JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 10, No. 2, April 2026, pp. 426-441

Combined Truncated Spline and Fourier series in Nonparametric Biresponse Regression: A Case of Agricultural Productivity

Hartina Husain[1]*, Rizki Aristyarini[2], Andi Oxy Raihan Machikami Rahman[1], Nur Rahmi[3], Muhammad Rifki Nisardi[4]
[1]Data Science Study Program, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia
[2]Food Technology Study Program, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia
[3]Mathematics Study Program, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia
[4]Metallurgical Engineering Study Program, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia
hartinahusain@ith.ac.id

Article History:
Received : 30-09-2025
Revised : 21-12-2025
Accepted : 31-12-2025
Online : 01-04-2026

Keywords:
Agricultural Productivity;
Biresponse;
Fourier Series;
Food Security;
Truncated Spline.

ABSTRACT

Agriculture plays a strategic role in supporting economic development and food security in Indonesia, particularly in South Sulawesi, one of the country's primary rice-producing regions. Existing studies on agricultural productivity commonly rely on parametric or single-response models, which are less effective in capturing the nonlinear, locally varying, and interrelated characteristics of agricultural indicators. Addressing this research gap, the present study applies a biresponse nonparametric regression approach that integrates truncated splines and Fourier series to simultaneously model rice productivity and the food security index. This quantitative observational research uses secondary regional agricultural statistics, and the analytical procedure includes formulating the biresponse model, conducting diagnostic checks of key nonparametric assumptions, and estimating parameters using the Weighted Least Squares (WLS) method. Model selection was conducted using the Generalized Cross Validation (GCV) criterion, which indicated that rainfall was better approximated with truncated splines and extension workers with Fourier series. The optimal knot points were obtained at 1207.096 for rice productivity variable and 1207.556 for food security index variable, with one oscillation applied in the Fourier series and one knot for the truncated spline. The results show that the best model was obtained with the smallest Generalized Cross Validation (GCV) value of 21.38, a coefficient of determination of 94.85%, and a Mean Absolute Percentage Error (MAPE) of 9.68%. These results demonstrate the methodological advantage of the combined biresponse nonparametric model in accommodating complex data structures and provide actionable insights for policymakers in optimizing resource allocation, strengthening extension services, and enhancing food security strategies in South Sulawesi.

https://doi.org/10.31764/jtam.v10i2.35319

A. INTRODUCTION

Agriculture is a strategic sector that plays a vital role in economic development and national food security (Gina et al., 2023). In addition to serving as a provider of food, this sector is also the main source of livelihood for millions of farmers worldwide (Viana et al., 2022). Indonesia, as an agrarian country endowed with abundant natural resources and a tropical climate, should have the capacity to achieve food self-sufficiency. However, the reality indicates otherwise (Susanti et al., 2024). In 2023, Indonesia was recorded as the world's largest rice importer, with an import value reaching USD 1.79 billion, or approximately 5.46% of total global rice imports (Safi'i, 2025). This position surpassed countries such as the Philippines, Saudi Arabia, and China, reflecting a disparity between agrarian potential and national production capacity

The reliance on food imports, particularly rice, demonstrates that domestic agricultural productivity still faces numerous challenges. The complexity of these problems is further exacerbated by climate change, land degradation, limited water resources, and the surge in food demand due to population growth. Addressing these challenges requires sustainable improvements in agricultural productivity. This aligns with the objectives of the Sustainable Development Goals (SDGs), particularly Goal 2 (Zero Hunger) and Goal 12 (Responsible Consumption and Production), which promote efficient and sustainable agricultural systems (Prihady & Djinar, 2022). Various government programs such as People's Business Credit, the Threefold Export Movement, and the Agricultural Development Strategic Command have been introduced (Marefanda et al., 2024). However, the effectiveness of these policies is still constrained by limited infrastructure, low technology adoption, and insufficient dissemination of information to farmers (Budiwiranto et al., 2025).

In the context of such complex and diverse challenges, data-driven approaches are crucial as a foundation for more targeted and contextual decision-making (Elragal & Elgendy, 2024). These approaches enable the identification of region-specific characteristics, provide insights into the key factors influencing productivity, and facilitate the design of more accurate and efficient interventions. Moreover, optimal data utilization enhances transparency, objectivity, and accountability in the planning and evaluation of agricultural programs at various levels (Weraikat et al., 2024). By integrating advanced analytical methods, policymakers can anticipate potential risks and formulate proactive strategies that align with local conditions. Data-driven model also support continuous monitoring, allowing interventions to be refined as new information become available. Ultimately, these approaches strengthen the resilience of agricultural systems and promote sustainable improvements in productivity and food security (Elseknidy et al., 2025).

Agricultural productivity is determined by interacting factors, both natural, such as rainfall, temperature, and soil fertility, and technical, such as fertilization, irrigation, and the use of technology. Previous studies have shown that agricultural productivity is influenced by a combination of harvested area, rainfall, and fertilizer application (Escobar et al., 2024; Wang, 2025 ; Tana & Woldesenbet, 2017). Therefore, analytical methods that are flexible and capable of capturing data variability patterns are required to obtain more accurate insights.

Nonparametric regression offers a relevant analytical solution as it does not require assumptions about the functional form of relationships among variables. The truncated spline estimator is effective in identifying local changes in data, while the Fourier series captures seasonal or periodic patterns commonly observed in agricultural systems, such as planting cycles and rainfall fluctuations (Husain et al., 2024; Ni'matuzzahroh & Dani, 2024). The combination of both estimators allows the development of adaptive and accurate models for estimating agricultural yields based on multiple influencing factors (Sauri et al., 2021).

South Sulawesi Province, as one of Indonesia's national rice granaries, contributes significantly to national rice production (Jamil et al., 2023). The region has relatively stable agro-climatic conditions, supportive rainfall, and extensive fertile agricultural land, making it highly potential in strengthening national food security. However, despite this potential, South Sulawesi also faces challenges such as climate variability, limited technology adoption, and unequal distribution of agricultural resources (Zulkifli, Besse Dahliana, 2025). These complex conditions make South Sulawesi a strategic region for the application of data-driven approaches. Through accurate and contextual modelling, agricultural development policies in the region can be formulated more effectively and have a direct impact on improving productivity and farmers' welfare.

