BAREKENG: Journal of Mathematics and Its Applications
March 2023 Volume 17 Issue 1 Page 0391–0400
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss1pp0391-0400

ROBUST STOCHASTIC PRODUCTION FRONTIER TO ESTIMATE TECHNICAL EFFICIENCY OF RICE FARMING IN SULAWESI SELATAN

Ismail Pranata[1*], Anik Djuraidah[2], Muhammad Nur Aidi[3]

[1,2,3]Department of Statistics, Faculty Mathematics and Natural Sciences, IPB University
Jl. Raya Darmaga, Bogor, Jawa Barat, 16680, Indonesia

[1]BPS-Statistics Indonesia, Jakarta
Jl. Dr. Sutomo 6-8, Jakarta, 10710, Indonesia

*Corresponding author's e-mail: * **09021984ismail@apps.ipb.ac.id***

ABSTRACT

Article History:

Received: 4th November 2022
Revised: 25th January 2023
Accepted: 8th February 2023

Keywords:

Student's T;
Gamma;
Outlier;
Stochastic Production Frontier;
Stochastic Frontier Analysis

The stochastic production frontier (SPF) is the stochastic frontier analysis (SFA) method used to estimate the production frontier by accounting for the existence of inefficiency. The standard SPF assumes that the noise component follows a Normal distribution and the inefficiency component follows a half-Normal distribution. The presence of outliers in the data will affect the inaccuracy in estimating the parameters, leading to an exaggerated spread of efficiency predictions. This study uses two alternative models, the first with SPF Normal-Gamma and the second with SPF Student's t-half Normal, and then the results are compared with standard SPF. This study used data from Statistics Indonesia on the cost structure of paddy cultivation household survey in 2014. This study examines the effect of changes in distribution assumptions on the standard SPF model in estimating parameter value and the technical efficiency score in the presence of outliers. The parameter coefficient estimates similar results that apply to three SPF models. Only the standard error value in the alternative SPF model tends to be smaller than the standard SPF model. The Normal-Gamma model performs better in assessing residual with smaller root mean square error (RMSE) than the others, but the results of the estimated technical efficiency still contain outliers. The Student's t-half Normal model estimates technical efficiency no longer has outliers, the range is shorter than the other models, and the results of estimating technical efficiency are not monotonous in the distribution of residual tails. The SPF Student's t-half Normal model is more robust in presence outliers than SPF Normal-half Normal and SPF Normal-Gamma.

How to cite this article:

I. Pranata , A. Djuraidah and M. N. Aidi tc., "ROBUST STOCHASTIC PRODUCTION FRONTIER TO ESTIMATE TECHNICAL EFFICIENCY OF RICE FARMING IN SULAWESI SELATAN," *BAREKENG: J. Math. & App.*, vol. 17, iss. 1, pp. 0391-0400, March 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

Stochastic production frontier (SPF) is a stochastic frontier analysis (SFA) method used to measure the performance of the decision-making unit (DMU) in estimating the production frontier by accounting for the existence of inefficiency. SPF assumes that the output of a DMU is not always efficient or does not always reach the frontier production [1]. The model also assumes that the deviation is due to statistical error factors (noise) [2]. Technical efficiency is a ratio between the production value produced in a DMU and the production frontier estimation with the same number of inputs. The production function in the SPF model has various equations, but the theoretical production function commonly used is the Cobb Douglas (CD) or Transcendental Logarithmic (Translog) production function [3]. The residual in the SPF model consists of two components, inefficiency and noise. The two components are assumed to have certain distribution and are independent. Inefficiency estimation can be measured using an output-oriented and an input-oriented. An output-oriented is a method that maximizes the output by using the same amount of input, while an input-oriented is a method that strives for the amount of production at an input to be achieved by minimizing the number of inputs [4]. The standard SPF model assumes the noise component follows a Normal distribution and the inefficiency component follows a half-Normal distribution. However, the standard SPF model has limitations in estimating technical efficiency when outliers are present in data [5]. The presence of outliers in the data will affect the inaccuracy in estimating the parameters, leading to an exaggerated spread of efficiency predictions [6]. To solve this problem, change the distribution assumption component residual with more flexible parameters, which the data can determine. This study employs two alternative models, the first with Normal-Gamma and the second with Student's t-half Normal, and then the results are compared with standard SPF.

The SPF Normal-Gamma model forms the SPF model by changing the assumption standard of the inefficiency component from half Normal distribution to Gamma distribution [7]. The SPF Normal-Gamma model estimates technical efficiency better than the SPF Normal-half Normal model and has a better distribution pattern of technical efficiency. The model yields a smaller standard deviation and a narrower range of technical efficiency estimates. But, the residual tails in the SPF Normal-Gamma give technical efficiency estimation similar to the SPF Normal-half Normal model. The Student's t-half Normal SPF model forms the SPF model by changing the assumption of the distribution of noise component from the Normal distribution to the Student's t distribution [6]. The SPF Student's t-half Normal model yields technical efficiency estimates better than the SPF Normal-half Normal model, with smaller standard deviation values and narrower ranges of inefficiency estimates. The model also yields a better distribution pattern of technical efficiency. The residual tails of the SPF Student's-half Normal give a different estimate of technical efficiency to the SPF Normal-half Normal Model, where The SPF Student's t-half Normal model can overcome the occurrence of an exaggerated spread of efficiency predictions.

