Forecast and soybean production base in East Java

Soetriono¹, A D Maharani¹* and D B Zahrosa¹

¹Department of Agribusiness, Faculty of Agriculture, University of Jember

E-mail: ariqdewi.faperta@unej.ac.id

Abstract. East Java Province as one base regions of the soybean commodity. Soybean also provides added value, when processed into tofu, tempeh, chips, soy sauce and other products. The fact is that agro-industry activities are still dependent on imported raw materials. The objectives of this study is to forecast soybean commodity production and find out soybean production base in East Java. The locations of the study are East Java Province, because it has the largest harvest area and highest production in Indonesia. The research method applies an analytical method. The data collection method uses secondary data, in which the available statistical data has been published by BPS. The data analysis method used is base and non-base sector analysis Location Quotient and Box-Jenkins Model (ARIMA) analysis. The results show that out of 38 district and city in East Java, there are 10 district soybeans production base. The district of soybean production base is Sampang Banyuwangi, Nganjuk, Bangkalan, Lamongan, Bojonegoro, Trenggalek, Blitar, Ponorogo and Jember District. The highest LQ value of soybean base production is Sampang district of 5.49; Banyuwangi district of 2.77 and Nganjuk district of 2.00. The development of soybean production forecasting in East Java in 2018-2027 fluctuations, tends to be constant and increase.

1. Introduction
Soybeans are the main source of vegetable protein and vegetable oils in the world. Soybeans are widely used as a staple in food processing. [1] The consumption of soybean in Indonesia reaches 2.2 million tons per year, but about 1.6 million tons must be imported. National soybean production for the 2013-2017 period decreased by an average of 6.37% per year. A significant decrease in soybean production occurred in 2017 of 36.90%, from 2016 production of 859,650 tonnes to 542,450 tonnes in 2017. The decline in production was the excess of the reduction in soybean harvested area in 2017 by 38.13% or an area of 220,010 hectares, from 2016 of 576,990 hectares to 356,980 hectares in 2017. East Java Province is one of the soybean producers in Indonesia. [2] East Java's soybean production contributed 37.33% of national production. East Java, especially in soybean production centers, has high hopes for economic viability, both for meeting local, inter-regional and export needs [3]. In fact, it has not been able to meet the local needs of soybeans every year, where soybean production in East Java continues to decline every year. Until now, soybean production in East Java is still insufficient for the consumption needs of people in East Java, which reaches 447,000 tons per year. In 2017, soybean production was 133,593 tons, resulting in a deficit of 313,407 tons [4].

Table 1. The Production of soybean in east java from 2013-2017

| Year | Land Area (Ha) | Production(Ton) | Produktivitas (Ton/Ha) |
|------|----------------|-----------------|------------------------|
| 2004 | 319.492        | 246.940,0       | 1,3                    |
| 2005 | 335.106        | 255.443,0       | 1,3                    |
| 2006 | 320.205        | 246.534,0       | 1,3                    |
| 2007 | 158.525        | 127.368,0       | 1,2                    |
| 2008 | 277.281        | 216.828,0       | 1,3                    |
| 2009 | 355.260        | 264.779,0       | 1,3                    |
Table 1 shows that, the area of soybean land in East Java has decreased starting in 2012. If soybean production in East Java is not increased, imported soybeans will continue to dominate the East Java soybean market and that means soybean farmers will have difficulty competing with soybeans from overseas. One aspect that must be observed in the development of soybean is how to regard the regional conditions in which each region has different characteristics. Certain studies are needed so that the development of soybean can be carried out in suitable areas. Therefore, the objectives of this study are: (1) to forecast soybean production and (2) find out the soybean production base in East Java.

2. Research Methodology
The areas of the study are East Java Province, because it has the largest harvest area and highest production in Indonesia. The research method applies an analytical method. The data collection method uses secondary data, in which the available statistical data has been published by BPS. The analysis data used is base and non-base sector analysis Location Quotient [5] and Box-Jenkins Model (ARIMA) analysis [6].