B. THEORITICAL REVIEW

Regression analysis is a fundamental statistical technique employed to examine the relationship between a dependent (response) variable and one or more independent (predictor) variables. Broadly, regression approaches can be classified into parametric, nonparametric, and semiparametric methods. Parametric regression is generally applied when the underlying data pattern is well defined such as linear, quadratic, or cubic forms because it assumes a specific functional relationship between the variables. Nevertheless, in many practical situations, data do not always conform to such rigid functional forms. To address this limitation, nonparametric regression offers a more flexible framework by allowing the relationship between response and predictor variables to be modelled without requiring prior assumptions about its exact shape (Husain et al., 2021). This flexibility makes nonparametric regression particularly useful for capturing complex, nonlinear, or locally varying patterns within data, thereby providing more accurate and adaptive representations of real-world phenomena. The general expression of a nonparametric regression model can be written as follows:

$$y_i = f(x_i) + \varepsilon_i \, , i = 1,2,3 \cdots n \tag{1}$$

where i denotes the $i-$th observation, y_i is the response variable, $f(x_i)$ represents the nonparametric regression function at the $i-$th observation, and ε_i is the error term assumed to be independently and identically distributed (IID) with mean 0 and variance σ^2.

In nonparametric regression, researchers are afforded greater flexibility in selecting analytical methods that align with the observed data patterns. Unlike parametric regression, which is restricted to predetermined functional forms, nonparametric regression does not impose such rigid assumptions. A variety of estimator techniques can be employed in this framework, including Kernel methods, Local Polynomial regression, Splines, Histograms, Fourier Series, and several other advanced approaches.

One of the widely used approaches in nonparametric regression is the Truncated Spline method. Truncated splines are characterized by changes in data behavior across specific sub-intervals. The advantage of splines lies in their ability to incorporate knot points while producing relatively smooth curves, whereas the term truncated refers to a piecewise-defined function. In a nonparametric spline regression model, a spline function of order s with knot points $K_1, K_2, \dots, K_M$ $(K_1 \leq K_2 \leq \cdots \leq K_M)$ can be expressed as follows (Sriliana et al., 2024):

$$f(x_i) = \sum_{r=0}^{s} \beta_r x_i^r + \sum_{k=1}^{M} \Phi_k (x_i - K_k)_+^s \tag{2}$$

For order s=1 the following equation is obtained

$$f(x_i) = \beta_0 + \beta_1 x_i + \sum_{k=1}^{M} \Phi_k (x_i - K_k)_+ \tag{3}$$

Truncated function given by:

$$(x_i - K_k)_+^s = \begin{cases} (x_i - K_k)_+^s & ;\ x_i \geq K_k \\ 0 & ;\ x_i < K_k \end{cases}$$

where β represents the model parameters and K refers to the spline knot points.

The fourier series is a trigonometric polynomial function that has levels of flexibility to deal with data that has repetitive patterns. The Fourier series is used to estimate the curve regression showing sinus and cosinus waveshe Fourier series is a trigonometric polynomial that provides flexibility in modelling data with recurring or periodic patterns. It is commonly employed to approximate regression curves represented through sine and cosine functions (Ramli et al., 2023). The nonparametric approximation function using the fourier series can be written as follows:

$$g(t_i) = bt_i + \frac{\alpha_0}{2} + \sum_{l=1}^{L} \gamma_l \cos lt_i \tag{4}$$

where b, α_0, γ_l is parameter of fourier series model and $l = 1,2,\cdots,L$ is oscillation parameters.

The optimal spline function is determined by the selection of knot points (K), which capture changes in the function across different intervals. In contrast, the Fourier series relies on the oscillation parameter (l), which represents the number of cosine wave oscillations incorporated into the model. If the values of $K$and l are set too low, the resulting curve will be under-smoothed, appearing rough and highly volatile (Sitohang et al., 2024). Conversely, excessively large values of K and l will lead to over-smoothing, producing a curve that is overly smooth and fails to adequately reflect the actual data pattern (Lestari et al., 2023; Sauri et al., 2021). In numerous nonparametric regression studies, the selection of optimal knot points and oscillation parameters is often carried out using the Generalized Cross Validation (GCV) approach. The general expression of GCV in the context of nonparametric regression is presented as follows (Eubank R L, 1999):

$$GCV(K, l) = \frac{MSE(K, l\,)}{(n^{-1} trace[\boldsymbol{I} - \boldsymbol{T}(K, l)])^2} \tag{5}$$

$$MSE\ (k, l) = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n} \tag{6}$$

The coefficient of determination (R^2) quantifies the proportion of variation in the response variable that is accounted for by the predictor variables in a statistical model. In other words, it reflects how well the regression model explains the variability of the response. The value of R^2 ranges between 0 and 1, and when expressed as a percentage, it represents the proportion of variance in the response variable that can be explained by the model (Joy et al., 2025). A higher R^2 value indicates a better fit of the regression model to the observed data (Vinod, 2022). The coefficient of determination can be calculated using the following formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{7}$$

C. METHODS

The data used in this study are secondary data for the year 2024 obtained from the publications of Badan Pusat Statistik (BPS), Global Satellite Mapping of Precipitation (GSMAP), and the National Food Agency. This study is conducted as quantitative research that employs nonparametric modelling techniques to analyze the determinants of agricultural productivity. The data type employed is cross-sectional, consisting of 24 regencies/municipalities in South Sulawesi. The variables used in this research include two response variables (y) and two predictor variables (x) as presented in Table 1 below.

