

Forecasting Demand of Agricultural Tractor, Riding Type Rice Transplanter and Combine Harvester by using an ARIMA Model

Byounggap Kim*, Seung-Yeoub Shin, Yu Yong Kim, Sunghyun Yum, Jinoh Kim

National Academy of Agricultural Science, Rural Development Administration, Suwon, Korea

Received: January 29th, 2013; Revised: February 27th, 2013; Accepted: February 28th, 2013

Abstract

Purpose: The goal of this study was to develop a methodology for the demand forecast of tractor, riding type rice transplanter and combine harvester using an ARIMA (autoregressive integrated moving average) model, one of time series analysis methods, and to forecast their demands from 2012 to 2021 in South Korea. **Methods:** To forecast the demands of three kinds of machines, ARIMA models were constructed by following three stages; identification, estimation and diagnose. Time series used were supply and stock of each machine and the analysis tool was SAS 9.2 for Windows XP. **Results:** Six final models, supply based ones and stock based ones for each machine, were constructed from 32 tentative models identified by examining the ACF (autocorrelation function) plots and the PACF (partial autocorrelation function) plots. All demand series forecasted by the final models showed increasing trends and fluctuations with two-year period. **Conclusions:** Some forecast results of this study are not applicable immediately due to periodic fluctuation and large variation. However, it can be advanced by incorporating treatment of outliers or combining with another forecast methods.

Keywords: Demand forecast, Tractor, Riding type rice transplanter, Combine harvester, an ARIMA model

Introduction

Demand forecast of agricultural machinery is important to manufacturers and the government. Manufacturers should set up their production plans based on the sales expectation. The government also needs to forecast demand of agricultural machinery, since it has a task of agricultural mechanization and is required to secure and execute the budget for it.

There have been many researches on the forecasting agricultural machinery demand by using various methodologies. Soto et al. (1989) developed a methodology to estimate the evolution of the farm tractor stock in Chile based on annual import reports and estimation of the appropriate normal distribution describing tractor replacement. Biondi et al. (1998) used both univariate and multivariate ARIMA models to forecast farm tractor

demand in Italy, France and the United States. Kang (1999) estimated demands of six kinds of agricultural machinery in Korea by adding new demand and replace demand. New demand was estimated from the difference of each year's stock which was forecasted by the logistic function. Replace demand was considered as the number of machines supplied durable years ago. Pawlak (1999) used single- and multiple-variable regression methods to survey the demand for agricultural tractors on the Polish market. Yun (2004) estimated mid and long term demands of a tractor, a riding type rice transplanter, a combine harvester and a grain dryer in Korea by using logistic curve function and Markov chain model. Kim et al. (2006) developed the demand functions of tractor, combine harvester and riding type rice transplanter in Korea by using an autoregressive time-series model, a multi-variable linear regression technique and SUR (seemingly unrelated regression) model. Kim et al. (2010) estimated the 2010's demand of seven kinds of agricultural machinery in Korea by surveying 233 retailers. Kim et al. (2012) estimated

*Corresponding author: Byounggap Kim

Tel: +82-31-290-1933; Fax: +82-31-290-1960

E-mail: kimbg@korea.kr

tractor's demand in Korea using two different time series models, an exponential smoothing model and an ARIMA model.

Nowadays, a time series analysis is widely used in many areas such as statistics, econometrics, mathematical finance, weather forecast and so on, since it provides various analysis techniques appropriate for the data characteristics and it reflects many factors simultaneously which can affect to the forecast. Recently, this analysis method has been used to forecast the demand of agricultural machinery.

The objective of this study was to develop a methodology for forecasting demands of tractor, riding type rice transplanter and combine harvester by using an ARIMA model and to forecast their demands from 2012 to 2021 in South Korea.

Materials and Methods

Construction of an ARIMA model

An autoregressive integrated moving average (ARIMA) model analyzes and forecasts equally spaced univariate time series data. It uses only current and past observations of the data, so it does not need exogenous variables as in another forecast models such as econometric ones or regression ones.

As its name implies, an ARIMA model is an integrated combination model of autoregressive (AR) model and moving average (MA) model. Integration means that it transforms the non-stationary data into stationary by differencing or logging. A mathematical model is represented as below (Wei, 1994).

If Y_t is stationary time series data, which means their average and variance are constant for all t and autocorrelation of two observations only depends on the time lag, the AR model is defined as

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (1)$$

where:

- ϕ_i = the parameters of the AR model ($i=1,2,\dots,p$)
- a_t = a well-behaved error term with zero mean and constant variance that is uncorrelated with

The MA model is defined as

$$Y_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2)$$

where:

- θ_i = the parameters of the MA model ($i=1,2,\dots,q$)

An ARMA model is a combination of AR model and MA model. It has the general form as follow.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3)$$

In case that Y_t is non-stationary, it can be made stationary by the use of successive differencing of a suitable degree or logging. If $\nabla^d Y_t$ is a stationary data series by differencing of degree d , an ARIMA model is represented as

$$\nabla^d Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (4)$$

Generally, the order of an ARIMA model is usually denoted by the notation ARIMA (p,d,q), where p is the order of the autoregressive part, d is the order of the differencing and q is the order of the moving-average process.

