Exponentially Weighted Moving Average (EWMA) Prediction in the Software Development Process

Thomas M. Fehlmann
Euro Project Office AG
Zurich, Switzerland
eMail: thomas.fehlmann@e-p-o.com

Eberhard Kranich
Euro Project Office
Duisburg, Germany
eMail: eberhard.kranich@e-p-o.com

***Abstract*—For some years, Statistical Process Controls (SPC) techniques such as traditional Shewhart control charts add value to monitor and to control the Software Development Process (SDP) efficiently. Nonetheless, the application of Shewhart control charts to the SDP involves a considerable problem, since the availability of a sufficiently large set of observations is essential when constructing traditional control charts. Especially at the start-up of each SDP phase such a set cannot be provided. To remedy this problem, Q control charts widely used when monitoring short-run manufacturing processes have been introduced successfully. This paper focuses on the predictive property of Exponentially Weighted Moving Average (EWMA) Q control charts and investigates whether the predictive property is attractive for monitoring and controlling the SDP. Results of initial experiments are also reported.**

Keywords Q-Statistic, Q control chart, Exponentially Weighted Moving Average (EWMA), software development process (SDP), statistical process control (SPC), forecasting

I. Introduction

Shewhart control charts, introduced by Shewhart [34] in the 1920's and further investigated by W. E. Deming [6] (see also Thompson and Koronacki [36]), are a highly valuable and well accepted Statistical Process Control (SPC) tool for the monitoring, controlling, and systematic improvement of mass production processes manufacturing identical parts, see also standard SPC textbooks such as Montgomery [25], Qiu [26], and Ryan [33] for further details. Shewhart control charts are mainly applied to investigate and understand the variation of a considered process. In view of the fundamental equation of SPC, the variation of the considered process is equal to the natural cause process variation plus the assignable cause process variation. Natural or common cause variation is unpreventable inherent to any process and can only be reduced by modifying the process itself, whereas assignable or special cause variation originates from external process events which make the process in-stable or out-of-control and, when detected, should be eliminated promptly to stabilize the process.

The Software Development Process (SPD) is evidently not a mass production process, since an actually developed software artifact is in general not identical to other previously developed software artifacts. Hence, a sufficiently large set of observations is not available which is required for the construction of Shewhart control charts. However, Florac and Carleton [10] provide guidelines how to monitor and control the SDP by means of Shewhart control charts.

But how to monitor and control the SDP when only a small set of individual observations such as defects at the start-up stage of testing is available? Chang and Tong [2] were the first who apply short-run statistical process control techniques along with associated short-run control charts, also termed self-starting control charts, to the monitoring of the SDP, see also Fehlmann and Kranich [8]. In particular, Chang and Tong [2] recommend to utilize Q control charts introduced by Quesenberry [27], [30]. These control charts are based on standardized, normally distributed Q-Statistics. In contrast to classic Shewhart control charts, Q control charts feature the notable property that they have constant control limits. This property enables, for instance, a visualization of the monitored indivudual observations of different process characteristics such as Chillarege's [3] Orthogonal Defect Classification (ODC) categorized defects in exactly one control chart.

In order to identify an out-of-control situation in the course of a process run, a set of decision rules is generally applied, see e.g. Hoyer and Ellis [17]. The main purpose of such rules is to signal a potential out-of-control situation before the actual in-control process reveals in fact an out-of-control event. Complementary to some of the decision rules or as an alternative to Shewhart control charts, Exponential Weighted Moving Average (EWMA) control charts which have individual observations as input can be utilized. Contrary to Shewhart control charts which can only detect large shifts in the observations, applying EWMA control charts is recommended when small shifts in the observations are to be identified, see e.g. Hunter [18], Montgomery [25], Qiu [26], or Ryan [33].

Quesenberry [28], [30] points out that the values of the various types of Q-Statistics can also serve as input to EWMA control charts resulting in EWMA Q control charts. This type of control charts is described in deatil and investigated by Fehlmann and Kranich [**?**] in view of monitoring and controlling the software development process. The focus of this paper is to analyze whether the predictive property of EWMA Q control charts is an attractive control mechanism in the context of the software development process.

The paper is organized as follows. Basic principles of Q-Statistics and Q control charts are described in Section II. The background of EWMA control charts along with an enhancement is given in Section III. EWMA Q control charts are introduced in Section IV. Section V shows how EWMA Q control charts can be utilized to control the software development process by means of forecasting. The results of this section are applied to an example in Section VI in order to illustrate the attractiveness of forecasting future individual observations. Finally, conclusions are pointed out in Section VII.

II. Q-Statistics and Q Control Charts

In a sequence of papers Quesenberry [27], [28], [30] introduces normalized statistics, the Q-Statistics, in order to construct Shewhart type control charts for individual observations from, for instance, a normal distribution when the process parameters are unknown at the start-up of the considered process.

A. Q-Statistics from Individual Observations

Quesenberry [27], [28], [30] considers a sequence of first k independent, identically and normally distributed (i. i. d.) random variables

$\{x_1, x_2, \ldots, x_k\}$ with mean μ and variance σ^2, i.e. $\{x_j\} \sim \mathcal{N}(\mu, \sigma^2)$ with $1 \leq j \leq k$. According to the parameters μ and σ^2 Quesenberry [27], [28], [30] distinguishes four cases:

(1) Case KK: $\mu = \mu_0$ and $\sigma = \sigma_0$ are known $(k = 1, 2, \ldots)$

$$Q_k(x_k) = \frac{x_k - \mu_0}{\sigma_0}. \tag{1}$$

(2) Case UK: μ is unknown, $\sigma = \sigma_0$ is known $(k = 2, 3, \ldots)$

$$Q_k(x_k) = \sqrt{\frac{k-1}{k}} \left(\frac{x_k - \overline{x}_{k-1}}{\sigma_0} \right), \tag{2}$$

with

$$\overline{x}_k = \frac{1}{k} \sum_{j=1}^{k} x_k. \tag{3}$$

(3) Case KU: $\mu = \mu_0$ is known, σ is unknown $(k = 2, 3, \ldots)$

$$Q_k(x_k) = \Phi^{-1} \left\{ G_{k-1} \left(\frac{x_k - \mu_0}{s_{k-1}} \right) \right\}, \tag{4}$$

where

$$s_k^2 = \frac{1}{k-1} \sum_{j=1}^{k} (x_j - \overline{x}_k)^2. \tag{5}$$

(4) Case UU: μ and σ unknown $(k = 3, 4, \ldots)$

$$Q_k(x_k) = \Phi^{-1} \left\{ G_{k-2} \left[\sqrt{\frac{k-1}{k}} \left(\frac{x_k - \overline{x}_{k-1}}{s_{k-1}} \right) \right] \right\}, \tag{6}$$

where $\overline{x}_{k-1}$ and s_{k-1} are defined in (3) and (5), respectively.

