Yayın Geliş Tarihi: *25.10.2016* *Dokuz Eylül Üniversitesi*
Yayın Kabul Tarihi: *18.06.2017* *...İktisadi ve İdari Bilimler Fakültesi Dergisi*
Online Yayın Tarihi: *04.12.2017* *Cilt:32, Sayı:2, Yıl:2017, ss. 143-182*

Detecting the Best Seasonal ARIMA Forecasting Model for Monthly Inflation Rates in Turkey[1][2]

Mehmet ÖZMEN[3] **Sera ŞANLI[4]**

Abstract

In this study, it has been aimed to find the best 'Seasonal Autoregressive Integrated Moving Average (SARIMA)' model for monthly inflation rates for Turkish economy over the period 1995:01-2015:03. Before the model identification based on Box Jenkins methodology, HEGY monthly seasonal unit root test has been applied. The orders of seasonal differencing have been detected through OCSB and CH tests. Finally, ARIMA(1,1,1)(1,0,2)[12] with drift model chosen by using stepwise selection method and ARIMA(1,1,1)(2,0,0)[12] with drift model chosen by using non-stepwise selection have been compared. The results have shown that the former model is better as the best fitted SARIMA model.

Keywords: *Inflation, Box Jenkins, SARIMA, HEGY, OCSB, CH, Stepwise Selection.*
JEL Classification Codes: *C01, C22, C51, E31.*

Türkiye İçin Enflasyon Oranının Uygun Mevsimsel ARIMA Modeli İle Belirlenmesi

Öz

Bu çalışmada Türkiye'nin 1995:01-2015:03 dönemine ilişkin aylık enflasyon oranları için en iyi 'Mevsimsel Otoregresif Bütünleşik Hareketli Ortalama (SARIMA)' modelini bulmak amaçlanmıştır. Box Jenkins metodolojisine dayalı model tanımlamasından önce HEGY mevsimsel birim kök testi uygulanmıştır. Mevsimsel fark alma dereceleri OCSB ve CH testleri kullanılarak saptanmıştır. Son olarak, sırasıyla adımsal (stepwise) ve adımsal olmayan (non-stepwise) seçim yöntemleri kullanılarak seçilen sürüklenmeli ARIMA(1,1,1)(1,0,2)[12] ve ARIMA(1,1,1)(2,0,0)[12] modelleri karşılaştırılmıştır. Sonuçlar adımsal yöntem kullanılarak seçilen sürüklenmeli ARIMA(1,1,1)(1,0,2)[12] modelinin en iyi uyan SARIMA modelini belirlemede adımsal olmayan modele göre daha iyi olduğunu göstermiştir.

Anahtar Kelimeler: *Enflasyon, Box Jenkins, SARIMA, HEGY, OCSB, CH, Adımsal Seçim.*
JEL Sınıflandırma Kodları: *C01, C22, C51, E31.*

---

[1] This study has been derived from the Master Thesis that has been prepared in consultancy of Assoc. Prof. Mehmet Ozmen called "The Econometric Analysis of Seasonal Time Series: Applications on Some Macroeconomic Variables (Sanlı, 2015)".

[2] This study has been supported by TUBITAK (The Scientific and Technological Research Council of Turkey) – BIDEB (Scientist Support Department) within the scope of 2211-E Direct National Scholarship Programme for PhD Students.

[3] Prof. Dr., Çukurova Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, mozmen@cu.edu.tr

[4] Arş. Gör., Çukurova Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, sanlis@cu.edu.tr

1. INTRODUCTION

In terms of policy makers, it is of great importance to have a reliable inflation rate forecast. In this context, the most suitable model should be accessed using 'Seasonal Autoregressive Integrated Moving Average (SARIMA)'. Since SARIMA models reveal more effective results in terms of handling the seasonal component of the series apart from the non-seasonal one when compared to the traditional ARIMA models. In this application, it has been aimed to find the best model for monthly inflation rates and therefore monthly (seasonally unadjusted) consumer price index (CPI) data have been utilized for Turkish economy over the period 1995:01-2015:03.

In modelling monthly inflation rates that are very crucial to design effective economic strategies, choosing a suitable seasonal ARIMA model which includes both seasonal and non-seasonal behaviours is not an easy task. Since such models give point to the recent past rather than distant past, primarily they are convenient for short term forecasting and this implies that long-term forecasts from ARIMA models are less reliable than short term forecasts (Aidoo, 2010: 3).

The study of Canova and Hansen (1995) presents Lagrange Multiplier (LM) tests of the null hypothesis of no unit roots at seasonal frequencies denoting the presence of deterministic seasonality contrary to the tests of Dickey, Hasza and Fuller (DHF) (1984) and Hylleberg, Engle, Granger and Yoo (HEGY) (1990) tests dealing with the null of presence of seasonal unit roots. They generalize the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) (1992) test framework.

Tam and Reinsel (1997) examine the locally best invariant unbiased (LBIU) and point optimal invariant test procedures for a unit root in the seasonal moving average (SMA) operator for SARIMA and make use of the monthly non-agricultural industry employment series for males age 16-19 modelled by Hillmer, Bell and Tiao (1983). The results for conducted simulations have revealed that for this series, seasonality is stochastic and therefore seasonal differencing is appropriate. They also apply their tests to different types of seasonal time series data and find some of these series to have deterministic seasonality.

In the study by Lim and McAleer (2000), the presence of stochastic seasonality is examined to clarify the nonstationary quarterly international tourist arrivals from Hong Kong and Singapore to Australia from 1975:Q1 to 1996:Q4 using HEGY (1990) procedure. Since the presence of seasonal unit roots gives an insight into a varying seasonal pattern that is against a constant seasonal pattern, the Box Jenkins SARIMA process is possible to be a more suitable model for tourist arrivals rather than a deterministic seasonal model with seasonal dummy variables.

Cosar (2006) has tried to examine the seasonal properties of the Turkish CPI through Beaulieu and Miron's (1993) extension of the classical HEGY test developed by Hylleberg et al. (1990) and the LM-type CH seasonal unit root test procedures with the aim to specify the seasonality accurately in econometric models. In Cosar's (2006) study, there has been an evidence of both deterministic and nonstationary stochastic seasonality in the CPI series of Turkey.

In their paper, Chang and Liao (2010) have aimed to forecast the monthly outbound tourism departures of three major destinations from Taiwan to Hong Kong, Japan and U.S.A. respectively using the SARIMA model.

Saz (2011) examines the efficacy of SARIMA models for forecasting Turkish inflation rates from 2003 to 2009 and presents a methodological approach for a combination of a systematic SARIMA forecasting structure and the stepwise selection procedure of the Hyndman-Khandakar (HK) algorithm. This combination is expressed to give rise to choosing a best single SARIMA model which is SARIMA(0,0,0)(1,1,1) model with one degree of seasonal integration, one seasonal autoregressive (AR) and one seasonal moving average (MA) part. According to a structural break analysis, the Turkish inflation rates have been found to display a range of structural breaks with the latest being in mid-2003 and stochastic nature of Turkish inflation has been found to outweigh its deterministic nature.

Our study has mainly focused on searching for the best-fitted SARIMA model for the monthly inflation rates in order to provide the best forecast. Therefore, following the Box-Jenkins approach, in the application part first model identification and estimation of parameters will be presented. Subsequent to this, diagnostic checking results based on the residuals of the possible model will be given place in order to make certain about the white-noise characteristic of residuals which becomes a vital assumption for a good ARIMA model.

The rest of this paper has been organized as follows: Section 2 provides the background for the analytical approach to SARIMA models, Beaulieu and Miron's (1993) extension of the classical HEGY test, OCSB and CH test; section 3 gives the information about the data set and discussions on the empirical results. Finally, section 4 presents the conclusions.

2. THEORETICAL BACKGROUND

2.1. Seasonal ARIMA Models (SARIMA)

The characterization of seasonal series occurs by a strong serial correlation at the seasonal lag. As known, the classical decomposition of the time series consists of a trend component, a seasonal component and a random noise component. But, in practice it may not be logical to assume that the seasonality component repeats itself exactly in the same way cycle after cycle. SARIMA models allow for randomness in the seasonal pattern from one cycle to the next (Brockwell and Davis, 2006: 320).

Box and Jenkins (1970) present an extension of the ARIMA model in order to take seasonal effects into consideration. At the core of idea, trying to adjust a cyclical effect takes place for adding this seasonal component. For example, in the case of monthly data, the observation y_t may depend in part on y_{t-12} accounting for an annual effect. In the same manner, for daily data, the dependence may be realized through y_{t-7} representing a weekly effect. Coping with these dependencies in order to remove the seasonal effect in question may be possible via differencing the data. However, one can also specify AR or MA relationships at the seasonal interval in question. For this case, Box and Jenkins (1970) define a general multiplicative SARIMA model shown as ARIMA (p,d,q) x $(P,D,Q)_s$, where the lower-case letters p,d,q indicate the nonseasonal orders and the upper-case letters P,D,Q

indicate the seasonal orders of the process with period *s* (that is, *s* is the number of observations per year). The parentheses mean that the seasonal and nonseasonal elements are multiplied (Hamaker and Dolan, 2009: 198-199; Pankratz, 1983: 281).

