

Real Time Statistical Process Control for Autocorrelated Serial Data: A Simulation Approach

ARTUR M. F. GRAXINHA¹, J. M. DIAS PEREIRA^{1,2}

¹ Instituto Politécnico de Setúbal, ESTSetúbal, Setúbal, 2914-761, Portugal

² Instituto de Telecomunicações, 1049-001 Lisboa, Portugal

Corresponding author: Artur M.F. Graxinha (e-mail: artur.graxinha@estsetubal.ips.pt).

ABSTRACT Computer measurement systems play an important role on process automation and Industry 4.0 implementation strategies. They can be easily integrated on modern production systems, enabling real time test and control of multiple product and process characteristics that need to be monitored. If for one side the big data provided by these systems is an important asset for production analytics and optimization, on the other hand, the high frequency data sampling, commonly used in these systems, can lead to autocorrelated data violating, this way, statistical independence requirements for statistical process control implementation. In this paper we present a simulation model, using digital recursive filters, to properly handle and deal with these issues. The model demonstrates how to eliminate the autocorrelation from data time series, creating and ensuring the conditions for statistical process control application through the application of real time control charts. A performance comparison between Shewhart of Residuals and Exponentially Weighted Moving Average (EWMA) of Individual Observations control charts is made for autocorrelated data time series with the presence of different mean shift amplitude perturbations.

KEYWORDS Autocorrelated time-series; ARIMA models; Digital Filters; Real Time; Simulation; Statistical Process Control.

I. INTRODUCTION

THE level of competitiveness experienced in modern factories are pushing even further the level of acuity and need for robust process control that ensure not only the quality of the end products but, simultaneously, the efficiency of the manufacturing processes. The Industry 4.0 [1] is a reflex of this environment where the digitalization and automation of manufacturing processes are part of the evolutionary path. If for one side the availability of the big data [2, 3] enables the possibility of information generation to scrutinize processes, improve product quality and process efficiency and ultimately improve competitiveness, on the other side the way data collection is made leads to another type of problem, the serial data autocorrelation [4]. Serial data autocorrelation can be found in many continuous and discrete manufacturing processes due to the high frequency data acquisition imposed by computer measurement systems and automated technology used for process sensing and measurement [5]. In consequence of that, modern manufacturing industries have changed the effectiveness of traditional statistical process control (SPC) techniques, namely when process data violates the requirement of statistical data independence. The high frequency data acquisition not only leads to autocorrelated data, thus violating the assumptions of statistical data independence, but also

makes it difficult to provide the real time statistical information.

Several authors have been suggesting methods to deal with autocorrelation, [6–9]. All of them propose the same strategy to eliminate the serial autocorrelation from process data, which includes the fitting of an appropriate time series model to the process observations, followed by the SPC application to the sequence of obtained residuals. Despite of the effectiveness of time series modeling to eliminate autocorrelation, the lack of recursiveness of these models makes them fail when real time data control is required.

To properly handle these issues, an approach for real time process control is proposed. A simulation model was developed using the LabVIEW applying recursive digital filters, state space models and Z transform. Through the simulation model it is demonstrated how to extract serial data autocorrelation from process data using recursive digital filters, restoring, this way, the statistical data independence necessary for SPC application. The simulation model developed implements the SPC through the implementation of real time processing control charts, such as Shewhart of residuals and EWMA of individual observations. A comparison of effectiveness to detect process perturbations between both control charts is presented.

II. THEORETICAL BACKGROUND

This section presents a brief theoretical background used in the paper development and simulation model.

A. TIME SERIES MODELING

To model the serial autocorrelated process, it's proposed the use of difference equations [10] designated by autoregressive integrated moving average (ARIMA) models. An ARIMA model of (p,d,q) order for a process variable Y_t can be represented by

$$\phi(B)\nabla^d Y_t = \theta(B)a_t, \quad (1)$$

where B is the backward shift operator defined as $B^m Y_t = Y_{t-m}$. $\phi(B)$ is the autoregressive (AR) operator of order p defined by $\phi(B) = 1 - \sum_{j=1}^p \phi_j B^j$ and $\theta(B)$ is the moving average (MA) operator of order q defined by $\theta(B) = 1 - \sum_{j=1}^q \theta_j B^j$; where $\nabla = Y_t - Y_{t-1}$ defines the backward difference operator and $\nabla^d = (1 - B)^d$ and d represents the order of differentiation. The process is said stationary for $d = 0$ and nonstationary when $d \geq 1$; a_t represents a white noise variable where $a_t \sim N(0, \sigma_a^2)$. Denoting the predicted value obtained from an appropriately fitted ARIMA model by \hat{Y}_t , the residuals $e_t = \hat{Y}_t - Y_t$ will behave like independent and identically distributed random variables.

B. THE Z TRANSFORM

The Z transform is an analytical tool extremely useful in the analysis of discrete-time systems and computer modeling based on difference equations and transfer functions [11]. Taking $Y(t)$ as a quality characteristic of a certain manufacturing process, the Z transform of the discrete variable Y_t is given by

$$Z[Y_t] = Y[Z] = \sum_{t=0}^{\infty} Y_t Z^{-t}, \quad (2)$$

where Z represents the transform operator and Z is the argument defined by a complex variable. On the other hand, the inverse Z transform of $Y(Z)$ is given by

$$Z^{-1}[Y(Z)] = Y_t, \quad (3)$$

The proof of the equivalence between the Z transform and the back shift operator $B^m Y_t = Y_{t-m}$ is presented in [12] as follows

$$E(Z) = a(Z)[Y_{t-m}] = Z^{-m} \sum_{t=0}^{\infty} Y_t \cdot Z^{-t} = Z^{-m} Y(Z). \quad (4)$$

As can be seen, the delay imposed by the backshift operator B^m to the variable Y_t in the time domain, corresponds, in the Z domain, to multiplying the Z transform of the original function Y_t by Z^{-m} .

C. RECURSIVE DIGITAL FILTERS

A digital filter $H(Z)$ can be described as the implementation of an algorithm that computes a sequence of outputs $Y(t)$ from a

sequence of inputs $a(t)$. The most common classes of digital filters are the Finite Impulse Response (FIR), also called moving average (MA) or all-zero filters, and the Infinite Impulse Response (IIR). The IIR filters can be split in autoregressive (AR) or all-pole filters and autoregressive moving average (ARMA) filters, which have both poles and zeros. ARMA digital filters can be represented using the Z transform, [13, 14].