Table 1. Research Variables

Variables	Description
y_1	Rice Productivity (quintals/hectare)
y_2	Food Security Index
x_1	Number of Agricultural Extension Workers (persons)
x_2	Rainfall (mm/year)

To estimate the biresponse nonparametric regression model, this study combines truncated splines and Fourier series because truncated splines provide flexible local fitting for nonlinear relationships, while Fourier series effectively capture periodic or cyclical patterns. The biresponse framework is used because the two response variables are correlated, allowing joint estimation with higher efficiency. Weighted Least Squares (WLS) is applied to address heterogeneity across observations. The analytical procedure in this study was carried out through the following steps:

1. Obtaining the parameter estimates of the biresponse nonparametric regression model combining truncated splines and Fourier series through the following steps:
   a. Given the response variable y_1 and y_2 with truncated spline nonparametric components $x_1, x_2, \cdots, x_P$ and the Fourier series component $t_1, t_2, \cdots, t_Q$.
   b. The function $f_j(x_{pi})$ is approximated using a linear truncated spline with M knots
   c. The function $g_j(t_{qi})$ is approximated using a Fourier series expansion
   d. The approximated curves $\vec{f}(x), and\ \vec{g}(t)$ can be expressed in matrix form

$$\vec{f}(x) = \boldsymbol{R}\vec{\beta}\ dan\ \vec{g}(t) = \boldsymbol{C}\vec{\gamma}$$

where $\vec{\beta}$ represents the truncated spline regression parameters and $\vec{\gamma}$ epresents the Fourier series regression parameters

e. Thus, the biresponse nonparametric regression model combining truncated splines and Fourier series can be written in matrix form as:

$$\vec{y} = \boldsymbol{R}\vec{\beta} + \boldsymbol{C}\vec{\gamma} + \vec{\varepsilon}$$

f. The optimization problem is solved using the Weighted Least Squares (WLS) method:

$$\underset{\vec{\beta},\vec{\gamma}}{\text{Min}} D(\vec{\beta},\vec{\gamma}) = \underset{\vec{\beta},\vec{\gamma}}{\text{Min}} \left\{ \left(\vec{y} - \boldsymbol{R}\vec{\beta} - \boldsymbol{C}\vec{\gamma}\right)^T \boldsymbol{W} \left(\vec{y} - \boldsymbol{R}\vec{\beta} - \boldsymbol{C}\vec{\gamma}\right) \right\}$$

g. Setting the partial derivatives equal to zero yields:

$$\frac{\partial D(\vec{\beta},\vec{\gamma})}{\partial \vec{\beta}} = 0; \frac{\partial D(\vec{\beta},\vec{\gamma})}{\partial \vec{\gamma}} = 0$$

h. The resulting combined estimator is then obtained as:

$$\hat{\vec{y}} = \boldsymbol{R}\hat{\vec{\beta}} + \boldsymbol{C}\hat{\vec{\gamma}}$$

2. Application of the combined estimator to agricultural productivity data in South Sulawesi, implemented in R Studio 2025 with R version 4.3.1, which involves:
   a. Conducting a correlation test between the two response variables
   b. Conducting data exploration and descriptive analysis of each research variable.
   c. Constructing scatter plots to observe the relationship patterns between the response variables and predictors.
   d. Determining the predictor variables through truncated spline regression curves and Fourier series regression curves using the minimum GCV criterion.
   e. Modelling agricultural productivity data with the combined biresponse nonparametric regression of truncated splines and Fourier series.
   f. Selecting optimal knot points, oscillation parameters using the minimum GCV criterion.
   g. Calculating the coefficient of determination (R^2) for the fitted model.
   h. Drawing conclusions based on the findings.

D. RESULT AND DISCUSSION

1. Parameter Estimates of the Biresponse Nonparametric Regression Model with Truncated Splines and Fourier Series

Given data with response variable Y_1 and Y_2 and predictor variable consisting of nonparametric components of a truncated spline x_1 , $x_2, \cdots, x_P$ and fourier series $t_1, t_2, \cdots, t_Q$. The form of a nonparametric regression model that contains these variables can be expressed as follows:

$$y_{ji} = \sum_{p=1}^{P} f_j(x_{pi}) + \sum_{q=1}^{Q} g_j(t_{qi}) + \varepsilon_{ji} \tag{8}$$

where $i = 1, 2, \cdots, n$ and ε_i assumed to be independent, identical, and normally distributed with zero mean and variance σ_i^2. The function $f_j(x_{pi})$ with p variables is approximated by a linear truncated spline function, namely:

$$\sum_{p=1}^{P} f_j(x_{pi}) = \beta_{0j} + \sum_{p=1}^{P} \left(\beta_{1jp} x_{pi} + \sum_{k=1}^{M} \Phi_{jpk} (x_{pi} - K_{jpk})_+ \right) \tag{9}$$

where M represents the number of knots with a truncated spline function given by:

$$(x_{pi} - K_{jpk})_+ = \begin{cases} (x_{pi} - K_{jpk})_+ & ; x_{pi} \geq K_{jpk} \\ 0 & ; x_{pi} < K_{jpk} \end{cases}$$

It can be expressed in matrix form as:

$$\vec{f}(x) = \boldsymbol{R}\vec{\beta}$$

where

$$\vec{f}(x) = \begin{pmatrix} \vec{f}_1(x) \\ --- \\ \vec{f}_2(x) \end{pmatrix}, \boldsymbol{R} = \begin{pmatrix} \boldsymbol{R_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R_2} \end{pmatrix}, \text{ and } \vec{\beta} = \begin{pmatrix} \vec{\beta}_1 \\ --- \\ \vec{\beta}_2 \end{pmatrix}$$

The **R** matrix contains matrix $\boldsymbol{R_1}, \boldsymbol{R_2}$**,** and **0** with the following description

$$\boldsymbol{R_1} = \begin{pmatrix} 1 & x_{11} & (x_{11} - K_{111})_+ & \cdots & (x_{P1} - K_{1PM})_+ \\ 1 & x_{12} & (x_{12} - K_{111})_+ & \cdots & (x_{P2} - K_{1PM})_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & (x_{1n} - K_{111})_+ & \cdots & (x_{Pn} - K_{1PM})_+ \end{pmatrix}$$

$$\boldsymbol{R_2} = \begin{pmatrix} 1 & x_{11} & (x_{11} - K_{211})_+ & \cdots & (x_{P1} - K_{2PM})_+ \\ 1 & x_{12} & (x_{12} - K_{211})_+ & \cdots & (x_{P2} - K_{2PM})_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & (x_{1n} - K_{211})_+ & \cdots & (x_{Pn} - K_{2PM})_+ \end{pmatrix}$$