Before giving a clear definition for SARIMA, assume that X_t $(t = 0, \pm 1, \pm 2, ...$ is an ARMA (p,q) process if $\{X_t\}$ is stationary and if for every *t*,

$$X_t - \phi_1 X_{t-1} - .... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q} \tag{1}$$

where $\{Z_t\} \sim WN(0, \sigma^2)$. (1) can be written symbolically as

$$\phi(L) X_t = \theta(L) Z_t, \quad (t = 0, \pm 1, \pm 2, ... \tag{2}$$

where ϕ and θ are the p^{th} and q^{th} degree polynomials

$$\phi(z) = 1 - \phi_1 z - .... - \phi_p z^p \tag{3}$$

and

$$\theta(z) = 1 + \phi_1 z + ... + \theta_q z^q \tag{4}$$

and *L* is the backward shift operator defined by $L^j X_t = X_{t-j}$, $j = 0, \pm 1, \pm 2, ...$ These ϕ and θ polynomials are mentioned as the AR and MA polynomials respectively of the difference equations (2) (Brockwell and Davis, 2006: 78). If we fit an ARMA (p,q) model $\phi(L) Y_t = \theta(L) Z_t$ to the differenced series $Y_t = (1 - L^S) X_t$, then the model for the original series becomes $\phi(L)(1 - L^S) X_t = \theta(L) Z_t$. This is a special case of the general SARIMA model which will be defined as follows:

Definition: If d and *D* are nonnegative integers, then { X_t } is a *seasonal ARIMA* $(p,d,q) \text{x} (P,D,Q)_S$ process with period *s* if the differenced series $Y_t = (1 - L)^d (1 - L^S)^D X_t$ is a causal ARMA process defined

$$\phi(L)\Phi(L^S)Y_t = \theta(L)\Theta(L^S)Z_t,\ \{Z_t\} \sim WN(0,\sigma^2) \quad (5)$$

where $\Phi(z) = 1 - \Phi_1 z - .... - \Phi_P z^P$ (seasonal AR(P) characteristic polynomial), $\Theta(z) = 1 + \Theta_1 z + ... + \Theta_Q z^Q$ (seasonal MA(Q) characteristic polynomial) with $\phi(z)$ and $\theta(z)$ expressed in (3) and (4) respectively (Brockwell and Davis, 2002: 203). On the other hand, a more general multiplicative SARIMA model can be expressed by adding a constant term δ to take the case of a deterministic trend into consideration as follows:

$$\phi(L)\Phi(L^S)Y_t = \delta + \theta(L)\Theta(L^S)Z_t \quad (6)$$

and substituting $Y_t = (1-L)^d(1-L^S)^D X_t = \Delta_S{}^D\Delta^d X_t$ into (6), it becomes

$$\phi(L)\Phi(L^S)\Delta_S{}^D\Delta^d X_t = \delta + \theta(L)\Theta(L^S)Z_t \quad (7)$$

(Shumway and Stoffer, 2011: 157).

As seen in the definition given above, derivation of $\{Y_t\}$ comes from the original series $\{X_t\}$ using both simple differencing (in order to remove trend) and seasonal differencing $\Delta_S = 1 - L^S$ to remove seasonality. For instance, when $d = D = 1$ and $s = 12$, then Y_t becomes

$$Y_t = \Delta\Delta_{12}X_t = \Delta_{12}X_t - \Delta_{12}X_{t-1} \quad = (X_t - X_{t-12}) - (X_{t-1} - X_{t-13}) \quad (8)$$

Now take a SARIMA model of order $(1,0,0)$ x $(0,1,1)_{12}$. Then this model can be written in the following equation:

$$(1-\phi L)Y_t = (1+\theta L^{12})Z_t \quad (9)$$

where $Y_t = \Delta_{12}X_t$. Then we find

$$X_t = X_{t-12} + \phi(X_{t-1} - X_{t-13}) + Z_t + \theta Z_{t-12} \quad (10)$$

so that X_t depends on X_{t-1}, X_{t-12} and X_{t-13} as well as the innovation at time $(t-12)$ (Chatfield, 1996: 60).

Now let us take an ARIMA $(1,0,0)$ x $(1,0,1)_{12}$ process with a periodicity of length 12 (since, s=12). In this example, it is obvious that X_t does not require seasonal and nonseasonal differencing at all since $d=0$ and $D=0$. On the other hand, the seasonal part of the process is composed of one AR $(P=1)$ and one MA $(Q=1)$ component at lag 12. In addition, there is a nonseasonal AR term at lag 1 $(p=1)$. The multiplication of the two AR operators in the lag operator form can be expressed as

$$(1-\phi_1 L)(1-\Phi_1 L^{12})X_t = (1+\Theta_1 L^{12})Z_t \qquad (11)$$

(Pankratz, 1983: 281).

In identifying SARIMA model, the first task is to find values d and D which reduce the series to stationarity and remove most of the seasonality. Then, we need to assess the values of p, P, q and Q by examining the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of the differenced series at lags which are multiples of *s* and choosing a SARIMA model in which ACF and PACF have a similar shape. Ultimately, the model parameters may be estimated through an appropriate iterative procedure. For details, see Box and Jenkins (1970, chap. 9) (Chatfield, 1996,: 60-61) (all AR and MA polynomial representations have been taken from Brockwell and Davis, 2006: 78).

2.1.1. Stationarity and Invertibility Conditions

Representing a model in a multiplicative form is a big convenience in terms of expressing the seasonal and nonseasonal components separately and controlling the stationarity and invertibility conditions. For instance, take an ARIMA $(2,0,1)$ x $(1,0,2)_s$ model and express it in a lag operator form as follows:

$$(1-\phi_1 L-\phi_2 L^2)(1-\Phi_1 L^s)X_t=(1+\Theta_1 L^s+\Theta_2 L^{2s})(1+\theta_1 L)Z_t \qquad (12)$$

The stationary requirement applies only to the AR coefficients. The nonseasonal part of (12) has the same stationarity conditions as for an $AR(2)$: $|\phi_2|<1,$ $\phi_2-\phi_1<1,$ and $\phi_2+\phi_1<1$. Now we need to apply a seperate stationary condition for the AR seasonal part. It is the same as for a nonseasonal AR(1) model, except in this case we have a seasonal $AR(1)_s$ component; so the condition becomes $|\Phi_1|<1$.

As in the case of stationarity, we need to consider invertibility condition which applies only to the MA part of (12) for nonseasonal and seasonal components separately. For the nonseasonal part, the condition is $|\theta_1|<1$. The conditions on the seasonal part are the same as for a nonseasonal MA(2) model, except in this case there exists an $MA(2)_s$ component. Therefore the joint conditions are given as $|\Theta_2|<1,$ $\Theta_2-\Theta_1<1$ and $\Theta_2+\Theta_1<1$ (Pankratz, 1983: 285). (AR and MA polynomial representations have been taken from Brockwell and Davis, 2006: 78).

2.1.2. The Expanded Model

It should be noted that all multiplicative SARIMA models can be telescoped into an ordinary ARMA (p,q) model in the variable

$$Y_t \overset{def}{=} \Delta_S{}^D \Delta^d X_t \qquad (13)$$

For instance, consider that the series $\{x_t\}_{t=1}^T$ follows a SARIMA $(0,1,1)$ x $(12,0,1,1)$ or ARIMA $(0,1,1)$ x $(0,1,1)_{12}$ model. For this process, we have

$$(1-L^{12})(1-L)X_t=(1+\theta_1 L)(1+\Theta_1 L^{12})Z_t \qquad (14)$$

and it becomes

$$Y_t = (1 + \theta_1 L + \Theta_1 L^{12} + \theta_1 \Theta_1 L^{13}) Z_t \tag{15}$$

where $Y_t = (1 - L^{12})(1 - L) X_t$. Hence, it can be said that this multiplicative SARIMA model has an ARMA (0,13) representation where only the coefficients θ_1 , $\theta_{12} \stackrel{def}{=} \Theta_1$ and $\theta_{13} \stackrel{def}{=} \theta_1 \Theta_1$ are not zero and all other coefficients of the MA polynomial are equal to zero. So, if the model in question is SARIMA $(0,1,1)$ x $(12,0,1,1)$ given in (14), only the two coefficients which are θ_1 and Θ_1 have to be estimated. However, for the ARMA (0,13), instead we have to estimate the three coefficients which are θ_1, θ_{12} and θ_{13}. Therefore, it is apparent that SARIMA models take a parsimonious model structure into account and a model specification such as (15) is called an expanded model. In addition, we can say that only an expanded multiplicative model can be estimated directly (Chen, Schulz and Stephan, 2003: 233-234).

2.1.3. Theoretical ACFs and PACFs for Seasonal Processes

In SARIMA models, estimated acfs and pacfs display the same expected behaviours as in the structure of nonseasonal models. For seasonal time series data, observations *s* time periods apart $(z_t, z_{t-s}, z_{t+s}, z_{t-2s}, z_{t+2s}, ...)$ have characteristics in common. So, observations *s* periods apart are expected to be correlated and in this manner, acfs and pacfs for seasonal series should have nonzero coefficients at one or more multiples of lag *s* $(s, 2s, 3s, ...)$. If we observe nonseasonal and purely seasonal acfs and pacfs, it is seen that the coefficients appearing at lags 1,2,…. in the former appear at lags $s, 2s, 3s, ...$ in the latter.

This similarity between nonseasonal and seasonal acfs and pacfs makes the seasonal analysis simpler. So, having knowledge of nonseasonal acfs and pacfs helps give a description of identical patterns occurring at multiples of lag *s* (Pankratz, 1983: 270-271). For more details, see Box and Jenkins (1976, chap. 9).

2.2. Testing for Seasonal Unit Roots in Monthly Data

Beaulieu and Miron (1992) make an extension of HEGY testing procedure for monthly data. $\varphi^{*}(L)$ is a polynomial associated with roots that are outside the unit circle and it can be expressed as

$$\varphi(L)^{*} y_{13,t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + \varepsilon_t \qquad (16)$$

where,

$$y_{1,t} = (1 + L + L^2 + L^3 + L^4 + L^5 + L^6 + L^7 + L^8 + L^9 + L^{10} + L^{11}) y_t \qquad (17)$$

$$y_{2,t} = -(1 - L + L^2 - L^3 + L^4 - L^5 + L^6 - L^7 + L^8 - L^9 + L^{10} - L^{11}) y_t \qquad (18)$$

$$y_{3,t} = -(L - L^3 + L^5 - L^7 + L^9 - L^{11}) y_t \qquad (19)$$

$$y_{4,t} = -(1 - L^2 + L^4 - L^6 + L^8 - L^{10}) y_t \qquad (20)$$

$$y_{5,t} = -\frac{1}{2}(1 + L - 2L^2 + L^3 + L^4 - 2L^5 + L^6 + L^7 - 2L^8 + L^9 + L^{10} - 2L^{11}) y_t \qquad (21)$$

$$y_{6,t} = \frac{\sqrt{3}}{2}(1 - L + L^3 - L^4 + L^6 - L^7 + L^9 - L^{10}) y_t \qquad (22)$$

$$y_{7,t} = \frac{1}{2}(1 - L - 2L^2 - L^3 + L^4 + 2L^5 + L^6 - L^7 - 2L^8 - L^9 + L^{10} + 2L^{11}) y_t \qquad (23)$$

$$y_{8,t} = -\frac{\sqrt{3}}{2}(1 + L - L^3 - L^4 + L^6 + L^7 - L^9 - L^{10}) y_t \qquad (24)$$

$$y_{9,t} = -\frac{1}{2}(\sqrt{3} - L + L^3 - \sqrt{3}L^4 + 2L^5 - \sqrt{3}L^6 + L^7 - L^9 + \sqrt{3}L^{10} - 2L^{11}) y_t \qquad (25)$$

$$y_{10,t} = \frac{1}{2}(1-\sqrt{3}L+2L^2-\sqrt{3}L^3+L^4-L^6+\sqrt{3}L^7-2L^8+\sqrt{3}L^9-L^{10})y_t \qquad (26)$$

$$y_{11,t} = \frac{1}{2}(\sqrt{3}+L-L^3-\sqrt{3}L^4-2L^5-\sqrt{3}L^6-L^7+L^9+\sqrt{3}L^{10}+2L^{11})y_t \qquad (27)$$

$$y_{12,t} = -\frac{1}{2}(1+\sqrt{3}L+2L^2+\sqrt{3}L^3+L^4-L^6-\sqrt{3}L^7-2L^8-\sqrt{3}L^9-L^{10})y_t \qquad (28)$$

$$y_{13,t} = (1-L^{12})y_t \qquad (29)$$

(Beaulieu and Miron, 1992: 2-4).