3. Soybean Production Base Area in East Java
The results of Location Quotient Analysis (LQ) are carried out using the indicator data on the amount of soybean production with the total production in East Java. Analysis conducted in this study was a study of the value of Location Quotient (LQ) for a period of five years starting from 2013 to 2017. Location Quotient Analysis (LQ) can also be used to determine the magnitude of the role of a sector in a region against the magnitude of the sector roles overall [7]. The following table is the result of the Location Quotient (LQ) analysis of soybean in East Java.

| No | Sub-district | 2013  | 2014  | 2015  | 2016  | 2017  | Average | Notes   |
|----|--------------|-------|-------|-------|-------|-------|---------|---------|
| 1  | Pacitan      | 0.54  | 0.48  | 0.69  | 0.58  | 0.28  | 0.51    | Non Basis |
| 2  | Ponorogo     | 1.12  | 1.14  | 1.61  | 1.35  | 0.58  | 1.16    | Basis    |
| 3  | Trenggalek   | 1.14  | 0.98  | 1.19  | 1.49  | 2.56  | 1.47    | Basis    |
| 4  | T. Agung     | 0.86  | 0.68  | 0.60  | 0.16  | 0.26  | 0.51    | Non Basis |
| 5  | Blitar       | 2.06  | 0.95  | 1.30  | 1.30  | 1.43  | 1.41    | Basis    |
| 6  | Kediri       | 0.02  | 0.02  | 0.13  | 0.10  | 0.01  | 0.06    | Non Basis |
| 7  | Malang       | 0.06  | 0.03  | 0.01  | 0.07  | 0.01  | 0.04    | Non Basis |
| 8  | Lumajang     | 0.27  | 0.40  | 0.37  | 0.18  | 0.10  | 0.26    | Non Basis |
| 9  | Jember       | 1.35  | 1.01  | 1.11  | 1.28  | 0.98  | 1.15    | Basis    |
| 10 | Banyuwangi   | 0.57  | 3.11  | 2.60  | 3.02  | 4.54  | 2.77    | Basis    |
| 11 | Bondowoso    | 0.01  | 0.01  | 0.00  | 0.00  | 0.01  | 0.01    | Non Basis |
| 12 | Situbondo    | 0.00  | 0.00  | 0.02  | 0.01  | 0.00  | 0.01    | Non Basis |
| 13 | Probolinggo  | 0.04  | 0.02  | 0.04  | 0.01  | 0.00  | 0.02    | Non Basis |
| 14 | Pasuruan     | 1.83  | 1.34  | 0.89  | 0.59  | 0.36  | 1.00    | Non Basis |
| 15 | Sidoarjo     | 0.27  | 0.43  | 0.58  | 0.23  | 0.33  | 0.37    | Non Basis |
| 16 | Mojokerto    | 0.82  | 0.56  | 0.55  | 0.38  | 0.53  | 0.57    | Non Basis |
| 17 | Jombang      | 1.15  | 0.98  | 0.94  | 0.72  | 0.99  | 0.96    | Non Basis |
Based on table 2, LQ soybean calculation results were obtained within a period of five years, from 2013 to 2017 in East Java. LQ values of less than one (LQ < 1) were found 18 district and 8 city, that is Pacitan, Tulungagung, Kediri, Malang, Lumajang, Bondowoso, Situbondo, Probolinggo, Pasuruan, Sidoarjo, Mojokerto, Jombang, Madiun, Magetan, Ngawi, Tuban, Gresik, Pamekasan, Sumenep district, Kediri city, Blitar city, Probolinggo city, Pasuruan city, Mojokerto city, Madin city, Surabaya and Batu city. LQ values of more than one (LQ > 1) were found 10 district, that is Ponorogo, Trenggalek, Blitar, Jember, Banyuwangi, Nganjuk, Bojonegoro, Lamongan, Bangkalan dan Sampang District. LQ value of soybean in Ponorogo District is 1.16; Trenggalek District is 1.47; Blitar District is 1.41; Jember District is 1.15; Banyuwangi District is 2.77; Nganjuk District is 2.00; Bojonegoro District is 1.52; Lamongan District is 1.68; Bangkalan district is 1.80 and Sampang district is 5.49. As seen from the results of the LQ analysis, 10 district for 38 District in East Java are soybean base areas in East Java. Graphically, the average LQ value of soybean in East Java for five years in a row can be determined through the following figure.

