{\rho}_m$ is an $M \times 1$ vector of spatial autoregressive coefficient, $\mathbf{X}_m$ is an $n_i \times k_i$ matrix of predictors, $\boldsymbol{\beta}_i$ is a $k_i \times 1$ parameter vector, and ε_i is an $n_i \times 1$ error vector. The vector of disturbances satisfies:

$$Var\left(\varepsilon\right) = \boldsymbol{\Sigma} \otimes \mathbf{I}_n, \tag{12}$$

indicating contemporaneous correlation among the equations.

The model captures two key features, Spatial autocorrelation through the lagged dependent variable $\mathbf{W}y_i$, reflecting that outcomes in one region may depend on those in neighboring regions and Cross-equation error correlation, which improves estimation efficiency when common shocks affect all equations. Parameter estimation is typically performed using MLE. Thus, it is often preferred due to its asymptotic efficiency and consistency [2, 4].

3. Results and Discussion

This section presents the empirical results of the SUR–SAR modeling of the HDI for 27 regencies and municipalities in West Java Province during 2022–2024, using data on regional income, poverty, education, economic growth, income inequality, gross regional domestic product, and open unemployment. The estimation proceeded sequentially using OLS, SUR, SAR, and SUR–SAR frameworks to assess efficiency, spatial effects, and model fit.

The Ordinary Least Squares (OLS) estimation provides a baseline assessment of the determinants of the HDI in West Java across 2022–2024. The results of the OLS coefficients are presented in Table 1.

According to Table 1, all OLS models demonstrate explanatory power (Adjusted R^2 between 0.875 and 0.893), indicating that the selected socio-economic variables collectively explain a large share of HDI variation across West Java's regions. Poverty (POV) shows a consistently negative and statistically significant effect ($p < 0.05$) in all years, suggesting that higher poverty rates are strongly associated with lower HDI outcomes. Population density (PDS) emerges as the most stable positive variables that significant ($p < 0.05$) across all years, reflecting that more urbanized areas tend to achieve higher HDI. Inequality (GNR) is positive and significant only in 2022. Meanwhile, Consumption level (CLI), GRDP per capita (GPC, log) and minimum wage (LMW, log) display positive yet statistically insignificant effects, implying that income-based indicators alone may not adequately capture differences in HDI across the region, while unemployment (UNE) has negative effects but insignificant throughout the period.

The Seemingly Unrelated Regression (SUR) model jointly estimates HDI equations for 2022–2024, allowing for correlated disturbances across years. This system-based estimation improves efficiency relative to single-year OLS

Table 1. OLS Results for HDI Determinants in West Java (2022–2024)

Variable	2022	2023	2024
(Intercept)	32.354	32.180	9.190
POV (X_1)	-0.445^*	-0.506^*	-0.570^*
NHC (X_2)	-0.0097	-0.0083	-0.0131
PDS (X_3)	0.000498^*	0.000552^*	0.000448^*
GNR (X_4)	19.603^*	14.475	19.509
UNE (X_5)	-0.160	-0.145	-0.090
CLI (X_6)	0.099	0.146	0.313
GPC (log) (X_7)	0.751	0.417	0.557
LMW (log) (X_8)	1.369	1.465	1.844
Adjusted R^2	0.893	0.875	0.886

models and accounts for inter-year linkages in regional development patterns. The results of the SUR coefficients are presented in Table 2.

Table 2. SUR Results for HDI Determinants in West Java (2022–2024)

Variable	2022	2023	2024
(Intercept)	42.847	44.952	41.059
POV (X_1)	-0.375^*	-0.385	-0.412^*
NHC (X_2)	-0.0160	-0.0158	-0.0169
PDS (X_3)	0.000629^*	0.000640^*	0.000597^*
GNR (X_4)	3.497	1.069	2.993
UNE (X_5)	-0.102	-0.084	-0.080
CLI (X_6)	0.031	0.032	0.055
GPC (log) (X_7)	0.886	0.755	0.791
LMW (log) (X_8)	1.304	1.316	1.413
Adjusted R^2	0.8649	0.8624	0.8686

