BIO Web of Conferences **201**, 02010 (2025) https://doi.org/10.1051/bioconf/202520102010
ICoN-BEAT 2025

Production risk and its determinants in soybean farming: evidence from three major production centers in East Java, Indonesia

Bambang Siswadi[1*], *Dina Kartika* Sari[1], and *Lia Rohmatul* Maula[1]

[1]Faculty of Agriculture, Universitas Islam Malang, Malang, East Java, Indonesia

Abstract. This study investigates the magnitude and determinants of soybean production risk across three central East Java, Indonesia districts : Pasuruan, Jember, and Sampang. Using cross-sectional survey data from 90 soybean farmers, the analysis employs a Cobb-Douglas production function and the Just and Pope variance function to estimate risk levels, alongside a Probit regression to identify key influencing factors. The results indicate that seed input, urea fertilizer, and labor significantly enhance productivity, while the Sampang district exhibits statistically lower yields than Pasuruan. Organic manure is the only input identified as a risk-reducing factor, significantly lowering output variability. In contrast, Sampang is associated with higher production variance, reflecting elevated exposure to production risks. Furthermore, farming experience, land ownership, and access to extension services significantly reduce the probability of farmers facing high production risk. These findings highlight the importance of both agronomic and institutional interventions in mitigating production risk, especially in structurally vulnerable regions like Sampang.

1 Introduction

Soybean is classified as a nationally strategic commodity in Indonesia due to its essential role in fulfilling the population's demand for plant-based protein, particularly as a primary ingredient in staple foods such as tempeh and tofu. Despite its significance, Indonesia continues to experience a substantial gap between domestic soybean production and consumption, resulting in heavy reliance on imports [1]. Two significant barriers to achieving soybean self-sufficiency are persistently low productivity and elevated levels of production risk [2]. On a global scale, several studies have identified climate variability, pest infestations, inefficient input utilization, and limited access to agronomic innovation as critical drivers of yield fluctuations in agricultural production, including soybeans [3] demonstrated that the early detection of drought stress through phenotyping technologies could significantly mitigate yield losses. However, adopting such advanced technological solutions remains largely confined to developed nations.

* Corresponding author: bsdidiek171@unisma.ac.id

In the Indonesian context, local research emphasizes that yield disparities among regions are influenced by agronomic factors such as seed quality, fertilizer application, land conditions, and the socio-economic status of farmers [4]. Globally, agricultural risk assessment has advanced through the application of quantitative approaches. Models such as Value at Risk (VaR), Expected Shortfall, and stochastic optimization are increasingly employed to manage risks associated with market volatility and climate uncertainty. In particular, the Cobb-Douglas production function, combined with the heteroskedastic residual specification introduced by Just and Pope, has been proven effective in modeling production variability.

East Java is among the leading soybean-producing provinces in Indonesia, with Pasuruan, Jember, and Sampang identified as key production centers [5] . Despite their importance, comparative studies examining production risk levels and their determinants across these districts remain limited. Existing research tends to be fragmented either focusing on a single region or lacking robust econometric risk modelling.

Moreover, essential determinants of production risk such as access to extension services, farmers' risk management behavior, and institutional support have not been extensively analyzed using quantitative regression models such as Probit [6]. A deeper understanding of the probability of farmers encountering high production risks is critical for developing targeted, location-specific policies, particularly under the current pressures of climate change and increasing input costs.

In response to these gaps, this study aims to comprehensively assess soybean production risk in three central districts of East Java and identify its key determinants using theoretically sound and empirically validated econometric models. Specifically, the study has two main objectives: (1) to estimate the magnitude of production risk in soybean farming using the Cobb-Douglas production function and the Just and Pope heteroskedastic variance approach, and (2) to identify the primary determinants of production risk through a Probit regression model that includes variables such as agronomic inputs, land type, local climate conditions, institutional support, and access to agricultural extension services.Formatting the title, authors and affiliations

2 Material and Methods

2.1 Research Approach

This study employs a quantitative, cross-sectional design based on a field survey of soybean farmers in Pasuruan, Jember, and Sampang regencies in East Java. This methodology enables a robust statistical examination of the relationships between production inputs and soybean output and estimates production risk and its determinants. Similar recent studies using household-level survey data have demonstrated that cross-sectional quantitative approaches facilitate the identification of socio-economic and institutional factors influencing farmers' exposure to risk [7]. These methods support nuanced policy development tailored to regional differences and institutional contexts.