With these transformations (y_i,s) of y_t, the seasonal unit roots are excluded at given frequencies while they are preserved at remaining frequencies. To give an example, consider the y_{1t} transformation. While it eliminates the seasonal unit roots, it preserves the long-run or zero frequency unit root. In table 1, the outline of long run and seasonal frequencies has been presented:

Table 1. Long Run and Seasonal Frequencies for Seasonal Unit Root Tests in Monthly Data

Frequency	**Period**	**Cycles/year**	**Root**	**Filter**	**Tested hypothesis** $H_0:$ Unit Root
0 **Long run**	∞	0	1	$(1-L)$	$\pi_1 = 0$
$\frac{\pi}{6}, \frac{11\pi}{6}$ **Annual**	12;1.09	1; 11	$\frac{1}{2}(\sqrt{3} \pm i)$	$(1-\sqrt{3}L+L^2)$	$\pi_{11} \cap \pi_{12} = 0$
$\frac{\pi}{3}, \frac{5\pi}{3}$ **Semiannual**	6; 1.2	2; 10	$\frac{1}{2}(1 \pm \sqrt{3}i)$	$(1-L+L^2)$	$\pi_7 \cap \pi_8 = 0$
$\frac{\pi}{2}, \frac{3\pi}{2}$	4; $\frac{4}{3}$	3; 9	$\pm i$	$(1+L^2)$	$\pi_3 \cap \pi_4 = 0$
$\frac{2\pi}{3}, \frac{4\pi}{3}$ **Quarterly**	3; 1.5	4; 8	$-\frac{1}{2}(1 \pm \sqrt{3}i)$	$(1+L+L^2)$	$\pi_5 \cap \pi_6 = 0$
$\frac{5\pi}{6}, \frac{7\pi}{6}$	2.4; 1.7	5; 7	$-\frac{1}{2}(\sqrt{3} \pm i)$	$(1+\sqrt{3}L+L^2)$	$\pi_9 \cap \pi_{10} = 0$
π **Bimonthly**	2	6	-1	$(1+L)$	$\pi_2 = 0$

Note. The information on first five columns have been obtained from Dıaz- Emparanza & López-de-Lacalle (2006).

If π_2 through π_{12} are significantly different from zero, there will be no seasonal unit roots and the pattern that the data display becomes deterministic or constant seasonal. Therefore, in this situation the dummy variable representation can be applied for modelling this pattern. The implication of the statement just given is that if there are seasonal unit roots, the corresponding π_i are zero. Due to the fact that pairs of complex unit root are conjugates, these roots will exist only in case pairs of

$\pi's$ are jointly equal to zero. If π_1 through π_{12} are all unequal to zero, we experience a stationary seasonal pattern and seasonal dummy variables can be used to model such a pattern. Also, when the coefficient for a given π is statistically not different from zero, it can be said that data have a varying seasonal pattern. If $\pi_1 = 0$, we cannot reject the presence of root 1 with long-run frequency and if all π_i are equal to zero, it becomes suitable to apply the $(1-L^{12})$ filter. If only some pairs of $\pi's$ are zero, the relevant operators can be used. In Abraham and Box (1978), it is exemplified that sometimes these operators may be adequate. (Franses, 1991: 101; Maddala and Kim, 1998: 370; Sørensen, 2001: 77).

2.3. OCSB Test

Osborn, Chui, Smith and Birchenhall (OCSB) (1988) have modified the Hasza and Fuller (1982) test framework to detect the presence of multiplicative differencing filter $\Delta_1\Delta_s$. That is, the OCSB test investigates whether $(1-L)$ or $(1-L^s)$ operators or both of them or none of them should be applied to data. The OCSB regression model in the original form is expressed as

$$\Delta_1\Delta_s y_t = \beta_1\Delta_s y_{t-1} + \beta_2\Delta_1 y_{t-s} + \varepsilon_t \tag{30}$$

and it can be generalized with deterministic components as follows:

$$\eta(L)\Delta_1\Delta_s y_t = \mu_t + \gamma_1\Delta_s y_{t-1} + \gamma_2\Delta_1 y_{t-s} + \varepsilon_t \tag{31}$$

where $\eta(L)$ is an AR polynomial (lag polynomial with roots outside the unit circle), $\Delta_s = (1-L^s), \Delta_1 = (1-L)$ and

$$\mu_t = \alpha_0 + \sum_{s=1}^{S-1}\alpha_s D_{s,t} + \beta_0 t + \sum_{s=1}^{S-1}\beta_s D_{s,t} t \tag{32}$$

Here, *t* is a deterministic trend. In the original study, the seasonal trend is not given place in μ_t i.e. $\beta_s = 0$ for $\forall$s. However, Franses and Koehler (1998) suggest the model (31) with the β parameters not being equal to zero in μ_t so that the test becomes applicable to y_t series showing increasing seasonal variation. In order to find out which filter is suitable for y_t, the significances of γ_1 and γ_2 are tested. When both γ_1 and γ_2 are equal to zero ($\gamma_1 = \gamma_2 = 0$), using $\Delta_1\Delta_s$ filter is suitable. When $\gamma_1 = 0$ and $\gamma_2 \neq 0$, Δ_1 filter should be selected; when $\gamma_1 \neq 0$ and $\gamma_2 = 0$, Δ_s filter is suitable. If both γ_1 and γ_2 are unequal to zero ($\gamma_1 \neq \gamma_2 \neq 0$), in that case no differencing filter is required (Franses, 1998: 563; Maddala and Kim, 1998: 366; Zhang, 2008: 11; Platon, 2010: 2-3).

2.4. Canova-Hansen (CH) Test

The study of Canova and Hansen (1995) presents Lagrange Multiplier (LM) tests of the null hypothesis of no unit roots at seasonal frequencies against the alternative of a unit root at either a specific seasonal frequency or a set of selected seasonal frequencies. So the test statistics of CH are derived from the LM principle that necessitates only the estimation of the model under the null using least square techniques and they are fairly simple functions of the residuals. These tests are also a framework for testing seasonal stability. CH tests complement the tests of Dickey et al. (1984) and Hylleberg et al. (1990) that examine the null of seasonal unit roots at one or more seasonal frequencies. So, it is clear that contrary to these seasonal unit root tests, the null hypothesis of CH test is that the process is stationary (that is, stationary seasonality rather than nonstationary seasonality). Here the rejection of the null hypothesis would imply the nonstationarity of the data. Although the null of CH test is stationary seasonality, for simplicity they refer to their tests as seasonal unit root tests. Since seasonal intercepts stand for the deterministic components of seasonality and they are assumed to be constant over the sample, under the null hypothesis of stationarity the tests by Canova and Hansen can also be introduced as the tests for constancy of seasonal intercepts over time (Canova and Hansen, 1995: 237-238).

3. APPLICATION

In this application, it has been aimed to find the best model for monthly inflation rates and therefore monthly (not seasonally adjusted) CPI data have been utilized for Turkish economy over the period 1995:01-2015:03 (Index 2010=1.00). Data have been obtained from Organization for Economic Co-operation and Development. This application has been carried out at the R Project for Statistical Computing-version 3.1.3. by using "forecast" and "uroot" packages. Since inflation is measured by the percentage change in CPI, inflation rates have been calculated by using the following transformation:

$$INF = \frac{CPI_t - CPI_{t-1}}{CPI_{t-1}}.100 \quad (33)$$

where INF denotes inflation rate, CPI_t denotes consumer price index at time *t* and CPI_{t-1} denotes consumer price index at time *t*-1.

The graph of inflation data has been presented in Figure 1:

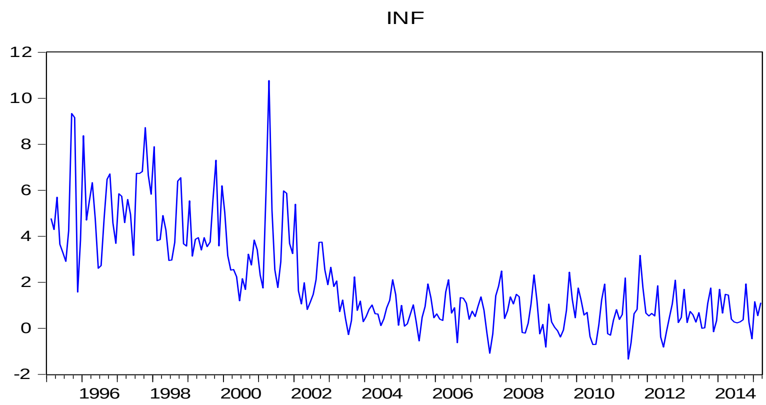


Figure 1. Graph of Inflation Series against Time

It is apparent from Figure 1 that inflation data are nonstationary with a non-constant mean and unsteady variance and follow some seasonal pattern. For this reason, first of all the series should be checked for seasonal unit roots at all seasonal frequencies and if INF series includes all seasonal unit roots, seasonal differencing operator has to be applied to this series. If INF series has seasonal unit roots only at some frequencies, filters corresponding to available unit roots at each given frequency have to be applied. Briefly, before constructing a suitable ARIMA model for our seasonal series, we should make a data transformation in a way to make the series stationary by taking Box-Jenkins methodology into consideration.