Construction of ARIMA model generally follows three stages, which are identification, estimation and diagnose.

In the first stage, tentative ARIMA models were identified. The order of differencing, d , is determined by examining the stationarity of the series. The standard way to check the stationarity is to plot the time series and the sample autocorrelation function (ACF). If the series is stationary, the time series has no trend, same variance in all time, and ACF plot tails off exponentially or cuts off after some lags (Wei, 1994; Lee, 2007). If the series is not stationary, the same examination for the plots of differenced series must be repeated until it becomes stationary. The other two orders, p and q , are determined by comparing patterns of ACF plots and sample partial autocorrelation function (PACF) plots with the theoretical characteristics given in the Table 1.

In the second stage, parameters of the tentative models were estimated from conditional least squares method. Some models which did not converge or of which all parameter's significance levels were under 95% were discarded.

Table 1. Characteristics of theoretical ACF and PACF for stationary processes

Process	ACF plot	PACF plot
AR (p)	Tails off as exponential decay or damped sine wave	Cuts off after lag p
MA (q)	Cuts off after lag q	Tails off as exponential decay or damped sine wave
ARMA (p,q)	Tails off after lag q	Tails off after lag p

Source: Park and Kim (2002):p97-98

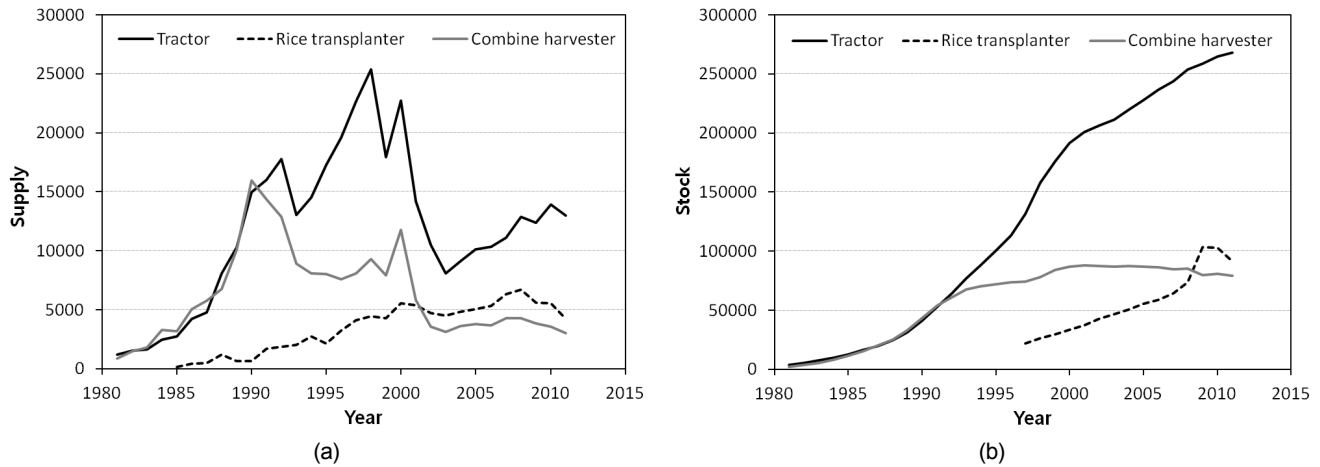


Figure 1. The supply series (a) and the stock series (b) for the selected agricultural machinery.

In the final stage, the adequacy of the models was diagnosed and the final models were confirmed. There are two diagnostic checking methods; analysis of the residual series and the portmanteau test (Wei, 1994). In this study, the latter was used since it was more widely used. After the diagnosis, the final models were confirmed by comparing Akaike's information criterion (AIC), Schwartz's Bayesian criterion (SBC) and the significance of parameter estimation. AIC and SBC are model selection criteria based on residuals. Smaller AIC and SBC statistics indicate the better fitting model (Wei, 1994). After this modeling sequence, the final model was used to forecast.

Time series data and analysis tool

Time series used for forecasting the demand were supply series and stock series of tractor, riding type rice transplanter and combine harvester. Annual supply can be considered as the annual demand. In the case of stock, the differences of two consecutive years can be used as a means of demand forecast. In some previous studies, the stock was used to forecast a demand (Soto et al. (1989); Kang (1999)). When a stock series was used to forecast, the annual demand was calculated by adding the number of machines supplied durable years ago with the stock

difference between that year and the previous year. It can be represented as equation 5. The survey results about the used years of agricultural machinery (Kim et al., 2011) were used as durable years of tractor, riding type rice transplanter and combine harvester for the calculation, which were 8.9, 5.5 and 6.5 years, respectively.

$$S_t = (X_t - X_{t-1}) + E_t \quad (5)$$

where:

S_t = annual supply at the year t ,

X_t = stock at the year t ,

E_t = the number of machines supplied durable years ago from the year t

The sources of the supply series and the stock series were annual data of tractor, riding type rice transplanter and combine harvester provided by the Korean government from 1981 to 2011 (MIFAFF, 2011). However, the supply and stock data for the riding type rice transplanter were not provided before 1985 and 1997, respectively. Hence, the analysis was done based on the provided data. The time plots for the six series are shown in the Figure 1.

