مجلّه‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **16** (2019): 409–446

دوره‌ی ۱۶، شماره‌ی ۲، پاییز و زمستان ۱۳۹۸، صص ۴۰۹–۴۴۶

DOI: 10.52547/jsri.16.2.409

DOR: 20.1001.1.25385771.2020.16.2.5.5

Best Linear Predictors in a Stationary Second Order Autoregressive process by means of near and far observations

Mohammad Mahdi Saber

Higher Education Center of Eghlid

Received: 2020/27/10 Approved: 2021/01/05

Abstract. In this paper, some predictors for prediction in a stationary second order autoregressive process are introduced. The paper attempts to find the best predictor for some cases such as circumstances there exist a fixed number of observations near or far from desired time. Pitman's measure of closeness and mean square error of prediction are used in order to comparison these predictors. The Gaussian and Gamma distributions have been used for distribution of errors. Finally analysis of two real data sets has also been presented for illustrative purposes.

Keywords. $AR(2)$ model; prediction performance; Pitman's measure of closeness.

MSC 2010: 62M10, 62M20, 62P12.

1 Introduction

Comparison between two linear predictors in a stationary autoregressive of order one ($AR(1)$) model has been performed in Hamaz and Ibazizen (2009). This work was done with respect to the well-known criterion of Pitman's Measure of Closeness (PMC). These predictors were based on previous observations and all observations, respectively. A more realistic case, when the parameters are unknown has been considered by Saadatmand et al (2016) for this model with Exponential innovations. For corresponding works in spatial domain, Saber and Nematollahi (2017) and Saber (2017) have studied on a stationary first order multiplicative spatial autoregressive model. In this work, we consider the stationary autoregressive of order two ($AR(2)$) model which is in form of

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \varepsilon_t \; t = 1, \ldots, n, \tag{1}$$

where ε_t s are uncorrelated random variables named as innovations. This process is stationary for parameters φ_1 and φ_2 which are in following set

$$\Theta_{\,stationary} = \{(\varphi_1, \varphi_2)\,;\; |\varphi_2| < 1 \text{ and } |\varphi_1| < 1 - \varphi_2\}\,. \tag{2}$$

A direct computation demonstrate that the autocorrelation function of process (1) is given by

$$\rho(h) = \begin{cases} \frac{\varphi_1}{1-\varphi_2} & h = 1 \\ \frac{\varphi_1^2+\varphi_2-\varphi_2^2}{1-\varphi_2} & h = 2 \\ \varphi_1 \,\rho(h-1) + \varphi_2 \,\rho(h-2) & h = 3, 4, \ldots, n, \end{cases} \tag{3}$$

Because of simplicity in notations, whenever we use ρ_h this means $\rho(h)$. The aim of paper is comparing predictors in two classes of predictors based on PMC and Mean Square Error of Prediction (MSP). These two classes of predictors are based on near observations and far observations to unsampled time which we need to predict our variable, respectively. For distribution of innovations we use from two well-known statistical distributions. The Gaussian distribution as a very general and applicable distribution is a suitable distribution for innovations which is our first choice. This distribution is applied for modeling symmetric data which their value can be in whole of real data ($\mathbb{R}$). *Hamaz and Ibazizen* (2009) had considered this distribution in their work, too. Another distribution which recently has been frequently used for distribution of errors in a time series is Gamma distribution whose a special case is Exponential distribution (*Ailliot* (2006), *Gourieroux and Jasiak* (2006), *Saadatmand et al.* (2016), *Senoglu and Bayrak* (2016)). Whenever time series data are positive and skewed to right, an appropriate choice for

distribution of innovations is Gamma distribution. Therefore, we have a motivation in order to use a Gamma distribution for distribution of innovations, too. By using of these two distribution which one of them is symmetric and in $\mathbb{R}$ and another is positive and skewed to right, a large class of data can be contained in this study.
The paper is organized as follows. Comparing these predictors which use from near observations is studied in Section 2. We peruse three classes of predictors based on far observations in Section 3. A real data set is analyzed for illustrative purposes in Section 4. Finally, a discussion and conclusion about matter has been included in Section 5.

2 Prediction of Missing Values by Using Near Observations

In this section, we use observations which are near to time which prediction must be done there. Notice that if all the past observations are used for prediction, the predictor equals with $\hat{Z}_t^{2,1} = \mathrm{E}\left(Z_t \mid Z_1 , \ldots , Z_{t-1}\right) = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \mathrm{E}(\varepsilon_t)$.
On the other hand, we are going to perform prediction by the near observations. Therefore, only 4 observations which are in distance 2 with unobserved time are used. First we find $15 = (2^4 - 1)$ linear predictors for Z_t in model (1). These predictors are different from using other observations point of view. Our predictors $\hat{Z}_t^{k,l}$ use observation $\boldsymbol{Z}_t^{k,l} = \{Z_s ; s \in \boldsymbol{O}_t^{k,l}\}$ for $k = 1, \ldots, 4$ and $l = 1, \ldots, \begin{pmatrix} 4 \\ k \end{pmatrix}$. These sets have described in following.

$O_t^{1,1} = \{t-2\}$; $O_t^{1,2} = \{t-1\}$; $O_t^{1,3} = \{t+1\}$; $O_t^{1,4} = \{t+2\}$. $O_t^{2,1} = \{t-2, t-1\}$; $O_t^{2,2} = \{t-2, t+1\}$; $O_t^{2,3} = \{t-2, t+2\}$; $O_t^{2,4} = \{t-1, t+1\}$; $O_t^{2,5} = \{t-1, t+2\}$; $O_t^{2,6} = \{t+1, t+2\}$. $O_t^{3,1} = \{t-2, t-1, t+1\}$; $O_t^{3,2} = \{t-2, t-1, t+2\}$; $O_t^{3,3} = \{t-2, t+1, t+2\}$; $O_t^{3,4} = \{t-1, t+1, t+2\}$ and $O_t^{4,1} = \{t-2, t-1, t+1, t+2\}$.
Since all indexes must be in set $\{1, \ldots, n\}$, there are some exceptions which $O_t^{k,l}$s are not same as defined set in above and in fact these sets are reduced for observations which are in borders. For instance, $O_1^{4,1} = \{2, 3\}$ and $O_n^{3,4} = \{n-1\}$. So $\hat{Z}_t^{k,l}s$ have more than one part based on $O_t^{k,l}$.
Before of any work notice that identifiability of two recommended models must to be checked. Since this matter is not directly concerned to our study, we show identifiability of these models in Appendix A.

Although a comparison can be done between all predictors, this is not more logical to compare two predictors which use of one and four observations, respectively. One the other hand, interpretation of comparison among 15 predictors is very complicated and difficult. So, we focus on comparing $\hat{Z}_t^{k,l}$ for a fixed value of $k = 1, \ldots, 3$ and between $l = 1, \ldots, \begin{pmatrix} 4 \\ k \end{pmatrix}$. Because of this fact that for $k = 4$, there exist only a predictor, we cannot have any comparison in this case. Therefore, we compare $\hat{Z}_t^{4,1}$ with the best predictors $\hat{Z}_t^{k,l}$ in cases $k = 1, \ldots, 3$. The introduced predictors are computed by conditional expectation Z_t on $\boldsymbol{Z}_t^{k,l}$, $\hat{Z}_t^{k,l} = E\left(Z_t \;\middle|\; \boldsymbol{Z}_t^{k,l}\right)$. Computation of all predictors is accomplished by the well-known Projection Theorem, see e.g. *Brockwell and Davis* (1991). In following theorem, we find mentioned predictors.

***Theorem* 1.** *Let Z_t comes from model* (1) *and $O_t^{k,l}$ are defined as above. If $E\,(Z_t\,) = 0$ and $Var\,(Z_t\,) < \infty$, then the best linear predictors for Z_t based on $\boldsymbol{Z}_t^{k,l}$ w.r.t MSP, are given by*

$$\hat{Z}_t^{1,1} = \begin{cases} \rho_2\ Z_{t+2} & t = 1, 2 \\ \\ \rho_2\ Z_{t-2} & t = 3, \ldots, n \end{cases}$$

$$\hat{Z}_t^{1,2} = \begin{cases} \rho_1\ Z_{t+1} & t = 1 \\ \\ \rho_1\ Z_{t-1} & t = 2, \ldots, n \end{cases}$$

$$\hat{Z}_t^{1,3} = \begin{cases} \rho_1\ Z_{t+1} & t = 1, \ldots, n-1 \\ \\ \rho_1\ Z_{t-1} & t = n \end{cases}$$

$$\hat{Z}_t^{1,4} = \begin{cases} \rho_2\ Z_{t+2} & t = 1, \ldots, n-2 \\ \\ \rho_2\ Z_{t-2} & t = n-1, n \end{cases}$$

$$\hat{Z}_t^{2,1} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1, 2 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = 3, \ldots, n \end{cases}$$

$$\hat{Z}_t^{2,2} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1, 2 \\ \\ \frac{(\rho_2 - \rho_1 \rho_3) Z_{t-2} \ + (\rho_1 - \rho_2 \rho_3) Z_{t+1}}{1 - \rho_3^2} & t = 3, \ldots, n-1 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

$$\hat{Z}_t^{2,3} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1, 2 \\ \\ \frac{\rho_2}{1 + \rho_4} (Z_{t-2} \ + Z_{t+2} \ ) & t = 3, \ldots, n-2 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n-1, n \end{cases}$$

$$\hat{Z}_t^{2,4} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \\ \frac{\rho_1}{1 + \rho_1} (Z_{t-1} \ + \ Z_{t+1} \ ) & t = 2, \ldots, n-1 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

$$\hat{Z}_t^{2,5} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \\ \frac{(\rho_1 - \rho_2 \rho_3) Z_{t-1} \ + (\rho_2 - \rho_1 \rho_3) Z_{t+2}}{1 - \rho_3^2} & t = 2, \ldots, n-2 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n-1, n \end{cases}$$

$$\hat{Z}_t^{2,6} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1, \ldots, n-2 \\ \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n-1, n \end{cases}$$

$$\hat{Z}_t^{3,1} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \psi_2 \left(aZ_{t-1} \ + \ b \ Z_{t+1} \ + \ c \ Z_{t+2} \ \right) & t = 2 \\ \psi_2 \left(cZ_{t-2} \ + \ b \ Z_{t-1} \ + \ a \ Z_{t+1} \ \right) & t = 3, \ldots, n-1 \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

where $\psi_2 = \frac{1}{1-\rho_1^2-\rho_2^2-\rho_3^2+2\rho_1\rho_2\rho_3}$, $a = \rho_1 + \rho_1\rho_2^2 + \rho_1^2\rho_3 - \rho_1\rho_2 - \rho_2\rho_3 - \rho_1^3$, $b = \rho_1 + \rho_3\rho_2^2 + \rho_1^2\rho_3 - 2\rho_1\rho_2 - \rho_1\rho_3^2$, $c = \rho_2 + \rho_2\rho_1^2 + \rho_1\rho_2\rho_3 - \rho_1^2 - \rho_1\rho_3 - \rho_2^3$.

$$\hat{Z}_t^{3,2} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \psi_2 \left(aZ_{t-1} \ + \ b \ Z_{t+1} \ + \ c \ Z_{t+2} \ \right) & t = 2 \\ qZ_{t-2} + \ r \ Z_{t-1} \ + \ s \ Z_{t+2} & t = 3, \ldots, n-2 \\ \psi_2 \left(cZ_{t-2} \ + \ b \ Z_{t-1} \ + \ a \ Z_{t+1} \ \right) & t = n-1 \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

where $s = \frac{\rho_1\rho_3(\rho_2-1)+\rho_1^2(\rho_4-\rho_2)+\rho_2(1-\rho_4)}{1-\rho_1^2-\rho_3^2-\rho_4^2+2\rho_1\rho_3\rho_4}$, $r = \frac{\rho_1(1-\rho_2)+(\rho_1\rho_4-\rho_3)s+\rho_2(1-\rho_4)}{1-\rho_1^2}$ *and* $q = \rho_2 - \rho_4 s - \rho_1 r$.

