International Journal of Business and Management Invention (IJBMI)
ISSN (Online): 2319 – 8028, ISSN (Print): 2319 – 801X
www.ijbmi.org || Volume 8 Issue 04 Series. V || April 2019 || PP 77-83

Quality Control and Outliers in Manufacturing Processes

Claude R. Superville, PhD, FRSS, FIMA
Texas Southern University JHJ School of Business Houston, TX 77004

ABSTRACT : *Manufacturing processes that consist of time series data are frequently monitored by forecast-based quality control schemes. These control schemes are based on the application of a time series forecast to the process and monitoring the resultant forecast errors with a control chart or tracking signal. This study compares the performance of the Individuals control chart, the Cumulative Sum (CUSUM), the Exponentially Weighted (EWMA) chart, the Smoothed Error (ETS) and Cumulative Sum (CTS) tracking signals in their ability to detect the presence of additive outliers in an autocorrelated process. The Individuals chart offers the greatest probability of early detection of an additive outlier in an autocorrelated process, based on the CDF criterion.*

KEYWORDS*: control charts, tracking signals, antocorrelation*

Date of Submission: 05-05-2019 Date of acceptance: 20-05-2019

I. INTRODUCTION

The occurence of large unusual observations is not uncommon in time series data. These outliers may be due to recording errors or to one-time unique situations such as an unexpected change in demand for a product or a change in a production system. Fox (1972) defines two types of outliers may occur in practice. An additive outlier corresponds to an external disturbance that affects the value of a single observation. An innovational outlier refers to an internal disturbance that changes the value of an observation and all other successive observations. Typically, in process control environments, monitoring schemes are compared based on their ability to detect step shifts or innovational outliers in the level of a process. However, which monitoring scheme detects the presence of an additive outlier most quickly is also of interest.

Autocorrelation implies the existence of a relationship between consecutive observations and can be of two types. A process that tends to drift over time is characteristic of positive autocorrelation and results when successive observations are similar in value. Negative autocorrelation is depicted by a sawtooth pattern and results when consecutive observations are dissimilar. High volume manufacturing processes along with an increased frequency of sampling by automated gages, gives rise to autocorrelated data.

The presence of autocorrelation creates unique problems for process monitoring schemes. Positive autocorrelation tends to increase the frequency of out-of-control signals that are detected by monitoring schemes. Positive autocorrelation occurs most often in production environments and chemical operations (Woodall and Faltin (1993)).

In the field of statistical process control (SPC), control charts have traditionally been used to monitor production processes. In the forecasting and time series fields, tracking signals perform a similar function, the monitoring of forecasting systems. The statistical tools are similar in that both are designed to monitor systems and provide information concerning changes in the systems.

Alwan and Roberts (1988) have proposed a method for monitoring autocorrelated data that involves the application of a time-series forecast to the process and monitoring the forecast errors. Unusual behavior in the process should result in a large error that is reflected as a signal on a control chart or tracking signal.

Traditionally, monitoring tools have been compared on the basis of Average Run Lengths (ARLs). The ARL is the expected number of observations required to detect an out-of-control situation. However, simple exponential smoothing forecasts recover quickly from step increases in the time series process that it monitors. This would suggest that the performance of forecast-based schemes should be based on the probability of "early detection". As an average measure that is inflated by long run lengths, the ARL is an inadequate measure of quick recovery, that is characterized by short run lengths. Hence the cumulative distribution function (CDF) of the run lengths is offered as an alternative criterion to the average run length (ARL) for the selection of an appropriate monitoring scheme. The CDF provides the cumulative probability of a signal occuring by the ith time period after a disturbance.

This paper compares the performance of tracking signals and control charts in monitoring residuals from exponential smoothing forecasts applied to autoregressive process data of order one, denoted by AR(1), in the presence of additive outliers. The study shows that the Individuals control chart offers the highest probability of early detection of an additive outlier in AR(1) processes.