Before the model identification, in order to detect at which frequencies INF series has unit roots and to decide about the appropriate order of differencing filter, we should recourse to HEGY monthly seasonal unit root test apart from CH test. The null hypotheses differ for CH and HEGY tests. In the former, the null hypothesis implies the stationarity case at all seasonal cycles while the latter implies the presence of seasonal unit root.

Figure 2 and Figure 3 show the ACF and PACF of the original inflation series for maximum lag numbers of 48 respectively. When looked at the correlogram of series in Figure 2, the autocorrelation coefficient is seen to decline very slowly towards zero with increasing lag length implying that the series is nonstationary. On the other hand, seasonal lags (12 24, 36,48) are clear to be significant. Thus, the presence of any seasonal unit root other than a zero (long-run) frequency unit root has to be detected.

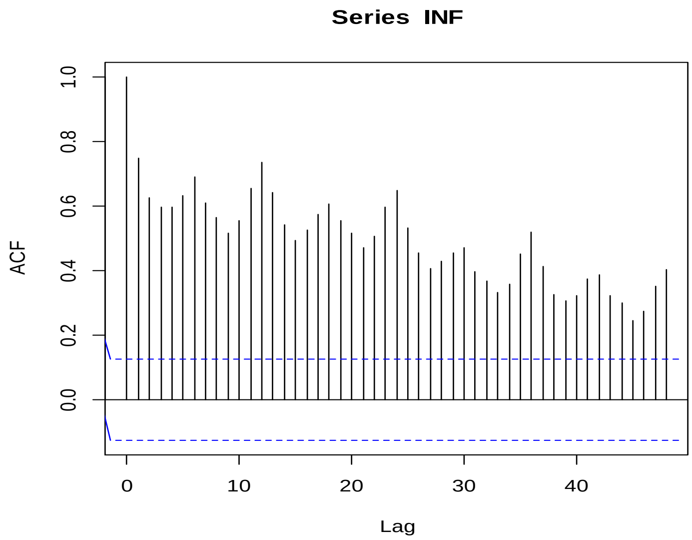


Figure 2. ACF of Inflation Series (for Lag.Max=48)

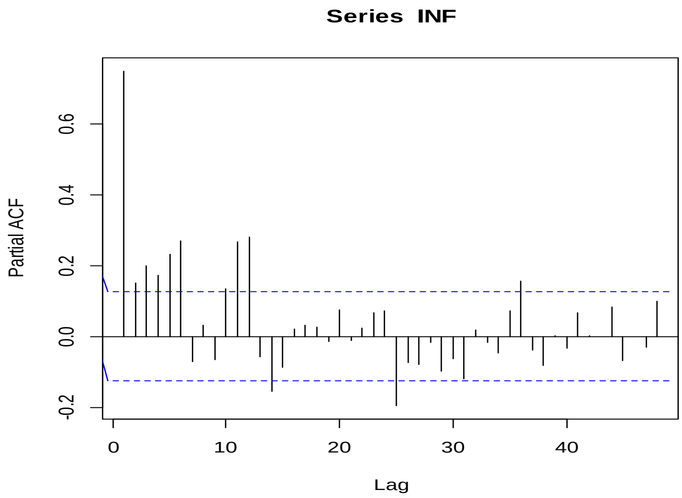


Figure 3. PACF of Inflation Series (for Lag.Max=48)

Table 1 has presented long-run and seasonal frequencies for monthly series in details. In this study, the monthly seasonal unit root analysis has been carried out by using three different lag order selection methods. First, significant lags have been added to the four deterministic regressions [with only constant (C); constant and trend (C, T); constant and dummies (C, D); constant, trend and dummies (C, T, D)] and one regression with no deterministic components (None) in order to make certain about that the residuals are white noise (that is, insignificant lags have been removed until all selected lags become significant). These test results have been given in Table 2. As mentioned before, the first two hypotheses which are $\pi_1 = 0$ and $\pi_2 = 0$ are tested by *t*-test and the other five joint hypotheses which are $\pi_3 = \pi_4 = 0$, $\pi_5 = \pi_6 = 0$, $\pi_7 = \pi_8 = 0$, $\pi_9 = \pi_{10} = 0$ and $\pi_{11} = \pi_{12} = 0$ are tested by F-test.

Table 2. HEGY Monthly Seasonal Unit Root Test Results for Inflation Series (by Using Significant Lags)

Auxiliary Regression Null Hypotheses	Seasonal Frequency	Estimates for the Model with Constant	Estimates for the Model with Constant and Trend	Estimates for the Model with Constant and Dummies	Estimates for the Model with Constant, Trend and Dummies	Estimates for the Model with None
$\pi_1 = 0$	0	**-1.537***	**-0.288***	**-1.294***	**-1.548***	-2.762
$\pi_2 = 0$	π	-2.348	-2.313	-3.588	-3.608	-2.347
$\pi_3 = \pi_4 = 0$	$\pi/2$	6.966	6.761	20.174	20.222	6.960
$\pi_5 = \pi_6 = 0$	$2\pi/3$	4.220	4.008	14.163	14.297	4.208
$\pi_7 = \pi_8 = 0$	$\pi/3$	**1.675***	**1.606***	9.036	9.132	**1.668***
$\pi_9 = \pi_{10} = 0$	$5\pi/6$	12.656	12.342	22.248	22.352	12.662
$\pi_{11} = \pi_{12} = 0$	$\pi/6$	5.461	5.236	14.104	14.524	5.435

Note. * denotes insignificant estimates (*p>.05) at 5% significance level.

For HEGY test applications, critical values have been obtained from Franses and Hobjin (1997) for S=12 and for 5% significance level (see pp. 29-33) for 20 years (that is, 240 observations).

When looked at Table 2, the results for the hypothesis $\pi_1 = 0$ have revealed that the presence of the zero (non-seasonal) frequency unit root is accepted depending on the non-rejection of the null hypothesis $\pi_1 = 0$ at all deterministic models (except none model). Thus, original INF series is not stationary at zero frequency. Having examined the other hypotheses, all other hypotheses implying the presence of a unit root at seasonal frequency except the hypothesis $\pi_7 = \pi_8 = 0$ are seen to be rejected for all deterministic models and therefore it is concluded that there are no seasonal unit roots at $\pi, \pm\frac{\pi}{2}, \pm\frac{2\pi}{3}, \pm\frac{5\pi}{6}$ and $\pm\frac{\pi}{6}$ frequencies. In other saying, there are conjugate complex seasonal unit roots only at $\pm\frac{\pi}{3}$ frequencies corresponding to (2, 10) cycles per year for "Constant", "Constant and Trend" and "None" models. From this point of view, seasonal cycles can be said to follow mostly a deterministic structure.

Table 3. HEGY Monthly Seasonal Unit Root Test Results for Inflation Series (by Using AIC for Lags)

Auxiliary Regression Null Hypotheses	Seasonal Frequency	Estimates for the Model with Constant	Estimates for the Model with Constant and Trend	Estimates for the Model with Constant and Dummies	Estimates for the Model with Constant, Trend and Dummies	Estimates for the Model with None
$\pi_1=0$	0	**-1.546***	**-0.579***	**-1.417***	**-0.935***	-2.542
$\pi_2=0$	π	-2.541	-2.515	-2.978	-2.991	-2.534
$\pi_3=\pi_4=0$	$\pi/2$	4.938	4.905	18.391	18.360	4.937
$\pi_5=\pi_6=0$	$2\pi/3$	3.373	3.310	7.305	7.267	3.359
$\pi_7=\pi_8=0$	$\pi/3$	**1.212***	**1.197***	**5.727***	**5.756***	**1.207***
$\pi_9=\pi_{10}=0$	$5\pi/6$	14.009	13.633	20.506	20.451	13.975
$\pi_{11}=\pi_{12}=$	$\pi/6$	3.897	3.842	13.631	13.624	3.860

Note. * denotes insignificant estimates (*p>.05) at 5% significance level.

Table 3 presents monthly HEGY seasonal unit root test results based on AIC. The results are almost the same as Table 2 with regard to statistical significance: Since the hypothesis $\pi_1=0$ could not be rejected at 5% significance level (meaning that non-rejection of the presence of root $+1$), the presence of the zero frequency unit root has been accepted. Thus, inflation series is nonstationary and seasonal unit roots have been detected only at $\pm\frac{\pi}{3}$ frequencies for all five models given in Table 3.

Table 4. HEGY Monthly Seasonal Unit Root Test Results for Inflation Series (by Using BIC for Lags)

Auxiliary Regression Null Hypotheses	Seasonal Frequency	Estimates for the Model with Constant	Estimates for the Model with Constant and Trend	Estimates for the Model with Constant and Dummies	Estimates for the Model with Constant, Trend and Dummies	Estimates for the Model with None
$\pi_1 = 0$	0	**-1.537***	**-0.288***	**-1.499***	**-1.315***	-2.762
$\pi_2 = 0$	π	-2.348	-2.313	-3.232	-3.278	-2.347
$\pi_3 = \pi_4 = 0$	$\pi/2$	6.966	6.761	15.593	15.816	6.960
$\pi_5 = \pi_6 = 0$	$2\pi/3$	4.220	4.008	9.53	9.773	4.208
$\pi_7 = \pi_8 = 0$	$\pi/3$	**1.675***	**1.606***	6.756	6.956	**1.668***
$\pi_9 = \pi_{10} = 0$	$5\pi/6$	12.656	12.342	17.772	17.988	12.662
$\pi_{11} = \pi_{12} = 0$	$\pi/6$	5.461	5.236	10.906	11.126	5.435

Note. * denotes insignificant estimates (*p>.05) at 5% significance level.