$$\hat{Z}_t^{3,3} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \psi_2 \left(aZ_{t-1} \ + \ b \ Z_{t+1} \ + \ c \ Z_{t+2} \ \right) & t = 2 \\ sZ_{t-2} + \ r \ Z_{t+1} \ + \ q \ Z_{t+2} & t = 3, \ldots, n-2 \\ \psi_2 \left(cZ_{t-2} \ + \ b \ Z_{t-1} \ + \ a \ Z_{t+1} \ \right) & t = n-1 \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

$$\hat{Z}_t^{3,4} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \psi_2 \left(aZ_{t-1} \ + \ b \ Z_{t+1} \ + \ c \ Z_{t+2} \ \right) & t = 2, \ldots, n-2 \\ \psi_2 \left(cZ_{t-2} \ + \ b \ Z_{t-1} \ + \ a \ Z_{t+1} \ \right) & t = n-1 \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

$$\hat{Z}_t^{4,1} = \begin{cases} \varphi_1 \ Z_{t+1} \ + \varphi_2 \ Z_{t+2} & t = 1 \\ \psi_2 \left(aZ_{t-1} \ + \ b \ Z_{t+1} \ + \ c \ Z_{t+2} \ \right) & t = 2 \\ \psi_1 (\alpha \left(Z_{t-2} \ + \ Z_{t+2} \ \right) + \beta \left(Z_{t-1} \ + \ Z_{t+1} \ \right)) & t = 3, \ldots, n-2 \\ \psi_2 \left(cZ_{t-2} \ + \ b \ Z_{t-1} \ + \ a \ Z_{t+1} \ \right) & t = n-1 \\ \varphi_2 \ Z_{t-2} \ + \varphi_1 \ Z_{t-1} & t = n \end{cases}$$

where $\psi_1 = \frac{1}{(1+\rho_2)(1+\rho_4)-(\rho_1+\rho_3)^2}$, $\alpha = \rho_2(1+\rho_2) - \rho_1(\rho_1+\rho_3)$, $\beta = \rho_1(1+\rho_4) - \rho_2(\rho_1+\rho_3)$. $\rho_1 = \frac{\varphi_1}{1-\varphi_2}$, $\rho_2 = \frac{\varphi_1^2+\varphi_2-\varphi_2^2}{1-\varphi_2}$, $\rho_3 = \frac{\varphi_1^3-\varphi_1\varphi_2^2}{1-\varphi_2}$ *and* $\rho_4 = \frac{\varphi_1^4+\varphi_2^2+\varphi_2\varphi_1^2-\varphi_1^2\varphi_2^2-\varphi_2^3}{1-\varphi_2}$.

Proof. First consider these predictors are in the form of

$$\hat{Z}_t^{k,l} = \sum_{i\in O_t^{k,l}} \alpha_{k,i}\ Z_i\ . \tag{4}$$

The coefficients $\alpha_{k,i}s$ are earned by minimizing MSP. By the projection theorem these coefficients for any $k = 1,2,3,4$ are resulted by

$$corr(Z_t\ - \hat{Z}_t^{k,l},\ Z_j\ ) = 0 \qquad\qquad \forall\ \ j \in O_t^{k,l}. \tag{5}$$

Therefore, by using of (4) in (5) we have $\sum_{i\in O_t^k} \alpha_{k,i}\ corr\ (\ Z_i\ ,\ Z_j\ ) = corr(Z_t\ ,\ Z_j\ )\ \forall\ \ j \in O_t^{k,l}$, and then

$$\sum_{i\in O_t^k} \alpha_{k,i}\ \rho\left(|i-j|\right) = \rho\left(|t-j|\right).\forall\ \ j \in O_t^{k,l}. \tag{6}$$

Depending on the number of members in $O_t^{k,l}$, (6) is a system of equations with the same set of unknowns. Solving these systems of linear equations and then using of Equation (3) for earning to $\rho(h)\ \ h = 1,\ldots,4$, complete the proof. □

The comparison between these predictors is the subject of continuity of this section. PMC criterion for comparing two predictors $\hat{Z}$ and $\tilde{Z}$ is given by $\mathrm{PMC}_Z\left(\hat{Z}\ \middle|\ \tilde{Z}\right) = P\left\{\left|\hat{Z}-Z\right| < \left|\tilde{Z}-Z\right|\right\}$. The predictor $\hat{Z}$ is said to be Pitman-closer to variable Z than the predictor $\tilde{Z}$ if $\mathrm{PMC}_Z\left(\hat{Z}\ \middle|\ \tilde{Z}\right) > 0.5$. In order to have more details on the concept and characteristics of this criterion see *Keating et al.* (1993). Various works on PMC is listed in a special issue of Communications in Statistics, Theory and Methods ((1991), Vol. 20, No. 11).

The predictors $\hat{Z}_t^{k,l}$s are too complicated to evaluate the MSP and PMC in the closed forms. Therefore, we will need to simulate the answer by

$$\widehat{\text{MSP}}_Z\left(\hat{Z}\right)=\frac{\sum_{i=1}^{n}\left(\hat{Z}_i-Z_i\right)^2}{n}, \tag{7}$$

and

$$\widehat{\text{PMC}}_Z\left(\hat{Z}\;\middle|\;\tilde{Z}\right)=\frac{\sum_{i=1}^{n}\;\;U(\left|\tilde{Z}_i-Z_i\right|-\left|\hat{Z}_i\;--Z_i\right|)}{n}, \tag{8}$$

where $U(t)$ is 1 for positive t and 0 otherwise.
For computing (7) and (8) a leave one out approach is applied. We simulate n observations form (1) and then the following algorithm is applied.
Algorithm A.

1. Let $t=1$. Remove Z_t and define $\boldsymbol{Z}_{-t}$ as the vector of remained observations.
2. Compute $\hat{Z}_t^{k,l}$ by using $\boldsymbol{Z}_{-t}$.
3. Compute $\left(\hat{Z}_t^{k,l}--Z_t\right)^2$, $U(\left|\hat{Z}_t^{k,l}-Z_t\right|-\left|\hat{Z}_t^1--Z_t\right|)$ and $U(\left|\hat{Z}_t^2-Z_t\right|-\left|\hat{Z}_t^3--Z_t\right|)$.
4. Repeat steps a, b and c for $t=2,3,\ldots,n$.

In order to have a comparison between introduced predictors, we generate 5000 realizations from $AR(2)$ model with Exponential innovations.
First of all, results demonstrate that predictors which use from symmetric observations around unsampled time t have approximately same performance. In the other words, using of $\hat{Z}_t^{1,1}$ or $\hat{Z}_t^{1,4}$, $\hat{Z}_t^{1,2}$ or $\hat{Z}_t^{1,3}$, $\hat{Z}_t^{2,1}$ or $\hat{Z}_t^{2,6}$, $\hat{Z}_t^{2,2}$ or $\hat{Z}_t^{2,5}$, $\hat{Z}_t^{3,1}$ or $\hat{Z}_t^{3,4}$ and $\hat{Z}_t^{3,2}$ or $\hat{Z}_t^{3,3}$ have a very neglible difference $w.r.t$ to PMC or MSP. (We indicate veracity of this claim by using of figures through paper). Therefore, comparison with any predictor is accomplished only with one of them.
In order to have a comparison between recommended predictors, we generate 1000 realizations from model (1) with Gaussian innovations. Figure 1 demonstrates some of these comparisons between $\hat{Z}_t^{1,l}$ $l=1,\ldots,4$ for some values of φ_2 and all values of φ_1 which satisfy stationary condition. Figure 1 denotes that $MSP\left(\hat{Z}^{1,1}\right)\cong MSP\left(\hat{Z}^{1,4}\right)$ and $\widehat{PMC}\;\left(\hat{Z}^{1,1}\;\middle|\;\hat{Z}^{1,4}\right)\cong$

$\widehat{PMC}\left(\hat{Z}^{1,2} \,\middle|\, \hat{Z}^{1,3}\right) \cong 0.5$, so as we mentioned previously the performances of $\hat{Z}^{1,1}$ and $\hat{Z}^{1,2}$ are approximately same as $\hat{Z}^{1,4}$ and $\hat{Z}^{1,3}$, respectively. From Figure 1 this comes that $MSP\left(\hat{Z}^{1,1}\right) < MSP\left(\hat{Z}^{1,2}\right)$ and $\widehat{PMC}\left(\hat{Z}^{1,1} \,\middle|\, \hat{Z}^{1,2}\right)$ > 0.5 whenever $|\varphi_2| > |\varphi_1|$. Therefore, $\hat{Z}^{1,1}$ and $\hat{Z}^{1,4}$ are better than $\hat{Z}^{1,2}$ and $\hat{Z}^{1,3}$ *w.r.t* both MSP and PMC for

$$\Theta^G_{1,1|1,2} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ;\ |\varphi_2| > |\varphi_1|\right\}. \tag{9}$$

This is logical from a statistically point of view. The predictors $\hat{Z}^{1,1}$ and $\hat{Z}^{1,4}$ use from Z_{t-2} and Z_{t+2} for prediction of Z_t which their coefficients in model (1) is φ_2, while $\hat{Z}^{1,2}$ and $\hat{Z}^{1,3}$ use from Z_{t-1} and Z_{t+1} for prediction of Z_t which their coefficients in model (1) is φ_1, so whenever $|\varphi_2| > |\varphi_1|$, Z_t is more dependent to Z_{t-2} and Z_{t+2} than Z_{t-1} and Z_{t+1}. This causes $\hat{Z}^{1,1}$ and $\hat{Z}^{1,4}$ have better performance than $\hat{Z}^{1,2}$ and $\hat{Z}^{1,3}$ in this case.

The results of comparison between $\hat{Z}^{2,l}_t$ $l = 1, \ldots, 6$ are long since 6 predictors have $\begin{pmatrix} 6 \\ 2 \end{pmatrix} = 15$ plots for comparison. Therefore, we first compare $\hat{Z}^{2,l}_t$ $l = 1, 2, 5, 6$ since $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,6}_t$ and also $\hat{Z}^{2,2}_t$ and $\hat{Z}^{2,5}_t$ as we mentioned previously have the same performance. Some of these results

have been represented in Figure 2. Both criteria MSP and PMC show that the predictors $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,2}_t$ have the same performance with predictors $\hat{Z}^{2,6}_t$ and $\hat{Z}^{2,5}_t$, respectively. So, comparison between $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,2}_t$ is enough for concluding. Our findings demonstrate the predictors $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,6}_t$ are approximately better than predictors $\hat{Z}^{2,2}_t$ and $\hat{Z}^{2,5}_t$ *w.r.t* both MSP and PMC for

$$\Theta^G_{2,1|2,2} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ;\ \varphi_2 \in (-1, -0.2) \text{ and } |\varphi_1| < 0.72 - \varphi_2\right\}. \tag{10}$$

However, the amount and strength of this being better is various for values of parameter φ_2 in $(-1, -0.2)$. Although $\widehat{PMC}\left(\hat{Z}^{2,1} \,\middle|\, \hat{Z}^{2,2}\right) > 0.5$ for all values of parameters φ_1 and φ_2 in set $\Theta^G_{2,1|2,2}$, this quantity is approximately a decreasing function of φ_2 whenever φ_1 and φ_2 are in set $\Theta^G_{2,1|2,2}$.

Notice that results show the predictors $\hat{Z}^{2,2}_t$ and $\hat{Z}^{2,5}_t$ are approximately better than predictors $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,6}_t$ *w.r.t* both MSP and PMC for $\Theta_{stationary} - \Theta^G_{2,1|2,2}$. However, difference between them is very inconsiderable for some

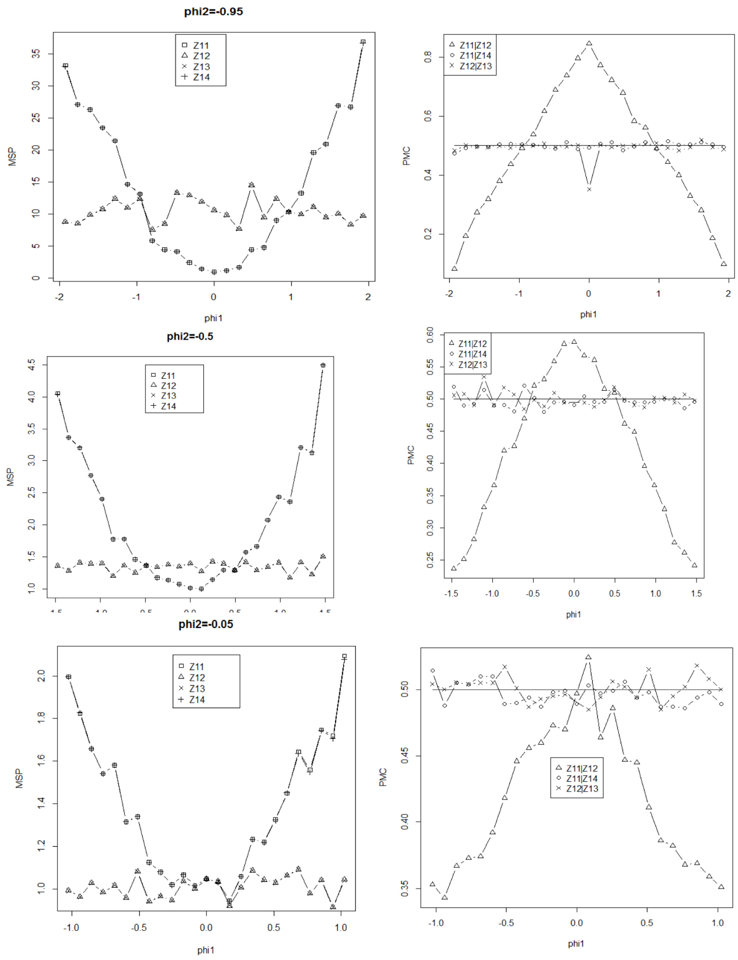

Figure 1. Comparison between predictors $\hat{Z}_t^{1,l}$ $l = 1, \ldots, 4$ for $\varphi_2 = -0.95,\ -0.5,\ -0.05$.

