International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 09 Issue 12 ǁ 2021 ǁ PP. 54-61

Time series analysis of horse population in the United States of America (USA)

Şenol Çelik
Bingöl University Faculty of Agriculture Biometry and Genetic Department Bingöl, Turkey

Abstract
In this study, time series analysis was aimed using horse count data for the period 1961-2019 in the United States of America.It was observed that the series became stationary after the first difference was taken.Model predictions were made from previously tested models whose parameter predictions were significant and Bayesian Information Criterion (BIC) values were the lowest. The most optimum prediction model defined for the horse population is the one called ARIMA (1,1,0) which is an integrated autoregressive model with a first degree. According to this model, it is predicted that the number of horse will increase continuous between the years 2020-2025 in USA and the number in 2020 will be 10 808 326 and 11 463 026 in 2025.
Keywords: *ARIMA, stationarity, time series, horse*

Date of Submission: 28-11-2021 Date of acceptance: 12-12-2021

I. Introduction

The equine animal, which is a member of the genus Equus (Lat. *Equidae*) in the equine family, is called a horse. Other members of the genus Equus are Wild Donkey (*Lat. Equusferus*), African Wild Donkey (*Lat. Equusafricanus*), Onager 'Asian' Donkey (*Lat. Equushemionus*), Kiang (*Lat. Equus kiang*), Grévy's Zebra (*Lat. Equus strike*), Burcell's 'Plain' Zebra (*Lat. Equusquagga*) and Mountain Zebra (*Lat. Equus zebra*) [1].

People have utilized horse's meat, milk and power. The horse's abilities such as transportation, tillage and the fast movement in war have made horses indispensable for human beings. Due to their skills in pacing horse breeding, the Turks gained an advantage in the wars and were successful. Until recently, there were cavalry units in Turkey. There are still cavalry units to maintain order in many developed European countries. Before mechanization, the horse's power was used in tillage and travels [2].

Horses of different breeds have various characteristics. Arabian horses are extremely durable. They can walk 80 km a day for 3-4 months, and walk for 1-2 days in the desert without any water [3]. The English horse is a fast walking, strong, solidly built, durable breed. They have a high build which provides them with high running ability. The withers height is around 165-170 cm [3]. The Midilli horse has a very small build. Withers height is around 116-120 cm. Ponies' head is short and coarse, their neck is short and strong, their chest is thick, their waist and rump are short, their legs are short and strong. In Turkey, Midilli horses are bred in the Ayvacık district of Çanakkale and they are called the AyvacıkMidilli [4].

According to FAO data, the total number of horses in the world is 59 041 725 in 2019. The USA ranks first in the world with 10 702 799 horses, Mexico is in second place with 6 382 699 and Brazil is in third place with 5 850 154 [5].

This study's aim is to model and predict horse existence with time series in the United States of America, with the highest breeding rate in the world.

II. Material and Method

The material of the research is 1961-2020 number of horses values supplied from the Food and Agriculture Organization of the United Nations [5]. The dependent variable was number of horses figures while the independent variable was year series. These variables were selected in order to be able to make reasonable estimations with the models to be performed using time series analysis method.

ARIMA Models

A pth-order **autoregressive model** AR(p) model is point out as [6].

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

AR(p) model uses a linear combination of past values of the target to make forecasts. A qth-order moving average process, expressed MA(q), is indicated by [7].

$$y_t = -\theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t$$

ARMA(p,q) model composed of apth-order **autoregressive and** qth-order moving average process and it is showed by [8].

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

In order for time series models to be applied, series must be stationary and white noise [9].
BIC is defined as:

$$BIC = -2\log(\hat{L}) + \log(N)\, K$$

Where $\hat{L}$ is the maximum likelihood, K the number of parameters to be estimated in the model, and N the sample size [10].

III. Results and Discussion

The graph presenting the number of horses bred in the USA in the 1961-2019 period is given in Figure 1.