2.2 Research Location and Sampling Technique

The research sites were deliberately selected based on their high soybean production volumes and recognition as central East Java producing areas. Our unit of analysis consisted of active soybean farmers who had cultivated the crop for at least the two most recent planting seasons. Employing a stratified random sampling design, stratification was carried out by districts, Pasuruan, Jember, and Sampang, to ensure equal representation of 30 farmers per district, resulting in a total sample size of 90 respondents. This approach enhances representativeness and improves estimation precision by reducing sampling error across heterogeneous subpopulations. Stratified random sampling stratifies a diverse population into homogeneous groups before random selection, ensuring balanced coverage and reliable inference across strata [8] .

2.3 Types and Sources of Data

This study utilizes primary data from structured questionnaires, in-depth interviews, and field observations. The data gathered include:
(a) Farmers' socio-economic characteristics,
(b) Use of production inputs (seeds, fertilizers, pesticides, labor, and land area),
(c) Output per planting season, and
(d) Farmers' perceptions of production risks and past farming experiences.

2.4 Analytical Method

The data analysis was conducted in three main stages:

2.4.1 Production Model (Cobb-Douglas Function)

To examine the effect of various production inputs on soybean yield, a multiple linear regression model in log-linear form was applied, based on the Cobb-Douglas production function:

$$\ln Yi = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \ldots\ldots\ldots \beta_k \ln X_{ki} + Z_1 D_1 + Z_2 D_2 + E \quad (1)$$

Explanation:
Yi = Soybean production (kg)
$X_1 \ldots\ldots X_{ki}$ = Production input factors (seed, urea, phosphorus, organic manure, pesticide, and labor)
D_1 = Dummy variable for Jember
D_2 = Dummy variable for Sampang
E_i = Error term

Pasuruan is used as the reference category (base group)
The coefficients Z_1 and Z_2 indicate whether there are statistically significant differences in productivity levels between Jember and Sampang compared to Pasuruan.

2.4.2. Production Risk Model (Just and Pope – Production Variance Approach)

To measure the level of production risk, the Just and Pope approach was employed using a two-stage estimation process:
1. Mean production function (as specified above).
2. Residual variance function.

$$LnV^2 i = a_0 + a_1 \ln X_{1i} + \ldots\ldots\ldots a_k \ln X_{ki} + Z_1 D_1 + Z_2 D_2 + U_i \quad (2)$$

Explanation:

$V^2 i$ = represents the variance of the production residuals, a proxy for production risk. Dummy variables D_1 and D_2 are again used to represent the locations of Jember and Sampang, respectively.
The coefficients Z_1 and Z_2 indicate whether the level of production risk in Jember and Sampang is significantly higher or lower than that in Pasuruan.

2.4.3 Determinants of Production Risk (Probit Model)

A Probit regression model was employed to identify the factors influencing the probability of farmers experiencing high production risk. The dependent variable is a binary (dummy) variable indicating the presence or absence of high production risk.

$$P(RISK_1) = 1 = \Phi(\theta_0 + \theta_{1i} Z_{1i} + \ldots\ldots\ldots \theta_m Z_{mi} + \gamma_1 D_1 + \gamma_2 D_2$$

Explanation:
Z_m = Independent variables (e.g., education, farming experience, land ownership, household size, extension service intensity, and farm diversification)
D_1, D_2 = District dummy variables (as defined above)
Φ = Cumulative normal distribution function
The coefficients of the district dummy variables indicate whether geographic location (District) has a statistically significant effect on the likelihood of farmers facing high production risk. Data analysis was performed using Stata 17, employing log-linear regression for the production function and Probit regression for the risk determinant model.

3 Results and Discussion

3.1. Respondent Profile

This study involved ninety soybean farmers equally selected from the three primary production regions in East Java: Pasuruan, Jember, and Sampang. Each District contributed 30 respondents, selected using stratified random sampling. As shown in Table 1, the majority of farmers were aged between 41 and 60 years, with most having completed only primary education. The average farming experience was 13.4 years. Monoculture was the dominant planting system, with a minority engaging in crop diversification, typically with maize or rice.

Table 1. Socioeconomic Characteristics of Soybean Farmers in Three District

No	Variable	Pasuruan (n=30)	Jember (n=30)	Sampang (n=30)	Average
1	Age (years)	52.3	50.7	54.1	52.4
2	Education (years of schooling)	6.8	7.4	5.9	6.7
3	Farming Experience (years)	13.2	14.1	12.8	13.4
4	Land Area (ha)	0.64	0.76	0.58	0.66
5	Land Ownership (% self-owned)	67%	73%	59%	66%
6	Extension Access (% with access)	40%	47%	27%	38%
7	Cropping System (Mono/Poly)	Mono 77%	Mono 70%	Mono 90%	Mono 79%

Source: Primary Data Processed, 2025

The socioeconomic characteristics of soybean farmers across Pasuruan, Jember, and Sampang reveal notable disparities. Farmers in Sampang are generally older (54.1 years) and less educated (5.9 years), compared to their counterparts in Jember, who have higher education levels (7.4 years) and slightly more farming experience (14.1 years). Jember also has the largest average landholding (0.76 ha) and the highest proportion of land ownership (73%), suggesting greater resource access and security. In contrast, farmers in Sampang cultivate smaller areas (0.58 ha) with lower ownership rates (59%), which may constrain productivity and investment.