D. ARIMA DIGITAL FILTERS

ARMA digital filters can be represented by Z transform using the following transfer function [13, 14]

$$H(Z) = \frac{Y(Z)}{a(Z)} = \frac{\sum_{m=0}^q (-\theta_m) Z^{-m}}{1 - \sum_{n=1}^p (\phi_n) Z^{-n}}. \quad (5)$$

In the previous equation, ϕ_n and θ_m are, respectively, the n^{th} and m^{th} order parameters of $AR(p)$ and $MA(q)$ processes, $Y(Z)$ is the output variable and $a(Z)$ is the input white noise stream. Therefore, the ARMA $H(Z)$ digital filter will be of the form

$$H(Z) = H_1(Z) \cdot H_2(Z), \quad (6)$$

where the AR component of the digital filter $H(z)$, called IIR, is given by

$$H_1(Z) = \frac{1}{1 - \sum_{n=1}^p (\phi_n) Z^{-n}}, \quad (7)$$

and the MA component of the digital filter $H(z)$, or FIR, will be

$$H_2(Z) = \sum_{m=0}^q (-\theta_m) Z^{-m}. \quad (8)$$

E. INVERSE DIGITAL FILTERS AND TIME SERIES MODELING

Recursive digital filters can be used effectively for modeling serial autocorrelated processes. On the other hand, when appropriate inverse recursive digital filters are applied to the autocorrelated observations Y_t , the autocorrelation can be recursively removed and a sequence of normal and independently distributed errors e_t are obtained enabling, this way, the application of real time SPC.

Considering the difference equation of an ARMA process defined as

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (9)$$

the respective recursive digital filter will be defined as

$$Y(Z) = \sum_{n=1}^p \phi_n Z^{-n} Y(Z) + \sum_{m=0}^q (-\theta_m) Z^{-m} a(Z), \quad (10)$$

with $-\theta_0 = 1$.

Considering now the difference equation of one-step-ahead prediction of an ARMA process [11] such as

$$\tilde{Y}_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 a_{t-1} - \theta_q a_{t-q}, \quad (11)$$

the inverse digital filter becomes

$$\tilde{Y}(Z) = \sum_{n=1}^p \phi_n Z^{-n} Y(Z) + \sum_{m=1}^q (-\theta_m) Z^{-m} a(Z). \quad (12)$$

The prediction error $E(Z)$ is found through the difference between the observed value and the one-step-ahead forecast as follows

$$E(Z) = Y(Z) - \tilde{Y}(Z), \quad (13)$$

$$E(Z) = \sum_{n=1}^p \phi_n Z^{-n} Y(Z) + \sum_{m=0}^q (-\theta_m) Z^{-m} a(Z) - \left[\sum_{n=1}^p \phi_n Z^{-n} Y(Z) + \sum_{m=1}^q (-\theta_m) Z^{-m} a(Z) \right] = (-\theta_0) a(Z), \quad (14)$$

as $-\theta_0 = 1$ then

$$E(Z) = a(Z). \quad (15)$$

The application of the inverse Z transform to (15) leads to

$$Z^{-1}[E(Z)] = Z^{-1}[a(Z)], \quad (16)$$

that is,

$$e_t = a_t. \quad (17)$$

F. ARIMA PROCESS REPRESENTED AS STATE-SPACE-MODEL

The space-state models (SSM) are based on recursive methods [15] and might be quite useful when recursive modeling and simulation of ARIMA processes are required. Due to their recursive properties, the SSM can be used in a real time applications once the model is updated each time a new observation becomes available. The SSM have been employed by several authors for simulation purposes [16, 17].

According to [18], an ARMA (p,q) process can be represented by the following discrete state-space equations

$$\begin{cases} X(t+1) = [A]X(t) + [B]a(t) \\ Y(t) = [C]X(t) + [D]a(t) \end{cases}, \quad t = 0,1,2, \dots, \quad (18)$$

where $[A]$ is the state matrix of $r \times r$ dimension and ϕ_r are the appropriate coefficients of $\phi(B)$; $[B]$ defines an $l \times r$ vector containing the coefficients of $\theta(B)$ and $[C]$ defines a $r \times l$ vector of known constants; $X(t)$ is an $r \times l$ vector representing the state of process variables, $a_t \sim N(0, \sigma_a^2)$ is the white-noise,

and $[D]$ equals 0 or 1 depending on whether the process is or isn't subject to an additional white-noise term respectively. In our case we assume $D = 0$. Finally, $r = \max(p, q + 1)$, where p and q represent, respectively, the order of the autoregressive $\phi(B)$ and moving average $\theta(B)$ polynomials.

SSM can also be used to represent ARIMA (p,d,q) processes [17]. For that the $\phi_1, \phi_2, \dots, \phi_p$ parameters must be replaced by the outcome parameters of the representation $(1 - B)^d \theta(B)$. The expanded transition and observation equations are defined respectively by (19) and (20).

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ x_3(t+1) \\ \vdots \\ x_{r-1}(t+1) \\ x_r(t+1) \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 & 0 & 0 \dots 0 \\ \phi_2 & 0 & 1 & 0 \dots 0 \\ \phi_3 & 0 & 0 & 1 \dots 0 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ \phi_{r-1} & 0 & 0 & 0 \dots 1 \\ \phi_r & 0 & 0 & 0 \dots 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_{r-1}(t) \\ x_r(t) \end{bmatrix} + \begin{bmatrix} 1 \\ -\theta_1 \\ -\theta_2 \\ \vdots \\ -\theta_{r-1} \\ -\theta_r \end{bmatrix} a(t) \quad (19)$$

$$Y(t) = [1 \ 0 \ 0 \ \dots \ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_{r-1}(t) \\ x_r(t) \end{bmatrix} + [D]a(t). \quad (20)$$

G. SPC STRATEGIES

Traditional SPC [19] assumes that the observations a certain process characteristic Y_t , obtained by the sample of order t are defined by

$$Y_t = \mu + a_t, \quad (21)$$

where μ defines the process mean and a_t represents the white noise error, where $a_t \sim N(0, \sigma_a^2)$ and $Cov(a_t, a_{t+j}) = 0$ for any $j \neq 0$, meaning that the sequence of errors a_t is statistically independent. This ensures that Y_t follow as iid process (independent and identically distributed), with $Y_t \sim N(\mu, \sigma_Y^2)$ and $Cov(Y_t, Y_{t+j}) = 0$ for any for any $j \neq 0$. This is a fundamental requirement for SPC application.

There is a close connection between control charts and hypothesis testing [19], null hypothesis against alternative hypothesis. Let's assume that the t^{th} sample of the process characteristic has an average of \bar{Y}_t . If the value \bar{Y}_t is bounded by the control limits $(\tilde{\mu} \pm 3\tilde{\sigma}_{\bar{Y}}^2)$ we accept null hypothesis ($H_0: \mu = \mu_0$). Under these conditions the process mean \bar{Y}_t is under control having only normal causes of variation. If the value \bar{Y}_t fall's outside of the control limits $(\tilde{\mu} \pm 3\tilde{\sigma}_{\bar{Y}}^2)$ we reject null hypothesis ($H_0: \mu = \mu_0$) and accept alternative hypothesis ($H_1: \mu \neq \mu_0$). These conditions highlight the present of special causes of variation on the process mean \bar{Y}_t .