Table 4 considers the results of monthly HEGY seasonal unit root test based on BIC (Bayesian Information Criterion). Table 4 and Table 2 results do not differ. In conclusion, three methods discussed in terms of different lag criteria have revealed only the presence of conjugate complex seasonal unit roots at $\pm\frac{\pi}{3}$ frequencies corresponding to (2, 10) cycles per year. The presence of all other seasonal unit roots with $\pi, \pm\frac{\pi}{2}, \pm\frac{2\pi}{3}, \pm\frac{5\pi}{6}$ and $\pm\frac{\pi}{6}$ has been rejected and it has been concluded that seasonal cycles mostly display a deterministic structure. Therefore, there is no need to take the seasonal difference of INF series. However, since the presence of zero frequency unit root cannot be denied; we have to take the first difference of INF series. In that case, INF series is not seasonally integrated and thus applying the seasonal difference filter $(1-L^{12})$ to the series is not required. Beaulieu and Miron (1992: 18) have also explained more clearly why applying $(1-L^{12})$ filter to the series is not required in that way: "The appropriateness of applying the filter $(1-L^d)$ to a series with a seasonal component, as advocated by Box and Jenkins (1970) depends on the series being integrated at zero and all of the seasonal frequencies". Briefly, this explanation holds since the presence of all seasonal unit roots has not been accepted and there is weak evidence of seasonal unit roots on monthly series.

Table 5. CH Test Results for Inflation Series

Tested Frequencies	L-Statistic	Critical Values		
		1%	5%	10%
$\frac{\pi}{6},\frac{\pi}{3},\frac{\pi}{2},\frac{2\pi}{3},\frac{5\pi}{6},\pi$	2.005	3.27	2.75	2.49

After applying to HEGY test, now Table 5 presents CH test results in order to make inference about the seasonal behaviour of INF series. Contrary to the HEGY test, the null hypothesis of CH is the stationarity of all seasonal cycles while the alternative hypothesis is the presence of seasonal unit root (indicating to the presence of stochastic seasonality). According to the results, since calculated L-statistic (2.005) is smaller than not only 5% critical value (2.75) but also 1% (3.27) and 10% (2.49) critical values, we fail to reject the null hypothesis saying that seasonal pattern is deterministic. Therefore it can be said that the result of CH test is consistent with the result of HEGY test and once again there is no need for seasonal differencing operator. However, there is one important thing that since the presence of only conjugate complex seasonal unit roots with $\pm\frac{\pi}{3}$ frequencies has been determined with the adoption of the hypothesis $\pi_7=\pi_8=0$, INF series should be transformed by the necessary filters corresponding to these frequencies. Filters corresponding to all frequencies have been presented in Table 1. Therefore, the necessary filter corresponding to $\pm\frac{\pi}{3}$ frequencies has been expressed as $(1-L+L^2)$. On the other hand, as expressed before, since the series includes zero (non-seasonal) frequency unit root, the first difference operator $(1-L)$ should also be applied. So, the necessary transformation that will be made in INF series will be $(1-L)(1-L+L^2)$. More precisely, if the new series to be obtained is called " $f\inf$ " (meaning filtered inflation), $f\inf$ will be formed as follows:

$$f\inf = \Delta(1-L+L^2) = \Delta(INF - INF(-1) + INF(-2))$$
$$= INF - 2INF(-1) + 2INF(-2) - INF(-3) \qquad (34)$$

The ACF function of the " $f\inf$ " series obtained after this transformation given above for maximum lags of 48 is given in Figure 4 and PACF function is given in Figure 5:

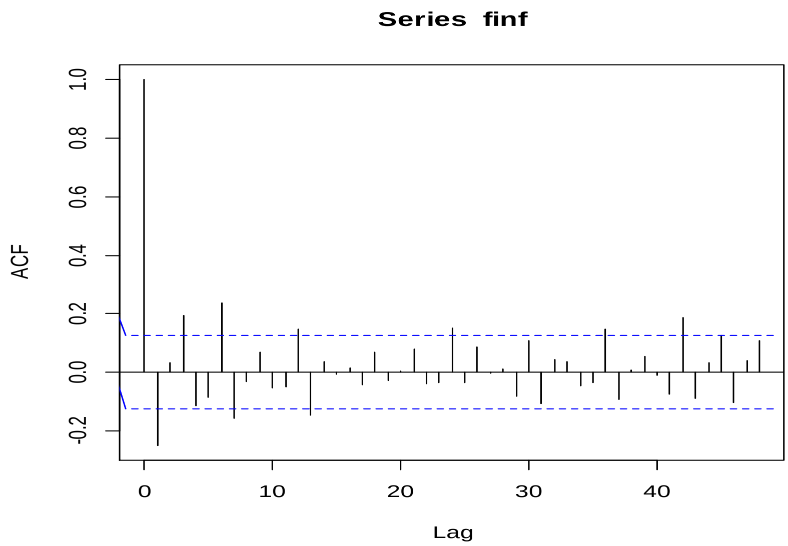


Figure 4. ACF of Filtered Inflation Series ($f\inf$) for Lag.Max=48

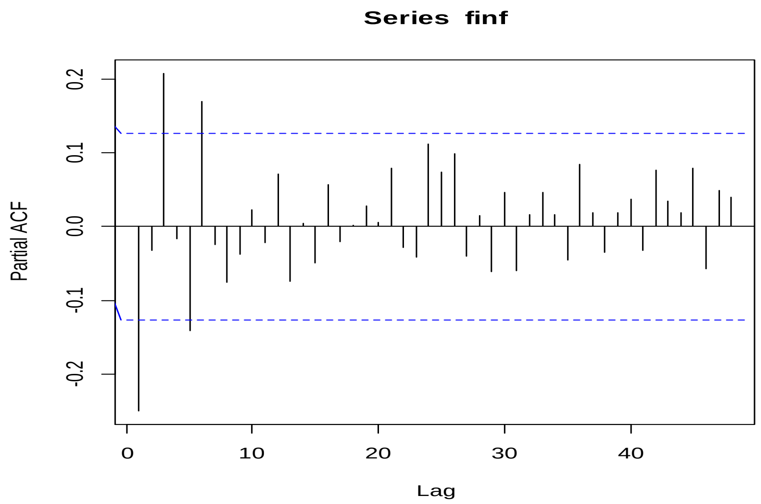


Figure 5. PACF of Filtered Inflation Series ($f\inf$) for Lag.Max=48

As seen in Figure 4 and Figure 5, the significant spikes at lag 1 in both ACF and PACF suggest a non-seasonal MA(1) and non-seasonal AR(1) components. When looked at the PACF correlogram, there has been found no significant spikes at seasonal lags 12, 24, 36, 48. However, 6th lag is seen to be significant. Therefore, it can be said again that series follows a semi-annual seasonal pattern (corresponding to the filter $(1-L+L^2)$ and thus to the hypothesis $\pi_7 \cap \pi_8 = 0$) as consistent with monthly seasonal unit root results and since there are no significant spikes at seasonal lags in PACF, once again it can be said that seasonal differencing is not required for the series.

"Forecast" package in R software offers us a very practical formula concerned with determining the order of both seasonal differencing and first-degree differencing benefiting from OCSB and CH tests. By running the following codes, we can compare the results that will obtained here with the results described above:

Table 6. R Codes and Outputs for Determining the Order of Seasonal Differencing by Using OCSB and CH Tests

R Codes and Outputs
>nsdiffs(INF,12,test="ocsb") [1] 0
>nsdiffs(INF,12,test="ch") [1] 0

Note. [1]The function "nsdiffs" estimates the order of seasonal differencing in a series to satisfy stationarity condition. Here "12" indicates the length of seasonal period of the series and "test" expresses the kind of seasonal unit root test to be applied (OCSB or CH).
[2]For more information, see (Hyndman, 2015).
[3]For OCSB test, the null hypothesis is H_0 **:** Seasonal unit root exists while H_0 **:** Seasonal cycles are stationary for CH test.

As seen in Table 6, the result "[1] 0" reveals the number of seasonal differencing for inflation series as "0 (zero)" as a result of carrying out both OCSB test and CH test. Thus, there has been no need to take any seasonal difference. These results show consistency with the results expressed before. Now with the codes given in Table 7, let us verify that original INF series is not stationary at zero frequency:

Table 7. R Codes and Outputs for Determining the Number of First Differences by Using KPSS and ADF Tests

R Codes and Outputs
>ndiffs(INF,test="kpss") [1] 1
>ndiffs(INF,test="adf") [1] 1

Note. [1] The function "ndiffs" estimates the number of first differences in order to make the series stationary.
[2] For more information, see (Hyndman, 2015).
[3]For KPSS test, the null hypothesis implies the stationarity of series (or the absence of unit root) while the null of ADF test implies the non-stationarity case of series in interest at the non-seasonal level (or the presence of unit root).

The results of practical codes that take place in Table 7 tell us that INF series should be first-degree differenced.

Another simple method for determining the optimal order of differencing comes from Box-Jenkins rule of thumb: The optimum order of differencing is the one with the smallest standard deviation (Akuffo and Ampaw, 2013: 15). In order to detect the optimal order, standard deviations corresponding to different orders of differencing are given in Table 8:

Table 8. Standard Deviations for Detecting the Optimal Order of Differencing by Box-Jenkins Rule of Thumb

Order of Differencing	Non	First	Second	Third
Standard Deviations	2.243578	**1.492247**	2.274733	3.806330

Hence, the minimum standard deviation is realized in first-degree differenced form with a value of 1.492247. Hence, once again we have verified the optimum order as 1.

Now after the orders of seasonal and non-seasonal differences are determined in order to satisfy the stationarity condition of original series (since the series should be stationary for SARIMA modelling), we should determine AR, SAR, MA and SMA (seasonal moving average) orders to construct the best model.