Access to agricultural extension services is limited, with only 38% of respondents reporting participation, and Sampang again showing the lowest access (27%). Limited exposure to advisory programs may hinder farmers' ability to adopt improved practices and manage risks [9]. Monoculture is prevalent in all regions, especially in Sampang (90%), increasing susceptibility to climatic and market-related shocks due to a lack of diversification.

3.2. Production Function Estimation Results

Table 2 presents the log-linear regression results from the Cobb-Douglas production function, including location dummy variables for Jember and Sampang, with Pasuruan as the base category.

Table 2. Estimation Results of the Soybean Production Function (Cobb-Douglas Model)

Variable	Coefficient	Std. Error	t-Stat	p-Value	Interpretation
Intercept (Constant)	1.245	0.184	6.77	0.000	Significant at 1%
ln(Seed, kg/ha)	0.273	0.067	4.07	0.000	Significant at 1%
ln(Urea, kg/ha)	0.352	0.089	3.95	0.000	Significant at 1%
ln(Manure, kg/ha)	0.058	0.041	1.42	0.160	Not significant
ln(Pesticide, ml/ha)	0.041	0.033	1.24	0.218	Not significant
ln(Labor, workdays/ha)	0.298	0.072	4.14	0.000	Significant at 1%
Dummy Jember (D1)	0.082	0.065	1.26	0.210	Not significant
Dummy Sampang (D2)	-0.134	0.058	-2.31	0.024	Significant at 5%

Source : Primary Data Processed, 2025

The results of the log-linear regression model indicate that seed usage, urea fertilizer, and labor inputs significantly enhance soybean productivity, as shown by their positive and statistically significant coefficients at the 1% level. Specifically, a 1% increase in seed application is associated with a 0.273% increase in output, while similar increases in urea and labor contribute 0.352% and 0.298% increases, respectively. These elasticities imply that these inputs are crucial for achieving optimal productivity in soybean farming and support previous findings on input responsiveness in smallholder agriculture [10] .

Conversely, the coefficients for manure and pesticide inputs are statistically insignificant, suggesting limited or inconsistent effects on output in the sampled context. The negative and significant coefficient for the Sampang dummy variable (−0.134, $p = 0.024$) indicates that soybean yields in this district are significantly lower than in the reference area (Pasuruan), possibly due to structural constraints or agroecological factors. The dummy variable for Jember was statistically insignificant, implying no meaningful difference in productivity relative to Pasuruan. The R^2 value of 86.3% confirms that the model provides a strong fit and explains most of the variability in soybean production outcomes, consistent with prior studies on input-output modeling in crop production [11].

3.3. Results of the Production Variance Function Estimation

Table 3 presents the estimation results of the production variance function using the Just and Pope framework. This model assesses how inputs variables and geographic location influence production risk, proxied by the variance of the residuals. The analysis focuses on three districts : Pasuruan (reference category), Jember, and Sampang.

Table 3. Estimation Results of the Just and Pope Production Variance Model

Variable	Coefficient	Std. Error	t-Statistic	p-Value	Remark
Intercept	-3.114	1.104	-2.82	0.006	Significant at 1%
ln(Seed, kg/ha)	-0.084	0.093	-0.90	0.371	Not Significant
ln(Urea, kg/ha)	0.109	0.097	1.12	0.265	Not Significant
ln(Organic manure, kg/ha)	-0.164	0.078	-2.10	0.038	Significant at 5%
ln(Pesticide, ml/ha)	0.073	0.058	1.26	0.212	Not Significant
ln(Labor, man-days/ha)	0.091	0.089	1.02	0.309	Not Significant
Dummy Jember (D1)	0.056	0.075	0.75	0.456	Not Significant
Dummy Sampang (D2)	0.187	0.069	2.71	0.008	Significant at 1%

Source : Primary data, processed (2025)
Note : The model uses squared residuals from the production function as the dependent variable.