In (4) and in (6), $G_{(\cdot)}$ denotes the Student t cumulative distribution function with $k-1$ resp. $k-2$ degrees of freedom and Φ^{-1} the inverse of the standard normal cumulative distribution function. For further details see Chang and Tong [2], Fehlmann and Kranich [8], Quesenberry [27], [30], Zantek [40], and Zantek and Nestler [41]. Obviously, Case UU is the most important and occurring case in practice and is therefore the only case investigated in this paper.

With respect to (6), Quesenberry [30] recommends to update $\overline{x}_k$ and s_k^2 each time a new observation is available by means of the following sequential updating formulas instead of calculating $\overline{x}_k$ and s_k^2 each time from scratch according to (3) and (5), respectivley:

(a) The sample mean $\overline{x}_k$ can be calculated sequentially by

$$\begin{aligned} \overline{x}_k &= \frac{1}{k} \sum_{j=1}^{k} x_k = \left(1 - \frac{1}{k}\right) \overline{x}_{k-1} + \frac{1}{k} x_k \\ &= \overline{x}_{k-1} + \frac{1}{k} (x_k - \overline{x}_{k-1}), \end{aligned} \tag{7}$$

for $k \geq 2$ and with $\overline{x}_1 = x_1$.

(b) The sample variance s_k^2 can be computed sequentially by

$$\begin{aligned} s_k^2 &= \frac{1}{k-1} \sum_{j=1}^{k} (x_j - \overline{x}_k)^2 \\ &= \left(\frac{k-2}{k-1} \right) s_{k-1}^2 + \frac{1}{k} (x_k - \overline{x}_{k-1})^2, \end{aligned} \tag{8}$$

for $k \geq 3$ and with

$$s_2^2 = \frac{1}{2} (x_2 - \overline{x}_1)^2 = \frac{1}{2} (x_2 - x_1)^2.$$

Note that the formulas (7) and (8) are numerically more stable than the fundamental formulas (3) and (5), respectively.

B. Q Control Charts

Quesenberry [27], and Zantek and Nestler [41] prove that each Q-Statistic $Q_k(x_k)$ in (1), (2), (4), and (6), respectively, produce a sequence of independent $\mathcal{N}(0,1)$ distributed random variables. Consequently, the 3σ upper control limit (UCL), the center line (CL), and the 3σ lower control limit (LCL) of a Q control chart are fixed:

$$UCL = +3, \qquad CL = 0, \qquad LCL = -3. \tag{9}$$

It is well known that traditional Shewhart control charts help to decide whether a considered process is under statistical control or in-control by using certain run rules or tests. A set of such run rules are listed in e.g. Champ et al. [1], Florac and Carleton [10], Hoyer and Ellis [17], and Montgomery [25]. Quesenberry [28] applies a subset of that run rules to Q control charts. For instance, one such (simple) rule is the *1-of-1* or outlier test, i.e., the process signals an outlier observation, if $Q_k(x_k) < LCL$ or $Q_k(x_k) > UCL$. Quesenberry [29] proves that Q control charts are optimal to detect outliers.

The occurrence of an outlier generally requires some action, since an outlier may strongly impact the sequence of parameter estimates by masking the outlier effect when further observations are taken, compare (7) and (8), respectively. Excluding an outlier from subsequent parameter estimates the sensitivity to detect further outliers will be improved. According to Q control charts based on the Q-Statistic (6) an effective way to automatically build a "new" Q control chart from scratch is to eliminate the outlier and all previous observations from subsequent calculations of the parameter estimates. Based on the example of Florac and Carleton [10, pp. 151], Figure 1 and Figure 2 illustrate the observations sequence without and with eliminating outliers automatically.

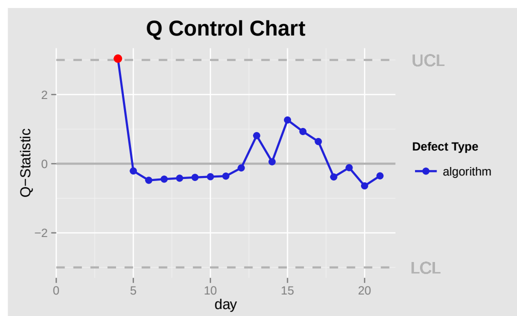

Fig. 1. Observations sequence without outlier elimination

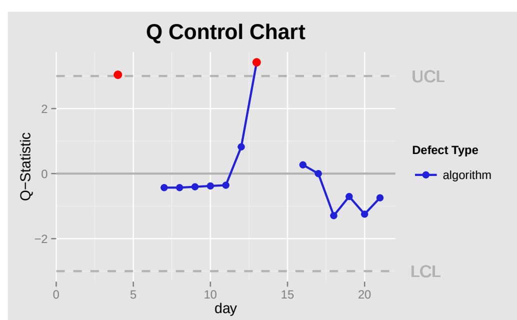

Fig. 2. Observations sequence with outliers elimination

In Figure 1 and Figure 2 an outlier is marked at day 4, and in Figure 2 an additional outlier at day 13. In addition, Figure 2 illustrates that each time an outlier has been detected subsequent Q-Statistic data points are calculated from scratch by means of (6).

A benefit of applying Q control charts is that the behavior of various process characteristics can be visualized in exactly one control chart, see Figure 3.