The analysis tool was SAS 9.2 for Windows XP (SAS Institute Inc.) and the used procedure was an “arima”.

Results and Discussion

Construction of an ARIMA model

Identification (Stage 1)

In the identification stage, the stationarities of the series were examined first. The time plots of the series (Fig. 1) showed that all of them looked like having a linear or quadratic trend. It meant that all series were not stationary. The ACF plots for the series in Figure 2 also showed that all series looked like not stationary since all ACF plots decreased slowly and linearly.

Since all series were not stationary, stationarities of their differenced series were examined. Plots of differenced series and their ACF plots are shown in the Figure 3 and

Figure 4. Time plots of the six differenced series in Figure 3 showed that all of them could be judged to have trend and their variations looked like different by some time lag. Hence, all series needed to be more differenced. However, their ACF plots in Figure 4 cut off or tailed off which implied that single differenced series could be considered as a tentative model.

The plots of double differenced series and their ACF plots in Figure 5 and Figure 6 showed similar characteristics with those of single differenced series. Since over differencing is not recommended (Wei, 1994), the degree of differencing of tentative ARIMA models, d , was determined as one and two.

After the degree of difference, d , was determined, the other two orders, p and q , were determined by inspecting the ACF and PACF plots. PACF plots of single and double differenced six series were shown in Figure 7 and Figure 8.

The determination of orders of ARMA model is not easy

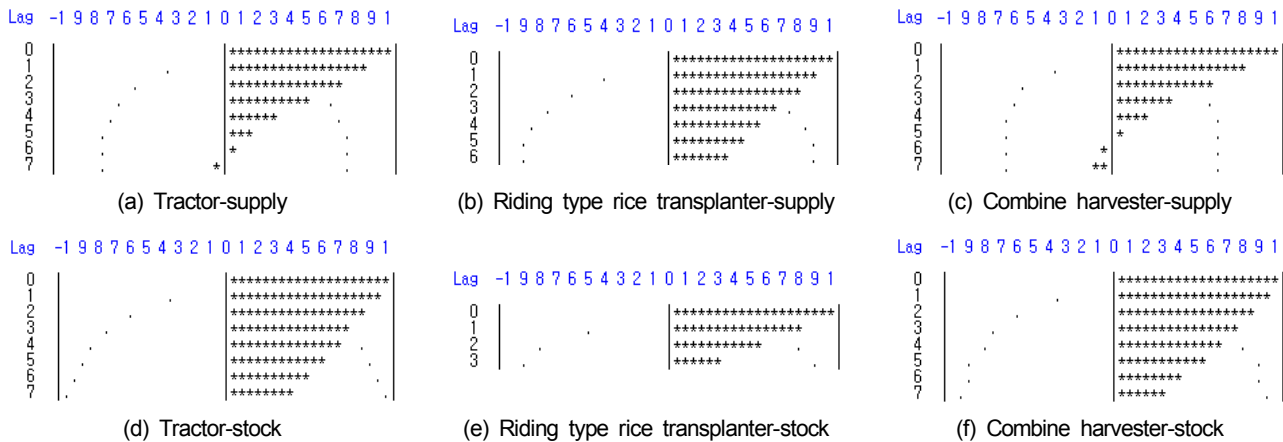


Figure 2. ACF plots of the supply series and the stock series for the selected agricultural machinery.

