Hindawi
Complexity
Volume 2022, Article ID 4997675, 14 pages
https://doi.org/10.1155/2022/4997675

WILEY | Hindawi

Research Article

Most Effective Sampling Scheme for Prediction of Stationary Stochastic Processes

Mohammad Mehdi Saber,[1] **Zohreh Shishebor**,[2] **M. M. Abd El Raouf**,[3] **E.H. Hafez**,[4] **and Ramy Aldallal**[5]

[1]*Department of Statistics, College of Sciences, Higher Education Center of Eghlid, Eghlid, Iran*
[2]*Department of Statistics, College of Sciences, Shiraz University, Shiraz, Iran*
[3]*Basic and Applied Science Institute, Arab Academy for Science, Technology and Maritime Transport (AASTMT), Alexandria, Egypt*
[4]*Faculty of science department of mathematics, Helwan University, Cairo, Egypt*
[5]*College of Business Administration in Hotat bani Tamim, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia*

Correspondence should be addressed to Mohammad Mehdi Saber; mmsaber@eghlid.ac.ir

Received 11 November 2021; Revised 14 April 2022; Accepted 28 April 2022; Published 9 July 2022

Academic Editor: Sheng Du

The problem of finding optimal sampling schemes has been resolved in two models. The novelty of this study lies in its cost efficiency, specifically, for the applied problems with expensive sampling process. In discussed models, we show that some observations counteract other ones in prediction mechanism. The autocovariance function of underlying process causes mentioned result. Our interesting result is that, although removing neutralizing observations convert sampling scheme to nonredundant case, it causes to worse prediction. A simulation study confirms this matter, too.

1. Introduction

Prediction is often the main goal in analyzing spatial processes and time series. It is widely used to making applicable decisions in a numerous scientific fields such as geology, biology, medicine, crime surveying, natural disasters, and so on (see for instance Isaaks and Srivastava [1] and Akselsson et al. [2]). The problem of finding an optimal sampling scheme is an important task of researchers, since in applied sciences often collecting of samples is a laborious and expensive process. We refer the interested readers to the works such as McBratney and Webster [3]; Zio et al. [4]; Xiao et al. [5]; and Sward et al. [6].

In this work, the authors attempt to consider some theoretical points in a branch of stochastic processes. In contrast, there are deep studies in existing stochastic processes. Among them, we refer to Xie et al. [7] and Cheng et al. [8]. In the first work, a memory-based event-triggered asynchronous control was addressed for semi-Markov switching systems. In the latter, the authors focus on static output feedback quantized control for fuzzy Markovian switching singularly perturbed systems with deception attacks. For other works, one can see Zhou et al. [9], Cheng et al. [10] and Xie et al. [7].

By Cressie [11], kriging is the most important prediction method for geostatistical data. For prediction in lattice data or discrete time series, we need to apply a model to data. Two predictors for a single missing value in a stationary autoregressive process of order one (AR(1)) model have been compared by Hamaz and Ibazizen [12] and Saadatmand et al. [13]. They considered the following model:

$$Z_t = \varphi Z_{t-1} + \varepsilon_t, t = 1, \ldots, n, |\varphi| < 1, \tag{1}$$

in which ε_t is a white noise process. Their comparison was done by Pitman's measure of closeness (PMC) criterion, They showed the best predictor depend only on two nearest observations. In spatial context, Saber and Nematollahi [14] and Saber [15] studied the stationary first-order multiplicative spatial autoregressive (MSAR(1)) model on a $m \times n$ lattice Λ, that is,

$$Z_{i,j} = aZ_{i-1,j} + bZ_{i,j-1} - abZ_{i-1,j-1} + \varepsilon_{ij}, = 1, \ldots, m, \quad j = 1, \ldots, n, |a| < 1, |b| < 1. \tag{2}$$

Saber and Nematollahi [14] performed a comparison among the three predictors in the above model. They showed that predictor which uses the quarter observations is better than predictor based on observations in the first neighborhood. Also, it is better than predictor based on observations in the nearest neighborhood wherever parameters a and b are near 0. However, for most values of a and b, the predictor which uses the nearest neighborhood observations is the best among three recommended predictors. Saber [15] compared interpolation and extrapolation in MSAR(1) model. Figure 1 represents schemes of used observations for these two predictors.

In Saber and Nematollahi [14], predictors use from all observations, while two predictors in Saber [15] do not use all observations. In fact, the interpolator which has been constructed on eight observations uses only two observations. Also, the extrapolator which is based on ten observations is constructed by just one observation. There are one or two observations which play an important role in the prediction process. It can be seen that the result of the prediction based on one basic and neutralizing observation is better than prediction based on a large number of nonessential observations. We show both models (1) and (2) are in this mentioned class.

So, this paper is organized as follows. In Section 2, the neutralizing observation for prediction in MSAR(1) model is achieved in some sampling schemes. A comparison between prediction by one neutralizing observation and prediction by eight usual observation has been performed in this section, too. Section 3 is the same as Section 2 in which the AR(1) model is in the role of MSAR(1) model. Finally, a discussion about redundant sampling is stated in Section 4.

2. Prediction in MSAR(1) Model

Suppose for any fixed values of i and j we want to predict $Z_{i,j}$ by a set of other observations $(\mathbf{Z}_{i,j})$ whose indices belong to $\mathbf{O}_{i,j}$ as a subset of lattice Λ. In other words, $\mathbf{O}_{i,j} = \{(k,l); Z_{k,l} \in \mathbf{Z}_{i,j}\}$ and $\mathbf{Z}_{i,j} = \{Z_{k,l}; (k,l) \in \mathbf{O}_{i,j}\}$. Also, the predictor of $Z_{i,j}$ in terms of the components of set $\mathbf{Z}_{i,j}$ is denoted by $\widehat{Z}_{i,j}$.

Theorem 1. *Let $Z_{i,j}$ satisfy (1) with $E(Z_{i,j}) = 0$ and $Var(Z_{i,j}) < \infty$. Then, the best linear predictor for $Z_{i,j}$ w.r.t mean square error (MSE) based on set $\mathbf{Z}_{i,j}$ is only dependent to Z_{s_0,t_0} and is given by*

$$\widehat{Z}_{i,j} = a^{|i-s_0|} b^{|j-t_0|} Z_{s_0,t_0}, \tag{3}$$

if there exists a location $(s_0, t_0) \in \mathbf{O}_{i,j}$ such that one of the following conditions is satisfied for all $(u, v) \in \mathbf{O}_{i,j}$.

$$i > u,\ j > v,\ s_0 \geq u, \text{and } t_0 \geq v, \tag{4}$$

$$i > u,\ j > v,\ s_0 \geq u, \text{and } t_0 \geq v, \tag{5}$$

$$i > u,\ j > v,\ s_0 \geq u, \text{and } t_0 \geq v, \tag{6}$$

$$i > u,\ j > v,\ s_0 \geq u, \text{and } t_0 \geq v, \tag{7}$$

$$i > u,\ j > v,\ s_0 \geq u, \text{and } t_0 \geq v, \tag{8}$$

$$i < u,\ s_0 \leq u, \text{and } t_0 = j, \tag{9}$$

$$i < u,\ s_0 \leq u, \text{and } t_0 = j, \tag{10}$$

$$i < u,\ s_0 \leq u, \text{and } t_0 = j. \tag{11}$$

Proof. Here, the best predictor is earned by projection theorem (see Brockwell and Davis [16] for details of this theorem). Consider this predictor in a linear form of observations $\widehat{Z}_{i,j} = \alpha \mathbf{Z}_{i,j}$ where $\mathbf{Z}_{i,j} = \{Z_{k,l}; (k,l) \in \mathbf{O}_{i,j}\}$ and $\alpha = \{\alpha_{k,l}; (k,l) \in \mathbf{O}_{i,j}\}$. By Saber [15], α is computed from the following equation:

$$\mathbf{A}\boldsymbol{\alpha} = \mathbf{B}, \tag{12}$$

in which $\mathbf{A} = \text{var}(\mathbf{Z}_{i,j})$ and $\mathbf{B} = \text{cov}(\mathbf{Z}_{i,j}, Z_{i,j})$; the first is covariance matrix of observations and the second is covariance vector between observations and unobserved variable, respectively. Regarding Gaetan and Guyon [17], $\text{cov}(Z_{s+h_1,t+h_2} Z_{s,t}) = \sigma_\varepsilon^2 / (b^2 - 1)(a^2 - 1)\, a^{|h_1|}\, b^{|h_2|}$ whose constants $\sigma_\varepsilon^2 / (b^2 - 1)(a^2 - 1)$ are removed from both sides of (12). Therefore, the general components of $\mathbf{A}$ and $\mathbf{B}$ are

$$a^{|s-u|} b^{|t-v|}, \quad a^{|s-i|} b^{|t-j|}, \quad \text{for all } (s,t) \text{and } (u,v) \in \mathbf{O}_{i,j}. \tag{13}$$

