Theoretical Mathematics & Applications, vol.4, no.3, 2014, 25-42
ISSN: 1792- 9687 (print), 1792-9709 (online)
Scienpress Ltd, 2014

Robust Test for detecting Outliers in Periodic Processes using Modified Hampel's Statistic

T.A. Lasisi[1] and D.K. Shangodouin[2]

Abstract

It is generally difficult to identify the exact outliers in periodic process and the suitable outlier detection method without a given underlying outlier process; also important is to discover the unusual data whose behavior is very exceptional when compared to the rest of the data set since the presence of the outlier can mar the model characterization techniques. Hampel suggested an identifier using the median to estimate data location and median absolute deviation to estimate the standard deviation. We apply the Modified Hampel Statistics by introducing the Jacknife method to the estimation of the parameters needed in Hampel detecting method to robust estimates. The two methods considered are implemented on-line and off-line points in finite samples taken for both real-life and simulated data using PAR (I) model. The results in both cases show that the Modified Hampel Statistic has higher rate of outlier identification for on-line detection. However, all the points identified by the Hampel method are also confirmed by the new Robust

[1] Department of Mathematics and Statistics, The Polytechnics Ibadan, Nigeria.
E-mail: lasisitaiwo40@yahoo.com
[2] Department of Statistics, University of Botswana, Botswana.
E-mail: shangodoyink@mopipi.ub.bw

Article Info: *Received:* March 13, 2014. *Revised:* May 12, 2014.
Published online : August 31, 2014.

Hampel Method. The robust method identified less off-line points than the Hampel Method, this further shows the effectiveness of our robust method in an attempt to reject off-line points that are falsely identified by Hampel method as outliers.

Mathematics Subject Classification: C52

Keywords: Outlier detection; Hampel test statistic; Robust Hampel method and Jacknife method

1 Introduction

In time series analysis, the first exploration of data pre-processing is to detect outliers. Following the work of Fox (1972) a number of research have been conducted on maximum likelihood based on their detections methods assuming known process (see Bianco et al 1996, Garcia et al 2001). It is generally difficult to know the exact outliers in periodic process and the suitable outlier's detection method without a given underlying outlier process. When data analysis is focused on either on-line or off-line points, operations are conducted without due cognizance to the presence of outliers and it could jeopardize the model characterization techniques. For instance, if a data set contains a single out-of-scale observation, most useful parameters and statistics for describing and characterizing data set deviate significantly (Liu et al, 2004).

In statistical analysis of time series, if multiple outliers exists, the single test is applied iteratively (Kelly 1988, and Hewitson 2006). However robust methods are reliable in detecting outliers in data set since they are designed to accommodate the influence of outliers and consequently make estimation resistant (Kniget and Way, 2011). The generalized extreme studentized deviate (Rosner, 1975) alleviates the primary limitations of Grubbs and Tietjen-Movre Test Statistic in that it relies on knowing the timing of the suspected outliers. However it is computationally difficult and has no visible truncation point. The study of influence function curve (Hampel, and Andrews et al, 1971) serves to deepen the understanding of estimators (e.g. of the relation between trimmed means, Winsorized means and Huber-estimators), and also serves to derive new estimators with specified robustness properties.

To robustly detect outliers in time series, Hampel (1994) suggested as identifier using the median to estimate data location and median absolute deviation to estimate the data standard deviation. An inclusion of Jacknife procedure to Hampel statistic when the outlying points are known may improve the detection rate of the statistic. In this paper, we introduce Jacknife method discussed in Lasisi, T.A et al (2013) section 4.1 PP 88 for the estimation of parameters needed in Hampel detecting method to give robust estimates. Since robust estimates are less susceptible to effect of outliers, we shall consider explicitly the Hampel test and its modifications to capture additive, innovative, level shift and transitory change outliers.

2 Statistical Methods of Detecting Outliers

Most statistical methods are generally not suitable to detect outliers in quantitative real- valued data except in situations when complex data transformations are required before actual data processing. In this section we consider Hampel method of detecting outliers and improve on this method to accommodate these limitations in terms of their applicability and robustness. We concentrate specifically on modifying this outlier test statistic to make it more resistant to effects of outliers.