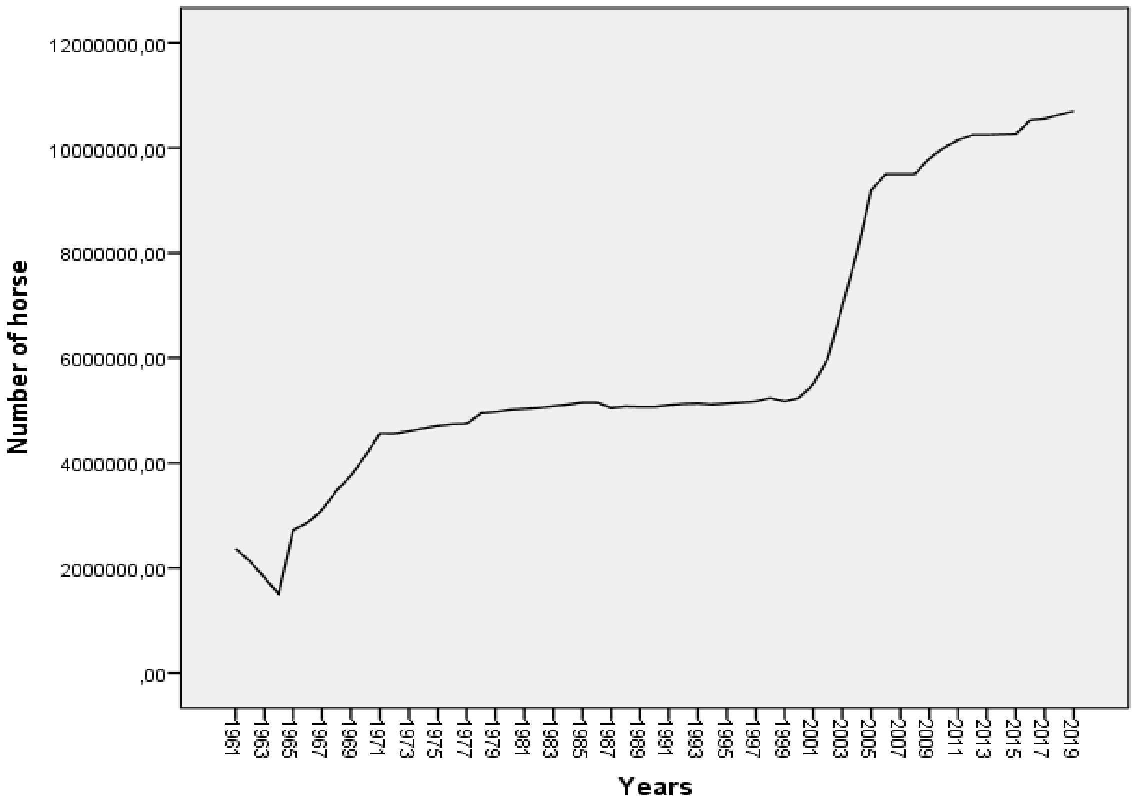

Figure 1. Number of horses in the USA between 1961 and 2019

When the graph presenting the number of horses is examined in Figure 1, a trend is seen. In order to see this trend more clearly, autocorrelation (ACF) and partial autocorrelation (PACF) graphs of the series were examined. In the ACF and PACF graphs, it is noticed that there is a trend in the series and it is not stationary. To make the series stationary, its first difference is taken. ACF and PACF graphs of the series whose first difference is taken are given in Figure 2 and Figure 3, respectively.