Equation (12) can be rewritten as follows:

$$\sum_{(k,l) \in O_{i,j}} \alpha_{k,l} \mathbf{C}_{k,l} = \mathbf{B}, \tag{14}$$

in which $C_{k,l}$ are the columns of $\mathbf{A}$. Since one of the columns of $\mathbf{A}$ is covariance between $\mathbf{Z}_{i,j}$ and Z_{s_0,t_0}, the latter equation can be written in the following form:

$$\alpha_{s_0,t_0} \mathbf{C}_{s_0,t_0} + \sum_{(k,l) \in O_{i,j}, (k,l) \neq (s_0,t_0)} \alpha_{k,l} \mathbf{C}_{k,l} = \mathbf{B}. \tag{15}$$

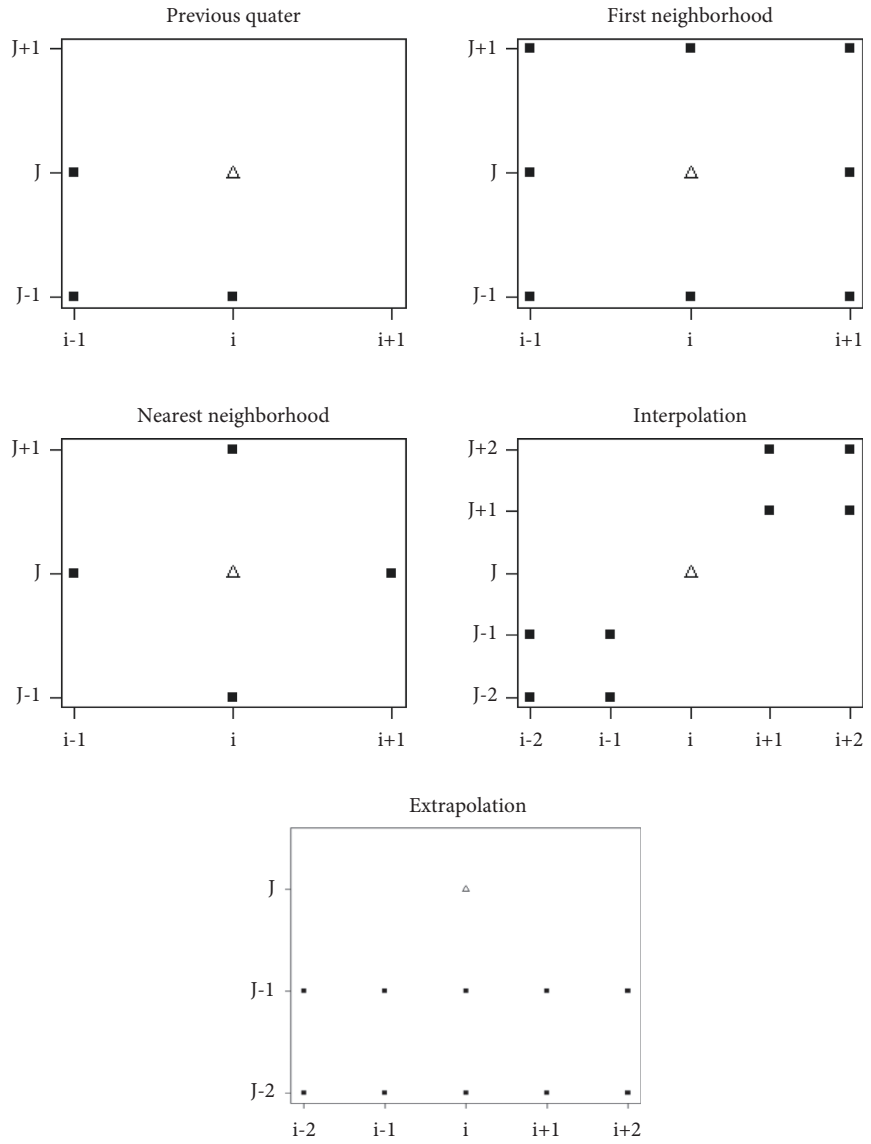

Figure 1: Locations of observations (■) and missing value (△). Upper figure shows three schemes for interpolation while lower figure shows interpolation and extrapolation.

On the other hand, assumptions (3)–(10) along with (13) lead to

$$cov\left(Z_{k,l}, Z_{i,j}\right) = cov\left(Z_{k,l}, Z_{s_0,t_0}\right)cov\left(Z_{s_0,t_0}, Z_{i,j}\right)\forall (k,l) \in \mathbf{O}_{i,j}, \tag{16}$$

which gives

$$\mathbf{B} = \mathbf{C}_{s_0,t_0} a^{|i-s_0|} b^{|j-t_0|}. \tag{17}$$

By substituting (17) in (15), we have

$$\alpha_{s_0,t_0}\mathbf{C}_{s_0,t_0} + \sum_{(k,l)\in\mathbf{O}_{i,j},(k,l)\neq(s_0,t_0)} \alpha_{k,l}\mathbf{C}_{k,l} = a^{|i-s_0|} b^{|j-t_0|}\mathbf{C}_{s_0,t_0}, \tag{18}$$

and the solution is

$$\alpha_{s_0,t_0} = a^{|i-s_0|} b^{|j-t_0|} \text{and } \alpha_{k,l} = 0. \forall (k,l) \neq (s_0,t_0). \tag{19}$$

This completes the proof.

To understand the concept of equations (4)–(7), some examples are given in Figure 2. Conditions (4)–(11) are available in a sampling process. In Saber [15], $\mathbf{O}_{i,j}^2$ has been defined as $\mathbf{O}_{i,j}^2 = \{(s,t); |s-i| \leq 2, t-j = -1 \text{or} -2\}$ which satisfies condition (11) for $(s_0,t_0) = (i, j-1)$. Therefore, we see that the predictor $\widehat{Z}_{i,j}^2$ is only dependent on $Z_{i,j-1}$ observation. □

2.1. Comparison with Unredundant Sampling. Some of the observations, which satisfy the mentioned cases, do not have an effect on prediction.Therefore, an important question is that with a fixed number of observations which sampling

