CSCanada
Studies in Mathematical Sciences
Vol. 6, No. 2, 2013, pp. [83–95]
DOI: 10.3968/j.sms.1923845220130602.3108

ISSN 1923-8444 [Print]
ISSN 1923-8452 [Online]
www.cscanada.net
www.cscanada.org

Specification of Periodic Autocovariance Structures in the Presence of Outliers

Lasisi, T. A.[a], *, Shangodoyin, D. K.[a] and Moeng, S. R. T.[a]

[a] Department of Statistics, University of Botswana, Botswana.

* Corresponding author.
Address: Department of Statistics, University of Botswana, Gaborone, Botswana;
E-Mail: shangodoyink@mopipi.ub.bw

Received: March 6, 2013/ Accepted: May 13, 2013/ Published: May 31, 2013

Abstract: This paper focuses on the specification of periodic autocovariance structures in the presence of outliers, we evaluate autocovariance structures using various outliers' generating models. The analytical results indicate that outliers affect the estimates of periodic autocovariance function (PACVF) due to biases and inflated standard errors. Robust autocovariance structures that accommodate the influence of outliers in different periodic processes are proposed. We fit AR (1) model using both the conventional and Jacknife autocovariance structures; the latter shows high precision in the standard errors of the estimates. We demonstrate our proposed methodology with the precipitation data from Maun Airport in Botswana, and the empirical study supports our theoretical findings.

Key words: Periodic processes; Outliers models; ARIMA; Jacknife and Autocovariance Structures

Lasisi, T. A., Shangodoyin, D. K., & Moeng, S. R. T. (2013). Specification of Periodic Autocovariance Structures in the Presence of Outliers. *Studies in Mathematical Sciences, 6*(2), 83–95. Available from http://www.cscanada.net/index.php/sms/article/view/j.sms.1923845220130602.3108 DOI: 10.3968/j.sms.1923845220130602.3108

1. INTRODUCTION

The adjustment of a time series model to a dataset is crucial in time series in modeling; and this should follow some basic steps as identification, estimation, diagnostic checking and model selection. Outliers can affect model identification tools, dependence structures of time series models and estimation of model parameters (Chernick *et al.*, 1982). The choice of suitable time series model is important when searching for outlying observations in periodic processes, because, on one hand, a large residual variance caused by overall lack of fit would result in underidentification of outliers, while on the other hand, a less adequate model unable to explain the local behavior of the series would yield a single large residuals, resulting in over-identification, Battaglia and Orfei (2005).

It is a fact that identification can be marred by the presence of more than one outlier in time series due to masking effect; to alleviate the aforementioned challenges this paper focuses on evaluation of autocovariance structures of periodic processes in the presence of additive outliers, innovational outliers, level shifts outliers and transitory change outliers; we shall propose some autocovariance structures that accommodate the impact of outliers in different periodic processes and the basic statistical properties of these structures will be derived. A robust method of estimating autocovariance in the presence of outliers is proposed with a test for checking the significance of autocovariance structures.

2. SPECIFICATION OF PERIODIC PROCESSES WITH OUTLIERS

Suppose that $\{Y_{t(r,m)}\}_{t=1,\dots,N,\ r\geq 1,\ m=1,\dots,s}$ is a periodic autoregressive process given as

$$Y_{t(r,m)} = \mu_m + \sum_{i=1}^{p_m} \left\{ \phi_i^{(m)} \left[Y_{t(r,m)} - \mu_{m-i} \right] \right\} + \varepsilon_{t(r,m)}$$

We define a mean deleted process given below as periodic autoregressive (PAR) process

$$y_{t(r,m)} = Y_{t(r,m)} - \mu_m = \sum_{i=1}^{p_m} \left\{ \phi_i^{(m)} y_{t(r,m)_{-i}} \right\} + \varepsilon_{t(r,m)} \quad (1)$$

and if $y_{t(r,m)}$ non-stationary, we can have the integrated periodic autoregressive process (IPAR) as

$$\nabla^m y_{t(r,m)} = \sum_{i=1}^{p_m} \phi_i^{(m)} \nabla^m y_{t(r,m)-i} + \varepsilon_{t(r,m)} \quad (2)$$

If $\{y_{t(r,m)}\}_{r\geq 1,\ m=1,\dots,s}$ follows a moving average process, then a periodic moving average (PMA) process is given as

$$y_{t(r,m)} = \sum_{j=0}^{q_m} \theta_j^{(m)} \varepsilon_{t(r,m)-j} \quad (3)$$

Also, we specify a periodic autoregressive moving average (PARMA) process as

$$y_{t(r,m)} = \sum_{i=1}^{p_m} \phi_i^{(m)} y_{t(r,m)-i} + \sum_{j=0}^{q_m} \theta_j^{(m)} \varepsilon_{t(r,m)-j} \quad (4)$$

And the integrated periodic autoregressive moving average (IPARMA) process is specified as

$$\nabla^m y_{t(r,m)} = \sum_{i=1}^{p_m} \phi_i^{(m)} \nabla^m y_{t(r,m)-i} + \sum_{j=0}^{q_m} \theta_j^{(m)} \varepsilon_{t(r,m)-j} \quad (5)$$

In Equations (1)–(5), for the sequence $\{y_{t(r,m)}\}$, the index t is an integer division as $t = t(r,m) = (r-1)s + m$, where $m = 1, 2, ..., s$ and $r = 1, 2, ...,$ for instance in the case of monthly data $s = 12$, m and r are respectively the month and year. $\varepsilon_{t(r,m)} \sim IIDN(0, \sigma^2_{(m)})$ and for all $r, r' \geq 1$ and $m, m' = 1, ..., s$, $\varepsilon_{t(r,m)}$ and $\varepsilon_{t(r',m')}$ are independent, p_m is order of autoregressive polynomial, q_m is the order of the moving average polynomial, $...\phi_i^{(m)}$, $i = 1, ..., p_m$ and $\theta_j^{(m)}$, $j = 1, ..., q_m$ are respectively the autoregressive and moving average parameters of the process.