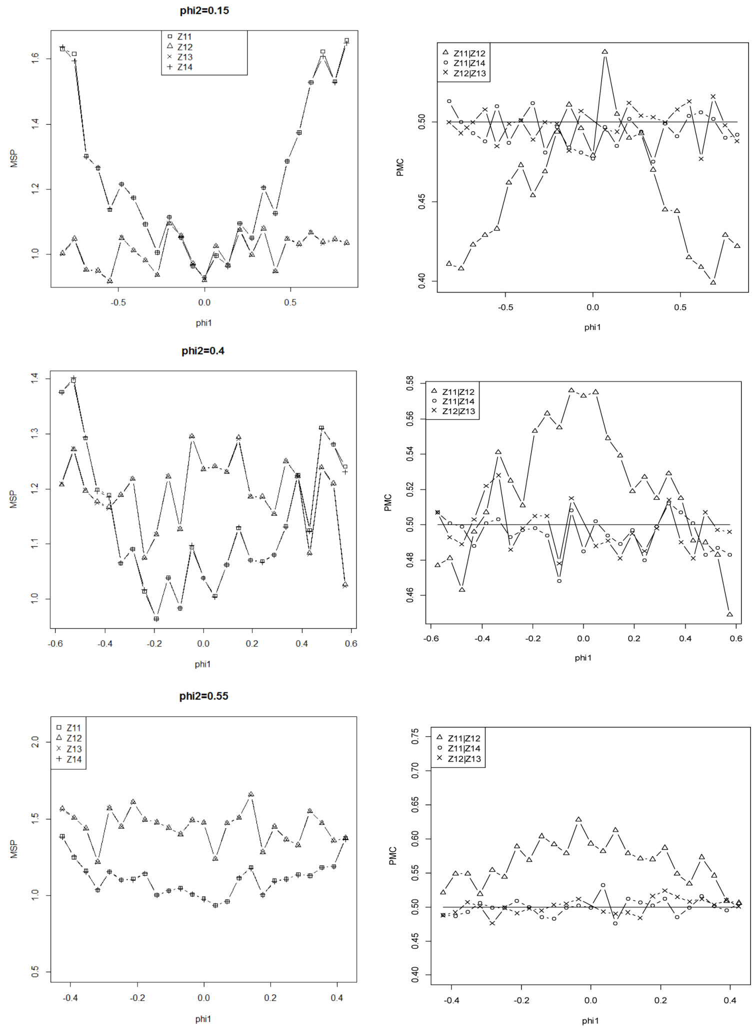

Figure 2. Comparison between predictors $\hat{Z}_t^{1,l}$ $l = 1, \ldots, 4$ for $\varphi_2 = 0.15,\ 0.4,\ 0.55$.

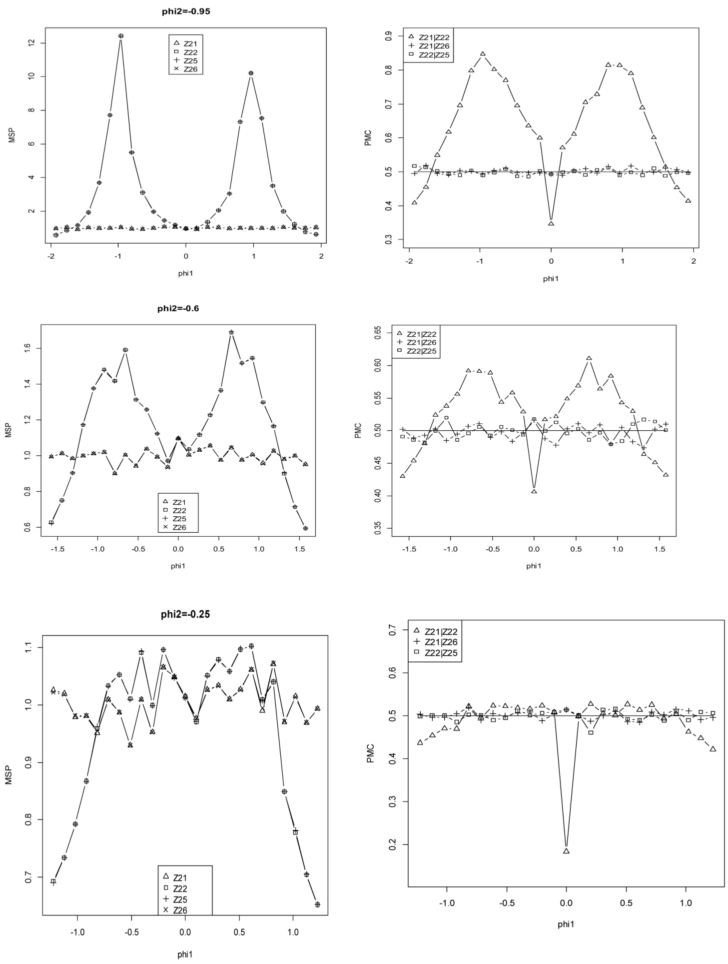

Figure 3. Comparison between predictors $\hat{Z}_t^{2,l}$ $l = 1, 2, 5, 6$ for $\varphi_2 = -0.95,\ -0.6,\ -0.25$.

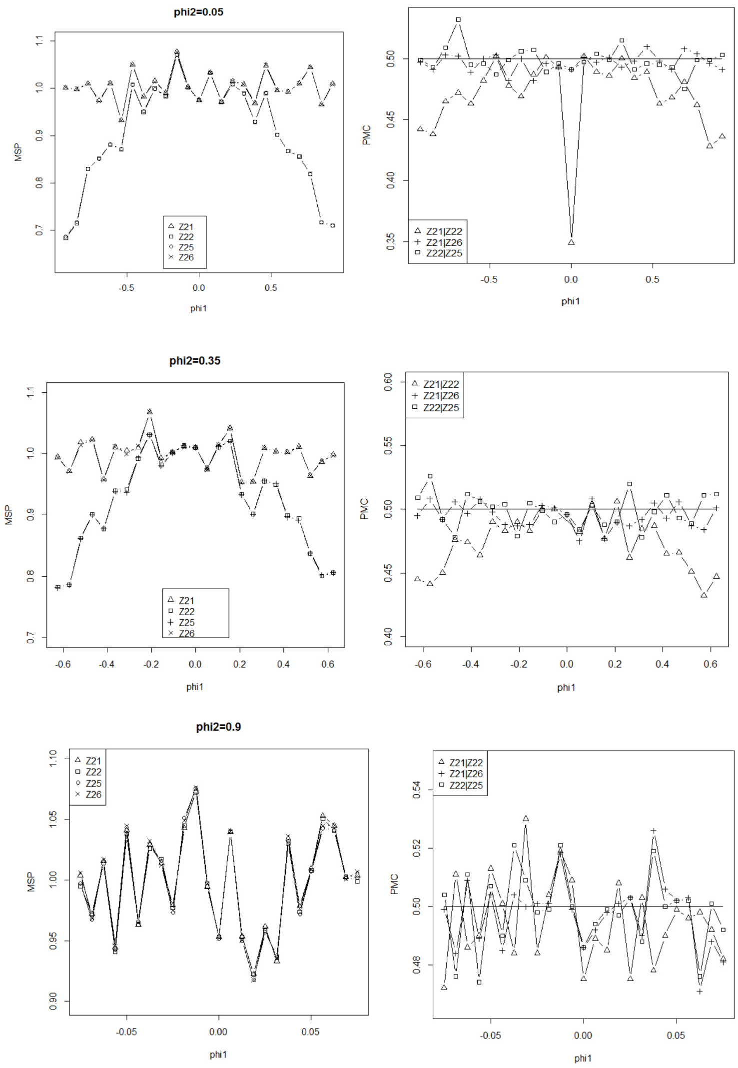

Figure 4. Comparison between predictors $\hat{Z}_t^{2,l}$ $l = 1, 2, 5, 6$ for $\varphi_2 = -0.25,\ 0.35,\ 0.9$.

values of parameters φ_1 and φ_2 in $\Theta_{stationary} - \Theta^G_{2,1|2,2}$. Indeed the predictors $\hat{Z}^{2,2}_t$ and $\hat{Z}^{2,5}_t$ are approximately better significantly than predictors $\hat{Z}^{2,1}_t$ and $\hat{Z}^{2,6}_t$ w.r.t both MSP and PMC for

$$\Theta^G_{2,2|2,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ;\ \varphi_2 \in (-1, -0.2) \text{ and } |\varphi_1| > 0.72 - \varphi_2\right\}. \tag{11}$$

For other values of φ_1 and φ_2 in $\Theta^G_{2,1256} = \Theta_{stationary} - \left(\Theta^G_{2,1|2,2} \cup \Theta^G_{2,2|2,1}\right)$, predictors $\hat{Z}^{2,k}_t$ $k = 1, 2, 5, 6$ have approximately same performance. By substitution (10) and (11) in $\Theta^G_{2,1256}$, it can be represented in following simpler form

$$\Theta^G_{2,1256} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ;\ \varphi_2 \in [-0.2,\ 1)\right\}, \tag{12}$$

which shows this set has only determined by parameter φ_2.

By earned results about comparing predictors $\hat{Z}^{2,l}_t$ s until now, we need to compare predictors $\hat{Z}^{2,l}_t$ $l = 1, 3, 4$ in set $\Theta^G_{2,1|2,2}$ and $\hat{Z}^{2,l}_t$ $l = 2, 3, 4$ in set $\Theta_{stationary} - \Theta^G_{2,1|2,2}$. In order to having simpler inference, these two works are done for $\varphi_2 < -0.2$ and $\varphi_2 \geq -0.2$, respectively.

Our results denote that MSP($\hat{Z}^{2,4}_t$) is very bigger than MSP($\hat{Z}^{2,l}_t$) $l = 1, 2, 3$ for

$$\Theta^G_{24:inadmissible} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ;\ \varphi_1 \in (\varphi_2 - 1,\ \varphi_2 - 0.7)\right\}. \tag{13}$$

That causes the plots of MSP($\hat{Z}^{2,l}_t$) $l = 1, \ldots, 4$ cannot show real difference between predictors. Since values of PMC($\hat{Z}^{k,l}_t$) cannot exceed of 1, comparison between $\hat{Z}^{2,l}_t$ $l = 1, \ldots, 4$ is only done by criteria PMC.

Figure 3 shows comparison between predictors $\hat{Z}^{2,l}_t$ $l = 1, 3, 4$ for $\varphi_2 < -0.2$. This figure shows $\widehat{PMC}\ \left(\hat{Z}^{2,1}\ \middle|\ \hat{Z}^{2,3}\right) > 0.5$ and $\widehat{PMC}\ \left(\hat{Z}^{2,1}\ \middle|\ \hat{Z}^{2,4}\right) > 0.5$ approximately for

$$\Theta^G_{2,1|(2,3),(2,4)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ; \varphi_2 < -0.2,\ \varphi_1 < \varphi_2\right\}. \tag{14}$$

Therefore, in this set $\hat{Z}^{2,1}$ is better than $\hat{Z}^{2,3}$ and $\hat{Z}^{2,4}$.