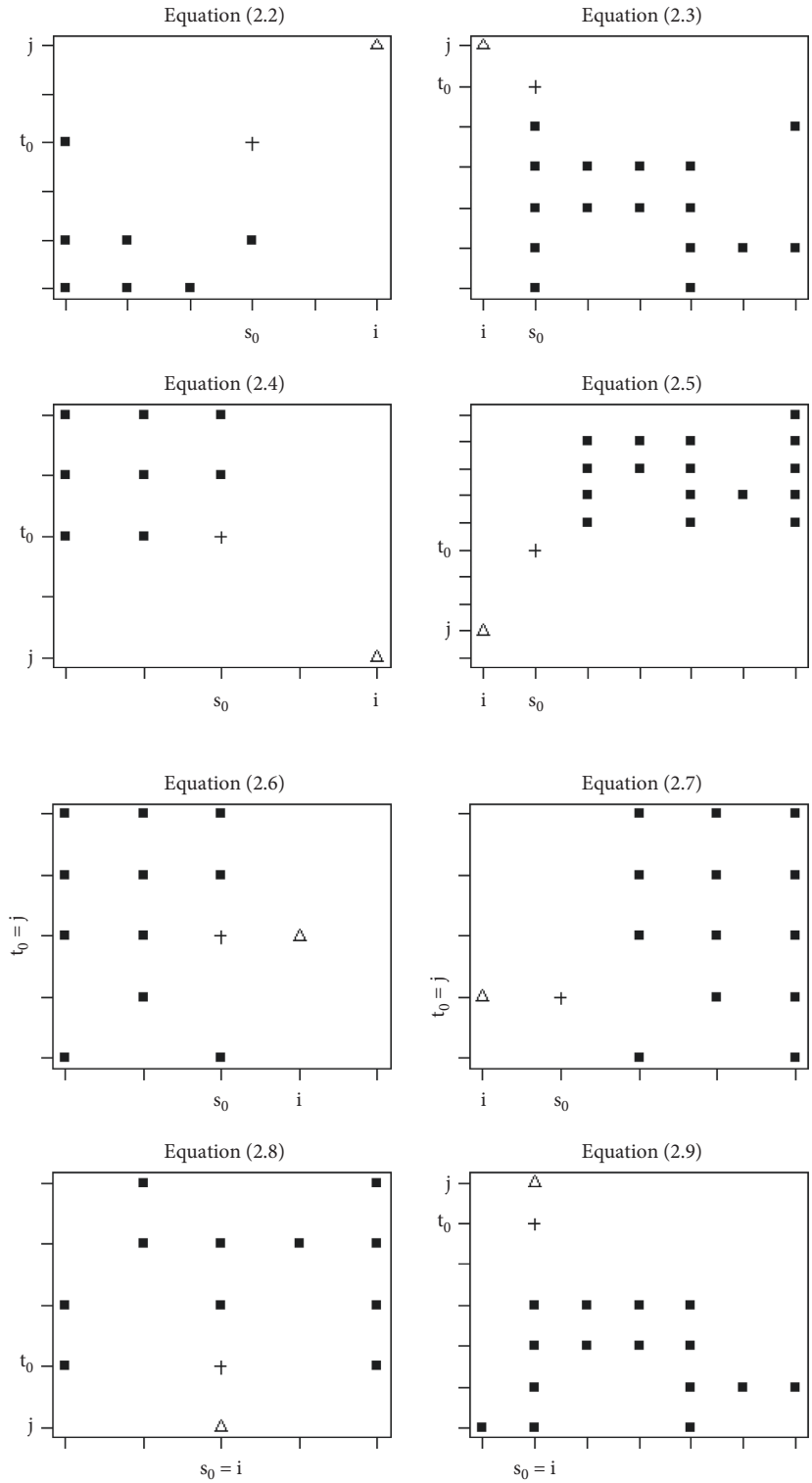

Figure 2: Locations of neutralized observations (△), neutralizing observation (+), and missing value (■).

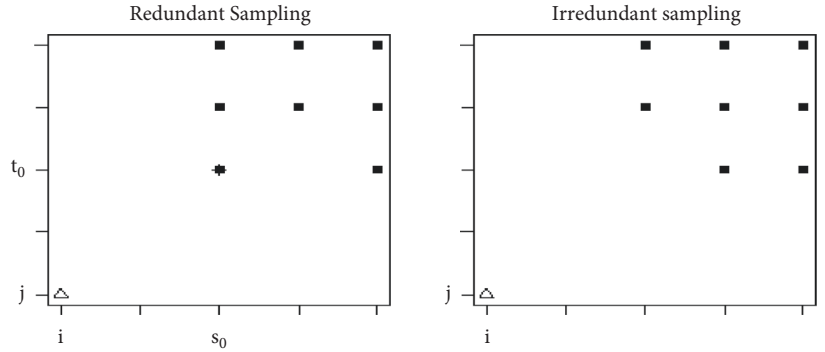

Figure 3: Two schemes for redundant sampling and unredundant sampling with eight observations.

scheme is better: sampling with redundant observation or the other?

In other words, is it logical that we remove neutralizing observations in the prediction process? In order to find answer of this question we compare these two methods in this section. To this end, two sampling schemes are considered in Figure 3.

$$\mathbf{A}=\begin{pmatrix}\gamma(0,0) & \gamma(0,1) & \gamma(1,1) & \gamma(1,0) & \gamma(1,1) & \gamma(2,1) & \gamma(2,0) & \gamma(2,1)\\ \gamma(0,1) & \gamma(0,0) & \gamma(1,2) & \gamma(1,1) & \gamma(1,0) & \gamma(2,2) & \gamma(2,1) & \gamma(2,0)\\ \gamma(1,1) & \gamma(1,2) & \gamma(0,0) & \gamma(0,1) & \gamma(0,2) & \gamma(1,0) & \gamma(1,1) & \gamma(1,2)\\ \gamma(1,0) & \gamma(1,1) & \gamma(0,1) & \gamma(0,0) & \gamma(0,1) & \gamma(1,1) & \gamma(1,0) & \gamma(1,1)\\ \gamma(1,1) & \gamma(1,0) & \gamma(0,2) & \gamma(0,1) & \gamma(0,0) & \gamma(1,2) & \gamma(1,1) & \gamma(1,0)\\ \gamma(2,1) & \gamma(2,2) & \gamma(1,0) & \gamma(1,1) & \gamma(1,2) & \gamma(0,0) & \gamma(0,1) & \gamma(0,2)\\ \gamma(2,0) & \gamma(2,1) & \gamma(1,1) & \gamma(1,0) & \gamma(1,1) & \gamma(0,1) & \gamma(0,0) & \gamma(0,1)\\ \gamma(2,1) & \gamma(2,0) & \gamma(1,2) & \gamma(1,1) & \gamma(1,0) & \gamma(0,2) & \gamma(0,1) & \gamma(0,0)\end{pmatrix}. \tag{20}$$

By Theorem 1, the predictor which has based on redundant sampling is

$$\widehat{Z}_{i,j}^{R}=a^2b^2Z_{i+2,j+2}, i=1,\ldots,m-2, j=1,\ldots,n-2. \tag{21}$$

By the same method as Saber [15], the other predictor based on unredundant sampling is

$$\widehat{Z}_{i,j}^{IR}=\mathbf{A}^{-1}\mathbf{B}\mathbf{Z}_{i,j}^{IR}, i=1,\ldots,m-4, j=1,\ldots,n-4, \tag{22}$$

where $\mathbf{Z}_{i,j}^{IR}=(Z_{i+2,j+3},Z_{i+2,j+4},Z_{i+3,j+2},Z_{i+3,j+3},Z_{i+3,j+4},Z_{i+4,j+2},Z_{i+4,j+3},Z_{i+4,j+4})$.

$\mathbf{B}^T=(\gamma(2,3)\gamma(2,4)\gamma(3,2)\gamma(3,3)\gamma(3,4)\gamma(4,2)\gamma(4,3)\gamma(4,4))$ and $\gamma(u,v)=a^u b^v$. After a cumbersome computation, we have

$$\begin{aligned}\widehat{Z}_{i,j}^{IR}&=a^2b^3Z_{i+2,j+3}+a^3b^2Z_{i+3,j+2}-,a^3b^3Z_{i+3,j+3},\\ &\quad i=1,\ldots,m-4, j=1,\ldots,n-4.\end{aligned} \tag{23}$$

A theoretical comparison of $\widehat{Z}_{i,j}^{R}$ and $\widehat{Z}_{i,j}^{IR}$ is not possible, so to compare them, we use $\widehat{\mathrm{MSP}}_Z(\widehat{Z})=\sum_{i,j}^{m,n}(\widehat{Z}_{i,j}-\mathrm{Z_{i,j}})^2/mn$ and $\widehat{\mathrm{PMC}}_Z(\widehat{Z}|\widetilde{Z})=\sum_{i,j}^{m,n}U(|\widetilde{Z}_{i,j}-Z_{i,j}|-|\widehat{Z}_{i,j}-\mathrm{Z_{i,j}}|)/\mathrm{mn}$ estimators. Here $U(t)$ is 1 for positive t and 0 otherwise.

Two well-known distributions normal ($N(0,1)$) and exponential ($E(1)$) for errors have been used. By Saber and Nematollahi [14], in case of $E(1)$ errors, we apply the mean-zero variables $Z_{i,j}^{*}=Z_{i,j}-1$.