The presence of serial autocorrelation has a negative effect on the SPC application effectiveness as decreases the ARL

(Average Run length) [6, 7]. This increases the level occurrences of Type I statistical error.

Some authors, [6–9] demonstrated the effect of serial autocorrelation and presented appropriated strategies to eliminate undesired autocorrelation effects on control charts.

H. RESIDUALS CONTROL CHARTS

One of the strategies followed was to model the autocorrelative structure applying the appropriated ARIMA model to the autocorrelated time series. When an appropriate ARIMA model is fitted to an autocorrelated data time series, a sequence of independent and identically distributed residuals $e_t = \tilde{Y}_t - Y_t$ is obtained. Once restored the iid conditions, the traditional Shewhart control charts can be applied to the residuals stream of data time series. These are called the residual control charts [19]. Therefore, whenever a shift in the mean occurs, highlighting the presence of a special cause of variation, it will be transferred to the residuals and detected by the control charts [6].

I. EWMA CONTROL CHARTS

The suggestion to use control charts based on Exponentially Weighted Moving Averages (EWMA) was made for the first time in [21]. Since then, several authors have been published their developments on this subject, [20–26].

The EWMA statistic [19] for a certain process characteristic Y_t , obtained at order t , is defined by

$$Z_t = \lambda Y_t + (1 - \lambda)Z_{t-1}, \quad (22)$$

where $t = 1, 2, \dots$, and $0 < \lambda < 1$ is the weighted constant working as a filter to evidence recent or past observation. If the Y_t data time series are uncorrelated, the control limits for the EWMA control chart for individual observations can be defined by the Lower and Upper Control Limit respectively, LCL and UCL, [8, 19].

$$LCL = \bar{Y} - k\sigma_y \sqrt{\frac{\lambda}{2-\lambda} [1 - (1 - \lambda)^{2t}]}, \quad (23)$$

$$UCL = \bar{Y} + k\sigma_y \sqrt{\frac{\lambda}{2-\lambda} [1 - (1 - \lambda)^{2t}]}, \quad (24)$$

with $k = 3$.

The EWMA can be used in situations where the data time series is autocorrelated, in specifically when the data follows a ARIMA (0,1,1) = IMA (1,1) process [8].

$$Y_t = Y_{t-1} + a_t - \theta_1 a_t. \quad (25)$$

According to these authors the EWMA with $\lambda = 1 - \theta_1$ is the optimal one-step-ahead forecast for these processes, where $\tilde{Y}_{t+1} = Z_t$ is the forecasted observation of instant $t + 1$ made at instant t . Therefore, we also have $\tilde{Y}_t = Z_{t-1}$. Using the above identities and replacing in (22) we have

$$\tilde{Y}_{t+1} = (1 - \theta_1)Y_t + \theta_1 \tilde{Y}_t = Y_t - \theta_1(Y_t - \tilde{Y}_t), \quad (26)$$

Also considering that $a_t = Y_t - \tilde{Y}_t$, it becomes

$$\tilde{Y}_{t+1} = Y_t - \theta_1 a_t. \quad (27)$$

If one, consider the instant t then we can re-write (27) as

$$\tilde{Y}_t = Y_{t-1} - \theta_1 a_{t-1}. \quad (28)$$

Once $\tilde{Y}_t = Y_t - a_t$ we can finally have an equivalent equation to IMA (1,1) process in (25)

$$Y_t = Y_{t-1} + a_t - \theta_1 a_{t-1}. \quad (29)$$

Under these circumstances, the one-step-ahead prediction errors $a_t = Y_t - \tilde{Y}_t$, are independent and identically distributed with zero mean and standard deviation σ_p . Control charts with control limits at $\pm 3\sigma_p$ can be applied to these one-step-ahead prediction errors.

According to [8], the EWMA can produce excellent one-step-ahead prediction for other ARIMA processes, if the observations are positively correlated, and the process mean does not drift too quickly.

Autocorrelated data requires a different approach for selecting λ to achieve a certain average run length, when compared with uncorrelated data. The authors [19, 27] proposed a method to select λ based on the minimization of the sum of squares of the one-step-ahead prediction errors (SSE). According to these authors, the estimate of the variance of one-step-ahead errors σ_p^2 can be calculated by the quotient between the sum of squares of prediction errors for the optimal λ and the n observations used for its determination. This method can be easily implemented in computer time series analysis.

$$\sigma_p^2 \cong \frac{SSE}{n}. \quad (30)$$

III. SIMULATION MODEL

The Real Time simulation model presented in this work (Figure 1) was developed in *LabVIEW*, a system engineering graphical programming language widely used in virtual instrumentation and computer measurement applications. This simulation model can simulate any ARIMA process using the SSM approach from a white noise $N(0,1)$ base line. This way once created the desired autocorrelation data structure, it is also possible to eliminate the autocorrelation using digital recursive filters (FIR and IIR), reconverting the data into i.i.d. residuals. The model explores the application of individual observations control charts to the residuals and proper set of EWMA charts to autocorrelated data time series. The mean shifts, simulated through a step degree function, enable the evaluation of the two types of charts for different process perturbation amplitude shifts imposed at instant $t = 2s$. The sum of squares errors (SSE) algorithm was also implemented to support the selection of the optimum λ to be used in the EWMA statistics.

