Spatial autoregressive stochastic frontier model with application to Indonesia’s aquaculture

R Zulkarnain\(^1\)*, A Djuraidah\(^2\) and I M Sumertajaya\(^2\)

\(^1\)BPS Statistics Indonesia, Jakarta, 10710, Indonesia
\(^2\)Department of Statistics, IPB University, Bogor, 16680, Indonesia

*E-mail: zulqarnaen@bps.go.id

**Abstract.** Stochastic frontier analysis (SFA) is one of the popular methods for measuring efficiency and productivity. SFA generally assumes that decision making units are independent. However, this assumption may be inappropriate since decision making units are probable interdependent depending on their spatial proximities. This paper aims to incorporate spatial lag into standard SFA model. The performance of spatial autoregressive stochastic frontier model (SAR-SFA) is compared to standard SFA using simulation and real data. Simulation study is undertaken using various degrees of spatial autocorrelation. Furthermore, SAR-SFA is applied to estimate technical efficiency of aquaculture in Indonesia. Simulation results showed that the SAR-SFA generally produces better performance than the standard SFA. If spatial dependencies are significant, then the estimation of production frontier from standard SFA could be misleading. It affects the distribution of technical efficiency and its ranking. Moran’s I test indicated that Indonesia’s aquaculture has significant positive spatial autocorrelation. If standard SFA is applied to estimate technical efficiency of Indonesia’s aquaculture, then the resulted technical efficiency is implausible. This issue is well handled by the SAR-SFA model. Moreover, SAR-SFA model serves procedure to estimate efficiency spillovers.

1. **Introduction**

Various methods have been proposed for measuring productivity. Stochastic frontier analysis (SFA) is one of the alternative methods that uses regression techniques to estimate productivity. The primary feature of SFA is that the error has two components, one component to explain inefficiency and another one to explain noise or random shocks such as engine failure and bad weather \([1,2]\). SFA generally assumes that inefficiency and noise terms between decision making units (DMUs) are independent \([3]\). However, the non-autocorrelation assumptions in SFA may be inappropriate. DMUs could be interdependent via supply-chain networks or various externalities. Economic activities in a specific area attracts labor and intermediate goods in its neighbor areas or vice versa. Moreover, spatial impedance between DMUs causes the knowledge spillovers.

Fusco \([4]\) stated that if spatial autocorrelation is significant, then conventional method used to estimate the parameters of SFA can produce considerable bias. In order to deal with these issues, several studies have been conducted to incorporate spatial lag into standard SFA model. Druska and Horrace \([5]\) integrated stochastic frontier model and spatial error model using generalized moments method. Fusco and Vidoli \([6]\) incorporated spatial lag of inefficiency term into standard SFA and estimating efficiency using maximum likelihood method. Adetutu et al. \([7]\) accounted for spatial dependencies by allowing spatial lag of inputs and exogenous variables into SFA model. Glass et al. \([8,9,10]\) proposed a
stochastic frontier model with spatial autoregressive structure. Spatial autoregressive stochastic frontier model has another advantage as it allows to study efficiency spillovers. Tsukamoto [11] developed a spatial autoregressive stochastic frontier model that can simultaneously estimates the production frontier and determinants of inefficiency.

This paper employs spatial autoregressive stochastic frontier model to estimate production frontier and technical efficiency using simulation and real data. Simulation is conducted to evaluate the performance of spatial autoregressive stochastic frontier model compared to non-spatial stochastic frontier model using various degrees of spatial autocorrelation. Furthermore, spatial autoregressive stochastic frontier model is applied to measure technical efficiency of Indonesia’s aquaculture. Efficiency spillovers for all provinces are also derived to determine whether specific province is net efficiency importer or net efficiency exporter.

2. Methodology

2.1. Stochastic Frontier Model

Stochastic frontier model was proposed by Aigner et al. [1] and Meeeusen and van Den Broeck [2] at the same time. Stochastic frontier model can be expressed as follows:

\[ \tilde{y}_i = f(\tilde{x}_i; \beta) + v_i - u_i ; i = 1, 2, ..., n \]

where \( \tilde{y}_i \) is the output of \( i \)th DMUs in natural logarithmic form, \( \tilde{x}_i \) is the vector of inputs of \( i \)th DMUs in natural logarithmic form, \( \beta \) is the vector of parameters, \( f(\tilde{x}_i; \beta) \) is the specified production function, \( v_i \sim iid N(0, \sigma_v^2) \) is noise term, and \( u_i \sim iid N^+(0, \sigma_u^2) \) is inefficiency term. The error structure can be written as \( \varepsilon_i = v_i - u_i \) and the log-likelihood function is given by [3]:

\[ LL = -\frac{n}{2} \ln \left( \frac{\pi \sigma^2}{2} \right) + \sum_{i=1}^{n} \ln \left[ \phi \left( -\frac{\lambda \varepsilon_i}{\sigma} \right) \right] - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \varepsilon_i^2 \]

where \( \sigma = \sqrt{\sigma_u^2 + \sigma_v^2} \), \( \lambda = \sigma_u / \sigma_v \), and \( \phi(.) \) is the cumulative distribution function of the standard normal distribution. The parameter \( \lambda \) is called inefficiency parameter. If \( \lambda \to 0 \) then the deviance to the frontier is caused by inefficiency. Conversely, if \( \lambda \to \infty \) then the deviance to the frontier is caused by noise and the model can be estimated by ordinary least squares.

Finally, technical efficiency score (TE score) for each DMUs can be estimated as follows [12]:

\[ E(u_i | \varepsilon_i) = \mu_i + \sigma_s \left( \frac{\phi(-\mu_i / \sigma_s)}{\phi(\mu_i / \sigma_s)} \right) \]

\[ TE_i = \exp \left(-E(u_i | \varepsilon_i)\right) \]

where \( \mu_i = -\varepsilon_i \sigma_n^2 / \sigma^2 \), \( \sigma_s = \sigma_u \sigma_v / \sigma \), and \( \phi(.) \) is the density function of the standard normal distribution.

2.2. Spatial Autoregressive Stochastic Frontier Model

This paper applies spatial autoregressive stochastic frontier model proposed by Glass et al. [10] for cross-section data as follows:

\[ \tilde{y}_i = f(\tilde{x}_i; \beta) + \rho \sum_{j=1}^{n} w_{ij} \tilde{y}_j + v_i - u_i \]

where \( \tilde{y}_i \) is the output of \( i \)th DMUs in natural logarithmic form, \( \tilde{x}_i \) is the vector of inputs of \( i \)th DMUs in natural logarithmic form, \( \beta \) is the vector of parameters, \( f(\tilde{x}_i; \beta) \) is the specified production function, \( \rho \) is spatial lag parameter, \( v_i \sim iid N(0, \sigma_v^2) \) is noise term, \( u_i \sim iid N^+(0, \sigma_u^2) \) is inefficiency term, and \( w_{ij} \) is the spatial weights. This paper uses threshold method to construct spatial weights where the distances are calculated using Great Circle Distance (GCD). The formula of GCD and threshold method are as below [13,14]:

\[ d_{ij} = r \times \arccos \left[ \cos |a_i - a_j| \cos(b_i) \cos(b_j) + \sin(b_i) \sin(b_j) \right] \]
\[ w_{ij} = \begin{cases} 1 & \text{if } 0 < d_{ij} \leq d_{\max} \\ 0 & \text{if } d_{ij} > d_{\max} \end{cases} \]

where \( r \) is the Earth’s radius, \((a_i, b_i)\) and \((a_j, b_j)\) are the respective coordinates. The value of \( r \) is 3959 if the distance in miles and \( r \) is 6371 if the distance in kilometres. The optimal upper distance bound \( d_{\max} \) is chosen when it produces the smallest p-value of Moran’s I test.

The log-likelihood function can be specified as [10]:

\[
LL = -\frac{n}{2} \ln \left( \frac{n \sigma^2}{2} \right) + \ln |I_n - \rho W| + \sum_{i=1}^{n} \ln \left( \Phi \left( -\frac{\lambda \tilde{e}_i}{\sigma} \right) \right) - \frac{1}{2 \sigma^2} \sum_{i=1}^{n} \tilde{e}_i^2
\]

where \( \sigma = \sqrt{\sigma_u^2 + \sigma_v^2} \), \( \lambda = \sigma_u / \sigma_v \), \( \Phi(.) \) is the cumulative distribution function of the standard normal distribution, \( I_n \) is the identity matrix, \( \rho \) is spatial lag parameter, and \( W \) is spatial weights matrix.