![Figure 1. Average LQ value of soybean in East Java](image-url)
Based on the results of the analysis, several soybean base areas in East Java, especially in Banyuwangi, Sampang and Nganjuk District. Each LQ value of these districts has the highest LQ value among the three other districts; they are respectively 2.77 and 2.00. Three district has the largest production in East Java.

4. Development of Soybean Production in the Next 10 Year

Efforts to increase soybean production in East Java Province need to be adjusted to the suitability of the region and land for the growth of soybean plants which can make soybeans the leading commodity of the region. Determining the superior commodities of a region is the first step towards agricultural development that is based on the concept of efficiency to gain competitive and comparative advantages in the face of trade globalization. Steps towards efficiency can be taken by developing commodities that have these advantages. The method used to identify leading commodities is to use the Location Quotient (LQ) method.

![Figure 2. The production of soybean in East Java in 2004-2017](image)

Based on the graph above, the production of soybean in East Java from 2004 to 2010 has fluctuated. The production of soybean in 2012 until 2017 tend to decrease. In 2012, the production of soybean in East Java is 220.815 ton. In 2013 until 2017 tend constant to decrease until 133.593 ton.

The selected model is used to predict soybean production in East Java from 2018 to 2027. The results obtained from forecasting soybean production in 2018 to 2027 tend to decrease. The results of this forecast accord with the data used and the resulting plot values of soybean production in East Java using annual data from 2000 to 2017 tends to fluctuate. The results of forecasting soybean production in East Java were obtained from the ARIMA method by conducting a stationary data test.

Stationary data testing is salient because the ARIMA method requires the criteria of the analyzed data must be stationary first [7]. It can translate data and economic models well because stationary data is not too varied and tends to approach the average value. The stationarity of the data used can be determined using the value of Augmented Dickey-Fuller (ADF). The ADF value of soybean production is presented in the following table.
Table 3. ADF value and critical value data of soybean production data in order 1 differencing

| Notes                              | t-Statistic | Prob.* |
|------------------------------------|-------------|--------|
| Augmented Dickey-Fuller test statistic | -4.842404  | 0.0009 |
| Test critical values:              |             |        |
| 1% level                           | -3.769597   |        |
| 5% level                           | -3.004861   |        |
| 10% level                          | -2.642242   |        |

The ADF value in table 1 is used for first-order differencing because the ADF t-statistic value is greater than 1%, 5%, and 10% on t-test critical value. After differentiating on the zero-order level, the value still did not meet stationary data criteria so that the first-order differencing is used to get stationary data. In this table, the ADF value for soybean production data is -4.842404. The ADF value is lower than the critical value of 1% (-3.769597). Moreover, the ADF value is lower than the critical value of 5% (-3.004861) and the ADF value is also lower than the critical value 10% (-2.642242). These values indicate that the soybean production data used is well stationary at critical values of 1%, 5%, and 10%.

The next step to be done after analyzing the plot and stationary data is eliminating inappropriate models. Estimation can be made if the ARIMA model has formed an AR (autoregressive) model from the autocorrelation function pattern and the MA (Moving Average) model is formed from partial autocorrelation. ACF-PACF plot estimations are carried out on previously identified production data. The results of the ACF-PACF plot analysis are presented in the following table.

