Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 2091-2099

http://www.ripublication.com

Modelling of Rice Production in South Sulawesi Province With Transfer Function Method

Andriana Ekawati[1], I Made Sumertajaya[2], Farit Mochamad Afendi[3]

Department of Statistics
Bogor Agricultural University
Jalan Pajajaran, Kampus IPB Baranangsiang , Bogor 1615, Indonesia
ekha04_stat@yahoo.co.id, 2.imsjaya@yahoo.com, 3.fmafendi@gmail.com

Abstract

South Sulawesi is a province that has an influence in economic of Indonesia, especially economic growth, and it's included into nine provinces which have the highest contribution to the national income. One of the leading sectors in South Sulawesi are agricultural crops include rice. Rice consumption level is still relatively high. With a population that is expected always increase every year, make the consumption of rice in the community increase too. So, we need important information related to rice, one of which is the production of rice. Multiple time series analysis methods that can be applied to analyze the data of rice production that is not only influenced by the behavior of past data but also involve one or more of the explanatory variables is the transfer function. Transfer function model is a multiple time series forecasting model which combines the characteristics of ARIMA model with characteristics of regression models. The purpose of this study was to perform modeling of rice production involving harvested area without involving the productivity of food crops. The method of this study use transfer function model. Data about rice production and harvested in South Sulawesi from 1st subround 1981 to 3rd subround 2007 are used as case study as a training data, and 1st subround 2008 to 3rd subround 2011 as testing data. The results show that rice production in the current period is affected by the production of one, two, and three previous subround. Besides that, it's also influenced by the area harvested at one and two previous subround, transfer function model has met the assumptions and fairly accurate in forecasting rice production in South Sulawesi Province.

Keywords: Rice Production, Transfer Function

1 Introduction

South Sulawesi is a province that has an influence in the Indonesian economy, especially in the economic growth. In 2013, the contribution of South Sulawesi is about 2.6 percent of national income, and it's included into nine provinces which have the highest contribution to the national income. South Sulawesi is the largest producer of food crops in eastern Indonesia. Predicate as national granary make South Sulawesi position as a manufacturer of considerable potential food crops, especially rice and corn (Herniwati and Kadir 2009). One of the important commodities in the agricultural sector is rice and it's also a main food of the people of Indonesia.

The total population of South Sulawesi has reached approximately 8.5 million people. This amount is expected to increase in subsequent years. Indonesia still includes countries with an average consumption of rice per capita per year is relatively high if compared with the average rice consumption per capita per year in other countries. Increasing of population, make the consumption of rice in the community increase too. Related to that, the policy on rice requires information that is important, accurate, and current. One of them related to rice production.

Multiple time series analysis methods that can be applied to analyze the data of rice production that is not only influenced by the behavior of past data but also involve one or more of the explanatory variables is the transfer function. Transfer function model is a multiple time series forecasting model which combines the characteristics of ARIMA model with characteristics of regression models. In the transfer function model, there is a response variable Yt (output series) are estimated to be affected by the explanatory variables Xt (input series) and other inputs are combined in one group as a disturbance series (noise) ηt.

2 Materials and Methods

Data

The data used in this research is secondary data obtained from annual crop production publications at Central Bureau of Statistics (BPS). They are:

- Rice Production Data in South Sulawesi (subround series), in 1981- 2011.
- Harvested Area Data in South Sulawesi (subround series), in 1981-2011.

The data used were divided into two, training data used to build the model and testing data to see the effectiveness of forecasting results.

Transfer function

Transfer function model is the development of ARIMA models known as multivariate ARIMA, ie, a model associated with one or more input series. If the time series Yt associated with one or more other time series (Xt), then it can be made a time series model to estimate the value of Yt based on Xt information, it is called transfer function model. Mathematically, the transfer function model has the following general form, Wei (2006):

$$y_t = v(B)x_t + \eta_t \quad (1)$$

With
y_t :stationary output series
x_t : stationary input series
η_t :disturbance series *(noise)*
$v(B)$: the weight of the impulse response where $v(B) = \sum_{j=0}^{k} v_j B^j = v_0 + v_1 B + v_2 B^2 + \cdots + v_k B^k$, and k is orde of transfer function.