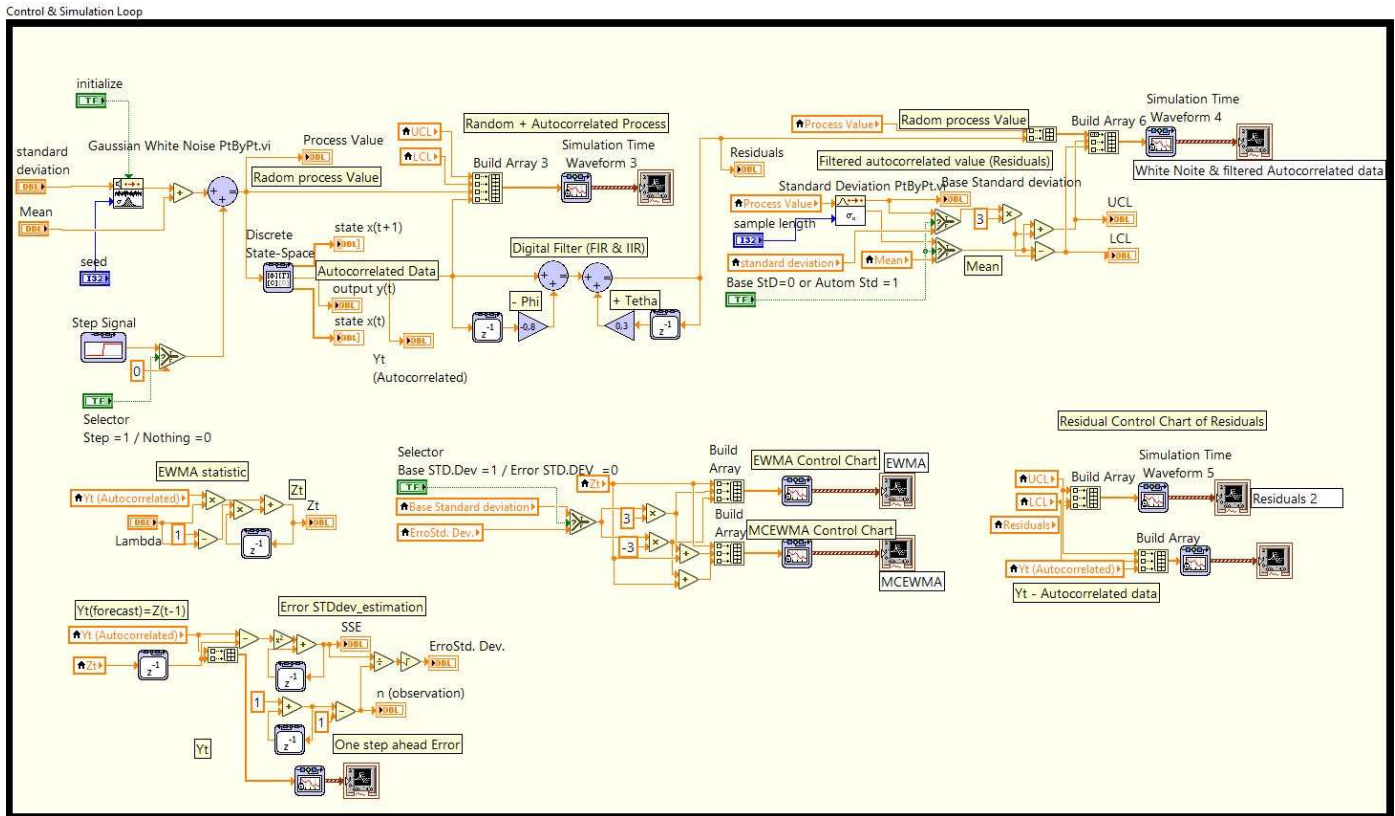


Figure 1. LabVIEW Simulation model for real time SPC.

A. SIMULATION CASES

This work explores the simulation of two practical examples of ARIMA processes, an AR (1) and an ARMA (1,1).

CASE 1 – AR (1), first order autoregressive model with $\phi_1 = 0.8$ is defined by

$$\tilde{Y}_t = a_t + 0.8Y_{t-1}, \quad (31)$$

and it has the following SSM representation

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 0.8 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} a_t, \quad (32)$$

$$y(t) = [1 \ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \quad (33)$$

CASE 2 – ARMA (1,1), autoregressive moving average model with parameters $\phi_1 = 0.8$ and $\theta_1 = 0.3$ can be written by

$$\tilde{Y}_t = 0.8Y_{t-1} + a_t - 0.3a_t, \quad (34)$$

with the following SSM representation,

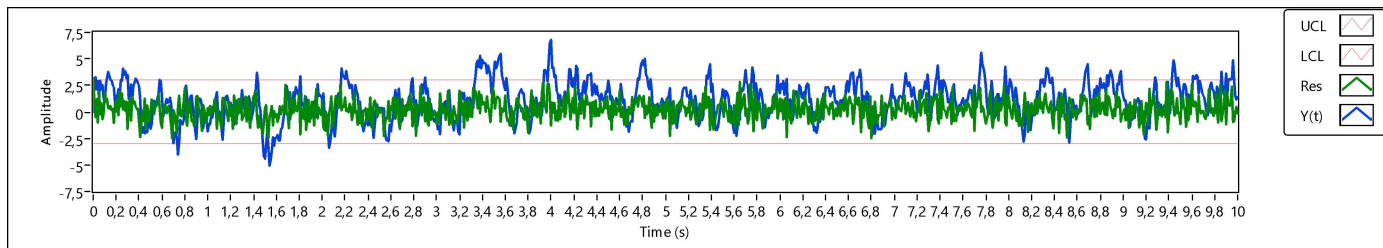
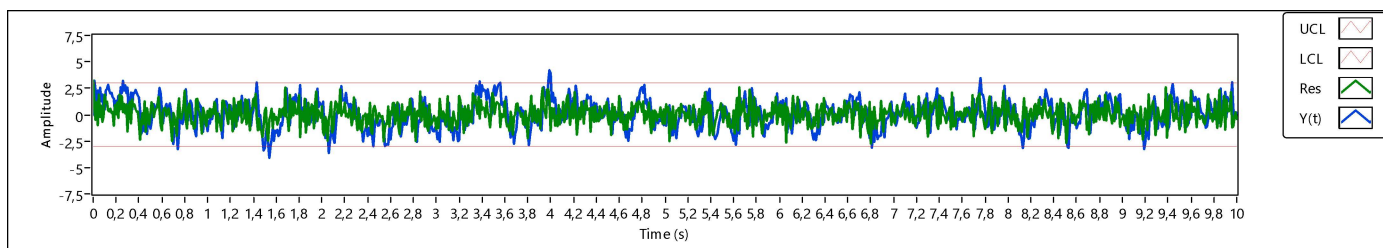
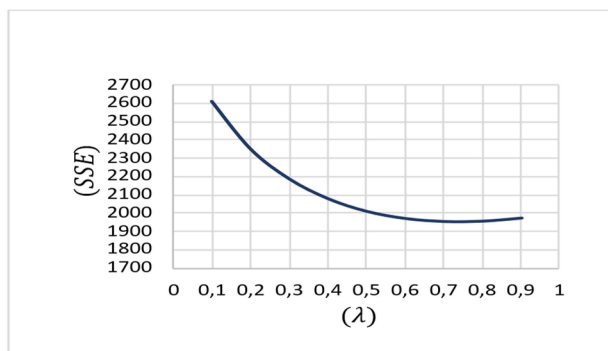
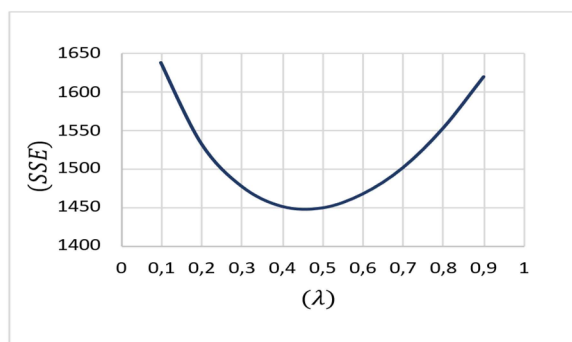
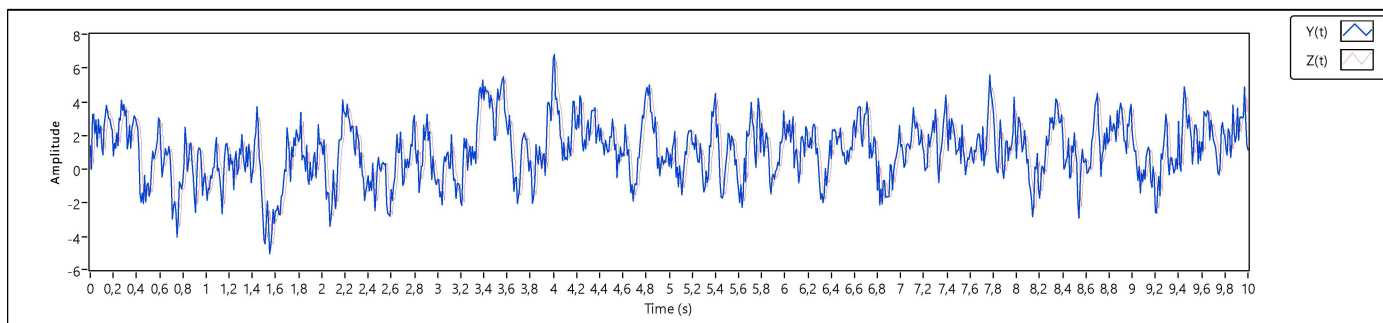
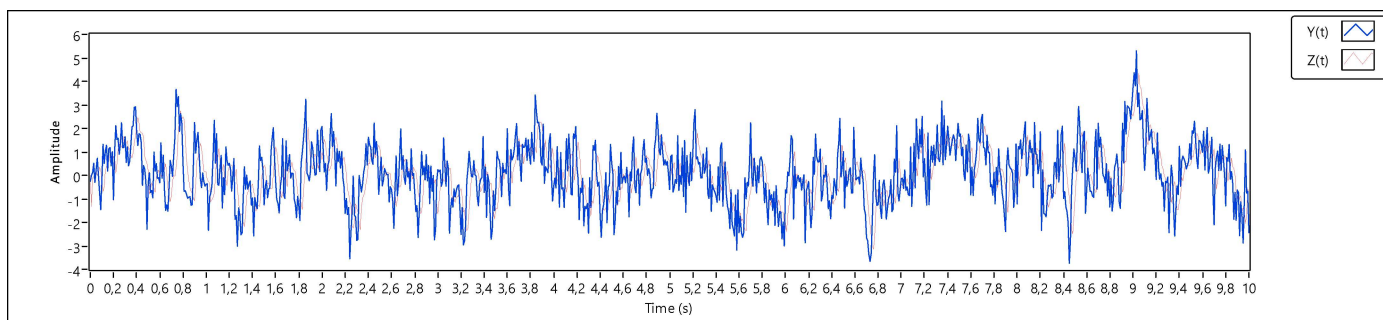
$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 0.8 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ -0.3 \end{bmatrix} a_t, \quad (35)$$

