

## FORECASTING OF PEACH AREA AND PRODUCTION WISE ECONOMETRIC ANALYSIS

A. Ullah<sup>1</sup>, D. Khan<sup>2</sup> and S. Zheng<sup>1,\*</sup>

<sup>1</sup>College of Economics and Management, Northwest A&F University, 3 Taicheng Road, Yangling, Shaanxi 712100, China

<sup>2</sup>Department of Economics, Kohat University of Science & Technology, Kohat 26000, Pakistan

\*Correspondence: zsfeng@nwfau.edu.cn

### ABSTRACT

The study was aimed at forecasting peach area and production in Pakistan using time series data for the period 1997-98 to 2014-15. Peach area and production were forecasted over a period of time 2015-2016 to 2025-26. The Box-Jenkins (1976) approach was applied to forecast area and production of peach. This study found ARIMA (1,1,0) as an appropriate model to forecast both area and production of peach. The best-forecasted model was determined based on the lowest values of Akaike information criterion, Bayesian information criterion and Hannan-Quinn criterion. However, the predictability power, performance and quality of the model was measured based on the lowest error value of the root mean square error, mean absolute error and mean absolute prediction error. The forecasted value of area and production of peach for the year 2025-26 were worked out as 11.05 thousand hectares and 65.05 thousand tonnes respectively. The minimum projection trend indicated declining area and production of peach in Pakistan. The author (s) may arrived at the conclusion that peach production can be increased by using improved peach cultivars, improved irrigation system and adequate cultural practices during the course of study in Pakistan. The strategy may be prioritized to bring more barren land under peach cultivation. Expansion of the area under peach would directly increase its production in the future.

**Key words:** ARIMA, Box- Jenkins, Forecasting, Peach.

### INTRODUCTION

Agriculture plays a major role in Pakistan's economy. This sector has traditionally sustained a satisfactory growth to ensure food security for the growing population. This sector contributes 19.8 percent to the gross domestic product (GDP) and employs 42.3 percent of country's labour force (GOP, 2016). Within agriculture, the horticulture is an important sub-sector of the agricultural economy. Pakistan horticulture sector produces approximately 12 million tonnes per year production of fruits, vegetables and spices. High value and great potential fruits are grown in different varieties and delicious in taste. These are apples, mangoes, peaches, grapes, citrus (kinnow), dates, and cherries. Other prominent fruits that have huge export potential are loquat, pears, plums, and guava (PHDEC, 2017).

Peach (*Prunus persica*) is the second most important stone fruit in Pakistan. It belongs to the family Rosaceae. It is the most important, among the stone fruit and is temperate in nature. It is a remarkable fruit having different attributes i.e. sweetness, juiciness, fleshiness, attractive in flavour and aroma. Due to these attributes, it is very delicious in taste (Yu *et al.*, 2015). Fresh peach is comprised of very healthy nutrient. It has a rich source of vitamins A and C and also contains potassium and fibre. The fruit has over 80% water and one average sized peach has 7% of the dietary fibre which is required each day (Habib, 2015). In spite of its various uses, its

cultivation is declining day by day both in acreage and yield.

In Pakistan and especially in northern parts of the country, there is huge scope for promotion of peaches which can be aimed at international marketing by promoting processing industry for value addition and export. Therefore, it has a huge potential for the home market as well as for the exports (Habib, 2015). However, Pakistan has experienced ups and downs in peach production. Moreover, farmers do not know the future prospect of peach production and prices while deciding to cultivate this and other crops. Fruits contribute 2.48% value addition to agriculture (GOP, 2016). The total area under fruits cultivation was 764.26 thousand hectares for the year 2014-15 in Pakistan. Out of the total area under fruits, 5.7 thousand hectares is under peach. The total fruits production was 7018 thousand tonnes among them peach fruit contributes 66.4 thousand tonnes for the year 2014-15 in Pakistan (GOP, 2015). Factually other fruits in general and peach, in particular, provide linkages through which it can stimulate economic growth in other sectors. Peach cultivation has been suffering from various problems, such as diseases attacks, low yields, marketing of produce, the high cost of inputs, energy crises, lack of cold storages, lack of capital, drought causes, and environmental changes that decreased the area under peach orchards.