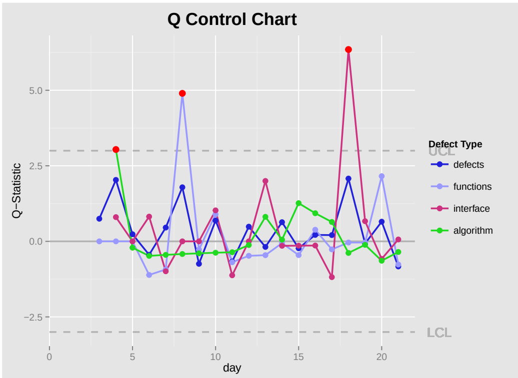

Fig. 3. Q control chart of four process characteristics

Such a control chart gives more insight into the process behavior. For instance, in Figure 3 the peaks of the aggregated number of defects are obviously caused by out-of-control events of the further three considered defect types.

III. EWMA Control Charts

Exponentially Weighted Moving Average (EWMA) control charts are an alternative to traditional Shewhart type control charts, see, e.g. Hunter [18], Montgomery [25], Qiu [26], or Ryan [33]. EWMA control charts are based on a time series model and thereby on a transfer function model, see e.g. Fehlmann and Kranich [7]. Hence, EWMA control charts hold the predictive characteristic of a transfer function and thereby may be applied to control and monitor a process.

A. EWMA Basics

EWMA control charts are based on a statistic of the form

$$z_k = \lambda x_k + (1-\lambda) z_{k-1}, \qquad k \geq 1, \tag{10}$$

where $\lambda \in (0,1]$ is an appropriately chosen weighting factor. Obviously, a $\lambda \in (0, 0.5)$ assigns more weight to the previous z_{k-1} as to an actual observation x_k. If $\lambda \in (0.5, 1)$, less weight is assigned to z_{k-1} and more to x_k. An EWMA control chart is simply a Shewhart control chart when $\lambda = 1$.

Apparently, the statistic (10) reflects that EWMA is a weighted average of all previous observations $x_1, x_2, \ldots, x_{k-1}$ and the current observation x_k. This implies that an EWMA control chart can be constructed, even though the distribution function of the observations is not known.

In order to specify the EWMA control chart limits the mean and variance of the statistic (10) have to be derived. This is accomplished by the recursive form of (10):

$$z_k = \lambda \sum_{j=0}^{k-1} (1-\lambda)^j x_{k-j} + (1-\lambda)^k z_0, \tag{11}$$

where the weights $\lambda(1-\lambda)^j$ geometrically decrease for increasing j.

If all the observations x_j $(1 \leq j \leq k)$ are independent and identically distributed with mean μ_x and variance σ_x^2, the mean of z_k in (10) is equal to

$$\mu_{z_k} = \mu_x \lambda \sum_{j=0}^{k-1} (1-\lambda)^j + (1-\lambda)^k \mu_{z_0} = \mu_x, \tag{12}$$

with $z_0 = \mu_x$ and the variance of the EWMA statistic z_k is

$$\sigma_{z_k}^2 = \sigma_x^2 \left(\lambda^2 \sum_{j=0}^{k-1} (1-\lambda)^{2j} \right) + (1-\lambda)^{2k} \sigma_{z_0}^2. \tag{13}$$

In view of $\sigma_{z_0}^2 = \sigma_{\mu_x}^2 = 0$, the variance of z_k is equal to

$$\sigma_{z_k}^2 = \frac{\lambda}{2-\lambda} \left(1 - (1-\lambda)^{2k} \right) \sigma_x^2, \tag{14}$$

with

$$\lim_{k \to \infty} \sigma_{z_k}^2 = \frac{\lambda}{2-\lambda} \sigma_x^2. \tag{15}$$

For more derivation details of μ_{z_k} in (12) and of $\sigma_{z_k}^2$ in (14), see e.g. Montgomery [25], and Qiu [26].

Hence, the center line CL of an EWMA control chart is equal to μ_x and the EWMA control chart limits are of the form

$$\mu_x \pm \rho \, \sigma_{z_k} = \mu_x \pm \rho \, \sigma_x \sqrt{\frac{\lambda}{2-\lambda} \left(1 - (1-\lambda)^{2k}\right)}, \tag{16}$$

where $\rho > 0$ is chosen appropriately. Montgomery [25] points out that $(\lambda, \rho) \in \{(0.05, 2.615), (0.10, 2.814), (0.20, 2.962), (0.25, 2.998)\}$ works well in practice, see likewise Lucas and Saccucci [24], and Qiu [26]. Evidently, a considered process is in-control, when z_k in (10) is located between the EWMA control chart limits (16).

B. An EWMA FIR Enhancement

In order to calculate the EWMA control chart limits in (16), the limiting variance given in (15) is often used in practice. A small λ motivates this approach since the variance of z_k in (13) slowly converges to its limiting variance (15) because of the slow convergence of $(1 - (1-\lambda)^{2k}) \to 1$. Taking the limiting variance of z_k impacts the start-up phase of a process since the sensitivity of an EWMA control chart to detect an out-of-control event in that phase gets lost.

The Fast Initial Response (FIR) feature compensates this problem, see e.g. Chiu [4], Haq, Brown and Moltchanova [13], Knoth [23], Rhoads, Montgomery and Mastrangelo [32], and Steiner [35]. The FIR feature narrows the EWMA control chart limits, at least in the course of the start-up phase of a process, and thereby increases the sensitivity of an EWMA control chart to detect an out-of-control event in that phase. The general FIR feature is defined by

$$\mathrm{FIR}_{adj} = \left(1 - (1-f)^{1+a(k-1)} \right)^b. \tag{17}$$

If $b = 1$, FIR_{adj} in (17) reflects Steiner's FIR adjustment [35], whereas the FIR adjustment of Haq, Brown and Moltchanova [13] results from setting $b = 1 + (1/k)$. Thus the adjusted EWMA control chart limits are of the form

$$\mu_x \pm \rho \times \mathrm{FIR}_{adj} \times \sigma_x \sqrt{\lambda/(2-\lambda)}. \tag{18}$$

According to (17), the parameter a is chosen (or calculated) such that FIR_{adj} has minor impact on the EWMA control chart limits (18) after a pre-defined (observation) index $k = k_0$, i.e., for indexes $k > k_0$ $\mathrm{FIR}_{adj} \approx 1$ is required. For $k = 1$, the parameter $f \in (0,1]$ reflects the proportion of the distance the FIR EWMA control limits (18) have from the center line to the limiting FIR EWMA control limits $\mu_x \pm \rho \times \sigma_x \sqrt{\lambda/(2-\lambda)}$. Steiner [35] recommends to set $a = 0.3$ and $f = 0.5$ in general.