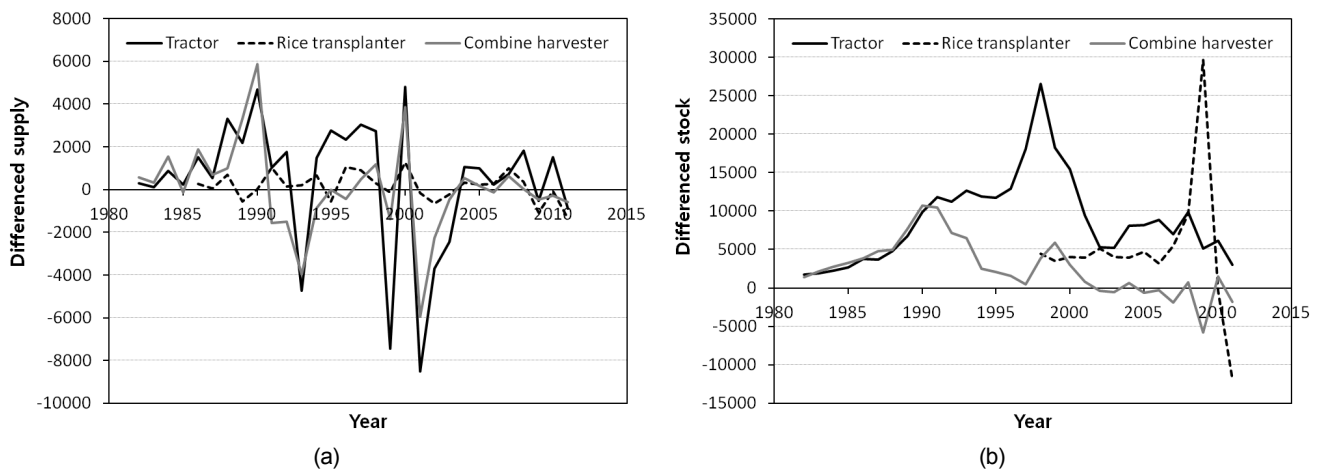


Figure 3. The single differenced supply series (a) and the single differenced stock series (b) for the selected agricultural machinery.

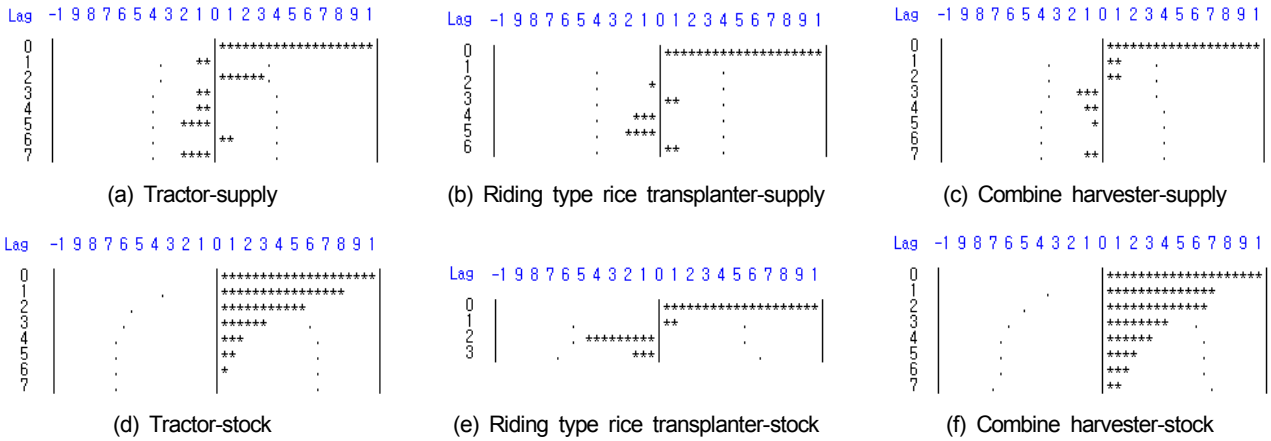


Figure 4. ACF plots of the single differenced supply and the single differenced stock series for the selected agricultural machinery.

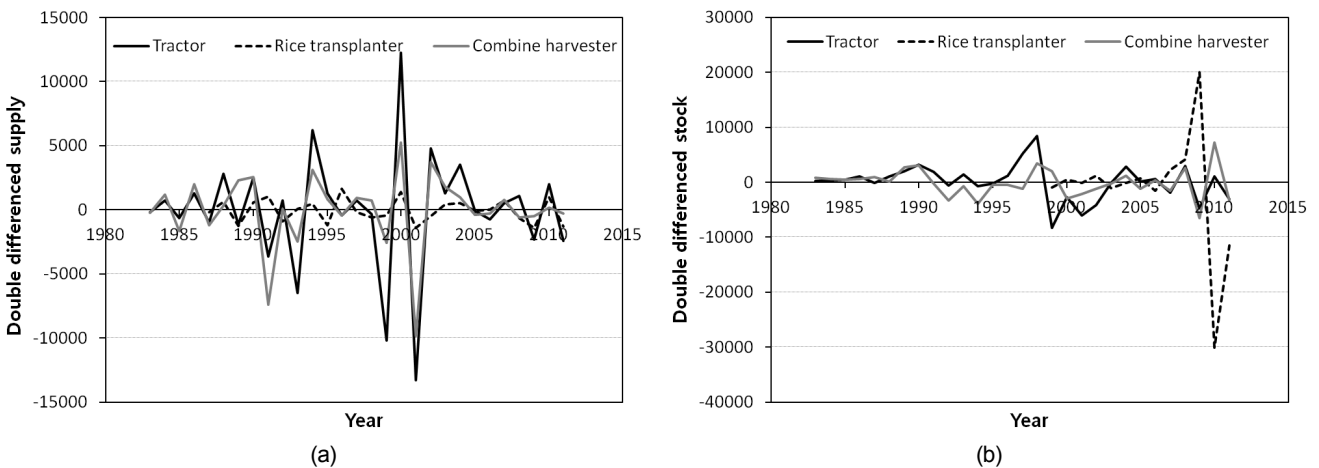


Figure 5. The double differenced supply series (a) and the double differenced stock series (b) for the selected agricultural machinery.