$$y(t) = [1 \ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \quad (36)$$

In both cases the LabVIEW simulation model was properly configured with parameters of (32), (33), (35) and (36). The FIR and IIR digital filters were setup with the appropriated values to enable the autocorrelation elimination. Figure 2 and Figure 3 represents, respectively, the control chart for individual observations for the autocorrelated data time series $Y(t)$ (blue line) and for the residuals $e(t)$ (green line) for case 1 and case 2. As we can observe by the number of points from the autocorrelated time series $Y(t)$ falling outside of control limits (UCL , LCL), the Type I statistical errors are significantly high on this case in comparison with the residuals $e(t)$.

The parameter λ of the EWMA control charts was selected, in both cases, using the minimum sum of square errors (SSE) of $a(t)$, obtained from the one-step-ahead prediction errors $a_t = Y_t - \tilde{Y}_t$, according to (28) and (29). Figure 4 and Figure 5 illustrate the SSE function of the λ for AR (1) and ARMA (1,1) cases, respectively. As it can be observed the values of λ that minimizes the SSE on the AR (1) data time series is $\lambda = 0.7$ while on ARMA (1,1) process is $\lambda = 0.45$.

Figure 6 and Figure 7 shows, respectively, the one-step-ahead EWMA forecast for the AR (1) and ARMA (1,1) autocorrelated data time series $Y(t)$ using the λ values obtained in each case.


 Figure 2. Individual control chart for autocorrelated data time series $Y(t)$ and residuals $e(t)$ of AR(1) model.

 Figure 3. Individual control chart for autocorrelated data time series $Y(t)$ and residuals $e(t)$ of ARMA (1,1) model.

 Figure 4. Sum of Square Errors of one-step-ahead prediction errors $a_t = Y_t - \tilde{Y}_t$ from AR (1) model.

 Figure 5. Sum of Square Errors of one-step-ahead prediction errors $a_t = Y_t - \tilde{Y}_t$ from ARMA (1,1) model.

 Figure 6. Autocorrelated data time series $Y(t)$ and one-step-ahead prediction EWMA $Z(t)$ for AR (1) process.

 Figure 7. Autocorrelated data time series $Y(t)$ and one-step-ahead prediction EWMA $Z(t)$ for ARMA (1,1) process.

The standard deviation of one-step-ahead error σ_p , required to setup the control limits of the EWMA control chart was computed in real time, applying the square root to (30). With autocorrelation eliminated from data time series and with the appropriated setting of the EWMA it was possible to run several simulations to establish a comparison between the effectiveness in detecting mean shifts of the residuals of individual observations control chart and EWMA control chart. The comparison between these two types of control charts was made for the following mean shifts, 0.25, 0.5, 0.75, 1.0, 1.25 and 1.5 units of standard deviation. The mean shifts were imposed to the model at instant $t = 2s$ using the step function. For easy comparison the data set of random generated data series was common for all the simulations executed. For simulation propose the time duration of each simulation is 10 seconds at a rate of 100 samples per second.

IV. DISCUSSION RESULTS

The data in Table 1 summarizes the results obtained during the simulations with the use of EWMA and Individual Observation of Residuals Control Charts for both processes, AR (1) and ARMA (1,1), for all the mean shifts. As expected, the EWMA Control Chart is more effective detecting small mean shifts in comparison with the Individual Observation of Residuals Control Chart for any of the processes AR (1) and ARMA (1,1) considered in the experience. Not only the detection is made earlier with the EWMA Control Chart, but it also provides,

consistently, more points out of control limits after the mean shift perturbation. For a small mean shift of 0.25 the Individual Observation of Residuals Control Chart is not able to detect the perturbation, while the EWMA Control Chart detects the first point out of control limits at instant $t = 3.38s$ and $t = 4.01s$ respectively, for AR (1) and ARMA (1,1) processes. The effectiveness of the Residual Control Chart detecting the first point out of control improves with the increment of the mean shifts. Even though for mean shifts equal or bigger than 0.75 the Residuals and EWMA Control Charts shown a similar capability in detecting the first point out of control, EWMA shown more consistence in detecting points out of control after the mean shift.