Now, we simulate random variables $Z_{i,j}$ with Normal distributed errors on a 200×200 lattice and compute $\widehat{\mathrm{MSP}}_Z(\widehat{Z}^R)$, $\widehat{\mathrm{MSP}}_Z(\widehat{Z}^{IR})$, and $\widehat{\mathrm{PMC}}_Z(\widehat{Z}^R|\widehat{Z}^{IR})$. All the findings are demonstrated in Figures 4 and 5, respectively. In both figures, one can see that $\widehat{\mathrm{MSP}}_Z(\widehat{Z}^R)\le\widehat{\mathrm{MSP}}_Z(\widehat{Z}^{IR})$ and $\widehat{\mathrm{PMC}}_Z(\widehat{Z}^R|\widehat{Z}^{IR})\ge0.5$ for almost all values of parameters a and b. For some values of these parameters which are less than 0.25, $\widehat{\mathrm{PMC}}_Z(\widehat{Z}^R|\widehat{Z}^{IR})<0.5$. In these cases, $\widehat{\mathrm{PMC}}_Z(\widehat{Z}^R|\widehat{Z}^{IR})>0.495$, so they are negligible. In other words, approximately we have $0\le\widehat{\mathrm{MSP}}_Z(\widehat{Z}^R)-\widehat{\mathrm{MSP}}_Z(\widehat{Z}^R)<1.75$ and $0.495\le\widehat{\mathrm{PMC}}_Z(\widehat{Z}^R|\widehat{Z}^{IR})<0.66$ for all values of parameters a and b. These both criteria are increasing functions w.r.t parameters a and b. These findings state that the predictor $\widehat{Z}^R$ is better than predictor $\widehat{Z}^{IR}$ with respect to both criteria MSP and PMC. Therefore, for predicting $Z_{i,j}$ in model (2), using one observation $Z_{i+2,j+2}$ leads to better results than using 8 observations in $\mathbf{Z}_{i,j}^{IR}$. When parameters a and b are near 1, recommendation for using of this one observation rather than 8 observations in $\mathbf{Z}_{i,j}^{IR}$ becomes more serious.

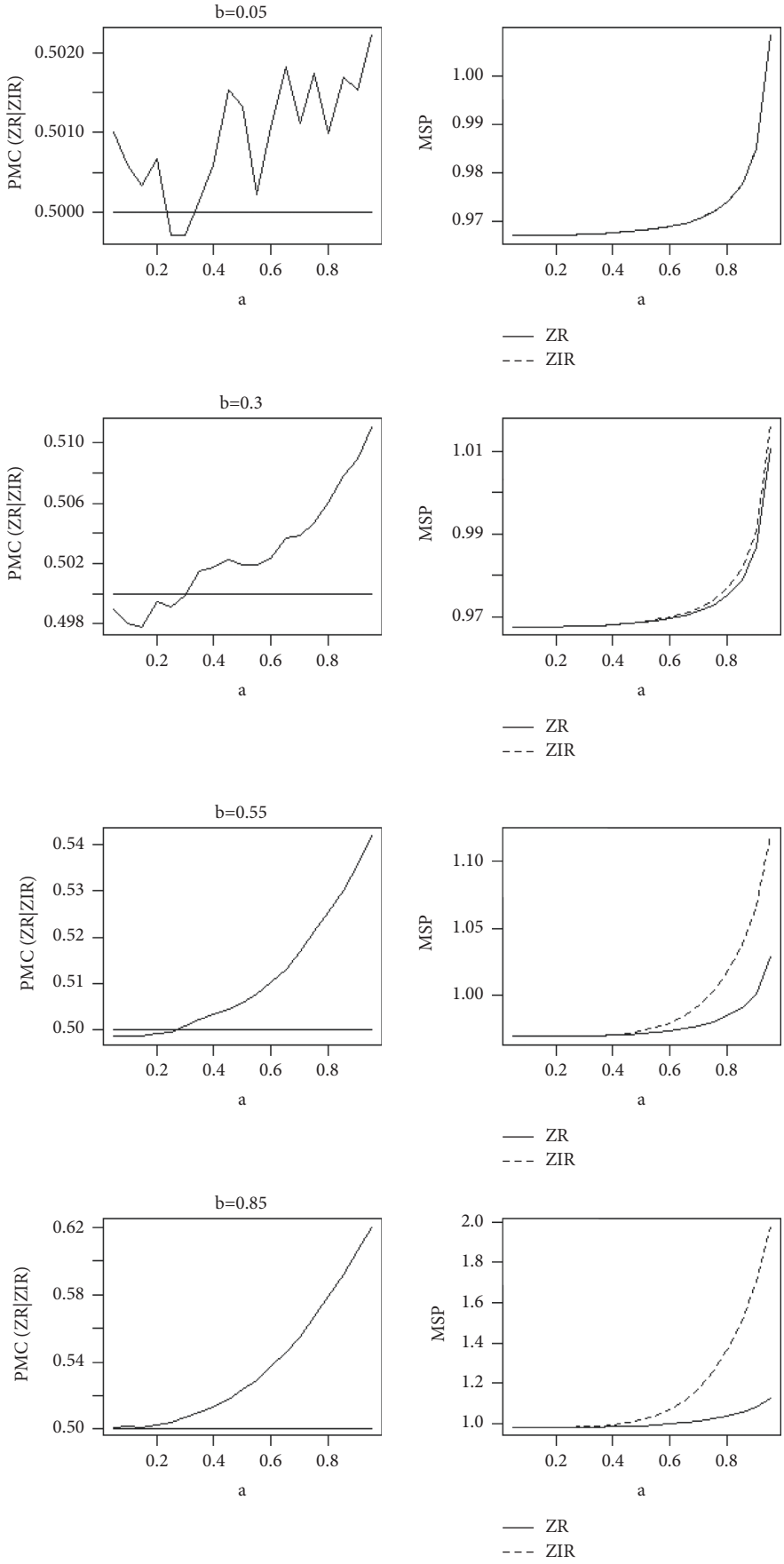

Figure 4: Comparison between $\widehat{Z}^{R}$ and $\widehat{Z}^{IR}$ for some fixed a and all b. Here, innovations have normal distribution.

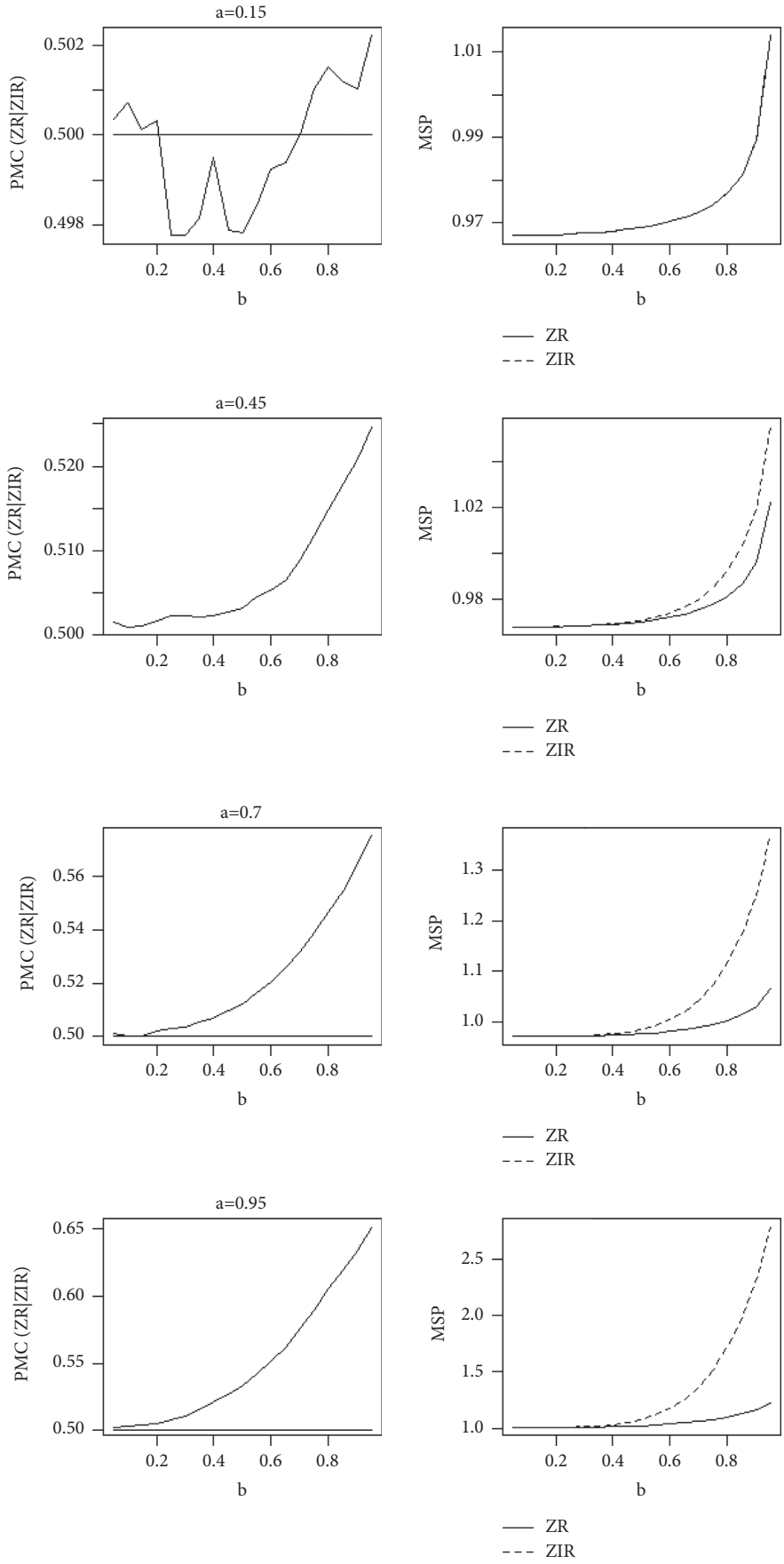

Figure 5: Comparison between $\widehat{Z}^{IR}$ and $\widehat{Z}^{IR}$ with normal errors for some fixed b and all a.