The measurement and forecasting of the volatility of agricultural commodity future is crucial for agricultural production, resource allocation, and risk

management (Tian *et al.*, 2017). Crop production prediction is important for advanced planning, formulation and implementation of policies related to food procurement, distribution and import-export decision. Long-term projections of agricultural produce and operational area help the policy makers regarding price fixation, procurement, marketing and storage. However, to fulfill this goal, there is a need to monitor trends of production in the country. Several attempts have been made in the past to develop yield forecast models for various commodities. Iqbal *et al.* (2005) studied the forecasting of wheat production in Pakistan. Rahman and Baten (2016) have also used the ARIMA models to forecast the area and production of black gram pulse in Bangladesh. Furthermore, Biswas *et al.* (2014) have studied the forecasting of wheat production in Punjab. Ahmad and Mustafa (2006) used the ARIMA model to forecast Kinnow production in Pakistan. However, Qureshi *et al.* (2014) have studied the production of the citrus fruit in Pakistan using ARIMA. Similarly, Khan *et al.* (2008) have used the ARIMA model to forecast mango production while Jam *et al.* (2013) used this model to forecast area of mango in Pakistan.

Forecasting is an essential part of the econometric analysis. Researchers, producers and policymakers are interested in examining the information regarding peach area under cultivation and production in Pakistan. A study is needed to provide recent information and estimation about the forecasting of peach area and production in Pakistan over the next few years. In econometrics, an autoregressive integrated moving

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Where  $Y_t$  is the original series at time  $t$ ,  $\varepsilon$  denotes error term,  $\varphi$  and  $\theta$  represent coefficients of the respective component of the model. Moreover,  $p$  is the component of autoregressive (AR) and  $q$  is the component of moving average (MA) process in the ARIMA model. The integrated component can be expressed with the term  $d$  shows the order of differencing of  $Y_t$ . The general form of the autoregressive integrated moving average model can be written as ARIMA ( $p, d, q$ ) (Anderson, 1971).

The development of an ARIMA model comprises three methodical steps. The first step is called identification of the model and it determines the stationarity and normality of the data. This option enables us to test for stationarity and tentative ARMA order identification. Generally, at this stage, the model is estimated. The analysis of the identify peach area and production output suggests one or more ARIMA models that can be a preferred or a best fit model. The second step of the ARIMA model is known as estimation and diagnostic checking. The adequacy of the best fit model is judged with the production of diagnostic statistics.

average (ARIMA) model developed by Box and Jenkin (1976) is a univariate time series model which is used for statistical forecasting of crop production. A number of researchers have used the Box- Jenkins methodology to forecast future demands regarding internal consumption and export to adopt appropriate measures. Thus, this study also uses the Box-Jenkins methodology to model and forecast peach area and production in Pakistan. The remainder of the paper proceeds as follows. The next section describes the materials and methods. Important empirical results and discussion are given in Section 3. Finally, Section 4 concludes the results.

## MATERIALS AND METHODS

**Data Source:** Two sets of time series data were used in this study. The series includes peach area and production for the periods 1997-98 to 2014-15. The data were taken from Federal Bureau of Statistics, Ministry of Food, Agriculture and Livestock (MINFAL). The data was processed and analysed by using the EViews software version 9. Finally, a forecast for the next 11 years (2015-2016 to 2025-26) was made.

**The ARIMA Model:** The ARIMA model, developed by Box and Jenkin (1976) was used for model building and forecasting of annual peach area and production. ARIMA is the most general class of model to forecast a time series data. The standard regressions of ARIMA model can be mathematically specified as below:

Finally, the forecasting of the peach area and production of the time series is performed.

## RESULTS AND DISCUSSION

**Descriptive Statistics:** Table 1. provides the descriptive statistics of peach area and production for Pakistan. The maximum area of peach was 15.8 thousand hectares in 2008-09 and was minimum 4.4 thousand hectares in 1997-98. The maximum production of peach in Pakistan was 83.7 thousand tonnes in the year 2008-09 and minimum were 32.9 thousand tonnes in the year 2000-01.