3. SPECIFICATION OF PERIODIC AUTOCOVARIANCES IN THE PRESENCE OF OUTLIERS

Assume that $y_{t(r,m)}$ and $d_{t=T}\xi_T^{t(r,m)}$ are mean-deleted observations such that at time point $t = T$ $\xi_T^{t(r,m)} = +1$ or -1 indicating the presence of outliers depending on the sign of the magnitude of outliers (d).

An infested time series data $\{z_t\}$ with outliers could be represented by

$$z_{t(r,m)} = y_{t(r,m)} + d_T\xi_T^{t(r,m)} \tag{6}$$

Now, writing Equation (6) with different outlier specifications, we have

$$\text{AO}: \quad z^{(1)}_{t(r,m)} = y_{t(r,m)} + d_{T(r,m)} \tag{7}$$

$$\text{IO}: \quad z^{(2)}_{t(r,m)} = y_{t(r,m)} + \alpha(B)d_{T(r,m)} \tag{8}$$

where $\alpha(B) = \dfrac{\theta(B)}{\phi(B)}$; $0 < \delta B < 1$ and $|\delta| \leqslant 1$.

$$\text{LS}: \quad z^{(3)}_{t(r,m)} = y_{t(r,m)} + (1-\delta)^{-1}d_{T(r,m)} \tag{9}$$

$$\text{TC}: \quad z^{(4)}_{t(r,m)} = y_{t(r,m)} + (1-\delta B)^{-1}d_{T(r,m)} \tag{10}$$

3.1. Proposition 1

Suppose that an outlier infested series follows a general liner difference equation of the form

$$z^{(i)}_{t(r,m)} = y_{t(r,m)} + \psi^{(i)}(B)d_{T(r,m)} \tag{11}$$

where $\psi^{(i)}(B)$ is the respective weights of the outlying model stated in (7)–(10). Assuming that series $y_{t(r,m)}$ and $d_{T(r,m)}$ the mean and variance of outlier infested series are represented as follows: The mean of $z^{(i)}_{t(r,m)}$ is $E\left(z^{(i)}_{t(r,m)}\right) = E\left[y_{t(r,m)} + \psi^{(i)}(B)d_{T(r,m)}\right] = 0$ since $y_{t(r,m)}$ and $d_{T(r,m)}$ are mean-deleted,

$$E\left(z^{(i)}_{t(r,m)}\right) = \mu_{Z^{(i)}_{t(r,m)}} = 0 \tag{12}$$

The variance of $z^{(i)}_{t(r,m)}$ is the obtained from the autocovariance specification as:

$$E\left(z^{(i)}_{t(r,m)}z^{(i)}_{t-k(r,m)}\right) = E\left[\left(y_{t(r,m)} + \psi^{(i)}(B)d_{T(r,m)}\right)\left(y_{t-k(r,m)} + \psi^{(i)}(B)d_{T-k(r,m)}\right)\right]$$
$$\gamma^{(i)}_{z_{t(r,m)}}(k) = \gamma_{y_{t(r,m)}}(k) + 2\psi^{(i)}(B)\gamma_{y_{t(r,m)}d_{t(r,m)}}(k) + \psi^{(i)2}(B)\gamma_{d_{t(r,m)}}(k) \tag{13}$$

Setting $k = 0$ in Equation (13) gives the variance of $z^{(i)}_{t(r,m)}$ as

$$\operatorname{Var}\left(z^{(i)}_{t(r,m)}\right) = \sigma^2_{y_{t(r,m)}} + \psi^{(i)2}(B)\sigma^2_{d_{T(r,m)}} + 2\psi^{(i)}(B)\gamma_{d_{T(r,m)}y_{T(r,m)}}(0) \qquad (14)$$

It is evident from Equation (14), that for a mean deleted series the magnitude (d) of the outlier does not produce any effect on the mean of the resultant contaminated series since it gives same value as outlier free series. However, from Equation (14), the variance of the contaminated series does not equal to the variance of the outlier free series; the variance of outlier-free series is affected by the quantity $\psi^{(i)2}(B)\sigma^2_{d_{T(r,m)}} + 2\psi^{(i)}(B)\gamma_{d_{T(r,m)}y_{T(r,m)}}(0)$.

We remark that using autocovariance structure specified in Equation (13) to model periodic processes, will not be appropriate due to the presence of the quantity $2\psi^{(i)}(B)\gamma_{y_{t(r,m)}d_{t(r,m)}}(k) + \psi^{(i)2}(B)\gamma_{d_{t(r,m)}}(k)$ couple with the missing effect of the magnitude of outliers in the computation of variance of $z^{(i)}_{t(r,m)}$. Furthermore, to compute an estimate of $\tilde{\gamma}^{(i)}_{z'_{t(r,m)}}(k)$ adjusting the effect of outliers we use the expression:

$$\begin{aligned}&\tilde{\gamma}^{(i)}_{z'_{t(r,m)}}(k)\\ =&\tilde{\gamma}_{z_{T(r,m)}}(k) - \tilde{\gamma}_{y_{T(r,m)}}(k) - 2\psi^{(i)}(B)\tilde{\gamma}_{\tilde{\gamma}_{d_{T(r,m)}}(k)y_{T(r,m)}}(k) - \psi^{(i)2}(B)\tilde{\gamma}_{d_{T(r,m)}}(k)\end{aligned} \qquad (15)$$

where $\tilde{\gamma}^{(i)}(\cdot)$ is the estimate of the autocovariance for the specific series of interest. $\tilde{\gamma}_{z_{t(r,m)}}(k)$ is a measure of the under/over estimation of the covariance structure.