While for the selected AR (1) and ARMA (1,1) processes the Residual Control Chart doesn't show differences in its performance as the residual series $e(t)$ will be equal in both processes after autocorrelation filtering, the EWMA seems to react a bit slowly to the mean shifts on ARMA (1,1) in comparison with AR (1) process. Regarding the number of points out of control, the EWMA Control Chart show a good consistence on AR (1) and ARMA (1,1) processes, although the number of out-of-control points is bigger in the AR (1). Also, we can observe that EWMA tend to react more quickly in detecting points out of control as mean shifts perturbation amplitude increases.

TABLE 1. Summary of Simulation Results

Deviation	AR(1)				ARMA(1,1)			
	OCL_Detection (s)		# Points OC		OCL_Detection (s)		# Points OC	
	Residuals	EWMA	Residuals	EWMA	Residuals	EWMA	Residuals	EWMA
0,25	ND	3,38 s	0	25	ND	4,01 s	0	2
0,5	3,95 s	2,19 s	3	106	3,95 s	3,38 s	3	33
0,75	2,18 s	2,19 s	10	300	2,18 s	2,19 s	10	134
1	2,18 s	2,16 s	18	550	2,18 s	2,19 s	18	356
1,25	2,18 s	2,16 s	31	693	2,18 s	2,17 s	31	585
1,5	2,18 s	2,15 s	52	774	2,18 s	2,16 s	52	721

OCL_Detection (s) - Out of control Limits Detection in seconds

ND - No Detection

Points OC - Number of points out of control in time series

Figure 8 and Figure 9 show simulated EWMA and Residual Control Charts for the AR (1) process with mean shift of 0.25, respectively. As referred, the Residual Control Chart is not able to perform any detection of points out of control after the perturbation at $t = 2s$, while EWMA Control Chart get the first point out of control at $t = 3.38s$.

The EWMA and Residual Control Charts for the AR (1) process with mean shift of 1 are shown on Figure 10 and Figure 11, respectively. As it can be seen and as per table 1, the detection of the mean shift in these two cases is made

immediately after the mean shift perturbation. Note as EWMA reaction time to mean shift perturbation increased with shift amplitude (Figure 8, Figure 10).

In Figure 12 and Figure 13 is shown the EWMA and Residual Control Charts for the ARMA (1,1) process with mean shift of 1.25, respectively. As observed, even though the ability of the Residual Control Chart to detect the mean shift at the same approximate time instant of the EWMA Control Chart, this one has consistently more points out of control outside the control limits after the perturbation.

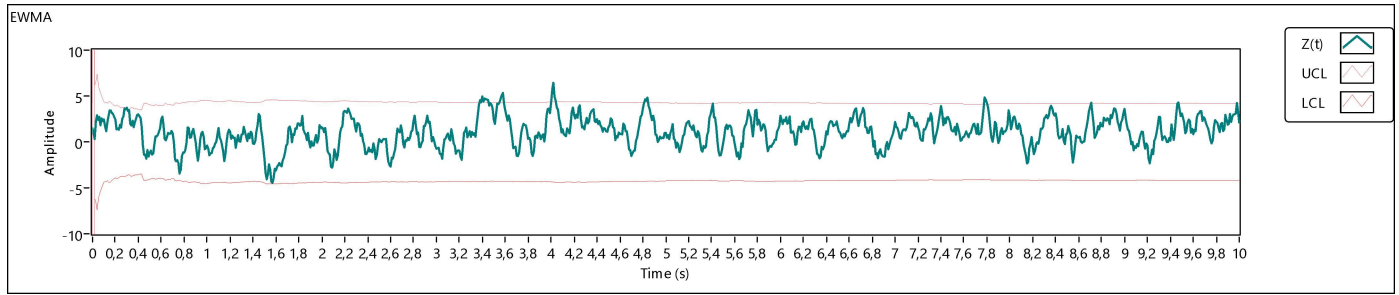


Figure 8. EWMA Control Chart for AR (1) process with mean shift amplitude of 0.25.

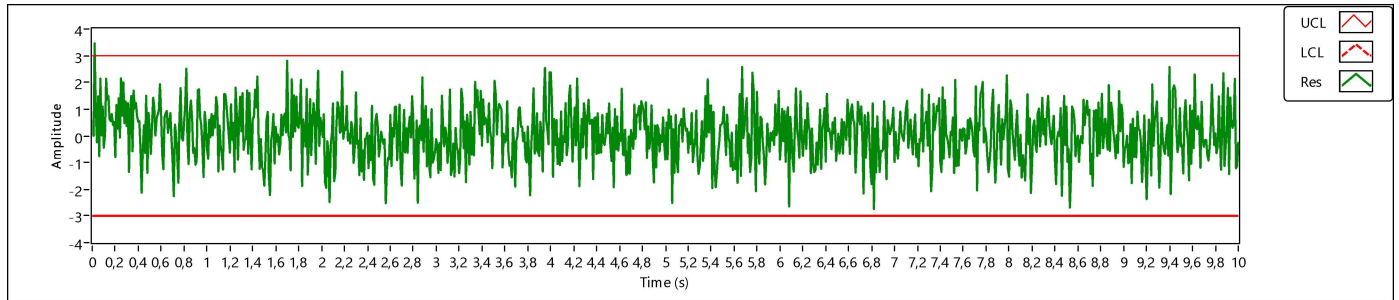


Figure 9. Residual of Individual Observations Control Chart for AR (1) process with mean shift amplitude of 0.25.

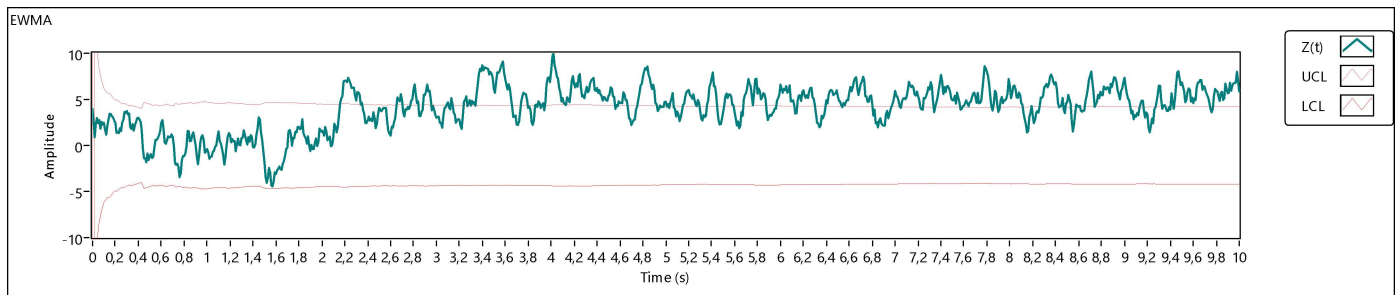


Figure 10. EWMA Control Chart for AR (1) process with mean shift amplitude of 1.