The weight of the impulse response also can be written in the form

$$v(B) = \frac{\omega_s(B)B^b}{\delta_r(B)} \tag{2}$$

While the disturbance series /noise (η_t) are assumed can be modeled by the ARIMA (p, d, q), so the combination of the model transfer function can be written in the following forms:

$$y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta(B)}{\emptyset(B)} a_t \tag{3}$$

with

$\omega_s(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_s B^s$

$\delta_r(B) = 1 - \delta_1 B - \delta_2 B^2 - \cdots - \omega_r B^r$

$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$

$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$

a_t is a residual at t^{th} time which follow the white noise process. b, s, r, p and q are constants.

Stages of modeling in this study:

a. Preparing the output series (Y_t) and the input series (X_t)

b. Check the stationarity of the output and input series. If the series is not stationary, then performed difrensing and or transformation so that the series becomes stationary.

c. Prewhitening of input series (X_t), ie, the process of transformation correlated series to white-noise behavior that is not correlated.

For example, if the input series Xt is modeled as a ARIMA (p, 0, q), then the model is

$$\emptyset_p(B) X_t = \theta_q(B) \alpha_t \;;\quad \alpha_t \text{ is a random residual}$$

So that the prewhite input series is:

$$\alpha_t = \frac{\emptyset_p(B)X_t}{\theta_q(B)} \tag{4}$$

d. Prewhitening of output series

$$\beta_t = \frac{\emptyset_p(B)Y_t}{\theta_q(B)} \tag{5}$$

e. Calculation of the cross-correlation between the prewhitening input series and prewhitening output series. Cross-correlation function between the αt and βt is

$$\hat{\rho}_{\alpha\beta}(k) = r_{\alpha\beta}(k) = \frac{C_{\alpha\beta}(k)}{S_\alpha S_\beta} \quad ; k=0, \pm1, \pm1, ... \tag{6}$$

with:
$\hat{\rho}_{\alpha\beta}(k)$ = cross-correlation between α_t and β_t at lag k.
$C_{\alpha\beta}(k)$ = covariance between α_t and β_t at lag-*k*
S_α = standard deviation of series α_t.
S_β = standard deviation of series β_t.

f. Determine the order of (b,r,s)
The constant b, r, s are determined based on the plot of cross-correlation function between the αt and βt. How to determine b, r, s are as follows:
- b is determined based on a first significant lag in the cross-correlation plot.
- s can be seen from the next lag that form a clear pattern. Or longer input affect the output after the first significant.
- r indicates how long the output series associated with the previous value of the output series. r seen from ACF plot of the stationary Yt, indicates the significant lag after the first lag. r value can also be determined based on the pattern lag of cross-correlation plot after (b + s). if it has an exponential pattern then r = 1 and if it has a sinus/cosinus wave pattern then r = 2.

g. Preliminary estimate of transfer function
Preliminary estimate of transfer function parameters are $\delta = (\delta_1, \delta_2, ..., \delta_r)$ and $\omega = (\omega_1, \omega_2, ..., \omega_s)$ performed by direct assessment of the weight of the impulse response and utilizing equation

$$V_j = 0 \qquad ; j < b \tag{7}$$

$$V_j = \delta_1 V_{j-1} + \delta_2 V_{j-2} + \cdots + \delta_r V_{j-r} + \omega_0 \qquad ; j = b \tag{8}$$

$$V_j = \delta_1 V_{j-1} + \delta_2 V_{j-2} + \cdots + \delta_r V_{j-r} - \omega_{j-b} \qquad ; j = b+1, \dots, b+s \tag{9}$$

$$V_j = \delta_1 V_{j-1} + \delta_2 V_{j-2} + \cdots + \delta_r V_{j-r} + \omega_0 \qquad ; j > b+s \tag{10}$$

with $\hat{V}_j = \frac{r_{\alpha\beta}(k) s_\beta}{s_\alpha}$

This initial estimate is used as the initial value of the final estimation algorithm to estimate the nonlinear and residual series.

h. ARIMA modeling of the disturbance series

i. Final parameter estimation of transfer function model.