3.2. Alternative to Proposition 1

Now, assuming that the sequence $\left\{z^{(i)}_{t(r,m)}\right\}_{r\geqslant 1, m=1,\ldots,s}$ satisfies the difference Equations (7)–(10) such that $\left\{\xi^{t(r,m)}_T\right\}$ are independent random variables with probabilities $\Pr\left(\xi^{t(r,m)}_T = -1\right) = \Pr\left(\xi^{t(r,m)}_T = +1\right) = \frac{P_i}{2}$ and $\Pr\left(\xi^{t(r,m)}_T = 0\right) = 1 - P_i$; where we take $P_T = \frac{1}{n} \in (0,1)\forall T$ since each observation has equal chance of having an outlier. For all r and m, $\xi^{t(r,m)}_T$ and $y_{t(r,m)}$ are independent variables but $\xi^{t(r,m)}_T$ depends on d_T then the alternate specification of periodic autocovariance in the presence of outliers is:

$$\begin{aligned}E\left[z^{(i)}_{t(r,m)}z^{(i)}_{t-k(r,m)}\right] &= E\left[y_{t(r,m)}y_{t-k(r,m)}\right] + E\left[\psi^{(i)2}(B)d_Td_{T-k)}\xi^{2t(r,m)}_T\right]\\ &+ E\left[\psi^{(i)}(B)d_T\xi^{t(r,m)}_T y_{t-k(r,m)}\right] + E\left[y_{t(r,m)}\psi^{(i)}(B)d_{T-k}\xi^{t(r,m)}_T\right]\end{aligned}$$

Taking an assumption that if $\xi^{t(r,m)}_T$ and $d_T\xi^{t(r,m)}_T$ are independent of y_t; then

$$\begin{aligned}E\left[z^{(i)}_{t(r,m)}d^{(i)}_{t-k(r,m)}\right] &= \gamma_{y_{t(r,m)}}(k) + \psi^{2(i)}\sum_T d_Td_{T-K}\xi^{2t(r,m)}_T P(\xi)\\ &+ \psi^{(i)}(B)E\left(d_T\xi^{t(r,m)}_T\right)E\left(y_{t-k(r,m)}\right) + \psi^{(i)}(B)E\left(d_{T-k}\xi^{t(r,m)}_T\right)E\left(y_{t(r,m)}\right)\end{aligned} \qquad (16)$$

Since $E\left(y_{t(r,m)}\right) = E\left(y_{t-k(r,m)}\right) = 0$ and letting $\xi_T^{t(r,m)} = -1, 0$ and $+1$; we have Equation (16) reducing to

$$\gamma^{(i)}_{z_{t(r,m)}}(k) = \gamma_{y_{t(r,m)}}(k) + \psi^{(i)2}(B)\sum_T d_T d_{T-K} P_T = \gamma_{y_{t(r,m)}}(k) + \psi^{2(i)}(B)\gamma_{d_T} \quad (17)$$

If we then assume independence of $y_{t(r,m)}$ and $d_{T(r,m)}$ in Equation (13), then Equations (13) and (17) will give us the same quantity; and Var $\left(z^{(i)}_{t(r,m)}\right)$ will be $\sigma^2_{y_{t(r,m)}} + \psi^{(i)2}(B)\sigma^2_{d_T}$, where $\sigma^2_{y_{t(r,m)}} = \gamma_{y_{t(r,m)}}(0)$ and $\sigma^2_{d_T} = \gamma_{d_T}(0)$ in Equation (17).

3.3. Proposition 11

We further examine the mean and variance of the autocovariance of outliers infested series. Let us assume that there are τ number of outliers in the observations at time $t = 1, ..., n$ for a particular periods; to derive the variance of $\gamma^{(i)}_{z_{t(r,m)}}(k)$, we rely on Barttlet (1946) expression on variance of autocovariance and cross-covariance estimates. Suppose that $\tilde{\gamma}_{(\cdot)}(k)$ is the estimate of $\gamma_{(\cdot)}(k)$, then the estimate of $\gamma^{(i)}_{Z_{t(r,m)}}(k)$ is

$$\tilde{\gamma}^{(i)}_{z'_{t(r,m)}}(k) = \tilde{\gamma}_{y_{t(r,m)}}(k) + 2\psi^{(i)}(B)\tilde{\gamma}_{d_{t(r,m)}y_{t(r,m)}}(k) + \psi^{(i)2}(B)\tilde{\gamma}_{d_{t(r,m)}}(k) \quad (18)$$

Equation (18) gives the sample estimates of covariance of $z^{(i)}_{t(r,m)}$ defined in Equation (11); taking the expectation of (18) gives

$$E\left(\tilde{\gamma}^{(i)}_{z'_{t(r,m)}}(k)\right) = \gamma_{y_{t(r,m)}}(k) + 2\psi^{(i)}(B)\gamma_{d_{t(r,m)}y_{t(r,m)}}(k) + \psi^{(i)2}(B)\gamma_{d_{t(r,m)}}(k) \quad (19)$$

Equation (17) gives a bias estimate of $\gamma^{(i)}_{Z_{t(r,m)}}(k)$ due to the influence of the outliers in the series.