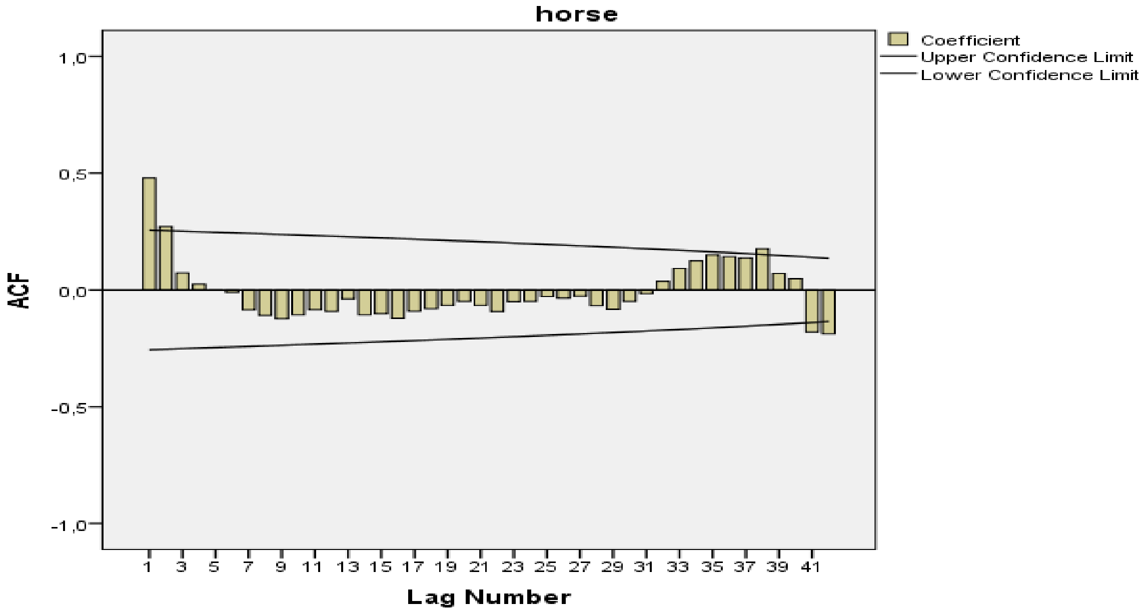

Figure 2. ACF plot of the 1st difference of the series

The ACF graph in Figure 2 shows that the series is stationary. Because after the 2nd delay, the lag values were seen to be within the confidence limits.