Presence Outliers in production data often occur, one of which is rice farming production. Rice is the primary commodity in the agricultural sector in Indonesia. Based on Statistics Indonesia (2021) in Social Economic National Survey, rice consumption in Indonesia reaches 6.75 kg per capita a month [8]. This amount of consumption exceeds the consumption of other staple foods such as corn, cassava, and sweet potatoes. Based on Statistics Indonesia (2021), Indonesia's most significant rice productions are in the provinces in Jawa, namely Jawa Barat, Jawa Tengah, and Jawa Timur, which have an average rice farming production of around 9 million tons per year [9]. Although rice farming production in Jawa is high, in the long term, rice farming production will decline due to high land use competition [10]. It must be overcome to maintain national food security by developing rice productivity in other provinces. Sulawesi Selatan is a province that has the potential resources to be developed in terms of increasing rice farming productivity, where the province is the fourth largest rice productivity nationally.

A study on the technical efficiency of rice farming in Sulawesi Selatan was conducted by [10] using data from Statistics Indonesia in Cost Structure of Paddy Cultivation Household Survey 2014. The study used data on only about a thousand rice farming households by minimizing outlier observations and building a model using SPF Normal-half Normal. In this study, no depreciation of outlier observations was not carried out. Because eliminating outlier observations would ignore information on the diversity of rice farming household production on the amount of input used, where the amount of data used was around four thousand rice farming households. This study uses an alternative model that can handle datasets that contain outliers. This study examines the effect of distribution assumption changes on the standard SPF model in estimating parameter value and the technical efficiency score in the presence of outlier observations.

2. RESEARCH METHODS

This study compares alternative SPF models' performance to the standard SPF model. Three SPF models regarding noise and inefficiency distributions: Normal-half Normal, Normal-Gamma, and Normal Student t-half Normal. These models estimate parameter and technical efficiency to data on rice farming in Sulawesi Selatan.

2.1 SPF Normal-half Normal

SPF model was proposed by [11] and [12], with the following Equation (1) form:

$$\ln y = \beta_0 + \sum_{j=1}^{J} \beta_j \ln x_j + \varepsilon,\ \ \varepsilon = v - u,\ \ u > 0 \tag{1}$$

where y is the output of producing unit, $x_j$is the input unit of jth, β_j is the parameter coefficient of jth, ε is the residual model, v is the noise term, and u is the inefficiency term.

Normal-half Normal model is a standard SPF model where the noise component follows a Normal distribution $N(0, \sigma_v^2)$ and the inefficiency component follows a half Normal distribution $N^+(0, \sigma_u^2)$. The probability density function for components v and u is as follows:

$$f(v) = \frac{1}{\sigma_v\sqrt{2\pi}} \exp\left(\frac{-v^2}{2\sigma_v^2}\right) \tag{2}$$

$$f(u) = \frac{2}{\sigma_u\sqrt{2\pi}} \exp\left(\frac{-u^2}{2\sigma_u^2}\right) \tag{3}$$

Since the two components are independent, the joint density function is the product of the probability density functions of the two components. The joint density function between v and u is as follows:

$$f(v, u) = \frac{1}{\sigma_v\sigma_u\pi} \exp\left\{-\frac{1}{2}\left(\frac{\sigma_u^2 v^2 + \sigma_v^2 u^2}{\sigma_v^2\sigma_u^2}\right)\right\} \tag{4}$$

The compound error term is $\varepsilon = v - u$, and the joint density function between ε and u is as follows:

$$f(\varepsilon, u) = \frac{1}{\sigma_v\sigma_u} \exp\left\{-\frac{1}{2}\left(\frac{\sigma_u^2\varepsilon^2 + 2\varepsilon u\sigma_u^2 + u^2(\sigma_v^2 + \sigma_u^2)}{\sigma_v^2\sigma_u^2}\right)\right\} \tag{5}$$

then the marginal probability density function of residual obtained by integrating the joint probability distribution, $f(\varepsilon, u)$, over u, as follows [13]:

$$f(\varepsilon) = \int_0^\infty f(\varepsilon, u)\, du = \frac{2}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \tag{6}$$

where $\sigma^2 = \sigma_v^2 + \sigma_u^2$, $\lambda = {}^{\sigma_u}/_{\sigma_v}$, $\phi(.)$ Is the probability density function of standard normal, and $\Phi(.)$ is the cumulative distribution function of standard normal. The parameters of the SPF standard are derived by maximizing the log-likelihood function as follows:

$$\ln L = \ln \prod_{i=1}^{n} f(\varepsilon) = \frac{n}{2}\ln\frac{2}{\pi} - n\ln\sigma - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}\varepsilon_i^2\right) + \sum_{i=1}^{n}\ln\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right) \tag{7}$$

Estimating the value of inefficiency can be calculated through the following equation, and The technical efficiency estimates are as follows [14]:

$$\hat{E}(u_i|\varepsilon_i) = \sigma_*\left(\frac{\phi\left(\frac{\lambda_{*i}}{\sigma_*}\right)}{\Phi\left(-\frac{\lambda_{*i}}{\sigma_*}\right)}\right) - \lambda_{*i},\quad \widehat{TE}_i = \exp\left(-\hat{E}(u_i|\varepsilon_i)\right) \tag{8}$$

where $\lambda_{*i} = \varepsilon_i\sigma_u^2/\sigma^2$, $\sigma_* = \sigma_u\sigma_v/\sigma$, $\phi(.)$ is the probability density function of standard normal, and $\Phi(.)$ is the cumulative distribution function of standard normal.