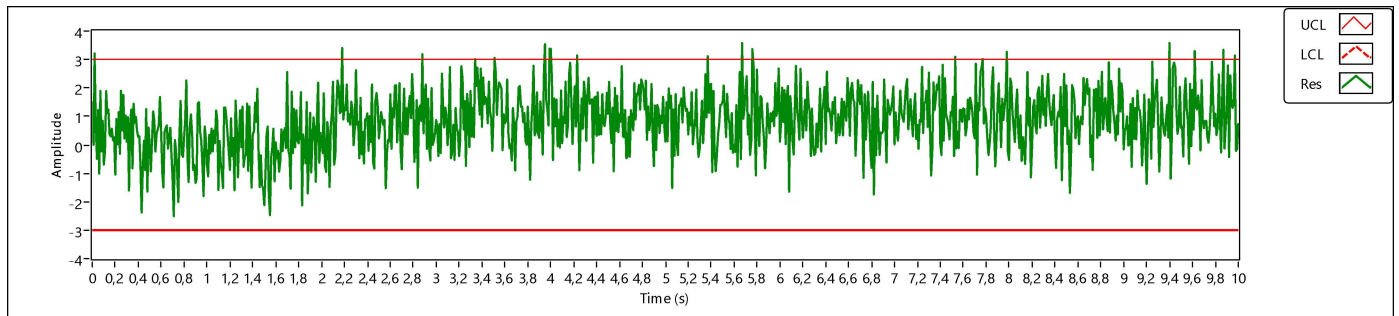


Figure 11. Residual of Individual Observations Control Chart for AR (1) process with mean shift amplitude of 1.

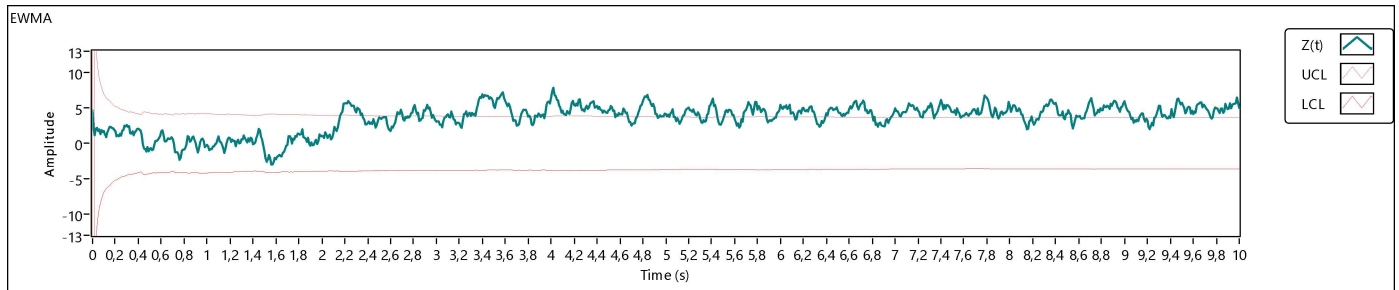


Figure 12. EWMA Control Chart for ARMA (1,1) process with mean shift amplitude of 1.25.

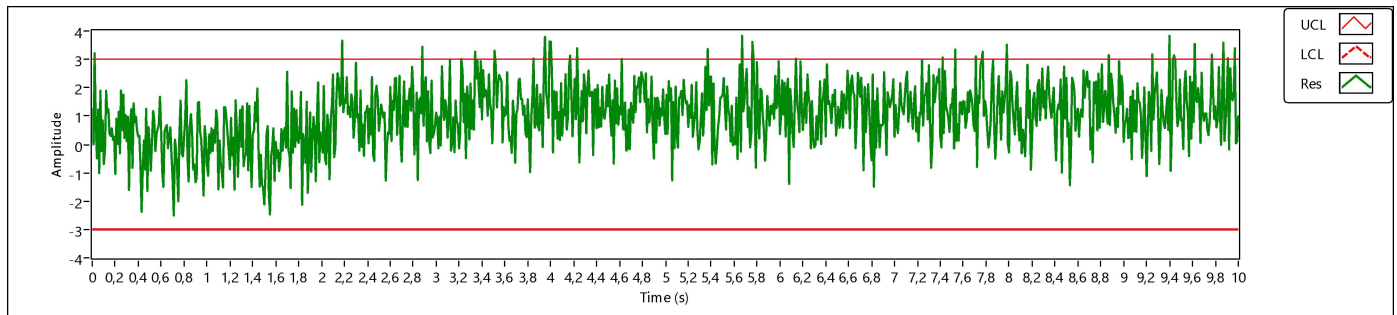


Figure 13. Residual of Individual Observations Control Chart for ARMA (1,1) process with mean shift amplitude of 1.25.

V. CONCLUSIONS

A simulation model was developed to illustrate an approach for dealing with real time statistical process control with the presence of time series data autocorrelation very often found in technological advanced industrial environments. The model enables the generation of any ARIMA process. In this paper two autocorrelated processes were evaluated, a first order autoregressive AR (1) and a first order autoregressive moving average ARMA (1,1). The results show that appropriate digital recursive filters could be effectively used to remove autocorrelation in a real time environment enabling, that way, further application of statistical process control without violating the principle of statistical independence. Though the application of different mean shift amplitude perturbations, imposed at a preset time to each one of the autocorrelated processes, we evaluated the effectiveness of the application of both, Residuals of Individual Observations and EWMA control charts. This evaluation of performance consisted in the comparison of the time response achieved with the two control charts after the mean shift perturbation has been imposed to the time series data and by the consistence of points out of control after the occurrence of this perturbation. The results obtained are in line with theoretical expectations. For small amplitude mean shifts EWMA control chart shown a quicker time reaction to perturbation, exposing consistently points out of control limits after that point. Residuals Individual Observation control charts tend to improve reaction time to mean shift perturbations as its amplitude increases. They show similar time responses to EWMA for bigger amplitude mean shifts, in spite its consistence exposing points out of control is much smaller in comparison with EWMA control charts.

