Hacettepe Journal of Mathematics and Statistics
Volume 43 (3) (2014), 543 – 552

Cases of residual types in diagnostic checking for ARMA model

Mehmet Guray UNSAL*† and Resat KASAP‡

Abstract

In this study, the residuals in time series analysis, which were classified in four different classes as "conditional residuals", "unconditional residuals", "innovation" and "normalized residuals", are calculated by a simulation study for the ARMA model under certain parameter values for different numbers of observation and their conditions in diagnostic checking are examined using the test statistic which belongs to Ljung Box.

Keywords: ARMA Models, Diagnostic Checking, Residual Types, Time Series.

1. Introduction

Box and Jenkins [2] worked on building and forecasting time series models and found out the method which is called Box-Jenkins Modelling Process in time series analysis. This approach has been improved by many studies for approximately forty years. Lutkepohl [8], Box et al. [3], Brockwell and Davis [4] and Wei [12] have popular text books about this subject. One of the most important points in the process of analyzing in time series is diagnostic checking. It can be determined which Box-Jenkins model is suitable for time series data in two ways. The first option is the examining of ACF and PACF plots of the time series. The second option tests the (H_0) null hypothesis where the autocorrelations of residuals are equal to zero in lag m. [1]. Thus, residual values play a significant role in diagnostic checking in time series models. Knowing the general structure of the residuals used in diagnostic checking is essential in order to apply this idea in practical applications effectively [9].

There are very little amount of study has been done about the residuals in literature. The residuals in time series analysis are classified in four different classes as, "conditional residuals", "unconditional residuals", "innovations" and "normalized residuals" in Mauricio [9]. Mauricio showed that when the considered model

*Gazi University, Faculty of Science, Statistics Department, Ankara, Turkey, Email: mgunsal@gazi.edu.tr
†Corresponding Author.
‡Gazi University, Faculty of Science, Statistics Department, Ankara, Turkey, Email: rkasap@gazi.edu.tr

contains moving average roots near the unit circle and the number of observation is small, the unconditional residuals and normalized residuals give different decisions in diagnostic checking. In this study, the residuals in time series analysis, which were classified in 4 different classes as "conditional residuals", "unconditional residuals", "innovation" and "normalized residuals", are calculated by a simulation study for the $ARMA(1,1)$ model under certain parameter values for different numbers of observation and their conditions in diagnostic checking are examined using the test statistic which belongs to Ljung-Box.

The rest of this article is organized as follows. In Section 2, a description of the $ARMA(p,q)$ models used for modeling the time series and the Ljung Box test statistic used for diagnostic checking for the time series models are presented. Section 3 gives the details of the residual types adopted by Mauricio. A simulation study results are given for different lags, number of observations and parameter values for $ARMA(1,1)$ model and the different decisions in diagnostic checking are given for the resudial types in the realization study in Section 4. Finally, Section 5 concludes this research.

2. AUTOREGRESSIVE AND MOVING AVERAGE ($ARMA$) MODEL AND MODEL DIAGNOSTIG CHECKING

Autoregressive and moving average ($ARMA$) model is used in modeling stationary time series. The mentioned model is a combination of AR and MA models. This model is expressed with a certain number of values preceding time and residual series of data. If the model is made up of a combination of AR and MA model with p and q degrees respectively, $ARMA$ model is symbolized as $ARMA(p,q)$ and theoretically written as $\phi(B)\widetilde{W}_t = \theta(b)A_t$, where,

$$\phi(B) = 1 - \phi_1(B) - \phi_2(B)^2 - ... - \phi_p(B)^p$$

$$\theta(B) = 1 - \theta_1(B) - \theta_2(B)^2 - ... - \theta_p(B)^p$$

Invertibility and stationarity conditions for the model are, the roots of $\phi(B) = 0$ and $\theta(B) = 0$ be in the unit circle. $ARMA(1,1)$ model is written theoretically as,

$$\widetilde{W} = \phi_1\widetilde{W} + A_t - \theta_1 A_{t-1} \quad or \quad (1-\phi_1 B)\widetilde{W}_t = (1-\theta_1 B)A_t$$