In the model identification, possible best models have been tried to be discovered by “auto.arima” function in “forecast” package of R software. The method for selecting the best-fitted model is based on choosing AIC, AICc (Corrected Akaike Information Criterion) and BIC with minimum values. Mostly, the model that provides minimum AIC (or AICc) rather than BIC is a candidate to be selected as the best-fitted one. In Table 9, suggested ARIMA models by utilizing from OCSB and ADF tests have been presented with AICc and AIC information criteria:

Table 9. AICc and AIC Values for Suggested ARIMA Models of INF Series by Using Stepwise Selection

Suggested ARIMA models	AICc	AIC
ARIMA(2,1,2)(1,0,1)[12] with drift	Inf	Inf
ARIMA(0,1,0) with drift	2560.113	2560.063
ARIMA(1,1,0)(1,0,0)[12] with drift	2494.328	2494.158
ARIMA(0,1,1)(0,0,1)[12] with drift	2466.34	2466.17
ARIMA(0,1,0)	2558.086	2558.069
ARIMA(0,1,1)(1,0,1)[12] with drift	Inf	Inf
ARIMA(0,1,1) with drift	2495.07	2494.969
ARIMA(0,1,1)(0,0,2)[12] with drift	2449.736	2449.481
ARIMA(1,1,1)(0,0,2)[12] with drift	2440.443	2440.084
ARIMA(1,1,0)(0,0,2)[12] with drift	2505.766	2505.511
ARIMA(1,1,2)(0,0,2)[12] with drift	2441.56	2441.079
ARIMA(0,1,0)(0,0,2)[12] with drift	2532.184	2532.015
ARIMA(2,1,2)(0,0,2)[12] with drift	2444.814	2444.194
ARIMA(1,1,1)(0,0,2)[12]	2440.654	2440.398
ARIMA(1,1,1)(1,0,2)[12] with drift	**2405.964**	**2405.484**
ARIMA(1,1,1)(1,0,1)[12] with drift	Inf	Inf
ARIMA(1,1,1)(0,0,1)[12] with drift	2453.309	2453.054
Suggested ARIMA models	AICc	AIC
ARIMA(0,1,1)(1,0,2)[12] with drift	Inf	Inf
ARIMA(2,1,1)(1,0,2)[12] with drift	Inf	Inf
ARIMA(1,1,0)(1,0,2)[12] with drift	Inf	Inf
ARIMA(1,1,2)(1,0,2)[12] with drift	Inf	Inf
ARIMA(0,1,0)(1,0,2)[12] with drift	2498.705	2498.449
ARIMA(2,1,2)(1,0,2)[12] with drift	Inf	Inf
ARIMA(1,1,1)(1,0,2)[12]	Inf	Inf
ARIMA(1,1,1)(2,0,2)[12] with drift	Inf	Inf

As shown in Table 9, the best model under the stepwise-selection method among other models has been chosen as ARIMA(1,1,1)(1,0,2)[12] model with drift with the smallest AICc value 2405.964 and the smallest AIC value 2405.484. All other models which have greater AIC values have been provided only for comparison purposes. After selecting the best model based on AIC and AICc, we need to estimate the significance of parameters:

Table 10. Estimates of Parameters for ARIMA (1,1,1)(1,0,2)[12] Model with Drift

	AR(1)	MA(1)	SAR(1)	SMA(1)	SMA(2)	DRIFT
Estimate	0.1750	-0.8857	0.8862	-0.7102	0.1813	-0.9323
Standard Error	0.0763	0.0375	0.0537	0.0847	0.0746	1.3789
Sigma^2 estimated: 1233 Log likelihood: -1194.59 AIC: 2405.48 **AICc: 2405.96** BIC: 2429.88						

As clearly seen in Table 10, the coefficients of *ARIMA (1,1,1)(1,0,2)[12] Model with Drift* are significantly different from zero.

Table 11. BIC Values for Suggested ARIMA Models of INF Series by Using Stepwise Selection

Suggested ARIMA models	BIC
ARIMA(2,1,2)(1,0,1)[12] with drift	Inf
ARIMA(0,1,0) with drift	2567.032
ARIMA(1,1,0)(1,0,0)[12] with drift	2508.098
ARIMA(0,1,1)(0,0,1)[12] with drift	2480.11
ARIMA(0,1,0)	2561.554
ARIMA(0,1,1)(1,0,1)[12] with drift	Inf
ARIMA(0,1,1) with drift	2505.423
ARIMA(0,1,1)(0,0,2)[12] with drift	2466.905
ARIMA(1,1,1)(0,0,2)[12] with drift	2460.993
ARIMA(1,1,0)(0,0,2)[12] with drift	2522.935
ARIMA(1,1,2)(0,0,2)[12] with drift	2465.473
ARIMA(0,1,0)(0,0,2)[12] with drift	2545.954
ARIMA(2,1,2)(0,0,2)[12] with drift	2472.072
ARIMA(1,1,1)(0,0,2)[12]	**2457.822**
ARIMA(1,1,1)(1,0,2)[12]	Inf
ARIMA(1,1,1)(0,0,1)[12]	2467.839
ARIMA(0,1,1)(0,0,2)[12]	2462.758
ARIMA(2,1,1)(0,0,2)[12]	2464.79

ARIMA(1,1,0)(0,0,2)[12]	2517.457
ARIMA(1,1,2)(0,0,2)[12]	2462.153
ARIMA(0,1,0)(0,0,2)[12]	2540.471
ARIMA(2,1,2)(0,0,2)[12]	2469.057

Table 11 presents BIC values for each suggested ARIMA model. If we take only BIC into account, the best model is seen to be ARIMA(1,1,1)(0,0,2)[12] model with a minimum value of 2457.822. The estimates of parameters of ARIMA(1,1,1)(0,0,2)[12] model are given in Table 12:

Table 12. Estimates of Parameters for ARIMA (1,1,1)(0,0,2)[12] Model

	AR(1)	MA(1)	SMA(1)	SMA(2)
Estimate	0.2412	-0.9183	0.2685	0.2295
Standard Error	0.0701	0.0249	0.0690	0.0569
Sigma^2: 1435 log-likelihood: -1219.39 AIC: 2448.78 AICc: 2449.03 BIC: 2466.2				

If ARIMA(1,1,1)(1,0,2)[12] model with drift chosen by AIC (or AICc) in Table 9 and ARIMA(1,1,1)(0,0,2)[12] model chosen by BIC in Table 11 are compared, ARIMA(1,1,1)(1,0,2)[12] model with drift is chosen because of having smaller information criteria.

For selecting the best-fitted model (to find out how well the model fits the data), we need to continue with the examination of residuals diagnostics (or diagnostic checking) in order to find out whether the residuals display a white noise process which is a vital assumption of a good ARIMA model (zero mean, constant variance, no serial correlation). In this stage, first we will have a look at Box-Ljung test results in order to make sure about residuals have no remaining autocorrelation. The null and alternative hypotheses are given respectively as follows:

$H_0:$ The residuals are random (independently distributed)

$H_1:$ The residuals are not random (not independently distributed, displaying serial correlation)

Table 13. Box-Ljung Test Results of ARIMA(1,1,1)(1,0,2)[12] Model with Drift at Seasonal Lags

Seasonal Lags	X-squared Statistics	p-value
12	10.6567	0.1543
24	21.996	0.2845
36	30.6726	0.4828
48	39.8145	0.6102

Table 13 presents the autocorrelation check results for the residuals of ARIMA(1,1,1)(1,0,2)[12] with drift model at seasonal lags and according to given results, we cannot reject the null hypothesis saying that residuals are independent and hence conclude about the absence of autocorrelation problem depending on the statistically insignificant chi-squared statistics (since p-values for Box-Ljung statistic are greater than 5% significance level for all seasonal lags 12,24,36,48). Therefore, this model can be said to fit the data well. This result is also verified by looking at the correlogram of residuals shown in Figure 6. All acf and pacf values in Figure 6 are within the significance limits and mean of the residuals seem to be randomly distributed around zero. Thus, the residuals appear to be white noise.

Now let us check the normality of ARIMA(1,1,1)(1,0,2)[12] model with drift residuals.

Table 14. Jarque - Bera Normality Test Results of ARIMA(1,1,1)(1,0,2)[12] Model with Drift

X-squared Statistic	Asymptotic p-value
3.2092	**0.201**

Table 14 shows the Jarque-Bera test Results. As well known, the null hypothesis for the test is that residuals are normally distributed and the alternative hypothesis is that residuals are not normally distributed. Insignificant X-squared statistic with an asymptotic p-value of 0.201 that is greater than 5% significance level reveals that the null hypothesis cannot be rejected concluding that residuals are normally distributed.

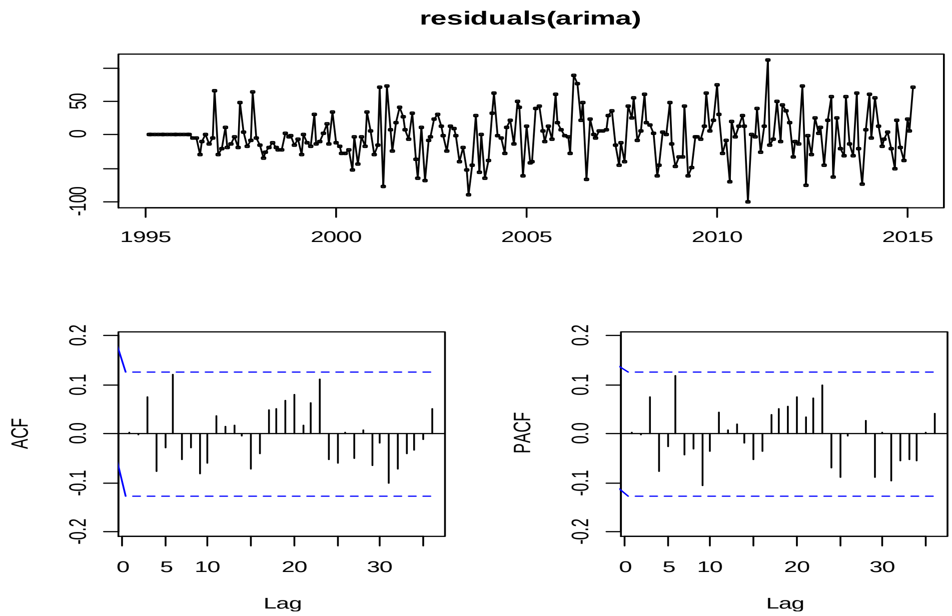


Figure 6. ACF and PACF Plots of the Residuals of ARIMA(1,1,1)(1,0,2)[12] Model with Drift

Table 15. ARCH-LM Test Results of ARIMA(1,1,1)(1,0,2)[12] Model with Drift

Chi-squared	p-Value
14.7563	**0.255**

After checking the normality assumption, now ARCH-LM (Autoregressive Conditional Heteroscedasticity-Lagrange Multiplier) test results are presented in Table 15 to find out if there is a heteroscedasticity problem. For this test, the null hypothesis says that there are no ARCH (Autoregressive Conditional Heteroscedasticity) effects (indicating to the constant variance). From ARCH-LM test results with the number of lags chosen as 12, it can be inferred that since p-value (0.255) exceeds 5% significance level, the null hypothesis of no ARCH effect (homoscedasticity) in the residuals of ARIMA(1,1,1)(1,0,2)[12] with drift model cannot be rejected and therefore concluding that the residuals of ARIMA(1,1,1)(1,0,2)[12] with drift model are homoscedastic (that is, the residuals have constant variance). Briefly, it can be said that all assumptions regarding diagnostic checking (no serial correlation, normality of residuals, constant variance) hold for this model.