$$\begin{aligned}\text{MSE}\left[\tilde{\gamma}^{(i)}_{z_{t(r,m)}}(k)\right] = (n-k)^{-1}\Bigg[\sum_{j=-\infty}^{\infty}\Big\{&\gamma^2_{z_{t(r,m)}}(j)\gamma^2_{z_{t(r,m)}}(0) + 2\gamma^2_{z_{t(r,m)}}(k)\gamma^2_{z_{t(r,m)}}(j) \\ &+ \gamma_{z_{t(r.m)}}(k+j)\left(\gamma^2_{z_{t(r,m)}}(0)\gamma_{z_{t(r,m)}}(k-j)\right. \\ &- \left.\left.\left.2\gamma_{z_{t(r,m)}}(k)\left(\gamma^2_{z_{t(r,m)}}(0)\left(\gamma_{z_{t(r,m)}}(i) + \gamma_{z_{t(r,m)}}(-j)\right)\right)\right)\right\}\Bigg]\end{aligned} \quad (20)$$

The simplest expression without loss of generality is when $j = 0,\ \forall k > 0$; then Equation (20) becomes

$$\begin{aligned}&\text{MSE}\left[\tilde{\gamma}^{(i)}_{z_{t(r,m)}}(k)\right] \\ &= (n-k)^{-1}\left[\gamma^4_{z_{t(r,m)}}(0) + 2\gamma^2_{z_{t(r,m)}}(k)\gamma^2_{z_{t(r,m)}}(0) + \gamma^2_{z_{t(r,m)}}(k)\left(\gamma^2_{z_{t(r,m)}}(0) - 4\gamma^2_{z_{t(r,m)}}(0)\right)\right] \\ &= (n-k)^{-1}\gamma^2_{z_{t(r,m)}}(0)\left[\gamma^2_{z_{t(r,m)}}(0) - \gamma^2_{z_{t(r,m)}}(k)\right]\end{aligned} \quad (21)$$

Equations (18) and (20) show clearly the influence of outliers in both the mean and mean square error derivations; indeed a jacknifying appearance is recognized in Equations (18) and (20), hence could be used as a measure of MSE for the outlier contaminated autocovariance.

Where $\psi_i(i = 1,,4)$ is the weight attached to the value of outliers d_t in the expression for the periodic processes in Equations (5)–(9); this weight gives a measure of bias and precision contributed by the various outliers' processes discussed. It is evident from Equation (19) that outliers can affect the statistical properties of outlier-periodic autocovariance function (OPEACVF); this observation will lead to under or over estimation of the true structure. It is therefore, necessary to find models that can accommodate outliers at model building stage; this is discussed in the next subsection.

4. THE ESTIMATION OF ROBUST PEACVF WITH OUTLIERS

In section 3, we have shown that the estimates of OPEACVF are biased under different outlier schemes; since according to Sarnaglia *et al.* (2010) classical sample autocovariance function is very sensitive to the presence of outliers in data band it will be more appropriate to obtain a robust estimator of autocovariance in the presence of outliers.

Several authors, Sarnaglia *et al.* (2010), Fajardo *et al.* (2009), Jeromir (1994) have proposed some robust methods of estimating autocovariance relying on scaling factors of distribution functions; but the generalized Jacknife estimator utilized in Transfer Function Model (Shangodoyin, 2012), will be more appropriate because of the partitioning of the sequence of observations into various categories and using pseudo-values to approximate the estimates of the parameters of interest. In this sub section, we propose a modified Jacknife estimator of $\gamma^{(i)}_{y_{t(r,m)}}$.

4.1. Proposition III

Suppose that $\gamma^{(i)}_{z(\cdot)}(k)$ is the unknown parameter and $z_1, ..., z_n$ are n independent identically distributed periodic observations with outliers, let $\gamma^{(i)}_{z_0}(k)$ be an estimate of $\gamma^{(i)}_{(z)}(k)$ based on all the n observations and let $\tilde{\gamma}^{(i)}_{(z)_J}(k)$, $J = 1, ..., P$ be the estimate of $\gamma^{(i)}_{(y)}(k)$ obtained after the deletion of J-th groups of observations. Then $\tilde{\gamma}^{(i)}_{(z)_J}(k)$ is the estimate of $\gamma^{(i)}_{(y)}(k)$ from the remaining $(p-1)(l_J)$ observations. By using Tukey (1957) pseudo-values, then $\hat{\gamma}^{(i)}_{(z)_J}(k) = p\tilde{\gamma}^{(i)}_{(z)_0}(k) - (p-1)\tilde{\gamma}^{(i)}_{(z)_J}(k)$; In this study, $p = 2$, therefore we have $\hat{\gamma}^{(i)}_{(z)_J}(k) = 2\tilde{\gamma}^{(i)}_{(z)_0}(k) - \tilde{\gamma}^{(i)}_{(z)_J}(k)$; $J = 1, 2$. The Jacknife estimate of $\gamma^{(i)}_{(y)}(k)$ is the average of the $\tilde{\gamma}^{(i)}_{(z)_J}(k)$, $J = 1, 2$, as

$$\tilde{\tilde{\gamma}}^{(i)}_{(z)}(k) = \frac{1}{2}\sum_{J}^{2}\tilde{\gamma}^{(i)}_{(z)_J} = 2\tilde{\gamma}^{(i)}_{(z)_0}(k) - \frac{1}{2}\sum_{J=1}^{2}\tilde{\gamma}^{(i)}_{(z)_J}(k) \tag{22}$$