Also, $\widehat{PMC}\ \left(\hat{Z}^{2,1}\ \middle|\ \hat{Z}^{2,4}\right) < 0.5$ and $\widehat{PMC}\ \left(\hat{Z}^{2,3}\ \middle|\ \hat{Z}^{2,4}\right) < 0.5$ approximately for

$$\Theta^G_{2,4|(2,1),(2,3)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ; \varphi_2 < -0.2,\ \varphi_1 > -\varphi_2\right\}. \tag{15}$$

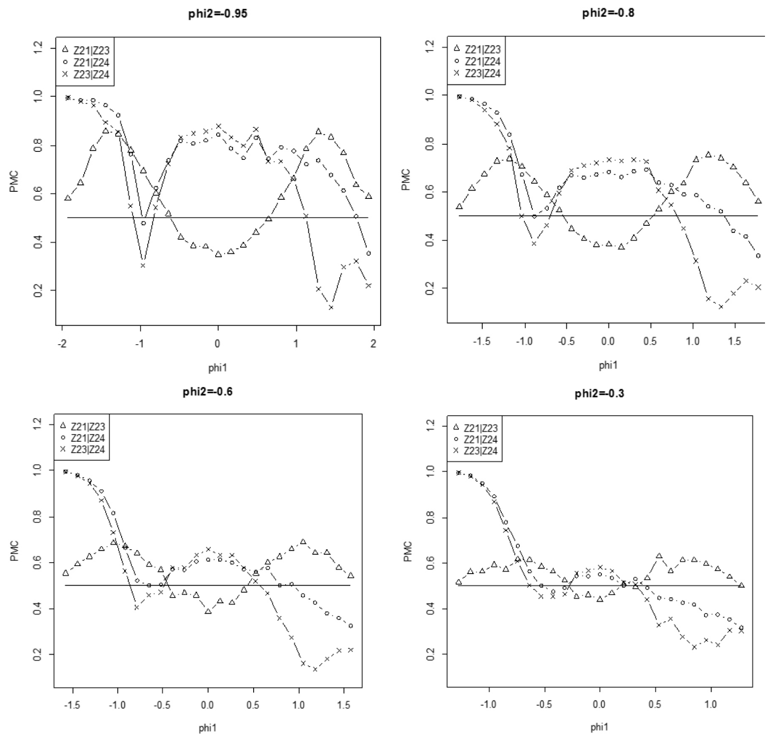

Figure 5. Comparison between predictors $\hat{Z}_t^{2,l}$ $l = 1, 3, 4$ for $\varphi_2 = -0.95, -0.8, -0.6, -0.3$.

Therefore, in this set $\hat{Z}^{2,4}$ is better than $\hat{Z}^{2,1}$ and $\hat{Z}^{2,3}$. Finally, $\hat{Z}^{2,3}$ is better than $\hat{Z}^{2,1}$ and $\hat{Z}^{2,4}$ in

$$\Theta^G_{2,3|(2,1),(2,4)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \,; \varphi_2 < -0.2,\ |\varphi_1| < -\varphi_2\right\}. \qquad (16)$$

Figure 4 shows comparison between predictors $\hat{Z}_t^{2,l}$ $l = 2, 3, 4$ for $\varphi_2 \geq -0.2$. From this figure we can see $\hat{Z}^{2,2}$ is better than $\hat{Z}^{2,3}$ and $\hat{Z}^{2,4}$ for

$$\Theta^G_{2,2|(2,3),(2,4)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \,; -0.2 \leq \varphi_2 < 0.45,\ \varphi_1 < 0\right\}, \qquad (17)$$

and $\hat{Z}^{2,4}$ is better than $\hat{Z}^{2,2}$ and $\hat{Z}^{2,3}$ for

$$\Theta^G_{2,4|(2,2),(2,3)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \,; -0.2 \leq \varphi_2 < 0.45,\ \varphi_1 > 0\right\}. \qquad (18)$$

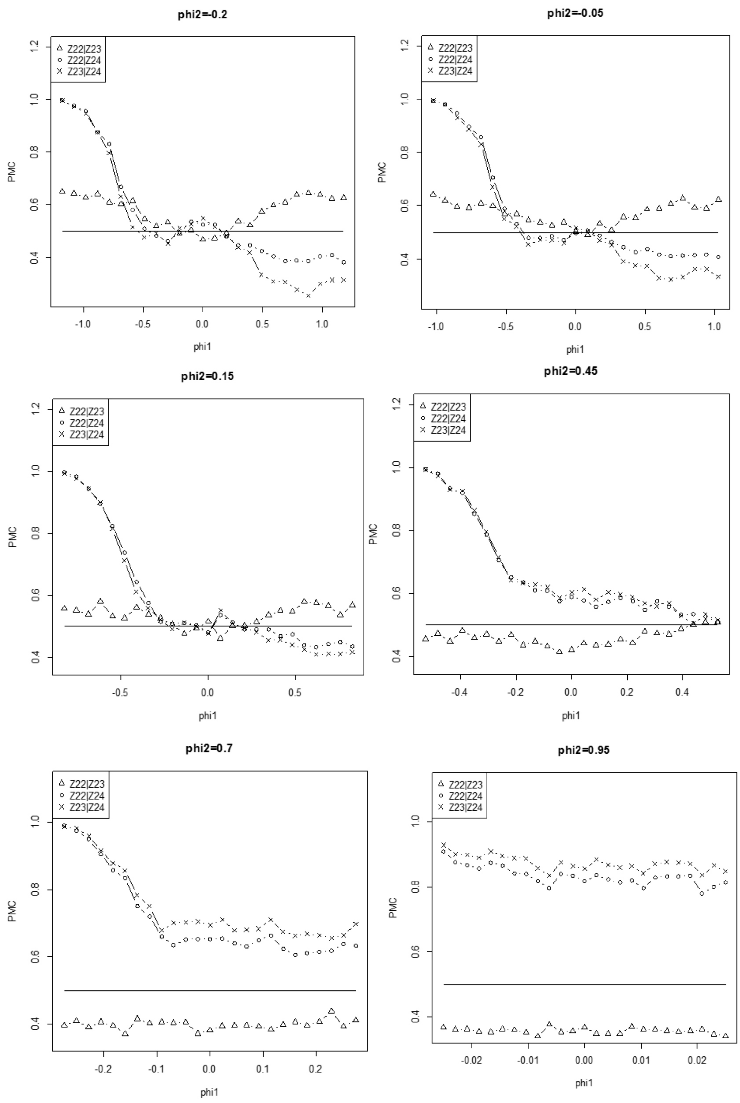

Figure 6. Comparison between predictors $\hat{Z}_t^{2,l}$ $l = 2, 3, 4$ for $\varphi_2 = -0.2,\ -0.05,\ 0.15,\ 0.45,\ 0.7,\ 0.95$.

Also, $\hat{Z}^{2,3}$ is better than $\hat{Z}^{2,2}$ and $\hat{Z}^{2,4}$ for

$$\Theta^{G}_{2,3|(2,2),(2,4)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; \varphi_2 \geq 0.45\right\}. \tag{19}$$

Consequently, from all results about comparison between predictors $\hat{Z}^{2,l}_t$ $l = 1, \ldots, 6$ (Equations (2.7)-(18)) we have following remarks.

1. The predictors $\hat{Z}^{2,1}_t$and $\hat{Z}^{2,6}_t$ are better than other predictors $\hat{Z}^{2,l}_t$ $l \neq 1, 6$ for values of parameters φ_1 and φ_2 which are in set

$$\Theta^{G}_{(2,1),(2,6)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; \varphi_2 < -0.2, \; \varphi_1 < \varphi_2\right\}.$$

2. The predictors $\hat{Z}^{2,2}_t$and $\hat{Z}^{2,5}_t$ are better than other predictors $\hat{Z}^{2,l}_t$ $l \neq 2, 5$ for values of parameters φ_1 and φ_2 which are in set

$$\Theta^{G}_{(2,2),(2,5)} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; -0.2 \leq \varphi_2 < 0.45, \; \varphi_1 < 0\right\}.$$

3. The predictor $\hat{Z}^{2,3}_t$ is better than other predictors $\hat{Z}^{2,l}_t$ $l \neq 3$ for values of parameters φ_1 and φ_2 which are in set

$$\Theta^{G}_{(2,3)} \\ = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; (\varphi_2 < -0.2, \; |\varphi_1| < -\varphi_2) \; or \; (\varphi_2 \geq 0.45)\right\}.$$

4. The predictor $\hat{Z}^{2,3}_t$ is better than other predictors $\hat{Z}^{2,l}_t$ $l \neq 3$ for values of parameters φ_1 and φ_2 which are in set

$$\Theta^{G}_{(2,4)} \\ = \{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; (-\varphi_1 < \varphi_2 < -0.2) \\ or (-0.2 \leq \varphi_2 < 0.45, \; \varphi_1 > 0)\}.$$

In continuation of this section, predictors $\hat{Z}^{3,l}_t$ $l = 1, \ldots, 4$ are compared in order to find the best predictor among predictors which use from 3 observations. The results of this comparison have been shown in Figure 5. As Figure 5 shows, predictors $\hat{Z}^{3,1}_t$ and $\hat{Z}^{3,4}_t$ have the same performance *w.r.t* MSP and PMC. This outcome holds for predictors $\hat{Z}^{3,2}_t$ and $\hat{Z}^{3,3}_t$, too. Comparison between predictors $\hat{Z}^{3,1}_t$ and $\hat{Z}^{3,2}_t$ demonstrates that predictor $\hat{Z}^{3,2}_t$ is better than predictor $\hat{Z}^{3,1}_t$ for

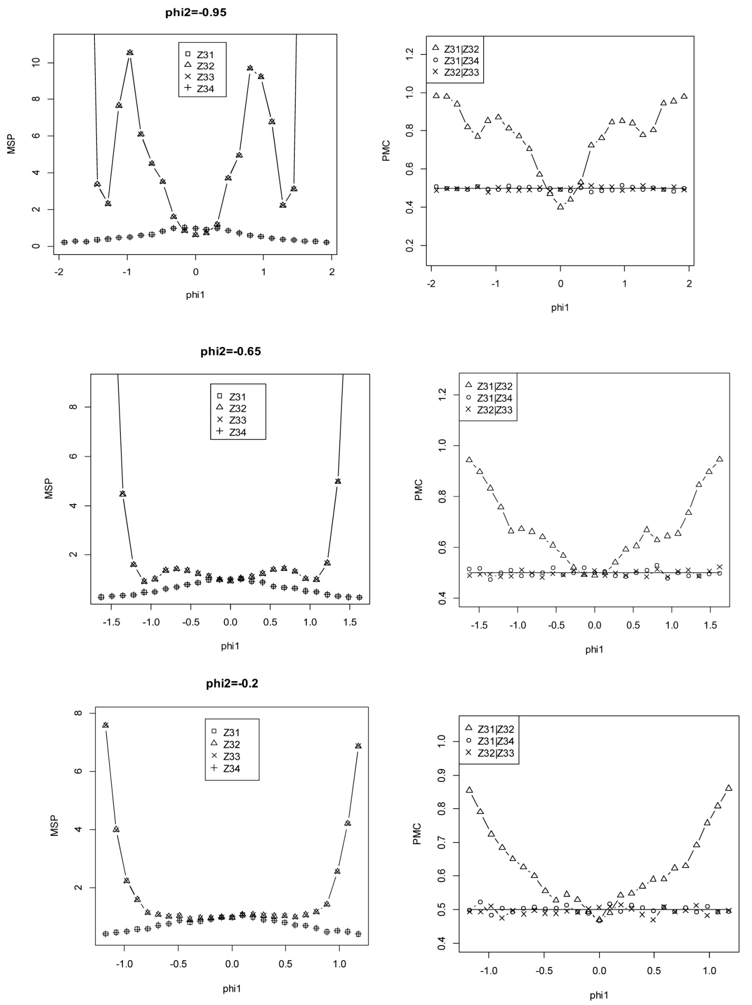

Figure 7. Comparison between predictors $\hat{Z}_t^{3,l}$ $l = 1, \ldots, 4$ for $\varphi_2 = -0.95,\ -0.65,\ -0.2$.