Here, the invertibility and stationarity conditions are possible in $-1 < \phi_1 < 1$ and $-1 < \theta_1 < 1$ (Box et al., [3], Wei citeref:Wei). As mentioned above one of the approaches for diagnostic checking is using of test statistic which is based on residuals' sample autocorrelation functions. The null and alternative hypothesis for the test statistics as below:

$$\begin{cases} H_0: & Model\ is\ appropriate\ (\rho_1 = \rho_2 = ... = \rho_k = 0) \\ H_1: & Model\ is\ inappropriate \end{cases}$$

(ρ_i: Autocorrelation coefficients, $i = 1, 2, \ldots, k$).

In 1970, Box-Pierce derived the following Q_{BP} statistic having chi square distribution with approximate m degree of freedom for diagnostic checking citeref:BoxP.

$$Q_{BP} = n\sum_{k=1}^{m} r_k^2,\ k = 1, ..., m \tag{2.1}$$

Where n, number of observations, r_k, residuals' sample autocorrelation coefficients and m is number of lags. Ljung and Box [6] suggested to use the Q_{LB} statistic as below,

$$(2.2) \qquad Q_{LB} = n(n+2)\sum_{k=1}^{m} r_k^2/(n-k)$$

The autocorrelation coefficient of the residuals, which is also used for deriving the Ljung-Box test statistic expressed in the equation 2.2 above is used.

$$(2.3) \qquad r_k = \sum_{t=k+1}^{n} u_t u_{t-k} / \sum_{t=1}^{n} u_t^2 \quad , \ k = 1, \ldots, m, \ t = 1, \ldots, n$$

Here, $r = (r_1, r_2, \ldots, r_m)$ has multivariate normal distribution with zero average, $V(r_k) = (n-k)/n(n+2)$ and $Kov(r_k, r_1) = 0 \ \ (k \neq 1)$, so autocorrelation coefficient of the residuals has chi-square distribution with m degree of freedom for large n values [2]. Since the distribution of residual autocorrelations is $N(0, n^{-1}I_m)$, then distributions of both Ljung-Box's Q_{LB} and Box-Pierce's Q_{BP} are chi-square distributions with m degree of freedom and their expected values are m and their variances are $2m$. The expected value of Q_{LB} is m for finite numbers of n values, whereas,

$$(2.4) \qquad E(Q_{BP}) = \frac{nm}{n+2}\left(1 - \frac{m+1}{2n}\right)$$

As long as n be less than m, $E(Q_{BP})$ will be less than m. Moreover, for n values which are greater than m, variances of Q_{BP} and Q_{LB} are written as follows,

$$V(Q_{BP}) = 2m(1 + (m-10)/n), \quad V(Q_{LB}) = 2m(1 + (m-10)/n)$$

Here, the variance of Q_{LB} exceeds $2m$ value, however Monte Carlo simulation method has shown that its distribution is much closer than the distribution of Q_{BP} to chi-square with m degree of freedom [2], [6]. The test statistics, which are used for diagnostic checking, are called Portmanteau test statistics. There are a lot of Portmanteau test statistics in the literature and detailed information about structures and distributions of other Portmanteau test statistics could be found in many studies [2], [6], [7] [5], [1].

3. DEFINING AND CALCULATING OF RESIDUAL TYPES FOR $ARMA(p, q)$

Let see stationary time series process $\{W_t\}$ following the model and let $w = [w_1, w_2, ..., w_n]'$ generated by $\{W_t\}$.The theoretical representation of ARMA (Autoregressive - Moving Average) model is given below:

$$(3.1) \qquad \phi(B)\widetilde{W}_t = \theta(B)A_t$$

Here, $\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$ are polynomials with degrees of p and q, also B is a lag operator,$\widetilde{W}_t = W_t - E[W_t]$ and $\{A_t\}$ is a white noise process with $\sigma^2 > 0$. Regarding the model (3.1) , $\widetilde{W} =$