IV. EWMA Q CONTROL CHARTS

Replacing x_k in (10) with a Q-Statistic $Q_k(x_k)$ defined in (1), (2), (4), or (6), results in the statistic

$$z_k = \lambda Q_k(x_k) + (1-\lambda) z_{k-1}, \qquad k \geq 1, \tag{19}$$

where $\lambda \in (0,1]$ is the appropriately chosen weighting factor. Since the initial value of z_k highly impacts the calculation of all subsequent values, z_{k-1} is set to the starting value of the considered Q-Statistic $Q_k(x_k)$. For instance, according to Case UU in (6) the initial value of the sequence z_k is $z_3 = Q_3(x_3)$.

As mentioned in Section II, the Q-Statistics $Q_k(x_k)$ are independent and $\mathcal{N}(0,1)$ distributed. Hence, by (12) $\mu_{z_k} = \mu_{Q_k(x_k)} = 0$, and in view of (14)

$$\sigma^2_{Q_k(x_k)} = \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2k}\right). \tag{20}$$

Then the EWMA Q control chart limits are given by

$$\pm\rho\sqrt{\frac{\lambda}{2-\lambda}\,(1-(1-\lambda)^{2k})}, \tag{21}$$

compare (16), and the FIR adjusted EWMA Q control chart limits are of the form

$$\pm\rho \times \mathrm{FIR}_{adj} \times \sqrt{\lambda/(2-\lambda)} \tag{22}$$

in view of (18).

In order to assure a good EWMA Q control chart performance, $\lambda = 0.25$ and $\rho = 2.998$ is chosen resulting in EWMA Q control chart limits of ± 1.133, or in $\pm 1.133 \times \mathrm{FIR}_{adj}$. Thus $|z_k| > 1.133$ (or, $|z_k| > 1.133 \times \mathrm{FIR}_{adj}$) indicates an out-of-control situation of a process under consideration. Figure 4 depicts an EWMA Q control chart according to the EWMA Q-Statistic (19) and the control limits (16) with $\mu_x = 0$ and $\sigma_x = 1$. Since these control limits are tighter than those in Figure 1 two additional out-of-control events at day 5 and 6 are detected. Note that the non-connected points in Figure 4 represent the EWMA Q-Statistic values.

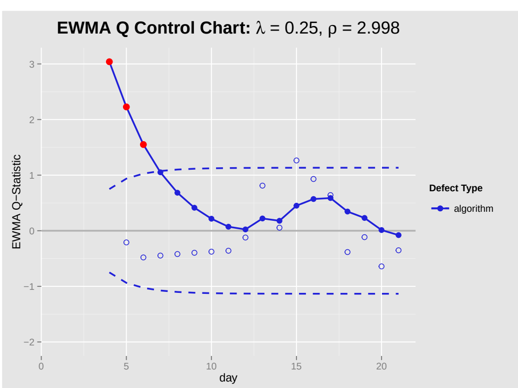

Fig. 4. An EWMA Q control chart (based on Eq. (16))

As mentioned in Section III-B, the sensitivity to detect out-of-control events can be increased at the start-up phase of a process, if the FIR enhancement defined in (17) is taken into account. Figure 5 and Figure 6 visualize the FIR enhancements of Steiner [35] and Haq, Brown and Moltchanova [13], respectively, with $a = 0.3$ and $f = 0.5$.

In both cases one additional out-of-control event at day 7 is detected.

Analog to the context of Figure 2, Figure 7 illustrates the modified FIR adjusted EWMA Q control chart when the out-of-control event at day 4 in Figure 6 is eliminated automatically.

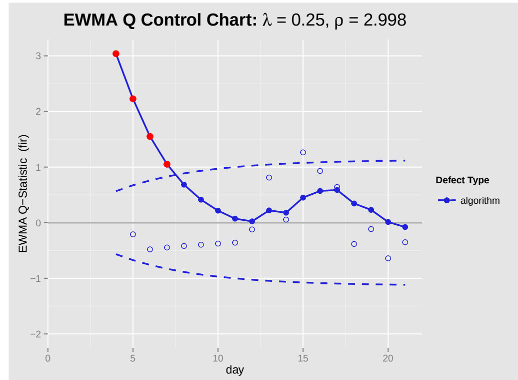

Fig. 5. A FIR adjusted EWMA Q control chart (Steiner [35])

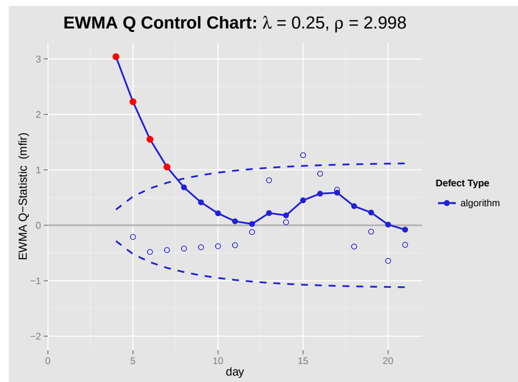

Fig. 6. A modified FIR adjusted EWMA Q control chart (Haq et al. [13])

Analog to Figure 2, the elimination of the out-of-control event at day 4 initiates the calculation of a new sequence of EWMA Q-Statistic values up to day 15 at which another out-of-control event occurs.

V. Q-STATISTICS PREDICTIONS BASED ON EWMA

Prediction or forecasting is a very important function in many business areas, see e.g. Fildes et al. [9]. The goal of forecasting is to predict values of a time series as reliable as possible in the (near) future. One of the most widely applied forecasting method that continually updates a forecast is exponential smoothing, see, for instance, Fildes et al. [9], Gardner [11], [12], Hyndman and Athanasopoulos [19], or Hyndman et al. [22].

A. Q-Statistics and Simple Exponential Smoothing

Simple exponential smoothing is appropriate for short-term prediction, e.g. for the forecast of the next, one-step-ahead time series value.