2.2 SPF Normal Gamma

SPF Normal-Gamma was proposed by [7] where the noise component follows a Normal distribution $N(0,\sigma_v^2)$ and the inefficiency component follows a Gamma distribution $G(\alpha,\theta)$ with $\theta = 1/\sigma_u$). The probability density function for components v follows **Equation (2)**, and u is as follows:

$$f(u) = \frac{\theta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} \exp(-\theta u), \quad \theta = {}^1/_{\sigma_u} \tag{9}$$

The joint density function between v and u is as follows:

$$f(v,u) = \frac{\theta^\alpha u^{\alpha-1}}{\sigma_v\sqrt{2\pi}\Gamma(\alpha)} \exp\left\{-\frac{v^2}{2\sigma_v^2} - \theta u\right\} \tag{10}$$

the joint density function between ε and u is as follows:

$$f(\varepsilon,u) = \frac{\theta^\alpha u^{\alpha-1}}{\sigma_v\sqrt{2\pi}\Gamma(\alpha)} \exp\left(\theta\varepsilon + \frac{\sigma_v^2\theta^2}{2}\right) \exp\left\{-\frac{1}{2\sigma_v^2}[u + (\varepsilon + \theta\sigma_v^2)]^2\right\} \tag{11}$$

the marginal probability density function of ε is as follows:

$$f(\varepsilon) = \frac{\theta^\alpha}{\Gamma(\alpha)} \exp\left\{\theta\varepsilon + \frac{\theta^2\sigma_v^2}{2}\right\} \Phi[-(\varepsilon + \theta\sigma_v^2)] h(\alpha-1,\varepsilon_i),\ h(\alpha-1,\varepsilon_i) = \int_0^\infty z^{\alpha-1} \frac{\frac{1}{\sigma_v}\phi\left(\frac{z-\mu_i}{\sigma_v}\right)}{\Phi\left(\frac{z-\mu_i}{\sigma_v}\right)} du \tag{12}$$

The solution **Equation (12)** is not a closed form where the computation of the $h(\alpha-1,\varepsilon_i)$ is complicated. An alternative approach to calculate $h(\alpha-1,\varepsilon_i)$ is approximated simulation technique. $h(\alpha-1,\varepsilon_i)$ is the expected value of $z^{\alpha-1}$ where z follows truncated Normal distribution.

$$h(\alpha-1,\varepsilon_i) = E(z^{\alpha-1}|z \geq 0), \quad \mu_i = -(\varepsilon_i + \theta\sigma_v^2) \text{ dan } z \sim N(\mu,\sigma_v^2) \tag{13}$$

and $h(\alpha-1,\varepsilon_i)$ would be consistently estimated by

$$\hat{h} = \frac{1}{Q}\sum_{q=1}^{Q} z_q^{\alpha-1} \tag{14}$$

where z_q is drawn from a truncated Normal distribution. This approach produces the simulated probability density function of ε and the function of simulated log-likelihood as follows:

$$\widehat{f(\epsilon)} = \frac{\theta^\alpha}{\Gamma(\alpha)} \exp\left\{\theta\varepsilon + \frac{\theta^2\sigma_v^2}{2}\right\} \Phi[-(\varepsilon + \theta\sigma_v^2)] \left(\frac{1}{Q}\sum_{q=1}^{Q} z_q^{\alpha-1}\right) \tag{15}$$

$$\ln L = n\left(\alpha \ln\beta + \frac{\sigma_v^2\beta^2}{2} - \ln\Gamma(\alpha)\right) + \beta\sum_{i=1}^{n}\varepsilon_i + \sum_{i=1}^{n}\ln\Phi[-(\varepsilon_i + \beta\sigma_v^2)] + \sum_{i=1}^{n}\ln\left\{\frac{1}{Q}\sum_{q=1}^{Q}[z_{iq}]^{\alpha-1}\right\} \tag{16}$$

Inefficiency and technical efficiency can be estimated through the following equation:

$$\hat{E}(u_i|\varepsilon_i) = \frac{h(\alpha,\varepsilon_i)}{h(\alpha-1,\varepsilon_i)} = \frac{\sum_{q=1}^{Q} z_{iq}^{\alpha}}{\sum_{q=1}^{Q} z_{iq}^{\alpha-1}}, \quad \widehat{TE}_i = \exp\left(-\hat{E}(u_i|\varepsilon_i)\right) \tag{17}$$