Table 16. Forecast Accuracy Measures for ARIMA (1,1,1)(1,0,2)[12] Model with Drift

ME	RMSE	MAE	MPE	MAPE	MASE
-0.3333779	34.08106	25.34299	-49.62708	70.50063	0.73495

Note. ME: Mean Error; RMSE: Root Mean Squared Error; MAE: Mean Absolute Error; MPE: Mean Percentage Error; MAPE: Mean Absolute Percentage Error; MASE: Mean Absolute Scaled Error (For more information about the accuracy measures, see Ord & Fildes, 2013, chap. 2).

In Table 16, various forecast accuracy measures for ARIMA(1,1,1)(1,0,2)[12] with drift model that is chosen under the stepwise-selection method have been presented. Afterwards, these results will be compared to the model that will be chosen under the non-stepwise selection method.

Subsequent to applying (faster) stepwise-selection method which provides a short-cut for selecting the best-fitted model, now let us try the same thing under the (slower) non-stepwise selection method which searches for all possible models. In this case, the best choice under the nonstepwise-selection method has been determined to be ARIMA(1,1,1)(2,0,0)[12] with drift model for inflation series. The estimates of parameters of this new model are given in Table 17:

Table 17. Estimates of Parameters for ARIMA (1,1,1)(2,0,0)[12] Model with Drift

	AR(1)	MA(1)	SAR(1)	SAR(2)	DRIFT
Estimate	0.2202	-0.9273	0.2961	0.3136	-0.4393
Standard Error	0.0752	0.0336	0.0610	0.0633	0.5195
Sigma^2: 1270 log-likelihood: -1200.75 AIC: 2413.51 AICc: 2413.87 BIC: 2434.42					

As it is apparent in Table 17, the coefficients of *ARIMA (1,1,1)(2,0,0)[12] Model with Drift* are seen to be significant.

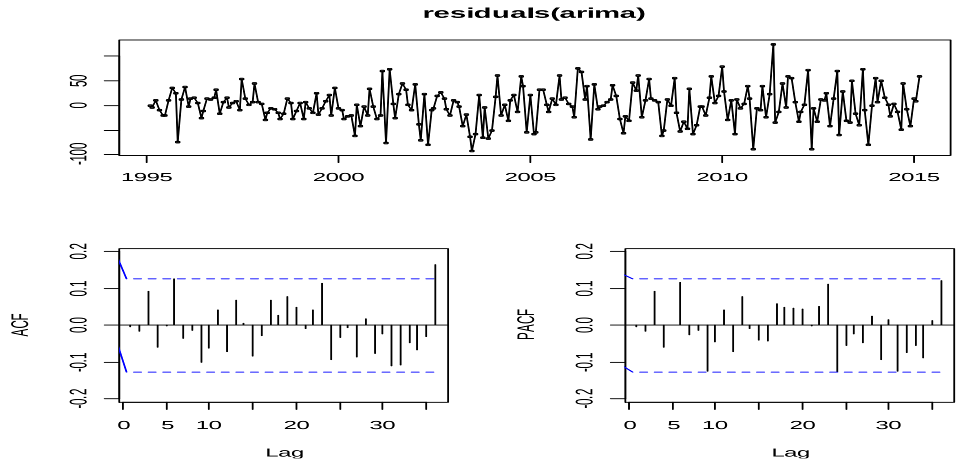


Figure 7. ACF and PACF Plots of the Residuals of ARIMA (1,1,1)(2,0,0)[12] Model with Drift

When looked at Figure 7, mean of the residuals of ARIMA(1,1,1)(2,0,0)[12] model with drift is seen to be distributed around zero. However, acf and pacf values are within the significance limits only up to 12 and 24 seasonal lags. Even though the absence of autocorrelation at seasonal lag 12 is sufficient to make a positive inference about no serially correlated residuals (since we are dealing with monthly inflation rates in which the length of seasonal period is 12), a spike is realized at 36th lag and therefore not all acf values are seen to take place within the significance limits because of this 36th lag. If ARIMA(1,1,1)(2,0,0)[12] model with drift is compared to ARIMA (1,1,1)(1,0,2)[12] model with drift that does not enable such a spike at 36th lag apart from other seasonal lags as observed in Figure 6, the latter (with stepwise-selection method) can be said to be a stronger model than the former (with non-stepwise selection method). Let us verify this with an examination on Box-Ljung test statistics at seasonal lags:

Table 18. Box-Ljung Test Results of ARIMA(1,1,1)(2,0,0)[12] Model with Drift at Seasonal Lags Based on the Non-stepwise Selection

Seasonal Lags	X-squared Statistics	p-value
12	12.6478	0.1246
24	25.7961	0.1727
36	46.7037	**0.04507**
48	58.2202	0.07392

Table 18 presents the autocorrelation check results for the residuals of ARIMA(1,1,1)(2,0,0)[12] with drift model at seasonal lags based on the non-stepwise selection. According to both the plot of ACF in Figure 7 and Table 18 results, no serial correlation has been detected except 36th lag with a probability value (p-value) of 0.04507 which is smaller than 5% significance level. Therefore p-values for Box-Ljung statistics at seasonal lags 12, 24, 48 are greater than 5% significance level indicating to the non-rejection of the null hypothesis of independently distributed residuals at these seasonal lags. Only 36th lag creates serially correlated residuals depending on the rejection of the null. Now let us check the normality of ARIMA(1,1,1)(2,0,0)[12] model with drift residuals:

Table 19. Jarque - Bera Normality Test Results of ARIMA(1,1,1)(2,0,0)[12] Model with Drift

X-squared Statistic	Asymptotic p-value
1.0074	**0.6043**

According to the Jarque-Bera test results given in Table 19, we fail to reject the null hypothesis saying that the residuals are normally distributed with an insignificant X-squared statistic having an asymptotic p-value of 0.6043 that is greater than 5% significance level.

Table 20. ARCH-LM Test Results of ARIMA(1,1,1)(2,0,0)[12] Model with Drift

Chi-squared	p-Value
15.6521	**0.2077**

From the ARCH-LM test results, it can be inferred that the null hypothesis of no ARCH effect in the residuals of ARIMA(1,1,1)(2,0,0)[12] model with drift cannot be rejected and hence the residuals of this model are said to be homoscedastic.

Briefly, all assumptions regarding normality of residuals, and constant variance hold for this model except autocorrelation check for 36th lag. Residuals of ARIMA(1,1,1)(2,0,0)[12] model with drift are independently distributed up to seasonal lags 12 and 24, however not independently distributed for seasonal lag 36.

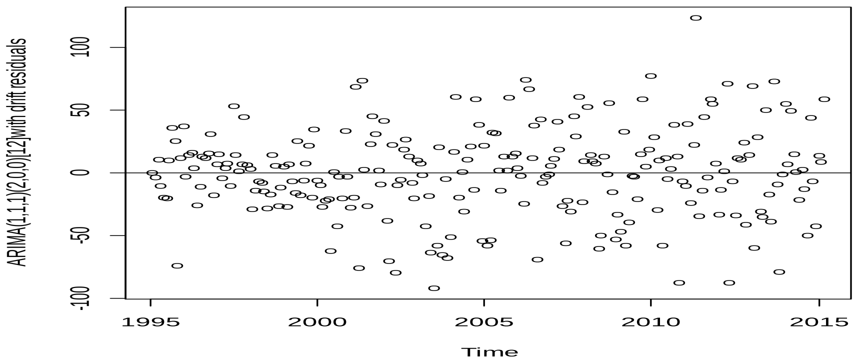


Figure 8. Plot of ARIMA (1,1,1)(2,0,0)[12] with Drift Residuals against Time

Table 21. Forecast Accuracy Measures for ARIMA (1,1,1)(2,0,0)[12] Model with Drift

ME	RMSE	MAE	MPE	MAPE	MASE
-0.3060675	35.49517	27.12352	-60.00625	80.78677	0.7865855

Note. (For more information about the accuracy measures, see Ord & Fildes, 2013, chap. 2.)

In Table 21, forecast accuracy measures for ARIMA(1,1,1)(2,0,0)[12] with drift model that is based on the non-stepwise selection method have been presented.

4. CONCLUSION

Now that we have identified two models based on both stepwise and non-stepwise selection, we can provide a summary of final results: In this application, ARIMA(1,1,1)(1,0,2)[12] with drift model chosen by using (faster) stepwise selection method and ARIMA(1,1,1)(2,0,0)[12] with drift model chosen by using (slower) non-stepwise selection which seeks for all possible models have been compared. Although we expect the latter model with non-stepwise selection to be better (since, stepwise selection offers short-cuts in selecting the best model), the results have shown that the former model with stepwise-selection is better as the best-fitted SARIMA model. A summary of the comparison of both models is given in Table 22:

Table 22. Comparison of ARIMA (1,1,1) (1,0,2) [12] with Drift and ARIMA (1,1,1) (2,0,0) [12] with Drift Models

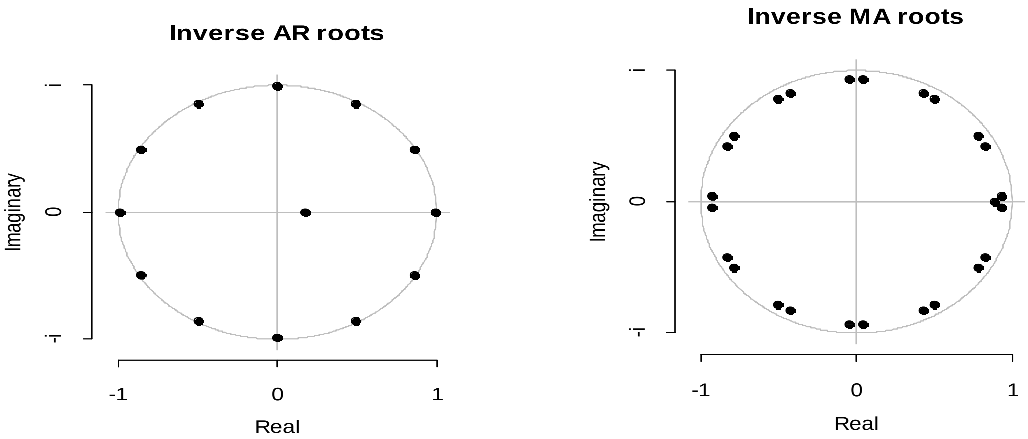
Model	Accuracy Measures	Significancy of Coefficients	AICc	Normality	ARCH-LM	ACF of Residuals (Autocorrelation check for residuals)
Model 1	RMSE: 34.08106 MAE: 25.34299 MAPE: 70.50063 MASE: 0.73495	All seasonal and non-seasonal AR and MA coefficients are significant.	2405.96	ok	ok	There is no spike (no autocorrelation at all seasonal lags 12,24,36,48.)
Model 2	RMSE: 35.49517 MAE: 27.12352 MAPE: 80.78677 MASE: 0.7865855	All seasonal and non-seasonal AR and MA coefficients are significant.	2413.87	ok	ok	There is a spike at 36^{th} lag (autocorrelation problem exists at 36^{th} lag).