The statistic (19) can be rewritten as

$$z_k = z_{k-1} + \lambda\,(Q_k(x_k) - z_{k-1}), \quad \lambda \in (0,1), \tag{23}$$

where z_{k-1} is interpreted as the forecast or prediction of the Q-Statistic $Q_k(x_k)$ for iteration k. The difference $Q_k(x_k) - z_{k-1}$ is termed the forecast error at iteration k. Therefore, the error correction form of (23) is equal to

$$z_k = z_{k-1} + \lambda\, e_k \ \text{ with } \ e_k = Q_k(x_k) - z_{k-1}. \tag{24}$$

In order to express that (23) resp. (24) represents a prediction or forecast, an alternative representation of z_k in (24) is

$$z_{k+1} = z_k + \lambda\, e_k \ \text{ with } \ e_k = Q_k(x_k) - z_k, \quad \lambda \in (0,1), \tag{25}$$

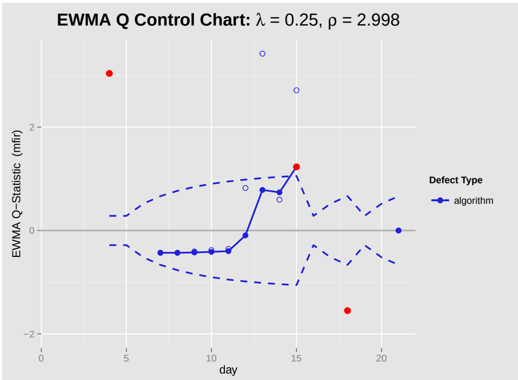

Fig. 7. Modified FIR EWMA Q control chart with outliers elimination

reflecting that a tomorrow's predicted value z_{k+1} is equal to the today's predicted value z_k plus the smoothing parameter λ times the today's prediction error e_k. This way to calculate EWMA based one-step-ahead forecasts is termed simple or single exponential smoothing due to the presence of the single smoothing parameter λ in (25).

The value of the parameter λ has a significant impact on smoothing the predicted values. Obviously, a (too) small value of λ entails that the new forecast z_{k+1} is quite equal to the previous forecast z_k, a (too) large λ focuses on the prediction error e_k. Thus a natural question arises: How to determine λ in order to get a reliable forecast effect? A simple trial and error approach seeks an (approximately) optimal λ that minimizes the sum of squared prediction errors $\sum_i e_i^2$. This can be accomplished by replacing λ with different values in (25), by calculating the corresponding sums of squared prediction errors for each value of λ and finally by selecting that value of λ which yields the minimum value of the sum. An alternative approach is to solve the quadratic programming problem $\min \sum_i e_i^2$ subject to λ by means of a mathematical programming software package.

Another alternative approach is to utilize the `R` Statistical Computing Environment [31], in particular the standard package `stats`, or the package `forecast` developed by Hyndman and Khandakar [20]. Both packages implement the forecast procedure described in Holt [15] and enhanced by Winters [39] to determine a solution to (25).

B. Q-Statistics and Double Exponential Smoothing

In general, time series data exhibit random variations, but in some cases the data may show a shift to higher or lower values over a certain time period. In this case a trend pattern exists. A trend reflects the long-term direction of the considered observations series.

According to the Q control chart run rules (see Section II-B), a positive trend in the observations exists if (at least) six consecutive out of the k actual observations reveal a monotone increasing pattern, i.e. if $Q_{j-5}(x_{j-5}) < Q_{j-4}(x_{j-4}) < \cdots < Q_j(x_j)$ with $j \leq k$, see e.g. Hoyer and Ellis [17, Rule 5]. A negative trend is defined analogously by a monotone decreasing pattern. In exponential smoothing a trend is of the form

$$t_k = \nu(z_k - z_{k-1}) + (1-\nu)t_{k-1}, \tag{26}$$

whereby $\nu \in (0,1)$ is a smoothing parameter. Obviously, the term t_k defines the local linear trend of an observation series for each $k \geq 1$. Alike the parameter λ in Section V-A, the value of ν impacts trend. In case of a small value of ν the new trend t_k is nearly equal to the previous one t_{k-1}, i.e., the difference $z_k - z_{k-1}$ in (26) affects the trend in a minor extent.

Incorporating the trend component (26) into the statistic (19) results in

$$z_k = \lambda Q_k(x_k) + (1-\lambda)(z_{k-1} + t_{k-1}), \quad \lambda \in (0,1) \tag{27}$$

which is equivalent to

$$z_k = z_{k-1} + t_{k-1} + \lambda(e_k), \quad \lambda \in (0,1), \tag{28}$$

with $e_k = Q_k(x_k) - z_{k-1} - t_{k-1}$ denotes the prediction error according to the trend, compare (25).

In view of (28) and the parameters λ and ν explicitly and implicitly given therein, the procedure to calculate a forecast or prediction z_k in (28) is termed double exponential smoothing. As in Section V-A the `R` implementation of Holt's procedure [15] is applied to determine the smoothing parameters λ and ν.

C. Measuring the Accuracy of Q-Statistics Predictions

Evidently, in order to decide whether an actual prediction calculated by (25) or by (28) is acceptable or not, some measure of forecast accuracy is required. A crucial component of some underlying metrics is the mean of the (absolute or squared) one-step-ahead forecast errors $e_j, 1 \leq j \leq k$, see e.g. Hyndman et al. [22].

For instance, the *tracking signal* TS_k is in general defined by

$$TS_k = \frac{\sum_{j=1}^k e_j}{\frac{1}{k}\sum_{j=1}^k |e_j|}. \tag{29}$$

TS_k sets the bias in relation to the average absolut forecast error and is recalculated each time a new individual observation has been made. Ideally, the various tracking signal values $TS_j, 1 \leq j \leq k$, fluctuate around zero within user-defined acceptable control limits, e.g. ± 4. If a tracking signal TS_j exceeds the control limits, then the prediction error may be nonrandom and the actual prediction is no longer beneficial. In this case, reset the forecasting and re-start.