The SUR results in Table 2 demonstrate consistently across all three years (Adjusted R^2 between 0.862 and 0.869), indicating that jointly estimating the equations improves efficiency relative to single-year OLS. Poverty (POV) remains a stable and significant negative determinant of HDI in 2022 and 2024 ($p < 0.05$), highlighting that regions with higher poverty systematically achieve lower HDI outcomes. Population density (PDS) is again the most consistently significant predictor ($p < 0.05$ for all years), confirming that more urbanized districts tend to attain higher HDI. Other socioeconomic factors, including inequality (GNR), unemployment (UNE), consumption level (CLI), GRDP per capita (GPC), and minimum wage (LMW), show positive or negative signs consistent with expectations but remain statistically insignificant, suggesting that their effects on HDI are either weak or overshadowed by structural differences across regions. The strong cross-equation residual correlations (0.98–0.99) further justify the use of SUR, indicating substantial shared shocks among HDI levels over time.

The Moran's I test is applied to examine the presence of spatial autocorrelation in the residuals of the OLS models for 2022–2024. A significant Moran's I statistic indicates that regions with similar HDI levels tend to cluster geographically, violating the assumption of independently distributed errors in OLS regression. The results of the Moran's I test are presented in Table 3.

The Moran's I result in Table 3 reveals that spatial autocorrelation in OLS residuals varies across the study period. In 2022, the Moran's I statistics are small (I = 0.0795) and statistically insignificant (p = 0.111), indicating that HDI prediction errors are spatially random. However, in 2023 and 2024, Moran's I become positive and significant

Table 3. Moran's I Test Results in OLS Residuals (2022–2024)

Year	Moran's I	p-value
2022	0.0795	0.111
2023	0.3806*	6.99×10^{-6}
2024	0.2631*	0.000801

($p < 0.05$), suggesting the presence of positive spatial dependence in the residuals, districts with under- or over-predicted HDI tend to cluster geographically. These findings imply that OLS is insufficient for capturing spatial interactions in 2023–2024, reinforcing the need for spatial econometric models such as SAR, SEM, or SDM to obtain unbiased and efficient estimates.

The Spatial Autoregressive (SAR) model extends the OLS framework by incorporating a spatial lag parameter (ρ) to capture interregional dependence in HDI outcomes. The model was estimated for each year (2022–2024) using MLE with an inverse-distance spatial weight matrix. The results of the SAR coefficients are presented in Table 4.

Table 4. SAR Results for HDI Determinants in West Java (2022–2024)

Variable	2022	2023	2024
(Intercept)	33.597*	36.915*	28.931
POV (X_1)	−0.394*	−0.440*	−0.456*
NHC (X_2)	−0.00145	−0.00058	−0.00516
PDS (X_3)	0.000389*	0.000389*	0.000418*
GNR (X_4)	19.153*	19.279*	13.762
UNE (X_5)	−0.130	−0.124	−0.131
CLI (X_6)	0.037	0.034	0.119
GPC (log) (X_7)	0.645	0.224	0.640
LMW (log) (X_8)	−0.995	−1.331	−0.876
ρ	0.551*	0.642*	0.527*

The SAR results in Table 4 indicate persistent spatial dependence in HDI across West Java, reflected by the significant spatial autoregressive parameter ρ in all years ($p < 0.05$). This confirms HDI spillover effects, where increases in HDI in one district are associated with increases in neighboring districts. Poverty (POV) remains consistently negative and significant ($p < 0.05$), showing that poverty reduction directly enhances local and neighboring HDI levels. Population density (PDS) also shows a stable and strongly positive influence ($p < 0.05$), indicating that more urbanized regions tend to have higher HDI. Inequality (GNR) is significant in 2022 and 2023 but not in 2024, suggesting its effect is less stable over time. Other variable unemployment, health centers, consumption level, GRDP per capita, and minimum wage, remain statistically insignificant, implying that socioeconomic spillovers are not driven by these factors.

Before estimating the SUR–SAR model, Lagrange Multiplier (LM) tests were performed to determine whether spatial effects are present in the system of HDI equations. The tests include LM_{SAR}^{SUR} (for spatial lag dependence) and LM_{SEM}^{SUR} (for spatial error dependence), along with their robust counterparts ($LM{*}_{SAR}^{SUR}$ and $LM{*}_{SEM}^{SUR}$). These diagnostics, proposed by Anselin [2] and extended by Lacambra [19], identify whether spatial interaction should be incorporated into the SUR framework. A significant LM_{SAR}^{SUR} statistic indicates that the SUR–SAR model is the appropriate specification. The result of LM test is presented in Table 5.