Then vector $\vec{\beta}$ contains $\vec{\beta}_1$ and $\vec{\beta}_2$ with the following description

$$\vec{\beta}_1 = (\beta_{01} \quad \beta_{111} \quad \cdots \quad \beta_{11P} \quad \Phi_{111} \quad \cdots \quad \Phi_{1PM})^T$$
$$\vec{\beta}_2 = (\beta_{02} \quad \beta_{121} \quad \cdots \quad \beta_{12P} \quad \Phi_{211} \quad \cdots \quad \Phi_{2PM})^T$$

with $\vec{f}(x)$ of size $2n \times 1$, matrix $\boldsymbol{R}$ of size $2n \times \big(2 + 2P(1 + M)\big)$, and vector $\vec{\beta}$ of size $\big(2 + 2P(1 + M)\big) \times 1$. Next, $g\big(t_{qi}\big)$ with variable q is approximated by a Fourier series, namely:

$$\sum_{q=1}^{Q} g_j\big(t_{qi}\big) = \sum_{q=1}^{Q}\left(b_{jq}t_{qi} + \frac{\gamma_{0jq}}{2} + \sum_{l=1}^{L}\gamma_{jlq}\cos lt_{qi}\right) \tag{10}$$

or in matrix notation it can be written as:

$$\vec{g}(t) = \boldsymbol{C}\vec{\gamma}$$

where

$$\vec{g}(t) = \begin{pmatrix} \vec{g}_1(t) \\ --- \\ \vec{g}_2(t) \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} \boldsymbol{C_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C_2} \end{pmatrix}, \text{and } \vec{\gamma} = \begin{pmatrix} \vec{\gamma}_1 \\ --- \\ \vec{\gamma}_2 \end{pmatrix}$$

The **C** matrix contains matrix $\boldsymbol{C_1}, \boldsymbol{C_2}$, and **0** with the following description

$$\boldsymbol{C_1} = \boldsymbol{C_2} = \begin{pmatrix} t_{11} & \frac{1}{2} & \cos t_{11} & \cdots & \cos Lt_{Q1} \\ t_{12} & \frac{1}{2} & \cos t_{12} & \cdots & \cos Lt_{Q2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{1n} & \frac{1}{2} & \cos t_{1n} & \cdots & \cos Lt_{Qn} \end{pmatrix}$$

Then vector $\vec{\gamma}$ contains $\vec{\gamma}_1$ and $\vec{\gamma}_2$ with the following description

$$\vec{\gamma}_1 = (b_{11} \quad \gamma_{01} \quad \gamma_{111} \quad \cdots \quad \gamma_{1LQ})^T \;;\; \vec{\gamma}_2 = (b_{21} \quad \gamma_{02} \quad \gamma_{211} \quad \cdots \quad \gamma_{2LQ})^T$$

with $\vec{g}(t)$ of size $2n \times 1$, matrix $\boldsymbol{C}$ of size $2n \times \big(2 + 2Q(1 + L)\big)$, and vector $\vec{\gamma}$ of size $\big(2 + 2Q(1 + L)\big) \times 1$. Furthermore, Combined of truncated spline and fourier series biresponse nonparametric regression model can be written as follows:

$$\vec{y} = \boldsymbol{R}\vec{\beta} + \boldsymbol{C}\vec{\gamma} + \tilde{\varepsilon} \tag{11}$$

to obtain parameter estimation results in model (10), it can be done using Weighted Least Square (WLS) optimization method:

$$\underset{\vec{\beta},\vec{\gamma}}{\text{Min}} D(\vec{\beta},\vec{\gamma}) = \underset{\vec{\beta},\vec{\gamma}}{\text{Min}}\, \vec{\varepsilon}^T \vec{\varepsilon}$$
$$= \underset{\vec{\beta},\vec{\gamma}}{\text{Min}} \left(\vec{y} - \boldsymbol{R}\vec{\beta} - \boldsymbol{C}\vec{\gamma}\right)^T \boldsymbol{W} \left(\vec{y} - \boldsymbol{R}\vec{\beta} - \boldsymbol{C}\vec{\gamma}\right)$$

where **W** is a weighting matrix, with the following equation:

$$\mathbf{W} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 & \vdots & \sigma_{12} & \sigma_{12} & \cdots & \sigma_{12} \\ 0 & \sigma_1^2 & \cdots & \vdots & \vdots & \sigma_{12} & \sigma_{12} & \cdots & \sigma_{12} \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_1^2 & \vdots & \sigma_{12} & \sigma_{12} & \cdots & \sigma_{12} \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{21} & \sigma_{21} & \cdots & \sigma_{21} & \vdots & \sigma_2^2 & 0 & \cdots & 0 \\ \sigma_{21} & \sigma_{21} & \cdots & \sigma_{21} & \vdots & 0 & \sigma_2^2 & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \sigma_{21} & \sigma_{21} & \cdots & \sigma_{21} & \vdots & 0 & 0 & \cdots & \sigma_2^2 \end{bmatrix}$$

with σ_1^2 is the variance of response y_1 , σ_1^2 is the variance of response y_2 , and $\sigma_{12} = \sigma_{21}$ is covariance of response y_1 and y_2. Estimated results are obtained

$$\hat{\vec{\beta}} = (\boldsymbol{R}^T\boldsymbol{W}\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{W}\vec{y} - (\boldsymbol{R}^T\boldsymbol{W}\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{W}\boldsymbol{C}\hat{\vec{\gamma}} \quad (12)$$
$$\hat{\vec{\gamma}} = (\boldsymbol{C}^T\boldsymbol{W}\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{W}\vec{y} - (\boldsymbol{C}^T\boldsymbol{W}\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{W}\boldsymbol{R}\hat{\vec{\beta}} \quad (13)$$

By example $\boldsymbol{P} = (\boldsymbol{R}^T\boldsymbol{W}\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{W}$ and $\boldsymbol{A} = (\boldsymbol{C}^T\boldsymbol{W}\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{W}$ then the form $\hat{\vec{\beta}}$ and $\hat{\vec{\gamma}}$ can be simplified to:

$$\hat{\vec{\beta}} = \boldsymbol{P}\tilde{y} - \boldsymbol{P}\boldsymbol{C}\hat{\vec{\gamma}} \quad (14)$$
$$\hat{\vec{\gamma}} = \boldsymbol{A}\tilde{y} - \boldsymbol{A}\boldsymbol{R}\hat{\vec{\beta}} \quad (15)$$