II. LITERATURE REVIEW

The presence of autocorrelation creates unique problems for process monitoring schemes. Positive autocorrelation tends to increase the frequency of out-of-control signals that are detected by monitoring schemes. Positive autocorrelation occurs most often in production environments and chemical operations (Woodall and Faltin (1993)).

The performance of control charts in the presence of autocorrelation has been explored by a number of authors. Superville and Adams (1994) compared the performance of an Individuals Chart, a Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) Chart in their ability to detect step shifts in autocorrelated process. Superville and Adams (1995) compared the performance of these charts to tracking signals in their ability to detect step shifts in autocorrelated process. Lu and Reynolds (1999) suggest the use of a combined Shewhart-EWMA for autocorrelated data. Lianjie, Daniel and Fugee (2002) suggest the use of a triggered CUSCORE on residuals. Lee et al. (2009) propose distribution-free charts for monitoring shifts in the mean of autocorrelated processes. Wu and Yu (2010) advocate a neural network approach for monitoring the mean and variance of an autocorrelated process. Chang and Wu (2011) suggest a Markov Chain approach to calculating the ARL for control charts on autocorrelated process data.

In the field of statistical process control (SPC), control charts have traditionally been used to monitor production processes. In the forecasting and time series fields, tracking signals perform a similar function, the monitoring of forecasting systems. The statistical tools are similar in that both are designed to monitor systems and provide information concerning changes in the systems.

Alwan and Roberts (1988) have proposed a method for monitoring autocorrelated data that involves the application of a time-series forecast to the process and monitoring the residuals. Unusual behavior in the process should result in a large error that is reflected as a signal on a control chart or tracking signal.

Traditionally, monitoring tools have been compared on the basis of Average Run Lengths (ARLs). The ARL is the expected number of observations required to detect an out-of-control situation. However, simple exponential smoothing forecasts recover quickly from step increases in the time series process that it monitors. This would suggest that the performance of forecast-based schemes should be based on the probability of "early detection". As an average measure that is inflated by long run lengths, the ARL is an inadequate measure of quick recovery, that is characterized by short run lengths. Hence the cumulative distribution function (CDF) of the run lengths is offered as an alternative criterion to the average run length (ARL) for the selection of an appropriate monitoring scheme. The CDF provides the cumulative probability of a signal occurring by the ith time period after a disturbance.

III. A MODEL FOR AUTOCORRELATED DATA

A time series model that has been found to be useful in production and quality control environments is the ARIMA(1,0,0), referred to as the first-order autoregressive model and denoted by AR(1). It is represented by

$$X_t = \xi + \phi X_{t-1} + \varepsilon_t . \qquad (1)$$

Without loss of generality it is assumed that $\varepsilon_t \sim N(0,1)$. It is also assumed that an AR(1) model is applicable in this article. Montgomery and Mastrangelo (1991) show that a number of chemical and manufacturing processes conform to this model.

The simple exponential smoothing forecast, also known as an exponentially weighted moving average (EWMA) forecast is given by

$$F_{t+1} = \alpha_F X_t + (1-\alpha_F)F_t , \qquad 0 \leq \alpha_F \leq 1 . \qquad (2)$$

where X_t represents the process observation at time period t, and F_{t+1} represents the one-step-ahead forecast for observation X_{t+1} at time period t. The forecast error at time period t, denoted by e_t, is defined as

$$e_t = X_t - F_t. \qquad (3)$$

Alwan and Roberts (1988) have observed that processes that do not drift too rapidly are well modeled by simple exponential smoothing. For the AR(1) model, Cox (1961) has shown that optimal simple exponential smoothing in terms of minimum mean square forecast error is given by

$$\alpha_F = 1 - ½[(1-\phi)/\phi], \qquad 1/3 < \phi \leq 1 \qquad (4)$$

where ϕ is the parameter of the AR(1) process. The simulation study on which this article is based, rely on this result.

IV. FORECAST-BASED QUALITY CONTROL SCHEMES

In this study, the Individuals, Cumulative Sum (CUSUM) and EWMA control charts and the Smoothed Error (ETS) and Cumulative Sum (CTS) tracking signals are applied to exponential smoothing forecast errors and their performances evaluated.