According to Table 5, the LM test results indicate that the LM_{SAR}^{SUR} is statistically significant ($p < 0.05$), while the LM_{SEM}^{SUR} statistics are not. This finding implies that spatial dependence arises primarily from the spatially laged dependent variable, not from spatial correlation in the residuals. Therefore, the SUR–SAR model is identified as the most appropriate specification for the HDI system.

Table 5. Lagrange Multiplier (LM) Tests for Spatial Dependence in the SUR System

Test Statistic	Hypothesis	p-value
$\mathbf{LM}^{\mathbf{SUR}}_{\mathbf{SAR}}$	$(H_0 : \rho = 0)$	0.0148^*
$\mathbf{LM}^{\mathbf{SUR}}_{\mathbf{SEM}}$	$(H_0 : \lambda = 0)$	0.2204
$\mathbf{LM}^{*\mathbf{SUR}}_{\mathbf{SAR}}$	$(H_0 : \rho = 0)$ (robust)	0.0915
$\mathbf{LM}^{*\mathbf{SUR}}_{\mathbf{SEM}}$	$(H_0 : \lambda = 0)$ (robust)	0.9469

The SUR–SAR model, estimated through MLE, combines spatial lag dependence with cross-equation correlation to improve efficiency in modeling HDI across multiple years. The spatial lag parameters (ρ) capture interregional effects, while the contemporaneous covariance structure captures correlations among the HDI equations for 2022, 2023, and 2024. This specification addresses both spatial and temporal dependencies simultaneously. The results of the SUR-SAR coefficients are presented in Table 6.

Table 6. SUR-SAR Results for HDI Determinants in West Java (2022–2024)

Variable	2022	2023	2024
(Intercept)	13.485	13.668	14.543
POV	-0.023	-0.057	-0.051
NHC	-0.009	-0.006	-0.007
PDS	0.000654^*	0.000616^*	0.000609^*
GNR	-2.525^*	0.952	0.473
UNE	0.017	0.017	0.033
CLI	0.0004	0.019^*	0.0099
GPC (log)	0.980	0.795	0.843
LMW (log)	1.657^*	1.544^*	1.606^*
ρ	0.325^*	0.339^*	0.324^*
R^2 (Equation)	0.891	0.9066	0.9059

Table 7. System Diagnostics for the SUR–SAR Model

Diagnostic Test	Statistic
Pooled R^2	0.9024
Breusch–Pagan Test	80.19^* (p = 2.25×10^{-17})
LMM Test (Spatial Dependence)	3.339 (p = 0.339)
Residual Correlations ($\rho_{12},\ \rho_{13},\ \rho_{23}$)	0.998 / 0.996 / 0.999

Table 6 shows that across all years, population density (PDS) consistently emerges as the positive and significant variable ($p < 0.05$), highlighting the advantage of urbanized areas in achieving higher HDI. Consumption level (CLI) becomes significant in 2023, while inequality (GNR) is significant only in 2022. Minimum wage (LMW) is positive and statistically significant across all years ($p < 0.05$), suggesting that higher regional wage standards support improvements in HDI. The spatial lag coefficients ($\rho \approx$ 0.32–0.34) are positive and significant, indicating spatial spillover effects where HDI improvements in one district positively influence neighboring districts. These results confirm that the SUR–SAR model provides an efficient and theoretically consistent representation of HDI dynamics in West Java.

According to Table 7, the SUR–SAR model achieves a strong overall pooled R^2 of 0.9024, indicating excellent explanatory and confirming that the combined spatial–temporal specification effectively accounts for HDI variation across all 27 regencies and cities in West Java. The Breusch–Pagan test is highly significant ($p < 0.05$), demonstrating inter-equation error correlations and validating the use of the SUR framework rather than separate OLS or SAR models. The LMM test shows no additional spatial dependence remaining in the residuals (p = 0.339), confirming that the spatial lag component successfully captures spatial interaction effects. Additionally, the variance–covariance and correlation matrices of the residuals show high correlations between equations (above 0.99), confirming that HDI across years is influenced by persistent, province-wide structural factors rather than localized shocks.