2.2 SPF Student's t-half Normal

SPF Student's t-half Normal was proposed by [6], the noise component follows a Student's t distribution $T(a)$ where a is a degree of freedom, and the inefficiency component follows a half Normal distribution $N^+(0,\sigma_u^2)$. The probability density function for components u follows **Equations (3)**, and v is as follows:

$$f(v) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\sqrt{\pi a}\sigma_v} \left[1 + \frac{1}{a}\left(\frac{v}{\sigma_v}\right)^2\right]^{\frac{a+1}{2}} \tag{18}$$

The joint density function between v and u is as follows:

$$f(v,u) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\sqrt{\pi a}\sigma_v}\left[1+\frac{1}{a}\left(\frac{v}{\sigma_v}\right)^2\right]^{\frac{a+1}{2}}\frac{2}{\sqrt{2\pi}\sigma_u}\exp\left\{-\frac{u^2}{2\sigma_u^2}\right\} \tag{19}$$

the joint density function between ε and u is as follows:

$$f(\varepsilon,u) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\sqrt{\pi a}\sigma_v}\left[1+\frac{1}{a}\left(\frac{\varepsilon+u}{\sigma_v}\right)^2\right]^{\frac{a+1}{2}}\frac{2}{\sqrt{2\pi}\sigma_u}\exp\left\{-\frac{u^2}{2\sigma_u^2}\right\} \tag{20}$$

the marginal probability density function of ε is as follows:

$$f(\varepsilon) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\sqrt{\pi a}\sigma_v}\int_0^\infty\left[1+\frac{1}{a}\left(\frac{\varepsilon+u}{\sigma_v}\right)^2\right]^{\frac{a+1}{2}}\frac{2}{\sqrt{2\pi}\sigma_u}\exp\left\{-\frac{u^2}{2\sigma_u^2}\right\}du \tag{21}$$

The solution **Equation (21)** is not a closed form in which the computation of the integral is complicated. An alternative approach to calculating is approximated simulation technique. Integration of **Equation (21)** is the expected value of $f(\varepsilon+u)$ where u follows half Normal distribution,

$$h(u) = \int_0^\infty f(\epsilon+u)f(u\geq 0)du = E_u[f(\epsilon+u)|u\geq 0], \quad u\sim N^+(0,\sigma_u^2) \tag{22}$$

and $h(u)$ would be consistently estimated by

$$\hat{h} = \frac{1}{Q}\sum_{q=1}^{Q} f(\varepsilon,u_q) \tag{23}$$

where u_q is drawn from a half Normal distribution. This approach produces the simulated probability density function of ε, and simulated log-likelihood function as follows:

$$\widehat{f(\varepsilon)} = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\sqrt{\pi a}\sigma_v}\frac{1}{Q}\sum_{q=1}^{Q}\left[1+\frac{1}{a}\left(\frac{\varepsilon+u_q}{\sigma_v}\right)^2\right]^{\frac{a+1}{2}} \tag{24}$$

$$\ln L = n\left[\ln\Gamma\left(\frac{a+1}{2}\right)-\ln Q-\ln\sigma_v-\ln\Gamma\left(\frac{a}{2}\right)-\frac{1}{2}\ln a\pi\right]+\sum_{i=1}^{n}\ln\sum_{q=1}^{Q}\left[1+\frac{1}{a}\left(\frac{\varepsilon_i+u_{iq}}{\sigma_v}\right)^2\right]^{-\left(\frac{a+1}{2}\right)} \tag{25}$$

Inefficiency and technical efficiency could be estimated through the following equation:

$$\hat{E}(u_i|\varepsilon_i) = \frac{\sum_{q=1}^{Q}u_{iq}\left[1+\frac{1}{a}\left(\frac{\varepsilon_i+u_{iq}}{\sigma_v}\right)^2\right]^{-\left(\frac{a+1}{2}\right)}}{\sum_{q=1}^{Q}\left[1+\frac{1}{a}\left(\frac{\varepsilon_i+u_{iq}}{\sigma_v}\right)^2\right]^{-\left(\frac{a+1}{2}\right)}}, \qquad \widehat{TE}_i = \exp\left(-\hat{E}(u_i|\varepsilon_i)\right) \tag{26}$$

2.3 Application to Rice Farming in Sulawesi Selatan

This study uses data from Statistics Indonesia in Cost Structure of Paddy Cultivation Household Survey 2014 (ST2013-SPD.S), a series of Agricultural Census 2013. The data covers about four thousand paddy cultivation households in 24 districts of Sulawesi Selatan Province. Table 1 shows the input and output variables used in this study. Independent variables such as production quantity, land, seed, fertilizer, labor, and capital are based on research conducted by **[15]**. The season dummy is based on research conducted by**[16]**. While researchers suspect the other dummy variables affect the amount of rice production.