The analysis indicates that organic manure acts as a risk-mitigating input, evidenced by its negative and statistically significant coefficient (–0.164, $p< 0.05$). This suggests that higher usage of organic manure proportionally decreases production variance, thereby reducing risk. This result supports existing literature highlighting the role of organic inputs in enhancing soil structure, cation exchange capacity, and water retention all essential for stabilizing crop growth, especially in suboptimal or marginal lands.

Furthermore, the Sampang dummy variable shows a positive and significant coefficient (0.187, $p < 0.01$), implying that farmers in Sampang experience significantly higher production risk than those in Pasuruan. This heightened risk is likely due to Sampang's relatively arid agroecological conditions, challenging topography, and the use of limited or non-standardized inputs. The vulnerability in Sampang is further intensified by inadequate agricultural extension services and low adoption of resilient technologies.

By contrast, the coefficient for Jember is statistically insignificant, indicating that production risk levels in Jember are not significantly different from those in Pasuruan.

Other production inputs, such as seed, urea, pesticide, and labor demonstrate mixed directions of influence but lack statistical significance. While these inputs play important roles in enhancing output, as highlighted in the Cobb-Douglas production function, they do not appear to influence yield variability directly . According to the Just and Pope framework, inputs can be classified as risk-neutral, risk-enhancing, or risk-reducing; not all yield-enhancing inputs necessarily affect risk levels.

These findings highlight the importance of increasing input usage and selecting inputs that contribute to production stability in soybean farming. In this context, encouraging the application of organic manure and strengthening institutional capacity in high-risk regions such as Sampang are critical strategies. The R^2 value of 29.3% is considered reasonable for a variance function model, as production risk is often influenced by external factors such as climate variability and extreme weather that are difficult to quantify.

3.4. Probit Model Estimation Output

Table 4 presents the Probit regression results analyzing the factors influencing high production risk among soybean farmers in Pasuruan, Jember, and Sampang districts. The dependent variable is a binary indicator of production risk coded as 1 for high risk and zero otherwise, classified based on the median value of residual variance from the production function.

Table 4. Probit Regression Results: Determinants of High Production Risk

Variable	Coefficient	Std. Error	z-value	p-value	Significance
Intercept (Constant)	0.913	0.722	1.26	0.209	Not significant
Farmer's Age (years)	0.008	0.015	0.53	0.594	Not significant
Formal Education (years)	-0.041	0.031	-1.32	0.188	Not significant
Farming Experience (years)	-0.064	0.027	-2.37	0.018	Significant at 5%
Land Ownership (1 = yes)	-0.492	0.212	-2.32	0.020	Significant at 5%
Extension Access (1 = yes)	-0.614	0.241	-2.55	0.011	Significant at 5%
Off-farm Diversification (1 = yes)	-0.328	0.256	-1.28	0.200	Not significant
Dummy Jember (D1)	-0.145	0.193	-0.75	0.454	Not significant
Dummy Sampang (D2)	0.503	0.198	2.54	0.011	Significant at 5%
Log-Likelihood	-41.76				
Pseudo R^2	0.238				
Prob > Chi^2	0.0001				Model is significant
Sample Size (N)	90				

Source : Primary data, processed (2025)

Farming experience has a negative and statistically significant effect on high production risk. Each additional year of experience reduces the probability of facing high risk, implying that experienced farmers are better equipped to make adaptive and efficient production decisions [12]. Land ownership also significantly reduces the likelihood of experiencing high production risk. Farmers cultivating their land tend to be more committed and manage their inputs more efficiently than renters. Access to agricultural extension services shows a significant adverse effect on production risk. Farmers who have received extension support are more likely to adopt agronomic practices that mitigate vulnerability to weather extremes and pest attacks.

The dummy variable for Sampang has a positive and significant coefficient, indicating that soybean farmers in Sampang face a statistically higher probability of encountering high production risk compared to those in Pasuruan. This may stem from less favorable agro-climatic conditions, limited resources, and weaker institutional support. On the other hand, the dummy variable for Jember is not statistically significant, suggesting that the probability of high production risk in Jember is not meaningfully different from Pasuruan. The model's Pseudo R^2 value of 23.8% is considered acceptable for a binary choice model in socio-agricultural studies.

4 Conclusion

This study successfully addressed its dual objectives by employing both the Cobb-Douglas production model and Pope's variance function approach to quantify soybean production risk and applying the Probit model to identify its determinants. First, the empirical findings reveal that seed quantity, urea fertilizer, and labor input significantly and positively affect soybean output, confirming their elasticity in enhancing productivity. However, the Sampang district dummy exhibits a negative and statistically significant coefficient, indicating lower productivity than Pasuruan. Second, only organic manure significantly reduces production variability among all agronomic inputs analyzed, thus serving as a risk-reducing input. Furthermore, Sampang also demonstrates a significantly higher production variance, highlighting a greater level of production risk in that region. Lastly, key structurale and institutional variables, farming experience, land ownership, and access to agricultural extension services significantly reduce the likelihood of high production risk. These results suggest that Sampang's production risk is driven by input use and underlying agroclimatic and institutional constraints.