This stage is done by combining the initial parameter estimates with the residual model.

j. Diagnosis of transfer function model with autocorrelation of residual.

k. Forecasting using transfer function model.

3 Results and Discussion

Output series used is the data of rice production in South Sulawesi province at subround I 1981 until subround III 2007. While the input series is rice harvested area in the same period. Based on time-series plot and the results of Augmented Dickey-Fuller test, it appears that the input series and output series has been stationary. Furthermore, identification ARIMA model of input series.

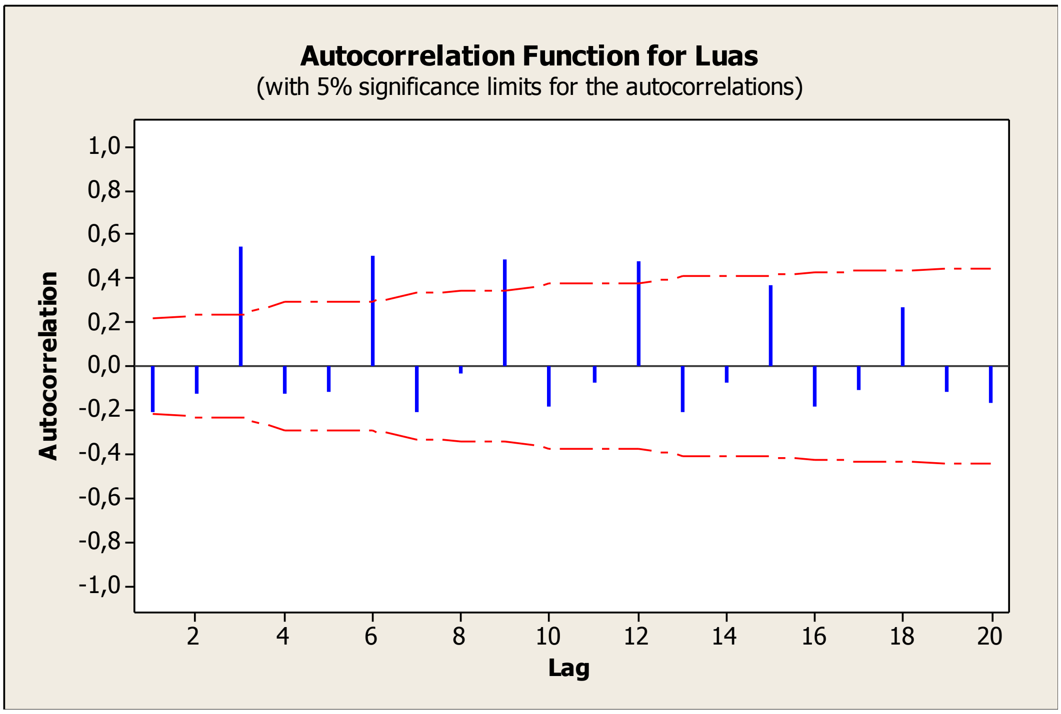

Figure1. Autocorrelation function of the harvested area

Based on ACF plot in figure 1, it appears that the data harvested area is seasonal data with season length s = 3. There are some tentative ARIMA models are formed from the input series. By attention to the smallest value of AIC and SBC, also the significant parameters, then ARIMA model of harvested area is SARIMA (1,0,1)