Table 1. List of input and output variables

Variable	Description	Unit
Y	Production quantity	Kg
X_1	Land	m^2
X_2	Seed	Kg
X_3	Fertilizer	Kg
X_4	Labor	Persons-hour
X_5	Capital	Thousand Rupiah
D_1	Dummy of Hybrid Seed	1: hybrid seeds, 0: not hybrid seeds
D_2	Dummy of paddy field	1: wetland paddy, 0: dryland paddy
D_3	Dummy of Pesticide	1: using pesticide, 0: not using pesticide
D_4	Dummy of season	1: rainy season, 0: dry season

This study uses the transcendental logarithmic (Translog) production function, where the SPF model can be written as follows:

$$\ln Y_i = \beta_0 + \sum_{j=1}^{5} \beta_j \ln X_{ij} + \sum_{j=1}^{5} \delta_j \ln X_{ij}^2 + \sum_{j=1}^{5} \sum_{k>j}^{5} \tau_j (\ln X_{ij})(\ln X_{ik}) + \sum_{l=1}^{4} \omega_l D_{il} + v_i - u_i \quad (27)$$

where Y_i is the output of producing unit for ith household, $X_{ij}$is the jth input used in the ith household, β_j is the linear parameter component, δ_j is the quadratic parameter component, τ_j is the parameter of interaction component, ω_l is the parameter of dummy variable, v_i is noise term for ith household, and u_i is inefficiency term for ith household.

Model SPF Normal-half Normal and Normal-Gamma estimate the parameter using sfaR package [17], while model SPF student's t-half Normal using rfrontier package [6].

3. RESULTS AND DISCUSSION

The estimation results of the SPF Normal-half Normal, Normal-Gamma, and Student's t-half Normal model parameters can be seen in Table 2. The root mean square error (RMSE) value of the Normal-Gamma and Student's t-half Normal model is smaller value than the Normal-half Normal model. It indicates that the alternative model can perform better than the standard SPF model. The SPF Normal-Gamma model is the model that has the best performance because it has the smallest RMSE value.

The independent variables that were significant at the 5% and 10% levels in the three SPF models tended to be the same, except for the independent variable for the interaction between labor and capital. The significant dummy variables in the model are the use of hybrid seeds, paddy fields, and pesticides, while the seasonal dummy variable is not sufficiently evidenced to say that the season's effect can affect rice farming production. Parameter coefficient values generally give almost the same estimated value. However, the standard error values generated by SPF Normal-Gamma and Student's t-half Normal model tend to be smaller than the SPF Normal-half Normal model. The estimated value of the variance inefficiency component and noise component in the SPF Student's t-half Normal model tends to be smaller than the other two SPF models.

In comparing the estimation inefficiency scores, plots between SPF Normal-half Normal and Normal Gamma models shown in Figure 1a are generally similar, and the estimation is monotonous. The two models cannot correct outliers on the estimated inefficiency score. The Normal-Gamma SPF model yields an estimated outlier value of inefficiency with a more extensive range than the Normal-Half Normal model with a smaller estimated inefficiency score. In addition, the distribution of the estimated inefficiency score on residuals is also compared between the Normal-half Normal model and the Student's t-half normal model in Figure 1b. The Normal-half Normal and Student's t-half Normal model plots have different patterns. The plot of the Student's t-half Normal model is non-monotonous, where the estimated inefficiency value at the end of the curve, which is the estimated inefficiency outlier, can be corrected with the Student's t-half Normal model.

Table 2. Estimation Results of SPF Normal-half Normal, Normal-Gamma, Student's t-half Normal model