**Identification of ARIMA Model:** The first phase of the ARIMA modelling is called identification. Tables 2-5 display the autocorrelation function (ACF) and partial autocorrelation function (PACF) for peach area and production. Autocorrelation indicates the order of the AR component of the  $p$  while the partial autocorrelation indicates the order of MA component of the  $q$  parameter (Pindyck *et al.*, 1991). Table 3 shows that the autocorrelation dies off smoothly and partial

autocorrelation was diminished drastically after one lag. However, Table 5 shows that partial autocorrelation cuts off gradually and autocorrelation is tappers off rapidly. Hence, both the series seems to be appropriate. Thus, Tables 3 & 5 is suitable for further analysis. Moreover, the Ljung-Box Q-Statistic having corresponding probability value is a test statistic with the null hypothesis of no autocorrelation for a specified order of autocorrelation lags. The autocorrelations and partial

autocorrelations at all lags should be nearly zero if there is no serial correlation, and all Q-Statistics should be insignificant having large probability-values. Both the tables show that there is no autocorrelation as the autocorrelation and partial autocorrelation at all lags are zero, and the corresponding Q-stat is insignificant at first difference. Thus, both the variables (area and production) are stationary at first difference.

**Table 1. Descriptive Statistics of the peach area and production in Pakistan from 1997- 98 to 2014-15.**

Parameters	Mean	Median	Max.	Min.	St. Dev	Kurtosis	Skewness
Area (A)	11.90	14.55	15.8	4.4	4.65	-0.79	1.76
Production (Y)	59.54	58.25	83.7	32.9	16.06	-0.19	1.97

**Table 2. Correlogram for the original series of peach area.**

Autocorrelation	Partial correlation	AC	PAC	Q-Stat	P value	
.  *****	.  *****	1	0.855	0.855	15.483	0.000
.  *****	.  **	2	0.651	-0.297	25.026	0.000
.  ***	.  *	3	0.439	-0.121	29.642	0.000
.  *	.  *	4	0.213	-0.199	30.808	0.000
.  .	.  *	5	-0.017	-0.195	30.817	0.000
.  *	.  .	6	-0.184	0.050	31.828	0.000
.  **	.  **	7	-0.235	0.221	33.633	0.000
.  **	.  *	8	-0.271	-0.193	36.278	0.000
.  **	.  *	9	-0.305	-0.154	39.995	0.000
.  **	.  *	10	-0.314	-0.094	44.429	0.000
.  **	.  *	11	-0.310	-0.102	49.376	0.000
.  **	.  .	12	-0.297	0.061	54.676	0.000

**Table 3. Correlogram for the differenced series of peach area.**

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	P value	
.  ***	.  ***	1	0.387	0.387	3.0183	0.082
.  .	.  *	2	0.012	-0.161	3.0215	0.221
.  .	.  *	3	0.012	0.084	3.0250	0.388
.  *	.  *	4	-0.072	-0.131	3.1546	0.532
.  .	.  .	5	-0.065	0.026	3.2676	0.659
.  *	.  *	6	-0.114	-0.136	3.6506	0.724
.  *	.  .	7	-0.114	-0.012	4.0701	0.772
.  *	.  .	8	-0.059	-0.037	4.1951	0.839
.  **	.  **	9	-0.245	-0.271	6.6114	0.678
.  *	.  .	10	-0.146	0.071	7.5957	0.668
.  *	.  *	11	-0.096	-0.174	8.0946	0.705
.  .	.  .	12	-0.065	0.050	8.3664	0.756

**Table 4. Correlogram for the original series of peach production.**

Autocorrelation	Partial correlation	AC	PAC	Q-Stat	P value	
.  *****	.  *****	1	0.646	0.646	8.8507	0.003
.  **	.  **	2	0.259	-0.272	10.365	0.006
.  .	.  *	3	-0.005	-0.079	10.366	0.016
.  *	.  .	4	-0.100	0.017	10.624	0.031

. *	.   .	. *	.   .	5	-0.164	-0.133	11.366	0.045
. **	.   .	. *	.   .	6	-0.249	-0.164	13.230	0.040
. ***	.   .	. **	.   .	7	-0.418	-0.318	18.955	0.008
. ***	.   .	. *	.   .	8	-0.353	0.158	23.443	0.003
. *	.   .	.   .	.   .	9	-0.186	-0.024	24.827	0.003
.   .	.   .	.   .	.   .	10	-0.022	-0.033	24.849	0.006
.   .	.   .	.   .	.   .	11	0.067	0.013	25.082	0.009
.   .	.   .	.   .	.   .	12	0.109	0.006	25.798	0.011