Following Fan et al. [15] approach, spatial autoregressive stochastic frontier model can be rewritten as \( \tilde{y}_i \equiv E(\tilde{y} | \beta, \rho) + g(\lambda, \sigma^2) + \epsilon_i \) where \( g(\lambda, \sigma^2) = (\sqrt{2} \sigma \lambda / [\pi (1 + \lambda^2)])^{1/2} \) and the spatial lag parameter (\( \rho \)) can be estimated using following concentrated log-likelihood function:

\[
LL_c = \Omega - \frac{n}{2} \ln \left( (e_0 - \rho e_1)'(e_0 - \rho e_1) \right) + \ln |I_n - \rho W|
\]

where \( \Omega \) is a constant, \( e_0 \) is residual of ordinary least squares from regressing \( \tilde{y} \) and \( \tilde{X} \), \( e_1 \) is residual of ordinary least squares from regressing \( W \tilde{y} \) and \( \tilde{X} \), \( I_n \) is identity matrix, \( \rho \) is spatial lag parameter, and \( W \) is spatial weights matrix.

Using the resulted estimation of spatial parameter (\( \hat{\rho} \)), the production function parameters (\( \beta \)) can be estimated as:

\[
\hat{\beta} = (\tilde{X}^{\prime} \tilde{X})^{-1} \tilde{X}^{\prime}[\tilde{y} - \hat{\rho} W \tilde{y}]
\]

Furthermore, using pseudo-likelihood method, the inefficiency parameter (\( \lambda \)) is estimated by maximizing the following log-likelihood function:

\[
LL = -n \ln \sigma + \sum_{i=1}^{n} \ln \left( \Phi \left( -\frac{\lambda \tilde{e}_i}{\sigma} \right) \right) - \frac{1}{2 \sigma^2} \sum_{i=1}^{n} \tilde{e}_i^2
\]

where \( \tilde{e}_i = \tilde{y}_i - E(\tilde{y} | \beta, \rho) - g(\lambda, \sigma^2) \) and \( \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (\tilde{y}_i - E(\tilde{y} | \beta, \rho))^2 / [1 - 2 \lambda^2 / \pi (1 + \lambda^2)] \). The solution of above log-likelihood function is the pseudo-likelihood estimate of inefficiency parameter (\( \hat{\lambda} \)). By substituting \( \hat{\lambda} \), the pseudo-likelihood estimates for \( \hat{\sigma}^2 \) and \( \hat{\epsilon}_i \) are also derived. Therefore, the estimates for \( \sigma_u^2 \) and \( \sigma_v^2 \) are calculated as below:

\[
\hat{\sigma}_u^2 = \frac{\hat{\sigma}^2 \hat{\lambda}^2}{1 + \hat{\lambda}^2}
\]
\[
\hat{\sigma}_v^2 = \frac{\hat{\sigma}^2}{1 + \hat{\lambda}^2}
\]

Following Jondrow et al. [12], the technical efficiency score for each DMUs can be estimated as:

\[
\hat{u}_i = \frac{\hat{\sigma}_u \hat{\sigma}_v}{\hat{\sigma}} \left( \frac{\phi(\hat{\lambda} \hat{\epsilon}_i / \hat{\sigma}) - \hat{\lambda} \hat{\epsilon}_i}{\phi(-\hat{\lambda} \hat{\epsilon}_i / \hat{\sigma})} \right)
\]

\[
\xi_i = TE_i = \exp(-\hat{u}_i)
\]

where \( \phi(.) \) is the density function of the standard normal distribution and \( \Phi(.) \) is the cumulative distribution function of the standard normal distribution.

Glass et al. [10] introduced the concept of efficiency spillover using spatial autoregressive stochastic frontier model. The relative direct, relative indirect and relative total efficiencies are calculated as below:

\[
(I_n - \rho W)^{-1} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \xi_{11}^{Tot} \\ \xi_{12}^{Tot} \\ \vdots \\ \xi_{nn}^{Tot} \end{pmatrix} = \begin{pmatrix} \xi_{11}^{Ind} + \xi_{12}^{Ind} + \cdots + \xi_{1n}^{Ind} \\ \xi_{21}^{Ind} + \xi_{22}^{Ind} + \cdots + \xi_{2n}^{Ind} \\ \vdots \\ \xi_{n1}^{Ind} + \xi_{n2}^{Ind} + \cdots + \xi_{nn}^{Ind} \end{pmatrix}
\]
where $RE_i^{dir}$ is efficiency spillover to the $i$th area from all other areas and $RE_i^{ind}$ is efficiency spillover from the $j$th area to all other areas.