	Normal-Half Normal			Normal-Gamma			t–Half Normal		
	Coefficient	Standard Error	p-value	Coefficient	Standard Error	p-value	Coefficient	Standard Error	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\beta}_0$	-6.829	1.265	0.00**	-7.412	1.209	0.00**	-6.913	1.1491	0.00**
$\hat{\beta}_1$ ($\ln X_1$)	3.469	0.476	0.00**	3.609	0.453	0.00**	3.496	0.429	0.00**
$\hat{\beta}_2$ ($\ln X_2$)	-2.024	0.352	0.00**	-2.065	0.335	0.00**	-2.126	0.311	0.00**
$\hat{\beta}_3$ ($\ln X_3$)	-0.471	0.224	0.036**	-0.534	0.216	0.01**	-0.544	0.219	0.01**
$\hat{\beta}_4$ ($\ln X_4$)	1.011	0.201	0,00**	1.010	0.198	0.00*	1.089	0.195	0.00**
$\hat{\beta}_5$ ($\ln X_5$)	-0.245	0.151	0.104	-0.230	0.145	0.11	-0.1967	0.147	0.18
$\hat{\delta}_1$ ($\ln X_1^2$)	-0.264	0.049	0,00**	-0.277	0.046	0.00**	-0.287	0.045	0.00**
$\hat{\delta}_2$ ($\ln X_2^2$)	-0.132	0.034	0.00**	-0.133	0.034	0.00**	-0.153	0.032	0.00**
$\hat{\delta}_3$ ($\ln X_3^2$)	-0.0001	0.012	0.992	0.004	0.012	0.71	-0.008	0.013	0.53
$\hat{\delta}_4$ ($\ln X_4^2$)	0.034	0.014	0.01**	0.037	0.013	0.00**	0.036	0.014	0.01**
$\hat{\delta}_5$ ($\ln X_5^2$)	0.055	0.010	0.00**	0.058	0.009	0.00**	0.052	0.009	0.00**
$\hat{\tau}_1$($\ln X_1$ x $\ln X_2$)	0.364	0.071	0.00**	0.376	0.068	0.00**	0.395	0.064	0.00**
$\hat{\tau}_2$($\ln X_1$ x $\ln X_3$)	0.166	0.044	0.00**	0.178	0.044	0.00**	0.214	0.046	0.00**
$\hat{\tau}_3$($\ln X_1$ x $\ln X_4$)	-0.161	0.039	0.00**	-0.166	0.039	0.00**	-0.182	0.038	0.00**
$\hat{\tau}_4$($\ln X_1$ x $\ln X_5$)	0.016	0.032	0.61	0.012	0.032	0.70	0.022	0.032	0.48
$\hat{\tau}_5$($\ln X_2$ x $\ln X_3$)	-0.067	0.034	0.06*	-0.083	0.033	0.01**	-0.084	0.034	0.01**
$\hat{\tau}_6$($\ln X_2$ x $\ln X_4$)	0.057	0.032	0.07*	0.054	0.031	0.08*	0.053	0.031	0.09 *
$\hat{\tau}_7$($\ln X_2$ x $\ln X_5$)	-0.011	0.025	0.66	-0.003	0.024	0.91	0.0008	0.024	0.97
$\hat{\tau}_8$($\ln X_3$ x $\ln X_4$)	0.024	0.022	0.27	0.032	0.021	0.14	0.021	0.022	0.32
$\hat{\tau}_9$($\ln X_3$ x $\ln X_5$)	-0.077	0.017	0.00**	-0.083	0.017	0.00**	-0.101	0.017	0.00**
$\hat{\tau}_{10}$($\ln X_4$ x $\ln X_5$)	-0.029	0.016	0.08*	-0.031	0.016	0.05*	-0.014	0.016	0.38
$\hat{\eta}_1$(D_1 [1])	0.129	0.027	0.00**	0.137	0.027	0.00*	0.138	0.028	0.00**
$\hat{\eta}_2$(D_2 [1])	0.153	0.043	0.00**	0.131	0.044	0.00**	0.145	0.046	0.00**
$\hat{\eta}_3$(D_3 [1])	-0.057	0.027	0.03**	-0.043	0.025	0.09*	-0.065	0.028	0.02**
$\hat{\eta}_4$(D_4 [1])	-0.010	0.014	0.48	-0.008	0.014	0.55	-0.009	0.014	0.79
$\hat{a}$	∞	-	-	-	-	-	2.850	0.263	0.00**
$\hat{P}$	-	-	-	0.765	0.099	0.00**	-	-	-
$\hat{\sigma}_u^2$	0.433	0.014	0.00**	0.432			0.338	0.02	0.00**
$\hat{\sigma}_v^2$	0.057	0.08	0.00**	0.083			0.034	0.011	0.00**
RMSE	0.699	-	-	0.577	-	-	0.665	-	-

Statistical significance at: * 10% ** 5%

The two black dots are the minimum and maximum outliers of the estimated inefficiency scores with the Student's t-half Normal model, and the two black dots are also plotted in the Normal-half Normal model. In the SPF Normal-half Normal plot, the blue dotted line represents the outlier range of the SPF Normal-half Normal model. SPF Normal outlier scores outside the black point range are corrected by the SPF Student's t-half Normal model so that the outlier values are only between the black dots.

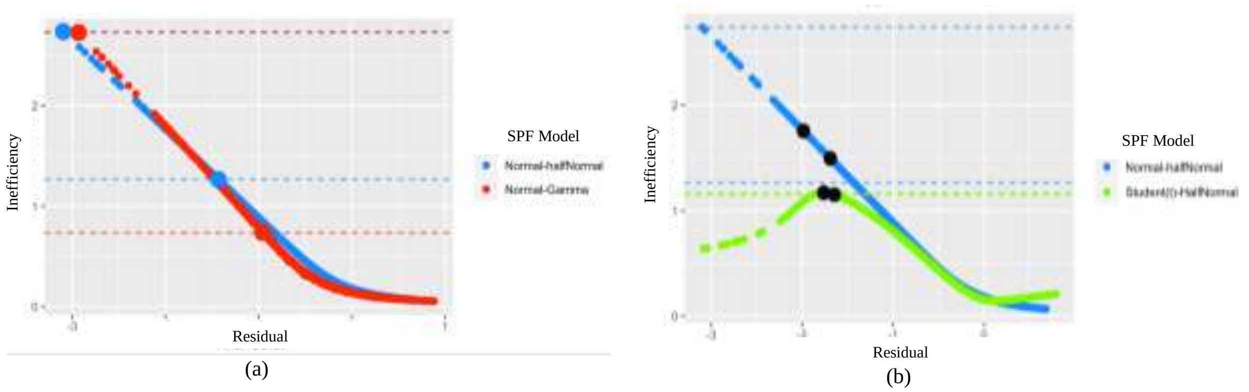
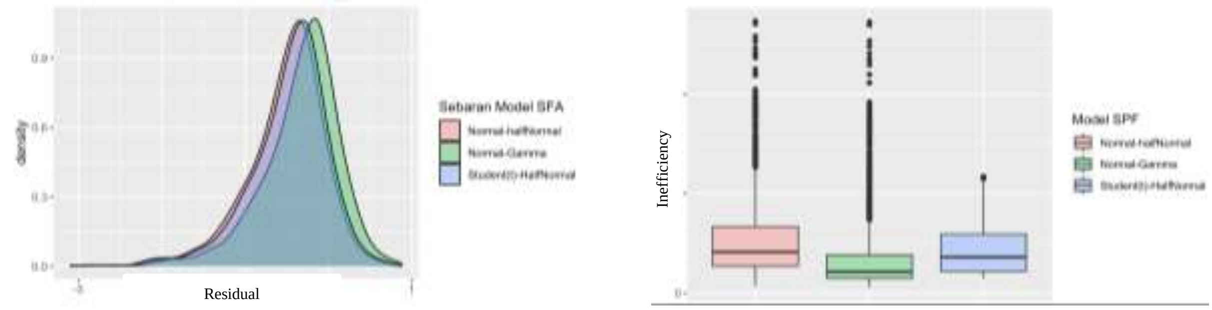