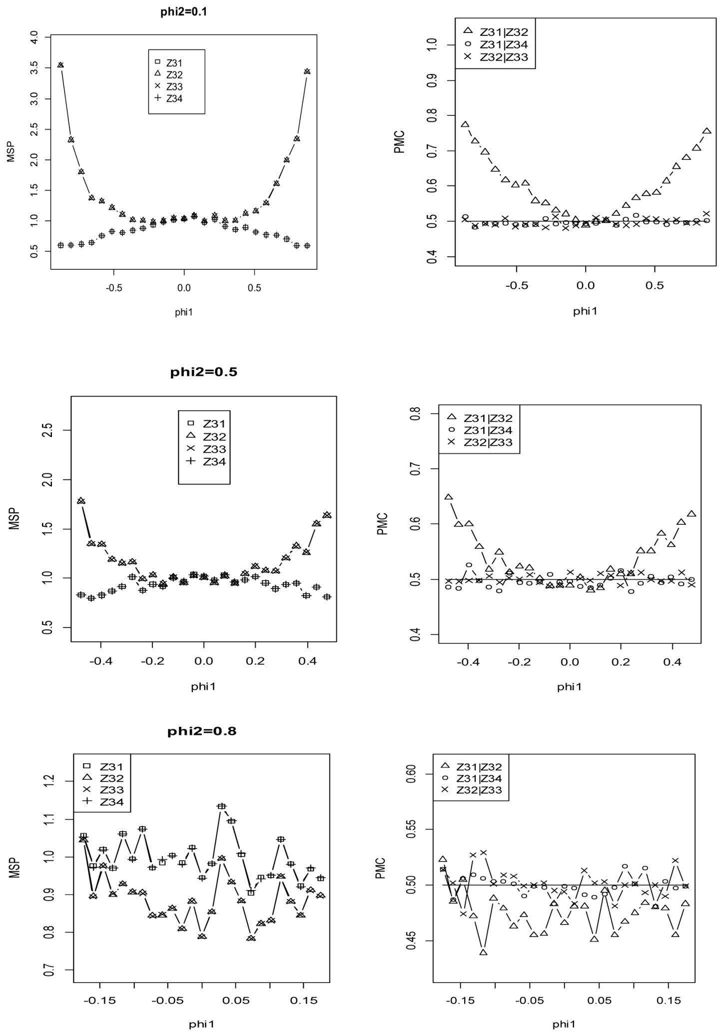

Figure 8. Comparison between predictors $\hat{Z}_t^{3,l}$ $l = 1, \ldots, 4$ for $\varphi_2 = -0.95, -0.65, -0.2.$

$$\Theta^{G}_{3,2|3,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; \; \varphi_2 > 0.75 \text{ or } |\varphi_1| < 0.15\right\}. \qquad (20)$$

Because of this fact that predictors $\hat{Z}^{3,2}_t$ and $\hat{Z}^{3,3}_t$ use from both Z_{t-2} and Z_{t+2} which their coefficients in model (1) is φ_2, these predictors have their best performance whenever φ_2 is big or φ_1 is near to 0.

Until now, we could find the best predictors among predictors $\hat{Z}^{k,l}_t$ $l = 1, \ldots, \binom{4}{k}$ for any fixed value of $k = 1, 2, 3$. Now, 14 comparisons are done between $\hat{Z}^{4,1}_t$ and all other 14 introduced predictors. We omit figures of these comparisons because of saving in space. However, these results have been brought in following. As the previous results, $\Theta^{G}_{k,l|4,1}$ shows the set of φ_1 and φ_2 which $\hat{Z}^{k,l}_t$ is better than $\hat{Z}^{4,1}_t$ or difference between $\hat{Z}^{k,l}_t$ and $\hat{Z}^{4,1}_t$ is negligible.

$$\Theta^{G}_{1,l|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; |\varphi_2| \leq 0.05 \text{ and } 0 < \varphi_1 < 0.1\right\}; \; l = 1, \ldots, 4$$

$$\Theta^{G}_{2,l|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; |\varphi_2| \leq 0.05 \text{ and } 0 < \varphi_1 < 0.1\right\}; \; l = 1, 2, 5, 6$$

$$\Theta^{G}_{2,3|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; (\varphi_2 \leq 0.7 \text{ and } |\varphi_1| < 0.05) or (\varphi_2 > 0.7) \right\};$$

$$\Theta^{G}_{2,4|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; -0.05 \leq \varphi_2 \leq 0.15 \text{ and } |\varphi_1| < 0.05\right\};$$

$$\Theta^{G}_{3,l|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; |\varphi_2| \leq 0.05\right\}; \; l = 1, 4$$

$$\Theta^{G}_{3,l|4,1} = \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary} \; ; |\varphi_2| \leq 0.05 \; and \; -0.05 \leq \varphi_1 \leq 0\right\}; \; l = 2, 3$$

These results show that the only predictor among all 14 predictors which can outpace predictor $\hat{Z}^{4,1}_t$ is predictor $\hat{Z}^{2,3}_t$ for big values of φ_2. All other predictors only for a small set of φ_1 and φ_2 which these parameters are near to 0, are better than predictor $\hat{Z}^{4,1}_t$.

2.1 Comparison in Case of Gamma Distributed Errors

Pace to pace similar to case of Gaussian innovation, the comparison has been done for Gamma distributed errors ($\varepsilon_t \sim Gamma(\alpha, \lambda)$). However, Theorem 2.1 is related to mean zero variables Z_t . So, for general case we need to use the centered variables ($Z^*_t = Z_t - E(Z_t)$). By *Saber and*

Nematollahi (2017) two criteria MSP and PMC are invariant with respect to location parameter which here is constant $E(Z_t\ )$. By stationarity we have $E\left(Z_t\ \right) = \frac{E(\varepsilon_t\ ).}{1-\varphi_1-\varphi_2}$ so we convert Z_t to

$$Z_t^* = Z_t\ - \frac{\alpha}{\lambda\left(1-\varphi_1-\varphi_2\right)}. \tag{21}$$

Results for this case are very similar with Gaussian distributed errors, so we do not present all of them in this study. For save of space, we bring only one case for parameters (α= 4, λ= 1). Our computations for another some parameters show approximately same results as mentioned parameters. Some of these results for $\alpha = 4$ and $\lambda = 1$ have been randomly shown in Figure 6. Although there is no more general explanation in this case than Gaussian case, this has worth
to notice on values of criterion MSP in cases of Gaussian and Gamma distributed errors. For latter case, MSP is very small in comparison with Gaussian case. For instance, for $\varphi_2 = 0.95$ maximum of criterion MSP for predictor $\hat{Z}_t^{3,1}$ is 0.0043 while its corresponding value for case of Gaussian distributed errors is 1.0805. This shows that prediction for Gamma distributed errors is very exact and exquisite.
Another little difference with Gaussian case is about similar performance for the predictors which use from symmetric observations around time t. As we showed for Gaussian case these predictors have really same performance *w.r.t* both MSP and PMC. However, for this case there is a bit difference between predictors $\hat{Z}_t^{3,1}$ and $\hat{Z}_t^{3,4}$. In fact for values of φ_1 and φ_2 in set

$$\Theta_{3,1|3,4}^{Gamm} = \left\{(\varphi_1,\varphi_2)\in\Theta_{stationary}\ ;(0.3\le\varphi_2<0.8\ and\ \varphi_1\ge 0.7-\varphi_2)\ or\ (0.8\le\varphi_2)\right\};$$

$\hat{Z}_t^{3,1}$ is better than $\hat{Z}_t^{3,4}$.

2.2 Prediction of Missing Values by Using of Far Observations

In this section, we find three linear predictors for Z_t in model (1) based on observations which are not near to non-sampled time. In fact, we consider the circumstance that there are only two observation which are far from unobserved time. The first predictor use observations in the past and future

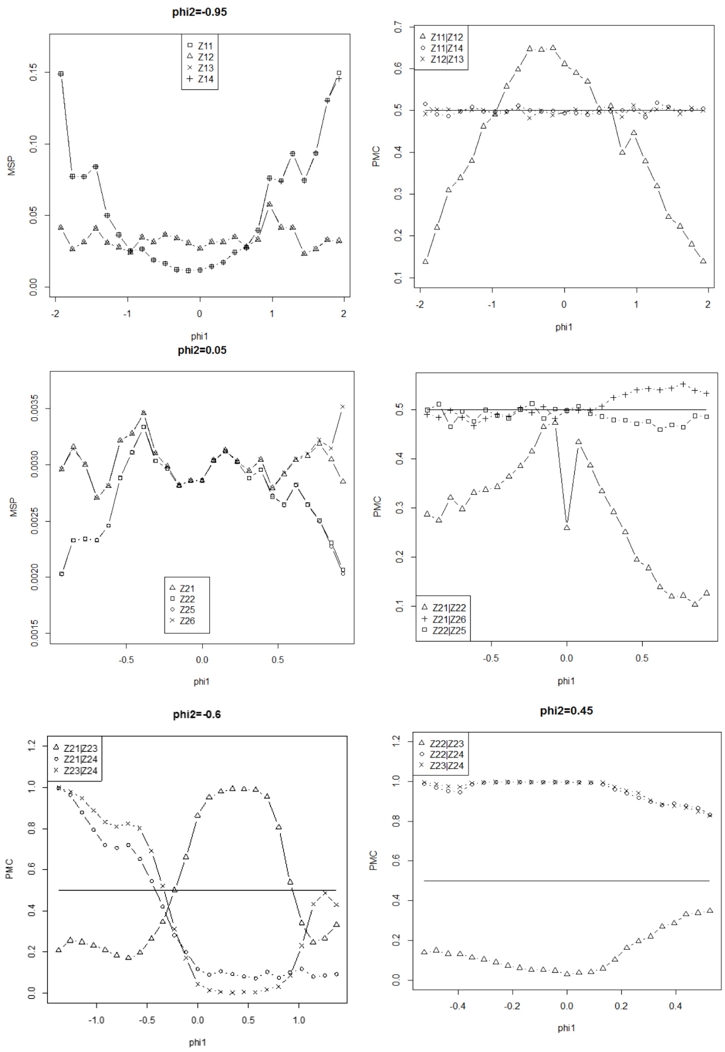

Figure 9. Comparison between predictors $\hat{Z}_t^{k,l}$ $k = 1, 2$ for Gamma distributed errors.

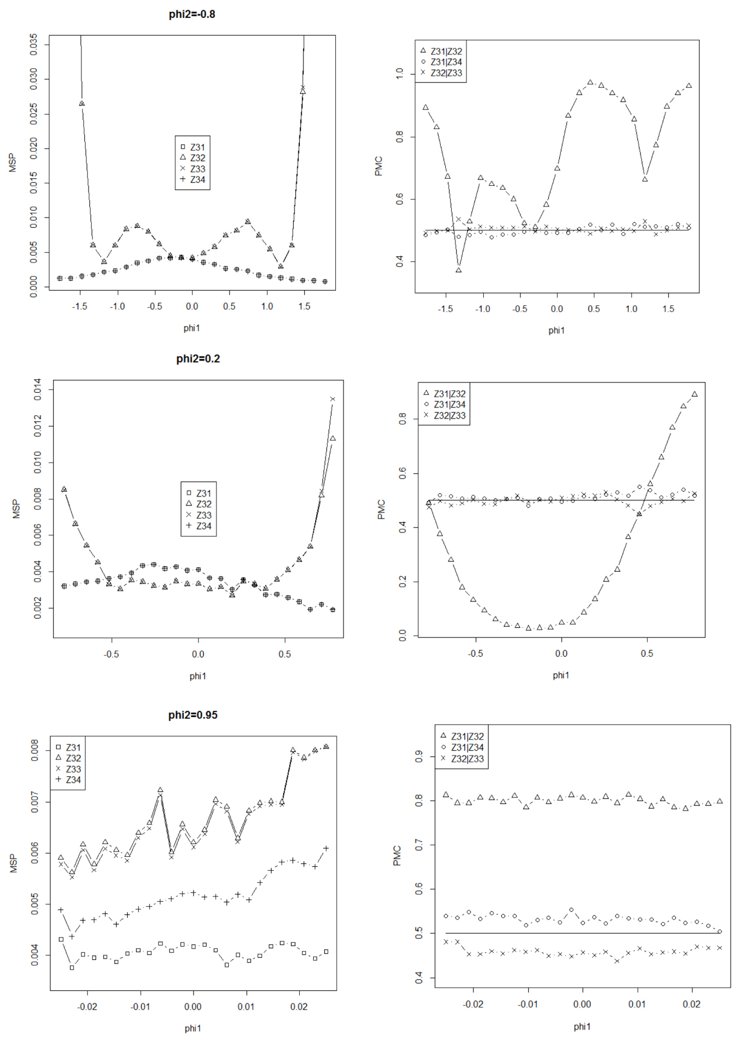

Figure 10. Comparison between predictors $\hat{Z}_t^{3,l}$ $l = 1, \ldots, 4$ for Gamma distributed errors.

(O_t^{pf}) while two other predictors use from observations which are only in past (O_t^{pp}) and future (O_t^{ff}), respectively. All of the symbols and notations are same as Section 2, so we only define

$$O_t^{pf} = \{t-u, t+v\}\,; O_t^{ff} = \{t+u, t+v\}\,; O_t^{pp} = \{t-v, t-u\} \qquad (22)$$

where $u \geq 3$ and $v > u$. This is obvious that $t-v \geq 1$ and $t+v \leq n$.
As in previous section we have $\hat{Z}_t^k = E\left(Z_t \;\middle|\; \boldsymbol{Z}_t^k\right),\ k = pf, pp, ff$. The following theorem gives the the best linear predictors for Z_t based on O_t^k w.r.t MSP.