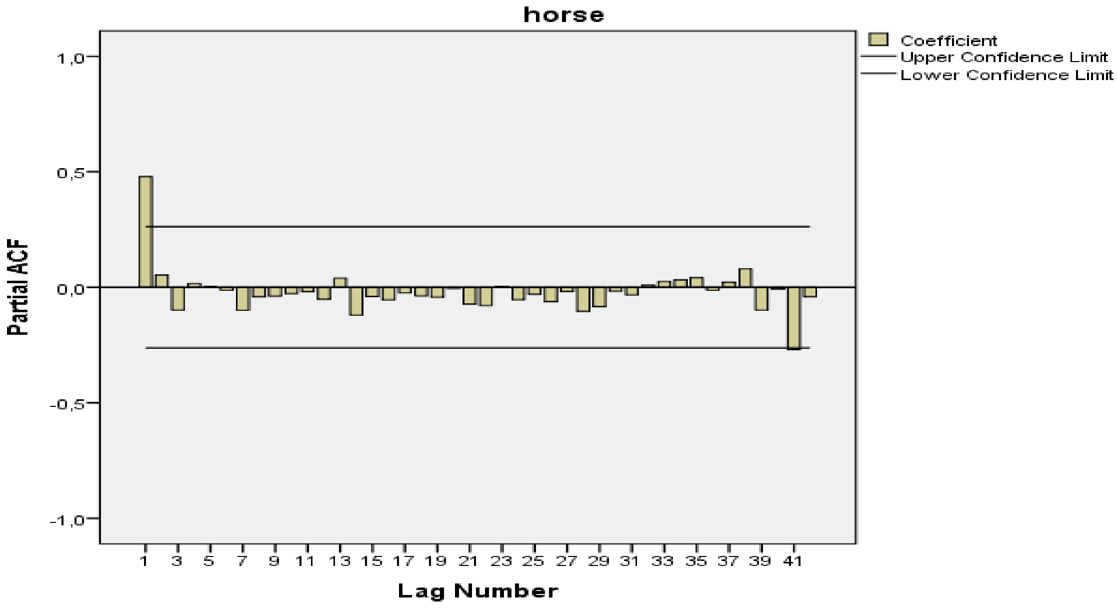

Figure 3. PACF plot of the 1st difference of the series

When Figure 2 and Figure 3 are evaluated together, the lag values in the ACF graph rapidly approached zero after the 2nd lag. On the other hand, in the PACF graph, the lag values rapidly approached zero after the 1st lag. The proximity to zero in the PACF graph occurred faster. Therefore, the ARIMA (1,1,0) model was deemed appropriate. However, in both ACF and PACF graphs, the lag values' proximity to zero obtained after significant lags, are close to each other. Considering this situation, ARIMA(1,1,0) and ARIMA(0,1,2) and ARIMA(1,1,2) models can be examined comparatively. The most suitable model was selected according to the values of the goodness of fit. Information regarding the goodness of fit statistics is presented in Table 1.

Table 1. Goodness-of-fit statistics for different models

Statistic	ARIMA(1,1,0)	ARIMA(0,1,2)	ARIMA(1,1,2)
R-squared	0.989	0.989	0.989
RMSE	270686.553	273179.634	273894.134
BIC	25.157	25.246	25.321

In Table 1 R^2 values are equal in all models. The ARIMA(1,1,0) model with the smallest RMSE and BIC statistics is the most suitable model.
Parameter estimation and prediction results will be evaluated with ARIMA(1,1,0), that is, the first-order integrated autoregressive time series model. Parameter estimation of the ARIMA(1,1,0) model is given in Table 2.

Table 2. ARIMA Model Parameters

	Estimate	SE	t	p
Constant	136634.42	67606.45	2.021	0.048
AR(1)	0.485	0.116	4.166	0.001

Ljung-Box Q=3.607, p=0.999

As seen in Table 2, the parameter coefficient of the model was found to be statistically significant ($p<0.001$). Since Ljung-Box Q=3.607 and $p>0.05$, the model is sufficient. Therefore, the model was found to be suitable. When the distribution of error terms in Figure 4 is examined, all error terms are found to be within confidence limits. Therefore, the series is white noised. According to the ARIMA(1,1,0) model, predictions were made for the years 2020-2025 and the results are presented in Table 3.