Then subtitute the equation (14) to (15) to get parameter values $\hat{\vec{\gamma}}$ that do not contain other parameters

$$\hat{\vec{\gamma}} = (\boldsymbol{I} - \boldsymbol{ARPC})^{-1}(\boldsymbol{A} - \boldsymbol{ARP})\vec{y} \quad (16)$$
$$\hat{\vec{\gamma}} = \boldsymbol{U}\vec{y} \quad (17)$$

where $\boldsymbol{U} = (\boldsymbol{I} - \boldsymbol{ARPC})^{-1}(\boldsymbol{A} - \boldsymbol{ARP})$

The results of equation (16) are then subtituted into equation (14) to obtain the value of parameter $\hat{\vec{\beta}}$

$$\hat{\vec{\beta}} = (\boldsymbol{P} - \boldsymbol{PC}((\boldsymbol{I} - \boldsymbol{ARPC})^{-1}(\boldsymbol{A} - \boldsymbol{ARP})))\vec{y} \quad (18)$$

$$\hat{\vec{\beta}} = \boldsymbol{T}\vec{y} \tag{19}$$

where $\boldsymbol{T} = \left(\boldsymbol{P} - \boldsymbol{PC}\big((\boldsymbol{I} - \boldsymbol{ARPC})^{-1}(\boldsymbol{A} - \boldsymbol{ARP})\big)\right)$

Based on equations (17) and (19), the combined estimator of truncated spline and fourier series in biresponse nonparametric regression model can be written as follows:

$$\begin{aligned}\hat{\vec{y}} &= \boldsymbol{R}\hat{\vec{\beta}} + \boldsymbol{C}\hat{\vec{\gamma}} \\ &= (\boldsymbol{RT} + \boldsymbol{CU})\vec{y} \\ \hat{\vec{y}} &= \boldsymbol{B}\vec{y}\end{aligned} \tag{18}$$

where $\boldsymbol{B} = \boldsymbol{RT} + \boldsymbol{CU}$

To determine the optimal knot point K and oscillation parameters l with the smallest Generalized Cross Validation (GCV) value criteria, namely:

$$GCV\ (k,l) = \frac{MSE\ (K,l)}{\left(n^{-1}tr\big(\boldsymbol{I} - \boldsymbol{B}(K,l)\big)\right)^2} \tag{19}$$

2. Agricultural Productivity Modelling

The initial step prior to modelling with the Biresponse Nonparametric Regression method is to conduct a correlation test between the response variables. The correlation test applied in this study is the Pearson Correlation. The hypotheses are as follows:

H_0 : There is no correlation between the response variables.

H_1 : There is a correlation between the response variables.

The test results indicated a Pearson correlation coefficient of 0.603 with a p-value of 0.02. At the 5% significance level ($\alpha = 0.05$), this leads to the rejection of H_0 , implying a significant correlation between the Rice Productivity (Y_1) and the Food Security Index (Y_2). This correlation coefficient of 0.603 indicates a moderately strong positive relationship, meaning that districts with higher rice productivity also tend to have better food security conditions. Substantively, this suggests that improvements in rice production contribute to greater availability and stability of food resources at the regional level. Since the correlation assumption between response variables is satisfied, the modelling process using the Biresponse Nonparametric Regression method can be appropriately continued, as shown in Table 2.

Table 2. Descriptive Statistics

Variables	Minimum	Maximum	Mean	Std. Deviation
Rice Productivity	40.05	61.07	49.69	5.229
Food Security Index	71.87	89.28	83.36	3.794
Number of Agricultural Extension Workers	22.00	160.00	95.79	36.60
Rainfall	854.00	5253.40	2304.64	1144.03

The descriptive statistics in Table 2 highlight substantial disparities across regencies in South Sulawesi. Pinrang recorded the highest rice productivity at 61.07 quintals/hectare, while Kepulauan Selayar showed the lowest at only 40.05, reflecting contrasting levels of agricultural performance. Food security also varied, with Pare-Pare achieving the highest index score (89.28), whereas Kepulauan Selayar again stood at the lowest (71.87), suggesting regional inequalities in both production and welfare outcomes. In terms of supporting factors, Bone had the largest number of agricultural extension workers (160), contrasting sharply with Pare-Pare which had only 22, and rainfall conditions ranged from extreme highs in Luwu Timur (5253.4 mm/year) to much lower levels in Pare-Pare (854 mm/year). These findings imply that agricultural capacity and resilience in South Sulawesi are strongly shaped by local resource endowment and institutional support, underscoring the importance of region-specific policies rather than a uniform approach.