2.3. Design of Simulation

This study uses coordinates of 34 provinces in Indonesia as spatial layout. The spatial weights matrix ($W$) is constructed using threshold method where the distances are calculated using Great Circle Distance (GCD). The data for input variable ($x_i$) are generated from uniform distribution with interval (0, 10). Inefficiency terms ($u_i$) are randomly chosen from half normal distribution $N^+(0,1)$ and noise components ($v_i$) are randomly drawn from normal distribution $N(0, 0.001^2)$. Furthermore, the intercept and slope are set to be $\beta_0 = 2$ and $\beta_1 = 0.3$, while spatial lag parameter ($\rho$) is simulated from 0.1 to 0.9 to reflect various degrees of spatial autocorrelation. Finally, the output variable is calculated as follows:

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = (I_n - \rho W)^{-1} [\beta_0 + \beta_1 \ln x + v - u]$$

For each setting, the calibration process is replicated 100 times with new sample of $u_i$ and $v_i$. The parameters $\beta_0$, $\beta_1$, $\sigma_u^2$, $\sigma_v^2$, input $x_i$, and spatial weights matrix $W$ are fixed for each replication. To evaluate the performance of the models, this study uses Root Mean Squared Error (RMSE) criteria as follows:

$$RMSE(\hat{\theta}) = \left[ \frac{1}{N} \sum_{r=1}^{N} (\hat{\theta}^{(r)} - \theta)^2 \right]^{1/2}$$

where $\hat{\theta}$ is the corresponding estimate to be evaluated and $N$ is the number of replications.

2.4. Application to Indonesia’s Aquaculture

Data was acquired for Indonesia’s aquaculture from the Ministry of Marine Affairs and Fisheries Republic of Indonesia. Units of analysis are 34 provinces in Indonesia. There are two input variables and one output variable that are used in this study. Table 1 lists the input and output variables.

| No. | Input/output | Variable | Description | Unit       |
|-----|--------------|----------|-------------|------------|
| 1   | Output       | $Y$      | Volume of production | Tonnes     |
| 2   | Input        | $X_1$    | Area of aquaculture | Hectares   |
| 3   | Input        | $X_2$    | Number of aquaculture | Unit of producers producers |

This study uses Cobb-Douglas (CD) production function for functional form. Thus, the stochastic frontier model could be written as:

$$\ln \left( \frac{y_i}{T_i} \right) = x_i' \beta + \ln \left( \frac{u_i}{T_i} \right) + v_i$$

where $u_i$ is inefficiency terms, $v_i$ is noise components, $y_i$ is output variable, $T_i$ is total production, $x_i$ is input variable, $\beta$ is parameters.
• Non-spatial stochastic frontier model (SFA):
\[ \ln y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + v_i - u_i \]

• Spatial autoregressive stochastic frontier model (SAR-SFA):
\[ \ln y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \rho \sum_{j=1}^{n} w_{ij} (\ln y_j) + v_i - u_i \]

where \( y_i \) is the aquaculture production of \( i \)th province, \( X_{1i} \) is the area of aquaculture of \( i \)th province, \( X_{2i} \) is the number of aquaculture producers of \( i \)th province, \( v_i \sim iid N(0, \sigma_v^2) \) is the noise term, \( u_i \sim iid N^+(0, \sigma_u^2) \) is the inefficiency term, and \( w_{ij} \) is the spatial weights.

3. Results

3.1. Simulation Results

Table 2 presents the comparison of RMSE between standard stochastic frontier model (SFA) and spatial autoregressive stochastic frontier model (SAR-SFA) for various values of \( \rho \). For intercept (\( \hat{\beta}_0 \)), RMSE of SFA is smaller than SAR-SFA until \( \rho = 0.3 \), but become larger since \( \rho = 0.4 \). If the data has high spatial autocorrelation, then the deviance of SFA estimate to the true intercept is substantial. For \( \hat{\beta}_1, \hat{\sigma}_v^2, \hat{\sigma}_u^2, \) and \( \hat{\lambda} ^2 \), RMSE of SFA is not much different with RMSE of SAR-SFA if the data has low spatial autocorrelation. If spatial autocorrelation is considerable (\( \rho = 0.8 \) and \( \rho = 0.9 \)), then RMSE of SAR-SFA is much smaller than the SFA. Interestingly, for inefficiency parameter (\( \hat{\lambda} \)). RMSE of SAR-SFA is always smaller than the SFA.