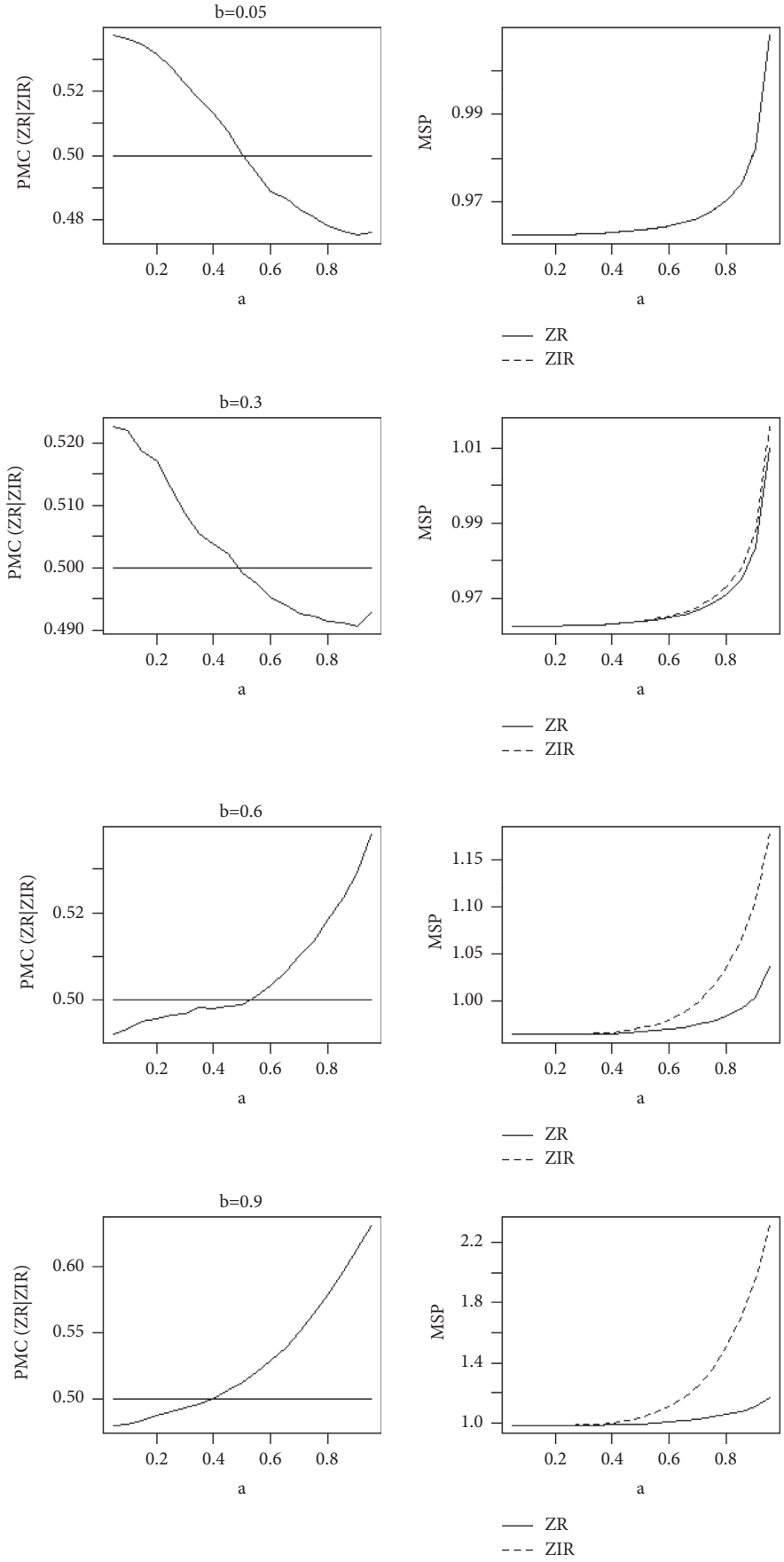

Figure 6: Comparison between $\widehat{Z}^R$ and $\widehat{Z}^R$ with exponential errors for some fixed a and all b.

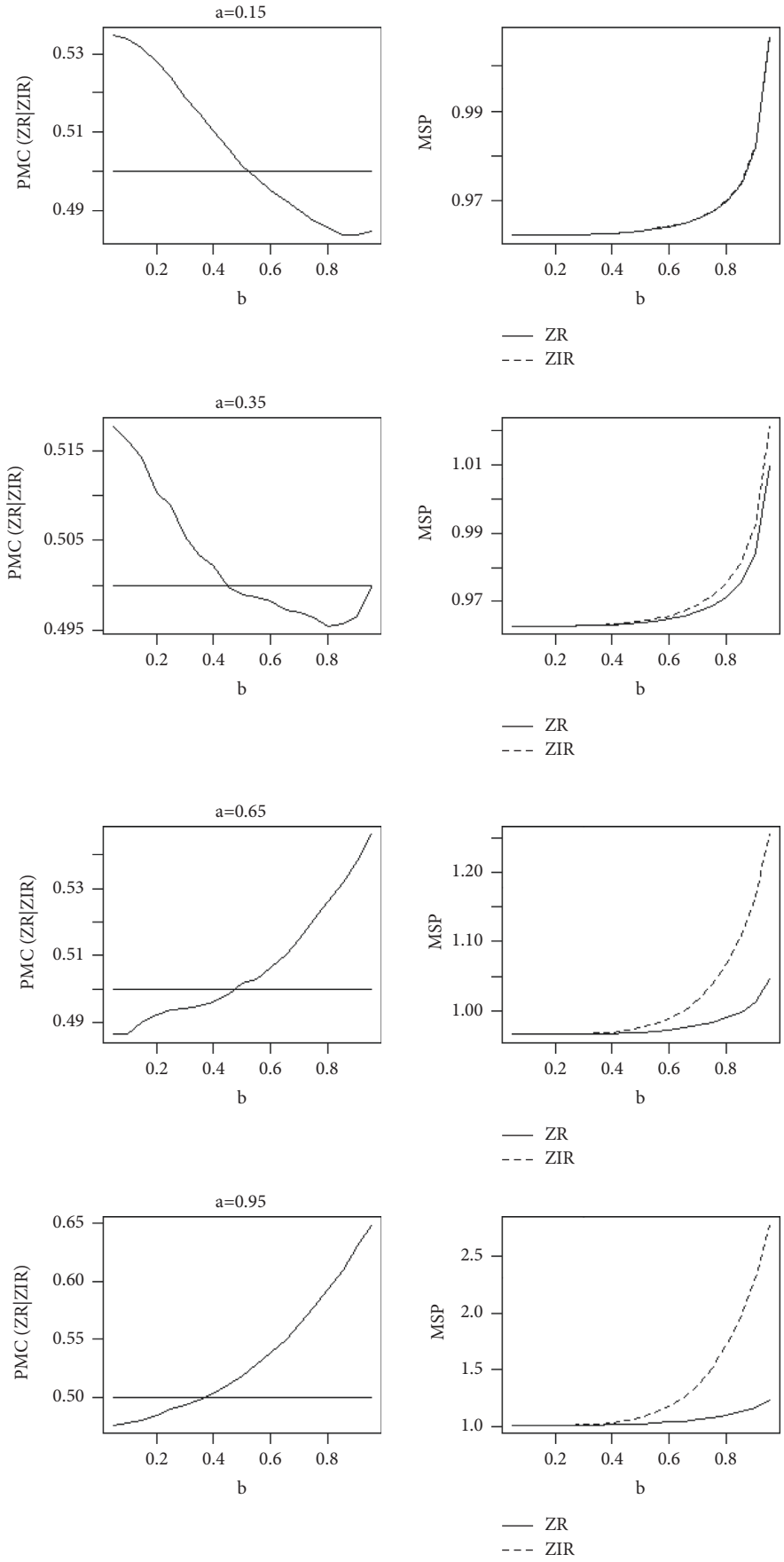

Figure 7: Comparison between $\widehat{Z}^{R}$ and $\widehat{Z}^{IR}$ with exponential errors for some fixed b and all a.