Figure 1. The distribution of inefficiency scores between the standard model and alternative model: (a) Normal-half Normal and Normal-Gamma, (b) Normal-half Normal and Student's t-half Normal

The goodness of the SPF model in estimating the production value can be seen in the residual density distribution graph in Figure 2a. The Normal-Gamma model yields a residual value whose distribution residual value is closer to zero than the Normal-half Normal and Student's t-half Normal models. Hence, the Normal-Gamma SPF model is a model that has a good performance based on a small root mean square error

(RMSE) value than the other two models. However, a smaller RMSE value does not mean that the Normal-Gamma SPF model will produce a better-estimated inefficiency score than the other two models.

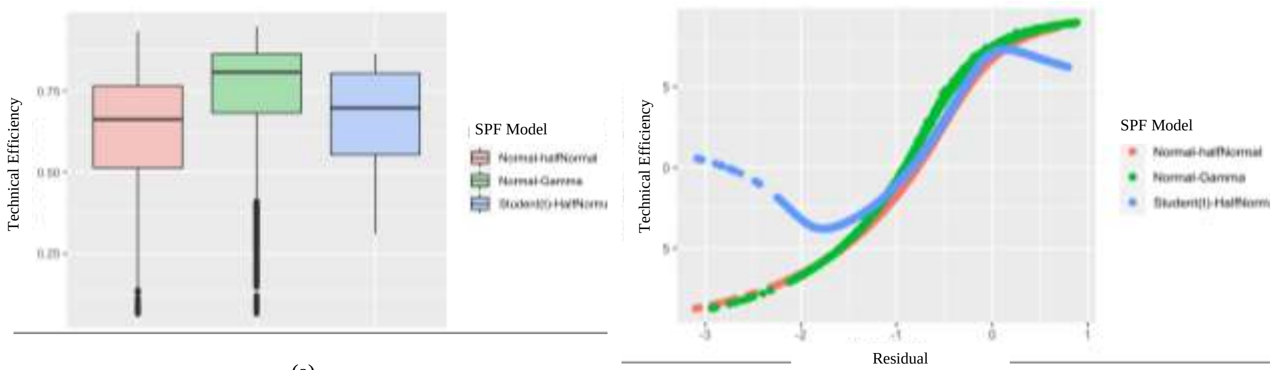
In the boxplot estimation of the inefficiency scores in **Figure 2b**, it can be seen that the estimated inefficiency scores in Normal-Gamma have a larger number of outliers than the other two models. The Student's t-half Normal SPF model is better than the other two models because it can estimate the inefficiency scores with fewer outlier observations. Based on these results, changes in the distribution assumption on the noise component have more effect on the estimation of inefficiency in outlier data than changing the distribution assumption on the inefficiency component.



(a) (b)

Figure 2. The density distribution of the residual using Normal-half Normal, Gamma-normal, and Student's t-half normal models (a), Boxplot estimates of inefficiency scores using Normal-half Normal, Normal-Gamma, and Student's t-half Normal model (b)

The Boxplot estimation of technical efficiency scores ($\widehat{TE}$) in **Figure 3a** shows that the Normal-Half Normal model estimates the technical efficiency score, which has outlier observations. The alternative model with the Normal-Gamma model estimates technical efficiency scores with larger outliers than the Normal-half Normal model. At the same time, the Student's t-half Normal model estimates technical efficiency scores with no outliers. The Student t-half Normal model can correct outliers in estimating technical efficiency scores better than the other two models.



(a) (b)

Figure 3. The distribution of estimates in the three SPF models: (a) Boxplot of estimated technical efficiency scores, (b) plot of the distribution of technical efficiency against residual

A comparison of estimated technical efficiency against residual for three SPF models shows in **Figure 3b**. The figure shows that using the Student's t-half Normal model, the outlier value at the tailed curve can be corrected so that it does not result in outliers on the estimated technical efficiency. It can happen because the noise component that follows Normal distribution estimates a monotonous value. In contrast, the resulting estimate is not monotonous if the distribution uses the Student's t distribution.