The *Mean Absolute Scaled Error* (MASE) prediction accuracy measure of Hyndman and Koehler [21] checks the mean of the absolute forecast errors $e_j, 1 \leq j \leq k$ against that of the absolute naïve one step forecast errors $Q_j(x_j) - Q_{j-1}(x_{j-1})$ for $2 \leq j \leq k-1$:

$$MASE_k = \frac{\frac{1}{k}\sum_{j=1}^k |e_j|}{\frac{1}{k-1}\sum_{j=2}^k |Q_j(x_j) - Q_{j-1}(x_{j-1})|}. \tag{30}$$

Clearly, if $MASE_k < 1$, then the forecast is better than the average naïve forecast, and the forecast is worser than the average naïve forecast when $MASE_k > 1$. In addition, each $MASE_k$ is undefined when all differences $|Q_j(x_j) - Q_{j-1}(x_{j-1})| = 0$ for $2 \leq j \leq k$.

Another forecast accuracy measure was proposed by the Dutch econometrician Henri Theil who mainly studied the inequality distribution of income and asset. Related to times series the Theil accuracy measure U_k is defined by

$$U_k = \sqrt{\frac{\sum_{j=1}^k e_j^2}{\sum_{j=2}^k (Q_j(x_j) - Q_{j-1}(x_{j-1}))^2}}. \tag{31}$$

Like $MASE_k$ in (30) Theil's measure U_k quantifies how well the series of forecast errors compares to the corresponding series of naïve one step forecast errors. Hence, if $U_k < 1$, then the forecast is better than the naïve one step forecast. The forecast is worser, when $U_k > 1$.

Trigg and Leach [38] modify (29) by replacing the prediction errors e_j and $|e_j|$ with the corresponding smoothed prediction errors resulting in the *smoothed error tracking signal* defined by

$$SETS_k = \frac{\tilde{e}_k}{\tilde{a}_k} = \frac{\gamma e_k + (1-\gamma)\tilde{e}_{k-1}}{\gamma|e_k| + (1-\gamma)\tilde{a}_{k-1}}. \tag{32}$$

Evidently, $SETS_k \in [-1, +1]$. With respect to Section III, the limiting variance of $SETS_k$ is equal to $c \times (\gamma/(2-\gamma))$, where $c \approx 1.5$ in general. In order to determine a value of γ, Trigg [37], and Trigg and Leach [38] propose to select a value of $\gamma \in [0.1, 0.3]$. Then the smoothed error tracking signal $SETS_k$ in (32) indicates an out-of-control forecast, when $SETS_k$ exceeds one of the corresponding control limits. For instance, when $\gamma = 0.1$ the control limits are ± 0.51, see Trigg [37]. When $SETS_k$ is in-control, Trigg and Leach [38] propose to set $\lambda = |SETS_k|$ in (23) and (28), respectively. When $SETS_k$ is out-of-control, Trigg [37] recommends to check the forecast in order to detect a potential assignable cause according to the transformed $Q_k(x_k)$ or to the real observation x_k.

A problem arises from the tracking signal given in (32), when e_k is a perfect prediction, i.e. when $e_k = 0$. In this case, $STS_k = STS_{k-1}$. In view of (30), this problem is solved when the denominator in (32) is replaced by a simple exponentially smoothing expression q_k which consists of the difference between the actual observation $Q_k(x_k)$ and its predecessor $Q_{k-1}(x_{k-1})$, i.e. $q_k = Q_k(x_k) - Q_{k-1}(x_{k-1})$. Then the new smoothed error tracking signal is

$$SETS_k = \frac{\tilde{e}_k}{\tilde{q}_k} = \frac{\gamma e_k + (1-\gamma)\tilde{e}_{k-1}}{\gamma q_k + (1-\gamma)\tilde{q}_{k-1}}, \tag{33}$$

compare (30). In order to determine a value of γ one can proceed by following the approach described in the previous paragraph. Or, simulations or practical experiments have to be conducted, which will be done by the authors in the near future.

When utilizing the `R` package `forecast` written by Hyndman and Khandakar [20] a first insight into the performance of the various tracking signals is given. In the next section a practical example is investigated by applying the `forecast` package.

VI. An Example

The results of Section IV indicate that EWMA Q control charts are attractive to control and to monitor a Software Development Process (SDP) and its various phases or stages. In Section V forecasts based on EWMA Q control charts and accuracy measures of such forecasts were introduced. This section exemplifies that forecasts are beneficial to control and to monitor an SDP, in this case the aggregated number of defects detected in a short-run test phase of the SDP, see Florac and Carleton [10, pp. 150].

When viewing Figure 8 a tester considers himself to be confident that the actual test process is under control.

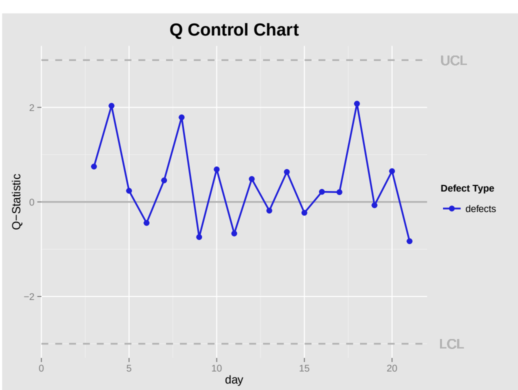

Fig. 8. Q Control chart of aggregated defects

But when applying an EWMA Q control chart, at least one out-of-control event occurred, see Figure 9 and Figure 10, respectively.

The `R` package `forecast` developed by Hyndman and Khandakar [20] implements the function `ses` for simple exponential

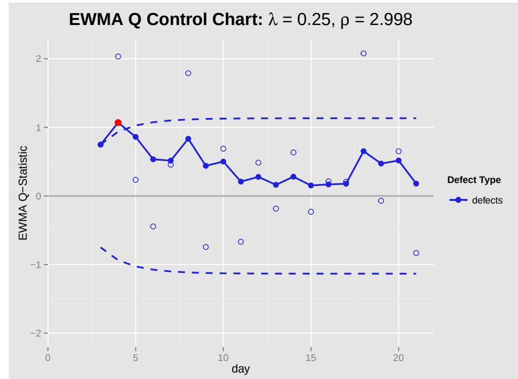

Fig. 9. EWMA Q control chart of aggregrated defects

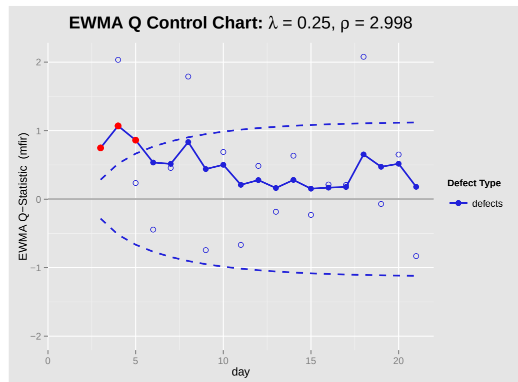

Fig. 10. Modified FIR EWMA Q control chart of aggregated defects

smoothing and the function `holt` for Holt's procedure [15]. The output of both functions can be controlled by some parameters, amongst others by the forecast horizon h. In this paper the forecast horizon h is always set to $h = 1$, since it is a beneficial test strategy to directly react to out-of-control events.