The similar work with normal errors has been done for case of exponential distributed errors with parameter 1. For exponential case, our findings demonstrate a bit difference with normal case. Results of this simulation are displayed in Figures 6 and 7. These figures show approximately $0 \leq \widehat{\mathrm{MSP}}_Z(\widehat{Z}^{IR}) - \widehat{\mathrm{MSP}}_Z(\widehat{Z}^{R}) < 1.6$ and $0.47 \leq \widehat{\mathrm{PMC}}_Z(\widehat{Z}^{R} | \widehat{Z}^{IR}) < 0.65$ for all values of parameters a and b. These findings state that the predictor $\widehat{Z}^{R}$ is better than predictor $\widehat{Z}^{IR}$ with respect to criterion MSP. Regarding PMC criterion, we can see that the predictor $\widehat{Z}^{IR}$ is better than predictor $\widehat{Z}^{R}$ whenever parameters a and b are near 0.5. However, the latter result is not significant.

Finally, we conclude that removing observation which has caused other observations be redundant in prediction does not lead to better results in almost all situations.

3. Prediction in AR(1) Model

In this section, we show that the best linear predictor at time t for stochastic process AR(1) uses at most one observation in every side of time t. First of all, for any fixed time t, define vector $\mathbf{Z}_T = \{Z_k; k \in \mathbf{O}_t\}$ where $\mathbf{O}_t$ is a subset of natural numbers ($\mathbb{N}$).

In the following theorem, we demonstrate the best linear predictor in this model based on n samples uses at least 2 observations.

Theorem 2. *Let Z_t come from model (1) with $E(Z_t) = 0$ and $Var(Z_t) < \infty$ and there exists a time $t_0 \in \mathbf{O}_t$ such that for all observed times $i \in \mathbf{O}_t$,*

$$i < t, \text{and } i \leq t_0, \tag{24}$$

or

$$i > t, \text{and } i \geq t_0. \tag{25}$$

Then, the best linear predictor for Z_t w.r.t criterion MSP based on variables $\mathbf{Z}_t = \{Z_k; k \in \mathbf{O}_t\}$ is only dependent to Z_{t_0} and

$$\widehat{Z}_t = \varphi^{|t-t_0|} Z_{t_0}. \tag{26}$$

Proof. The proof is similar with the proof of Theorem 1 in which $\widehat{Z}_t = \alpha \mathbf{Z}_t$ where $\alpha = \{\alpha_k; k \in \mathbf{O}_t\}$, $\mathbf{A} = \mathrm{var}(\mathbf{Z}_t)$, $\mathbf{B} = \mathrm{cov}(\mathbf{Z_t}, Z_t)$, and $\mathrm{cov}(Z_{s+h}, Z_s) = \sigma_\varepsilon^2 / 1 - \varphi^2 \varphi^{|h|}$.

This matter shows that this is not required to do a corresponding comparison with Section 2.1. By Theorem 2 and above paragraph, the best linear predictors based on observations $\mathbf{Z}_t$ and $\mathbf{Z}_t^{\mathrm{new}} = \{Z_k; k \in \mathbf{O}_t - t_0\}$ are $\widehat{Z}_t = \varphi^{|t-t_0|} Z_{t_0}$ and $\widehat{Z}_t^{\mathrm{new}} = \varphi^{|t-t_1|} Z_{t_1}$, respectively. Here, t_1 is the nearest time to time t among all times in $\mathbf{O}_t - t_0$. Clearly, $|t - t_1| > |t - t_0|$, and hence the predictor $\widehat{Z}_t$ is better than predictor $\widehat{Z}_t^{\mathrm{new}}$.

In the end, we give a theorem with which Theorems 1 and 2 can be expressed as its special cases. □

Theorem 3. *Let $\{Z_s; \mathbf{s} \in \mathbf{S}\}$, $\mathbf{S} \subseteq R^k, k \in \mathbb{N}$ be a stationary stochastic process with $E(Z_s) = 0$ and $Var(Z_\mathbf{s}) < \infty$. Then, the best linear predictor for $Z_\mathbf{s}$ based on $\mathbf{Z_s} = \{Z_\mathbf{v}; \mathbf{v} \in \mathbf{O_s}\}$ with respect to MSE is $\widehat{Z}_\mathbf{s} = cov(Z_{\mathbf{s}_0}, Z_s) Z_{\mathbf{s}_0}$ if there exists a point $\mathbf{s}_0 \in \mathbf{O_s}$ such that*

$$\mathrm{cov}(Z_u, Z_s) = \mathrm{cov}\left(Z_u, Z_{\mathbf{s}_0}\right)\mathrm{cov}\left(Z_{\mathbf{s}_0}, Z_s\right) \forall \mathbf{u} \in \mathbf{O_s}. \tag{27}$$

The proof of theorem is not presented, since it is point to point similar to the proof of Theorem 1. Notice that both Theorems 1 and 2 are special cases of Theorem 3 when $k = 2$ and $k = 1$, respectively.

4. Conclusion

In this paper, three theorems were presented which are useful for deriving optimal sampling scheme. In Theorem 1, the design of an optimal sampling scheme is given for model (2). A similar result for model (1) is found in Theorem 2.

In fact, we have tried to give an answer to the following question: for a fixed set of data collected through observations, if we would like to find the best linear predictor of a missing value, which of the following sampling schemes is logical?

(1) Should we use neutralizing observations?

(2) Should we remove neutralizing observations and use remaining observations?

The results show that the first scenario is more efficient in prediction process. Indeed, we have shown that some samples have no effect on prediction in discussed models. The achievement of this work will be useful, since it might not be possible to access a list of elements in many practical situations and several types of populations, so the use of an element as a sampling unit is not applicable.

There are other complicated situations than those described in Theorem 1. They lead to redundant sampling in mechanism of prediction, too. This is left as future work. Also, in Saber and Khorshidian [18], the best predictor in stationary first-order moving average model has been represented. So, another interesting study may be searching and exploring neutralizing and neutralized samples in moving average models.

