

# Improvement in Parameter Estimation for a Gaussian AR(1) Process with an Unknown Drift and Additive Outliers: A Simulation Study

Wararit Panichkitkosolkul

## ABSTRACT

This paper presents a new parameter estimation for a Gaussian first-order autoregressive (AR(1)) process with an unknown drift and additive outliers. Recursive median adjustment was applied based on an  $\alpha$ -winsorized mean to the weighted symmetric estimator. The following estimators were considered: the weighted symmetric estimator ( $\hat{\rho}_W$ ), the recursive mean adjusted weighted symmetric estimator ( $\hat{\rho}_{R-W}$ ), the recursive median adjusted weighted symmetric estimator ( $\hat{\rho}_{Rmd-W}$ ), and the weighted symmetric estimator using the adjusted recursive median, based on the  $\alpha$ -winsorized mean ( $\hat{\rho}_{W-Rmd-W}$ ). Using Monte Carlo simulations, the mean square error (MSE) of estimators were compared. Simulation results showed that the proposed estimator,  $\hat{\rho}_{W-Rmd-W}$ , provided an MSE lower than those of  $\hat{\rho}_{W-Rmd-W}$ , and  $\hat{\rho}_{W-Rmd-W}$  for almost all situations.

**Keywords:** parameter estimation, AR(1) process, recursive median, winsorized mean, additive outliers

## INTRODUCTION

Economic time series are frequently affected by special events, such as policy interventions, strikes, the outbreak of war and sudden changes in the market structure of a commodity (Nielsen, 2004). Such aberrant observations are usually referred to as outliers. Because outliers are known to wreck havoc on parameter estimation, it is important to have procedures that will deal with such outlier effects. The detection of time series outliers was first studied by Fox (1972), who introduced two statistical models for times series contaminated by outliers, namely, additive outliers (AO) and innovations outliers (IO). Additive outliers

correspond to the situation in which a gross error of observation or recording affects a single observation (Fox, 1972). An innovations outlier affects not only the particular observation, but also subsequent observations (Fox, 1972). The current study focuses solely on additive outliers, as they are the most common type found in time series, due to their association with human errors, such as typing and recording mistakes (Zaharim, 2009). Furthermore, additive outliers are more harmful than innovations outliers (Chatfield, 2001). A time series that does not contain any outliers is called an outlier-free series.

Suppose an outlier-free time series  $\{X_t ; t = 2, 3, \dots, n\}$  follows an AR(1) model as shown in Equation 1:

$X_t = \mu + \rho(X_{t-1} - \mu) + e_t$ , (1)  
 where  $\mu$  is the mean of the process,  $\rho$  is an autoregressive parameter,  $\rho \in (-1, 1)$ ,  $e_t$  are unobservable independent errors and identically  $N(0, \sigma_e^2)$  distributed.

For  $|\rho| = 1$ , model (1) in Equation 1 is called the random walk model, otherwise it is called a stationary AR(1) process when  $|\rho| < 1$ . For  $\rho$  close to one or near a non-stationary process, the mean, variance and autocorrelation function of this process are not constant through time. Let the observed time series be denoted by  $\{Y_t\}$ . In the simple case when  $\{X_t\}$  has a single additive outlier at time point  $T$  ( $1 < T < n$ ), model (1) can be modified as shown in Equation 2:

$$Y_t = X_t + \delta I_t^{(T)}, \quad (2)$$

where  $\delta$  represents the magnitude of the additive outlier effect and  $I_t^{(T)}$  is an indicator variable such that

$$I_t^{(T)} = \begin{cases} 1, & t = T, \\ 0, & t \neq T. \end{cases}$$

It is known that the ordinary least squares estimator of  $\rho$ , which is denoted by  $\hat{\rho}_{OLS}$ , for Equation 1 is biased (see, for example, Shaman and Stine, 1988). For case of additive outliers, the ordinary least squares (OLS) estimator not only lacks robustness in terms of variability but also suffers from severe bias problems (Guo, 2000). Furthermore, Conover (1980) pointed out that the OLS estimator was sensitive to outliers (see Section 5.5). Therefore, statisticians have suggested methods to reduce the bias. Park and Fuller (1995) proposed the weighted symmetric estimator of  $\rho$ , which is denoted by  $\hat{\rho}_W$ . So and Shin (1999) applied a recursive mean adjustment to the OLS estimator (R-OLS) and they concluded that the mean square error of the R-OLS estimator, which is denoted by  $\hat{\rho}_{R-OLS}$ , is smaller than the OLS estimator for  $\rho \in (0, 1)$ . They also showed that the  $\hat{\rho}_{R-OLS}$  estimator has a coverage probability that is close to the nominal value. Niwitpong (2007) applied the recursive mean adjustment to the weighted symmetric