$[\widetilde{W_1}, ..., \widetilde{W_n}]'$, $A = [A_1, ..., A_n]'$ and $U_* = [\widetilde{W}_{1-p}, ..., \widetilde{W_0}, A_{1-q}, ..., A_0]'$, observed time series $w = [w_1, w_2, ..., w_n]'$ can be seen as a particular realization of a random vector $W = [W_1, W_2, ..., W_n]'$ following the model,

$$D_\phi \widetilde{W} = D_\theta A + VU_* \tag{3.2}$$

Where D_ϕ and D_θ are $n \times n$ parameter matrices with ones as diagonal elements and $-\phi_j$ and $-\theta_j$ as elements that constitute the jth subdiagonal, respectively, and V is a $n \times (p+q)$ matrix with $V_{ij} = \phi_{p+i-j}(i = 1, ..., p; j = 1, ..., p)$ and $V_{ij} = -\theta_{q+i-j+p}(i = 1, ..., q; j = p+i, ..., p+q)$, where the remaining elements are zero [9].

Let us assume that the theoretical autocovariance matrix is $\sum_w = \sigma^{-2}E[\widetilde{W}\widetilde{W}']$, and $\hat{\sum}_w$ is an estimation of $\sum_w$. The autocovariance matrix can be given as follows from the equation (3.2) [9];

$$\sum_w = D_\phi^{-1}(D_\phi D_\theta' + V\Omega V')(D_\phi^{-1})' = K^{-1}(I + Z\Omega Z)(K^{-1})' \tag{3.3}$$

In equation (3.3) , $K = D_\theta^{-1}D_\phi, Z = -D_\theta^{-1}V$ and $\Omega = \sigma^{-2}E[U_*U_*']$ are parameter matrices of dimensions $n \times n$, $n \times (p+q)$ and $(p+q) \times (p+q)$, respectively, with Ω being readily expressible in terms of $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$, for example, in Ljung and Box [7]. Besides, here $\sum_0 = I + Z\Omega Z' = [I - Z(\Omega^{-1} + Z'Z)^{-1}Z']^{-1}$.

Using the relation (3.3), $\widetilde{w}' \hat{\sum}_w^{-1} \widetilde{w}$ can be written as,

$$\widetilde{w}' \hat{\sum_w}^{-1} \widetilde{w} = \widetilde{w}' \hat{K}' (I + \hat{Z}\hat{\Omega}\hat{Z})\hat{K}\widetilde{w} \tag{3.4}$$

$\hat{K}, \hat{Z}$ and $\hat{\Omega}$ symbolize the estimations of parameter matrices defined in equation (3.3). According to these theoretical information, residuals have been grouped by Mauricio [9] in 4 different classes as below:

3.1. Conditional Residuals. Conditional residuals are associated with relation (3.4) and defined as the elements of the $n \times 1$ vector $\hat{a}_0 = \hat{K}\tilde{w}$.

3.2. Unconditional Residuals. Unconditional residuals are associated with (3.4) and defined as the elements of the $n \times 1$ vector $\hat{a} = (I + \hat{Z}\hat{\Omega}\hat{Z}'^{-1}\hat{K}\tilde{w} = \hat{\sum}_0^{-1}\hat{a}_0$.

3.3. Innovations. Innovation residuals are associated with (3.4) and defined as the elements of the $n \times 1$ vector $\hat{e} = \hat{L}^{-1}\tilde{w} = (\hat{K}\hat{L})^{-1}\hat{a}_0$. Here, $\hat{L}$ is the estimation of the $n \times n$ unit lower- triangle matrix L in below factorization,

$\sum_w = LFL'$, or,

$$\sum_w = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ L_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} F_1 & 0 & \cdots & 0 \\ 0 & F_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_n \end{pmatrix} \begin{pmatrix} 1 & L_{12} & \cdots & L_{1n} \\ 0 & 1 & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

where $F_t > 0$, and $t = 1, 2, ..., n$.