|          | RMSE   | \( \rho \) | \( 0.1 \) | \( 0.2 \) | \( 0.3 \) | \( 0.4 \) | \( 0.5 \) | \( 0.6 \) | \( 0.7 \) | \( 0.8 \) | \( 0.9 \) |
|----------|--------|------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| SFA      | \( \hat{\beta}_0 \) | 0.17 | 0.37 | 0.63 | 1.04 | 1.55 | 2.35 | 3.68 | 6.37 | 14.35 |
|          | \( \hat{\beta}_1 \) | 0.09 | 0.08 | 0.09 | 0.09 | 0.10 | 0.11 | 0.13 | 0.15 | 0.18 |
|          | \( \hat{\sigma}_v^2 \) | 0.31 | 0.32 | 0.33 | 0.32 | 0.33 | 0.36 | 0.42 | 0.58 | 0.96 |
|          | \( \hat{\sigma}_u^2 \) | 0.02 | 0.04 | 0.07 | 0.04 | 0.04 | 0.07 | 0.10 | 0.15 | 0.26 |
|          | \( \hat{\lambda} \) | 6241.42 | 6883.37 | 6689.48 | 6468.33 | 5382.12 | 5566.60 | 5115.42 | 5264.95 | 5405.95 |
| SAR-SFA  | \( \hat{\beta}_0 \) | 0.75 | 0.79 | 0.81 | 0.91 | 1.07 | 1.33 | 1.77 | 2.66 | 5.24 |
|          | \( \hat{\beta}_1 \) | 0.13 | 0.11 | 0.12 | 0.11 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
|          | \( \hat{\sigma}_v^2 \) | 0.34 | 0.34 | 0.37 | 0.34 | 0.34 | 0.34 | 0.35 | 0.35 | 0.35 |
|          | \( \hat{\sigma}_u^2 \) | 0.04 | 0.04 | 0.07 | 0.04 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 |
|          | \( \hat{\lambda} \) | 1046.12 | 1370.57 | 1028.39 | 1033.66 | 1473.50 | 965.54 | 962.15 | 983.48 | 1181.08 |

Spatial autocorrelation affects the estimate of \( \hat{\beta}_0 \) and hence the production frontier considerably. Figure 1 shows that the estimate of \( \hat{\beta}_0 \) from standard SFA tends to be biased as its expected value deviates from the true parameter (red dashed line). The bias become larger as the spatial autocorrelation increases. On the other hand, the estimate of \( \hat{\beta}_0 \) from SAR-SFA tends to be unbiased. The estimate of \( \beta_0 \) from SAR-SFA model has higher probability to be closed to the true parameter as the spatial autocorrelation increases. The estimate of \( \hat{\beta}_0 \) in stochastic frontier model is important since it reflects the technology index. The estimate of \( \beta_0 \) determines the position of production frontier and hence affects the technical efficiency scores.
Figure 1. Distribution of $\beta_0$ estimates from SFA and SAR-SFA models

Figure 2 depicts the comparison of production frontier estimates between SFA and SAR-SFA for various $\rho$. If spatial autocorrelation is insignificant, then SFA and SAR-SFA frontiers are not much different. Both are close but below the true frontier line. If spatial autocorrelation is highly significant, that is if $\rho = 0.8$ or $\rho = 0.9$, then the deviance of SFA frontier to the true frontier are larger. Meanwhile, SAR-SFA frontier is closer to the true frontier. This affects the technical efficiency score of each DMUs. All DMUs are in SFA frontier line if spatial autocorrelation is highly significant. As a result, all DMUs are recorded as fully efficient if standard SFA is employed. Conversely, SAR-SFA frontiers are always above the DMUs.

The distribution of technical efficiency scores of SFA and SAR-SFA for any degrees of spatial autocorrelation ($\rho$) are presented in Figure 3. For the range of $\rho$ from 0.1 to 0.6, the technical efficiency scores of SFA and SAR-SFA are not substantial different in median and interquartile range. For $\rho = 0.7$, the interquartile range of SFA is much smaller than the SAR-SFA, but with a slightly different in median. If spatial autocorrelation is very high, represented by $\rho = 0.8$ and $\rho = 0.9$, then the difference of technical efficiency scores between SFA and SAR-SFA are large. SFA produces high technical efficiency scores for all DMUs and hence are recorded as fully efficient. As a result, the median of technical efficiency scores of SFA is larger than the SAR-SFA, while the interquartile range of SFA is smaller than the SAR-SFA. The technical efficiency scores of SAR-SFA are relatively persistent for any degrees of spatial autocorrelation.
Figure 2. Estimation of production frontier of SFA and SAR-SFA for various $\rho$