Theorem 2. *Let Z_t comes from model (1) and O_t^k, $k = pf, pp, ff$ are defined as (22). If $E(Z_t) = 0$ and $Var(Z_t) < \infty$, then the best linear predictors for Z_t based on O_t^k w.r.t MSP, $k = pf, pp, ff$ are given by*

$$\hat{Z}_t^{pf} = \frac{1}{1-\rho_{u+v}^2}\left(\left(\rho_u - \rho_v\rho_{u+v}\right)Z_{t-u} + \left(\rho_v - \rho_u\rho_{u+v}\right)Z_{t+v}\right),$$

$$\hat{Z}_t^{ff} = \frac{1}{1-\rho_{v-u}^2}\left(\left(\rho_u - \rho_v\rho_{v-u}\right)Z_{t+u} + \left(\rho_v - \rho_u\rho_{v-u}\right)Z_{t+v}\right)$$

and

$$\hat{Z}_t^{pp} = \frac{1}{1-\rho_{v-u}^2}\left(\left(\rho_v - \rho_u\rho_{v-u}\right)Z_{t-v} + \left(\rho_u - \rho_v\rho_{v-u}\right)Z_{t-u}\right).$$

Proof. All parts of proof in Theorem 1 is remained same here. □

Throughout this section we compare three predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $(u,v) = (3,4)$, $(5,8)$ and $(50,80)$. The results for $(u,v) = (3,4)$ have been represented in Figure 7. From this figure one can see that predictors $\hat{Z}_t^{pp}$ and $\hat{Z}_t^{ff}$have the same performance. Also, comparison between predictors $\hat{Z}_t^{pf}$ and $\hat{Z}_t^{ff}$ illustrate that the first predictor is better than latter *w.r.t* both MSP and PMC in set

$$\Theta_{pf|ff}^{G,\ (u,v)=(3,4)}$$

$$= \left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\,; (0.15 < \varphi_2 < 0.6\,)\ or\ (0.6 < \varphi_2\ and\ |\varphi_1| > 0.1\right\}. \qquad (23)$$

Here, we used of 6000 observation instead 5000 observation (in Section 2) since number of observations which are in borders is much. In fact, for case of

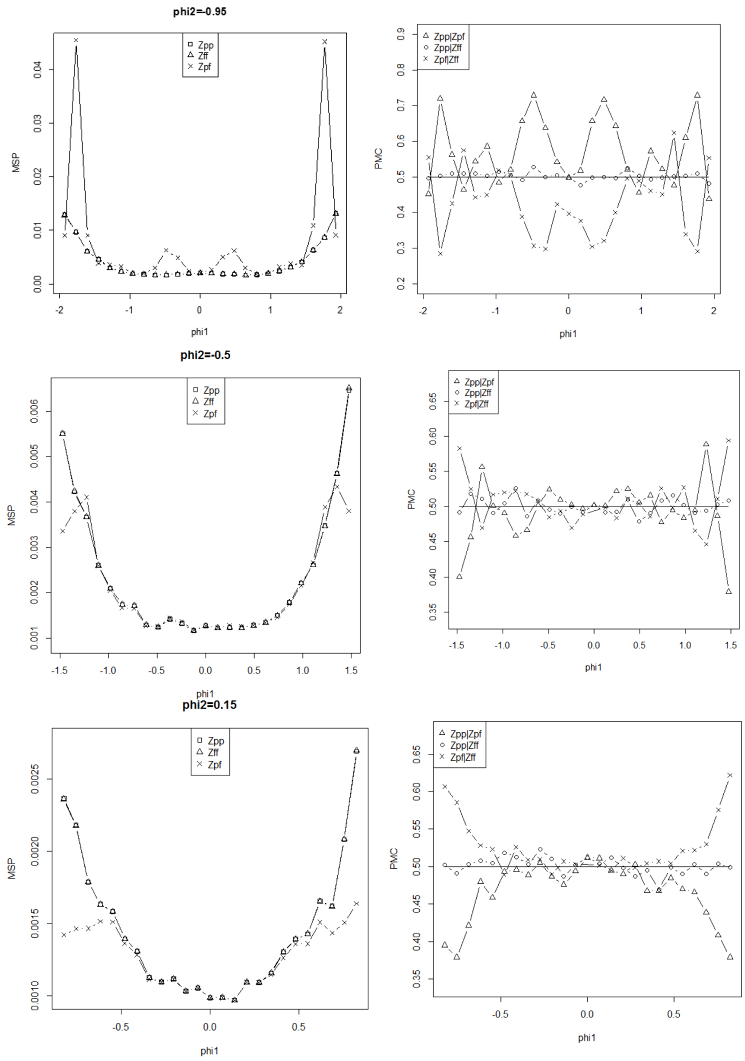

Figure 11. Comparison between predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $u = 3$ and $v = 4$ with Gaussian errors.

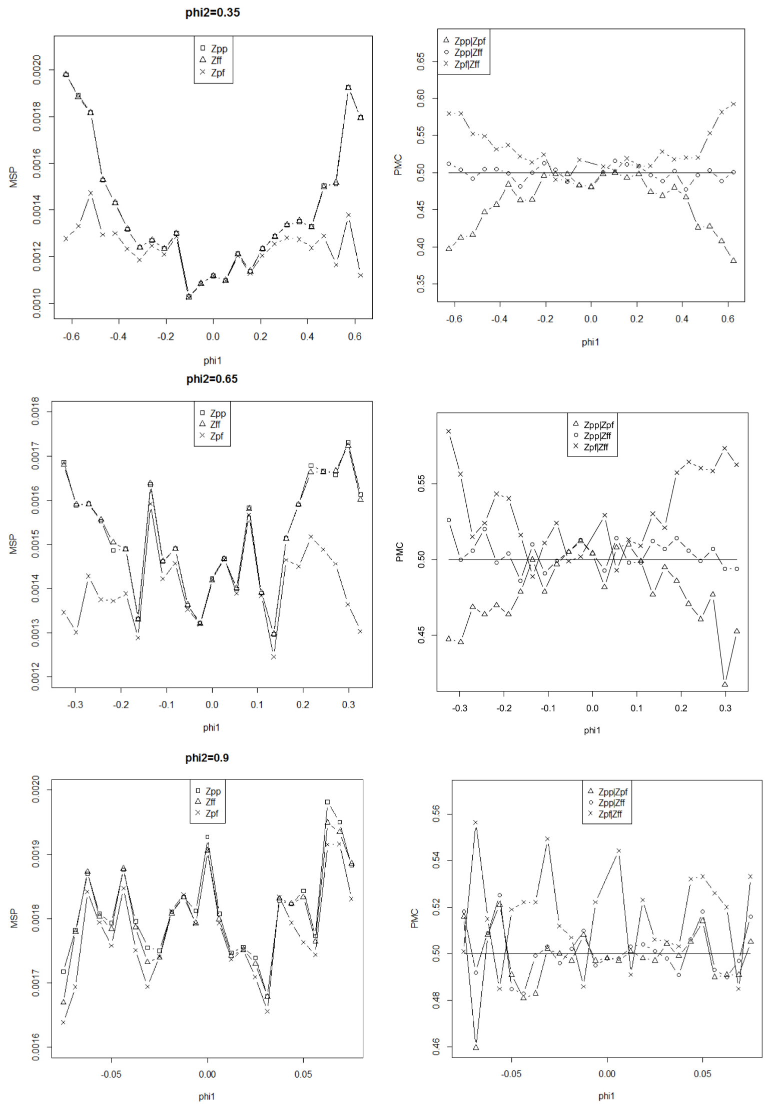

Figure 12. Comparison between predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $u = 3$ and $v = 4$ with Gaussian errors.

$(u, v) = (50, 80)$, prediction is not done for 160 observations. Our findings in Figure 8 express that predictors $\hat{Z}_t^{pp}$ and $\hat{Z}_t^{ff}$ are similar *w.r.t* both criteria MSP and PMC, similar with previous case. Also, applying predictor $\hat{Z}_t^{pf}$ instead $\hat{Z}_t^{ff}$ has better preponderance in following set.

$$\Theta_{pf|ff}^{G,\ (u,v)=(5,8)} \tag{24}$$
$$=\Theta_{stationary}\ \cap \{(-0.1 < \varphi_2 < 0.2\ and\ |\varphi_1| > |0.5 - \varphi_2|)\ or\ (0.2 \le \varphi_2 < 0.85\ )\}.$$

Finally, Figure 9 represent the achieved results in case $(u, v) = (50, 80)$. Again, there is no difference in using of one of the predictors $\hat{Z}_t^{pp}$ and $\hat{Z}_t^{ff}$. Also, predictor $\hat{Z}_t^{pf}$ has more reliable advantages in comparison with $\hat{Z}_t^{ff}$ in set.

$$\Theta_{pf|ff}^{G,\ (u,v)=(50,80)} \tag{25}$$
$$=\left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ; (-0.425 < \varphi_2 < -0.375\ )\ or\ \varphi_2 > 0.875\right\}.$$

Note 3.1. Figure 9 points out $\widehat{PMC}\ \left(\hat{Z}^{pp}\ \middle|\ \hat{Z}^{pf}\right) = \widehat{PMC}\ \left(\hat{Z}^{pp}\ \middle|\ \hat{Z}^{pf}\right) = 0$ for values of φ_1 and φ_2 which are in following set

$$\Theta_{NA}^{G,\ (u,v)=(50,80)} \tag{26}$$
$$=\left\{(\varphi_1, \varphi_2) \in \Theta_{stationary}\ ; -0.375 < \varphi_2 < 0.375 \text{ and } |\varphi_1| < |0.5 - \varphi_2|\ \right\}.$$

In fact, the values of $\left(\rho_u - \rho_v\rho_{u+v}\right)$, $\left(\rho_v - \rho_u\rho_{u+v}\right)$, $\left(\rho_u - \rho_v\rho_{v-u}\right)$ and $(\rho_v - \rho_u\ \rho_{v-u})$ are near to 0 for $(u, v) = (50, 80)$ in set (26). These values are coefficients of observations $Z_{t-u},\ Z_{t+v},\ Z_{t+u}$ and Z_{t+v} in predictors $\hat{Z}_t^{pf}$, $\hat{Z}_t^{ff}$ and $\hat{Z}_t^{pp}$.

Comparison among predictors $\hat{Z}_t^{pf}$, $\hat{Z}_t^{ff}$ and $\hat{Z}_t^{pp}$ with Gamma innovations has been accomplished, too. The results of this comparison are very similar with case of Gaussian distributed errors. Therefore, we omit them of paper in order miss of redundancy and saving space.

Remark 1. In application, parameters of model φ_1 and φ_2 and parameters in distribution of innovations are unknown. The achieved results in this study on the comparison among predictors are based on known parameters. Therefore, these findings cannot be used in the case of unknown parame-

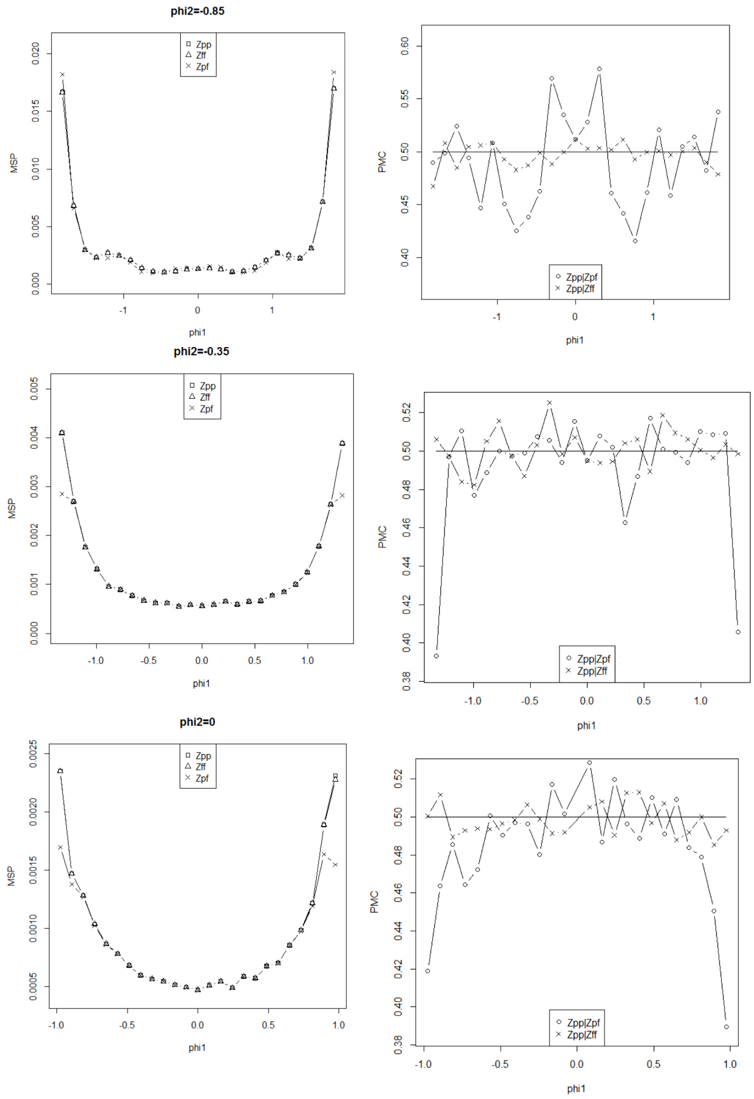

Figure 13. Comparison between predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $u = 5$ and $v = 8$ with Gaussian errors.