Summary statistics of estimated technical efficiency shows in **Table 3**. Estimating technical efficiency values in the Student's t-half Normal model can estimate a narrower range of technical efficiency values than the other two SPF models, even though there are outliers in the data. Based on the results of the study, the SPF Student's t-half Normal model is a model that is more robust to the presence of outliers than the Normal-half Normal model and the Normal-Gamma model.

Table 3. Summary statistics of estimated technical efficiency

	Normal-half Normal	Normal-Gamma	Student's t-half Normal
Mean	0.62937	0.74705	0.6684
Minimum	0.06479	0.06534	0.3107
Maximum	0.93318	0.94841	0.8650

4. CONCLUSIONS

The estimated parameter values of the three SPF models were similar, but the standard error of the alternative models tends to be smaller than the Normal-half Normal model. The Normal-gamma model performs better in estimating residuals with smaller RMSE than the others. However, the results of the estimated technical efficiency using the Normal-Gamma model still contain outliers. Meanwhile, using the Student's t-half Normal model, the estimation of technical efficiency no longer contains an outlier. This study shows that the Student's t-half Normal model is more robust in estimating the technical efficiency score when the presence outlier in the data

REFERENCES

[1] T. J. Coelli, D. S. P. Rao, C. J. O'Donnell, and G. E. Battese, *An Introduction to Efficiency and Productivity Analysis*. Springer US, 2005. [Online]. Available: https://books.google.co.id/books?id=V2Rpu8M6RhwC
[2] C. J. O'Donnell, *Productivity and Efficiency Analysis*. Singapore: Springer, 2018.
[3] R. Zulkarnain, "Inefisiensi Perikanan Budi Daya Indonesia dengan Pendekatan Model Stochastic Frontier Spasial [Tesis]," Institut Pertanian Bogor, Bogor, 2021.
[4] C. S. Kumbhakar, J. H. Wang, and P. A. Horncastle, *A Practitioner's Guide to: Stochastic Frontier Analysis Using Stata*. New York: Cambridge University Press, 2015.
[5] M. S. Campos, M. A. Costa, T. S. Gontijo, and A. L. Lopes-Ahn, "Robust Stochastic Frontier Analysis Applied to The Brazilian Electricity Distribution Benchmarking Method," *Decis. Anal. J.*, vol. 3, no. April, p. 100051, 2022, doi: 10.1016/j.dajour.2022.100051.
[6] P. Wheat, A. D. Stead, and W. H. Greene, "Robust Stochastic Frontier Analysis: a Student's t-half Normal Model with Application to Highway Maintenance Costs in England," *J. Product. Anal.*, vol. 51, no. 1, pp. 21–38, 2019, doi: 10.1007/s11123-018-0541-y.
[7] W. H. Greene, "A Gamma-Distributed Stochastic Frontier Model," *J. Econom.*, vol. 46, pp. 141–163, 1990.
[8] [BPS] Badan Pusat Statistik, *Ringkasan Eksekutif Pengeluaran dan Konsumsi Penduduk Indonesia berdasarkan Survei Sosial Ekonomi Nasiomal (Susenas) September 2021*. Jakarta: BPS, 2021.
[9] Statistics Indonesia, "Luas Panen, Produksi, dan Produktivitas Padi menurut Provinsi 2018-2020," 2021. https://www.bps.go.id/indicator/53/1498/2/luas-panen-produksi-dan-produktivitas-padi-menurut-provinsi.html (accessed Aug. 06, 2021).
[10] A. A. Muhammad, "Efisiensi Teknis Usaha Tani Padi di Provinsi Sulawesi Selatan [tesis]," Bogor: Institut Pertanian Bogor, 2018.
[11] D. Aigner, C. A. Lovell, and P. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," *J. Econom.*, vol. 6, pp. 21–37, 1977.
[12] W. Meeusen and J. Van Den Broeck, "Efficiency Estimation from Cobb-Douglas Production Function with Composed Error," *Int. Econ. Rev. (Philadelphia).*, vol. 18, 1977.
[13] S. Sangeethamani and L. Mary Louis, "Maximization of Technical Efficiency of a Normal- Half Normal Stochastic Production Frontier Model," vol. 13, no. 10, pp. 7353–7364, 2017.
[14] J. Jondrow, C. A. Knox Lovell, I. S. Materov, and P. Schmidt, "On The Estimation of Technical Inefficiency in The Stochastic Frontier Production Function Model," *J. Econom.*, vol. 19, no. 2–3, pp. 233–238, 1982, doi: 10.1016/0304-4076(82)90004-5.
[15] R. Zulkarnain, A. Djuraidah, I. M. Sumertajaya, and I. Indahwati, "Utilization of Student'S T Distribution To Handle Outliers in Technical Efficiency Measurement," *Media Stat.*, vol. 14, no. 1, pp. 56–67, 2021, doi: 10.14710/medstat.14.1.56-67.
[16] A. R. Syahputra, "Analissi Efisiensi Teknis Usahatani Padi di Kalimantan Tengah: Pendekatan Stochastic Frontier [Tesis]," Bogor, Institut Pertanian Bogor, 2022.
[17] K. H. D. and Y. D. and L. Latruffe, "{sfaR}: Stochastic Frontier Analysis using R." 2022. [Online]. Available: https://cran.r-project.org/package=sfaR/