Figure 3. Boxplot of technical efficiency of SFA and SAR-SFA for various $\rho$
3.2. Application to Indonesia’s Aquaculture

Spatial distribution of Indonesia’s aquaculture productions is shown in Figure 4. Some provinces in the middle of Indonesia such as South Sulawesi, Central Sulawesi, Southeast Sulawesi, East Nusa Tenggara, and West Nusa Tenggara along with some provinces in Java such as West Java and East Java are the largest aquaculture producers in Indonesia. Provinces in the west side of Indonesia are generally moderate aquaculture producers, while provinces in the east of Indonesia (Papua and West Papua) are small aquaculture producers.

Figure 4. Spatial distribution of aquaculture in Indonesia

This study chooses optimal upper distance bound before calculating spatial weights using threshold method. Optimal upper distance bound is chosen when it produces smallest p-value of Moran’s I test. Figure 5 presents p-value of Moran’s I test for various upper distance bound. Smallest p-value is produced when upper distance bound is 1000 kilometres. Figure 6 depicts Moran’s scatterplot of aquaculture production in Indonesia using optimal upper distance bound. The resulted p-value is 0.0017, thus indicates that there is significant positive spatial autocorrelation in Indonesia’s aquaculture.

Table 3 presents the comparison of estimation results using SFA and SAR-SFA models. Both are resulting insignificant intercept. Coefficient of number of producers ($\hat{\beta}_2$) is not significant in SFA model but become significant in SAR-SFA model. Moreover, $\hat{\sigma}_u^2$ is insignificant in SFA model but is significant in SAR-SFA model. Conversely, $\hat{\sigma}_v^2$ is significant in SFA model but is insignificant in SAR-SFA model. This indicates that in SFA model the deviance to the frontier is caused by noise, while in SAR-SFA model the deviance to the frontier is caused by inefficiency. The estimate of spatial lag
parameter ($\rho$) in SAR-SFA is quite large (0.6968), thus indicates that the spatial autocorrelation is significant. It is consistent with the result of Moran’s I test. SAR-SFA model produces better performance than the SFA model, indicated by smaller Akaike Information Criterion (AIC).

Table 3. Estimation results

|              | SFA  | SAR-SFA |
|--------------|------|---------|
| Coefficient |      |         |
| $\beta_0$   | 2.9068 | 3.9761  |
| $\beta_1$   | 0.6234 | 0.4747  |
| $\beta_2$   | 0.2933 | 0.4384  |
| $\sigma_u^2$| 0.0003 | 0.0426  |
| $\sigma_v^2$| 0.8829 | 0.3216  |
| $\sigma_\gamma^2$| 0.8832 | 0.0233**|
| $\lambda$   | 0.0191 | 1.7825  |
| $\rho$      | -     | 0.3072  |

LL: -46.1292, AIC: 102.2583

Statistical significance at the: *10% level, **5% level, ***1% level

Input coefficients in Cobb-Douglas production function could be interpreted as elasticities. However, coefficients in SAR-SFA model should be interpreted carefully since the marginal effect of inputs is a function of the SAR variable. Therefore, elasticities in SAR-SFA model are calculated from direct effect only. Table 4 presents the elasticities for each input from SFA and SAR-SFA models. In SFA model, elasticity of area of aquaculture is 0.6234 (if area of aquaculture increases in 1%, aquaculture production will increase 0.62%). Elasticity of number of aquaculture producers is 0.2933 (if number of aquaculture producers increases in 1%, aquaculture production will increase 0.29%). Therefore, in SFA model, the coefficient of return to scale is 0.9167 (decreasing return to scale). Elasticity of area of aquaculture from SAR-SFA model is lower than SFA model, that is 0.5237 (if area of aquaculture increases in 1%, aquaculture production will increase 0.52%). On the other side, elasticity of number of aquaculture producers from SAR-SFA model is higher than SFA model, that is 0.4836 (if number of aquaculture producers increases in 1%, aquaculture production will increase 0.48%). Therefore, in SAR-SFA model, the coefficient of return to scale is 1.0073 (increasing return to scale).