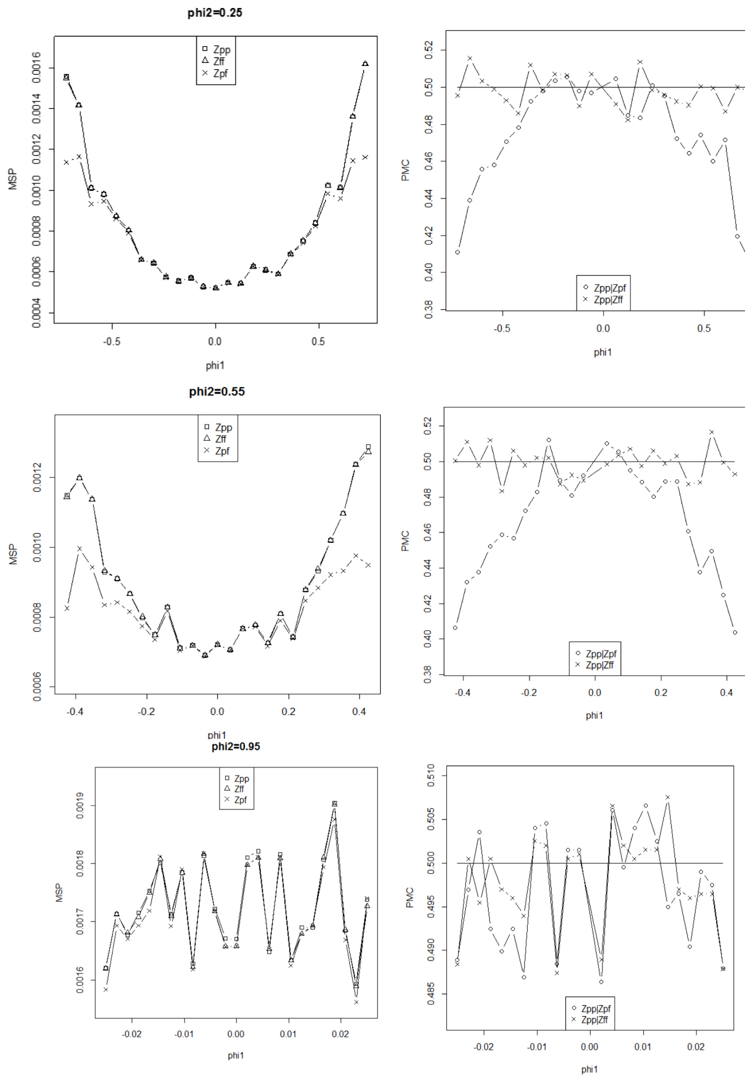

Figure 14. Comparison between predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $u = 5$ and $v = 8$ with Gaussian errors.

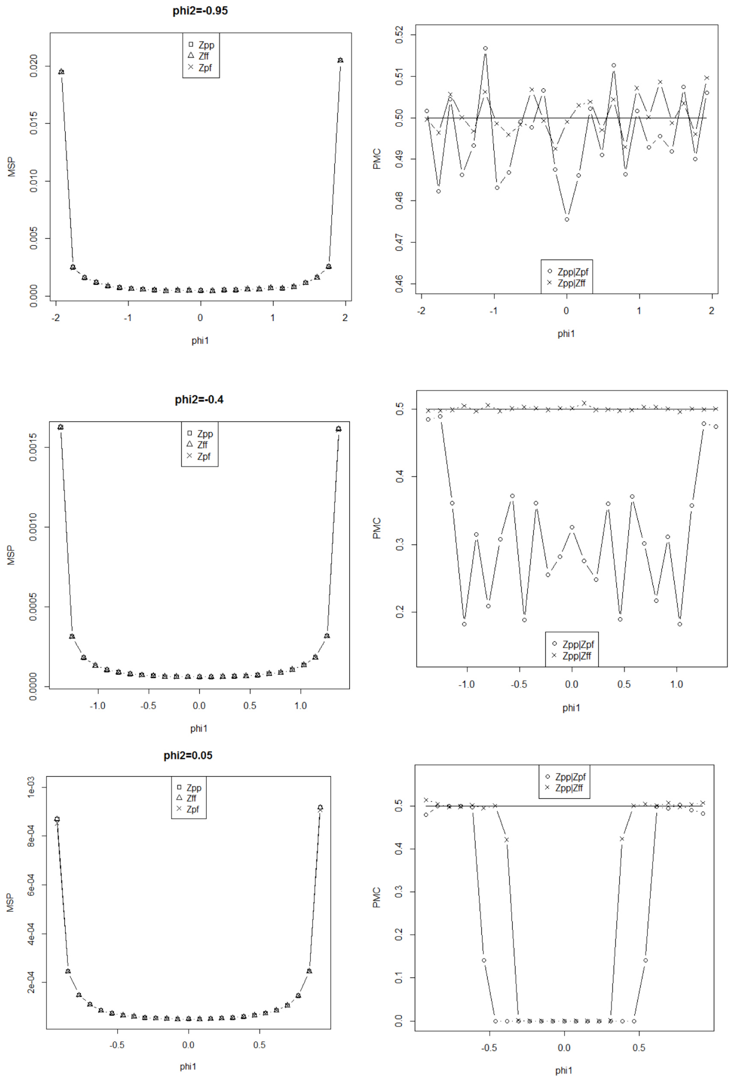

Figure 15. Comparison between predictors $\hat{Z}_t^k,\ k = pf, pp, ff$ for $u = 50$ and $v = 80$ with Gaussian errors.

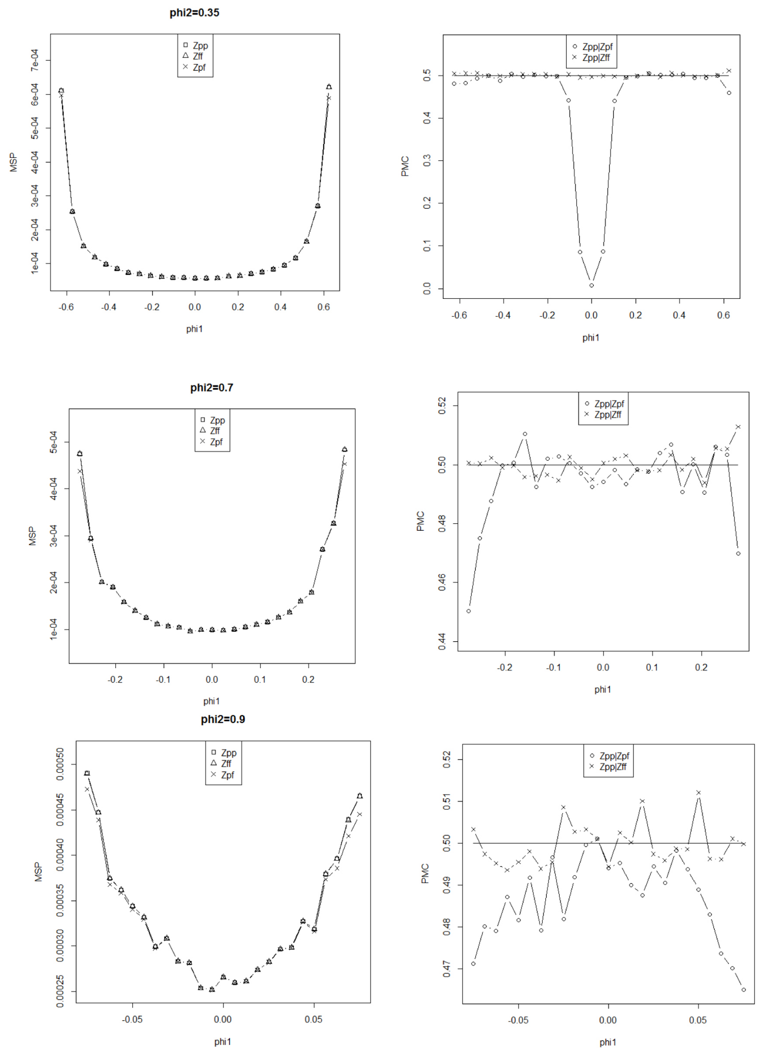

Figure 16. Comparison between predictors $\hat{Z}_t^k$, $k = pf, pp, ff$ for $u = 50$ and $v = 80$ with Gaussian errors.

Table 1. MSE and estimated parameters for $AR(2)$ model with Normal and Gamma distributed errors.

Errors	Data set	Number	LSE	φ_1	φ_2	α	λ	μ	σ^2
Normal	A	56	2.2732	1.0000	-0.2833	-	-	4.3333	0.25
Gamma			6.5861	1.4833	-0.5833	8.5000	5.6667	-	-
Normal	B	600	16.4461	1.4892	-0.5697	-	-	14.9931	1
Gamma			5.0287	0.5000	0.3500	4.2500	1.9000	-	-

ters. However, if estimated parameters are close to real parameters, they are approximately similar with previous case. As an instance, we compared predictors $\hat{Z}_t^{1,1}$, $\hat{Z}_t^{1,2}$, $\hat{Z}_t^{1,3}$and $\hat{Z}_t^{1,4}$ for unknown values of φ_1 and φ_2 and $\varepsilon \sim N(0\ ,\ 1)$ but without estimation of φ_1 and φ_2. We generated observations from model (1) with two fixed values of φ_1 and φ_2. Then, we performed prediction by predictor but with values $\varphi_1 + 0.05$ and $\varphi_2 + 0.05$. The results of comparison is approximately same as known parameters case.

3 Application to Real Data

In this section, we use two real data sets for illustrative and comparative purposes. Our used data are annual barley yields per acre in England and Wales (data set A) and monthly Lake Erie levels (data set B), respectively (See https://datamarket.com/data for link of data). Fitting these data to a $AR(2)$ model have been done based on Least Square Error (LSE) in R software. Our computations show that these two data sets satisfy $AR(2)$ model with Normal and Gamma distributed errors, respectively. The results of fitted models have been brought in Table 1.
The plots of ACF and PACF have been shown in Figure 10, too.
For applying Theorem 2.1 and Theorem 3.1 for prediction we must covert data to zero mean data. After converting to centered data, we can perform some predictions. Our findings in previous section denoted that predictor $\hat{Z}_t^{4,1}$ is better than all predictors $\hat{Z}_t^{k,l}$ and only exception is predictor $\hat{Z}_t^{2,3}$ for big values of φ_2. Since values of parameter φ_2 are not big here for both data sets, we use of predictor $\hat{Z}_t^{4,1}$ for prediction by near observations. By a similar argument, using of predictor $\hat{Z}_t^{pf}$ has preference than predictors $\hat{Z}_t^{ff}$ and $\hat{Z}_t^{pp}$ for prediction by far observation in these two data sets.

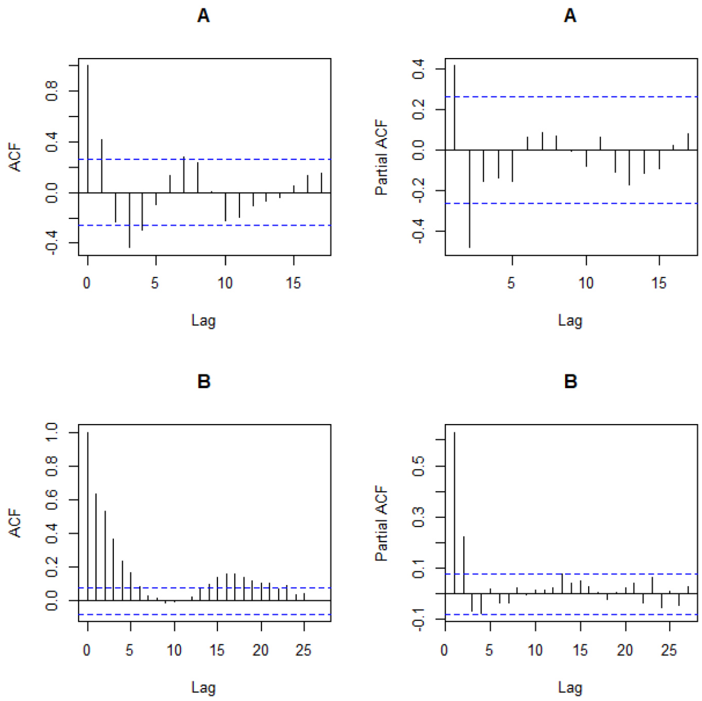

Figure 17. ACF and PACF plots of data sets A and B.