From relation (3.4), the below equations will be result [9]:

$$\tilde{w}'\hat{\sum}_w^{-1}\tilde{w} = \hat{a}_0'\hat{\sum}_0^{-1}\hat{a}_0 = \hat{e}'\hat{F}^{-1}\hat{e}$$

where $\hat{F}$, is the estimation of matrix F.

3.4. Normalized Residuals. If we define lower-triangle matrix P as, $\sum_0 = I + Z\Omega Z' = PP'$, the definition of vector $\hat{v}$(normalized residuals) is,

$$\hat{v} = \hat{P}^{-1}\hat{a}_0 = \hat{P}'\hat{a} = \hat{F}^{-\frac{1}{2}}\hat{e}$$

where $\hat{P}$ is the estimation of matrix P [9], [10], [11].

4. SIMULATION, REALIZATION AND REAL WORLD DATA APPLICATION STUDIES

4.1. Simulation Study.

In this part, conditions of the residual types in diagnostic checking are compared by using Ljung-Box test statistic for $ARMA(1,1)$ model. Here, the main aim is not to test whether the model is appropriate or not, but it is to compare whether the types of residuals lead us to the same results for diagnostic checking or not. Data are generated under assumption of $ARMA(1,1)$ models which have autoregressive and moving average parameter values with $\phi = 0.5$ *and* $\theta = 0.6$, $\phi = 0.9$ *and* $\theta = 0.6$, $\phi = 0.1$ *and* $\theta = 0.9$, $\phi = 0.2$ *and* $\theta = 0.9$. By considering these four different parameter conditions, the simulation study has been done and types of residuals are calculated for all generated time series. In this simulation study, to see the rejection ratio of the null hypothesis "H_0: Model is appropriate", trials are repeated 100 times and total number of rejections are divided by 100. Furthermore, to see interaction between number of observation and rejection ratio, n=10, n=25, n=50, n=100, n=250 and n=500 observation numbers are taken and result are in Figure 1-4. Lags(m) are taken as 2,3,5,10.