Table 4. Elasticities of inputs

| Input            | SFA     | SAR-SFA |
|------------------|---------|---------|
| Area of aquaculture | 0.6234 | 0.5237 |
| Number of aquaculture producers | 0.2933 | 0.4836 |
| Return to scale   | 0.9167 | 1.0073 |

Spatial autocorrelation affects the score of technical efficiency. Figure 7 and Figure 8 visualize the distribution of technical efficiency scores of SFA and SAR-SFA, respectively. Technical efficiencies of SFA are distributed in very high scores, while technical efficiencies of SAR-SFA are varies from low to high scores. If traditional SFA method is employed, then all aquaculture producers are recorded as fully efficient. This result is implausible since according to Ministry of Marine Affairs and Fisheries report, the productivity of aquaculture in Indonesia is still low as the majority of aquaculture producers are small scale/traditional with the limitation to capital and technology access. Conversely, SAR-SFA produces more reasonable technical efficiency scores. The average of technical efficiency scores from SAR-SFA model is 0.51. This shows that aquaculture in Indonesia is still inefficient. There is opportunity to enhance aquaculture production up to 49% by eliminating the inefficiency effects.
The comparison of spatial distributions between SFA and SAR-SFA technical efficiency scores are presented in Figure 9 and Figure 10. The figures show that several provinces experience considerable shifting in their ranking after spatial dependencies are considered. The ranking of West Sumatera increases from middle quantile to higher quantile after SAR-SFA is applied. Meanwhile, the ranking of South Kalimantan decreases from middle quantile to low quantile.

The advantage of SAR-SFA is that it enables to study efficiency spillovers. Table 5 shows the relative direct, indirect and total efficiencies of aquaculture in all Indonesian provinces. The average estimate of direct relative efficiency is 0.613. The estimates of $RE^{ind}_i$ are generally higher than the $RE^{ind}_j$, thus
indicates that most provinces are net efficiency importer. There are few provinces that are net exporter of efficiency. They are West Sumatera, Central Sulawesi, South Sulawesi, and Southeast Sulawesi.

4. Conclusion

Stochastic frontier model incorporating spatial effects has better performance than the traditional SFA, especially if spatial autocorrelation is significant. If the data has serious spatial autocorrelation, then the production frontier estimates using traditional SFA could be misleading. As a result, the resulted technical efficiency score is implausible. This issue is well handled by the spatial autoregressive stochastic frontier model (SAR-SFA).

SAR-SFA provides another advantage that it enables to study efficiency spillovers. Application to Indonesia’s aquaculture shows that most provinces in Indonesia are net efficiency importer. However, there are several provinces that are net efficiency exporter such as West Sumatera, Central Sulawesi, South Sulawesi, and Southeast Sulawesi. Finally, this paper recommends that stochastic frontier model incorporating spatial effect should be used to compute technical efficiency because economic activities are likely interdependent depending on their geographical proximities.