estimator (R-W) of Park and Fuller (1995). Panichkitkosolkul (2010) suggested an estimator for an unknown mean Gaussian AR(1) process with additive outliers by applying the recursive median adjustment to the weighted symmetric estimator (Rmd-W). He found that the  $\hat{\rho}_{Rmd-W}$  estimator produces a mean square error lower than those of  $\hat{\rho}_W$  and  $\hat{\rho}_{R-W}$  for almost all situations. New recursive median adjustment based on an  $\alpha$ -winsorized mean (see Serfling, 1980) is applied to the weighted symmetric estimator (W-Rmd-W) for model (1) when there are additive outliers in the time series data. Because the outliers do not affect the  $\alpha$ -winsorized mean and median values, the recursive mean adjustment is replaced by a new recursive median adjustment based on an  $\alpha$ -winsorized mean to the weighted symmetric estimator. The aim of the current study was to compare four estimators,  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ ,  $\hat{\rho}_{Rmd-W}$  and  $\hat{\rho}_{W-Rmd-W}$ , in terms of the mean square error (MSE) of the estimators.

## METHODOLOGY

Park and Fuller (1995) proposed that the weighted symmetric estimator of  $\rho$  is given by Equation 3:

$$\hat{\rho}_W = \frac{\sum_{t=2}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=3}^n (Y_{t-1} - \bar{Y})^2 + n^{-1} \sum_{t=1}^n (Y_t - \bar{Y})^2}. \quad (3)$$

Niwitpong (2007) replaced  $\bar{Y}$  by  $\bar{Y}_t = \frac{1}{t} \sum_{i=1}^t Y_i$  in Equation 3. The estimator of  $\rho$  obtained as a result of this recursive mean adjustment is given by Equation 4:

$$\hat{\rho}_{R-W} = \frac{\sum_{t=2}^n (Y_t - \bar{Y}_t)(Y_{t-1} - \bar{Y}_{t-1})}{\sum_{t=3}^n (Y_{t-1} - \bar{Y}_{t-1})^2 + n^{-1} \sum_{t=1}^n (Y_t - \bar{Y}_t)^2}. \quad (4)$$

When there are outliers in time series

data, it affects the recursive mean  $\bar{Y}_t$  in Equation 4. Panichkitkosolkul (2010) replaced the recursive mean in Equation 4 by the recursive median. The estimator of  $\rho$  obtained as a result of the recursive median adjustment is given by Equation 5:

$$\hat{\rho}_{Rmd-W} = \frac{\sum_{t=2}^n (Y_t - \tilde{Y}_t)(Y_{t-1} - \tilde{Y}_{t-1})}{\sum_{t=3}^n (Y_{t-1} - \tilde{Y}_{t-1})^2 + n^{-1} \sum_{t=1}^n (Y_t - \tilde{Y}_t)^2}, \quad (5)$$

where  $\tilde{Y}_t = \text{median}(Y_1, Y_2, \dots, Y_t)$ .