2.1 Hampel's Methods and its Modification

The influence curve is essentially an earlier derivative of an estimator viewed as functional distribution as it has been shown that it can be used not only to derive asymptotic variances, but also to study several local robustness properties which are defined and intuitively in terpreted (see Hampel F.R., 1971). Let $\mathbb{R}$ be the real line and V be a real-valued functions defined on some subset of the of all probability measures on $\mathbb{R}$. Suppose that F is the probability measure on $\mathbb{R}$ for which V is defined. Denote by δ_y the probability measure determined by point mass 1 in any given point $y \in \mathbb{R}$. According to Hampel (1974), the mixture of F and some δ_y are written as $(1-\varepsilon)F + \varepsilon\delta_y$, $\forall\ 0<\varepsilon<1$. The influence curve $IC_{V,F}(\cdot)$ of the estimator V at the underlying probability distribution F is defined pairwise by

$$IC_{T,F(y)} = \lim_{\varepsilon\to\infty}\left\{V\left[(1-\varepsilon)F + \varepsilon\delta_y\right] - V(F)\right\}\Big/\varepsilon \tag{1}$$

Assuming that the limit in equation (5.1) is defined for every point $y \in \mathbb{R}$, Shangodoyin (1998), and Moeng et al (2009) have used equation (5.1) to determine the time occurrence of an outlier and the order of ARIMA model for a given series.

Suppose that we define $V = \int y dF(y)$ for all probability measure with existing first moment and let the mean of F exist and equal to μ; the influence curve of V is defined as

$$IC_{T,F(y)} = \lim_{\varepsilon\to 0}\left[(1-\varepsilon)\mu + \varepsilon y - \mu\right]\Big/\varepsilon = y - \mu \ (\forall y \in R) \tag{2}$$

The influence of an outlier on the value y say has been measured graphically using equation (5.2) (Gnanadesican and kettering, 1975) and given an estimator of autocorrelation function of ARMA model using a modification of equation (5.2). The scale estimate which is the counterpart of the mean as the most robust estimator of location for accessing the influence of outliers was proposed by Hampel(1974), and he suggested an identifier using the median to estimate the data, location and mean absolute deviation to estimate standard deviation. To robustly estimate the location and shape parameters the median and the median absolute deviation (MAD) are often recommended. The median is given as

$$Median(Z_N) = \frac{Z\left[(N+1)/2\right] + Z\left[N/2\right]}{2} \tag{3}$$

and the median absolute deviation is given as

$$MAD[Z_N] = median \text{ of } \left[\left|Z_1 - median(Z_N)\right|, \left|Z_2 - median(Z_N)\right|, \ldots, \left|Z_N - median(Z_N)\right|\right] \tag{4}$$

If we assume the robust estimator of the median as to be m_R obtained by using the u methodology derived in chapter four; then a modified Hampel method is Robust Median Absolute Deviation (RMAD) from median of

$$\left[\left|Z_1 - m_R\right|, \left|Z_2 - m_R\right|, \ldots, \left|Z_N - m_R\right|\right] \tag{5}$$

To obtain robust the parameter, suppose that the unknown parameter is $\mathcal{F}$, may be the median. Let $\mathcal{F}$ be the unknown parameter and $Z_1, \ldots, Z_n$ are n independent identically distributed periodic observations with outliers, let $\hat{\mathcal{F}}_{(Z)_0}$ be an estimate of $\mathcal{F}_{(Z)}$ based on all the n observations and let $\tilde{\mathcal{F}}_{(Z)_J}$, J=1,...,P be the estimate of $\mathcal{F}_{(Z)}$ obtained after the deletion of J-th groups of observations. Then $\tilde{\mathcal{F}}_{(Z)_J}$ is the estimate of $\mathcal{F}_{(Z)}$ from the remaining $(p-1)(l_J)$ observations. By using Turkey (1957) pseudo-values, then $\hat{\mathcal{F}}_{(Z)} = p\tilde{\mathcal{F}}_{(Z)_0} - (p-1)\tilde{\mathcal{F}}_{(Z)_J}$ In this study, $p = 2$, therefore we have $\hat{\mathcal{F}}_{(Z)_J} = 2\tilde{\mathcal{F}}_{(Z)_0} - \tilde{\mathcal{F}}_{(Z)_J}$ The Jacknife estimate of $\mathcal{F}_{(Z)}$ is the average of the $\hat{\mathcal{F}}_{(Z)}$, J=1,2, as

$$\hat{\hat{\mathcal{F}}}_{(Z)} = \frac{1}{2}\sum_{J}^{2} \hat{\mathcal{F}}_{(Z)} = 2\tilde{\mathcal{F}}_{(Z)_0} - \frac{1}{2}\sum_{J}^{2} \tilde{\mathcal{F}}_{(Z)_J} \tag{6}$$

A given observation $Z(.)$ is identified as outlier if

$$|Z(.)-MEDIAN\,(Z_N)| \geq g(N,\alpha_N)MAD(Z_N) \tag{7}$$

Where g a function is related to the number of data points and α_N is specified rejection level (Davies and Gather, 1996).