| Code | Province            | REi  | REj  | REi  | REj  | REi  | REj  |
|------|---------------------|------|------|------|------|------|------|
| 11   | Aceh                | 0.196| 0.196| 0.865| 0.065| 0.665| 0.110|
| 12   | North Sumatera      | 0.660| 0.660| 0.782| 0.385| 0.815| 0.490|
| 13   | West Sumatera       | 0.872| 0.872| 0.754| 0.754| 0.891| 0.841|
| 14   | Riau                | 0.716| 0.716| 0.769| 0.641| 0.832| 0.691|
| 15   | Jambi               | 0.682| 0.682| 0.767| 0.710| 0.815| 0.728|
| 16   | South Sumatera      | 0.580| 0.580| 0.784| 0.620| 0.781| 0.631|
| 17   | Bengkulu            | 0.519| 0.519| 0.799| 0.437| 0.764| 0.481|
| 18   | Lampung             | 0.480| 0.480| 0.791| 0.464| 0.742| 0.488|
| 19   | Bangka Belitung Island | 0.521| 0.521| 0.783| 0.632| 0.754| 0.620|
| 20   | Riau Island         | 0.677| 0.677| 0.774| 0.476| 0.818| 0.561|
| 21   | Jakarta             | 0.503| 0.503| 0.788| 0.560| 0.749| 0.563|
| 22   | West Java           | 0.686| 0.686| 0.761| 0.702| 0.813| 0.724|
| 23   | Central Java        | 0.477| 0.477| 0.783| 0.489| 0.735| 0.504|
| 24   | Yogyakarta          | 0.573| 0.573| 0.778| 0.543| 0.774| 0.574|
| 25   | East Java           | 0.587| 0.587| 0.788| 0.654| 0.786| 0.658|
| 26   | Banten              | 0.651| 0.651| 0.770| 0.675| 0.803| 0.693|
| 27   | Bali                | 0.748| 0.748| 0.786| 0.737| 0.857| 0.769|
| 28   | West Nusa Tenggara | 0.796| 0.796| 0.810| 0.670| 0.894| 0.737|
| 29   | East Nusa Tenggara | 0.918| 0.918| 0.830| 0.567| 0.962| 0.704|
| 30   | West Kalimantan     | 0.386| 0.386| 0.771| 0.462| 0.687| 0.456|
| 31   | Central Kalimantan  | 0.351| 0.351| 0.807| 0.479| 0.695| 0.456|
| 32   | South Kalimantan    | 0.443| 0.443| 0.806| 0.547| 0.736| 0.535|
| 33   | East Kalimantan     | 0.245| 0.245| 0.818| 0.235| 0.655| 0.247|
| 34   | North Kalimantan    | 0.649| 0.649| 0.778| 0.574| 0.808| 0.621|
| 35   | North Sulawesi      | 0.702| 0.702| 0.832| 0.576| 0.868| 0.640|
| 36   | Central Sulawesi    | 0.680| 0.680| 0.806| 1.000| 0.929| 1.000|
| 37   | South Sulawesi      | 0.680| 0.680| 0.794| 0.828| 0.885| 0.851|
| 38   | Southeast Sulawesi  | 0.819| 0.819| 0.813| 0.992| 0.907| 0.974|
| 39   | Gorontalo           | 0.275| 0.275| 0.865| 0.247| 0.700| 0.266|
| 40   | West Sulawesi       | 0.268| 0.268| 0.831| 0.297| 0.674| 0.299|
| 41   | Maluku              | 0.949| 0.949| 0.867| 0.670| 1.000| 0.788|
| 42   | North Maluku        | 0.938| 0.938| 0.868| 0.662| 0.996| 0.778|
| 43   | West Papua          | 1.000| 1.000| 0.821| 0.538| 0.992| 0.710|
| 44   | Papua               | 0.304| 0.304| 1.000| 0.051| 0.803| 0.136|
References

[1] Aigner D, Lovell CAK and Schmidt P 1977 Formulation and estimation of stochastic frontier production function models J. Econom. 6(1) 21–37
[2] Meeusen W and van Den Broeck J 1977 Efficiency estimation from cobb-douglas production functions with composed error Int. Econ. Rev. 18(2) 435–444
[3] Coelli TJ, Prasada Rao DS, O’Donnell CJ and Battese GE 2005 An Introduction to Efficiency and Productivity Analysis (New York: Springer)
[4] Fusco E 2017 New Developments in Frontier Models for Objective Assessments [Ph.D. thesis] (Rome: Sapienza – University of Rome)
[5] Druska V and Horrace W 2004 Generalized moments estimation for spatial panel data: Indonesian rice farming Amer. J. Agr. Econom. 86 185–198
[6] Fusco E and Vidoli F 2013 Spatial stochastic frontier models: controlling spatial global and local heterogeneity Int. Rev. Appl. Econom. 27(5) 679–694
[7] Adetutu M, Glass AJ, Kenjegalieva K and Sickles RC 2015 The effects of efficiency and TFP growth on pollution in Europe: a multistage spatial analysis J. Product. Anal. 43(3) 307–326
[8] Glass A, Kenjegalieva K and Paez-Farrell J 2013 Productivity growth decomposition using a spatial autoregressive frontier model Econom. Lett. 119(3) 291–295
[9] Glass A, Kenjegalieva K and Sickles RC 2014 Estimating efficiency spillovers with state level evidence for manufacturing in the US Econom. Lett. 123(2) 154–159
[10] Glass AJ, Kenjegalieva K and Sickles RC 2016 A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers J. Econom. 190(2) 289–300
[11] Tsukamoto T 2019 A spatial autoregressive stochastic frontier model for panel data incorporating a model of technical inefficiency Japan World Econom. 50 66–77
[12] Jondrow J, Lovell CAK, Materov IS and Schmidt P 1982 On the estimation of technical inefficiency in the stochastic frontier production function model J. Econom. 19(2–3) 233–238
[13] Kells L, Kern W and Bland J 1940 Plane and Spherical Trigonometry (New York: McGraw-Hill)
[14] Arbia G 2006 Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence (New York: Springer)
[15] Fan Y, Li Q and Weersink A 1996 Semiparametric estimation of stochastic production frontier models J. Bus. Econom. Stat. 14(4) 460–468