4.1 The Individuals Control Chart

The Individuals control chart applied to forecast errors requires an estimate of the variance of the forecast errors. Defining the ith moving range to be

$$MR_i = |e_i - e_{i-1}|, \quad i = 2, 3, ..., m \tag{5}$$

and

$$\overline{MR} = \frac{1}{m-1}\sum_{i=2}^{m} MR_i, \tag{6}$$

the control limits are

$$\overline{X} \pm C_1 MR/d_2 \tag{7}$$

where the constant C_1 is set to achieve a desired in-control ARL. Montgomery (1991) has tabulated values for C_1 and d_2.

4.2 The Cumulative Sum Control Chart

An alternative to the Shewhart control chart is the Cumulative Sum (CUSUM) control chart. The CUSUM control chart may be represented by either a V-mask representation or equivalently by the use of two one-sided cumulative sums.

The V-mask form of the CUSUM applied to forecast errors requires plotting the quantity

$$S_i' = \sum_{j=1}^{i} e_j, \quad i = 1, 2, \ldots \tag{8}$$

against the sample number i.

The 'two one-sided cumulative sums' procedure also known as the Tabular CUSUM requires calculating:

$$S_i = \max\ 0, (e_i/\sigma_e) - K + S_{i-1} \tag{9}$$

$$T_i = \min\ 0, (e_i/\sigma_e) + K + T_{i-1} \tag{10}$$

where S_0 = w and T_0= –w($0 \le w < K$). The value $_e$ represent the standard deviation of the forecast errors, which is typically estimated in practice. A head start value, w, is recommended for earlier detection of out of control situations. In this study w = 0 is used. The reference value K is usually set to be $\delta/2$, where δ is the smallest shift in the mean (measured in forecast error standard deviations) considered important to be detected quickly. If $S_i > h$ or $T_i < -h$ (where h is a critical value) the chart signals. The critical values, h used in this study were determined through simulation.

4.3 The Exponentially Weighted Moving Average Control Chart

The Exponentially Weighted Moving Average (EWMA) control chart is another alternative to the Shewhart control chart that has been found to be more sensitive to small process disturbances. Also known as a Geometric Moving Average control chart, the EWMA applied to forecast errors is a weighted average of past and present data given by

$$Z_i = \lambda e_i + (1-\lambda) Z_{i-1}, \quad i = 1, 2, ... \tag{11}$$

where $\lambda(0 < \lambda < 1)$ is a smoothing constant and $Z_0 = 0$ usually. Assuming that the process is in control and the observations are independent then

$$\sigma^2_{Zi} = \frac{\sigma^2}{n} \frac{\lambda}{2-\lambda} [1-(1-\lambda)^{2i}] \tag{12}$$

with control limits determined by

$$\overline{X} \pm c\sigma_{Zi} \tag{13}$$

where c is a constant designed to achieve a desired in-control ARL. The value of c required for this study was determined through simulation. Note if $\lambda = 1$, the EWMA control chart becomes a Shewhart control chart.

4.4 The Smoothed Error Tracking Signal

Trigg's (1964) Smoothed Error (ETS) tracking signal is given by

$$ETS_t = |E_t / MAD_t| \tag{14}$$

where

$$E_t = \alpha_1 e_t + (1-\alpha_1)E_{t-1}, \quad 0 \le \alpha_1 \le 1 \tag{15}$$

and

$$MAD_t = \alpha_2 |e_t| + (1-\alpha_2)MAD_{t-1}, \quad 0 \le \alpha_2 \le 1. \tag{16}$$

Typically, $E_0 = 0$ and MAD_0 is set equal to its expected value which is approximately equal to $0.8\sigma_e$ (where σ_e is the standard deviation of the forecast errors). A signal occurs if ETS_t exceeds a critical value K_1. Gardner (1983) suggests that the value of K_1 should be set to achieve a desired in-control ARL.