The effect of outliers on an estimator of  $\rho$  in model (1) can be reduced by using a new recursive median adjustment based on the  $\alpha$ -winsorized mean, which is a robust estimator of a mean (David and Nagaraja, 2003). The  $\alpha$ -winsorized mean, which is shown by Epstein and Sobel (1953 cited in Staudte and Sheather, 1990), is a winsorized statistical measurement of central tendency, like the mean and median and even more similar to the trimmed mean (Mahir and Al-Khazaleh, 2009; Wu and Zuo, 2009). Additionally, the  $\alpha$ -winsorized mean eliminates the outliers at both ends of an ordered set of observations. Unlike the trimmed mean, the  $\alpha$ -winsorized mean replaces the outliers with observed values, rather than discarding them. Apart from that, the  $\alpha$ -winsorized mean is more efficient than the trimmed mean (Johnson *et al.*, 1994). In addition, this statistic has various applications in detecting outliers (Barnett and Lewis, 1994). Thus, the improved recursive median values adjusted using an  $\alpha$ -winsorized mean are derived from computing the  $\alpha$ -winsorized mean of the recursive median. There are two steps in computing the new recursive median. First, the recursive median ( $\tilde{Y}_t$ ) is computed using the time series data  $Y_t$ . Second, the  $\alpha$ -winsorized mean is calculated using the recursive median obtained from the first step. Therefore, the recursive median in Equation 5 is replaced by a new recursive median. The new estimator of  $\rho$  obtained as a result of this new recursive median adjustment is given by Equation 6:

$$\hat{\rho}_{W-Rmd-W} = \frac{\sum_{t=2}^n (Y_t - \bar{\tilde{Y}}_t)(Y_{t-1} - \bar{\tilde{Y}}_{t-1})}{\sum_{t=3}^n (Y_{t-1} - \bar{\tilde{Y}}_{t-1})^2 + n^{-1} \sum_{t=1}^n (Y_t - \bar{\tilde{Y}}_t)^2}, \quad (6)$$

where  $\bar{\tilde{Y}}_t = \frac{1}{t} \left( [t\alpha] \tilde{Y}_{([t\alpha]+1)} + \sum_{i=[t\alpha]+1}^{t-[t\alpha]} \tilde{Y}_{(i)} + [t\alpha] \tilde{Y}_{(t-[t\alpha])} \right)$ ,

$$\tilde{Y}_t = \text{median}(Y_1, Y_2, \dots, Y_t), \quad \tilde{Y}_{(1)} \leq \tilde{Y}_{(2)} \leq \dots \leq \tilde{Y}_{(t)}$$

denote the ordered values of the recursive median  $\tilde{Y}_t$ ,  $\alpha$  denotes the fraction of time series data to be trimmed,  $0 < \alpha < 0.5$  and  $[t\alpha]$  denotes the greatest integer not greater than  $t\alpha$ .

## RESULTS

The performance of the proposed estimator for a Gaussian AR(1) process with additive outliers was examined (via Monte Carlo simulations), with particular emphasis on comparisons between the new and existing approaches. Data were generated from a Gaussian AR(1) process with additive outliers. It should be noted that while this process generates  $Y_t \sim N(0, \frac{\sigma_e^2}{1-\rho^2})$  and simulates the time series of length  $n + 50$ , the time series used in the calculations were  $\{Y_{51}, Y_{52}, \dots, Y_{n+50}\}$ . The following parameter values were used:  $(\mu, \sigma_e^2) = (0, 1)$ ;  $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$  and  $0.9$ ; sample sizes  $n = 25, 50, 100$  and  $250$ ; the magnitude of the AOs effect  $\delta = 3\sigma_e$  and  $5\sigma_e$ ; the percentage of additive outliers  $p = 5\%$  and  $10\%$ ; the fraction of data to be trimmed  $\alpha = 0.05$ . All simulations were performed using programs written in the R statistical software package (The R Development Core Team, 2009a, 2009b) with the number of simulation runs,  $M = 10,000$ . In addition, the additive outliers occurred randomly.

The simulation results, including the estimated mean square error (MSE) of all estimators,  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ ,  $\hat{\rho}_{Rmd-W}$  and  $\hat{\rho}_{W-Rmd-W}$ , are shown in Tables 1 and 2. Table 3 showed the standard error of all estimators. As can be seen from Tables 1 and 2, the MSE of  $\hat{\rho}_W$  was larger

**Table 1** The estimated mean square error (MSE) of  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ ,  $\hat{\rho}_{Rmd-W}$  and  $\hat{\rho}_{W-Rmd-W}$  for the percentage of additive outliers;  $p = 5\%$ .