In order to apply the `ses` and `holt` functions to the number of defects the corresponding Q-Statistic values are calculated, see Table I. Recall that in view of (6), $Q_1(x_1)$ and $Q_2(x_2)$ are not defined.

TABLE I. Q-Statistic values

[1]	NA	NA	0.7481477	2.0335408	0.2355839
[6]	-0.4445491	0.4559483	1.7892421	-0.7438131	0.6884329
[11]	-0.6677563	0.4852904	-0.1842182	0.6343398	-0.2298550
[16]	0.2133119	0.2073223	2.0780123	-0.0701442	0.6508527
[21]	-0.8308747				

In a first step, the functions `ses` and `holt` are applied to predict the Q-Statistic value $Q_{10}(x_{10})$. The results are listed in Table II.

TABLE II. Predicting $Q_{10}(x_{10})$

fnct.	forecast	$Q_{10}(x_{10})$	80% CI	95% CI	MASE
`ses`	0.581	0.688	$[-0.66, 1.82]$	$[-1.31, 2.48]$	0.999
`holt`	-0.096	0.688	$[-1.26, 1.07]$	$[-1.87, 1.68]$	0.957

Since $MASE < 1$ in both cases, the accuracy of the forecasts are better than those of the average naïve forecasts. Furthermore, a closer look at the `summary` function of the `ses` and `holt` forecasts reveals that a trend cannot be detected.

When predicting $Q_{18}(x_{18})$ the `holt` function detects a trend defined in (26). The trend parameter ν is equal to 0.1222. The results of the `ses` and `holt` forecasts of $Q_{18}(x_{18})$ are listed in Table III.

TABLE III. PREDICTING $Q_{18}(x_{18})$

fnct.	forecast	$Q_{18}(x_{18})$	80% CI	95% CI	MASE
`ses`	0.348	2.078	$[-0.64, 1.33]$	$[-1.16, 1.85]$	1.000
`holt`	0.182	2.078	$[-0.86, 1.22]$	$[-1.41, 1.77]$	1.131

Obviously, the accuracy of both forecasts is not better than that of the average naïve forecast.

Q-Statistic forecasts can be controlled and monitored by means of control charts. The associated control limits can be calculated by taking into account the standard deviation of the forecast errors e_k. The `summary` function of the `ses` and `holt` procedures lists the standard deviation. The control limits are in general ± 2 or ± 3 times the standard deviation. The Q-Statistic forecasts listed in Table II and in Table III do not reveal an out-of-control forecast.

In view of (6) a forecast of the data point x_k can readily be determined:

$$x_k = \bar{x}_{k-1} + \sqrt{\frac{k}{k-1}} \times s_{k-1} \times G_{k-2}^{-1}\Big\{\Phi\Big[Q_k(x_k)\Big]\Big\}, \qquad (34)$$

where G_{k-2}^{-1} denotes the inverse of the Student t cumulative distribution function with $k-2$ degrees of freedom, and Φ the standard normal cumulative distribution function. Thus, replacing the term $Q_k(x_k)$ in (34) with a corresponding Q-Statistic forecast results in a forecast of x_k. A forecast of the range of x_k is calculated by replacing the term $Q_k(x_k)$ with the endpoints of the corresponding confidence interval of 80% or 90% confidence level. For instance, in view of the values listed in Table III one obtains:

TABLE IV. PREDICTING EVENT x_{18}

fnct.	forecast x_{18}	x_{18}	80% CI	95% CI
`ses`	24	36	$[17, 31]$	$[14, 35]$
`holt`	23	36	$[16, 30]$	$[12, 35]$

Suppose that the forecast x_k has been calculated and the "true" x_k has been observed. If the true x_k does not lie in one of the confidence intervals of the forecast x_k, then an investigation why this event occurred is appropriate in order to potentially decrease the number of aggregated defects immediately.

VII. CONCLUSION

Traditional Shewhart control charts are a successfully applied Statistical Process Control (SPC) tool for controlling and monitoring the variation of long-run mass production processes. To ascertain robust and valid control chart limits, a sufficiently large set of observations drawn from a process under consideration is required. Hence, Shewhart control charts cannot readily be adopted to the Software Development Process (SDP), since the SDP does not provide in general the needed large set of observations. As indicated in this paper, Q control charts are a highly appropriate alternative to Shewhart control charts in the SDP context, because they enable early detection of nonrandom process behavior and the controlling and monitoring of the SDP as a short-run process in real time. These Q control chart properties originate from pre-defined, constant control limits making the Phase I as part of classic Shewhart control charts in order to stabilize the control chart limits redundant.

This paper focused on the performance of Q-Statistics based EWMA control charts, inclusive the FIR adjustment enhancement, and their forecasting feature and on the quality of the forecasts. Although only limited experiments have been conducted, the results are promising. So further practical experiments have to be conducted in order to confirm the findings obtained so far.

REFERENCES

[1] C. W. Champ, W. H. Wood, R. W. Hoyer, and W. C. Elliall, "Exact results for Shewhart control charts with supplementary runs rules," *Technometrics*, vol. 29, no. 4, pp. 393–399, 1987.

[2] C.-W. Chang and L.-I. Tong, "Monitoring the software development process using a short-run control chart," *Software Quality Journal*, vol. 21, no. 3, pp. 479–499, 2013.