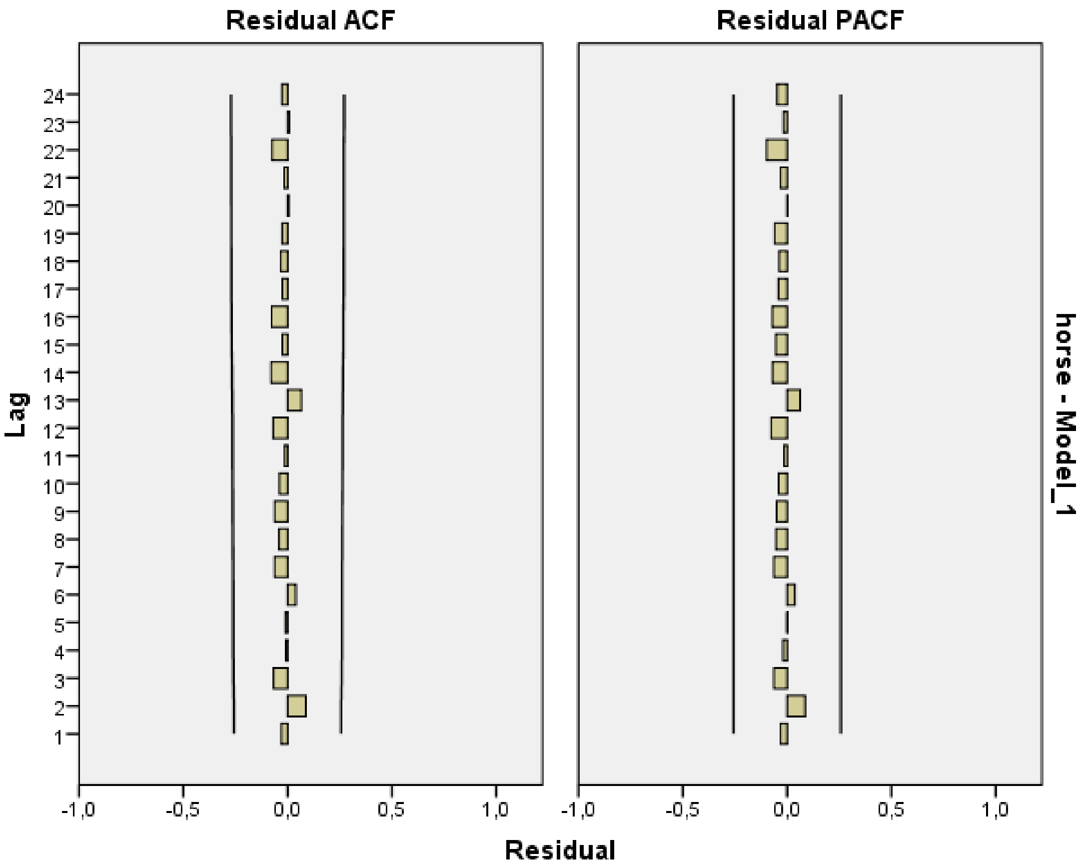

Figure 4. ACF and PACF plots of residual terms

Table 3. Forecast of the number of horses in the USA by years

Years	2020	2021	2022	2023	2024	2025
Forecast	10 808 326	10 929 884	11 059 212	11 192 305	11 327 223	11 463 026

As seen in Table 3, an increase is expected in the 2020-2025 period. It is estimated that the number of horses will be 10 808 326 in 2020 and this will be increased to 11 463 026 by the year 2025. Observed and estimated horse numbers by years are in harmony (Figure 5).

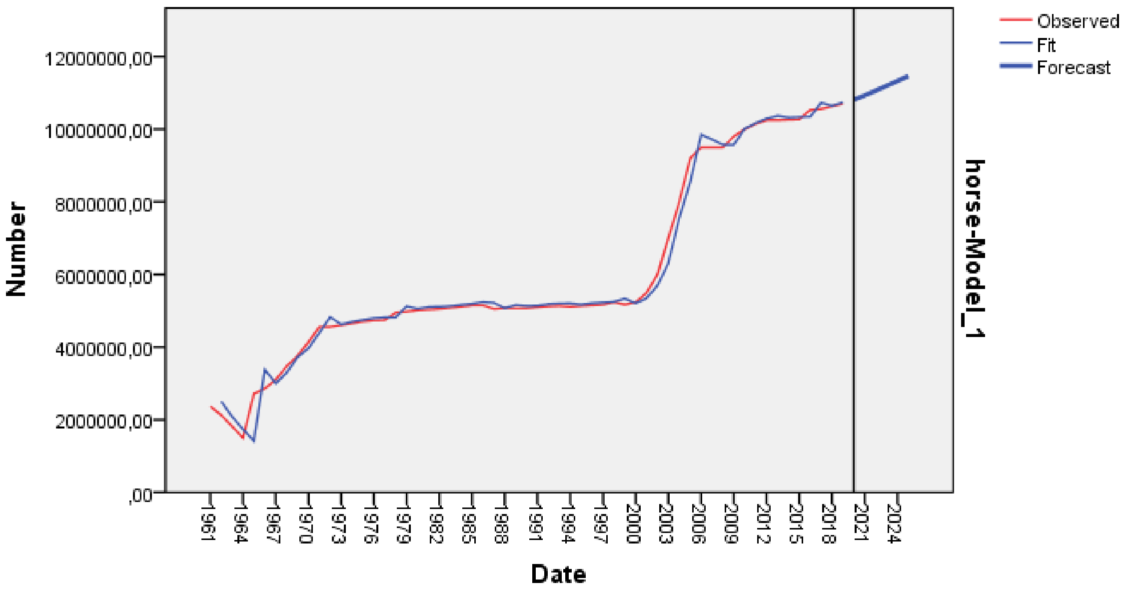

Figure 5. Distribution of observed and estimated values by years

Regarding time series models in livestock, the number of pigs [11], mules [12], camels [13], donkeys [14], the milk yield in ovine animals [15], the presence of ovine animals [16] and honey production amount [17] was studied and predicted. However, all of the mentioned studies were conducted for Turkey.