Now, without loss of generality, we use the quantity

$$|Z(.)-RMEDIAN(Z_N)| \geq g(N,\alpha_N)RMAD(Z_N) \tag{8}$$

for the robust Hampel test statistic.

According to Davies and Gather (1996), to select $g(\alpha_N, N)$, we assume that the observation Z follows normal distribution $N(\mu_Z, \sigma_Z^2)$; an outlier identifier is defined by specifying a lower bound $L(Z_N, \alpha_N)$ and an upper bound $R(Z_N, \alpha_N)$ that depend on the sample and chosen the value of $\alpha_N = 1-(1-\alpha)^{1/N}$. All points less than $L(Z_N, \alpha_N)$ or greater than $R(Z_N, \alpha_N)$ will be defined as lying in the outlier region $OUT(\alpha_N, \mu_Z, \sigma_Z^2)$; thus the set of all Z_T identified to be α_N outliers by identifier is given by $OR(Z_N, \alpha_N) = [-\infty, L(Z_N, \alpha_N)]$ or $[R(Z_N, \alpha_N), \alpha_N]$. Davies and Gather (1996) have suggested by simulations, a way of specifying the function $g(\alpha_N, N)$ in equations (5.4) and (5.5); the method is to standardize an outlier identifier for a given for given α_N such that

$$P\left(OR(Z_N,\alpha_N) \subset out\left(\alpha_N, \mu_Z, \sigma_Z^2\right)\right) = 1-\alpha = 0.95 \tag{9}$$

this probability is calculated under assumptions of iid sample Z_N. The simulations resulted into various values of $g(\alpha_N, N)$ for $N = 20, 50$ and 100. Davies and Gather (1996).

3 Empirical Illustrations

The two methods considered in this chapter are implemented on-line and off-line outlier points in finite samples taken for both real-life and simulated PAR (1) data. The outlier detection rates in both on-time and off-time are evaluated. In declaring there is an outlier at time point t, for series Z_t, the condition is that all Z_t for which we have

$$\left|Z_t - MED(Z_N)\right| \geq g(N,\alpha_N)MAD(Z_N)$$

and

$$\left|Z_t - RMED(Z_N)\right| \geq g(N,\alpha_N)RMAD(Z_N)$$

is confirmed as outliers for Hampel and Robust identifier respectively.

3.1 Real-life Data

The data collected on precipitation for Maun Airport were utililized for three periods, namely, January, February and October; these being the only months for which h the initial check-up of outliers in series using SPSS did not show the presence of outliers. For these periods, the actual outliers are added at time points 2,6,11,16,18,21,24,26,27 and 30 in the series using $1.75(y_t)$ is Z_t. The functional form of the critical values of the test statistic used is based on Davies and Gather (1993) simulated values of $g(30,0.05)$ is 2.88.

In Table 1, we have generated the results of the test statistic for January; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.6 than Hampel method with maximum of 0.15. All the points identified by Hampel method are also confirmed by our new robust Hampel method. Although the robust method identified more off-time points than the Hampel method but the rate of identification is with a maximum of 0.15. The AO, TC and IO re spectively perform better with a collective maximum of 0.6 and minimum of 0.4 on-time point’s identification for robust method; while these outlier models record a maximum of 0.3 and minimum of 0.2 on-time points confirmation for Hampel method.

The LS fails to identify any time points as outliers for both robust Hampel and conventional Hampel methods; one may conclude that the January period is prone to AO, TC, and IO outlier generating models.