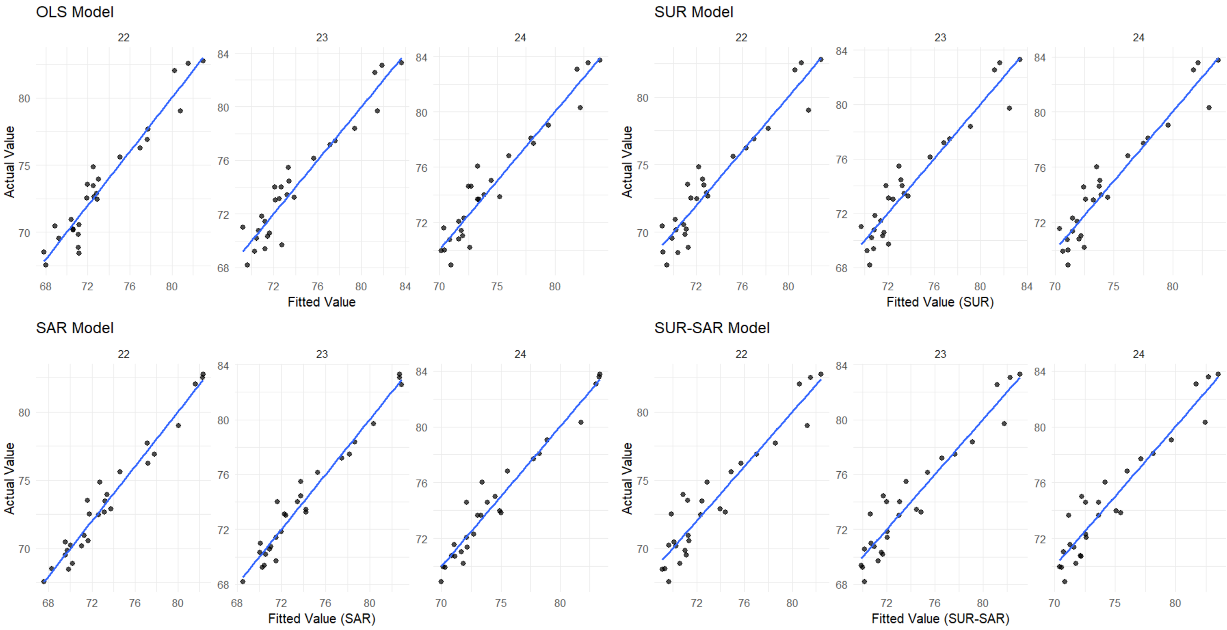

Figure 3. Fitted vs Actual HDI (2022–2024)

The fitted-value plot (Figure 3) shows that the predicted HDI values from the SUR–SAR model aligns closely with the actual observations across all 27 regencies and municipalities for 2022–2024. The fitted lines follow the observed HDI pattern with minimal deviation, confirming the model's excellent in-sample predictive performance. Regions with higher observed HDI scores also exhibit higher fitted values, suggesting that the model accurately captures inter-regional variation in HDI. This alignment demonstrates that the SUR–SAR specification effectively represents both the temporal and spatial structure of HDI dynamics in West Java.

Table 8 provides a comparative overview of significant variables across OLS, SUR, SAR, and SUR–SAR models for the 2022–2024 period, revealing several consistent patterns in the determinants of HDI in West Java. Across all model families, population density (PDS) emerges as the most stable and highly significant predictor ($p < 0.05$), reinforcing the notion that more urbanized districts benefit from better access to education, healthcare, and economic opportunities, which translate into higher HDI outcomes. Poverty (POV) shows a persistent negative effect in OLS, SUR, and especially in the SAR models, where the coefficients become significant ($p < 0.05$), highlighting the strong and direct influence of poverty reduction on HDI.