| n   | $\rho$ | $\delta = 3\sigma_e$ |        |        |         | $\delta = 5\sigma_e$ |        |        |         |
|-----|--------|----------------------|--------|--------|---------|----------------------|--------|--------|---------|
|     |        | W                    | R-W    | Rmd-W  | W-Rmd-W | W                    | R-W    | Rmd-W  | W-Rmd-W |
| 25  | 0.1    | 0.0429               | 0.0384 | 0.0365 | 0.0446  | 0.0387               | 0.0340 | 0.0317 | 0.0336  |
|     | 0.2    | 0.0500               | 0.0428 | 0.0404 | 0.0409  | 0.0554               | 0.0469 | 0.0435 | 0.0368  |
|     | 0.3    | 0.0611               | 0.0514 | 0.0474 | 0.0412  | 0.0773               | 0.0657 | 0.0622 | 0.0467  |
|     | 0.4    | 0.0701               | 0.0583 | 0.0551 | 0.0414  | 0.1021               | 0.0869 | 0.0834 | 0.0595  |
|     | 0.5    | 0.0812               | 0.0676 | 0.0657 | 0.0451  | 0.1342               | 0.1162 | 0.1136 | 0.0779  |
|     | 0.6    | 0.0932               | 0.0785 | 0.0770 | 0.0494  | 0.1625               | 0.1420 | 0.1399 | 0.0942  |
|     | 0.7    | 0.1036               | 0.0886 | 0.0870 | 0.0545  | 0.1936               | 0.1701 | 0.1680 | 0.1123  |
|     | 0.8    | 0.1125               | 0.0975 | 0.0973 | 0.0591  | 0.2106               | 0.1876 | 0.1874 | 0.1237  |
|     | 0.9    | 0.1186               | 0.1059 | 0.1069 | 0.0642  | 0.2227               | 0.2005 | 0.2014 | 0.1315  |
| 50  | 0.1    | 0.0214               | 0.0198 | 0.0192 | 0.0238  | 0.0228               | 0.0209 | 0.0195 | 0.0202  |
|     | 0.2    | 0.0256               | 0.0231 | 0.0218 | 0.0228  | 0.0335               | 0.0300 | 0.0274 | 0.0225  |
|     | 0.3    | 0.0308               | 0.0273 | 0.0253 | 0.0220  | 0.0499               | 0.0449 | 0.0418 | 0.0306  |
|     | 0.4    | 0.0384               | 0.0341 | 0.0321 | 0.0236  | 0.0682               | 0.0618 | 0.0584 | 0.0411  |
|     | 0.5    | 0.0433               | 0.0386 | 0.0370 | 0.0252  | 0.0897               | 0.0820 | 0.0790 | 0.0542  |
|     | 0.6    | 0.0483               | 0.0435 | 0.0421 | 0.0264  | 0.1039               | 0.0956 | 0.0928 | 0.0623  |
|     | 0.7    | 0.0513               | 0.0466 | 0.0453 | 0.0276  | 0.1181               | 0.1094 | 0.1073 | 0.0706  |
|     | 0.8    | 0.0493               | 0.0455 | 0.0445 | 0.0261  | 0.1178               | 0.1099 | 0.1081 | 0.0696  |
|     | 0.9    | 0.0441               | 0.0416 | 0.0413 | 0.0237  | 0.1063               | 0.1002 | 0.0991 | 0.0606  |
| 100 | 0.1    | 0.0113               | 0.0107 | 0.0101 | 0.0121  | 0.0137               | 0.0129 | 0.0113 | 0.0111  |
|     | 0.2    | 0.0152               | 0.0140 | 0.0128 | 0.0117  | 0.0243               | 0.0226 | 0.0194 | 0.0154  |
|     | 0.3    | 0.0202               | 0.0187 | 0.0171 | 0.0130  | 0.0394               | 0.0370 | 0.0330 | 0.0247  |
|     | 0.4    | 0.0265               | 0.0246 | 0.0229 | 0.0158  | 0.0592               | 0.0560 | 0.0514 | 0.0383  |
|     | 0.5    | 0.0317               | 0.0296 | 0.0280 | 0.0184  | 0.0791               | 0.0755 | 0.0712 | 0.0527  |
|     | 0.6    | 0.0354               | 0.0332 | 0.0317 | 0.0200  | 0.0947               | 0.0907 | 0.0866 | 0.0627  |
|     | 0.7    | 0.0361               | 0.0342 | 0.0330 | 0.0200  | 0.1040               | 0.0999 | 0.0966 | 0.0688  |
|     | 0.8    | 0.0317               | 0.0302 | 0.0292 | 0.0172  | 0.0971               | 0.0933 | 0.0910 | 0.0609  |
|     | 0.9    | 0.0232               | 0.0225 | 0.0219 | 0.0122  | 0.0719               | 0.0696 | 0.0678 | 0.0425  |
| 250 | 0.1    | 0.0052               | 0.0050 | 0.0046 | 0.0047  | 0.0074               | 0.0071 | 0.0057 | 0.0052  |
|     | 0.2    | 0.0080               | 0.0076 | 0.0068 | 0.0055  | 0.0160               | 0.0154 | 0.0127 | 0.0101  |
|     | 0.3    | 0.0121               | 0.0116 | 0.0106 | 0.0076  | 0.0295               | 0.0286 | 0.0250 | 0.0201  |
|     | 0.4    | 0.0165               | 0.0158 | 0.0146 | 0.0102  | 0.0458               | 0.0446 | 0.0405 | 0.0328  |
|     | 0.5    | 0.0201               | 0.0194 | 0.0184 | 0.0125  | 0.0626               | 0.0612 | 0.0571 | 0.0463  |
|     | 0.6    | 0.0228               | 0.0220 | 0.0211 | 0.0143  | 0.0757               | 0.0742 | 0.0707 | 0.0564  |
|     | 0.7    | 0.0222               | 0.0216 | 0.0209 | 0.0136  | 0.0792               | 0.0777 | 0.0748 | 0.0580  |
|     | 0.8    | 0.0177               | 0.0173 | 0.0167 | 0.0103  | 0.0687               | 0.0675 | 0.0656 | 0.0483  |
|     | 0.9    | 0.0101               | 0.0100 | 0.0097 | 0.0056  | 0.0412               | 0.0406 | 0.0394 | 0.0261  |