Table 1: January Results For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUE	161.84	185.92	249.13	162.12
RMAD CUT-OFF VALUE	188.16	134.54	253.19	187.80
MAD ON-TIMING POINTS	12,19,31(3)	12,31(2)	NIL	12,31(2)
MAD OFF-TIMING POINTS	NIL	15(1)	NIL	15,19(2)
RMAD ON - TIMING POINTS	12,17,19,25,27 31 (6)	12,17,25,31(4)	NIL	12,17,25,27,31 (5)
RMAD OFF-TIMIG POINTS	10,15,21(3)	10,15(2)	NIL	15,19,21(3)

In Table 2, we also generated the results of the test statistic for February; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.7 than Hampel method with maximum of 0.5. All the points identified by Hampel's method are also confirmed by our new robust Hampel method. Although the robust method identified more off-time points than the Hampel method, but the rate of identification is with a maximum of 0.25. It is also evident that for February period, all the outlier generating models perform better with an overall maximum of 0.7 and minimum 0.4 on-time points identification for robust method, whereas these outlier models record a maximum of 0.5 and minimum of 0.2 on-time confirmation for Hampel method. The AO, LS and TC fail to identify off-time points as outliers for both robust Hampel and conventional Hampel method; this make these approaches more acceptable is in practice.

Table 2: February Results For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUE	141.96	132.72	185.36	163.52
RMAD CUT-OFF VALUE	127.96	175.84	136.08	116.48
MAD ON-TIMING POINTS	12, 17, 27, 31(4)	12, 17, 19, 27, 31(5)	17,27,31(3)	17,27,31(3)
MAD OFF-TIMING POINTS	NIL	29(1)	NIL	NIL
RMAD ON - TIMING POINTS	12,17,22,25,27,31 (6)	12,17,19,22,25,27,31 (7)	12,17,27,28, 31(5)	12,17,27,31(4)
RMAD OFF-TIMING POINTS	2,10,21(3)	10(1)	NIL	NIL

In Table 3, we also generated the results of the test statistic for October period; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.6 than Hampel method with maximum of 0.4. All the points identified by Hampel method are also confirmed by our new robust Hampel method. Although the robust method identified more off-time points than the Hampel method, the rate of identification is with a maximum of 0.15. For October period, all the outlier generating models perform better with an overall max of 0.6 and minimum 0.4 on-time points identification for robust method, whereas these outlier models record a maximum of 0.4 and minimum of 0.2 on-time confirmation for Hampel method. All the generating models identify off-time points as outliers for both robust Hampel and conventional Hampel method but with LS and TC having more identifying prowess than both AO and IO.

Table 3: October Results For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUE	27.16	35.28	40.04	162.12
RMAD CUT-OFF VALUE	36.12	33.32	27.72	187.80
MAD ON-TIMING POINTS	3,7,12,22,28(4)	3,7,12,18(4)	3,7(2)	12,31(2)
MAD OFF-TIMING POINTS	30(1)	5(1)	4,8(2)	15,19(2)
RMAD ON - TIMING POINTS	3,7,12,27,28(5)	3,7,12,18,22,28(6)	3,7,12,22,28(5)	12,17,25,27,32(5)
RMAD OFF-TIMING POINTS	30(1)	5(1)	4,8,30(3)	15,19,21(3)

3.2 Simulated Studies

The simulation results are obtained under 500 replications of normal random series of size 50 for PAR (1) processes generated using the following equations:

$$Y_{t(r,1)} = 0.15Y_{t(r,m)-1} + \varepsilon_{t(r,1)}$$
$$Y_{t(r,2)} = -0.25Y_{t(r,2)-1} + \varepsilon_{t(r,2)}$$
$$Y_{t(r,3)} = 0.35Y_{t(r,3)-1} + \varepsilon_{t(r,3)}$$
$$Y_{t(r,4)} = 0.45Y_{t(r,4)-1} + \varepsilon_{t(r,4)}$$

The autoregressive model parameters are;

$$\phi_1^{(1)}=0.15,\ \phi_1^{(2)} = -0.25,\ \phi_1^{(3)} = 0.35 \text{ and } \phi_1^{(4)} = 0.45$$

The simulation model for the periodic series $y_{t(r,m)}$ is Normal Rand ()*(b-a) $+\varepsilon$ for a=0.1 and b=2; we generate the error terms using simulation model Normal Rand (sigma=1, miu=0). The magnitude of outliers are injected using $D_{t(r,m)} = 0.75Y_{t(r,m)}$ (for various outlier models) at time points $t = \{T : 5,10,15,20,25,30,35,40,45 \textit{ and } 50\}$

We have generated the results of the test statistic for first quarter; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.6 than Hampel method with maximum of 0.5. All the points identified by Hampel method are also confirmed by our new robust Hampel method. Although the robust method identified more off-time points than the Hampel method but the rate of identification is with a maximum of 0.125. The AO, TC and IO respectively perform better with a collective maximum of 0.6 and minimum of 0.5 on-time point's identification for robust method; these outlier models however, record a flat rate of 0.5 on-time point's confirmation for Hampel method. The LS fails to identify any on-timing points as outliers for robust Hampel, it however, identifies four on-timing points for conventional Hampel method. They both identify two off-timing points each. One may again conclude that the first quarter period is prone to AO, TC, and IO outlier generated models.