The role of income-related factors, particularly minimum wage (LMW) and GRDP per capita (GPC) varies across models. While both appear statistically insignificant in OLS and SUR, LMW becomes consistently significant in the SUR–SAR model, that once spatial dependence and inter-equation correlations are accounted, wage standards are revealed as an important structural determinant of HDI. This indicates the presence of spatial spillover effects, increases in wage levels in one district may indirectly stimulate improvements in bordering areas. In contrast, GNR (inequality) exhibits mixed effects, positive in OLS and SAR but negative in SUR–SAR for 2022,

Table 8. Summary of Significant Variables Across Models (2022–2024)

Variable	OLS			SUR			SAR			SUR-SAR		
	2022	2023	2024	2022	2023	2024	2022	2023	2024	2022	2023	2024
POV	*	*	*	*		*	*	*	*			
NHC												
PDS	*	*	*	*	*	*	*	*	*	*	*	*
GNR	*						*	*		*		
UNE												
CLI											*	
GPC (log)												
LMW (log)										*	*	*
ρ							*	*	*	*	*	*

reflecting differences in how each model treats error structures and spatial interactions. This is in line with research conducted by Wibowo [25].

The SAR and SUR–SAR models also highlight the significance of the spatial lag parameter (ρ), which is positive and highly significant across all years, indicating that HDI levels exhibit strong spatial clustering, districts with high HDI tend to be surrounded by districts with similarly high HDI. This spatial spillover further justifies the shift from traditional OLS and SUR models toward spatial specifications.

Collectively, Table 8 demonstrates that while classic models (OLS, SUR) identify PDS and POV as the dominant drivers of HDI, the inclusion of spatial dependence in SAR and SUR–SAR models uncover deeper structural influences, particularly those related to wage policy and spatial interactions. The SUR–SAR specification therefore provides the most comprehensive representation of HDI dynamics, integrating both spatial spillovers and cross-equation correlations to yield more robust and policy-relevant insights.

4. Conclusion

This study successfully addresses its three research objectives. First, it specifies and estimates a SUR–SAR model for the HDI across 27 regencies and municipalities in West Java for the 2022–2024 period, incorporating eight key socio-economic variables such as poverty, health facility availability, population density, inequality, unemployment, consumption levels, GRDP per capita, and regional minimum wage. Second, the application of MLE ensures statistical efficiency and methodological rigor, fully aligned with established spatial econometric standards. Third, the comparative assessment of OLS, SUR, SAR, and SUR–SAR models demonstrates that the SUR–SAR framework provides the most comprehensive specification, effectively capturing both spatial dependence and inter-year correlations.

The empirical findings reveal that population density (PDS) consistently emerges as the most influential and statistically robust determinant, while poverty (POV) exhibits stable negative effects across multiple models. The regional minimum wage (LMW) becomes significant only after accounting for spatial interactions and inter-equation correlations in the SUR–SAR framework, indicating its strategic role in broader regional development dynamics. Meanwhile, inequality, health centers, unemployment, consumption levels, and GRDP per capita show varying effects depending on model specification. The significant spatial lag parameter confirms meaningful HDI spillover effects across neighboring districts, emphasizing that regional development is shaped not only by local conditions but also by spatial interdependencies.

Despite these contributions, the study remains subject to several limitations. Its temporal scope is restricted to three years and one province, limiting broader generalization. The spatial weight matrix is based solely on economic distance, potentially omitting alternative spatial linkages such as transportation, demographic interaction, or administrative boundaries. Furthermore, reliance on secondary data prevents inclusion of additional relevant factors such as governance quality, health service capacity, and digital infrastructure. Future research may extend temporal coverage, incorporate multi-dimensional spatial weight matrices, or adopt Bayesian and dynamic SUR–SAR approaches to capture evolving spatial and temporal development patterns more comprehensively.

Acknowledgement

The authors would like to thank the West Java Provincial Government for providing access to the Open Data Jabar portal, and to the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