**Table 2** The estimated mean square error (MSE) of  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ ,  $\hat{\rho}_{Rmd-W}$  and  $\hat{\rho}_{W-Rmd-W}$  for the percentage of additive outliers;  $p = 10\%$ .

| n   | $\rho$ | $\delta = 3\sigma_e$ |        |        |         | $\delta = 5\sigma_e$ |        |        |         |
|-----|--------|----------------------|--------|--------|---------|----------------------|--------|--------|---------|
|     |        | W                    | R-W    | Rmd-W  | W-Rmd-W | W                    | R-W    | Rmd-W  | W-Rmd-W |
| 25  | 0.1    | 0.0441               | 0.0391 | 0.0366 | 0.0417  | 0.0445               | 0.0394 | 0.0345 | 0.0354  |
|     | 0.2    | 0.0553               | 0.0471 | 0.0424 | 0.0400  | 0.0666               | 0.0574 | 0.0498 | 0.0428  |
|     | 0.3    | 0.0723               | 0.0612 | 0.0556 | 0.0445  | 0.0967               | 0.0836 | 0.0738 | 0.0582  |
|     | 0.4    | 0.0925               | 0.0781 | 0.0723 | 0.0522  | 0.1395               | 0.1222 | 0.1106 | 0.0844  |
|     | 0.5    | 0.1143               | 0.0973 | 0.0915 | 0.0628  | 0.1862               | 0.1648 | 0.1533 | 0.1153  |
|     | 0.6    | 0.1367               | 0.1176 | 0.1136 | 0.0756  | 0.2343               | 0.2097 | 0.1978 | 0.1472  |
|     | 0.7    | 0.1538               | 0.1339 | 0.1302 | 0.0846  | 0.2745               | 0.2462 | 0.2353 | 0.1720  |
|     | 0.8    | 0.1686               | 0.1484 | 0.1455 | 0.0936  | 0.3159               | 0.2861 | 0.2767 | 0.1997  |
|     | 0.9    | 0.1736               | 0.1555 | 0.1532 | 0.0971  | 0.3347               | 0.3044 | 0.2978 | 0.2094  |
| 50  | 0.1    | 0.0240               | 0.0221 | 0.0202 | 0.0232  | 0.0261               | 0.0239 | 0.0193 | 0.0202  |
|     | 0.2    | 0.0334               | 0.0299 | 0.0261 | 0.0231  | 0.0454               | 0.0412 | 0.0323 | 0.0276  |
|     | 0.3    | 0.0469               | 0.0420 | 0.0374 | 0.0284  | 0.0740               | 0.0678 | 0.0554 | 0.0444  |
|     | 0.4    | 0.0623               | 0.0561 | 0.0508 | 0.0359  | 0.1100               | 0.1019 | 0.0865 | 0.0675  |
|     | 0.5    | 0.0786               | 0.0714 | 0.0661 | 0.0447  | 0.1489               | 0.1391 | 0.1225 | 0.0956  |
|     | 0.6    | 0.0926               | 0.0849 | 0.0800 | 0.0518  | 0.1913               | 0.1799 | 0.1630 | 0.1251  |
|     | 0.7    | 0.1009               | 0.0931 | 0.0888 | 0.0572  | 0.2237               | 0.2110 | 0.1961 | 0.1475  |