Appendix

Comparison between redundant and nonredundant sampling.

```
rm(list = ls())
m < −200; n < −200
z0 < -rnorm(m + n + 1)
epsilon < -matrix(rnorm((m + 1)∗(n + 1)),m + 1,n + 1)
A < -seq(.05,.96,.05)
B < -seq(.05,.96,.05)
msezr < -msezir < -matrix(0,length(A),length(B))
pzrzir < -matrix(0,length(A),length(B))
for(a in A){
for(b in B){
g < -function(x,y) âx∗bŷ
```

```
zAR < -matrix(0,m + 1,n + 1)
zAR[1,]<-z0[1:(n + 1)]
zAR[−1,1]<-z0[(n + 2):(n + m + 1)]
for(i in 2:(m + 1)){
for(j in 2:(n + 1)){
zAR[i,j]<-zAR[i − 1,j]∗a + zAR[i,j − 1]∗b + zAR
[i − 1,j − 1]∗-(a∗b)
}}
z < -zAR + epsilon
z < -z[−1,−1]
fZR < -function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) y < −((a∗b)^2/2)∗z
[i + 2,j + 2]
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
return(y)}
fZIR < −function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) {
A1<-
matrix(c(g(0,0),g(0,1),g(1,1),g(1,0),g(1,1),g(2,1),g(2,0),-
g(2,1),
  g(0,1),g(0,0),g(1,2),g(1,1),g(1,0),g(2,2),g(2,1),g(2,0),
  g(1,1),g(1,2),g(0,0),g(0,1),g(0,2),g(1,0),g(1,1),g(1,2),
  g(1,0),g(1,1),g(0,1),g(0,0),g(0,1),g(1,1),g(1,0),g(1,1),
  g(1,1),g(1,0),g(0,2),g(0,1),g(0,0),g(1,2),g(1,1),g(1,0),
  g(2,1),g(2,2),g(1,0),g(1,1),g(1,2),g(0,0),g(0,1),g(0,2),
  g(2,0),g(2,1),g(1,1),g(1,0),g(1,1),g(0,1),g(0,0),g(0,1),
  g(2,1),g(2,0),g(1,2),g(1,1),g(1,0),g(0,2),g(0,1),g(0,0)),
8,8)
B1<-
matrix(c(g(2,3),g(2,4),g(3,2),g(3,3),g(3,4),g(4,2),-
g(4,3),g(4,4)),8,1)
Q1<-c(z[i + 2,j + 3],z[i + 2,j + 4],z[i + 3,j + 2],z
[i + 3,j + 3],z[i + 3,j + 4],z[i + 4,j + 2],
z[i + 4,j + 3],z[i + 4,j + 4])
y < -t(solve(A1,B1))%∗%Q1 }
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
y < -as.vector(y)
return(y)}
fZIR1<-function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) y < −(a^2∗b^2)∗(b∗z
[i + 2,j + 3] + a∗z[i + 3,j + 2]−a∗b∗z[i + 3,j + 3])
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
return(y)}
zr < -matrix(0,m,n)
zir < -matrix(0,m,n)
for(i in 1:m){
for(j in 1:n){
zr[i,j]<-fZR(i,j)
zir[i,j]<-fZIR1(i,j) }}
msezr[abs(a)∗20,abs(b)∗20]<-mean((zr-z)^2)
msezir[abs(a)∗20,abs(b)∗20]<-mean((zir-z)^2)
pzrzir[abs(a)∗20,abs(b)∗20]<-sum((abs(zr-z))<(abs(-
zir-z)))/sum((abs(zr-z))!=(abs(zir-z))) }}
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj
BB < -grid$yj
persp(A,B,msezr, theta = 135,phi = 5,scale = TRUE,
expand = 1,col = "green," ltheta = 0,
lphi = 0,box = T,ticktype = "simple,"xlab = "a,"
ylab = "b," zlab = "MSE(Z1)")
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj
BB < -grid$yj
persp(A,B,msezir, theta = 155,phi = 5,scale = TRUE,
expand = 1,col = "green," ltheta = 0,
lphi = 0,box = T,ticktype = "simple," xlab = "a,"
ylab = "b," zlab = "MSE(Z2)")
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj
BB < -grid$yj
persp(A,B,pzrzir, theta = 115,phi = 5,scale = TRUE,
expand = 1,col = "green," ltheta = 0,
lphi = 0,box = T,ticktype = "simple," xlab = "a,"
ylab = "b," zlab = "PMC(z1|z2)")
par(mfrow = c(2,2)) # mean on b similar to b fixed
plot(A,rowMeans(pzrzir), "l,"xlab = "a,"
ylab = "PMC(ZR|ZIR)," main = "Mean on b")
lines(A,rep(.5,length(A)), "l")
minn < -min(c(rowMeans(msezr),rowMeans(msezir)))
maxx < -
max(c(rowMeans(msezr),rowMeans(msezir)))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA, "n," xlab = "a," ylab = "MSP")
lines(A,rowMeans(msezr), "l")
lines(A,rowMeans(msezir), "l", lty = 2)
legend("topleft," legend = c("ZR," "ZIR"),lty = c(1,2),
merge = TRUE)
plot(A,colMeans(pzrzir), "l,"xlab = "b,"
ylab = "PMC(ZR|ZIR)," main = "Mean on a")
lines(A,rep(.5,length(A)), "l")
minn < -min(c(colMeans(msezr),colMeans(msezir)))
maxx < -max(c(colMeans(msezr),colMeans(msezir)))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA, "n," xlab = "b", ylab = "MSP")
lines(A,colMeans(msezr), "l")
```

```
lines(A,colMeans(msezir), “l”, lty = 2)
legend(“topleft”, legend = c(“ZR”, “ZIR”),lty = c(1,2), merge = TRUE)
par(mfrow = c(2,2))
plot(A,pzrzir[,j], “l”, xlab = “a”,ylab = “PMC(ZR|ZIR)”, main = “b = 0.95”)
lines(A,rep(.5,length(A)), “l”)
minn < -min(c(msezr[,j],msezir[,j]))
maxx < -max(c(msezr[,j],msezir[,j]))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA, “n”, xlab = “a”, ylab = “MSP”)
lines(A,msezr[,j], “l”)
lines(A,msezir[,j], “l”, lty = 2,xlab = “a”)
legend(“topleft”, legend = c(“ZR,” “ZIR”),lty = c(1,2), merge = TRUE)
```

More complicated forms:

```
a<-.5
b < -.13
g < -function(i,j) aˆi∗bˆj
A < -matrix(c(1,a,g(3,2),-g(3,3),a,1,g(2,2),g(2,3),g(3,2),g(2,2),1,b,g(3,3),g(2,3),-b,1),4,4)
B < -c(g(2,1),g(1,1),g(1,1),g(1,2))
solve(A,B)
A < -matrix(c(1,a,g(2,2),a,1,g(1,2),g(2,2),g(1,2),1),3,3)
B < -c(g(2,1),g(1,1),g(1,1))
solve(A,B)
```

Figure 3:

```
par(mfrow = c(1,2))
y < -c(1,5)
x < -c(1,5)
y1<-c(1,2,3,2,3,1,2,3) + 2
x1<-c(1,1,1,2,2,3,3,3) + 2
plot(x,y,xlab = “”ylab = “”, “n,” main = “Redundant Sampling,” axes = FALSE).
points(x1,y1,xlab = ““ylab = ,”” pch = 15).
points(3,3,pch = 3). points(1,1,pch = 2). axis(1,1:5,c(expression(i),”,expression(s[0]),““,””)). axis(2,1:5,c(expression(j),“,expression(t[0]),““,””))
box()
y < -c(1,5)
x < -c(1,5)
y1<-c(2,3,1,2,3,1,2,3) + 2
x1<-c(1,1,2,2,2,3,3,3) + 2
plot(x,y,xlab = “”,ylab = ””, “n,”main = “Irredundant sampling,”axes = FALSE)
points(x1,y1,xlab = ““,ylab = ””,pch = 15)
points(1,1,pch = 2)
axis(1,1:5,c(expression(i),“,”,““,””))
axis(2,1:5,c(expression(j),“,”,““,””))
box()
```