4.5 The Cumulative Sum Tracking Signal

Brown's (1959) Cumulative Sum (CTS) tracking signal is given by

$$CTS_t = |SUM_t / MAD_t| \tag{17}$$

where

$$SUM_t = e_t + SUM_{t-1}. \tag{18}$$

The value of MAD_0 is set equal to its expected value as with ETS_0. The value of SUM_0 is set equal to zero. A signal occurs if the value of CTS_t exceeds a critical value K_2. Gardner (1983) suggests that the value of K_2 should be set to achieve a desired in-control ARL.

Concerning the choice of parameters for the forecast model (α_F), and tracking signals (α_1 and α_2), McKenzie (1978) and Gardner (1985) recommend that $\alpha_F \ge \alpha_1$, with $\alpha_1 = 0.1$ commonly used in practice. Small values of α_1 allow the ETS to respond more quickly to small disturbances in the demand process. Traditionally, the smoothing parameters in the numerator and denominator of the ETS have been set equal to each other, that is,

$\alpha_1 = \alpha_2$. More recently, McClain (1988) has suggested that the smoothing parameter in the MAD, α_2, be smaller than the parameter in the numerator, α_1, so that the variance of the forecast errors may be stabilized.

V. PERFORMANCE COMPARISONS: ARL vs. CDF

Concerning forecast recovery in the presence of additive outliers, consider an exponential smoothing forecast applied to an AR(1) process in which an additive outlier occurs in the process. Figure 1 shows a sequence of fifty observations from an AR(1) process (with $\phi = 0.9$) and the optimal exponentially smoothed forecasts (with $\alpha_F = 0.9444$). Figure 2 displays the resulting forecast errors.

At time period 31, a step increase in the level of the time-series occurs. The forecast lags behind the observed data at time period 32 resulting in a large forecast error. By time period 33, the forecast has adjusted to the new level of the process. The forecast errors have returned to values close to zero, as they were prior to the step increase. Notice that the 'window of opportunity' available for detection of this time-series disturbance is quite small.

The use of the cumulative distribution functions (CDF) as an evaluation criterion is not new. Barnard (1959), Bissell (1968) and Gan (1991) recommend its use for control charts on independent observations.

Referred to as a 'response to a change in demand', McClain (1988) advocates its use for forecast-based schemes which incorporate tracking signals. The CDF measures the cumulative percentage of disturbances in a time series that are detected early.

Five monitoring schemes were compared by simulation. They are the ETS, CTS, EWMA, CUSUM and the Shewhart Individuals control charts. ARLs and CDFs are provided for each monitoring scheme for outliers of size $3.0\sigma_p$, where $\sigma^2_p = \sigma^2/(1-\phi^2)$, is the variance of an AR(1) process.

The CUSUM control chart was designed to detect a shift of 1σ, with the reference value k = 0.5. The EWMA control chart was constructed with $\lambda = 0.10$, for quick detection of small disturbances as suggested by Montgomery (1991, p.306). The initial value of the EWMA is set to zero, as this is the expected value of a forecast error, if the forecast is correct. The initial values of the smoothed-error for the ETS (equation 14) and the sum of errors for the CTS (equation 17) are set to zero as suggested by Gardner (1985) and McClain (1988). The smoothing constants

$\alpha_1$and α_2 were set to 0.10 as suggested by McKenzie (1978). Further details on the simulation study are available on request from the author.