|     | 0.8    | 0.1038               | 0.0965 | 0.0930 | 0.0586  | 0.2401               | 0.2268 | 0.2153 | 0.1566  |
|     | 0.9    | 0.0905               | 0.0852 | 0.0828 | 0.0505  | 0.2207               | 0.2087 | 0.1996 | 0.1364  |
| 100 | 0.1    | 0.0129               | 0.0121 | 0.0109 | 0.0122  | 0.0160               | 0.0150 | 0.0109 | 0.0114  |
|     | 0.2    | 0.0200               | 0.0185 | 0.0157 | 0.0133  | 0.0319               | 0.0299 | 0.0216 | 0.0185  |
|     | 0.3    | 0.0305               | 0.0283 | 0.0246 | 0.0180  | 0.0571               | 0.0542 | 0.0422 | 0.0347  |
|     | 0.4    | 0.0440               | 0.0412 | 0.0368 | 0.0257  | 0.0892               | 0.0854 | 0.0710 | 0.0579  |
|     | 0.5    | 0.0566               | 0.0534 | 0.0491 | 0.0341  | 0.1243               | 0.1198 | 0.1041 | 0.0848  |
|     | 0.6    | 0.0656               | 0.0623 | 0.0587 | 0.0398  | 0.1580               | 0.1526 | 0.1374 | 0.1106  |
|     | 0.7    | 0.0705               | 0.0672 | 0.0639 | 0.0425  | 0.1819               | 0.1759 | 0.1633 | 0.1278  |
|     | 0.8    | 0.0644               | 0.0617 | 0.0594 | 0.0381  | 0.1818               | 0.1761 | 0.1665 | 0.1245  |
|     | 0.9    | 0.0463               | 0.0448 | 0.0433 | 0.0261  | 0.1426               | 0.1382 | 0.1322 | 0.0898  |
| 250 | 0.1    | 0.0064               | 0.0061 | 0.0050 | 0.0050  | 0.0093               | 0.0089 | 0.0051 | 0.0052  |
|     | 0.2    | 0.0128               | 0.0122 | 0.0098 | 0.0077  | 0.0243               | 0.0235 | 0.0150 | 0.0131  |
|     | 0.3    | 0.0223               | 0.0215 | 0.0183 | 0.0138  | 0.0473               | 0.0462 | 0.0340 | 0.0295  |
|     | 0.4    | 0.0335               | 0.0325 | 0.0291 | 0.0221  | 0.0767               | 0.0752 | 0.0606 | 0.0532  |
|     | 0.5    | 0.0439               | 0.0428 | 0.0396 | 0.0302  | 0.1091               | 0.1074 | 0.0918 | 0.0804  |
|     | 0.6    | 0.0510               | 0.0498 | 0.0469 | 0.0354  | 0.1378               | 0.1358 | 0.1212 | 0.1052  |
|     | 0.7    | 0.0526               | 0.0514 | 0.0494 | 0.0362  | 0.1565               | 0.1542 | 0.1423 | 0.1203  |
|     | 0.8    | 0.0448               | 0.0439 | 0.0424 | 0.0298  | 0.1500               | 0.1478 | 0.1399 | 0.1131  |
|     | 0.9    | 0.0255               | 0.0252 | 0.0244 | 0.0154  | 0.1002               | 0.0987 | 0.0950 | 0.0691  |