The Jacknife eliminates exactly n^{-1} bias term (Jeromir, 1994) and the result holds for all values of k if we can establish that:

$$E\left[\tilde{\gamma}^{(i)}_{(z)_0}(k)\right] = \gamma^{(i)}_{(y)}(k) + a(p(l_1+l_2))^{-1} + b(p(l_1+l_2))^{-2} + ...$$

where a and b are real constants. This is true for proposition II as shown in Equations (17) and (18). Since $\tilde{\gamma}_J^{(i)}(k)$ are approximately independently, identically distributed (Tukey, 1957), then $\{p(p-1)\}^{-1}\sum_{J=1}^{k}\left(\widehat{\gamma}_{(z)_J}^{(i)}(k)-\widehat{\overline{\gamma}}_{(z)}^{(i)}(\text{k})\right)^2$ is an approximate estimate of the Var $\left(\tilde{\overline{\gamma}}_{(z)}^{(i)}(k)\right)$ and

$$\left(\widehat{\overline{\gamma}}_{(z)}^{(i)}(k)-\gamma_{(y)}^{(i)}(k\right)\left\{\{p(p-1)\}^{-1}\sum_{J=1}^{K}\left(\widehat{\gamma}_{(z)_J}^{(i)}(k)-\widehat{\overline{\gamma}}_{(z)}^{(i)}(k)\right)^2\right\}^{-\frac{1}{2}}\sim t_{p-1} \tag{23}$$

for a specified level of α. Equation (21) will be used to determine the significance of $\widehat{\overline{\gamma}}_{(z)}^{(i)}(k)$.

5. ESTIMATING THE OUTLIER-PERIODIC PROCESSES

The maximum likelihood estimates has potential to handle complex models. In this paper we illustrate how to estimate the parameters of the proposed model using ML method and indeed the substitution of the proposed Jacknife autocovariance.

Let $z_{t(r,m)}=y_{t(r,m)}+\psi_i(B)d_{t(r,m)}$ as defined in Equation (6) through (9), then

$$z_{t(r,m)}=\sum_{i=1}^{p_m}\phi_i^{(m)}z_{t(r,m)-i}+\varepsilon_{t(r,m)} \qquad \text{(AR)} \tag{24}$$

$$z_{t(r,m)}=\sum_{j=1}^{q_m}\theta_i^{(m)}\varepsilon_{t(r,m)-j}+\varepsilon_{t(r,m)} \qquad \text{(MA)} \tag{25}$$

$$z_{t(r,m)}=\sum_{i=1}^{p_m}\phi_i^{(m)}z_{t(r,m)-i}+\sum_{j=1}^{q_m}\theta_j^{(m)}\varepsilon_{t(r,m)-j}+\varepsilon_{t(r,m)} \qquad \text{(ARMA)} \tag{26}$$

By using the conditional maximum likelihood estimation method, the parameters of the models (24)–(26) are obtainable.

Now consider the Equation (26); the complete form of model incorporating white noise into the series gives:

$$z_{t(r,m)}=\sum_{i=1}^{p_m}\phi_i^{(m)}z_{t(r,m)-i}+\sum_{j=0}^{q_m}\theta_j^{(m)}\varepsilon_{t(r,m)-j} \tag{27}$$

then, Equation (27) can be re-written as

$$\begin{aligned}z_{t(r,m)}=&\phi_1^{(m)}z_{t(r,m)-1}+...+\phi_{p_m}^{(m)}z_{t(r,m)-p_m}\\&+\theta_0^{(m)}\varepsilon_{t(r,m)}+\theta_1^{(m)}\varepsilon_{t(r,m)-1}+...+\theta_m^{(m)}\varepsilon_{t(r,m)-q_m}\end{aligned} \tag{28}$$

Let us assumed that $\{\varepsilon_{t(r,m)}\}$ are iid $N(0,\sigma_\varepsilon^2)$, and the joint density probability distribution $\underline{\in}^{(r,m)}=\left(\varepsilon_1^{(r,m)},....,\varepsilon_1^{(r,m)}\right)'$ is given by $P\left(\underline{\in}/\phi,\theta,\sigma_\varepsilon^2\right)=\left(2\pi\sigma_\varepsilon^2\right)^{\frac{n}{2}}\exp\left[-\frac{1}{2\pi\sigma_\varepsilon^2}\sum_{i=0}^{n}\varepsilon_t^2\right]$

By re-writing Equation (28) as:

$$\begin{aligned}\varepsilon_{t(r,m)} = &- \theta_1^{(m)}\varepsilon_{t(r,m)-1} - \ldots - \theta_m^{(m)}\varepsilon_{t(r,m)-q_m} \\ &+ z_{t(r,m)} - \phi_1^{(m)} z_{t(r,m)-1} - \ldots - \phi_{p_m}^{(m)} z_{t(r,m)-p_m}\end{aligned} \tag{29}$$

Now let $\mathbf{Y}^{*(r,m)} = (Z_1^{t(r,m)}, ..., Z_n^{t(r,m)})'$ and assume same in our initial conditions $\mathbf{Z}^* = (Z_{1-p}^*, ..., Z_{-1}^*, Z_{-0}^*)$ and $\mathbf{a}^* = (\varepsilon_{1-p}^*, ..., \varepsilon_{-1}^*, \varepsilon_{-0}^*)'$ then the conditional log − ln likelihood function is $\ln L_* \left(\phi, \theta, \sigma_\varepsilon^2\right) = -\frac{n}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{S_*\left(\underline{\phi}, \underline{\theta}\right)}{2\sigma_\varepsilon^2}$ where $S_*\left(\underline{\phi}, \underline{\theta}\right) = \sum_{i=0}^{n} \varepsilon_{t(r,m)}^2 \left(\underline{\phi}, \underline{\theta}/Z*, a, *Y*\right)$. Hence,

$$\ln L_* \left(\phi, \theta, \sigma_\varepsilon^2\right) = -\frac{n}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{i=0}^{n} \varepsilon_{t(r,m)}^2 \left(\underline{\phi}, \underline{\theta}/Z*, a, *Y*\right) \tag{30}$$