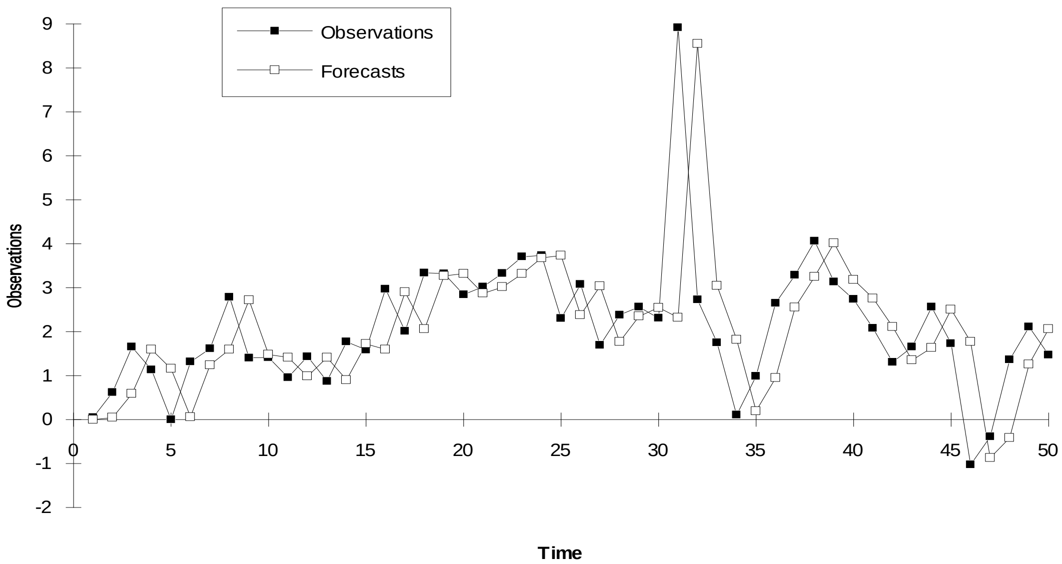

Figure 1. Observations from an AR(1) process with $\phi = 0.9$ and exponentially smoothed forecasts with λ = 0.9444. An additive outlier of size $3\sigma_p$ occurs at observation 31.

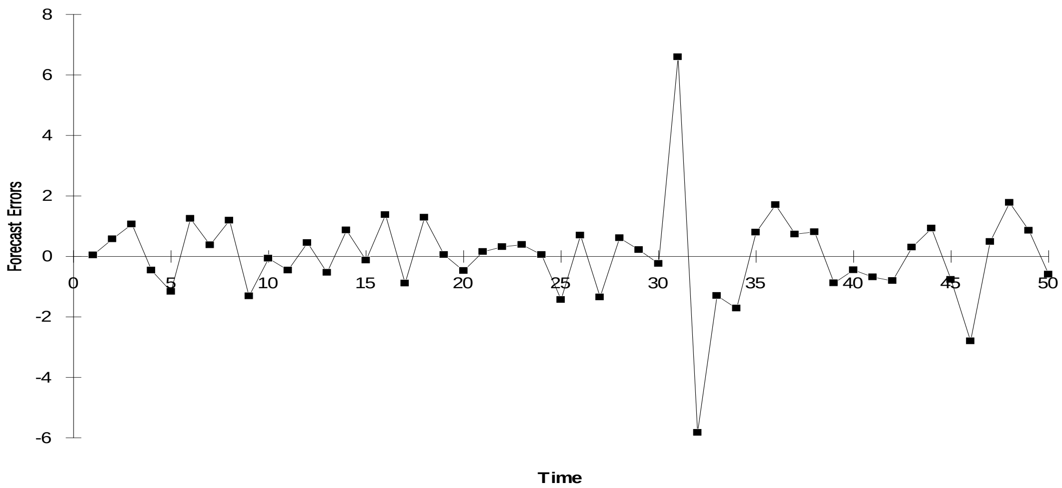

Figure 2. Forecast Errors from an AR(1) process with $\phi = 0.9$ and exponentially smoothed forecasts with λ = 0.9444. An additive outlier of size $3\sigma_p$ occurs at observation 31.

VI. RESULTS AND IMPLICATIONS

ARLs and CDFs for the Individuals, CUSUM, and EWMA control charts and the ETS and CTS tracking signals applied to the optimal exponential smoothing forecast errors from an AR(1) process with ϕ ranging from 0.0 to 0.9 are shown in Table 1. Outliers are simulated as $3\sigma_p$.