References

- [1] R.S. Nakayama, M. M Spinola and J. R. Silva, "Towards I4.0: A comprehensive analysis of evolution from I3.0," *Computers and Industrial Engineering*, vol. 144, pp. 106453, 2020, <https://doi.org/10.1016/j.cie.2020.106453>.
- [2] K. W. Chun, H. Kim and K. Lee, "A study on research trends of technologies for Industry 4.0; 3D printing, artificial intelligence, big data, cloud computing, and Internet of Things," *Advanced Multimedia and Ubiquitous Engineering, Lecture Notes in Electrical Engineering*, vol. 518, pp. 397-403, 2019, https://doi.org/10.1007/978-981-13-1328-8_51.
- [3] M. S. Reis, G. Gins, "Industrial process monitoring in the big data/industry 4.0 era: From detection, to diagnosis, to prognosis," *processes*, vol. 35, no. 5, pp. 1-16, 2017, <https://doi.org/10.3390/pr5030035>.
- [4] M. Kulahci, F.D. Frumosu, A.R. Khan, G. Ø. Rønsch, M.P. Spooner, "Experiences with big data: Accounts from a data scientist's perspective," *Quality Engineering*, vol. 32, no. 4, pp. 529-542, 2020, <https://doi.org/10.1080/08982112.2019.1686641>.
- [5] E. Vanhatalo, M. Kulahci, "Impact of autocorrelation on principal components and their use in statistical process control," *Quality and Reliability Engineering International*, vol. 32, pp. 1483-1500, 2015, <https://doi.org/10.1002/qre.1858>.
- [6] L. C. Alwan, and H. V. Roberts, "Time-series modelling for statistical process control," *Journal of Business & Economic Statistics*, vol. 6, no. 1, pp. 87-95, 1988, <https://doi.org/10.1080/07350015.1988.10509640>.
- [7] H. V. Roberts, and R. S. Tsay, "Making control charts more effective by time series analysis: Two illustrative applications," *Communications in Statistics – Theory and Methods*, vol. 26 no. 11, pp. 2767-2796, 2007, <https://doi.org/10.1080/03610929608831868>.
- [8] D. C. Montgomery, and C. M. Mastrangelo, "Some statistical control methods for autocorrelated data," *Journal of Quality Technology*, vol. 23, no. 3, pp. 179-193, 1991, <https://doi.org/10.1080/00224065.1991.11979321>.
- [9] J. E. Wieringa, *Statistical Process Control for Serially Correlated Data*, Ph.D. Dissertation, University of Groningen, Netherlands, 1999.
- [10] G. E. P. Box and G. M. Jenkins, *Linear Stationary Models, Linear Nonstationary Models, in Time Series Analysis, Forecasting and Control*, 5th edition, Hoboken, New Jersey, US: John Wiley & Sons, 2016, pp.47-125.
- [11] S. Jack Hu and C. Roan, "Change patterns of time series-based control charts," *Journal of Quality Technology*, vol. 28, no. 3, pp. 302-312, 1996, <https://doi.org/10.1080/00224065.1996.11979680>.
- [12] E. Del Castillo, *Z-transforms in Statistical Process Adjustment for Quality Control*, New York, US, John Wiley & Sons, 2002, pp. 60-69.
- [13] J. G. Webster, Digital Filters, *Wiley Encyclopaedia of Electrical and Electronics Engineering II*, pp. 1-18, 2007, <https://doi.org/10.1002/047134608X.W2233.pub2>.
- [14] D. G. Childers, *Ergodicity, Statistics, Estimation and Simulation in Probability and Random Process*, Boston, US, McGraw-Hill, 1997, pp. 304-309.
- [15] R.E. Kalman, and R.S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95-108, 1961, <https://doi.org/10.1115/1.3658902>.
- [16] H. B. Nembhard, C. M. Mastrangelo, "Integrated process control for startup operations," *Journal of Quality Technology*, vol. 30, no. 3, pp. 201-211, 1998, <https://doi.org/10.1080/00224065.1998.11979849>.
- [17] H. B. Nembhard, C. M. Mastrangelo, and M. S. Kao, "Statistical monitoring performance for startup operations in a feedback control system," *Quality and Reliability Engineering International*, vol. 17, no. 5, pp. 379-390, 2001, <https://doi.org/10.1002/qre.421>.
- [18] B. J. F. Murteira, D. A. Muller and K. F. Turkman, "Modelos em Espaço de Estados" in *Análise de Sucessões Cronológicas*, Lisboa, McGraw-Hill, 1993, pp. 275-281.
- [19] D. C. Montgomery, "Methods and Philosophy of Statistical Process Control", "Cumulative Sum and Exponentially Weighted Moving Average Control Charts" in *Introduction to Statistical Quality Control*, 7th edition. New York, US, John Wiley & Sons, 2013, pp. 187-221, pp. 413-442.
- [20] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 42, no. 1, pp. 97-101, 2000, <https://doi.org/10.1080/00401706.2000.10485986>.
- [21] P. B. Robison and T. Y. Ho, "Average run lengths of geometric moving average charts by numerical methods," *Technometrics*, vol. 20, no.1, pp: 85-93, 1978, <https://doi.org/10.1080/00401706.1978.10489620>.
- [22] J. M. Lucas and M. S. Saccucci, "Exponentially weighted moving average control schemes: Properties and enhancements," *Technometrics*, vol. 32, no.1, pp. 1-12, 1990, <https://doi.org/10.1080/00401706.1990.10484583>.
- [23] S. V. Crowder, "A simple method for studying run length distributions of exponentially weighted moving average charts," *Technometrics*, vol. 29, no. 4, pp. 401-407, 1987, <https://doi.org/10.2307/1269450>.
- [24] S. V. Crowder, "Design of exponentially weighted moving average schemes," *Journal of Quality Technology*, vol. 21, no.3, pp. 155-162, 1989, <https://doi.org/10.1080/00224065.1989.11979164>.

- [25] M. C. Testik, O. Kara, and S. Knoth, "An algorithmic approach to outlier detection and parameter estimation in phase I for designing phase II EWMA control chart," *Computers & Industrial Engineering*, vol. 144, 2020, <https://doi.org/10.1016/j.cie.2020.106440>.
- [26] H. H. Ko, J. Kim, S. H. Park, J. G. Baek, S. S. Kim, "Advanced semiconductor fabrication process control using dual filter exponentially weighted moving average," *Journal of Intelligent Manufacturing*, vol. 23, pp. 443-455, 2010, <https://doi.org/10.1007/s10845-010-0383-6>.
- [27] T. Pyzdek, "EWMA Charts" in *The Six Sigma Handbook*, New York, US, McGraw-Hill, 2001, pp. 438-451.



ARTUR M. F. GRAXINHA is an Adjunct Professor at Escola Superior de Tecnologia do Instituto Politécnico de Setúbal at Computer and Systems Department, scientific área of Process Control. He's technical fellow on Electronics and Automation Engineering by Instituto Politécnico de Setúbal, Instituto Politécnico de Lisboa and Instituto Politécnico de Leiria, since 2020. He obtained his Master's degree in Industrial Engineering from Faculdade de Ciências e Tecnologia of Universidade Nova de Lisboa (FCT/UNL) in 2003 and he is post graduated in Lean Six

Sigma by Instituto Superior Técnico (IST). He has a vast and consolidated experience in the automotive electronics industry and he has a track record of goals and challenges achieved throughout his professional life. His main research interests include dynamical systems, real time processing and control.



J. M. DIAS PEREIRA was born in Portugal in 1959. He received his degrees in Electrical Engineering from the Instituto Superior Técnico (IST) of the Technical University of Lisbon (UTL) in 1982. He starts his professional activities as a teacher in IST and during almost eight years he worked for Portugal Telecom in digital switching and transmission systems. In 1992, he returned to teaching activities as an Assistant Professor in Escola Superior de Tecnologia of Instituto Politécnico de Setúbal, where he is, at present, a Principal Coordinator Professor. His main research interests are included in the Instrumentation and Measurements fields.

...