Now, by substituting Equation (28) into (29), we have

$$\begin{aligned}\ln L_* \left(\phi, \theta, \sigma_\varepsilon^2\right) = &- \frac{n}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\Sigma\Big[-\theta_1^{(m)}\varepsilon_{t(r,m)-1} - \ldots - \theta_m^{(m)}\varepsilon_{t(r,m)-q_m} \\ &+ z_{t(r,m)} - \phi_1^{(m)} z_{t(r,m)-1} - \ldots - \phi_{p_m}^{(m)} z_{t(r,m)-p_m}\Big]^2\end{aligned} \tag{31}$$

By differentiating partially the log-likelihood function L_* with respect to the parameters $\left(\hat{\theta}, \hat{\phi}\right)$, $\left(\frac{\partial \ln L_*}{\partial \theta_1^{(m)}} = \ldots = \frac{\partial \ln L_*}{\partial \theta_{q_m}^{(m)}} \text{ and } \frac{\partial \ln L_*}{\partial \phi_1^{(m)}} = \ldots = \frac{\partial \ln L_*}{\partial \phi_{q_m}^{(m)}}\right)$ and equating to zero will result in the following set of normal equations:

$$\begin{aligned}
&\theta_1^2 \sum \varepsilon_{t(r,m)-1}^2 + \ldots + \theta_1\theta_q \sum \varepsilon_{t(r,m)-1}\varepsilon_{t(r,m)-q} + \theta_1\phi_1 \sum \varepsilon_{t(r,m)-1} z_{t(r,m)} + \\
&\qquad \ldots + \theta_1\phi_p \sum \varepsilon_{t(r,m)-1} z_{t(r,m)-p} = 0 \\
&\vdots \\
&\theta_q\theta_1 \sum \varepsilon_{t(r,m)-1}\varepsilon_{t(r,m)-q} + \ldots + \theta_q^2 \sum \varepsilon_{t(r.m)-q}^2 + \theta_q\phi_1 \sum \varepsilon_{t(r,m)-q} z_{t(r,m)} + \\
&\qquad \ldots + \theta_q\phi_p \sum \varepsilon_{t(r,m)-1} z_{t(r,m)-p} = 0 \\
&\theta_1\phi_1 \sum \varepsilon_{t(r,m)-1} z_{t(r,m)-1} + \ldots + \theta_q\phi_1 \sum \varepsilon_{t(r,m)-q} z_{t(r,m)-1} + \phi_1^2 \sum z_{t(r,m)}^2 + \\
&\qquad \ldots + \theta_q\phi_p \sum \varepsilon_{t(r,m)-1} z_{t(r,m)-p} = 0 \\
&\vdots \\
&\theta_1\phi_P \sum \varepsilon_{t(r,m)-1} z_{t(r,m)-P} + \ldots + \theta_q\phi_P \sum \varepsilon_{t(r,m)-q} z_{t(r,m)-P} + \\
&\qquad \phi_1 \sum z_{t(r,m)-1} z_{t(r,m)-p} + \ldots + \phi_1^2 \sum z_{t(r,m)}^2 = 0
\end{aligned} \tag{32}$$

In matrix form, Equation (32) becomes

$$\begin{pmatrix}
\theta_1^2 \sum \varepsilon_{t-1}^2 & \theta_1\phi_q \sum \varepsilon_{t-1}\varepsilon_{t-q} & \theta_1\phi_1 \sum \varepsilon_{t-1}\varepsilon_{t-q} & \theta_1\phi_q \sum \varepsilon_{t-1}\varepsilon_{t-q} \\
\vdots & \vdots & \vdots & \vdots \\
\theta_q\phi_1 \sum \varepsilon_{t-1}\varepsilon_{t-q} & \theta_q^2 \sum \varepsilon_{t-q}^2 & \theta_q\theta_1 \sum \varepsilon_{t-q}\varepsilon_t & \theta_1\phi_q \sum \varepsilon_{t-1}\varepsilon_{t-q} \\
\theta_1\phi_q \sum \varepsilon_{t-1} z_{t-1} & \theta_q\phi_1 \sum \varepsilon_{t-1}\varepsilon_{t-q} & \phi_1^2 \sum z_t^2 & \phi_1\phi_q \sum z_{t-1} z_{t-p} \\
\vdots & \vdots & \vdots & \vdots \\
\theta_1\phi_p \sum \varepsilon_{t-1} z_{t-p} & \theta_q\phi_p \sum \varepsilon_{t-q} z_{t-p} & \theta_1\phi_p \sum z_{t-1} z_{t-p} & \theta_p^2 \sum z_{t-p}^2
\end{pmatrix}
\begin{pmatrix} n \\ \vdots \\ n \\ n \\ \vdots \\ n \end{pmatrix} = 0 \tag{33}$$

We can write the expression for the sum of cross products in matrix from as:

$$\begin{pmatrix} n\theta_1^2\sigma_\varepsilon^2 & \cdots & 0 & 0 & n\theta_1\phi_1\sigma_\varepsilon^2 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & n\theta_q^2\sigma_q^2 & 0 & 0 & \cdots & n\theta_q\phi_q\sigma_\varepsilon^2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ n\theta_1\phi_1\sigma_\varepsilon^2 & \cdots & 0 & n\phi_1\gamma_z(1) & n\phi_1^2\gamma_z(0) & \cdots & n\phi_1\phi_p\gamma_z(p-1) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & n\theta_q\phi_q\sigma_\varepsilon^2 & n\phi_1\gamma_z(p) & n\phi_1^2\gamma_z(p-1) & \cdots & n\theta_p^2\gamma_z(0) \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \underline{0} \tag{34}$$

Equation (34) gives system of equations to solve for θ and ϕ, then the compact form, where $S_{(\cdot)}$ means sum of squares of the products of variables, we have

$$\begin{pmatrix} \theta_i^2 S_{\varepsilon\varepsilon} & \theta_i\phi_j S_{z\varepsilon} \\ \vdots & \vdots \\ \phi_j\theta_i S_{\varepsilon z} & \phi_j^2 S_{zz} \end{pmatrix} \begin{pmatrix} nI \\ \cdots \\ nI \end{pmatrix} = \left[\underline{0}\right]$$