**Table 5. Correlogram for the differenced series of peach production.**

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	P value
.   .	.   .	1	0.067	0.067	0.0912	0.763
. **	. **	2	-0.255	-0.260	1.4859	0.476
. **	. **	3	-0.268	-0.247	3.1487	0.369
.   .	. *	4	-0.031	-0.080	3.1728	0.529
.   .	. *	5	0.214	0.102	4.4106	0.492
.   .	.   .	6	0.050	-0.053	4.4844	0.611
. ***	. ***	7	-0.395	-0.406	9.5136	0.218
. *	. *	8	-0.130	-0.077	10.124	0.256
.   .	. *	9	0.007	-0.191	10.126	0.340
.   .	. *	10	0.168	-0.129	11.424	0.325
.   .	. *	11	0.060	-0.165	11.620	0.393
.   .	. *	12	0.058	0.108	11.839	0.459

**Stationarity Test:** The results of Augmented Dickey and Fuller (1979) (ADF) test are presented in Table 6. The corresponding P value of the ADF test for both the

variables (area and production) are significant at 5% level. Therefore, we conclude that both the variables are stationary at first difference.

**Table 6. Stationarity tests for peach area and production**

Parameter		ADF	P value	Conclusion
Area	Level	-1.99	0.29	I(1)
	Differenced	-2.07*	0.04	
Production	Level	-2.07	0.26	I(1)
	Differenced	-3.23*	0.04	

Note: Asterisk (\*) indicates statistical significance at 5% level.

**Diagnostic Checking and Model Estimation:** Usually, the different models for both area and production of peach were fitted using different values of p and q. ARIMA (1,1,0) was selected as the best model for both

area and production of peach based on the minimum values of Akaike information criterion (AIC), Schwarz information criterion (BIC) and the Hannan-Quinn criterion (HQ) (Table 7).

**Table 7. ARIMA models fitted for time series data on peach area and production and corresponding selection criterion, i.e. AIC, BIC and HQ.**

Parameter	ARIMA Model	AIC	BIC	HQ
Area	ARIMA (1,1,0)	3.77	3.87	3.79
	ARIMA (1,1,1)	3.81	4.01	3.83
	ARIMA (0,1,1)	5.12	5.27	5.14
Production	ARIMA (1,1,0)	7.99	8.09	8.00
	ARIMA (1,1,1)	8.17	8.36	8.19
	ARIMA (0,1,1)	8.19	8.33	8.21

Numerous attempts have been made in the past to forecast area and production for various commodities in different countries. However, the current literature is still inadequate. To the best of our knowledge, none of the scholars has forecasted the peach area and production by using ARIMA model. In the similar context, Khan *et al.* (2008) forecasted the mango production for Pakistan over the period 1982-2004. Results of their study reveal that ARIMA (1,1,1) was a best-forecasted model. However, Jam *et al.* (2013) found ARIMA (0,1,0) as the best model among five forecasted models for the area of mangoes over the period 1961-2009 in Pakistan. They also estimated ARIMA (1,1,0) model, which is in line with the results of our ideal ARIMA model. Nonetheless, the preferred model they selected was ARIMA (0,1,0) based on the AIC, BIC and HQ criteria. Similarly, Ahmad and Mustafa (2006) used ARIMA model for the forecasting of Kinnow production of Pakistan covering the period 1990-91 to 2001-02. They choose ARIMA (3,1,2) as the most suitable model. Moreover, Rahman and Baten (2016) reported the area and production of black gram pulse in Bangladesh from the period 1967-68 to 2013-14. They found ARIMA (0,1,0) as a best suitable model for their respective series. While Iqbal *et al.* (2005) estimated two model ARIMA (1,1,1) and ARIMA (2,1,2) for the wheat area and production in Pakistan respectively.

Furthermore, Biswas *et al.* (2014) analyzed the wheat yield, production and area for Punjab using the forecasting approach. The optimal forecasted models they selected for their respective series were ARIMA (0,1,1), ARIMA (0,1,1) and ARIMA (0,1,0) in the study. After the selection of best forecasted ARIMA (1,1,0) model for both peach area and production, we proceed our analysis to check the model residual diagnostics.