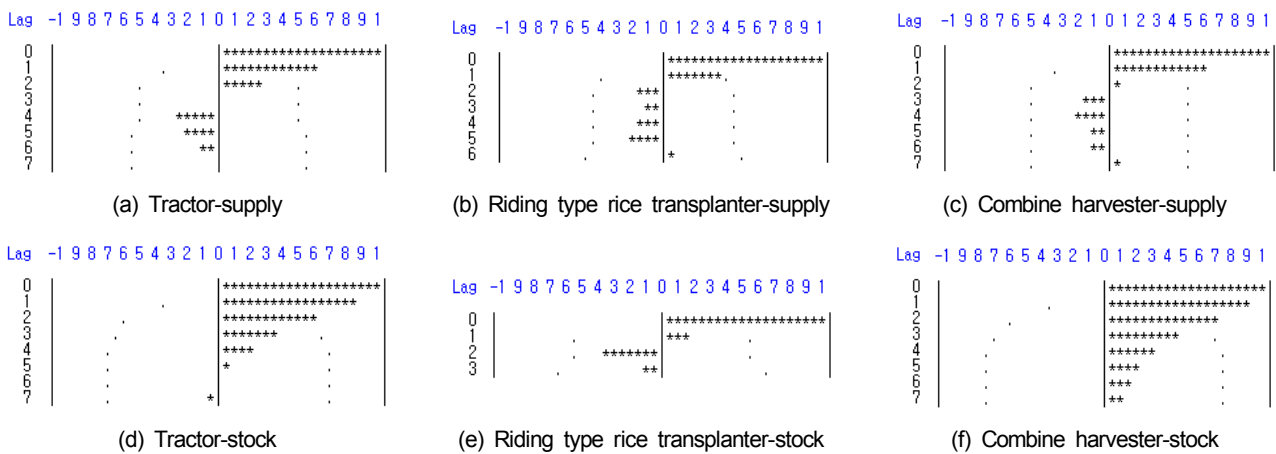


Figure 6. ACF plots of the double differenced supply series and the double differenced stock series for the selected agricultural machinery.

since the characteristics of ACF and PACF plots of real series are different from theoretical ones suggested in the Table 1. Hence, possible all model process, AR, MA and

ARMA, were examined and the results were summarized in the Table 2, resulting in total 32 tentative models.

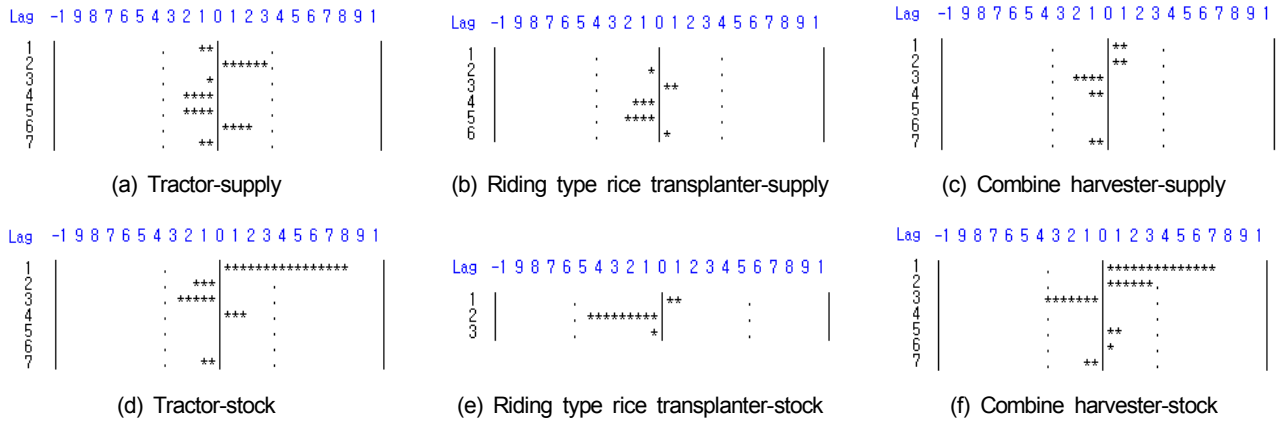


Figure 7. PACF plots of the single differenced supply series and the single differenced stock series for the selected agricultural machinery.