Table 2. Comparison between prediction in Normal and Gamma fitted models based on MSP and PMC criteria.

Data set	Used model	$\widehat{MSP}(\hat{Z}^{1,4})$	$\widehat{PMC}\left(\hat{Z}^{1,4,N}\vert\hat{Z}^{1,4,G}\right)$	$\widehat{MSP}(\hat{Z}^{pf})$	$\widehat{PMC}\left(\hat{Z}^{p,f,N}\vert\hat{Z}^{p,f,G}\right)$
A	Normal	1.733119	0.7143	1.1408	0.5417
	Gamma	2.042484		1.4877	
B	Normal	0.789951	0.4967	20.8240	0.1722
	Gamma	0.227606		1.6513	

In Table 2, the results of these predictions have been demonstrated. As this table demonstrate both criteria MSP and PMC denote Gamma distributed error $AR(2)$ model is better than Normal distributed error $AR(2)$ model in prediction of data set B. Although $\widehat{MSP}(\hat{Z}^{1,4})$ show bitterness of Gamma model with a high confidence, $\widehat{PMC}\left(\hat{Z}^{1,4,N}|\hat{Z}^{1,4,G}\right)$ has less confidence since it is near to 0.5.

This matter is different for prediction by means of far observation. In this case, both criteria absolutely show exacter prediction for Gamma model. This finding is reversed for data set A where Normal model has better results for prediction.

4 Discussion and Conclusion

Our results express that predictors $\hat{Z}_t^{1,1}$and $\hat{Z}_t^{1,4}$ have similar performances *w.r.t* both criteria MSP and PMC. This finding is remained same for other predictors which use of observations with symmetric indices around unsampled time t. These similar predictors are $\hat{Z}_t^{1,2}$and $\hat{Z}_t^{1,3}$, $\hat{Z}_t^{2,1}$and $\hat{Z}_t^{2,6}$, $\hat{Z}_t^{2,2}$and $\hat{Z}_t^{2,5}$, $\hat{Z}_t^{3,1}$and $\hat{Z}_t^{3,4}$ and $\hat{Z}_t^{3,2}$and $\hat{Z}_t^{3,3}$. In fact, for almost all values of φ_1 and φ_2 which satisfy stationary condition (2), we have $0.465 \leq \widehat{PMC}\left(\hat{Z}^{1,1}\,\middle|\,\hat{Z}^{1,4}\right) \leq 0.534$, $0.468 \leq \widehat{PMC}\left(\hat{Z}^{1,2}\,\middle|\,\hat{Z}^{1,3}\right) \leq 0.535$, $0.463 \leq \widehat{PMC}\left(\hat{Z}^{2,1}\,\middle|\,\hat{Z}^{2,6}\right) \leq 0.531$, $0.461 \leq \widehat{PMC}\left(\hat{Z}^{2,2}\,\middle|\,\hat{Z}^{2,5}\right) \leq 0.539$, $0.467 \leq \widehat{PMC}\left(\hat{Z}^{3,1}\,\middle|\,\hat{Z}^{3,4}\right) \leq 0.535$ and $0.467 \leq \widehat{PMC}\left(\hat{Z}^{3,2}\,\middle|\,\hat{Z}^{3,3}\right) \leq 0.533$.

After finding the best predictor in any cases between $\hat{Z}_t^{k,l}$ for fixed $k = 1, 2, 3$, we have compared the best predictors with $\hat{Z}_t^{4,1}$. Our results demonstrated predictor $\hat{Z}_t^{4,1}$ is approximately better than all predictors $\hat{Z}_t^{k,l}$ except for values of φ_1 and φ_2 which are near to 0. The only predictor which can compete with predictor $\hat{Z}_t^{4,1}$ is predictor $\hat{Z}_t^{2,3}$ for big values of φ_2. This result is logic since this predictor use from observations Z_{t-2} and Z_{t+2} whose coefficients in model (1) is φ_2.

Another work which has been done in this paper was comparison between predictors which use of far observations. The goal of paper in this part is comparing three classes of predictors. First two groups use from observations which are in the past or in future and last group use from observations both in past and future. Our findings have expressed that predictors which only use from observations in the past have similar performance with predictors which only use from observations in the future. Also, these two classes of predictors are better than the predictors which use from observations both in past and future approximately for only values of φ_2 which are near to -1. Therefore, using of third predictor is recommended in application.

This study is regarding to model $AR(2)$ while in the most of previous works model $AR(1)$ had been studied. A corresponding perusal on Moving Average (MA) models can be done, too. We are going to do this work for model $MA(1)$ in future work. A more general work which cover all previous works is study on ARMA models which we left them for future.

As we mentioned in Remark 1, in application, parameters of model φ_1 and φ_2 and parameters in distribution of innovations are unknown. Comparison in this realistic case can be done although results may be not same as present findings. By joint distribution of observations earned in Appendix A, MLEs of parameters can be computed and then a plug-in estimate is replaced in our introduced predictors. This is out of scope of this paper and needs to large space for presenting results. A new article may be written for this end.

References

Ailliot, P. (2006). Some Theoretical Results on Markov-switching Autoregressive Models with Gamma Innovations. *Statistics*, **343**, 271-274.

Brockwell, P.J., and Davis, R.A. (1991). *Time Series: Theory and Methods.* **2**nd ed. Springer-Verlag, New York.

Gourieroux, C., and Jasiak, J. (2006). Autoregressive Gamma Processes. *Journal of Forecasting*, **25**, 129-152.

Hamaz, A., and Ibazizen, M. (2009). Comparison of Two Estimation Methods of Missing

Values Using Pitman-closeness Criterion. *Communication in Statistics: Theory and Method*, **38**, 2210-2213..

Keating, J.P., Mason, R.L., and Sen, P.K. (1993). *Pitman's Measure of Closeness: A Comparison of Statistical Estimators,* SIAM, Philadelphia.

Saber, M.M., and Nematollahi, A.R. (2017). Comparison of Spatial Interpolation Methods in the First Order Stationary Multiplicative Spatial Autoregressive Models. *Communication in Statistics: Theory and Method*, **18**, 9230-9246.

Saber, M.M. (2017). Performance of Extrapolation Based on Pitman's Measure of Closeness in Spatial Regression Models with Extended Skew t Innovations. *Communication in Statistics: Theory and Method*, **0**, 1-18.

Saadatmand, A., Nematollahi, A.R., and Sadooghi-Alvandi, S. M. (2016). On the Estimation of Missing Values in AR(1) Model with Exponential Innovations. *Communication in Statistics: Theory and Method*, 3393-3400.

Senoglu, B., and Bayrak, O.T. (2016). Linear Contrasts in One-way Classification AR(1) Model with Gamma Innovations. *Hacettepe Journal of Mathematics and Statistics*, **45**, 1743-1754.

Appendix A

Here we present proof of identifiability of model (1) for both Normal and Gamma distributed errors. First of all model (1) is written in a matrix form as following.
$\boldsymbol{Z} = \boldsymbol{A}\ \boldsymbol{Z} + \boldsymbol{B}\ \boldsymbol{Z}_0 + \boldsymbol{\varepsilon}$ (A.1)
where $\boldsymbol{Z}^T = (\ Z_1\ , \ldots, Z_n\ )$ is vector of observations, $\boldsymbol{Z}_0^T = (\ Z_{-1}\ , Z_0\ )$ is vector of initial values, $\boldsymbol{\varepsilon}^T = (\varepsilon_1\ , \ldots, \varepsilon_n\ )$ and $\boldsymbol{A}$ and $\boldsymbol{B}$ are two $n \times n$ and $n \times 2$ matrices whose components are φ_1, φ_2 and 0. To see construction of these matrices more clearly, they have been demonstrated for case of $n = 4$ in following.

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \varphi_1 & 0 & 0 & 0 \\ \varphi_2 & \varphi_1 & 0 & 0 \\ 0 & \varphi_2 & \varphi_1 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} \varphi_2 & \varphi_1 \\ 0 & \varphi_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

From (A.1) we have $(\boldsymbol{I} - \boldsymbol{A}\ )\boldsymbol{Z} = \boldsymbol{B}\ \boldsymbol{Z}_0 + \boldsymbol{\varepsilon}$ and then
$\boldsymbol{Z} = \boldsymbol{M} + \boldsymbol{W}\ \boldsymbol{\varepsilon}$ (A.2)

where $\boldsymbol{M} = \boldsymbol{W}\ \boldsymbol{B}\ \boldsymbol{Z}_0$ is an $n-$dimensianal deterministic vector and $\boldsymbol{W} = (\boldsymbol{I} - \boldsymbol{A}\ )^{-1}$.
For case of Normal distributed errors which $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \boldsymbol{I})$ we have $\boldsymbol{Z} \sim N_n(\boldsymbol{M}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \boldsymbol{W}\boldsymbol{W}^T$. Let $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2)$ and suppose $f_{\boldsymbol{Z}|\boldsymbol{\varphi}}(\boldsymbol{z}) = f_{\boldsymbol{Z}|\boldsymbol{\theta}}(\boldsymbol{z})\ \forall\ \mathbf{z} \in \mathbb{R}^n$. Therefore, we have $\boldsymbol{M}(\boldsymbol{\varphi}) = \boldsymbol{M}(\boldsymbol{\theta})\ \forall\ \mathbf{z}_0 \in \mathbb{R}^2$ and $\boldsymbol{\Sigma}(\boldsymbol{\varphi}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})$ which results in $\boldsymbol{A}(\boldsymbol{\varphi}) = \boldsymbol{A}(\boldsymbol{\theta})$ and so $\boldsymbol{\varphi} = \boldsymbol{\theta}$.
Let $\varepsilon_i\ \sim Gamma(\alpha, \lambda)$ for known values of parameters α and λ. First, distribution of $\boldsymbol{\varepsilon}$ is computed by independent assumption, so

$$
\begin{aligned}
f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) = \prod_{i=1}^{n} f_{\varepsilon_i}\ (\varepsilon_i\ ) &= \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)^n \left(\prod_{i=1}^{n} \varepsilon_i\right)^{\alpha-1} e^{-\lambda \sum_{i=1}^{n} \varepsilon_i} \\
&= \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)^n e^{\mathbf{1}^T[(\alpha-1)\ln(\boldsymbol{\varepsilon})\ -\lambda\boldsymbol{\varepsilon}]}.
\end{aligned}
$$

Now, from (A.2) this is obviously seen that $\boldsymbol{\varepsilon} = \boldsymbol{W}^{-1}(\boldsymbol{z} - \boldsymbol{M}) = (\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{z} - \boldsymbol{B}\ \boldsymbol{z}_0$. Therefore,

$$
\begin{aligned}
f_{\boldsymbol{Z}}(\boldsymbol{z}) &= f_{\boldsymbol{\varepsilon}}((\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{z}-\boldsymbol{M}))\ \ |\boldsymbol{I}-\boldsymbol{B}| \\
&= \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)^n e^{\mathbf{1}^T[(\alpha-1)\ln((\boldsymbol{I}-\boldsymbol{A})\boldsymbol{z}-\boldsymbol{B}\ \boldsymbol{z}_0)\ -\lambda((\boldsymbol{I}-\boldsymbol{A})\boldsymbol{z}-\boldsymbol{B}\ \boldsymbol{z}_0)]}
\end{aligned}
$$

Again begin from this assumption that $f_{\boldsymbol{Z}|\boldsymbol{\varphi}}(\boldsymbol{z}) = f_{\boldsymbol{Z}|\boldsymbol{\theta}}(\boldsymbol{z})\ \forall\ \mathbf{z} \in \mathbb{R}^n$ which by some algebraic computation leads to $\boldsymbol{\varphi} = \boldsymbol{\theta}$.

Mohammad Mahdi Saber
Department of Statistics,
Higher Education Center of Eghlid,
Eghlid, Iran.
email: *mmsaber@eghlid.ac.ir*