We consider a case when $p = q = 1$; using the first row of (34), we have $n\theta_1^2\sigma_\varepsilon^2 = -n\theta_1\phi_1\sigma_\varepsilon^2$ which gives the $\theta_1 = -\phi_1$; Also using the equation $n\phi_1\theta_1\sigma_\varepsilon^2 + n\phi_1\gamma_z(1) + n\phi_1^2\gamma_z(0) = 0$ to find ϕ_1, we have to divide the last expression by $n\phi_1$; this gives

$$\phi_1\gamma_z(0) + \theta_1\sigma_\varepsilon^2 = \gamma_z(1) \tag{35}$$

By using Equation (35) in the first row of (35) we have $\phi_1\left(\gamma_z(0) - \sigma_\varepsilon^2\right) = \gamma_z(1)$ and this gives

$$\left.\begin{aligned} \hat{\phi}_1 &= \frac{\gamma_z(1)}{\sigma_\varepsilon^2 - \gamma_z(0)} \\ \hat{\theta}_1 &= -\frac{\gamma_z(1)}{\sigma_\varepsilon^2 - \gamma_z(0)} \end{aligned}\right\} \tag{36}$$

The covariances of the outlier infested series shall be replaced with the robust estimate given in Equation (22); therefore, the robust estimates of ϕ_1 and θ_1 will be

$$\left.\begin{aligned} \tilde{\phi}_1 &= \frac{\hat{\tilde{\gamma}}_z^{(i)}(1)}{\tilde{\sigma}_\varepsilon^2 - \hat{\tilde{\gamma}}_z^{(i)}(0)} \\ \tilde{\theta}_1 &= \frac{-\hat{\tilde{\gamma}}_z^{(i)}(1)}{\tilde{\sigma}_\varepsilon^2 - \hat{\tilde{\gamma}}_z^{(i)}(0)} \end{aligned}\right\} \tag{37}$$

We shall use the conventional and Jacknife autocovariances in Equation (37) to demonstrate the efficacy of the proposed method.

6. EMPIRICAL REAL LIFE STUDIES

The data utilized in this study is precipitation data from Maun Airport in Botswana. The data was collected by METS and it is on monthly bases covering the period 1980–2010; a period of 30 years duration with 12 observations per each period. Precipitation in metrology (also known as one of the classes of hydrometers) is any product of the condensations of the atmospheric water vapors that falls under gravity. The main forms of precipitations include drizzle (sometimes called mist), rain, snow, graupel and hail. Precipitation occurs when a local portion of the atmosphere becomes saturated with water vapor, so that the water condenses and precipitates. The methods described in our propositions I to III above were applied to the mean deleted data, the procedure employed the use of Excel, SPSS and E-views at various stages of data analysis. The precipitations data is confirmed to have a number of outliers at some periods (see Table 1), thus it is an ideal data in this respect, however, periods without outliers are used for the study.

We first sought to artificially inject outliers into the series at randomly selected time points $T = 2, 6, 11, 18, 21, 24, 26, 27$ and 30. The magnitude of outliers taking arbitrarily is of order $0.75y_t$, this is done to circumvent the procedure of having to test for outliers at time T and estimate the magnitude of d_T; The basic assumption we have used in this empirical study is that since we know the magnitude of the outliers, then AO outliers series is taken as observed outlier contaminated series, thus we compare other outlying generating models with this and see which of these models best represents the periodic system under study. Secondly, we generate the AO, IO, LS and TC outliers' series using Equations (7)–(10). The generating models used are as follows:

$$
\begin{aligned}
\text{AO}: \quad z_t^{(1)} &= y_t + d_T \quad \text{where} \quad d_T = 0.75y_t;\\
&\quad T = 2,6,11,16,18,21,24,26,27,30 \quad \text{and} \quad d_T = 0 \ \forall\ t \neq T\\
\text{IO}: \quad z_t^{(2)} &= y_t + \left((1-0.904B)^{-1}(1-0.742B)\right) d_T\\
&= y_t + (1 + 0.16B + 0.67B^2)d_T\\
&= y_t + d_T + 0.16d_{T-1} + 0.67d_{T-2}\\
\text{LS}: \quad z_t^{(3)} &= y_t + (1-\delta)^{-1}d_T\\
&= y_t + (1+\delta)d_T\\
&= y_t + (1+0.001)d_T; \ \delta = 0.001\\
\text{TC}: \quad z_t^{(4)} &= y_t + (1-\delta B)^{-1}d_T\\
&= y_t + (1+0.001B)d_T\\
&= y_t + d_T + 0.001d_{T-1}
\end{aligned}
$$

Subsequently, we disaggregated these series into outlier-free series (y_t) and outlier-infested series (y_T). To empirically study our proposed methods, we obtained the autocovariances with the respective standard errors of AO, IO, LS and TC using Equations (13) and (21). In Table 2, we see that all the estimates of these autocovariances are not significantly different from zero (see the t-statistic), this may be due to the presence of outliers. Although if we purely use standard errors of the estimates which has a bench-mark for their significance, we see that AO, LS, TC have three (3) covariance estimates that seem significant while IO has four (4) significant estimates. The proposed Jacknife estimates accommodate the influence of the outliers, since almost all the estimates are significantly greater than zero in the Table 3. The results in Table 3 show that LS and TC perform in the same manner has AO, as evident in the lag at which the covariance estimates are significant, the IO has lesser number of estimates that significantly different from zero but the proposed Jacknifying has greater improvement on the relative number of estimates that are significant.