Table 4: Simulated Period 1 For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUES	1.3839	1.3839	2.0928	1.6692
RMAD CUT-OFF VALUES	1.5185	1.5185	2.0455	1.6769
MAD ON-TIMING POINTS	6,11,26,36,51(5)	6,11,26,36,51(5)	6,26,36,51(4)	6,11,26,36,51(5)
MAD OFF-TIMING POINTS	3,15,20,39,38(5)	3,15,38,39(4)	37,39(2)	15,19,39(3)
RMAD ON-TIMING POINTS	6,11,26,36,41,51 (6)	6,11,26, 36,51(5)	NIL	6,11,26,36,51(5)
RMAD OFF-TIMINGPOINTS	3,15,20,38,39(5)	3,15,39(3)	13,30(2)	3,15,39(3)

We have generated the results of the test statistic for second quarter; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.5 than Hampel method with maximum of 0.3. All the points identified by Hampel method are also confirmed by our new robust Hampel method. Although the robust method identified no off-timing points, the Hampel method identified three and one off-timing points respectively for both LS and IO. Only the LS appears to give the best performance in this second quarter period with a maximum of 0.6 and minimum of 0.3 on-time points identification for robust method. The AO, IO and TC fail to identify any off-time points for robust method but the Hampel method slightly identifies at the rate of 0.001 off-time points for the LS. At this juncture, it is evident that LS is considerably affected by the outlier generating models.

Table 5: Simulated Period 2 For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUES	3.0983	3.0026	3.2316	3.0977
RMAD CUT-OFF VALUES	3,7589	3.1780	2.8262	3.7549
MAD ON-TIMING POINTS	11,36,51(3)	11,36,51(3)	36(1)	11,36,51(3)
MAD OFF-TIMING POINTS	NIL	9(1)	9,12,37(3)	NIL
RMAD ON-TIMING POINTS	11,36,51(3)	11,31,36,46,51(5)	11,31,36,46,51(5)	11,36,46,51(4)
RMAD OFF-TIMINGPOINTS	NIL	NIL	12,37(2)	NIL

We have generated the results of the test statistic for quarter three; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.6 than Hampel method with maximum of 0.4. Although the Hampel method identifies more off-time points than the robust, the rate of identification is with a maximum of 0.1. All the points identified by Hampel method are also confirmed by our new robust Hampel method. AO and LS collectively perform better with a maximum of 0.6 and minimum of o.3 for on-time point's identification for robust, these outliers, however, record a maximum of 0.4 and minimum of 0.2 on-time points confirmation for Hampel method. The IO identifies more off-time points at the maximum rate of 0.1 and minimum of 0.04. The robust method only identifies off-time points at the rate of 0.06 maximum and 0.04 minimum.

Here, LS, AO and IO are seen to respectively perform better with LS taking the lead while TC is mostly affected by the outlier generating models in this context.

Table 6: Simulated Period 3 For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUES	2.8435	2.8167	4.1317	3.1398
RMAD CUT-OFF VALUES	3.4524	4.1450	3.5578	2.9478
MAD ON-TIMING POINTS	11,26,41(3)	11,16,26,41(4)	11,41(2)	26,41(2)
MAD OFF-TIMING POINTS	4,9,42(3)	2,4,9,15,42(5)	4,15(2)	2,4,13,15(4)
RMAD ON-TIMING POINTS	6,11,26,41,51(5)	11,16,26,41(4)	11,21,31,26,41, 51(6)	16,26,41(3)
RMAD OFF-TIMING POINTS	4,9,42(3)	9,42(2)	4,15(2)	4,15(2)

We have generated the results of the test statistic for quarter four; the robust method has a higher rate of outlier identification for on-time detection with a maximum of 0.8 than Hampel method with maximum of 0.7. All the points identified by Hampel method are also confirmed by our new robust Hampel method. Although the robust method identified less off-time points than the Hampel method but the rate of identification is with a maximum of 0.08. All the generating models appear to perform equally likely with a collective maximum rate of 0.8 and minimum of 0.7 for on-time identification for the robust method, while all the models record a maximum of 0.7 and minimum of 0.4 on-time confirmation for Hampel method. The IO fails to identify any off-time points for robust, it however, identifies a maximum of a maximum off-time points rate of 0.02 for Hampel method .The LS identifies a off-time points rate at the maximum of 0.08 and maximum of 0.06 respectively for both robust and Hampel methods. AO and TC have off-time points with maximum rate of 0.04 and minimum of 0.01 for the two methods.