IV. Conclusion

In this study, the series regarding the horse existence in the USA for the period 1961-2019 is not stationary however it becomes stationary after the first difference is taken. As a result of time series analysis, the most suitable model was determined as ARIMA(1,1,0). In line with the determined ARIMA model, a prediction regarding the number of horses until 2025 was made. According to the forecast results, the expected number of horses in 2022 is 11 059 212 which by the year 2024 will be increased to 11 327 223 and by the year 2025 to 11 463 026 . As a result of this, the number of horses in the USA will increase continuously in the coming years and will gain significant momentum. This is necessary in terms of the contribution the horse population will provide to the country's economy and maintaining its position in the world countries ranking.

References

[1]. Yılmaz, O. 2016. Atçılık. VeniVidiVici Yayınları: 12, Ankara, ISBN: 978-605-89397-6-9
[2]. Boztepe, S., Aytekin, İ. 2013. At yetiştirme. Selçuk Üniversitesi Basımevi, Konya
[3]. Yarkın, İ. 1962. Atçılık. Ankara Üniversitesi Ziraat Fakültesi Yayınları: 40. Ders kitabı: 20, Ankara
[4]. Arpacık, R. 1994. At yetiştiriciliği. Şahin Matbaası, Ankara
[5]. FAO, 2019. Food and Agriculture Organization of the United Nations. Live animal. https://www.fao.org/faostat/en/#data/QCL
[6]. Cooray, T.M.J.A. 2008. Applied Time Series. Analysis and Forecasting. Narosa Publishing House Pvt. Ltd., New Delhi.
[7]. Cryer, J. D. 1986. Time Series Analysis. PWS Publishing, USA.
[8]. Hamilton, J.D. 1994. Time Series Analysis. Princeton University Press Princeton, New Jersey
[9]. Kadılar, C., Çekim, H.Ö. 2020. SPSSve R Uygulamalı Zaman Serileri Analizine Giriş. Seçkin Yayıncılık San. ve Tic. A. Ş., Ankara.
[10]. Schwarz, G. 1978. Estimating the dimension of a model. Ann Stat., 6:461-464
[11]. Çelik, Ş. 2021. Modeling the number of pigs in Turkey through ARIMA models and Artificial Neural Networks. Quest Journals. Journal of Research in Agriculture and Animal Science, 8(4):39–45
[12]. Çelik, Ş. 2021. Modeling the number of mules in Turkey through Time Series Analysis and Artificial Neural Networks. European Journal of Advances in Engineering and Technology, 8(9):22–28
[13]. Çelik, Ş. 2021.Modeling and Estimation of Camel Population in Turkey with Time Series Analysis and Artificial Neural Networks. International Journal of Research in Engineering and Science (IJRES), 9(5):38–44

[14]. Çelik, Ş. 2021.Estimation of the quantity of donkeys in Turkey via Artificial Neural Network and Time-Series Analysis. Research Inventy: International Journal of Engineering and Science, 11(9):7–20
[15]. Çelik, Ş. 2015. Investigation milk yield in small ruminants by region in Turkey. International Journal of Innovation Sciences and Research, 4(10):516-519
[16]. Çelik, Ş. 2015. Estimation of number of small cattle through ARIMA models in Turkey. Journal of Mathematics and System Science 5:464-473
[17]. Çelik, Ş. 2015. Türkiye'de balüretiminin zaman serileri ile modellenmesi. Sakarya Üniversitesi Fen Bilimleri Dergisi, 19(3):377-382