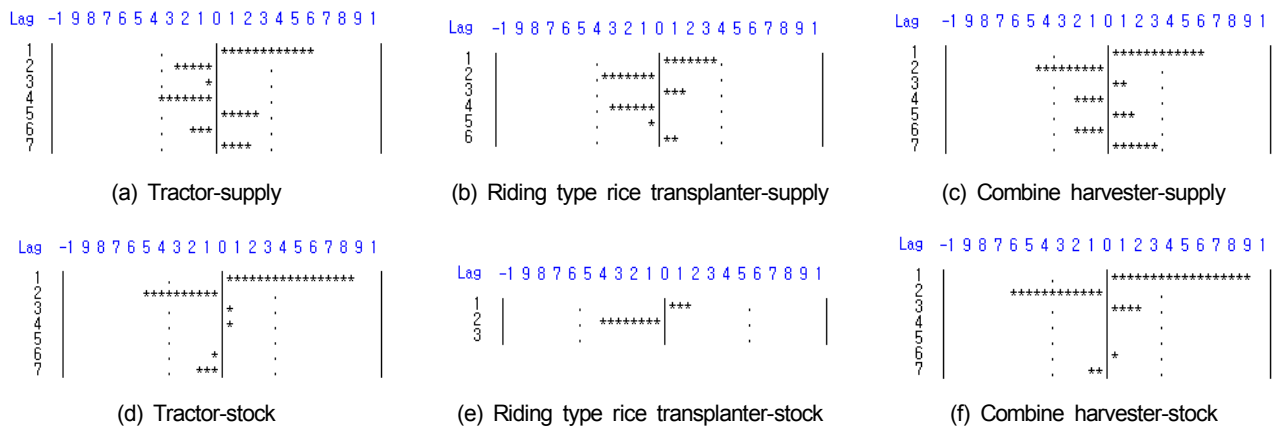


Figure 8. PACF plots of the double differenced supply series and the double differenced stock series for the selected agricultural machinery.

Table 2. Tentative models for the six series determined in the identification stage

Degree of difference	Process	The orders of p and q in the tentative model					
		Tractor supply	Riding type rice trans-planter supply	Combine harvester supply	Tractor stock	Riding type rice trans-planter stock	Combine harvester stock
Single ($d=1$)	AR	2, 0	-*	-	1, 0	2, 0	3, 0
	MA	0, 2	-	-	0, 2	0, 2	0, 2
	ARMA	1, 1	4, 4	2, 2	2, 1	1, 1	3, 2
Double ($d=2$)	AR	4, 0	1, 0	2, 0	2, 0	2, 0	2, 0
	MA	0, 1	0, 1	0, 1	0, 2	0, 2	0, 2
	ARMA	3, 1	1, 1	1, 1	2, 1	1, 1	1, 1

Note: The symbol ‘-’ denotes that the process was not applicable to the series.

Estimation (Stage 2)

Results of each parameter estimation of 32 models in Table 2 conducted by using SAS program were shown in the Table 3. Among the 32 tentative models, 16 models did not converge. The ratio of not converged models to the total models was relatively high in the riding type rice transplanter stock, which was 5 of 6. It might be caused by

the small number of the series as shown in the Figure 1 (b). The number of models which converged but had low significance levels was two; single differenced tractor supply AR model and single differenced tractor stock ARMA model. Hence, 14 tentative models were treated in the next stage.

Table 3. Estimation results of model parameters

Degree of difference	Process	The orders of p and q in the tentative model, ARIMA (p, d, q)		
		Tractor supply	Riding type rice-transplanter supply	Combine harvester supply
Single ($d=1$)	AR	$\phi_1=-0.0677$ (-0.37) $\phi_2=0.31355$ (1.71*)	-	-
	MA	$\theta_1=0.03869$ (0.22) $\theta_2=-0.45192$ (-2.61**)	-	-
	ARMA	Not converged	Not converged	Not converged
Double ($d=2$)	AR	$\phi_1=0.74742$ (3.88***) $\phi_2=-0.27069$ (-1.12) $\phi_3=0.21253$ (0.88) $\phi_4=-0.33275$ (-1.72*)	$\phi_1=-0.45192$ (2.15**)	$\phi_1=0.83865$ (4.77***) $\phi_2=-0.44419$ (-2.52**)
	MA	$\theta_1=-0.94145$ (-13.19***)	Not converged	Not converged
	ARMA	Not converged	Not converged	Not converged

Degree of difference	Process	The orders of p and q in the tentative model, ARIMA (p, d, q)		
		Tractor stock	Riding type rice-transplanter stock	Combine harvester stock
Single ($d=1$)	AR	$\phi_1=0.92224$ (12.1***)	$\phi_1=0.07529$ (0.25) $\phi_2=-0.71404$ (-2.29**)	$\phi_1=0.59231$ (3.16***) $\phi_2=0.54359$ (2.31**) $\phi_3=-0.39863$ (-1.71*)
	MA	$\theta_1=-0.84285$ (-5.39***) $\theta_2=-0.60294$ (-3.84***)	Not converged	$\theta_1=-0.40560$ (-5.27***) $\theta_2=-0.97622$ (-13.89***)
	ARMA	$\phi_1=1.28192$ (1.15) $\phi_2=-0.36402$ (-0.36) $\theta_1=0.23034$ (0.20)	Not converged	Not converged
Double ($d=2$)	AR	$\phi_1=1.56005$ (10.09***) $\phi_2=-0.62369$ (-4.03***)	Not converged	$\phi_1=1.45742$ (8.88***) $\phi_2=-0.60846$ (-3.60***)
	MA	$\theta_1=-1.29465$ (-8.55***) $\theta_2=-0.70387$ (-4.66***)	Not converged	Not converged
	ARMA	Not converged	Not converged	$\phi_1=0.85142$ (7.90***) $\theta_1=-0.72299$ (-4.44***)

Note: The symbols *, ** and *** denotes that the t value is 90, 95 and 99% significant, respectively.