than the MSEs of the other estimators, especially when  $\rho$  is close to one and for small sample sizes. These values decreased as the sample sizes increased. The  $\hat{\rho}_W$  performed well for  $n \geq 50$ . On the other hand, the new estimator,  $\hat{\rho}_{W-Rmd-W}$ , provided the lowest MSE in all scenarios that were considered, except when the parameter  $\rho$  was small ( $\rho = 0.1$  or  $0.2$ ). Additionally, the  $\hat{\rho}_{W-Rmd-W}$  performed very well with respect to the other three estimators. The proposed estimator,  $\hat{\rho}_{W-Rmd-W}$  in Equation 6, dominated all estimators, since the MSE of the proposed estimator was the lowest for almost all cases. For the rest, the MSE of  $\hat{\rho}_{Rmd-W}$  was less than that of  $\hat{\rho}_{R-W}$  and  $\hat{\rho}_W$  for almost all situations. The  $\hat{\rho}_{Rmd-W}$  often ranked second best following the proposed estimator. Furthermore, the MSEs shown in Table 1 are less than those reported in Table 2, because the time series data of Table 1 had less outliers.

## DISCUSSION AND CONCLUSIONS

A new parameter estimation for a Gaussian AR(1) process with an unknown drift and additive outliers has been proposed in this paper. This proposed estimator of  $\rho$  was obtained by applying the recursive median adjustment based

on an  $\alpha$ -winsorized mean to the weighted symmetric estimator. The adjusted recursive median values were derived from computing the  $\alpha$ -winsorized mean of the recursive median. Furthermore, the weighted symmetric estimator ( $\hat{\rho}_W$ ), the recursive mean adjusted weighted symmetric estimator ( $\hat{\rho}_{R-W}$ ), the recursive median adjusted weighted symmetric estimator ( $\hat{\rho}_{Rmd-W}$ ) and the new estimator ( $\hat{\rho}_{W-Rmd-W}$ ) were compared in this study. The new estimator,  $\hat{\rho}_{W-Rmd-W}$ , performed better than  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ , and  $\hat{\rho}_{Rmd-W}$  in terms of the MSE for almost all scenarios. One reason behind this is that the additive outliers do not affect the median and  $\alpha$ -winsorized mean values. Moreover, the adjusted recursive median values applied in the formula for  $\hat{\rho}_{W-Rmd-W}$  in Equation 6 could also reduce the mean square error (MSE) of the estimator. Therefore, the proposed estimator ( $\hat{\rho}_{W-Rmd-W}$ ), which is based on the recursive median adjusted by an  $\alpha$ -winsorized mean, is superior to the existing estimators.

There is a problem for further research, which goes beyond the scope of the present paper, but is of practical interest. In practice, a statistician or an econometrician has one time series set that is contaminated by various kinds of outliers (i.e., additive outliers (AO) and innovations outliers

**Table 3** Standard errors of all estimators,  $\hat{\rho}_W$ ,  $\hat{\rho}_{R-W}$ ,  $\hat{\rho}_{Rmd-W}$  and  $\hat{\rho}_{W-Rmd-W}$ .

| Standard error of estimators  |   |
|---|---|
| $SE(\hat{\rho}_W) = \frac{\hat{\sigma}_W}{\sqrt{\sum_{t=2}^n (Y_{t-1} - \bar{Y})^2}}$ ,                               | $\hat{\sigma}_W^2 = \frac{\sum_{t=2}^n (Y_t - \bar{Y} - \hat{\rho}_W(Y_{t-1} - \bar{Y}))^2}{n-2}$   |
| $SE(\hat{\rho}_{R-W}) = \frac{\hat{\sigma}_{R-W}}{\sqrt{\sum_{t=2}^n (Y_{t-1} - \bar{Y}_{t-1})^2}}$ ,                 | $\hat{\sigma}_{R-W}^2 = \frac{\sum_{t=2}^n (Y_t - \bar{Y}_t - \hat{\rho}_{R-W}(Y_{t-1} - \bar{Y}_{t-1}))^2}{n-2}$                         |
| $SE(\hat{\rho}_{Rmd-W}) = \frac{\hat{\sigma}_{Rmd-W}}{\sqrt{\sum_{t=2}^n (Y_{t-1} - \tilde{Y}_{t-1})^2}}$ ,           | $\hat{\sigma}_{Rmd-W}^2 = \frac{\sum_{t=2}^n (Y_t - \tilde{Y}_t - \hat{\rho}_{Rmd-W}(Y_{t-1} - \tilde{Y}_{t-1}))^2}{n-2}$                 |
| $SE(\hat{\rho}_{W-Rmd-W}) = \frac{\hat{\sigma}_{W-Rmd-W}}{\sqrt{\sum_{t=2}^n (Y_{t-1} - \bar{\tilde{Y}}_{t-1})^2}}$ , | $\hat{\sigma}_{W-Rmd-W}^2 = \frac{\sum_{t=2}^n (Y_t - \bar{\tilde{Y}}_t - \hat{\rho}_{W-Rmd-W}(Y_{t-1} - \bar{\tilde{Y}}_{t-1}))^2}{n-2}$ |