Table 4. ACF-PACF value of soybean production data in east java

| Autocorrelation | Partial Correlation | AC  | PAC  | Q-Stat | Prob  |
|-----------------|---------------------|-----|------|--------|-------|
| . **** | . **** | 1 | 0.587 | 0.587 | 10.719 | 0.001 |
| . * * * | . | 2 | 0.236 | -0.166 | 12.513 | 0.002 |
| . * | . * | 3 | 0.202 | 0.217 | 13.887 | 0.003 |
| . * | . . | 4 | 0.185 | -0.023 | 15.087 | 0.005 |
| . . | . | 5 | 0.110 | -0.001 | 15.532 | 0.008 |
| . | * * | 6 | 0.003 | -0.097 | 15.533 | 0.016 |
| . * | . . | 7 | -0.093 | -0.092 | 15.879 | 0.026 |
| . * | . | 8 | -0.066 | 0.055 | 16.061 | 0.042 |
| . * | . * * | 9 | 0.114 | 0.221 | 16.638 | 0.055 |
| . * | . * | 10 | 0.097 | -0.111 | 17.080 | 0.073 |
| . * | . * | 11 | -0.100 | -0.158 | 17.570 | 0.092 |
| . * | . | 12 | -0.168 | -0.065 | 19.058 | 0.087 |

The step taken in estimated is by looking at ACF and PACF tables so that the ARMA or ARIMA model will be formed. The estimation process is analyzed using the formula $Z = 0.05 \sqrt{1/n} = \pm 1.96\sqrt{1/n}$ which the calculation result is $\pm 0.49$. This value is used to eliminate the AC and PAC models where the best model is seen from values outside $\pm 0.49$. ARIMA models that have been obtained include ARIMA (1,0,1), ARIMA (1,0,0), ARIMA (0,0,1), ARIMA (4,2,2), ARIMA (2,2,2), ARIMA (2,2,0), ARIMA (4,2,0), dan ARIMA (0,2,2).

The next step is conducting a diagnostic test to see whether the model is still feasible to use by looking at the criteria of Adjusted R-squared, Akaike info criterion, Schwarz criterion, S.E of regression, and Sum Square Residua [8]. The adjusted R squared value must be higher than all selected models. On the other hand, the value of Akaike info criterion, Schwarz criterion, S.E of regression, and Sum Square
Residual must be lower than all existing models. The model that best meets the criteria will be selected as the best model for forecasting the soybean production in East Java. Based on these criteria, the best forecasting model was chosen, namely ARIMA (1,0,1). The results of the forecasting test with the ARIMA model (1,0,1) are presented in the following table.

| Model       | Adj R-Square (terbesar) | Akaike info criterion (terkecil) | S.E Regression (terkecil) | Sum Squared Residual (terendah) |
|-------------|-------------------------|----------------------------------|---------------------------|---------------------------------|
| ARIMA (1,0,1) | 0.550792                | 23.53023                         | 28187.21                  | 1.91E+10                        |
| ARIMA (1,0,0) | 0.414025                | 23.72579                         | 32193.48                  | 2.59E+10                        |
| ARIMA (0,0,1) | 0.509544                | 23.58344                         | 29452.94                  | 2.17E+10                        |
| ARIMA (4,2,2) | -0.011412               | 24.28507                         | 42925.35                  | 4.29E+10                        |
| ARIMA (2,2,2) | -0.027623               | 24.29791                         | 42632.94                  | 4.36E+10                        |
| ARIMA (2,2,0) | 0.005490                | 24.23389                         | 41940.45                  | 4.40E+10                        |
| ARIMA (4,2,0) | -0.013615               | 24.25695                         | 42341.37                  | 4.48E+10                        |
| ARIMA (0,2,2) | -0.0182161              | 24.25445                         | 42438.30                  | 4.50E+10                        |

Based on table 5, the ARIMA model ARIMA (1,0,1) was selected as the best model because the value is suitable with the Adjusted R-squared criteria, Akaike info criterion, Schwarz criterion, S.E of regression, and Sum Square Residuals. The Adjusted R-squared ARIMA (1,0,1) value is 55.07% with R2 of 44.93% and this is the greatest value. Akaike info criterion is seen in the ARIMA model (1,0,1) of 23.53023 which is smaller than the value of Akaike info criterion of other models. The values of S.E of regression and Sum Squared Residual in the ARIMA model (1,0,1) are 28187.21 and 1.91E+10 respectively. These values are smaller than the other models.