Diagnosis (Stage 3)

Results of diagnose for the 14 tentative models were shown in Table 4. Final models for each series were determined by comparing AIC and SBC values and checking the Portmanteau test result. They are as follows; ARIMA (0,2,1), ARIMA (1,2,0), ARIMA (2,2,0), ARIMA (2,2,0), ARIMA (2,1,0) and ARIMA (1,2,1) for tractor supply, riding type rice transplanter supply, combine harvester supply, tractor stock, riding type rice transplanter stock and combine harvester stock, respectively. Mathematical representations for each model are shown in the equation (6)~(11).

$$\nabla^2 Y_t = a_t + 0.94145a_{t-1} \quad (6)$$

$$\nabla^2 Y_t = -0.45192Y_{t-1} + a_t \quad (7)$$

$$\nabla^2 Y_t = 0.83865Y_{t-1} - 0.44419Y_{t-2} + a_t \quad (8)$$

$$\nabla^2 Y_t = 1.56005Y_{t-1} - 0.62369Y_{t-2} + a_t \quad (9)$$

$$\nabla Y_t = 0.07529Y_{t-1} - 0.71404Y_{t-2} + a_t \quad (10)$$

$$\nabla^2 Y_t = 0.85142Y_{t-1} + a_t + 0.72299a_{t-1} \quad (11)$$

Demand forecast of tractor, riding type rice transplanter and combine harvester

Using the six models determined above, the demand of

Table 4. The diagnose results for the 14 tentative models selected in the previous stages

Series	Model	AIC	SBC	Portmanteau test result
Tractor supply	ARIMA (0,1,2)	567.1659	571.3695	Fit
	ARIMA (4,2,0)	555.9694	562.8059	Fit
	ARIMA (0,2,1)	549.4637	552.1983	Fit
Riding type Rice- transplanter supply	ARIMA (1,2,0)	407.9768	410.4146	Fit
Combine harvester supply	ARIMA (2,2,0)	537.8047	541.9066	Fit
Tractor stock	ARIMA (1,1,0)	572.3191	575.1215	Fit
	ARIMA (0,1,2)	578.5265	582.7301	Fit
	ARIMA (2,2,0)	565.0356	569.1375	Fit
	ARIMA (0,2,2)	581.8918	585.9937	Not fit
Riding type Rice- transplanter stock	ARIMA (2,1,0)	292.6714	294.5886	Fit
Combine harvester stock	ARIMA (3,1,0)	554.6312	560.236	Fit
	ARIMA (0,1,2)	552.7866	556.9902	Fit
	ARIMA (2,2,0)	537.4688	541.5707	Fit
	ARIMA (1,2,1)	535.5202	539.6221	Fit

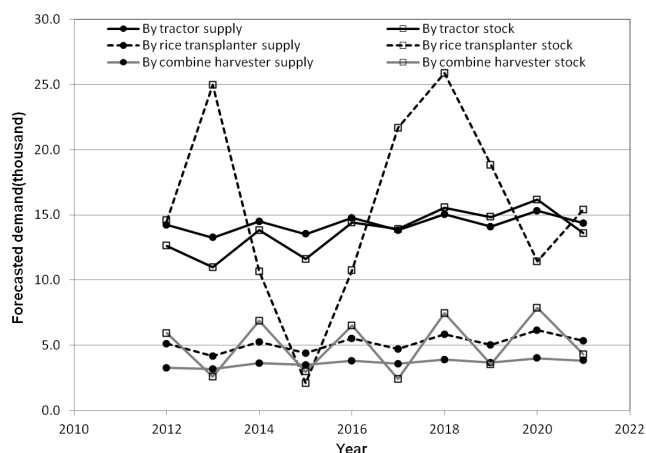


Figure 9. Forecast results of tractor, riding type rice transplanter and combine harvester demand.

tractor, riding type rice transplanter and combine harvester were forecasted and the results are shown in the Figure 9.

The forecasted tractor demand by using supply series was 14,217 in 2012 to 14,361 in 2021, that of riding type rice transplanter 5,104 to 5,330, and that of combine harvester 3,261 to 3,834. The three demand series forecasted by using the supply series commonly showed slow increase and a fluctuation with two-year period. It could be caused by the order of the ARIMA model.

The forecasted demand of tractor by using stock series was 12,632 in 2012 and 13,606 in 2021, and that of combine harvester was 5,948 in 2012 and 4,323 in 2021. The differences between forecasted value by using supply series and stock series were relatively small. The two demand series also showed slow increase and a fluctuation with two-year period which were similar with

the cases of forecast by supply series. The range of fluctuation, however, was wider than that by using supply series.