**Residual Diagnostics:** Figures 1 & 2 display descriptive statistics and a histogram of the standardised residuals for normality test of the estimated models. This test concludes that there is no autocorrelation among the residuals of the fitted ARIMA (1,1,0) models for both area and production of peach at 5% level of significance. Furthermore, Jarque-Bera (1987) statistic (Table 8) shows that standardised residuals of both the series are normally distributed.

**Evaluation of ARIMA Models for Peach Area and Production:** The performance, quality and predictability power of the ARIMA (1,1,0) models was measured for both the series based on the lowest errors values of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE) (Table 9).

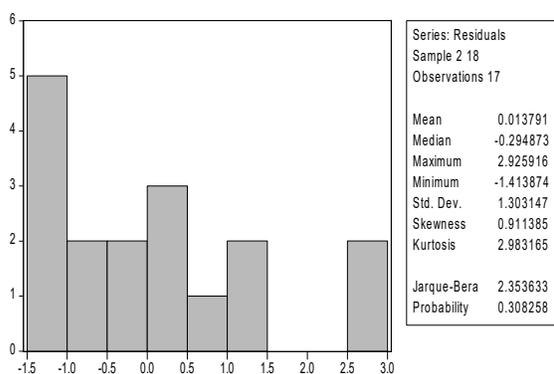


Figure 1. Normality test for peach area

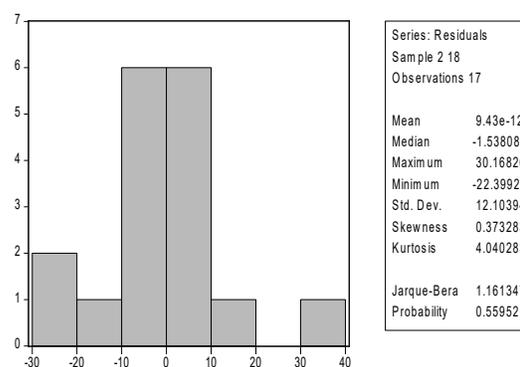


Figure 2. Normality test for peach production

Table 8. Diagnostic tests of residuals Jarque-Bera test.

Parameter	Jarque-Bera	P Value	Sum	Sum Sq. Dev.	Observations
Area	2.35	0.31	214.2	368	18
Production	1.16	0.56	1071.7	4384.62	18

Table 9. Performance of ARIMA models for peach area and production.

Parameter	RMSE	MAE	MAPE
Area	2.65	2.28	19.98
Production	11.18	7.75	14.96

**Forecast for Peach Area and Production Using ARIMA Model:** The 11 years ahead and forecasts for the peach area and production in Pakistan based on ARIMA

(1,1,0) are calculated and presented in Table 10. The projected results of the peach area and production indicated declining trends in area and production in the

coming years. For a 2015-16 forecast of peach area was about 13.39 thousand hectares. A peach area forecast for the year 2025-26 was 11.05 thousand hectares. Similarly, the forecasted value for peach production for the year 2015-16 was 65.05 thousand tonnes. The peach production forecast for the year 2025-26 was estimated

62.53 thousand tonnes. The minimum projection trend showed that peach area and production may decrease in the coming years. This may occur due to the high cost of inputs, energy crises, lack of cold storages, government attention to the crop and high variation in environmental changes.

**Table 10. Forecast for the peach area and production up to 2025-26 for Pakistan.**

Years	Area (000 hectares)	Production (000 tonnes)
2015-16	13.39	65.05
2016-17	12.89	64.17
2017-18	12.49	63.59
2018-19	12.16	63.21
2019-20	11.89	62.96
2020-21	11.67	62.80
2021-22	11.49	62.69
2022-23	11.35	62.63
2023-24	11.23	62.58
2024-25	11.13	62.55
2025-26	11.05	62.53

**Conclusions:** ARIMA model was used for forecasting the peach area and production of Pakistan. The model demonstrated the fact that the technique is appropriate for time series analysis with any pattern of change. In this study, a short run forecast of peach area and production showed a clear indication of declining trend in the following decade. On the basis of these results, it may be determined that ARIMA model could be efficiently used for predicting peach area and production of Pakistan. These projections can help both peach farmers as well as the planner for future planning in Pakistan.