Comparison between redundant and nonredundant sampling with exponential error.

```
rm(list = ls())
m < −200; n < −200
z0<-rnorm(m + n + 1). epsilon < -matrix(rexp((m + 1)∗(n + 1)),m + 1,n + 1). A < -seq(.05,.96,.05). B < -seq(.05,.96,.05)
msezr < -msezir < -matrix(0,length(A),length(B)). pzrzir < -matrix(0,length(A),length(B)). for(a in A){
for(b in B){
g < -function(x,y) aˆx∗bˆy. zAR < -matrix(0,m + 1,n + 1)
zAR[1,]<-z0[1:(n + 1)]
zAR[−1,1]<−z0[(n + 2):(n + m + 1)]
for(i in 2:(m + 1)){. for(j in 2:(n + 1)){. zAR[i,j]<−zAR[i − 1,j]∗a + zAR[i,j − 1]∗b + zAR[i − 1,j − 1]∗ − (a∗b) }}
z < -zAR + epsilon
z < −z[−1,−1]
z < −z−1
fZR < -function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) y < −((a∗b)ˆ2/2)∗z[i + 2,j + 2]
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
return(y)}
fZIR < -function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) {
A1<-
matrix(c(g(0,0),g(0,1),g(1,1),g(1,0),g(1,1),g(2,1),g(2,0),-g(2,1),
  g(0,1),g(0,0),g(1,2),g(1,1),g(1,0),g(2,2),g(2,1),g(2,0),
  g(1,1),g(1,2),g(0,0),g(0,1),g(0,2),g(1,0),g(1,1),g(1,2),
  g(1,0),g(1,1),g(0,1),g(0,0),g(0,1),g(1,1),g(1,0),g(1,1),
  g(1,1),g(1,0),g(0,2),g(0,1),g(0,0),g(1,2),g(1,1),g(1,0),
  g(2,1),g(2,2),g(1,0),g(1,1),g(1,2),g(0,0),g(0,1),g(0,2),
  g(2,0),g(2,1),g(1,1),g(1,0),g(1,1),g(0,1),g(0,0),g(0,1),
  g(2,1),g(2,0),g(1,2),g(1,1),g(1,0),g(0,2),g(0,1),g(0,0)),8,8)
B1<-
matrix(c(g(2,3),g(2,4),g(3,2),g(3,3),g(3,4),g(4,2),-g(4,3),g(4,4)),8,1)
Q1<-c(z[i + 2,j + 3],z[i + 2,j + 4],z[i + 3,j + 2],z[i + 3,j + 3],z[i + 3,j + 4],z[i + 4,j + 2],
z[i + 4,j + 3],z[i + 4,j + 4])
y < -t(solve(A1,B1))%∗%Q1 }
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
y < -as.vector(y)
return(y)}
```

```
fZIR1<-function(i,j){
if((j<=(n − 4)) &(i<=(m − 4))) y < −(aˆ2∗bˆ2)∗(b∗z
[i + 2,j + 3] + a∗z[i + 3,j + 2]−a∗b∗z[i + 3,j + 3])
if((j > (n − 4))|(i > (m − 4))) y < −z[i,j]
return(y)}
zr < -matrix(0,m,n). zir < -matrix(0,m,n). for(i in 1:m){.
for(j in 1:n){
zr[i,j]<-fZR(i,j)
zir[i,j]<-fZIR1(i,j) }}
msezr[abs(a)∗20,abs(b)∗20]<-mean((zr-z)ˆ2)
msezir[abs(a)∗20,abs(b)∗20]<-mean((zir-z)ˆ2)
pzrzir[abs(a)∗20,abs(b)∗20]<-sum((abs(zr-z))<(abs(-
zir-z)))/sum((abs(zr-z))!=(abs(zir-z))) }}
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj. BB < -grid$yj
persp(A,B,msezr, theta = 135,phi = 5,scale = TRUE,
expand = 1,col = "green,"ltheta = 0,
lphi = 0,box = T,ticktype = "simple,"
xlab = "a,"ylab = "b," zlab = "MSE(Z1)")
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj
BB < -grid$yj
persp(A,B,msezir, theta = 155,phi = 5,scale = TRUE,
expand = 1,col = "green,"ltheta = 0,
lphi = 0,box = T,ticktype = "simple,"xlab = "a,"
ylab = "b," zlab = "MSE(Z2)")
grid < -expand.grid(xj = A,yj = B)
AA < -grid$xj
BB < -grid$yj
persp(A,B,pzrzir, theta = 115,phi = 5,scale = TRUE,
expand = 1,col = "green,"ltheta = 0,
lphi = 0,box = T,ticktype = "simple," xlab = "a,"
ylab = "b," zlab = "PMC(z1|z2)")
par(mfrow = c(2,2)) # mean on b similar to b fixed.
plot(A,rowMeans(pzrzir),
"l",xlab = "a",ylab = "PMC(ZR|ZIR)",main = "Mean on
b")
lines(A,rep(.5,length(A)),"l"). minn < -min(-
c(rowMeans(msezr),rowMeans(msezir))), maxx < -
max(c(rowMeans(msezr),rowMeans(msezir)))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA,"n",xlab = "a",ylab = "MSP")
lines(A,rowMeans(msezr),"l")
lines(A,rowMeans(msezir),"l",lty = 2). legend("topleft",
legend = c("ZR","ZIR"),lty = c(1,2), merge = TRUE),
plot(A,colMeans(pzrzir),"l",xlab = "b",ylab = "
PMC(ZR|ZIR)",main = "Mean on a")
lines(A,rep(.5,length(A)),"l"). minn < -min(-
c(colMeans(msezr),colMeans(msezir))). maxx < -
max(c(colMeans(msezr),colMeans(msezir)))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA,"n",xlab = "b",ylab = "MSP")
lines(A,colMeans(msezr),"l")
lines(A,colMeans(msezir),"l",lty = 2)
legend("topleft", legend = c("ZR","ZIR"),lty = c(1,2),
merge = TRUE)
j < −19; B[j]
par(mfrow = c(2,2))
plot(A,pzrzir[,j],"l",xlab = "a",ylab = "PMC(ZR|
ZIR)",main = "b = 0.95")
lines(A,rep(.5,length(A)),"l")
minn < -min(c(msezr[,j],msezir[,j]))
maxx < -max(c(msezr[,j],msezir[,j]))
AA < -seq(minn, maxx,length = length(A))
plot(A,AA,"n",xlab = "a",ylab = "MSP")
lines(A,msezr[,j],"l")
lines(A,msezir[,j],"l",lty = 2,xlab = "a")
legend("topleft", legend = c("ZR","ZIR"),lty = c(1,2),
merge = TRUE)
j < −18; A[j]
```

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The idea of the paper was conceived and planned by Mohammad Mehdi Saber and Zohreh Shishebor. M. M. Abd El Raouf, E.H. Hafez, and Ramy Aldallal took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

References

[1] E. Isaaks and R. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

[2] C. Akselsson, J. Holmqvist, M. Alveteg, D. Kurz, and H. Sverdrup, "Scaling and mapping regional calculations of soil chemical weathering rates in Sweden," *Water, Air, & Soil Pollution: Focus*, vol. 4, no. 2/3, pp. 671–681, 2004.

[3] A. B. McBratney and R. Webster, "The design of optimal sampling schemes for local estimation and mapping of regionalized variables-II," *Computers & Geosciences*, vol. 7, no. 4, pp. 335–365, 1981.

[4] S. D. Zio, L. Fontanella, and L. Ippoliti, "Optimal spatial sampling schemes for environmental surveys," *Environmental and Ecological Statistics*, vol. 11, no. 4, pp. 397–414, 2004.

[5] X. Xiao, G. Gertner, G. Wang, and A. B. Anderson, "Optimal sampling scheme for estimation landscape mapping of vegetation cover," *Landscape Ecology*, vol. 20, no. 4, pp. 375–387, 2005.

[6] J. Sward, F. Elvander, and A. Jakobsson, "Designing Optimal Sampling Schemes," in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, pp. 912–916, Kos, Greece, September 2017.

[7] L. Xie, J. Cheng, H. Wang, J. Wang, M. Hu, and Z. Zhou, "Memory-based event-triggered asynchronous control for semi-Markov switching systems," *Applied Mathematics and Computation*, vol. 415, p. 126694, Article ID 126694, 2022.

[8] J. Cheng, Y. Wang, J. H. Park, J. Cao, and K. Shi, "Static output feedback quantized control for fuzzy markovian switching singularly perturbed systems with deception attacks," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 4, pp. 1036–1047, 2022.

[9] W. Zhou, J. Fu, H. Yan, X. Du, Y. Wang, and H. Zhou, "Event-triggered approximate optimal path-following control for unmanned surface vehicles with state constraints," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2021.

[10] J. Cheng, L. Liang, J. H. Park, H. Yan, and K. Li, "A dynamic event-triggered approach to state estimation for switched memristive neural networks with nonhomogeneous sojourn probabilities," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 12, pp. 4924–4934, 2021.

[11] N. A. C. Cressie, *Statistics for Spatial Data*, Wiley, New York, 1993.

[12] A. Hamaz and M. Ibazizen, "Comparison of two estimation methods of missing values using Pitman-closeness criterion," *Communications in Statistics - Theory and Methods*, vol. 38, no. 13, pp. 2210–2213, 2009.

[13] A. Saadatmand, A. R. Nematollahi, and S. M. Sadooghi-Alvandi, "On the estimation of missing values in AR(1) model with exponential innovations," *Communications in Statistics - Theory and Methods*, vol. 46, no. 7, pp. 3393–3400, 2016.

[14] M. M. Saber and A. R. Nematollahi, "Comparison of spatial interpolation methods in the first order stationary multiplicative spatial autoregressive models," *Communications in Statistics - Theory and Methods*, vol. 46, no. 18, pp. 9230–9246, 2017.

[15] M. M. Saber, "Performance of extrapolation based on Pitman's measure of closeness in spatial regression models with extended skew innovations," *Communications in Statistics - Theory and Methods*, vol. 0, pp. 1–18, 2017.

[16] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1991.

[17] C. Gaetan and X. Guyon, *Spatial Statistics and Modelling*, Springer-Verlag, New York, 2009.

[18] M. M. Saber and K. Khorshidian, "Optimum predictor in stationary first-order moving average process," *Iranian Journal of Science and Technology, Transaction A: Science*, vol. 45, no. 425, pp. 1757–1764, 2021.