The riding type rice transplanter demand by using stock series was 14,576 in 2012 and 15,418 in 2021, much larger than that by using supply series. Furthermore, the range of fluctuation was so wide that the ratio of difference to the forecasted demand was over 100%. Such results were due to a sharp change of stock and the number of stock series. The stock of riding type rice transplanter in 2009 was 103,412 which was drastically jumped from 73,735 in 2008. The number of riding type rice transplanter supplied by government in 2008 was only 6,691, much smaller than the difference of stocks 29,667. Generally the recommended number in the application of ARIMA model is over 30. However, since the number of used stock series was 15, it might cause the problem worse.

Conclusions

In this study, an ARIMA model for forecasting the demands of tractor, riding type rice transplanter and combine harvester were constructed and the demands of three machines from 2012 to 2021 was forecasted. Two kinds of series, supply series and stock series of each machine, were used for the forecast. The results and significances of this study can be summarized as follows.

- (1) The stationarity of six series, the supply series and the stock series of tractor, riding type rice transplanter, were examined by checking its time plots

and ACF plots and the degree of difference was determined one and two for all series.

- (2) ARIMA models were constructed by following the three stage; identification, estimation and diagnosis. In the identification stage, 32 tentative models were identified based on the ACF and PACF plots of single and double differenced supply and stock series for three machines. In the estimation stage, the parameters of each model were estimated and the tentative models were decreased to 14 by checking their convergence and significance. Finally, diagnosing the 14 models, the final models were determined.
- (3) The forecasted demands of tractor, riding type rice transplanter and combine harvester by using supply series were 14,217 in 2012 to 14,361 in 2021, 5,104 to 5,330 and 3,261 to 3,834, respectively. Those by stock series were 12,632 to 13,606, 14,576 to 15,418 and 5,948 to 4,323, respectively.

The demand forecast results showed fluctuations with two-year period or large variation. This is caused by the policy change of agricultural machinery supply, presence of outlier, and insufficiency of data. It is expected to eliminate fluctuation and to reduce variation by using more advanced ARIMA models, which are incorporating treatment of outliers or combination with another forecast method such as a regression.

Conflict of Interest

The authors have no conflicting financial or other interests.

Acknowledgement

This study was carried out with the support of "Research Program for Agricultural Science & Technology Development (Project No. PJ0066672013)" of National Academy of Agricultural Science, Rural Development Administration, Republic of Korea.

References

- Biondi, P., D. Monarca and A. Panaro. 1998. Simple forecasting models for farm tractor demand in Italy. France and the United States. *Journal of Agricultural Engineering Research*. 71:25-35.
- Kang, J. I. 1999. Problem and direction of agricultural mechanization policy. Seoul: Korea Rural Economic Institute (In Korean).
- Kim, B., S. Y. Shin, Y. Y. Kim, H. K. Kim, Y. H. Cho and K. H. Cho. 2010. The demand trend of agricultural machinery in 2010. Suwon: Rural Development Administration (In Korean).
- Kim, B., S. Y. Shin, Y. Y. Kim, H. Y. Kim, Y. H. Cho and Y. B. Lee. 2011. Utilization of agricultural machinery and mechanized ratio. Suwon: Rural Development Administration (In Korean).
- Kim, B., S. Y. Shin, Y. Y. Kim and H. Y. Kim. 2012. Forecast of the tractor demand by using time series analysis. In; *Proceedings of the 6th International Symposium on Machinery and Mechatronics for Agriculture and Biosystems Engineering (ISMAB)*, Jeonju: KSAM-CIAM-JSAM.
- Kim, K., K. U. Kim, B. Kim and C. K. Park. 2006. Development of a demand forecasting system for major agricultural machinery. Suwon: Korean Society for Agricultural Machinery (In Korean).
- Lee, J. H. 2007. Time series analysis and application. Paju: Freedom Academy Pub. Co. (In Korean)
- MIFAFF. 2011. Working reference on agricultural machinery. Governmental Publication Registration No. 11-1541000-000338-14. Gwacheon: Ministry for Food, Agriculture, Forestry and Fisheries (In Korean).
- Park, Y. S. and K. H. Kim. 2002. Time series analysis using SAS/ETS. Seoul. Republic of Korea: Freedom Academy Pub. Co. (In Korean).
- Pawlak, J. 1999. Impact of some selected factors on the sale of agricultural tractors. *Problemy Inzynierii Rolniczej* 7(4):89-94.
- Soto, R., E. Saavedra and G. Edwards. 1989. An estimate of the demand for tractors in Chile, 1961-86. *Ciencia e Investigacion Agraria* 6(1-2):99-105.
- Yun, Y. D. 2004. Study on demand estimation of agricultural machinery by using logistic curve function and Markov chain model. *Journal of Biosystems Eng.* 29(5):331-450. (In Korean, with English abstract).
- Wei, W. W. S. 1994. Time series analysis. USA: Addison-Wesley Publishing Company.