We fitted autoregressive model of order one, AR (1) to each of the AO, IO, LS and TC using the conventional autocovariance and Jacknife autocovariance estimates. In Table 4, we find that the estimates are not significant for conventional method whereas for Jacknife method, they are conspicuously significant, showing that the Jacknifying is resistant to the presence of outlying observations.

7. CONCLUSION

This paper has demonstrated the importance of evaluating autocovariance estimates in the specification of periodic processes. We evaluate autocovariance structures using various outliers' generating models. The analytical results indicate that outliers affect the estimates of periodic autocovariance function (PACVF) due to biases and inflated standard errors. Robust autocovariance structures that accommodate the influence of outliers in different periodic processes are proposed. We fit an AR (1) model using both the conventional and Jacknife autocovariance structures; the latter show high precision in the standard errors of the estimates. The proposed Jacknife autocovariance structures are more efficient than the conventional autocovariance structures in the presence outliers as evident from the real data analyzed.

REFERENCES

[1] Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Ser. B, 8*(Suppl.), 27–41.

[2] Jeromir, K. B. (1994). The linear model variance-covariance components and Jackknife estimation. *Application of Mathematics, 39*(2), 111–125.

[3] Chernick, M. R., Downing, D. J., & Pike, D. H. (1982). Detecting outliers in time series data. *Journal of the American Statistical Association, 77*, 743–747.

[4] Farjardo, F., Reisen, V. A., & Cribar-Neto, F. (2009). Robust estimation in long-memory processes under additive outliers. *Journal of Statistical Planning and Inference, 139*, 2511–2525.

[5] Battaglia, F., & Orfei, L. (2005). Outlier detection and estimation in nonlinear time series. *J. Time Ser. Anal., 26*, 107–121.

[6] Sarnaglia, A. J. Q., Reisen, V. A., & Lévy-Leduc, C. (2010). Robust estimation of periodic autoregressive processes in the presence of additive outliers. *Journal of Multivariate Analysis, 101*, 2168–2183.

[7] Shangodoyin, D. K. (2012). *Robust estimation of crosscovariance and specification of transfer function model in the presence of outliers leading indicator.* Seminar Presentation, University of Botswana.

[8] Tukey, J. W. (1957). Variances of variance components: II. The unbalanced single classification. *Ann. Maths. Statist., 28*, 43–56.

APPENDIX

Table 1
Periodic Evaluation of Outliers Using Spss

Period (M)	JAN (M=1)	FEB (M=2)	MAR (M=3)	APR (M=4)	MAY (M=5)	JUN (M=6)
Number of outliers	0	0	1	1	3	3
Period (M)	**JUL (M=7)**	**AUG (M=8)**	**SEP (M=9)**	**OCT (M=10)**	**NOV (M=11)**	**DEC (M=12)**
Number of outliers	6	6	14	0	1	1

Table 2
Periodic Evaluation of Outliers Using Spss

	AO	AO(SE)	IO	IO(SE)
Lag	COV(AO)	SE(C0V)	COV(IO)	SE(IO)
1	-2603.93	1761.024	-2629.29	2176.116
2	3478.142	1730.396	5198.243	2135.437
3	-1152.76	1699.215	-2120.51	2093.968
4	1300.386	1667.451	1965.109	2051.661
5	-1119.95	1635.07	-2266.21	2008.464
6	1872.194	1602.035	1822.794	1964.316
7	-1340.53	1568.304	-992.513	1919.154
	LS	**LS(SE)**	**TC**	**TC(SE)**
Lag	COV(LS)	SE(LS)	COV(TC)	SE(TC)
1	-2605.78	1762.285	-2725.8	1850.844
2	3480.5	1731.635	3586.797	1817.493
3	-1153.79	1700.431	-1240.23	1783.518
4	1301.168	1668.645	1319.055	1748.883
5	-1120.39	1636.241	-1177.64	1713.549
6	1873.602	1603.182	1893.412	1677.47
7	-1340	1569.427	-1430.75	1640.598

Table 3
Proposed Jacknife Estimates with Standard Errors

Lag	JKAO	JKAO(SE)	JKIO	JKIO(SE)
1	-3983.618838	346.3892587	-1159.54	3045.033
2	8967.041875	2085.537011	10674.39	569.0771
3	-1138.701891	224.7899154	-4090.38	929.2655
4	2154.89988	2660.934495	1395.097	2959.367
5	-1754.014414	1359.213892	-5017.11	1592.435
6	2224.128621	1337.207261	5231.874	840.2366
7	-1726.753626	463.9025323	-3080.95	2488.502
Lag	**JKLS**	**JKLS(SE)**	**JKTC**	**JKTC(SE)**
1	-3013.71	771.3389	-3221.28	733.8556
2	8325.289	892.6178	8500.364	927.9045
3	-1949.57	537.1987	-2114.41	544.0047
4	2157.142	2568.01	2062.075	2695.93
5	-2909.5	2317.115	-3023.68	2315.108
6	4720.689	804.203	4824.627	737.5807
7	-3893.25	1613.824	-4075.33	1609.61

Table 4
AR (1) Model Estimates Using Conventional(C) and Jacknife(JK) Autocovariances

	CAO	JKAO	CIO	JKIO
ϕ_1	-0.257	-0.393	-0.217	-0.096
S.E.	0.174	0.028	0.179	0.035
t-dist	-1.477	-13.952	-1.212	-2.700
Significance at $\alpha = 0.05$	Not sign	Sign	Not sign	Sign
	CLS	**JKLS**	**CTC**	**JKTC**
ϕ_1	-0.257	-0.248	-0.26	-0.318
S.E.	0.174	0.031	0.179	0.031
t-dist	-1.477	-7.938	-1.477	-10.248
Significance at $\alpha = 0.05$	Not sign	Sign	Not sign	Sign