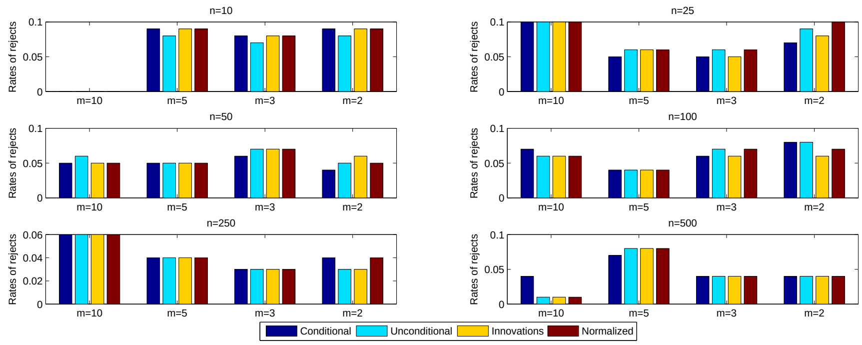

Figure 1. H_0 rejection ratio for $\phi = 0.5$ and $\theta = 0.6$,

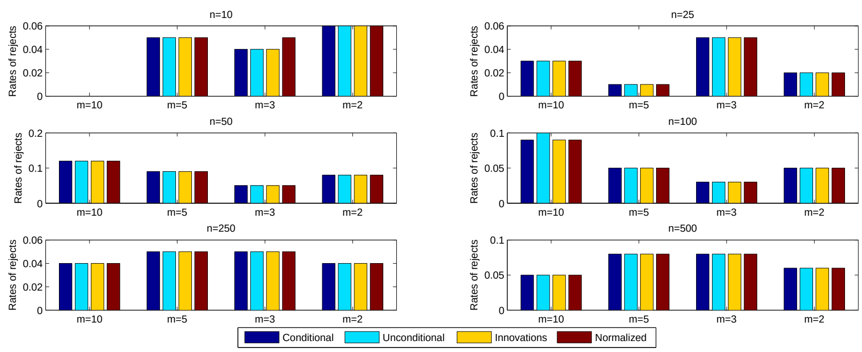

Figure 2. H_0 rejection ratio for $\phi = 0.9$ and $\theta = 0.6$

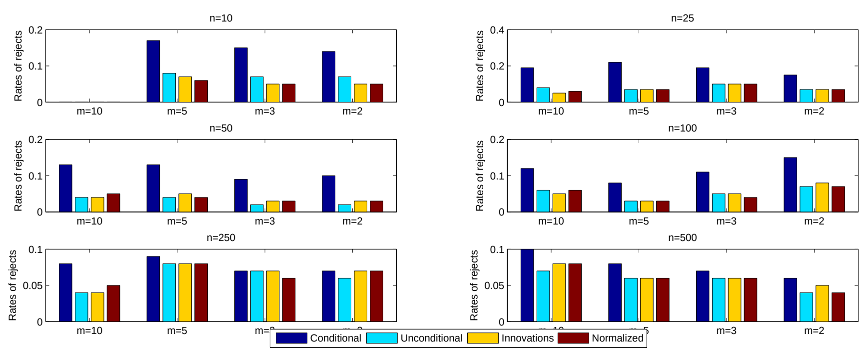

Figure 3. H_0 rejection ratio for $\phi = 0.1$ and $\theta = 0.9$

As seen in Figure 1, when autoregressive and moving average parameter values are $\phi = 0.5$ and $\theta = 0.6$, there is no difference between rejection ratio in four residual types with different number of observations and lags. Also in Figure 2, with $\phi = 0.9$ and $\theta = 0.6$, the same sitiation is valid. In Figure 3, $\phi = 0.1$ and

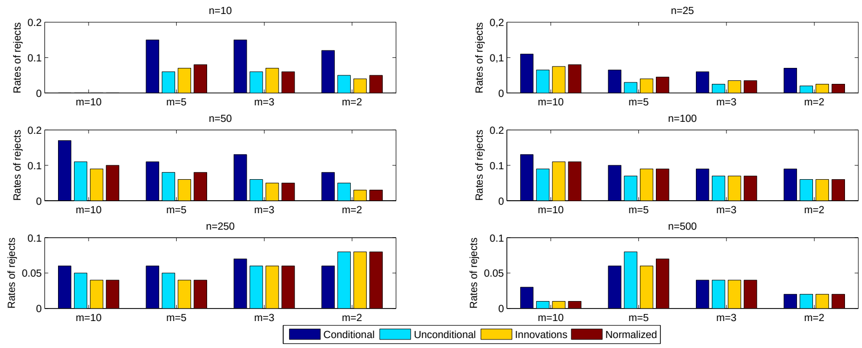

Figure 4. H_0 rejection ratio for $\phi = 0.2$ and $\theta = 0.9$

$\theta = 0.9$, the rejection ratio of conditional residuals is higher than the other types of residuals. And when the number of observation increases, the rejection ratios of four different types of residuals, are close to each other for all of the lags. In Figure 4, $\phi = 0.2$ and $\theta = 0.9$, the same situations as Figure 3 appear, rejection ratios of the conditional residuals have high values, rather to other residual types. Also the differences of the rejection ratios vanish when the number of observation increases.

Similar to Mauricio [9], as it can be seen from the bar charts (Figure 1-4), when the number of observations is small and $ARMA(1,1)$ model is close to non invertibility, conditional residuals have higher rejection ratios than the other three residual types. As mentioned above, these differences vanish when the number of observation increases. And when the number of observation values are large, the ratios of rejection, which are calculated from the different types of residual, are close to each other for all of the lags and parameter values. This situation is also lead us to use large number of observations in the modeling process.