(IO)). Thus, it would be interesting to see whether, in this context, the proposed approach still maintains an edge over the other methodologies.

### ACKNOWLEDGEMENTS

The author would like to acknowledge the numerous helpful comments and suggestions provided by Dr Gareth Clayton.

### LITERATURE CITED

Barnett, V. and T. Lewis. 1994. **Outliers in Statistical Data**. 3rd ed. John Wiley and Sons. Chichester. 604 pp.

Chatfield, C. 2001. **Time-Series Forecasting**. Chapman & Hall. New York. 267 pp.

Conover, W.J. 1980. **Practical Nonparametric Statistics**. John Wiley and Sons. New York. 592 pp.

David, H.A. and H.N. Nagaraja. 2003. **Order Statistics**. John Wiley and Sons. 3rd ed. New York. 458 pp.

Epstein, B. and B. Sobel. 1953. Life testing. **J. Amer. Statist. Assoc.** 48: 486-502.

Fox, A.J. 1972. Outliers in time series. **J. Roy. Statist. Soc.** 34: 350-363.

Guo, J.H. 2000. Robust estimation for the coefficient of a first order autoregressive process. **Commun. Stat. Theory** 29: 55-66.

Johnson, N.L., S. Kotz and N. Balakrishnan. 1994. **Continuous Univariate Distribution, Volume 1**. 2nd ed. John Wiley and Sons. New York. 756 pp.

Nielsen, H.B. 2004. Cointegration analysis in the presence of outliers. **Econometrics J.** 7: 249-271.

Niwitpong, S. 2007. **Predictive Inference for Times Series**. Ph.D. Thesis, La Trobe University.

Mahir, R. and A.M.H. Al-Khazaleh. 2009. New method to estimate missing data by using the asymmetrical winsorized mean in a time series. **Appl. Math. Sci.** 3: 1715-1726.

Panichkitkosolkul, W. 2010. New estimator for an unknown mean Gaussian AR(1) process with additive outliers. **Chiang Mai J. Sci.** 37: 14-20.

Park, H.J. and W.A. Fuller. 1995. Alternative estimators and unit root tests for the autoregressive process. **J. Time Series Analysis.** 16: 415-429.

Serfling, R.J. 1980. **Approximation Theorems of Mathematical and Statistics**. John Wiley and Sons. New York. 371 pp.

Shaman, P. and R.A. Stine. 1988. The bias of autoregressive coefficient estimators. **J. Amer. Statist. Assoc.** 83: 842-848.

So, B.S. and D.W. Shin. 1999. Recursive mean adjustment in time series inferences. **Statists. Prob. Lett.** 43: 65-73.

Staudte, R.G. and S.J. Sheather. 1990. **Robust Estimation and Testing**. John Wiley and Sons. New York. 351 pp.

The R Development Core Team. 2009a. **An Introduction to R**. R Foundation for Statistical Computing. Vienna. 102 pp.

The R Development Core Team. 2009b. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing. Vienna. 1847 pp.

Wu, M. and Y. Zuo. 2009. Trimmed and winsorised means based on a scaled deviation. **J. Statist. Plann. Inference.** 139: 350-365.

Zaharim, A., R. Rajali, R.M. Atok, K. Ibrahim and A.M. Razali. 2009. A study on the nature of an additive outlier in ARMA(1,1) models. **Eur. J. Sci. Res.** 32: 362-368.