The selected model is used to forecast soybean production from 2018 to 2027. The results of the forecasting of soybean production in 2018 to 2027 fluctuations tend to increase. The following graph is the result of soybean production forecasting in East Java in 2018 until 2027:

![Figure 3. Chart of forecasting of soybean production in east java in 2018-2027](image-url)

Figure 3 shows that the development of soybean production in 2018-2027 tends to be increased. It is consistent with the previous year of fluctuating and inclining production. The results of soybean production forecasting in East Java in 2018-2027 are presented in the following table.
Table 6. Forecasting of soybean production in east java in 2018-2027

| Year | Production (ton) |
|------|------------------|
| 2018 | 211.014          |
| 2019 | 255.575          |
| 2020 | 274.292          |
| 2021 | 282.154          |
| 2022 | 285.456          |
| 2023 | 286.843          |
| 2024 | 287.425          |
| 2025 | 287.670          |
| 2026 | 287.773          |
| 2027 | 287.816          |

Table 6 reveals that soybean production in 2018-2027 tends to be fluctuating. The production of soybean in East Java in 2018 was 211.014 ton, it increased in 2019 to 255.575 ton. In 2020, soybean production is 274.292 ton and then it increased again in 2021 to 282.154 ton. Soybean production in East Java continues to increased until 2027 to 287.816 ton. The available data is graphed to find out the patterns of the data used more clearly. The graph of the development of soybean production in East Java is presented as follows.

Figure 4. The development of soybean production in east java

Figure 4 shows that the results of the forecasting of soybean production in East Java from 2018 to 2027 have increase. In 2018, production of soybean is 211.014 ton, tend to increase until 255.575 ton in 2019. The estimated production of soybean in 2020 until 2017 tend to increase, however the increase was only slight. This can be seen on the graph that the line from left to right tends to increase from bottom to top every year.
5. Conclusion
Potential production bases of soybean in East Java are Banyuwangi, Sampang and Nganjuk. The three district have highest LQ values of more than one. The development of soybean production in East Java in 2018-2027 is fluctuating, but start in 2018 tends to increase until 2027. The estimated of soybean production in East Java is 211.014 ton, increase until 287.816 ton in 2027.

Acknowledgment
We gratefully acknowledge the support from LP2M and Research and Development Agency East Java.

References
[1] Badan Pusat Statistik, “Statistik Indonesia dalam Infografis 2018,”.
[2] Kementan, “Outlook Kedelai 2016.” http://epublikasi.pertanian.go.id/download/file/306-outlook-kedelai-2016.
[3] S. Soetiriono, “Rancang Bangun Hulu Hilir, Pemodelan dan Kebijakan Pemerintah pada Agribisnis Kedelai,” JSEP (Journal Soc. Agric. Econ., vol. 4, no. 3, pp. 44–59, 2010.
[4] Badan Pusat Statistik, “Produksi Kedelai di Indonesia,” Dalam angka Indonesia, 2017. https://www.bps.go.id.
[5] W. A. Schaffer, “Regional Models of income determination: simple economic-base theory,” Reg. impact Model., 2010.
[6] T. H. D. Ngo and W. Bros, “The box-jenkins methodology for time series models,” in Proceedings of the SAS Global Forum 2013 conference, 2013, vol. 6, pp. 1–11.
[7] A. Wicaksono and M. A. Rudhito, “The simulation of three crossroad traffic queueing systems using petri nets and colored petri nets program,” in Journal of Physics: Conference Series, 2019, vol. 1307, no. 1, p. 12012.
[8] R. Wibowo, “Konsep, Teori dan Landasan Analisis Wilayah.” Malang: Bayumedia Publishing, 2004.