For the negative values of and , the same situations are valid in diagnostic checking, because the differences of calculated residual types are related with non-invertible MA parameter value, so non-invertibility condition is provided not only MA parameter close to 1, but also close to -1. The invertibility condition is satistified when $-1 < \theta_1 < 1$.

Simulation results are consistent with the findings of Mauricio [9].

4.2. Realization Study.

The realization data is produced under the assumptation of suitable model, we consider the suitable model is $ARMA(1,1)$ model which have autoregressive and moving average parameter values with $\phi = 0.1$ and $\theta = 0.9$ in data production process. The number of observations is taken as 25. The realization data and calculated values of residual types are in Table 1.

Unit Number	Residual Types				
	Data	Conditional	Unconditional	Innovations	Normalized
1	2.6795	2.6795	0.9476	2.6795	2.0882
2	-1.3419	0.8017	-0.757	-0.1452	-0.1264
3	0.3349	1.1906	-0.2122	0.3699	0.3384
4	-0.3621	0.6768	-0.5866	-0.1171	-0.11
5	1.0322	1.6768	0.5405	0.9754	0.9322
6	1.0101	2.4159	1.3933	1.7087	1.6518
7	-2.0718	0.0014	-0.9189	-0.7357	-0.717
8	1.9296	2.1381	1.3098	1.5079	1.4781
9	-0.4807	1.2506	0.5051	0.6303	0.6205
10	0.4625	1.6361	0.9652	1.0604	1.0474
11	-0.6746	0.7517	0.1479	0.2103	0.2083
12	1.0333	1.7772	1.2338	1.2863	1.2764
13	-0.3375	1.1587	0.6696	0.699	0.6947
14	1.6726	2.7492	2.309	2.3277	2.3162
15	-2.673	-0.366	-0.7622	-0.766	-0.7629
16	0.6017	0.5396	0.1831	0.1851	0.1845
17	-2.2197	-1.7942	-2.1151	-2.1144	-2.1089
18	1.2676	-0.1253	-0.4141	-0.4036	-0.4027
19	0.7519	0.5124	0.2525	0.2634	0.2629
20	-0.1922	0.1938	-0.0401	-0.0311	-0.031
21	0.8995	1.0931	0.8825	0.8908	0.8898
22	0.9274	1.8212	1.6317	1.6373	1.6359
23	-1.8327	-0.2864	-0.4569	-0.4544	-0.4541
24	2.3883	2.3139	2.1604	2.1632	2.162
25	-1.4132	0.4305	0.2923	0.2926	0.2925

Table1. Realization Data and Calculated Residual Values for Residual Types

Lag	Residual Types			
	Conditional	Unconditional	Innovations	Normalized
m=10	43.42	9.55	8.88	9.87
m=5	25.89	3.63	4.2	4.47
m=3	18.33	3.38	3.71	4.17
m=2	14.31	3.33	3.66	4.11

Table2. Calculated Box-Jenkins Test Statistic Values for Different Residual Types and Lags

Lag	Residual Types			
	Conditional	Unconditional	Innovations	Normalized
m=10	Rejected	Not Rejected	Not Rejected	Not Rejected
m=5	Rejected	Not Rejected	Not Rejected	Not Rejected
m=3	Rejected	Not Rejected	Not Rejected	Not Rejected
m=2	Rejected	Not Rejected	Not Rejected	Not Rejected

Table 3. Decision Cases of H_0 Hypothesis For Different Residual Types and Lags

In Table 2, there are the calculated chi square values for different types of residual and different lags. The calculated chi square values from the conditional residuals give significant chi square test statistic value. This situation leads us to reject the "H_0: Model is appropriate" hypothesis when the conditional residuals are considered in diagnostic checking contrary to the other residuals as seen in Table 3.

4.3. Real World Data Application. In real world data application study, we used Gross Domestic Product data of Turkey (between 1998,Q1 and 2010,Q2), and considered that the estimated autocorrelation and moving average parameters are $\phi = 0.9$ and $\theta = 0.5$, respectively.