The results can be summarized as follows:

1. With the exception of the case where ϕ=0.9, the magnitude of the ARLs for the autocorrelated cases (ϕ>0) are significantly larger than for the independent case (ϕ=0). The difference in ARL magnitudes can be attributed to the quick recovery of the EWMA forecast. Recall that the ARL, as an average measure, is inflated by long run lengths. It is unable to adequately reflect short run lengths that are indicative of quick forecast recovery. For forecast-based schemes, ARLs are not informative.
2. Based on CDFs, the Individuals control chart provides a higher probability of early detection of an outlier for the autocorrelated cases where ϕ=0.5 and 0.7. This occurs although the Individuals control chart may have a longer ARL than any other monitoring scheme. As an example, consider the case where ϕ=0.5. The Individuals control chart provides a higher probability of early detection on the first observation after the outlier (60.5%) despite having a longer ARL (92.7) than the other monitoring schemes. The detection of an outlier early, that is, within the first few observations after the occurance of an outlier is critical since the forecast recovers quickly. This suggests the use of the Individuals control chart for the autocorrelated cases.

VII. CONCLUSIONS

This paper has compared forecast-based quality control schemes for monitoring autocorrelated observations in the presence of additive outliers. The quick recovery property of forecasting tools suggests that comparisons of control charts and tracking signals applied to forecast errors be based on the CDF on the run lengths and not on the ARL. The Individuals control chart is recommended over the CUSUM and EWMA control charts and the Smoothed Error and CUSUM tracking signals as it offers the highest probability of early detection of an additive outlier in an AR(1) process.

TABLE I. Average Run Lengths and Percentage of Signals detected by the ith observation after an outlier of size $3\sigma_p$. Residuals are from AR(1) processes with autoregressive parameters ϕ and an in-control ARL of 250.

ϕ	Monitoring Scheme	ARL	Number of time periods after an outlier 1	2	3	4	5	6
0	Individuals	1.8	54.2	79.0	90.5	96.1	98.2	98.9
	CUSUM	2.2	11.5	73.3	97.4	99.9	100	100
	EWMA	2.5	12.9	52.5	86.7	97.5	99.4	100
	ETS	4.3	3.9	12.7	30.5	53.9	75.7	92.2
	CTS	21.9	0.0	0.0	0.0	0.0	0.0	0.0
0.5	Individuals	92.7	60.5	64.0	64.3	64.8	64.9	65.2
	CUSUM	49.4	27.0	60.6	72.5	77.0	78.4	80.4
	EWMA	12.4	26.8	64.6	77.7	83.4	88.3	90.8
	ETS	43.1	12.6	32.8	48.3	60.0	67.5	72.6
	CTS	4.5	5.4	21.6	40.3	58.7	73.0	83.2
0.7	Individuals	42.9	84.9	85.2	85.4	85.4	85.5	85.5
	CUSUM	47.9	56.3	72.1	76.3	79.4	80.8	81.3
	EWMA	19.8	49.3	68.2	75.8	79.7	82.9	85.0
	ETS	90.0	20.2	34.1	40.5	46.6	49.4	52.4
	CTS	6.4	9.9	22.6	34.5	45.9	54.2	61.7
0.9	Individuals	1.0	100	100	100	100	100	100
	CUSUM	2.5	99.3	99.4	99.5	99.6	99.6	99.6
	EWMA	10.5	85.0	88.4	88.8	89.4	89.8	90.1
	ETS	61.9	30.2	35.6	38.0	39.5	39.8	40.3
	CTS	52.6	0.9	2.0	3.1	5.0	7.6	10.3

REFERENCES

[1]. Alwan, L, and Roberts, H. V. (1988). "Time Series Modeling for Statistical Process Control," Journal of Business and Economic Statistics, 6, 87-95.

[2]. Barnard, G. A. (1959). "Control Charts and Stochastic Processes," Journal of the Royal Statistical Society, B21, 239-271.

[3]. Bissell, A. F. (1968). "CUSUM Techniques for Quality Control," Applied Statistics, 18, 1-30.

[4]. Brown, R. G. (1959). Statistical Forecasting for Inventory Control, McGraw-Hill, New York, NY.