$(1,0,1)^3$. The model also had to meet the assumptions of nonautocorrelation residual. Based Ljung-Box test showed that all lag has p-value> 5% which means that there is no autocorrelation in the residual value of the area harvested data. Identification ARIMA model for output series production of rice is also has a similar model to the model in harvested area. SARIMA (1,0,1) $(1,0,1)^3$ has the following form:

$$\varphi_1(B^3)\emptyset_1(B)(1-B)^0(1-B^3)^0 Z_t = \theta_0 + \theta_1(B)\vartheta_1(B^3)a_t$$

$$\varphi_1(B^3)\emptyset_1(B)Z_t = \theta_1(B)\vartheta_1(B^3)a_t$$

$$Z_t = \emptyset_1 Z_{t-1} + \varphi_1 Z_{t-3} - \varphi_1\emptyset_1 Z_{t-4} + a_t - \theta_1 a_{t-1} - \vartheta_1 a_{t-3} + \vartheta_1\theta_1 a_{t-4}$$

The next stage is to prewhitening the input series of harvested area based on the ARIMA model identification of the input series. So we get a prewhitening input series of harvested area (Xt) is:

$$a_t = \frac{\varphi_1(B^3)\emptyset_1(B)}{\theta_1(B)\vartheta_1(B^3)} X_t$$

While the prewhitening output series of production output (Yt) is

$$\beta_t = \frac{\varphi_1(B^3)\emptyset_1(B)}{\theta_1(B)\vartheta_1(B^3)} Y_t$$

Cross-correlation carried out between input and output series that has been through the process prewhitening. This cross-correlation shows the relationship between the harvested area and rice production. In addition, the cross-correlation pattern will be used to identify the initial model transfer function by determining the order of b, s, and r. Order b symbolizes the period before harvesting area series start to affect the rice production series, the value of b is determined based on a real lag for the first time in the cross-correlation plot, it's b=1. s shows how long input series affecting output series after the first real, s values obtained from subsequent lag that form a clear pattern. Based on cross-correlation plot there is no clear pattern after the first lag so identified s = 0. Orde r indicates how long production series associated with the production value of the series itself. Based on the pattern of the cross-correlation, is assumed to form a pattern of a sinus/cosinus wave so it's identified r = 2. Based on the cross correlation and overfitting that has been done, the obtained tentative model of the transfer function is b = 1, s = 0, r = 2. So the initial model for the transfer function is:

$$Y_t = \mu + \frac{\omega_0(B)}{\delta_2(B)} X_{t-1} + \eta_t$$

$$Y_t = \mu + \frac{\omega_0(B)}{1 - \delta_1 B - \delta_2 B^2} X_{t-1} + \eta_t$$

$$Y_t = 1172943 - \frac{0{,}53407}{1 + 0{,}81365B + 0{,}92865B^2} X_{t-1} + \eta_t$$

Identification of residual models (η_t):

$$\eta_t = Y_t - 1172943 + \frac{0{,}53407}{1 + 0{,}81365B + 0{,}92865B^2} X_{t-1}$$

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	2887461.2	707017.7	4.08	<.0001	0	y	0
AR1,1	0.79913	0.10034	7.96	<.0001	1	y	0
NUM1	-4.51534	0.60718	-7.44	<.0001	0	x	1
DEN1,1	0.66789	0.14104	4.74	<.0001	1	x	1
DEN1,2	-0.32602	0.10801	-3.02	0.0025	2	x	1

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	--------------------Autocorrelations--------------------					
6	7.04	5	0.2180	0.131	-0.134	-0.003	-0.137	-0.064	0.161
12	16.78	11	0.1145	-0.048	0.091	0.214	-0.049	0.022	0.216
18	22.15	17	0.1791	-0.049	0.022	0.161	-0.095	-0.111	-0.059
24	28.57	23	0.1949	0.010	0.035	0.233	0.020	-0.022	0.047