5 Policy Implication

Based on the findings, several policy recommendations can be proposed to mitigate production risks among soybean farmers. First, location-specific policies should be implemented, particularly in high-risk areas such as Sampang Regency, through developing drought-tolerant crop varieties and improving basic irrigation infrastructure. Second, central and local governments should expand adaptive agricultural extension programs, emphasizing cultivation techniques that reduce yield uncertainty. Community-based extension approaches, such as farmer-to-farmer learning, enhance the diffusion of innovations. Third, providing subsidies or incentives for the use of organic fertilizers (e.g., manure) should be considered, especially on marginal lands, as empirical evidence indicates that such inputs help stabilize yields. Fourth, strengthening farmers' land tenure whether through agrarian reform or long-term lease schemes could reduce production risks by encouraging greater commitment and long-term investment in land management. Lastly, enhancing farmers' managerial capacity through technical and financial training is essential to improve their adaptive capabilities in facing climate-related and market-driven uncertainties.

References

1. Bibi, Z., Khan, D. (2021). Technical and environmental efficiency of agriculture sector in South Asia: A stochastic frontier analysis approach. *Environment, Development and Sustainability*, *23*(6), 9260–9279.
2. Fattah, A., & Tahir, G. (2018). Analisis efisiensi produksi kedelai di Sulawesi Selatan. *Jurnal Ilmu Pertanian Indonesia*, *23*(1), 10–19. https://doi.org/https://doi.org/10.18343/jipi.23.1.10
3. Akhavizadegan, H., Shahraki, A., & Mohammadi, H. C. A. (2022). Risk-averse stochastic optimization for crop planning under uncertainty. *Computational Agriculture*, *11*(3), 205–220. https://doi.org/https://doi.org/10.1016/j.compag.2022.106113
4. Cotter, J., & Hanly, J. (2021). Extreme risk in agricultural commodities: Value at Risk, Expected Shortfall, and SRM. *Journal of Risk Finance*, *22*(2), 123–140. https://doi.org/https://doi.org/10.1108/JRF-07-2020-0152
5. Sitanggang, A. P., & Silalahi, L. (2020). Efisiensi teknis usahatani kedelai pada lahan pasang surut di Jambi. . *Agro Ekonomi*, *38*(1), 25–36. https://doi.org/https://doi.org/10.21082/jae.v38n1.2020.25-36
6. Krisdiana, I., Purwanti, R., & Rachman, B. (2021). Financial feasibility and competitiveness of soybean varieties in Indonesia. *Sustainability*, *13*(15). https://doi.org/https://doi.org/10.3390/su13158334
7. Roessali, W., Budiraharjo, K., & Nurfadillah, D. (2023). Farmers' behavior in facing soybean production risks in Grobogan. *Jurnal Sosial Ekonomi Pertanian*, *10*(2), 150–161. https://doi.org/https://doi.org/10.14710/jsep.2023.150-161

8. Kalogiannidis, S., Papadopoulou, C.-I., Loizou, E., & Chatzitheodoridis, F. (2023). Risk, Vulnerability, and Resilience in Agriculture and Their Impact on Sustainable Rural Economy Development: A Case Study of Greece. *Agriculture*, *13*(6). https://doi.org/https://doi.org/10.3390/agriculture13061222
9. Särndal, C.-E., Swensson, B., & Wretman, J. H. (2020). *Model Assisted Survey Sampling (with discussion).* Springer. https://doi.org/https://doi.org/10.1007/978-3-30-45736-1
10. Chandra, A., Djanibekov, U., & Finger, R. (2020). Adaptive capacity and climate change adaptation in agriculture: Evidence from rice farming in Indonesia. *Ecological Economics*. https://doi.org/https://doi.org/10.1016/j.ecolecon.2020.106874
11. Rahman, S., & Barmon, B. K. (2021). Determinants of technical efficiency and production risk of maize farming in Bangladesh. *Journal of the Saudi Society of Agricultural Sciences*, *20*(3), 180–187. https://doi.org/https://doi.org/10.1016/j.jssas.2020.01.004
12. Bhatt, D., Sharma, M., & Sharma, V. (2022). Nutritional and health benefits of soybean: A review. *Frontiers in Nutrition*. https://doi.org/https://doi.org/10.3389/fnut.2022.878957