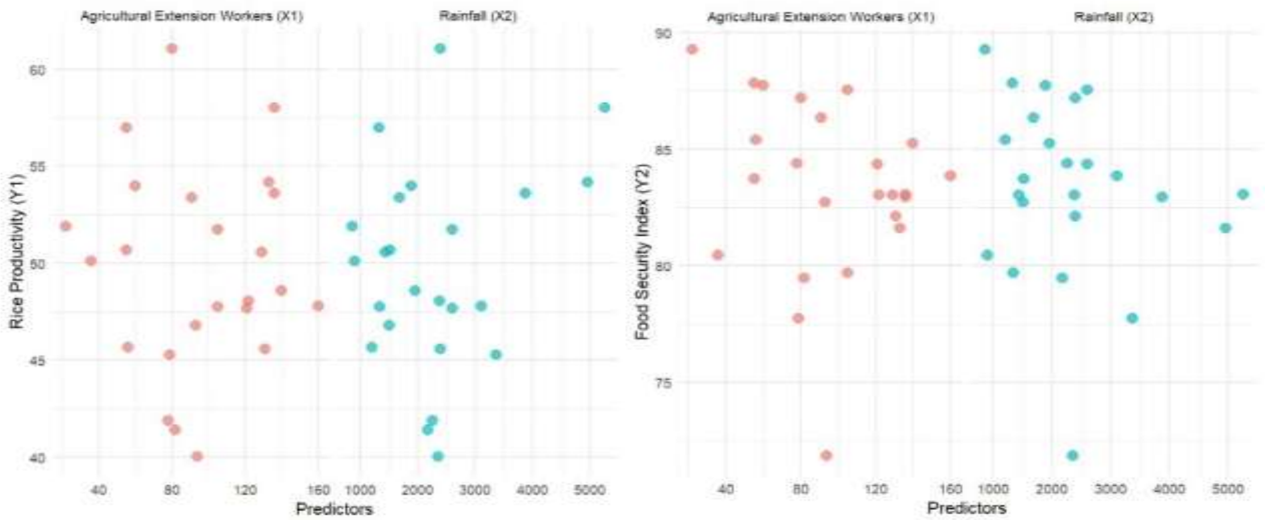

Figure 1. Scatterplot Data

Figure 1 illustrates the scatterplots between the response variables and their predictors. It can be observed that the data patterns do not follow a clear or specific functional form, as the distribution of points appears irregular and does not align with simple linear or polynomial trends. This irregularity suggests that imposing a strict parametric assumption may not capture the true relationship among the variables. Therefore, a more flexible modelling framework, such as nonparametric regression, is more appropriate to accommodate local variations and potentially nonlinear relationships between rice productivity, food security, and their predictors. In this context, the truncated spline estimator is suitable for capturing localized changes in the data pattern, while the Fourier series estimator is effective in representing cyclical or oscillatory tendencies that may arise from seasonal or periodic agricultural phenomena. Hence, combining truncated splines and Fourier series provides an adaptive modelling strategy that aligns well with the characteristics observed in the data.

In this context, the truncated spline estimator is suitable for capturing localized threshold effects, such as sudden increases or decreases in productivity when rainfall surpasses certain critical levels. Meanwhile, the Fourier series estimator is effective in representing cyclical or oscillatory tendencies, which may arise from seasonal agricultural cycles, variability in extension worker performance across periods, or recurring climatic patterns. Hence, combining truncated splines and Fourier series provides an adaptive modelling strategy that aligns well with the characteristics observed in the data and allows the model to reflect both abrupt and periodic variations in agricultural conditions. Subsequently, the selection was carried out using the Generalized Cross Validation (GCV) criterion to determine which predictors are better approximated by truncated splines and which are more suitably represented by the Fourier series. This step ensures that each predictor is modelled with the most appropriate estimator, thereby improving the overall flexibility and accuracy of the biresponse nonparametric regression model, as shown in Table 3.

Table 3. Comparison of GCV Value

No	Variable Truncated spline	Variable Fourier Series	GCV
1	x_1	x_2	25.70
2	x_2	x_1	**21.38**

Table 3 shows that the model with X_2 approximated by a truncated spline and X_1 by a Fourier series yields a lower GCV value is 21.38 compared to the alternative 25.70. This indicates that such a specification provides a better fit for the data. Substantively, this finding suggests that rainfall exhibits localized nonlinear effects best captured by spline knots, while the influence of extension workers follows a smoother, possibly cyclical pattern that is more effectively represented by the Fourier series. Based on this specification, the next step is to estimate the model parameters for both the truncated spline and Fourier series components. The results of the parameter estimation are presented in Table 4.

Table 4. Parameter Estimation

Parameter $\hat{\hat{\beta}}$	Estimation	Parameter $\hat{\hat{\gamma}}$	Estimation
β_{01}	49.7582	b_{11}	-0.0400
β_{111}	0.0009	γ_{01}	5.12×10^{-11}
Φ_{111}	0.0008	γ_{111}	-2.6972
β_{02}	81.2248	b_{11}	-0.0197
β_{121}	0.0034	γ_{02}	6.50×10^{-11}
Φ_{211}	-0.0039	γ_{211}	-2.5764

From Table 4, the parameter estimates can be used to construct the regression models for rice productivity and the food security index. In this specification, one oscillation is applied to approximate the Fourier series estimator, while a single knot is used for the truncated spline estimator. The knot represents a threshold value of rainfall at which the relationship with the response variable changes, allowing the model to capture localized nonlinear effects that would not be detected under a global functional form. Meanwhile, the single oscillation in the Fourier series reflects a smooth cyclical variation in the influence of extension workers, indicating that their effect follows a gradual and possibly repetitive pattern rather than abrupt changes. The optimal knot points are obtained at 1207.096 for Rice Productivity (y_1) and 1207.556 for Food Security Index (y_2), which serve as the basis for the nonparametric modelling of both response variables.

$$\hat{y}_{1i} = 49.7582 + 0.0009x_{2i} + 0.0008(x_{2i} - 1207.096)_+ - 0.0400x_{1i} + \frac{5.12 \times 10^{-11}}{2} - 2.6972 \cos x_{1i}$$

$$\hat{y}_{2i} = 81.2248 + 0.0034x_{2i} - 0.0039(x_{2i} - 1207.556)_+ - 0.0197x_{1i} + \frac{6.50 \times 10^{-11}}{2} - 2.5764 \cos x_{1i}$$

The obtained smallest Generalized Cross Validation (GCV) value is 21.38 , coefficient of determination (R^2) is 94.85%, and the Mean Absolute Percentage Error (MAPE) is 9.68%. These indicators collectively confirm that the combined truncated spline and Fourier series specification provides a highly accurate and adaptive representation of agricultural productivity dynamics in South Sulawesi. Figure 2 presents the comparison between the actual and predicted values for both response variables. For rice productivity (y_1), the predicted values generally follow the overall pattern of the actual observations, although some fluctuations are not fully captured, particularly at extreme points. Meanwhile, for the food security index (y_2), the predicted values align more closely with the actual data, indicating a better fit. Overall, the model demonstrates reasonable accuracy in representing both response variables.