Crosscorrelation Check of Residuals with Input x

To Lag	Chi-Square	DF	Pr > ChiSq	--------------------Crosscorrelations--------------------					
5	5.32	4	0.2556	0.180	-0.042	-0.144	-0.114	0.006	0.038
11	8.06	10	0.6234	-0.103	0.049	0.079	-0.022	-0.016	0.124
17	10.40	16	0.8450	-0.008	-0.001	0.042	0.008	-0.020	-0.168
23	12.72	22	0.9407	0.089	-0.019	0.100	0.098	0.036	0.030

Figure 2. Output of the Final estimate transfer function and assumptions check of residual

Identification final transfer function model is by combining the previous model with the residual model. The final transfer function model with b=1, s=0, r=2, and residual model with AR(1) :

$$Y_t = \mu + \frac{\omega_0(B)}{\delta_2(B)} X_{t-1} + \frac{a_t}{\emptyset_1(B)}$$

$$Y_t = 2887461{,}2 - \frac{4{,}51534}{1 - 0{,}66789B + 0{,}32602B^2} X_{t-1} + \frac{a_t}{1 - 0{,}79913B}$$

Final transfer function model indicate that production at t-th subround is affected by one, two and three previous production. And also affected by harvested area at one and two previous subround.

The transfer function model also had to meet the assumptions of nonautocorrelation residual, and there is no correlation between residual and input series of harvested area. Beside that, all parameter estimates in this model are significant. Based on the autocorrelation check of residual at figure 2 showed that all lag has p-value> 5%, which means that there is no autocorrelation in the residual value. So, this model can be used to forecasting the rice production.

Forecasting result can be use to measure the validation of model by comparing forecasting value with the real data. Accurate of model can be know from MAPE (Mean absolute persentage error). The next table give information about that. We have MAPE about 15,4 %, less than 20 %. So, we can conclude that transfer function model is fairly accurate in forecasting rice production in South Sulawesi Province.

Year	Subround (t)	Y	$\hat{Y}$
2008	I	1476126	1087817,7
2008	II	1664297	1290159,4
2008	III	942933	1087926,5
2009	I	1772281	1226042,3
2009	II	1522133	1328263,2
2009	III	1029764	1062127,9
2010	I	1404862	1190752,6
2010	II	1725965	1310821,6
2010	III	1251616	1064564,5
2011	I	1754238	1199447,1
2011	II	1568631	1319146,6
2011	III	1188836	1074308,9
MAPE		0,154	

4 Conclusion

According the the result, we can conclude:

1. Transfer function model is a multiple time series forecasting model which combines the characteristics of ARIMA model with characteristics of regression models .
2. Final transfer function model indicate that production at t^{th} subround is affected by one, two and three previous production. And also affected by harvested area at one and two previous subround..
3. The transfer function model met the assumptions of nonautocorrelation residual. So, we can conclude that transfer function model is fairly accurate in forecasting rice production in South Sulawesi Province.

References

[1] Arifin B., 2007, “Diagnosis Ekonomi Politik Pangan dan Pertanian”. Jakarta(ID): Raja Grafindo Persada.

[2] Box GEP, Jenkins GM, Reinsel GC., 1994, “Time Series Analysis Forecasting and Control”, third eddition, New Jersey(US): Prentice Hall.

[3] Herniwati, Kadir S., 2009, “Potensi Iklim, Sumber Daya Lahan, dan Pola Tanam di Sulawesi Selatan”, Prosiding Seminar Nasional Serealia, Balai Pengkajian Teknologi Pertanian Sulawesi Selatan.

[4] Montgomery DC, Jennings CL, Kulahci M., 2008, “Introduction to Time Series Analysis and Forecasting”, Canada(CA): Wiley Interscience.

[5] Suryana A., 2008, “Menelisik Ketahanan Pangan, Kebijakan Pangan, dan Swasembada beras”, Bogor(ID): Pusat Analisis Sosial Ekonomi dan Kebijakan Pertanian Bogor.

[6] Wei WWS., 2006, “Time Series Analysis Univariate and Multivariate Methods”, Second Edition, Canada (CA): Addison-Wesley.