Based on the findings of the study key policy implication emerge. In order to increase peach production, the government of Pakistan should extend the agricultural knowledge for the better management of the peach farm. In this framework, adoption of best management practices for peach fruit is the most suitable approach. The government should make a concrete plan to develop or produce new cultivars and hybrid varieties for more peach production. Institutional services should be strengthened to provide the managerial and technical skill to the peach farmers for better production. Most importantly, integrated strategy is needed to rescue peach cultivation land from increasing urbanization. In conclusion, peach production can be increased by using improved peach cultivars, improved irrigation system and adequate cultural practices. The strategy may be prioritized to bring more barren land under peach cultivation. Expansion of the area under peach would directly increase its production in the future.

**Acknowledgements:** This research was funded by Study on Effectiveness and Operational Mechanism Based on Self-Organizational Control of Regional Fruit Quality.

Key Projects of Humanities and Social Science. Research supported by Northwest A&F University, China (Grant No. Z109021414).

## REFERENCES

- Ahmad, B. and K. Mustafa (2006). Forecasting kinnow production in Pakistan : An econometric analysis. *Int J Agric Biol.* 8: 455–458.
- Anderson, T.W. (1971). *The statistical analysis of time series.* John Wiley, New York.
- Biswas, B., L.K. Dhaliwal, S. Singh and S.K. Sandhu (2014). Forecasting wheat production using ARIMA model in Punjab. *Int J Agric Sci.* 10: 158–161.
- Box, G. and G. Jenkin (1976). *Time series of analysis, forecasting and control.* San Franscico, Holden-Day, California, USA.
- Dickey, D.A. and W.A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc.* 74: 427–431.
- Government of Pakistan (GOP) (2016). *Economic Survey (2015- 2016).* Finance Division, Economic Adviser's Wing, Islamabad.
- Government of Pakistan (GOP) (2015). *Fruit, Vegetables and Condiments Statistics of Pakistan (2014-2015).* Ministry of National Food Security & Research Economic Wing, Islamabad.
- Habib, S. (2015). Peach: Queen of fruits. *Pakistan Food J.* September - October 2015. pp: 26 –27.
- Iqbal, N., K. Bakhsh, A. Maqbool and A.S. Ahmad (2005). Use of the ARIMA model for forecasting wheat area and production in

- Pakistan. *J Agric Soc Sci.* 1: 120–122.
- Jam, F.A., S, Mehmood and Z. Ahmad (2013). Time series model to forecast area of mangoes from Pakistan: An application of univariate ARIMA model. *Acad Contemp Res.* 2:10–15.
- Jarque, C.M. and A.K. Bera (1987). A test for normality of observations and regression residuals. *Int. Stat. Rev.* 55: 163–172.
- Khan, M., K. Mustafa, M. Shah, N. Khan and J.Z. Khan (2008). Forecasting mango production in Pakistan an econometric model approach. *Sarhad J. Agric.* 24(2): 363–70.
- PHDEC (2017). Pakistan Horticulture Development & Export Company. Ministry of Commerce. <http://www.phdec.org.pk/phs.php> Accessed April 5 2017.
- Pindyck, R., S. Daniel and L. Rubinfeld (1991). *Economic Models and Economic Forecasts*, 3rd ed. McGraw Hill International editions (Economic Survey) New York, USA.
- Qureshi, M. N., M. Bilal, M.K Ali and R.M. Ayyub (2014). Modelling of citrus fruit production in Pakistan. *Sci Int.* 26: 733–737.
- Rahman, N. F. and A. Baten (2016). Forecasting area and production of black gram pulse in Bangladesh using ARIMA models. *Pakistan J Agric Sci.* 53: 759–765.
- Tian, F., K. Yang and L. Chen (2017). Realised volatility forecasting of agricultural commodity futures using the HAR model with time-varying sparsity. *Int J Forecast.* 33: 132–152.
- Yu, F., X. Shao, L. Yu, F. Xu and H. Wang (2015). Proteomic analysis of postharvest peach fruit subjected to chilling stress or non-chilling stress temperatures during storage. *Sci Hortic.* 197: 72–89.