[3] R. Chillarege, *Orthogonal Defect Classification*. McGraw-Hill Book Company, 1996, ch. 9, pp. 359–400.

[4] W. C. Chiu, "General weighted moving average control chart with fast initial response features," *Journal of Applied Statistics*, vol. 36, no. 3, pp. 255–275, 2009.

[5] S. A. DeLurgio, *Forecasting Principles and Applications*. Boston, MA, USA: Irwin McGraw-Hill, 1998.

[6] W. E. Deming, *Out of the Crisis*. Cambridge, MA, USA: The MIT Press, 2000, (originally published in 1986).

[7] T. M. Fehlmann and E. Kranich, "Transfer functions, eigenvalues and QFD in concert," in *Proceedings of the 17th International QFD Symposium (ISQFD Stuttgart 2011)*, 2011.

[8] ——, "Introducing short-run control charts for monitoring the software developoment process," in *MetriKon 2013 - Praxis der Software-Messung (DASMA Software Metrik Kongress, Fraunhofer IESE, Kaiserslautern, Germany, November 14-15, 2013)*, 2013, pp. 213–234.

[9] R. Fildes, K. Nikolopoulus, S. F. Crone, and A. A. Syntetos, "Forecasting and Operational Research: A review," *Journal of the Operational Research Society*, vol. 59, no. 9, pp. 1150–1172, 2008.

[10] W. A. Florac and A. D. Carleton, *Measuring the Software Process - Statistical Process Control for Software Process Improvement*. Boston, MA, USA: Addison-Wesley, 1999.

[11] E. S. Gardner, "Exponential smoothing: The state of the art," *Journal of Forecasting*, vol. 4, no. 1, pp. 1–28, 1985.

[12] ——, "Exponential smoothing: The state of the art - part ii," *International Journal of Forecasting*, vol. 22, no. 4, pp. 637–666, 2006.

[13] A. Haq, J. A. Brown, and E. Moltchanova, "Improved fast initial response features for exponentially weighted moving average and cumulative sum control charts," *Quality and Reliability Engineering International*, 2014, in press, online version: http://dx.doi.org/10.1002/qre.1521.

[14] F. He, W.Jiang, and L. Shu, "Improved self-starting control charts for short runs," *Quality Technology & Quantitative Management*, vol. 5, no. 3, pp. 289–308, 2008.

[15] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.

[16] R. W. Hoyer and W. C. Ellis, "A graphical exploration of SPC - Part 1: SPC's definitions and methods," *Quality Progress*, vol. 29, no. 5, pp. 65–73, 1996.

[17] ——, "A graphical exploration of SPC - Part 2: The probability structure of rules for interpreting control charts," *Quality Progress*, vol. 29, no. 6, pp. 57–64, 1996.

[18] J. S. Hunter, "The exponentially weighted moving average," *Journal of Quality Technology*, vol. 18, no. 4, pp. 203–210, 1986.

[19] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. OTexts, 2014, available online at https://www.otexts.org/fpp.

[20] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The `forecast` package for `R`," *Journal of Statistical Software*, vol. 27, no. 3, pp. 1–22, 2008.

[21] R. J. Hyndman and A. B. Koehler, "Another look at measaures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[22] R. J. Hyndman, A. B. Koehler, J. K. Ord, and R. D. Snyder, *Forecasting with Exponential Smoothing - The State Space Approach*. Berlin, Heidelberg, Germany: Springer Verlag, 2008.

[23] S. Knoth, "Fast initial response features for EWMA control charts," *Statistical Papers*, vol. 46, no. 1, pp. 47–64, 2005.

[24] 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.

[25] D. C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. New York, NY, USA: John Wiley & Sons, Inc., 2009.

[26] P. Qiu, *Introduction to Statistical Process Control*. Boca Raton, FL, USA: CRC Press, 2014.

[27] C. P. Quesenberry, "SPC Q Charts for start-up processes and short or long runs," *Journal of Quality Technology*, vol. 23, no. 3, pp. 213–224, 1991.

[28] ——, "On properties of Q Charts for variables," *Journal of Quality Technology*, vol. 27, no. 3, pp. 184–203, 1995.

[29] ——, "On optimality of Q-charts for outliers in statistical process control," in *Statistical Theory and Applications : Papers in the Honor of Herbert A. Davis*. New York, NY, USA: Springer Verlag, 1996, pp. 305–310.

[30] ——, *SPC Methods for Quality Improvement*. New York, NY, USA: John Wiley & Sons, Inc., 1997.

[31] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014, url: http://www.R-project.org.

[32] T. R. Rhoads, D. C. Montgomery, and C. M. Mastrangelo, "A fast initial response scheme for the exponentially weighted moving average control chart," *Quality Engineering*, vol. 9, no. 2, pp. 317–327, 1996.

[33] T. P. Ryan, *Statistical Methods for Quality Improvement*, 3rd ed. New York, NY, USA: John Wiley & Sons, Inc., 2011.

[34] W. A. Shewhart, *Statistical Method from the Viewpoint of Quality Control*. New York, NY, USA: Dover Publications, Inc., 1986, (unabridged republication of the book originally published in 1939).

[35] S. H. Steiner, "EWMA control charts with time-varying control limits and fast initial response," *Journal of Quality Technology*, vol. 31, no. 1, pp. 75–86, 1999.

[36] J. R. Thompson and J. Koronacki, *Statistical Process Control - The Deming Paradigm and Beyond*, 2nd ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2002.

[37] D. W. Trigg, "Monitoring a forecasting system," *Operational Research Quarterly*, vol. 15, no. 3, pp. 271–274, 1964.

[38] D. W. Trigg and A. G. Leach, "Exponential smoothing with an adaptive response rate," *Journal of the Operational Research Society*, vol. 18, no. 1, pp. 53–59, 1967.

[39] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol. 6, no. 3, pp. 324–342, 1960.

[40] P. F. Zantek, "Run-length distributions of Q-Chart schemes," *IIE Transactions*, vol. 37, no. 11, pp. 1037–1045, 2005.

[41] P. F. Zantek and S. T. Nestler, "Performance and properties of Q-Statistic monitoring schemes," *Naval Research Logistics*, vol. 56, no. 3, pp. 279–292, 2009.