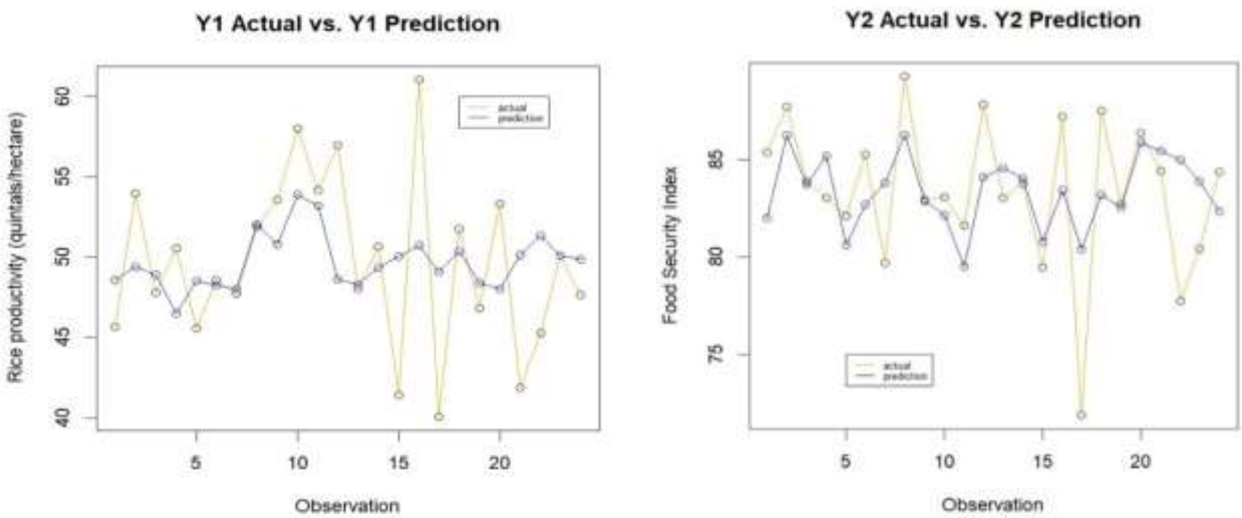

Figure 2. The comparison between actual value and predicted value

The results of the modelling provide important insights for agricultural productivity in South Sulawesi. By capturing both local variations through the truncated spline and seasonal or cyclical patterns through the Fourier series, the model helps identify critical thresholds of rainfall and the role of extension workers in influencing rice productivity and food security.

This information is valuable for policymakers and local governments to design more targeted interventions, such as optimizing irrigation systems, allocating agricultural extension services more effectively, and anticipating the impact of extreme rainfall on crop yields. Ultimately, the model not only improves statistical accuracy but also serves as a decision-support tool to strengthen agricultural planning and food security strategies in South Sulawesi.

The findings of this study are consistent with previous research (Escobar et al., 2024; Wang, 2025; Tana & Woldesenbet, 2017), which similarly identified rainfall and other agro-technical factors as key determinants of agricultural productivity. However, by employing a biresponse nonparametric regression model that integrates truncated spline and Fourier series, this study offers a new contribution by capturing local variations and cyclical patterns that earlier parametric approaches were unable to detect. These results also reinforce the perspectives of Weraikat et al., (2024) and Elragal & Elgendy, (2024) regarding the importance of data-driven analytical methods, demonstrating that adaptive modelling techniques are essential for understanding the complex agricultural dynamics in regions such as South Sulawesi.

E. CONCLUSION AND SUGGESTIONS

This study demonstrates that the biresponse nonparametric regression model combining truncated splines and Fourier series is an effective analytical approach for capturing the complex, nonlinear relationships influencing agricultural productivity and food security in South Sulawesi. By modelling rainfall and the number of agricultural extension workers simultaneously, the results reveal that both factors play significant roles: rainfall exhibits nonlinear threshold effects on productivity, while extension workers contribute consistently to improving regional food security. These findings underscore the importance of flexible modelling approaches for understanding agricultural systems characterized by local variation and cyclical environmental patterns.

The research contributes to the literature by offering an adaptive modelling framework capable of uncovering patterns that traditional parametric methods may overlook. Practically, the results provide an evidence-based foundation for policymakers to optimize irrigation planning, enhance the deployment of extension workers, and anticipate climate-related risks to crop production. Future research is recommended to incorporate additional environmental and socio-economic variables, explore spatial or spatiotemporal extensions of the model, and apply the approach to other key agricultural commodities to further support data-driven agricultural planning in Indonesia.

ACKNOWLEDGEMENT

This research was funded by the Government of Indonesia, with primary support from the Directorate of Research and Community Service under the Ministry of Higher Education, Science, and Technology. We would also like to express our sincere gratitude to all parties who provided support and constructive feedback during the preparation of this study. Their contributions have been invaluable in improving the overall quality of this research.

REFERENCES

Budiwiranto, B., Maryam, D., & Zaimuddin, L. (2025). Building Inclusive Communication in Empowering Farmers : Opportunities and Challenges for Sustainability in the Digital Era. *Jurnal Ilmu Sosial Dan Pendidikan (JISIP)*, *14*(1), 68–79. https://doi.org/10.23887/jish.v14i1.86105

Cláudia M. Viana, Freire, D., Abrantes, P., Rocha, J., & Pereira, P. (2022). Agricultural land systems importance for supporting food security and sustainable development goals: A systematic review. *Science of The Total Environment*, *806*(3), 150718. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.150718

Elragal, A., & Elgendy, N. (2024). A data-driven decision-making readiness assessment model : The case of a Swedish food manufacturer. *Decision Analytics Journal*, *10*(January), 100405. https://doi.org/10.1016/j.dajour.2024.100405

Elseknidy, M., Al-mhdawi, M. K. S., Qazi, A., & Ojiako, U. (2025). Developing a sustainability-driven risk management framework for green building projects : A literature review. *Journal of Cleaner Production*, *519*(May), 145891. https://doi.org/10.1016/j.jclepro.2025.145891

Escobar, D., Arata, L., Mausch, K., Sckokai, P., Fasse, A., Rommel, J., & Chopin, P. (2024). Linking innovations adoption with farm sustainability : Empirical evidence from rainwater harvesting and fertilizer micro-dosing in Tanzania. *World Development*, *183*(July), 106732. https://doi.org/10.1016/j.worlddev.2024.106732

Eubank R L. (1999). *Spline Smoothing and Nonparametric Regression Second Edition*. Marcel Dekker. https://doi.org/10.1201/9781482273144