1. Afolayan, R. B., Banjoko, A. W., Garba, M. K., & Yahya, W. B. (2020). On Seemingly Unrelated Regression and Single Equation Estimators Under Heteroscedastic Error and Non-Gaussian Responses. *FUOYE Journal of Engineering and Technology*, *5*(2).
2. Anselin, L. (1988a). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
3. Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1).
4. Baltagi, B. H., & Pirotte, A. (2011). Seemingly unrelated regressions with spatial error components. *Empirical Economics*, *40*(1).
5. Bartels, R., & Fiebig, D. G. (1992). Efficiency of Alternative Estimators in Generalized Seemingly Unrelated Regression Models. In *Contributions to Consumer Demand and Econometrics*.
6. Bera, A. K., & Yoon, M. J. (1993). Specification testing with locally misspecified alternatives. *Econometric Theory*, *9*(4).
7. Statistics Indonesia. (2025). Indeks Pembangunan Manusia Indonesia 2024. *Badan Pusat Statistik*, *19*(73).
8. Cruz, N. A., Romero, D. A., & Melo, O. O. (2025). *SAR models with specific spatial coefficients and heteroskedastic innovations*.
9. Getis, A., & Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis*, *36*(2).
10. Hubert, M., Verdonck, T., & Yorulmaz, Ö. (2017). Fast robust SUR with economical and actuarial applications. *Statistical Analysis and Data Mining*, *10*(2).
11. Jesús, M., Fernando, L., & Herrera, M. (2010). Testing for spatial effects in seemingly unrelated regressions. *Spatial Economic Analysis*, *5*(4).
12. Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, *17*(1).
13. Lang, S., Adebayo, S. B., Fahrmeir, L., & Steiner, W. J. (2003). Bayesian geoadditive seemingly unrelated regression. *Computational Statistics*, *18*(2).
14. Liang, X., Gao, J., & Gong, X. (2022). Semiparametric Spatial Autoregressive Panel Data Model with Fixed Effects and Time-Varying Coefficients. *Journal of Business and Economic Statistics*, *40*(4).
15. López, F. A., Mur, J., & Angulo, A. (2014). Spatial model selection strategies in a SUR framework. The case of regional productivity in EU. *Annals of Regional Science*, *53*(1).
16. Maekawa, K., Setiawan, K., & Mada, U. G. (2014). Estimation of vector error correction model with garch errors: Monte Carlo Simulation And Application. *International Conference on Economic Modeling- Ecomod 2014*.
17. Mendes, D. A., & Mendes, V. M. (2015). Parametric Models in Spatial Econometrics: A Survey. In *Dynamic Modeling and Econometrics in Economics and Finance* (Vol. 19).
18. Mur, J., & Angulo, A. (2009). Model selection strategies in a spatial setting: Some additional results. *Regional Science and Urban Economics*, *39*(2).
19. Mur Lacambra, J., & López Hernández, F. A. (2010). *Spatial SUR models: Specification, Testing and Selection*.
20. Olamide, E., Olamide, E. I., & Adepoju, A. A. (2013). Estimating Seemingly Unrelated Regressions with First Order Autoregressive Disturbances. *Studies in Mathematical Sciences*, *6*(2), 40–57.
21. Olayemi, O. S. (2020a). Effects of Correlation between the Error Term and Autocorrelation on Some Estimators in a System of Regression Equations. *Global Journal of Science Frontier Research*.
22. Olayemi, O. S. (2020b). Effects of Multicollinearity and Correlation between the Error Terms on Some Estimators in a System of Regression Equations. *Global Journal of Science Frontier Research*, 77–94.
23. Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, *70*(349).
24. Tobler, W. (2004). On the first law of geography: A reply. In *Annals of the Association of American Geographers* (Vol. 94, Issue 2).
25. Wibowo, D. A., Hidajat, M. S., & Widyatmoko, W. (2023). Poverty Modeling in East Java Province Using the Spatial Seemingly Unrelated Regression (Sur) Method. *Journal of Applied Intelligent System*, *8*(2).
26. Wickens, M. R., & Kmenta, J. (1972). Elements of Econometrics. *Economica*, *39*(154).
27. Wrigley, N., Cliff, A. D., & Ord, J. K. (1982). Spatial Processes: Models and Applications. *The Geographical Journal*, *148*(3).
28. Xu, K., Sun, L., Liu, J., Zhu, X., & Wang, H. (2020). A spatial autoregression model with time-varying coefficients. *Statistics and Its Interface*, *13*(2).
29. Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, *57*(298).
30. Zhang, Z., Li, Z., & Song, Y. (2024). On ignoring the heterogeneity in spatial autocorrelation: consequences and solutions. *International Journal of Geographical Information Science*, *38*(12), 2545–2571.