Table 7: Simulated Period 4 For Hampel And Hampel Modified Test Statistic

TEST STATISTIC	AO	IO	LS	TC
MAD CUT-OFF VALUES	3.2872	3.6098	3.7926	3.2844
RMAD CUT-OFF VALUES	3.9767	4.5599	4.3105	3.9721
MAD ON-TIMING POINTS	16,21,31,41,46,51(6)	16,21,39,51(4)	6,16,21,31,41,46(6)	6,16,21,31,41,46,51(7)

MAD OFF-TIMING POINTS	5,10(2)	5(1)	5,10,22,32(4)	5,10(2)
RMAD ON-TIMING POINTS	6,11,16,21,31, 41,51(7)	6,11,16,21,26,3 1,51(7)	11,16,21,26,3 1,41,51(7)	6,11,16,21,31 , 41,46,51(8)
RMAD OFF-TIMING POINTS	5(1)	NIL	5,22,32(3)	51(1)

4 Conclusion

In using Hampel modified approach to detect outliers, the two methods considered are implemented on-line and off-line points in finite samples taken for both real-life and simulated data using PAR (I) model. The results in both cases show that the Modified Hampel Statistic has higher rate of outlier identification for on-line detection. However, all the points identified by the Hampel method are also confirmed by the new Robust Hampel Method. The robust method identified less off-line points than the Hampel Method; this further shows the effectiveness of our robust method in an attempt to reject off-line points that are falsely identified by Hampel method as outliers.

References

[1] Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H. and Tukey, J.W., *Robust Estimates of Location,* Survey and Advances, Princeton Univ. Press, Princeton NJ., 1972.

[2] Bianco, A.M, Garcia. B.M.G., Martinez, E.J., & Yohai, V.J., Robust procedure for regression models with ARIMA errors, in COMPSTAT 96, *Proceedings in Computational Statistics Part* ABerlin, Physica- Verlags, (1996), 27-38.

[3] Bianco, A.M, Garcia. B.M.G., Martinez, E.J., & Yohai, V.J., Outlier detection in regression models with ARIMA errors using robust estimations, *Journal of forecasting*, **20**, (2001), 365-579.

[4] Devlin, S.J. Gnanadeskan, R. and Kettenring, J.R, Robust Estimation and Outlier Detection with correlation Coefficients, *Biometrika*, **62**(3), (1975), 531-545.

[5] Fox, A.J, Outliers in time series, *Journal of Royal Stat. Soc.*, **34**, (1972), 350-363.

[6] Hampel, F.R., A general qualitative definition of robustness, *Annals of mathematics Statistics*, **42**, (1971), 1887-1896.

[7] Hampel, F.R., The influence curve and its role in robust estimation, *Journal of American Statistical Association,* **69**, (1974), 382-393.

[8] Hewitson, S.W.J., GNSS Receiver Autonomous Integrity Monitoring (RAIM), *6 Performance Analysis, GPS Solut*., **10**(3), (2006), 155-170.

[9] Kelly, R.J., The Linear Model, RNP, and the Near-Optimum Fault Detection and Exclusion Algorithim, in: Global Positioning System, **5**, *The Institute of Nvigation*, Fairefax,Virginia, (1998), 277-263.

[10] Lasisi, T.A., Shangodoyin, D.K., and Moeng, S.R.T., Specification of Periodic Autocovariance Structures in the Presence of Outliers, *Studies in Mathematical Sciences*, **6**(2), (2013),83-95.

[11] Liu, H., Shah, S. and Jiang,W., On-line outlier detection and data cleaning, *Computer and Chemical Engineering*, **28**, (2004), 1635-1647.

[12] Rosner, B., On the detection of many outliers, *Technometrics*, **17**(2), (1975), 221-227.

[13] Shangodoyin, D.K., The use of correlation function for detecting outliers and model order determination, *Journal of Nigerian Statisticians*, (1998), 22-28.