Gina, G. A., Ana Mariya, Charita Natalia, Sirat Nispuana, M. Farhan Wijaya, & M. Yoga Phalepi. (2023). The Role of the Agricultural Sector on Economic Growth in Indonesia. *Indonesian Journal of Multidisciplinary Sciences (IJoMS)*, *2*(1), 167–179. https://doi.org/10.59066/ijoms.v2i1.325

Husain, H., Budiantara, I. N., & Zain, I. (2021). Mixed estimator of spline truncated, Fourier series, and kernel in biresponse semiparametric regression model. *IOP Conference Series: Earth and Environmental Science*, *880*(1), 1–11. https://doi.org/10.1088/1755-1315/880/1/012046

Husain, H., Irmayani, & Andi Oxy MR. (2024). Modelling Stunting Prevalence in Indonesia : Mixed Spline Truncated and Fourier Series Nonparametric Regression. *Inferensi*, *7*(3), 215–222. https://doi.org/10.12962%2Fj27213862.v7i3.20518

Jamil, A., Ali, M. S. S., Fahmid, I. M., Salman, D., & Rahmadanih, R. (2023). Subsidized fertilizer management in the rice production centers of South Sulawesi, Indonesia: Bridging the gap between policy and practice. *Open Agriculture*, *8*(1), 20220233. https://doi.org/10.1515/opag-2022-0233

Joy, F., Tamba, R., Ranly, L. O., Tri, A., Dani, R., Fauziyah, M., & Adrianingsih, N. Y. (2025). Comparison of Mean Centering Regression and Spline Truncated Nonparametric Regression on Factors Affecting the Number of Crimes in Indonesia. *Multica Science and Technology*, *5*(1), 36–45. https://doi.org/10.47002/fpp74f96

Lestari, B., Chamidah, N., Nyoman Budiantara, I., & Aydin, D. (2023). Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator. *Journal of King Saud University - Science*, *35*(5), 102664. https://doi.org/10.1016/j.jksus.2023.102664

Marefanda, N., Husna, C. A., & Hajad, V. (2024). The Effectiveness of Local Government Policies in Gayo Coffee Farmers ' Welfare during the COVID-19 Pandemic. *Journal of Governance and Public Policy*, *11*(1), 71–85. https://doi.org/10.18196/jgpp.v11i1.18619

Mohamad Rapi Safi'i. (2025). The Influence of Consumption, Production, Price, and Exchange Rate on Indonesia's Rice Imports. *Indonesian Interdisciplinary Journal of Sharia Economics*, *8*(3), 9507–9515. https://doi.org/10.31538/iijse.v8i3.7396

Ni'matuzzahroh, L., & Dani, A. T. R. (2024). Nonparametric Regression Modelling with Multivariable Fourier Series Estimator on Average Length of Schooling in Central Java in 2023. *Inferensi*, *7*(2), 73. https://doi.org/10.12962/j27213862.v7i2.20219

Prihady, I. T., & Setiawina Nyoman Djinar. (2022). The Effect Of Production, International Rice Prices, And Foreign Exchange Reserves On The Volume Of Indonesian Rice Imports In 2008-2020. *Eurasia: Economics & Business*, *3*(57). https://doi.org/https://doi.org/10.18551/econeurasia.2022-03

Ramli, M., Budiantara, I. N., & Ratnasari, V. (2023). A method for parameter hypothesis testing in nonparametric regression with Fourier series approach. *MethodsX*, *11*(October), 1–13. https://doi.org/10.1016/j.mex.2023.102468

Sauri, M. S., Hadijati, M., & Fitriyani, N. (2021). Spline and Kernel Mixed Nonparametric Regression for Malnourished Children Model in West Nusa Tenggara. *Jurnal Varian*, *4*(2), 99–108. https://doi.org/10.30812/varian.v4i2.1003

Sitohang, F. K. S., Sifriyani, S., & Mahmuda, S. (2024). Application of Nonparametric Regression Spline Truncated for Modelling the Height of Yeop Chagi Kicks of Taekwondo Athletes in Samarinda City. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *18*(2), 0657–0666. https://doi.org/10.30598/barekengvol18iss2pp0657-0666

Sriliana, I., Budiantara, I. N., & Ratnasari, V. (2024). The performance of mixed truncated spline-local linear nonparametric regression model for longitudinal data. *MethodsX*, *12*(March), 102652. https://doi.org/10.1016/j.mex.2024.102652

Susanti, W. I., Cholidah, S. N., & Agus, F. (2024). Agroecological Nutrient Management Strategy for Attaining Sustainable Rice Self-Sufficiency in Indonesia. *Sustainability*, *16*(2), 1–29. https://doi.org/10.3390/su16020845

Tana, T., & Woldesenbet, M. (2017). Effect of Combined Application of Organic and Mineral Nitrogen and Phosphorus Fertilizer on Soil Physico-chemical Properties and Grain Yield of Food Barley (Hordeum vulgare L .) in Kaffa Zone , South-western Ethiopia. *Momona Ethiopian Journal of Science (MEJS)*, *9*(2), 242–261. https://doi.org/10.4314/mejs.v9i2.8

Vinod, H. D. (2022). Kernel Regression Coefficients for Practical Significance. *Journal of Risk and Financial Management*, *15*(1), 32. https://doi.org/10.3390/jrfm15010032

Wang, Y. (2025). A Path to Agricultural Fertilizer Non-Point Source Pollution Control Enabled by Big Data and Machine Learning. *Polish Journal of Environmental Studies*, *XX*(X), 1–14. https://doi.org/10.15244/pjoes/210392

Weraikat, D., Sori, K., Zagar, M., & Soka, M. (2024). Data Analytics in Agriculture : Enhancing Decision-Making for Crop Yield Optimization and Sustainable Practices. *Sustainability*, *16*(17), 1–12. https://doi.org/10.20944/preprints202406.1042.v1

Zulkifli, Besse Dahliana, S. R. (2025). Food Security Challenges And Policy Responses In South Sulawesi : Toward Sustainable Self-Sufficiency. *Multidisciplinary Indonesian Center Journal (MICJO)*, *2*(4), 4311–4318. https://doi.org/10.62567/micjo.v2i4.1249