This identification yields the approximate model of order (1,1)

$$(1 - 0.9B)\widetilde{w}_t = (1 - 0.5B)a_t$$

The results of diagnostic checking (approximately values of test statistic and decision cases) could be found in Table 4-5.

Lag	Residual Types			
	Conditional	Unconditional	Innovations	Normalized
m=10	230	290	240	244
m=5	130	160	138	145
m=3	80	90	86	87
m=2	55	60	57	58

Table 4. Calculated Box-Jenkins Test Statistic Values for Different Residual Types and Lags

Lag	Residual Types			
	Conditional	Unconditional	Innovations	Normalized
m=10	Rejected	Rejected	Rejected	Rejected
m=5	Rejected	Rejected	Rejected	Rejected
m=3	Rejected	Rejected	Rejected	Rejected
m=2	Rejected	Rejected	Rejected	Rejected

Table 5. Decision Cases of H_0 Hypothesis For Different Residual Types and Lags

According the results in Table 4-5, it could be seen that the residuals lead us to have same decision in diagnostic checking as in the real word application study.

5. CONCLUSION

According to simulation study, especially when the number of observations are small and ARMA model's parameters are close to non invertibility, the differences between the calculated chi square test statistic values are becoming significant, and the ratios of rejection differ. The values of test statistics which are calculated from conditional residuals are bigger than the other three types of residuals. For ARMA(1,1) models which have parameter values close to non invertibility situations, using of conditional residuals give big ratio of the rejection than the other residuals types in diagnostic checking, but as the number of observation n increases, rates of rejections between the different types of residuals are in tendency to decrease. When the trials are done with the other parameter values which are not close to non invertible or stationary situation, four types of residuals give similar rates of rejection in the simulation studies. The moving average parameter in realization study is close to non-invertibility situation, thus "H_0: Model is appropriate" hypothesis is rejected, when the conditional residuals are considered in diagnostic checking contrary to the other residuals as seen in Table 3. On the contrary, there is no difference between four residual types according to diognostic checking results as in the real world data. As it can be seen from the bar charts (Figure 1-4), when the number of observations is small and $ARMA(1,1)$ model is close to non invertibility, conditional residuals have higher rejection ratios than the other three residual types.

References

[1] Arranz, M.A., Portmanteau Test Statistics in Time Series. Time-Oriented Language, 25-32, 2005.

[2] Box, G.E.P., Pierce, D.A., Distribution of the residual autocorrelations in autoregressive integrated moving average time series. Journal of the American Statistical Association, **65**, 1509-1526, 1976.

[3] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Time Series Analysis.Forecasting and Control. (Prentice-Hall, Englewood Cliffs, NJ, 1994).

[4] Brockwell, P.J., Davis, R.A., Introduction to Time Series and Forecasting. (Springer, NewYork, 2002).

[5] Davies, N., Triggs, C. M., Newbold, P., Significance levels of the Box-Pierce "Portmanteau statistic in finite samples. Biometrika, **64**, 517-522, 1977.

[6] Ljung, G.M. and Box, G.E.P., On a measure of lack of fit in time series models. Biometrika, **65**, 297-303, 1978.

[7] Ljung, G.M. and Box, G.E.P., The likelihood function of stationary autoregressive moving average models. Biometrika, **66**, 265-270, 1979.

[8] Lutkepohl, H., Introduction to Multiple Time Series Analysis. (Springer-Verlag, 1991).

[9] Mauricio, J.A., Computing and Using Residuals in Time Series Models. Computational Statistics-Data Analysis, **52**, 1746-1763, 2008.

[10] Unsal, M.G., Kasap, R., Residual Types in Time Series and Their Applications. Gazi University Journal of Science, **25**(2), 409-416, 2012.

[11] Unsal, M.G., Kasap, R., Residual Types in Time Series Analysis. (Lambert Academic Publishing, 2013).

[12] Wei, W.W.S., Time Series Analysis, Univariate and Multivariate Methods. (Pearson Education, 2006).