[5]. Change, Y.M and Wu, T.L. (2011). "On Average Run Length of Control Charts for Autocorrelated Processes," Methodology and Computing in Applied Probability, 1392), 419-431.

[6]. Cox, D. R. (1961). "Prediction by Exponentially Weighted Moving Average and Related Methods," Journal of the Royal Statistical Society, B23, 414-422.

[7]. Fox, A. J. (1972). "Outliers in Time Series," Journal of the Royal Statistical Society, Series B, 3, 350-363.

[8]. Gan, F. F. (1991). "Computing the Percentage Points of the Run Length Distribution of an EWMA Control Chart," Journal of Quality Technology, 23, 359-365.

[9]. Gardner, E. S. Jr. (1983). "Automatic Monitoring of Residuals," Journal of Forecasting, 2, 1-21.

[10]. Gardner, E. S. Jr. (1985). "CUSUM vs. Smoothed-Error Forecast Monitoring Schemes: Some Simulation Results," Journal of the Operation Research Society, 36, 43-47.

[11]. Lee, J, Alexopoulos, C, Goldsman, D, Kim, S.H, Tsui, K.L, and Wilson, J.R. (2009). "Monitoring Autocorrelated Processes Using A Distribution-Free tabular CUSUM Chart with Automated Variance Estimation," IIE Transactions, 41(11), 979-994.

[12]. Lianjie, S, Daniel, W.A and Fugee, T. (2002). "Autocorrelated Process Monitoring Using Triggered CUSCORE Charts," Quality and Reliability Engineering International, 18, 411-421.

[13]. Lu, C.W and Reynolds, M.R, Jr. (1999). "Control Charts for Monitoring the Mean and Variance of Autocorrelated Process," Journal of Quality Technology, 31(3), 259-274.

[14]. McClain, J. O. (1988). "Dominant Tracking Signals," International Journal of Forecasting, 4, 563-572.

[15]. McKenzie, E. (1978). "The Monitoring of Exponentially Weighted Forecasts," Journal of Operation Research Society, 29, 449-458.

[16]. Montgomery, D. C. (1991). Introduction to Statistical Quality Control, 2nd Edition. John Wiley and Sons, Inc., New York, NY.

[17]. Montgomery, D. C., Johnson, A. L., and Gardiner, J. A. (1990). Forecasting and Time Series Analysis, 2nd Edition. McGraw-Hill, Inc.,New York, NY.

[18]. Montgomery, D. C, and Mastrangelo, C. M. (1991). "Some Statistical Process Control Methods for Autocorrelated Data," (with discussion), Journal of Quality Technology, 23, 179-204.

[19]. Superville, C. R., and Adams, B. M. (1994). "An Evaluation of Forecast-Based Quality Control Schemes," Communications in Statistics - Simulation and Computation, 23(3), 645-661.

[20]. Superville, C. R., and Adams, B. M. (1995). "The Performance of Control Charts and Tracking Signals in Process Monitoring" International Journal of Operations and Quantitative Management, 1(2),131-143.

[21]. Trigg, D. W. (1964). "Monitoring a Forecasting System," Operational Research Quarterly, 15, 271-274.

[22]. Wardell, D.G, Moskowitz, H. and Plante, R. D. (1994). "Run Length Distributions of Special Cause Control Charts for Correlated Data," Technometrics, 36(1), 2-19.

[23]. Woodall, W. H., and Faltin, F. W. (1993). "Autocorrelated Data and SPC," American Society for Quality Control (ASQC) Statistics Division Newsletter, 13, 18-21.

[24]. Wu, B. and Yu, J-bo. (2010). "A Neural Network Ensemble Model for On-Line Monitoring of process Mean and Variance Shifts in Correlated Process," Expert Systems with Applications, 37(6), 4058-4056.

Claude R. Superville" Quality Control and Outliers in Manufacturing Processes" International Journal of Business and Management Invention (IJBMI), vol. 08, no. 04, 2019, pp 77-83