Note. Model 1 represents ARIMA(1,1,1)(1,0,2)[12] with Drift. Model 2 represents ARIMA(1,1,1)(2,0,0)[12] with Drift.

As seen in Table 22, forecast accuracy measures of model 1 are smaller than the ones of model 2. In the light of given information, it is possible to say that model 1 satisfies all necessary assumptions (no serial correlation, constant variance and normality) and is better in all respects than model 2 with the smallest AICc, significant parameters, no spike at ACF etc. Therefore having satisfied all model assumptions, model 1 can be regarded as the best-fitted model for forecasting monthly inflation rates in Turkish economy.

For ARIMA(1,1,1)(1,0,2)[12] with drift (stepwise) model, apart from all required checks, we need to check also the causality, stationarity and invertibility condition. For ARIMA(1,1,1)(1,0,2)[12] with drift (stepwise) model to be causal, stationary and invertible, all roots of the characteristic polynomial of AR, MA, SAR and SMA operators should be greater than 1 in absolute value.

A causal invertible model should have all the roots outside the unit circle. Equivalently, the inverse roots should lie inside the unit circle (Hyndman, 2014). Here, all inverse roots lie inside the unit circle as shown in the figures given as follows



(Hyndman, 2015: 53): our ARIMA(1,1,1)(1,0,2)[12] with drift (stepwise) model can be said to satisfy causality, stationarity and invertibility conditions.

Figure 9. Plots of Inverse AR and MA Roots

REFERENCES

ABRAHAM, B., BOX, G. E. P. (1978), "Deterministic and Forecast-Adaptive Time-Dependent Models", *Applied Statistics*, 27(2), 120-130.

AIDOO, E. (2010), *Modelling and forecasting inflation rates in Ghana: An application of SARIMA models.* Master's Thesis, Högskolan Dalarna School of Technology & Business Studies, Sweden.

AKUFFO, B., AMPAW, E. M. (2013), "An Autoregressive Integrated Moving Average (ARIMA) Model for Ghana's Inflation (1985-2011)", *Mathematical Theory and Modelling*, 3(3), 10-26.

BEAULIEU, J. J., MIRON, J. A. (1992), *Seasonal Unit Roots in Aggregate U.S. Data* (NBER Technical Paper No. 126). Cambridge: National Bureau of Economic Research.

BEAULIEU, J. J., MIRON, J. A. (1993), "Seasonal Unit Roots in Aggregate U.S. Data", *Journal of Econometrics*, 55(1-2), 305-328.

BOX, G. E. P., JENKINS, G. M. (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

BOX, G. E. P., JENKINS, G. M. (1976), *Time Series Analysis: Forecasting and Control* (2nd ed.), Holden-Day, San Francisco.

BROCKWELL, P. J., DAVIS, R. A. (2002), *Introduction to Time Series and Forecasting* (2nd ed.), Springer-Verlag, New York.

BROCKWELL, P. J., DAVIS, R. A. (2006), *Time Series: Theory and Methods* (2nd ed.), Springer, New York.

CANOVA, F., HANSEN, B. E. (1995), "Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability", *Journal of Business and Economic Statistics,* 13(3), 237-252.

CHANG, Y. W., LIAO, M. Y. (2010), "A Seasonal ARIMA Model of Tourism Forecasting: The Case of Taiwan", *Asia Pacific Journal of Tourism Research,* 15(2), 215-221.

CHATFIELD, C. (1996), *The Analysis of Time Series: An Introduction* (5th ed.), Chapman & Hall/CRC, London, UK.

CHEN, R., SCHULZ, R., STEPHAN, S. (2003), "Multiplicative SARIMA Models", *Computer-Aided Introduction to Econometrics,* (Ed. J.R. Poo), Berlin: Springer-Verlag, 225-254.

COSAR, E. E. (2006), "Seasonal Behaviour of the Consumer Price Index of Turkey", *Applied Economics Letters,* 13(7), 449-455.

DIAZ-EMPARANZA, I., LOPEZ-de-LACALLE, J. (2006), “Testing for Unit Roots in Seasonal Time Series with R: The Uroot Package”, http://www.jalobe.com:8080/doc/uroot.pdf, (10.05.2015).

DICKEY, D., HASZA, D., FULLER, W. (1984), “Testing for Unit Roots in Seasonal Time Series”, *Journal of the American Statistical Association*, 79(386), 355-367.

FRANSES, P. H. (1991), *Model Selection and Seasonality in Time Series.* Doctoral dissertation, Erasmus University Rotterdam, Netherlands. Retrieved from http://hdl.handle.net/1765/2047.

FRANSES, P. H. (1998), “Modeling Seasonality in Economic Time Series”, *Handbook of Applied Economic Statistics,* (Eds. A. Ullah and D.E.A. Giles), New York: Marcel Dekker, 553-577.

FRANSES, P. H., HOBIJN, B. (1997), “Critical Values for Unit Root Tests in Seasonal Time Series”, *Journal of Applied Statistics,* 24(1), 25-48.

FRANSES, P. H., KOEHLER, A. B. (1998), “A Model Selection Strategy for Time Series with Increasing Seasonal Variation”, *International Journal of Forecasting,* 14(3), 405-414.

HAMAKER, E. L., DOLAN, C. V. (2009), “Idiographic Data Analysis: Quantitative Methods - from Simple to Advanced”, *Dynamic Process Methodology in the Social and Developmental Sciences,* (Eds. J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra and N. Chaudhary), New York: Springer-Verlag, 191-216.

HASZA, D. P., FULLER, W. A. (1982), “Testing for Nonstationary Parameter Specifications in Seasonal Time Series Models”, *The Annals of Statistics*, 10(4), 1209-1216.

HILLMER, S. C., BELL, W. R., TIAO, G. C. (1983), “Modeling Considerations in the Seasonal Adjustment of Economic Time Series”, *Applied Time Series Analysis of Economic Data*, (Ed. A. Zellner), Washington, DC: U.S. Bureau of the Census, 74-100.

HYLLEBERG, S., ENGLE, R., GRANGER, C., YOO, S. (1990), “Seasonal Integration and Cointegration”, *Journal of Econometrics*, 44(1), 215-238.

HYNDMAN, R., J. (2014), "Plotting the Characteristic Roots for ARIMA Models", http://robjhyndman.com/hyndsight/arma-roots/, (01.08.2015).

HYNDMAN, R., J. (2015, May), "Package 'Forecast'", http://cran.r-project.org/web/packages/forecast/forecast.pdf, (02.06.2015).

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P., SHIN, Y. (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?", *Journal of Econometrics*, 54(1-3), 159-178.

LIM, C., MCALEER, M. (2000), "A Seasonal Analysis of Asian Tourist Arrivals to Australia", *Applied Economics*, 32(4), 499-509.

MADDALA, G. S., KIM, I. M. (1998), *Unit Roots, Cointegration and Structural Change,* Cambridge University Press, Cambridge.

ORD, K., FILDES, R. (2013), *Principles of Business Forecasting,* South-Western, Cengage Learning.

OSBORN, D.R., CHUI, A. P. L., SMITH, J. P., BIRCHENHALL, C. R. (1988), "Seasonality and the Order of Integration for Consumption", *Oxford Bulletin of Economics and Statistics,* 50(4), 361-377.

PANKRATZ, A. (1983), *Forecasting with Univariate Box-Jenkins Model: Concepts and Cases,* John Wiley & Sons, New York.

PLATON, V. (2010), "Application of Seasonal Unit Roots Tests and Regime Switching Models to the Prices of Agricultural Products in Moscow 1884-1913", http://www.hse.ru/data/2010/10/22/1222675037/Seasonal%20unit%20roots%20and%20regime%20switch.pdf, (04.01.2015).

SANLI, S. (2015), *The Econometric Analysis of Seasonal Time Series: Applications on Some Macroeconomic Variables*, Master's Thesis, Cukurova University, Adana.

SAZ, G. (2011), "The Efficacy of SARIMA Models for Forecasting Inflation Rates in Developing Countries: The Case for Turkey", *International Research Journal of Finance and Economics,* 62, 111-142.

SHUMWAY, R. H., STOFFER, D. S. (2011), *Time Series Analysis and Its Applications - with R Examples* (3rd ed.), Springer, New York.

SØRENSEN, N. K. (2001), "Modelling the Seasonality of Hotel Nights in Denmark by County and Nationality", *Seasonality in Tourism,* (Eds. T. Baum and S. Lundtrop), Oxford: Elsevier, 75-88.

TAM, W. K., REINSEL, G. C. (1997), "Tests for Seasonal Moving Average Unit Root in ARIMA Models", *Journal of the American Statistical Association,* 92(438), 725-738.

ZHANG, Q. (2008), *Seasonal Unit Root Tests: A Comparison.* Doctoral Dissertation. North Carolina State University, Raleigh.